

AVALIAÇÃO DO CONHECIMENTO FONOLÓGICO DE CRIANÇAS COM DISTÚRBO ESPECÍFICO DA LINGUAGEM E MODELAGEM ESTATÍSTICA¹

Christina Abreu Gomes

Gastão Coelho Gomes

Aline de Azevedo Ferreira

Resumo

Esse artigo aborda o tipo de conhecimento fonológico que caracteriza crianças com Distúrbio Específico da Linguagem (DEL). Os dados foram obtidos através dos testes de nomeação e repetição que fazem parte do pacote ABFW com o objetivo de checar a acuidade da produção das crianças e sua relação com o tamanho do léxico. Um modelo de regressão foi aplicado para a modelagem dos dados. Os resultados obtidos indicaram que as crianças DEL mostraram altos índices de processos fonológicos envolvendo aspectos mais abstratos das formas das palavras como estrutura silábica se comparadas ao grupo controle e que esse comportamento está relacionado ao tamanho do léxico da criança.

Palavras-chave: conhecimento fonológico, léxico, distúrbio específico da linguagem.

Abstract

Evaluation of phonological knowledge of children with Specific Language Impairment and statistical modeling

This paper addresses the type of phonological knowledge that characterizes children with Specific Language Impairment (SLI). Data were obtained through the naming and repetition of words tests that are part of the ABFW Package, in order to check accuracy in children's production and its relationship with lexicon size. A regression model was applied for modeling the data. The results obtained revealed that SLI children showed higher scores of phonological processes involving abstract aspects of the word form such as syllable structure if compared to the control group, and that it is related to the size of the child's lexicon.

Keywords: phonological knowledge, lexicon, specific language impairment.

1. Introdução

O objetivo desse artigo é apresentar a aplicação de alguns métodos estatísticos no tratamento de dados lingüísticos obtidos através de método experimental, a partir de hipóteses que comparam o comportamento de crianças com diagnóstico de Distúrbio Específico da

Linguagem (DEL) e crianças com desenvolvimento típico, o grupo controle, com idades entre 4 e 6 anos de idade.

As pesquisas realizadas na área da psicolinguística precisam lidar com dados de um certo número de indivíduos, em diferentes condições de teste, em alguns casos de diferentes tipos de experimentos. Em todas essas situações, é preciso checar a significância e a confiabilidade dos resultados obtidos através de diferentes técnicas quantitativas. É necessário definir o tipo de observação ou a variável a ser explicada (variável dependente), as variáveis explicativas (ou variáveis independentes), e o método estatístico a ser utilizado. O método experimental pressupõe o controle das variáveis explicativas no desenho experimental, isto é, estímulos serão estabelecidos e organizados em função das variáveis explicativas e os resultados são observados para uma certa amostra. Esta amostra vai ser a base para a tomada de decisão das hipóteses de pesquisa. As variáveis explicativas podem ser uni ou multi-dimensionais, situação em que uma ou mais variáveis influenciam a variável resposta (variável dependente).

Esse artigo se estrutura da seguinte maneira: na seção a seguir apresentamos o quadro teórico de referência bem como as hipóteses gerais que norteiam o trabalho, na seção seguinte são apresentadas algumas noções básicas da ferramenta estatística para a tomada de decisão em relação às hipóteses levantadas e em seguida a análise linguística relativa à avaliação do conhecimento fonológica de crianças DEL.

2. Quadro teórico

Pierrehumbert (2003) propõe um modelo de representação fonológica com uma base fonética forte e diferentes níveis de representação que emergem da representação da forma sonora das palavras no léxico e que contemplam aspectos articulatórios finos. Os níveis propostos são: 1) nível fonético paramétrico: representação fonética paramétrica é o mapa

quantitativo do espaço articulatório e acústico. Na percepção, descreve a codificação perceptual do sinal de fala para cada ocasião. Na produção, descreve cada gesto articulatório que se desenvolve/ocorre no tempo e no espaço; 2) nível da codificação fonética: abstrai sobre o espaço fonético paramétrico, definindo o inventário disponível na língua para a codificação da forma das palavras. Corresponde à fonologia nos modelos clássicos, mas na proposta de Pierrehumbert fonemas seriam categorias menos abstratas que o conceito difundido nos modelos estruturalista e gerativista. Não são categorias mínimas porque contêm informação redundante e não podem ser equacionados via contexto. Codificação fonética inclui aspectos da prosódia e da intonação; 3) representação da forma sonora das palavras no léxico: cada palavra do léxico do falante tem uma representação que reflete sua experiência de produzir e ouvir os itens lexicais. Adota-se a representação de acordo com o modelo de exemplares (Johnson, 1997), em que diversas realizações da mesma instância são representadas capturando a gradiência entre elas; 4) gramática fonológica: representa generalizações sobre a forma das palavras no léxico, que são por sua vez, generalizações sobre a fala. A fonologia não abstrai sobre a fala propriamente, mas indiretamente através da abstração das formas das palavras. Os aprendizes/falantes têm à disposição uma quantidade de dados muito grande para adquirir generalizações sobre a fala, como a pronúncia nativa. Inclui a estrutura prosódica e a fonotática e descreve o conjunto de itens lexicais possíveis na língua; 5) nível das correspondências morfofonológicas: correspondências entre formas fonológicas e morfológicas. Nem todas as alternâncias morfológicas se dão com base em um padrão fonológico geral. Muitas dessas correspondências precisam ser aprendidas com base na relação específica entre palavras ou no mesmo paradigma. Essas correspondências são estabelecidas no léxico, organizado em redes de relações lexicais (Bybee, 1998).

Beckman, Munson & Edwards (2007) argumentam que pelo menos dois níveis de representação postulados por Pierrehumbert (2003) podem se desenvolver de forma independente em populações clínicas. Os autores demonstraram, através de testes com pseudopalavras, observando a

acuidade de crianças na produção de difones de alta e baixa frequência do inglês, que as crianças com DEL têm maior dificuldade global com a tarefa. Além disso, são afetadas pela dificuldade de produzir as sucessões de baixa frequência. Percentuais maiores ou menores de sucesso estão relacionados com o tamanho do léxico. Isto é, quanto maior o tamanho do léxico da criança melhor o desempenho destas em relação à acuracidade na repetição dos difones. De acordo com Beckman et al. (2007), esses resultados sugerem que o DEL seja associado a dificuldades em fazer generalizações fonológicas abstratas através da nuvem de representações episódicas de itens lexicais representados no léxico mental. As generalizações abstratas estão relacionadas ao conjunto de itens armazenados no léxico. Por conseguinte, as crianças com DEL são mais pobres que as de idade cronológica semelhante para generalizar a produção correta do fonema para as seqüências pouco freqüentes. Este mesmo déficit deve ser parcialmente responsável pela dificuldade em adquirir novas palavras uma vez que crianças com DEL não têm robustez nas representações fonológicas abstratas e, portanto, têm dificuldade de processar e armazenar novos itens lexicais. O aumento gradativo do léxico, dentro dos padrões esperados por idade, permite o aumento da acuracidade no desempenho das crianças, já que uma quantidade maior de itens vai demandar um refinamento da informação fonética, ao mesmo tempo em que permite abstrações de mais padrões estruturais. Esses padrões, por sua vez, serão utilizados na interpretação e incorporação de novos itens lexicais. Portanto, mais itens estocados possibilitam mais abstrações de tipos estruturais fonológicos.

3. Método de análise de dados: teste de hipóteses e análise de regressão.

Além de usarmos a análise exploratória, que é um processo de tirar informação de uma massa de dados através de representações gráficas, numéricas ou organizações em tabelas, também são utilizados testes de hipóteses para uma tomada de decisão. O teste de hipóteses é um procedimento de decisão baseado em dados. Em função dos dados, vamos aceitar ou rejeitar uma hipótese conservadora (H_0) pré-estabelecida versus uma hipótese inovadora (H_1). H_0 ou hipótese nula é a hipótese de referência do teste.

Há, portanto, duas possibilidades em relação à hipótese conservadora: aceitar ou rejeitar H_0 . Nessa tomada de decisão, existem duas possibilidades de erro. Erro I, que consiste em rejeitar H_0 quando H_0 é verdadeira e o Erro II, aceitar H_0 , quando H_0 é falsa. O erro I vai ser o erro a ser controlado e sua probabilidade deve ser baixa (.05 ou .01). O erro I é o mais importante a ser evitado, isto é, as hipóteses são formuladas de forma que H_0 seja aquela hipótese cuja rejeição equivocada constitui o erro de maior importância. Este erro supõe H_0 verdadeira, e, para que tenha uma única distribuição de probabilidade sobre essa hipótese, usaremos hipóteses simples com um único valor a ser testado.

Uma ferramenta importante nessa tomada de decisão é o p-valor. O p-valor, que é uma probabilidade, é uma medida da concordância dos dados com a hipótese H_0 , isto é, se o p-valor é pequeno devemos rejeitar a hipótese H_0 e quando ele é grande os dados não indicam evidência para a rejeição de H_0 . Tudo funciona como uma prova por absurdo. Supondo H_0 verdadeira e observando os dados que temos, se eles indicarem que a estatística de teste cai numa região rara de acontecer, devemos rejeitar o que foi suposto verdadeiro, isto é, se o p-valor é pequeno, devemos rejeitar H_0 . Aceitar H_0 significa que não houve evidência para rejeitá-la.

Portanto, no processo de inferência vamos procurar explicar o comportamento de uma variável dependente em função de variáveis independentes. Um método estatístico para dar conta desse ajuste linear é a análise de regressão, que usa o método de mínimos quadrados que nesse trabalho foi usado para explicar a percentagem de ocorrência de processos fonológicos em função da variável independente tamanho do léxico da criança.

Na análise de regressão simples, com uma variável independente, os dados podem ser representados num plano com eixo x (variável independente) e eixo y (variável dependente). A reta dos mínimos quadrados é a reta que ajusta esses dados da melhor maneira possível por esse critério. A equação da reta, $y = \alpha + \beta x + \varepsilon$, é composta por dois parâmetros: o intercepto e

a inclinação. O intercepto (α) corresponde ao ponto y quando x tem o valor zero. A inclinação (β) corresponde ao aumento no valor de y quando x é aumentado de uma unidade. O erro ε corresponde a uma flutuação aleatória, já que a reta teórica (não observada), referente à relação populacional entre x e y, não passa por todos os pontos. O que teremos acesso é uma reta estimada em função da amostra observada.

4. Análise de dados

Nesta seção, procuramos explicar, através de modelos de análise de regressão, o comportamento observado em população clínica, crianças com diagnóstico de Distúrbio Específico da Linguagem (DEL), em dados obtidos em situação de teste (Ferreira, 2007). O objetivo é identificar em que medida as alterações na forma sonora das palavras, típicas de crianças DEL, se relacionam a hipóteses de armazenagem da forma sonora abstrata das palavras em função do léxico estocado da criança. Constitui o conjunto de observações:

1. a variável dependente, $y = \%P$, as percentagens de realização, pelas crianças, dos vários processos fonológicos na palavra-alvo na nomeação e repetição de palavras da Prova Fonológica do Teste ABFW, teste padrão utilizado em avaliação clínica.
2. as variáveis explicativas o tamanho do léxico das crianças e as variáveis indicadoras, representando alguns dos processos fonológicos estabelecidos no teste ABFW, a saber, simplificação de encontro consonantal (**SEC**) (p. ex. plástico como pástico); simplificação de líquida (**SL**) (p.ex. jacaré como jacayé); plosivização de fricativas (**CF**) (p. ex. [f] realizado como [p]) e posteriorização de oclusiva (**PV**) (p. ex. [t] realizado como [k]). Todas essas alterações são esperadas em percentual inferior a 25% em crianças com desenvolvimento típico.

Os dados foram coletados a partir da prova de avaliação fonológica do protocolo do ABFW que contém 39 palavras no teste de imitação (Tabela A1 em anexo), repetição da palavra falada pelo avaliador, e 34 no teste de nomeação, figuras representando as palavras do teste (Tabela A1 em anexo). O teste foi aplicado a 9 crianças DEL e 11 do grupo controle. Todas as sessões de aplicação do teste ABFW foram gravadas. A identificação da forma usada pelas crianças foi estabelecida com base nas gravações e avaliada por mais de um julgador. Os dados obtidos nos dois testes foram agrupados na análise estatística. Todas as palavras existem no léxico da língua e estão adequadas às faixas etárias das crianças da amostra. No teste, a possibilidade de ocorrência dos processos fonológicos não é homogênea, não havendo, portanto, um mesmo número de observações (N) para cada um deles (Tabela A2 em anexo). A %P de cada processo foi obtida em função do total de possibilidades e do número de vezes em que a produção da criança era discrepante em relação à forma da palavra alvo. Por exemplo, há 12 possibilidades (palavras) de simplificação de grupo consonantal, como na palavra livro/livu, no teste de imitação, e 8 no de nomeação, totalizando 20 possibilidades. O teste estabelece a faixa de até 25% de ocorrência de um processo como parâmetro para a aquisição normal.

O tamanho do léxico da criança foi obtido através do teste Peabody, que é um teste padronizado adaptado para o português e utilizado com aplicação clínica. O teste de vocabulário receptivo de Peabody (Dunn & Dunn, 1997) é um teste de compreensão que visa quantificar o conhecimento do vocabulário em crianças (a partir de 2 anos e meio) e adultos. Consiste de uma série de cartões, cada qual contendo 4 gravuras diferentes, num total de 204 cartões. A criança é solicitada a identificar a gravura que melhor representa o significado da palavra enunciada pelo examinador. O teste foi administrado conforme as instruções especificadas no manual, com as palavras traduzidas para o português, e os scores já estabelecidos para cada idade utilizados como referência para a análise (ver tabela A3 em

anexo). Para a faixa etária das crianças da amostra, o score esperado para as crianças com desenvolvimento típico está entre 51 e 79 pontos. O score pode estar adequado, abaixo ou acima do esperado.

A amostra foi constituída de dois grupos de crianças na faixa de 4 a 6 anos de idade: um grupo com diagnóstico de DEL, e o outro controle, crianças com desenvolvimento típico. Na seleção das crianças do grupo controle foram mantidas as características sócio-econômicas das crianças do grupo experimental. Crianças DEL têm dificuldades na aquisição da linguagem, manifestando alterações em diversos níveis lingüísticos, embora não apresentem problemas auditivos, neurológicos ou de inteligência. O comportamento observado para variável dependente (%P) em função da relação entre variáveis explicativas (tamanho do léxico e as variáveis indicadoras: processos fonológicos e grupo experimental x grupo controle) serve para comparar e avaliar o conhecimento fonológico das crianças DEL e do grupo controle para fins clínicos e da Lingüística teórica.

Observar se há relação entre a porcentagem de ocorrência dos processos e o tamanho do léxico pode subsidiar a hipótese lingüística segundo a qual o conhecimento fonológico emerge das formas sonoras das palavras armazenadas no léxico, de acordo com Pierrehumbert (2003). O tamanho do léxico adequado para a idade possibilita desenvolver abstrações fonológicas suficientes para dar conta da representação da forma sonora das palavras. Os processos de simplificação de encontro consonantal (**SEC**) e simplificação de líquida (**SL**) envolvem alteração da estrutura abstrata da palavra como a estrutura silábica e o número de segmentos. A plosivização de fricativas (**CF**) e a posteriorização de oclusiva (**PV**) envolvem alteração da forma fonética sem alterar a estrutura abstrata da palavra no que diz respeito a número de segmentos, sílabas e tipo de sílaba.

4. 1. Análise Estatística

A técnica estatística que usamos é Análise de Regressão, sendo a %P, porcentagem de realizações dos processos, a variável a explicar. As variáveis explicativas são o tamanho do

léxico (x) e algumas variáveis indicadoras, ou seja, variáveis que assumem o valor um quando a característica está presente e o valor zero quando ela está ausente. As 4 variáveis indicadoras aqui usadas são:

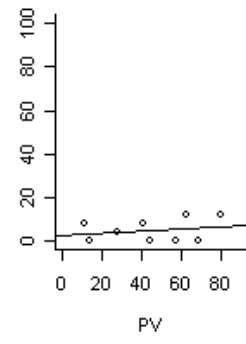
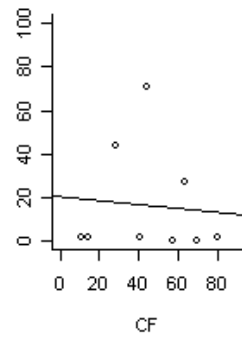
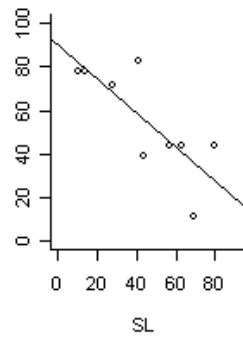
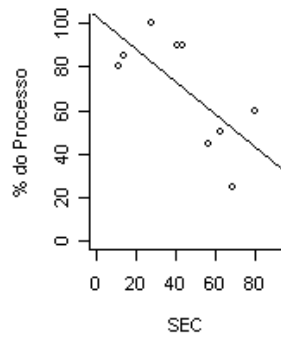
- I_1 para indicar se a observação corresponde ao estudo do processo SEC,
- I_2 para indicar se a observação corresponde ao estudo do processo SL,
- I_3 para indicar se a observação corresponde ao estudo do processo CF,
- I_4 para indicar se a observação corresponde ao estudo do processo PV,

As indicadoras I_1, I_2, I_3, I_4 são excludentes entre si: quando para uma observação uma delas assume o valor um, as outras assumem o valor zero.

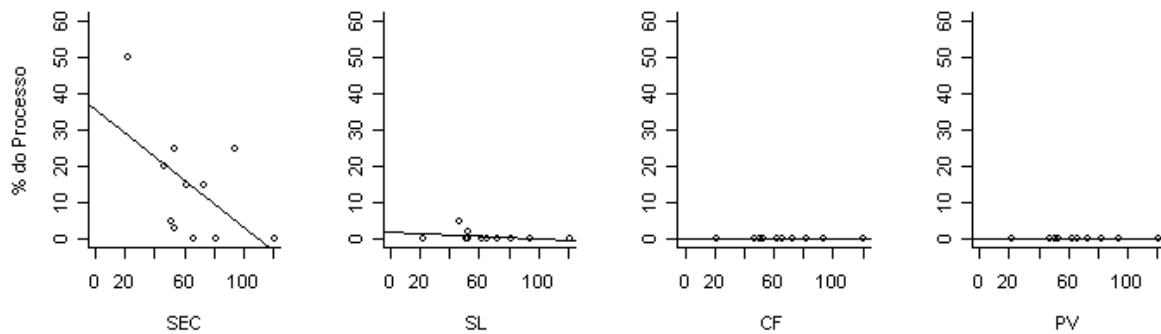
4.2. Análise Exploratória

A seguir apresentamos, para efeito descritivo das observações, os gráficos de cada um dos processos para o grupo das crianças com DEL e para o grupo das crianças controle (CON). Foram tomados alguns cuidados para facilitar a comparação entre os grupos em cada caso. No eixo vertical, em que é medida a variável resposta, %P, foi padronizada a mesma escala, de 0 a 100. No eixo horizontal, a variável x também foi padronizada de 0 a 80 para todos os processos no grupo DEL. No grupo CON a padronização foi de 0 a 60 no eixo y , %P, e 0 a 120 no eixo x , tamanho do léxico. Em todos os quatro gráficos dos dois grupos, DEL e CON, as retas de mínimos quadrados foram ajustadas aos pontos e traçadas.

2.1.1 Gráficos de %P versus Tamanho do Léxico para o grupo DEL



2.1.2 Gráficos de %P versus Tamanho do Léxico para o grupo controle



4.2.1 Comparação do tamanho do léxico dos grupos DEL e CON

Com a intenção de verificar a diferença do tamanho do léxico nos dois grupos estudados foi realizado um teste.t para comparar as médias obtidas no teste Peabody, que verifica o tamanho do léxico adquirido em função da idade da criança, conforme explicado na introdução. O esperado é que as crianças com diagnóstico de DEL apresentem um léxico menor se comparado ao de crianças com desenvolvimento típico na mesma idade, uma vez que essa é uma das manifestações desse distúrbio. O teste.t foi realizado considerando a hipótese $H_0: \mu_{DEL} = \mu_{CON}$ versus $H_1: \mu_{DEL} < \mu_{CON}$. O resultado obtido foi $t=1.7988$, $df=18$, $p\text{-valor}=0.044$, o que indica, a um nível de significância de 5%, diferença entre as médias do tamanho do léxico, sendo o das crianças do grupo controle significativamente maior do que o das crianças DEL.

4.3 Análise de Regressão

Começamos, para o grupo DEL, o processo de estimação com o modelo que ajusta 4 retas de regressão para cada um dos 4 processos. Neste modelo fizemos uma análise de resíduo e avaliamos os resíduos studentizados externos. Assim foi diagnosticado que no grupo controle há uma criança que apresenta o valor de 5.24 cujo resíduo é muito maior e se destaca perante as demais crianças. Foi também avaliada a distância de Cook e o Dfits, que apresentaram valores altos nesta criança, sendo os restantes valores não indicando nenhum

ponto de influência e discrepante. Optamos então pela retirada desta criança e recalculamos os coeficientes de alfa e beta e as estatísticas correspondentes. A Tabela 1 a seguir apresenta os coeficiente e os testes t, como em Montgomery (1982), da equação 1 referente ao modelo completo para os dados do grupo DEL, cuja hipótese H_0 é que o coeficiente é nulo, contra a hipótese alternativa dele ser diferente de zero.

Tabela 1. Estimação dos α e β da equação 1 correspondente ao grupo CON eliminando o ponto de influência (Criança 5)

| | Indicador Do Grupo | Coef. | Estat. - t | p.valor |
|----------------------------|----------------------------|--------------|-------------------|----------------|
| α | I₁ (SEC) | 18.39 | 3.38 | 0.0018 |
| | I₂ (SL) | 1.76 | 0.41 | 0.6878 |
| | I₃ (CF) | 0.00 | 0.00 | 1.0000 |
| | I₄ (PV) | 0.00 | 0.00 | 1.0000 |
| β | I₁X | -0.11 | -1.46 | 0.1523 |
| | I₂X | -0.02 | -0.28 | 0.7831 |
| | I₃X | 0.00 | 0.00 | 1.0000 |
| | I₄X | 0.00 | 0.00 | 1.0000 |

Medidas sobre o ajuste:

Erro Padrão do resíduo = 5.16, $R^2 = 0.5686$
 N = 35, F = 5.78 com 8 e 36 gl, p-valor = 0.0001
 SQR = 933.03

Conforme os resultados observados na tabela acima, nenhum processo rejeita a hipótese dos β 's serem iguais a 0. Em todos os processos (SEC, SL, CF e PV), os valores dos β 's correspondentes não são significativamente diferentes de 0. Do ponto de vista do processo aquisitivo, esses resultados revelam que embora o alfa do processo SEC seja diferente de 0, e, portanto, superior aos demais, para nenhum deles há interferência do tamanho do léxico.

Na tabela 2 estão os coeficiente da equação 1 referente ao modelo completo para os dados do grupo CON.

Tabela 2. Estimaco dos α e β da equaco 1 correspondente ao grupo CON

| | Indicador Do Grupo | Coef. | Estat. - t | p.valor |
|----------------------------|-------------------------------|--------------|-------------------|----------------|
| α | I₁ (SEC) | 35.51 | 6.19 | 0.0000 |
| | I₂ (SL) | 1.76 | 0.31 | 0.7602 |
| | I₃ (CF) | 0.00 | 0.00 | 1.0000 |
| | I₄ (PV) | 0.00 | 0.00 | 1.0000 |
| β | I_{1X} | -0.32 | -3.95 | 0.0004 |
| | I_{2X} | -0.02 | -0.21 | 0.8344 |
| | I_{3X} | 0.00 | 0.00 | 1.0000 |
| | I_{4X} | 0.00 | 0.00 | 1.0000 |

Medidas sobre o ajuste:

Erro Padro do resduo = 6.80, $R^2 = 0.6428$

N = 36, F = 8.10 com 8 e 36 gl, p-valor = 0

SQR = 1666

Quando comparamos as inclinaes inter-grupos (DEL e CON), observamos que, para o grupo CON, as inclinaes em todos os processos correspondem a zero, o que no se observa no grupo DEL. Observamos tambm que os α 's correspondentes para as crianas DEL nos processos SEC, SL e CF so todos significativamente diferentes de zero, sendo o processo SEC das crianas CON tambm diferente de 0. Os dois grupo se igualam somente no processo PV em que a hiptese dos alfas e dos betas serem iguais a zero no foi rejeitada. Essas desigualdades caracterizam as diferenas de representao fonolgica entre as populaes estudadas, com as crianas DEL apresentando incidncia mais acentuada dos processos SEC, SL e CF que os indivduos normais do grupo controle.

Em funo da retirada da criana do grupo controle com comportamento atpico, foi refeito o teste.t da seo 2.1.3 e o resultado foi o seguinte, $t=2.2913$, $df=17$, $p\text{-valor}=0.018$ evidenciando a diferena entre os dois grupos com relao ao tamanho do lxico.

4.4 Anlise Estatstica visando  identificao dos processos com comportamento equivalente

Alm desta formulao geral h tambm interesse juntar processos que apresentam comportamentos equivalentes conforme sugerido pela anlise exploratria precedente. Para

isso consideraremos modelos aninhados (Weisberg, 1985). Para decidirmos sobre a semelhança dos dois modelos em questão, o completo e o reduzido, com menos variáveis explicativas, será testada uma hipótese de igualdade dos parâmetros (inclinação e/ou intercepto) para os diferentes processos. A estatística de teste F estará baseada nas SQR (soma dos quadrados dos resíduos) dos modelos em questão com os seus respectivos gl (graus de liberdade):

$$F_{CR} = \frac{(SQR_R - SQR_C)/(gl_R - gl_C)}{SQR_C / gl_C} \quad (\text{eq. 2})$$

onde, o índice C significa modelo completo e o R significa reduzido, assim.

O objetivo é juntar grupos sem que aumente significativamente a SQR, sendo a medida para esta verificação a estatística F_{CR} acima. A tradução de F_{CR} foi expressa numa função feita no software “R”.

Procuraremos identificar processos para os quais haja uma grande proximidade dos β 's (inclinação), o que indicaria o mesmo comportamento para processos pertencentes ao mesmo nível de representação no léxico da criança. Testaremos então a hipótese de que os β 's correspondentes a esses processos são iguais. Caso a hipótese nula (de igualdade dos β 's) seja aceita, poderemos concluir que o modelo reduzido é mais indicado que o modelo completo para aquele particular Problema. O p-valor correspondente à estatística F_{CR} acima mencionada é o instrumento que nos permitirá decidir pela aceitação ou pela rejeição dessa hipótese.

O mesmo também foi tentado para os α 's (intercepto), porém, como estes são muito sensíveis a pequenas variações dos β 's, neste caso obtve-se pouco sucesso em termos da redução dos modelos.

4.4.1 Em Busca de um modelo adequado para o Problema I (DEL)

Sugerido pela análise dos gráficos 1 a 4 procuramos checar se as inclinações observadas dos processos SEC e SL podem ser consideradas iguais no modelo reduzido. O mesmo procedimento foi adotado para os processos CF e PV, que também apresentaram valores semelhantes para os β 's das retas correspondentes. O modelo reduzido que proporíamos é o seguinte:

$$\%P = \alpha_1 I_1 + \alpha_2 I_2 + \alpha_3 I_3 + \alpha_4 I_4 + \beta_1 (I_1 + I_2) X + \beta_2 (I_3 + I_4) X \quad (\text{eq. 3})$$

Os resultados obtidos com a função f.encaixante foram:

$F = 16.19$ com 27 e 29 graus de liberdade, correspondendo a um p-valor = $2.391943e-05$, o que indica que o modelo reduzido não é equivalente ao modelo completo. Buscamos resolver o problema agrupando os processos em dois conjuntos, um com os processos SEC e SL e outro agrupando PV e CF. Assim, os modelos completo e reduzido na primeira etapa serão respectivamente:

$$\%P = \alpha_1 I_1 + \alpha_2 I_2 + \beta_1 I_1 X + \beta_2 I_2 X \quad (\text{MC}) \quad (\text{eq. 4})$$

$$\%P = \alpha_1 I_1 + \alpha_2 I_2 + \beta_1 (I_1 + I_2) X \quad (\text{MR}) \quad (\text{eq. 5})$$

Os resultados obtidos para essa etapa foram:

$F = 1.49$ com 31 e 32 graus de liberdade, correspondendo a um p-valor = 0.23, que indica que as duas inclinações dos grupos SEC e SL são equivalentes com um p-valor bastante alto significando que o segundo modelo reduzido é equivalente ao modelo completo tendo a vantagem de ser mais parcimonioso, com menos parâmetros. Ou seja, os dois processos são equivalentes em relação ao parâmetro inclinação.

O mesmo procedimento foi adotado para os processos PV e CF. Nesse caso o modelo completo e o reduzido serão:

$$\%P = \alpha_1 I_3 + \alpha_2 I_4 + \beta_1 I_3 X + \beta_2 I_4 X \quad (\text{MC}) \quad (\text{eq. 6})$$

$$\%P = \alpha_1 I_3 + \alpha_2 I_4 + \beta_1 (I_3 + I_4) X \quad (\text{MR}) \quad (\text{eq. 7})$$

e os resultados dispostos a seguir:

F= 4.03 com 31 e 32 graus de liberdade e p-valor=0.053. Como acima, esse resultado indica que segundo o modelo reduzido proposto é equivalente ao modelo completo.

4.4.2 Em Busca de um modelo adequado para o Problema II (CON)

Sugerido pela análise apresentada na seção 2.2.1, procuramos checar igualmente se as inclinações para os processos SEC, SL, CF e PV podem ser consideradas iguais no modelo reduzido. Inicialmente juntamos os α 's para os processos SL, CF e PV, deixando o processo SEC isolado.

O modelo reduzido é o seguinte:

$$\%P = \alpha_1 I_1 + \alpha_2 (I_2 + I_3 + I_4) + \beta (I_1 + I_2 + I_3 + I_4) X \quad (\text{eq. 8})$$

Seguindo os mesmos procedimentos adotados nesse trabalho, os resultados obtidos na função f.encaixante foram os seguintes

F=17.526 com 34 e 39 graus de liberdade e p-valor=0.0000, mostrando que este modelo reduzido ainda não é equivalente ao modelo completo. Assim, testamos ainda a possibilidade de considerar o α do processo SL isoladamente. Nesse caso, o modelo reduzido é o seguinte:

$$\%P = \alpha_1 I_1 + \alpha_2 I_2 + \alpha_3 (I_3 + I_4) + \beta (I_1 + I_2 + I_3 + I_4) X \quad (\text{eq. 9})$$

Seguindo os mesmo procedimentos, procuramos testar o modelo reduzido como na equação 9 e os resultados obtidos com a função f. encaixante foram os seguintes:

F=1.101 com 34 e 38 graus de liberdade e p-valor=0.3719. Esses resultados indicam que o modelo expresso pela equação 9 é o mais adequado, o que mostra que este é equivalente ao modelo completo.

5. Conclusões

A busca dos modelos estatísticos mais adequados, no grupo DEL, para explicação da %P, a variável resposta, revelou um modelo reduzido mais parcimonioso que os modelos completos em função das regressões apresentadas em duas janelas, SEC+SL e PV+CF. No caso do grupo CON, o modelo mais parcimonioso é o que agrupa os processos da seguinte maneira SEC, SL, CF+PV para os α 's e tem em todos os processos β 's zero.

Através dos coeficientes angulares de β observamos que, para as crianças DEL, os processos relacionados à forma mais abstrata da estrutura da palavra, SEC e SL, têm relação com o tamanho do léxico. Já para o grupo CON, os coeficientes angulares de β obtidos indicam não haver relação entre os processos e o tamanho do léxico. Esses resultados podem ser indicativos de que abstrações de certos tipos estruturais fonológicos dependem de uma armazenagem maior de itens, uma vez que as crianças DEL apresentaram, em geral, um score de léxico abaixo do esperado para a idade. Já aqueles processos que envolvem a forma fonética fina não apresentaram essa inclinação (β 's zero) para o grupo DEL.

Os coeficientes lineares de α revelaram que para os processos SEC, SL, CF, o %P é maior entre as crianças DEL que no grupo controle. Observa-se que as crianças DEL apresentam, em relação à forma fonética fina das palavras, uma distância maior em relação à forma alvo, o que pode reforçar a hipótese de interdependência entre os níveis de representação fonológica, já que o aumento do léxico também demanda maior detalhamento fonético das palavras estocadas.

5. Referências Bibliográficas

- Bybee, J. The emergent lexicon. *CLS 34: The panels*. Chicago: Chicago Linguistics Society. 1998, p. 421-435.
- Beckman, M.E., Munson, B., & Edwards, J. The influence of vocabulary growth on developmental changes in types of phonological knowledge. In: Cole, J. & Hualde, J. (Eds.), *Laboratory Phonology 9*. New York: Mouton de Gruyter, 2005, p. 241-264.
- Cook, R. D. & Weisberg, S. *Residuals and Influence in Regression*. New York: Chapman and Hall, 1982.
- Dunn, L. M., Dunn, D., Capovilla, F. C. & Capovilla, A.G. S. *Teste de Vocabulário por Figuras Peabody – Versão Brasileira*. São Paulo: Casa do Psicólogo, 1997.
- Ferreira, A. de A. Avaliação Fonológica de Crianças com Distúrbio Específico da Linguagem sob a ótica da Fonologia de Uso. Dissertação (Mestrado), Universidade Federal do Rio de Janeiro (UFRJ), 2007.
- Johnson, K. Speech perception without speaker normalization: An exemplar model. In Johnson, K. & Mullennix, J. W.(eds). *Talker Variability in Speech Processing*. San Diego: Academic Press, 1997, p. 145-165.
- Lamprecht, R. (org) *Aquisição Fonológica do Português*. Porto Alegre: Artes Médicas, 2003.
- Pierrehumbert, J. Probabilistic Phonology: discrimination and robustness. In: Bod, R., Hay, J., Jannedy, S. (eds). *Probabilistic Linguistics*. Cambridge/Massachussets: MIT Press, 2003, p.177-228.
- Montgomery, D. C. & Peck, E. A. *Introduction to Linear Regression Analysis*, John Wiley and Sons, 1982.
- Weisberg, S. *Applied Linear Regression*, 2nd ed, New York: John Wiley and Sons, 1985.

TABELA A1. Lista de Palavras do Teste de Avaliação Fonológica do ABFW

| Nomeação | Transcrição | Repetição | Transcrição |
|-----------------|--------------------|------------------|--------------------|
| 1. Peteca | | 1. Palhaço | |
| 2. Bandeja | | 2. Bolsa | |
| 3. Tigela | | 3. Tesoura | |
| 4. Doce | | 4. Cadeira | |
| 5. Cortina | | 5. Galinha | |
| 6. Gato | | 6. Vassoura | |
| 7. Foguete | | 7. Cebola | |
| 8. Vinho | | 8. Xícara | |
| 9. Selo | | 9. Mesa | |
| 10. Zero | | 10. Navio | |
| 11. Chuva | | 11. Livro | |
| 12. Jacaré | | 12. Sapo | |
| 13. Machado | | 13. Tambor | |
| 14. Nata | | 14. Sapato | |
| 15. Lama | | 15. Balde | |
| 16. Ônibus | | 16. Faca | |
| 17. Pregoeiro | | 17. Fogão | |
| 18. Café | | 18. Peixe | |
| 19. Alface | | 19. Relógio | |
| 20. Raposa | | 20. Cama | |
| 21. Borracha | | 21. Anel | |
| 22. Abelha | | 22. Milho | |
| 23. Carro | | 23. Cachorro | |
| 24. Branco | | 24. Blusa | |
| 25. Travessa | | 25. Garfo | |
| 26. Droga | | 26. Trator | |
| 27. Cravo | | 27. Prato | |
| 28. Grosso | | 28. Pasta | |
| 29. Fraco | | 29. Dedo | |
| 30. Plástico | | 30. Braço | |
| 31. Bloco | | 31. Girafa | |
| 32. Clube | | 32. Zebra | |
| 33. Globo | | 33. Planta | |
| 34. Flauta | | 34. Cruz | |
| 35. Pastel | | | |
| 36. Porco | | | |
| 37. Nariz | | | |
| 38. Amor | | | |
| 39. Roupas | | | |

TABELA A2. POSSIBILIDADES DE OCORRENCIA DOS PROCESSOS FONOLÓGICOS
(ABFW, 2004)

| PROCESSOS FONOLÓGICOS | Prova de Imitação | Prova de Nomeação |
|--|-------------------|-------------------|
| 1. Redução de sílaba | 52 | 45 |
| 2. Harmonia consonantal | 52 | 45 |
| 3. Plosivação de fricativas | 22 | 23 |
| 4. Posteriorização para velar | 13 | 12 |
| 5. Posteriorização para palatal | 7 | 11 |
| 6. Frontalização de velars | 17 | 9 |
| 7. Frontalização de palatal | 6 | 5 |
| 8. Simplificação de líquida | 8 | 11 |
| 9. Simplificação de encontro consonantal | 12 | 8 |
| 10. Simplificação de consoante final | 7 | 5 |
| 11. Sonorização de plosivas | 29 | 21 |
| 12. Sonorização de fricativas | 13 | 14 |
| 13. Ensurdimento de plosivas | 17 | 14 |
| 14. Ensurdimento de fricativas | 9 | 9 |

Tabela A3. Peabody – Tabela de Idades

| IDADE EQUIVALENTE A PONTUAÇÃO | | | | | | | |
|-------------------------------|-------------|--------|-------------|--------|-------------|--------|-------------|
| Pontos | Idade Média | Pontos | Idade Média | Pontos | Idade Média | Pontos | Idade Média |
| 0-22 | <1-09 | 61 | 4-09 | 100 | 7-06 | 139 | 11-04 |
| 23 | 1-9 | 62 | 4-10 | 101 | 7-07 | 140 | 11-05 |
| 24 | 1-10 | 63 | 4-10 | 102 | 7-08 | 141 | 11-07 |
| 25 | 1-11 | 64 | 4-11 | 103 | 7-09 | 142 | 11-09 |
| 26 | 2-00 | 65 | 5-00 | 104 | 7-10 | 143 | 11-11 |
| 27 | 2-01 | 66 | 5-01 | 105 | 7-11 | 144 | 12-00 |
| 28 | 2-02 | 67 | 5-02 | 106 | 8-00 | 145 | 12-02 |
| 29 | 2-04 | 68 | 5-03 | 107 | 8-01 | 146 | 12-04 |
| 30 | 2-05 | 69 | 5-03 | 108 | 8-02 | 147 | 12-06 |
| 31 | 2-06 | 70 | 5-04 | 109 | 8-03 | 148 | 12-08 |
| 32 | 2-07 | 71 | 5-05 | 110 | 8-04 | 149 | 12-11 |
| 33 | 2-08 | 72 | 5-06 | 111 | 8-05 | 150 | 13-01 |
| 34 | 2-09 | 73 | 5-07 | 112 | 8-06 | 151 | 13-03 |
| 35 | 2-10 | 74 | 5-08 | 113 | 8-07 | 152 | 13-05 |
| 36 | 2-11 | 75 | 5-08 | 114 | 8-08 | 153 | 13-08 |
| 37 | 3-00 | 76 | 5-09 | 115 | 8-09 | 154 | 13-10 |
| 38 | 3-00 | 77 | 5-10 | 116 | 8-10 | 155 | 14-01 |
| 39 | 3-01 | 78 | 5-11 | 117 | 8-11 | 156 | 14-03 |
| 40 | 3-02 | 79 | 6-00 | 118 | 9-00 | 157 | 14-06 |
| 41 | 3-03 | 80 | 6-01 | 119 | 9-01 | 158 | 14-09 |
| 42 | 3-04 | 81 | 6-01 | 120 | 9-02 | 159 | 15-00 |
| 43 | 3-05 | 82 | 6-02 | 121 | 9-03 | 160 | 15-03 |
| 44 | 3-06 | 83 | 6-03 | 122 | 9-04 | 161 | 15-06 |
| 45 | 3-07 | 84 | 6-04 | 123 | 9-06 | 162 | 15-09 |
| 46 | 3-08 | 85 | 6-05 | 124 | 9-07 | 163 | 16-01 |
| 47 | 3-09 | 86 | 6-06 | 125 | 9-08 | 164 | 16-04 |
| 48 | 3-10 | 87 | 6-06 | 126 | 9-09 | 165 | 16-08 |
| 49 | 3-11 | 88 | 6-07 | 127 | 9-11 | 166 | 17-00 |
| 50 | 3-11 | 89 | 6-08 | 128 | 10-00 | 167 | 17-05 |
| 51 | 4-00 | 90 | 6-09 | 129 | 10-01 | 168 | 17-09 |
| 52 | 4-01 | 91 | 6-10 | 130 | 10-03 | 169 | 18-02 |
| 53 | 4-02 | 92 | 6-11 | 131 | 10-04 | 170 | 18-07 |
| 54 | 4-03 | 93 | 7-00 | 132 | 10-05 | 171 | 19-01 |
| 55 | 4-04 | 94 | 7-01 | 133 | 10-06 | 172 | 19-09 |
| 56 | 4-05 | 95 | 7-02 | 134 | 10-08 | 173 | 20-04 |
| 57 | 4-05 | 96 | 7-02 | 135 | 10-09 | 174 | 20-11 |
| 58 | 4-06 | 97 | 7-03 | 136 | 10-11 | 175 | 21-10 |
| 59 | 4-07 | 98 | 7-04 | 137 | 11-00 | 176 | 22+ |
| 60 | 4-08 | 99 | 7-05 | 138 | 11-02 | 177 | 22+ |

¹ Os resultados apresentados neste artigo fazem parte do trabalho de dissertação de mestrado de Ferreira (2007) do Programa de Pós Graduação em Linguística da UFRJ, sob a orientação de Christina Abreu Gomes e co-orientação de Gastão Coelho Gomes.