

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE FILOSOFIA E CIÊNCIAS HUMANAS

INVESTIGAÇÃO PSICOMÉTRICA DE UMA TAREFA DE LEITURA EM VOZ
ALTA DE PALAVRAS ISOLADAS

Patrícia Silva Lúcio

Dissertação de Mestrado

Belo Horizonte, maio de 2008

**INVESTIGAÇÃO PSICOMÉTRICA DE UMA TAREFA DE LEITURA EM VOZ
ALTA DE PALAVRAS ISOLADAS**

Patrícia Silva Lúcio

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Psicologia, sob a orientação da Prof^a. Dr.^a. Ângela Maria Vieira Pinheiro e Co-orientação da Prof^a. Dr.^a. Elizabeth do Nascimento

Universidade Federal de Minas Gerais
Programa de Pós-Graduação em Psicologia

Belo Horizonte, maio de 2008

150 Lúcio, Patrícia Silva
L937i Investigação psicométrica de uma tarefa de leitura em voz alta
2008 de palavras isoladas / Patrícia Silva Lúcio. – 2008.

179 f. : il.

Orientador: Ângela Maria Vieira Pinheiro

Co-orientadora: Elizabeth do Nascimento

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Faculdade de Filosofia e Ciências Humanas.

1. Psicologia – Teses. 3. Psicologia do Desenvolvimento – Teses 3. Psicometria – Teses. 4. Leitura - Teses I. Pinheiro, Ângela Maria Vieira. II. Nascimento, Elizabeth. III. Universidade Federal de Minas Gerais. Faculdade de Filosofia e Ciências Humanas. III. Título

A todas as pessoas que amo,
dedico esta dissertação.

AGRADECIMENTOS

Gostaria não somente de agradecer, como também de dedicar esta dissertação, a todas as instituições e pessoas que abaixo relaciono:

À Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG), que financiou o projeto do qual esta dissertação faz parte. Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), que me concedeu uma bolsa de mestrado durante dois anos de execução desse projeto.

Às minhas orientadoras, Dra. Ângela Maria Vieira Pinheiro e Dra. Elizabeth do Nascimento, que me deram apoio e orientação durante todo o percurso da execução desse trabalho.

Ao professor Dr. José Francisco Soares, pela solicitude que sempre me recebeu e pelas sugestões de análise de dados fornecidas na qualificação e também em outros encontros. Ao professor Dr. Rui Rothe-Neves, pelas instruções informais de lingüística que tanto me auxiliaram na compreensão do fenômeno com o qual estou lidando.

Aos colegas estagiários que passaram por mim durante todo o período em que participo das atividades do projeto. Dedico esta dissertação a toda a equipe de coleta e de correção dos dados, bolsistas e voluntários, principalmente as estagiárias Gilcência Goulart, Patrícia Curto e Renata Alves. Em especial, agradeço aos bolsistas Lucas Braga e Ricardo Moura, meus companheiros de dia-a-dia que, além de contribuírem diretamente para o andamento do projeto, tornaram todas as dificuldades mais leves devido ao bom humor com que sempre trabalhavam.

Agradeço aos amigos Carolina Rezende, Daniel Silva, Fábio Diniz, Marmília Lambertucci e Ricardo Napoleão que me acompanharam e incentivaram durante todo o período de graduação e de pós-graduação. Em especial, agradeço ao amigo Hudson de

Carvalho, companheiro, confidente e apoio intelectual e emocional; sem você tudo teria sido mais difícil.

À minha família, que sempre me incentivou nos estudos, especialmente a meu pai, que dedicou sua vida para que este objetivo se concretizasse. Também gostaria de agradecer em particular à minha irmã Walquíria, que me iniciou no processo de leitura e de escrita, assunto que hoje trato nesta dissertação.

Por último, gostaria de dar meus sinceros agradecimentos a todas as escolas que nos abriram as portas, aos pais que nos concederam seus filhos e às crianças que participaram desta pesquisa; sem a esta colaboração, nada disto teria sido realizado.

SUMÁRIO

LISTA DE TABELAS	09
LISTA DE FIGURAS	11
RESUMO	12
ABSTRACT	14
1. INTRODUÇÃO.....	16
2. DEFINIÇÃO DO CONSTRUTO E REVISÃO TEÓRICA.....	20
2.1 Psicologia cognitiva e abordagem do processamento de informação	20
2.2 O modelo de dupla-rota de leitura de palavras isoladas	29
2.3 Estudos sobre a tarefa de leitura em voz na em língua portuguesa.....	38
2.4 Avaliação cognitiva com o subteste de leitura do TDE.....	50
3. FASES DE CONSTRUÇÃO DE UM TESTE PSICOLÓGICO	60
3.1 Etapas na construção de testes.....	60
3.2 Procedimentos anteriores.....	63
3.3 Proposta de trabalho	66
4. CLASSIFICAÇÃO DAS PALAVRAS QUANTO AO NÍVEL DE REGULARIDADE E HIPÓTESES GERADAS A PARTIR DA CLASSIFICAÇÃO ..	73
5. OBJETIVOS.....	77
5.1 Objetivo geral	77
5.2 Objetivos específicos.....	77
5.2.1 Análise de itens.....	77
5.2.2 Efeito de fadiga.....	78
5.2.3 Efeitos na Leitura.....	78

5.2.4 Investigação da fidedignidade a partir de simulação de amostra de itens-candidatos	78
6. MÉTODO	79
6.1 Características do estudo	79
6.2 Amostra	79
6.3 Instrumentos	82
6.4 Procedimento de coleta de dados.....	85
6.5 Procedimento de análise estatística	86
7. RESULTADOS	87
7.1 Análise de itens.....	87
7.1.1 Procedimento	87
7.1.2 Resultados.....	87
7.2 Efeito de treino e fadiga.....	94
7.2.1 Procedimento	94
7.2.2 Resultados.....	95
7.3 Efeitos na leitura.....	99
7.3.1 Procedimento	99
7.3.2 Resultados.....	102
7.3.2.1 Modelo de tempo de reação.....	102
7.3.2.2 Modelo de precisão.....	104
7.4 Simulação de seleção de amostra final de itens a partir da amostra de itens-candidatos: procedimentos psicométrico	107
7.4.1 Análise de itens	108
7.4.1.1 Procedimentos e resultados	108
7.4.2 Consistência interna e temporal.....	112

7.4.2.1 Procedimento	112
7.4.2.2 Resultados.....	113
7.4.2.2.1 Consistência interna.....	113
7.4.2.2.2 Estabilidade temporal	115
7.4.3 Concordância entre avaliadores.....	116
7.4.3.1 Procedimento	116
7.4.3.2 Resultados.....	117
8. CONSIDERAÇÕES FINAIS	118
9. REFERÊNCIAS	134
10. APÊNDICES	149
APÊNDICE A. Cópia da aprovação do projeto pelo Comitê de Ética em Pesquisa com Seres Humanos (COEP) da UFMG	149
APÊNDICE B. Modelo de termo de consentimento da escola	150
APÊNDICE C. Critérios de participação na pesquisa.....	151
APÊNDICE D. Modelo de termo de consentimento livre e esclarecido entregue aos pais dos participantes (frente e verso)	152
APÊNDICE E. Questionário de auto-avaliação (Modelo de entrevista com as crianças).....	154
APÊNDICE F. Escala de Avaliação de Competência em Leitura pelo Professor para Crianças de 1ª série (APÊNDICE F-1) e de 2ª à 4ª séries (APÊNDICE F-2)	155
APÊNDICE G. Valores dos índices de dificuldade, discriminação e correlação item-total apresentados pelas séries e pela amostra geral nas palavras do teste (distribuídas em função da classificação de regularidade para a leitura e para a escrita, do número de letras e da Lista de aplicação)	158
APÊNDICE H. Classificação das palavras adotada no presente estudo	178

LISTA DE TABELAS

Tabela 1. Resultados para o efeito de regularidade na leitura em diversos estudos com a língua portuguesa, em função da direção da classificação de regularidade das palavras, e estatísticas por eles utilizadas	48
Tabela 2. Valores das Médias e Desvios Padrões (D.P.) por série para o subteste de leitura do TDE (dados da padronização)	52
Tabela 3. Resumo das conclusões de trabalhos que relataram os desempenhos de participantes no subteste de leitura do TDE	55
Tabela 4. Etapas de construção de um teste descritas por Adánez (1999)	61
Tabela 5. Exemplos de palavras em cada categoria de classificação para a leitura, em função da classificação para a escrita	75
Tabela 6. Perfil socioeconômico das escolas participantes do estudo	81
Tabela 7. Quantidade de alunos por série e escola na prova que leram o banco de itens testado.....	82
Tabela 8. Valores médios da dificuldade, discriminação e correlação item-total nas Listas 1 e 2 e no total para cada série separadamente e toda a amostra	89
Tabela 9. Distribuição da proporção de palavras (%) em cada nível de dificuldade, discriminação e correlação item-total.....	90
Tabela 10. Itens selecionados por meio da análise, separados em função do nível de dificuldade e da lista de aplicação	91
Tabela 11. Porcentagens de palavras em cada classificação de regularidade para a leitura presentes para os níveis de dificuldade especificados	93

Tabela 12. Valores de média (M) e desvio padrão (D.P.) para a proporção de acertos (%) e para o tempo de reação (mseg) nas três partes de cada uma das listas de palavras e para todo o corpo de itens aplicados às crianças da amostra para cada série separadamente e no total.....	96
Tabela 13. Níveis de significância para as comparações entre os três terços das Listas 1 e 2, valores do tamanho do efeito (η^2), do poder observado (β) e de p para as comparações significativas encontradas no teste de Bonferroni, para cada série e no geral	98
Tabela 14. Equação de estimação do tempo de reação dos alunos.....	103
Tabela 15. Equação de estimação da precisão dos alunos.....	105
Tabela 16. Palavras selecionadas a partir da simulação de análise dos itens-candidatos distribuídos em termos de dificuldade pelo critério de Pasquali (1996)	110
Tabela 17. Valores dos coeficientes alfa de Cronbach para as Formas de Exemplo 1 e 2 e valores de correlação e de Spearman-Brown para as metades de cada uma dessas formas	114
Tabela 18. Valores de correlação de Pearson para a precisão das respostas dos participantes nas palavras selecionadas das Listas 1 e 2 em cada intervalo de tempo selecionado	115
Tabela 19. Descrição das características dos participantes que fizeram parte da análise de concordância entre avaliadores.....	117

LISTA DE FIGURAS

Figura 1. Estágios geralmente presentes em um modelo de processamento de informações	25
Figura 2. Estrutura básica do modelo de dupla-rota em cascata para o reconhecimento de palavras e a leitura em voz alta.....	30
Figura 3. Exemplo de <i>output</i> do programa <i>Soundforge 4.5</i> , utilizado para a gravação da leitura das crianças	83
Figura 4. Distribuição de tempo de reação para as respostas dos participantes da amostra	102

RESUMO

Uma das formas pelas quais a psicologia cognitiva estuda o reconhecimento de palavras é por meio da tarefa de leitura em voz alta de palavras isoladas, que pode ser utilizada para testar os componentes do modelo de dupla-rota. Este estudo conduziu a investigação psicométrica de um banco de itens composto de 323 palavras isoladas e de baixa frequência de ocorrência. Um grupo de crianças de 1ª à 4ª série do Ensino Fundamental da Rede de Ensino de Belo Horizonte (N = 333) leu individualmente as palavras isoladas apresentadas por meio de microcomputadores. Medidas de tempo de reação (TR) e de precisão foram coletadas. Os resultados mostraram que, de uma maneira geral, os itens não impuseram muitas dificuldades aos leitores. Dos itens que causaram dificuldades para a amostra ($p < 0,40$), todos eram irregulares para a leitura. A análise de itens selecionou 112 itens-candidatos com níveis adequados de dificuldade e discriminação. Em relação aos efeitos na leitura do banco de itens, houve um efeito de fadiga no TR, mas não foi confirmado um efeito de treino na precisão. Uma análise de regressão hierárquica (modelo GEE) mostrou a existência de um forte efeito de regularidade grafema-fonema e fonema-grafema na leitura e de um impacto do número de categorias utilizadas na análise na emergência desse efeito. Uma análise simulada de redução de itens mostrou que o mesmo padrão de regularidade e extensão das palavras selecionadas pela análise de itens foi observado para as Formas de Exemplo obtidas. Os resultados alcançados a partir do método das formas alternativas com intervalo mostraram que ambas as Formas de Exemplo apresentaram estabilidade temporal satisfatória em todas as faixas de tempo investigadas. A análise de fidedignidade do avaliador mostrou que houve concordância entre os juízes quanto à correção dos protocolos das Formas de Exemplo. Os resultados obtidos pela análise da dificuldade e

discriminação dos itens foram discutidos à luz das diferenças entre as pesquisas de representação do construto e as pesquisas de intervalo nomotético para a validação de construto propostas por Embretson (1983). Concluiu-se que os indicadores psicométricos produzidos no presente trabalho corroboraram para a validade de construto da tarefa em questão.

Palavras-chave: Análise de itens; reconhecimento de palavras; efeito de regularidade grafema-fonema e fonema-grafema; fidedignidade; validade.

ABSTRACT

The reading word aloud task is generally used in Cognitive Psychology to evaluate word recognition and to test the components of dual-route models. The present study reports the psychometrical inquiry of a word reading aloud task composed by 323 low frequency words. A group of 1st to 4th grade school children from Belo Horizonte City (N = 333) read aloud the single words that were presented in a computer screen. Measures of reaction time (RT) and error scores were collected. The results indicated that in general, items were not difficult to the sample. The items that cause some difficulty to the sample ($p < .40$) were irregular ones. The Item analysis selected 112 potential items that presented satisfactory levels of discrimination and difficulty. There was a fatigue effect in RT, but was not demonstrated a training effect in error scores. The Generalized Estimating Equations (GEE) method exhibit the grapheme-phoneme and phoneme-grapheme regularity effect in reading and also showed an impact of the number of categories of regularity in this effect. A simulated item analysis used for data reduction showed that the pattern of the items obtained was similar to that found in the item analysis, in terms of levels of regularity and size of the words. The item analysis simulation produced two versions of a task-example. The coefficients obtained by the method of alternate-form administrated in separated occasions indicated a satisfactory level of temporal stability for both tasks-examples. There was also reliability between the judges in conferring scores of those tasks. The results obtained for the item difficulty and discrimination analysis are discussed in relation to the differences between the two types of construct validity research proposed by Embretson (1983) – the construct representation research and the nomothetic span research. We concluded

that the psychometrical indicators produced in the present study contributed to construct validity under inquire.

Key words: Item analysis; word recognition; grapheme-phoneme and phoneme-grapheme regularity effect; reliability; validity.

1 INTRODUÇÃO

Em psicologia, a avaliação e o diagnóstico de competências psicológicas específicas como a inteligência e as habilidades de leitura e de escrita, necessitam da utilização de instrumentos que sejam válidos, confiáveis e que apresentem normas para a população a que se destina. No Brasil, a carência de sistematização da maioria dos instrumentos de avaliação psicológica conduziu a uma resolução por parte do Conselho Federal de Psicologia (CFP, 2003) que define e regulamenta o uso, a elaboração e a comercialização de testes psicológicos. O artigo quatro dessa resolução é dedicado ao estabelecimento dos requisitos mínimos necessários para o reconhecimento oficial de instrumentos que buscam estabelecer uma avaliação objetiva das propriedades psicológicas. Dentre outras especificações, a resolução determina que os instrumentos devam passar pelos procedimentos de validação, precisão e estabelecimento de normas, além de apresentarem relatos de dados empíricos sobre as propriedades psicométricas de seus itens. Por fim, o CFP ressalta a necessidade de especificação da fundamentação teórica do instrumento, com ênfase na definição do construto, dos propósitos do instrumento e dos principais contextos para os quais ele foi desenvolvido.

A preocupação recente do CFP com a condução e o relato de pesquisas que demonstrem a cientificidade dos instrumentos de avaliação psicológica torna nítida a importância das iniciativas que são voltadas para esta finalidade em nosso contexto. É nesse sentido que se justifica a relevância do presente trabalho, cujo objetivo central é relatar a investigação psicométrica de uma tarefa de leitura em voz alta de palavras isoladas. Esta tarefa constituirá parte de uma bateria computadorizada de avaliação cognitiva da leitura e da escrita para crianças cursando as séries iniciais do Ensino Fundamental denominada “Avaliação das Capacidades Cognitivas de Leitura e Escrita –

AVACLE”, que está sendo desenvolvida pelo grupo de pesquisa da professora Ângela Maria Vieira Pinheiro do Departamento de Psicologia da Universidade Federal de Minas Gerais (UFMG).

A investigação psicométrica da tarefa de leitura em voz alta em questão terá como primeiro passo fornecer subsídios para a validade do banco de itens que a compõe. Isso será feito por meio de uma análise que selecionará os itens que apresentarem níveis adequados de dificuldade e de discriminação para a amostra de crianças do presente estudo. Os itens assim selecionados (chamados aqui de itens-candidatos) serão, em uma segunda etapa, aplicados em uma nova amostra de participantes a fim de se selecionar o conjunto final de itens que irão compor o teste de leitura de palavras isoladas que se pretende criar. Este será um procedimento validação cruzada a ser conduzido em outro estudo (assunto que não será tratado nessa dissertação).

Os outros procedimentos sob investigação estarão centrados na coleta de informações acerca da validade de construto da tarefa, o que será feito por meio da verificação de efeitos na leitura (os quais dizem respeito aos pressupostos subjacentes à execução da tarefa, conforme veremos) e de uma investigação preliminar da fidedignidade. Todos esses procedimentos, assim reunidos, certamente contribuirão para uma maior compreensão do fenômeno que será investigado (isto é, o reconhecimento palavras por crianças) ao constituir uma importante fonte de dados empíricos cuja magnitude não está presente em nenhum estudo anteriormente conduzido no Brasil.

Esta dissertação está organizada em seções, as quais buscam fornecer uma descrição do construto em que se baseia a tarefa destinada a avaliá-lo, assim como das diretrizes utilizadas para construí-la. O construto sob investigação, a saber, a habilidade de ler palavras fora de contexto, é estudado na psicologia cognitiva pela área de

reconhecimento de palavras. Esta área tem na abordagem do processamento de informação, nos modelos verbais e computacionais e nos achados clínicos da neuropsicologia cognitiva os principais pontos para a construção da teoria sobre como os indivíduos reconhecem as palavras, desde a análise sensorial do estímulo impresso pelo sistema visual (*input*) até o momento em que uma resposta (*output*) acústica é produzida. Nesse sentido, primeiramente (seção 2.1) serão descritos os principais pressupostos da psicologia cognitiva e as características centrais da abordagem do processamento de informação, que constituem as bases teóricas que sustentam o trabalho. Em seguida (seção 2.2), será apresentado o modelo de dupla-rota da leitura, que foi desenvolvido inicialmente para a língua inglesa, mas tem demonstrado aplicações em diversas línguas, inclusive o português.

Na seção 2.3 será feita uma revisão da produção científica sobre o tema *avaliação cognitiva da leitura no contexto brasileiro*, buscando avaliar as características metodológicas dos estudos aqui produzidos e a aplicação do modelo de dupla rota ao contexto da língua portuguesa, em especial o português brasileiro. Na seção 2.4 será feita uma revisão sobre a investigação científica da leitura conduzida por trabalhos que utilizaram o subteste de leitura do Teste de Desempenho Escolar (TDE) de Stein (1994). Esta descrição se justifica pelo fato de este subteste ser o único instrumento de avaliação de leitura em voz alta de palavras isoladas que apresenta normas para a nossa realidade, além de evidências psicométricas de validade e precisão.

Na terceira seção serão descritas as fases de construção de um teste psicológico que são comumente apresentadas pelos pesquisadores e, a partir dessas etapas, serão descritos todos os procedimentos anteriormente efetuados para a construção da tarefa cuja análise de itens será aqui desenvolvida. Ao final desta seção, será descrito o plano de análise para as etapas que foram efetuadas durante a execução deste trabalho.

Na quarta seção parte-se para a descrição da classificação de regularidade das palavras adotada nesta pesquisa e, a partir dessa classificação, são expostas as hipóteses da dificuldade das palavras.

Por fim, são expostos com um maior grau de especificidade os objetivos gerais e os específicos da pesquisa, seguidos do método e da apresentação dos resultados. As considerações finais ou conclusões do estudo estão dispostas na última seção.

2 DEFINIÇÃO DO CONSTRUTO E REVISÃO TEÓRICA

2.1 Psicologia cognitiva e abordagem do processamento de informação

A psicologia cognitiva se interessa pelas habilidades mentais que estão relacionadas à aquisição do conhecimento como, por exemplo, a memória, a atenção e a linguagem. Segundo a definição clássica de Ulrick Neisser (1967), a psicologia cognitiva se refere a todos os processos pelos quais o *input* sensorial é transformado, reduzido, elaborado, armazenado, recuperado e utilizado. Por esta definição, podemos ter uma noção de como os psicólogos cognitivos estudam a cognição. Assim, a referência a um *input* sensorial implica que a cognição começa com o contato com o mundo externo. A transformação do *input* sensorial implica que a representação não é registrada de uma maneira passiva, mas ocorre por meio uma construção ativa que pode envolver tanto a redução (descarte de informações irrelevantes pelo filtro da atenção) quanto a elaboração (adição de novas informações e integração com outras informações anteriormente armazenadas). A separação entre os processos de armazenamento e de recuperação deixa claro que, para que uma informação seja recordada, não basta apenas que esteja armazenada, sendo que o esquecimento pode ser devido tanto a um processo quanto a outro. O último estágio implica que, após processada, a informação deve ser utilizada pelo indivíduo em um processo ativo de integração com as experiências anteriores, como na resolução de problemas ou na tomada de decisões (Reed, 1982).

A definição proposta por Neisser (1967) evidencia alguns pontos fundamentais para o estudo da cognição, que constituem os pressupostos da abordagem do processamento da informação mais comumente aceitos. O primeiro deles é a noção básica de que podemos traçar o progresso da informação através do sistema cognitivo

desde o ponto em que o estímulo é decodificado até a execução da resposta (Massaro & Cowan, 1993). A ideia de decodificação do estímulo implica que os dados que estão no ambiente devem ser transformados em um tipo de representação que possa ser trabalhada pelo sistema cognitivo, passando a constituir uma forma de conhecimento ou informação. As representações são estados do sistema nos quais a informação está encaixada, enquanto que as operações utilizadas para transformar as representações são chamadas de processos (Palmer & Kimchi, 1986; citados por Massaro & Cowan, 1993). Outra característica é que o sistema cognitivo é composto de séries sequenciais de estágios de processamento, em que o *output* de um processamento serve de *input* para o outro (Miller, 1982). Nos estágios também ocorrem operações sobre as representações, sendo que a informação flui dos estágios iniciais aos estágios posteriores com o decorrer do tempo.

O estudo dos estágios de processamento para a execução de diferentes tarefas cognitivas foi impulsionado pela retomada do tempo de reação como uma medida dos processos mentais (Galera & Lopes, 1995; citado por Rodrigues & Lopes, 2002). A pressuposição de que cada estágio de processamento cognitivo leva tempo para ser executado levou à utilização do tempo de reação como forma de inferir os processos mentais que ocorrem durante o desempenho de tarefas. Nesse sentido, podemos ressaltar o papel do método da subtração proposto por Donders no século XIX, que foi muito utilizado pelos psicólogos cognitivos experimentais e ofereceu *insights* importantes sobre o processamento cognitivo em diversas áreas. No entanto, tal método tem a limitação de necessitar de uma conceituação precisa dos componentes processuais envolvidos na realização da tarefa e de pressupor que o processamento da mesma é descontínuo e serial. Isso implica admitir que a série de estágios que compõe a tarefa é invariável e que cada um desses estágios deve ser finalizado antes que o próximo se

inicie (Eysenck & Keane, 1994; Massaro & Cowan, 1993; Miller, 1982). Nesse sentido, o método não abrange as tarefas que se enquadram no modelo em cascata, que veremos mais adiante. Assim, apesar de a utilidade do método da subtração ser limitado a um conjunto reduzido de tarefas mais simples que se enquadram no modelo serial clássico, o uso do tempo de reação em si permaneceu como uma fonte importante de informação sobre os processos cognitivos humanos e continua a ser uma das principais variáveis dependentes utilizadas na psicologia.

Segundo Massaro e Cowan (1993), existem evidências de que o processamento da informação ocorra em estágios e questões importantes para a abordagem emergem desse fato. Dado um único estímulo, e para cada estágio de processamento, além das representações de *input* e de *output* podem ser distinguidos os processos de transformação e de transmissão da informação (respectivamente, a modificação da representação de um estímulo em uma nova forma de representação e o envio da informação através do sistema). Essas representações e esses processos podem ser caracterizados como discretos ou contínuos e, se múltiplos códigos de representações coexistirem em um determinado estágio, estes podem ser processados de maneira serial ou paralela. A divergência entre processamento discreto e contínuo é relevante para determinar como um único item-estímulo é processado em cada estágio, enquanto que a dicotomia entre processamento serial e paralelo refere-se ao processamento de um estímulo como um todo dentro de cada estágio. No primeiro caso, e segundo Miller (1988, 1990, citado por Massaro & Cowan, 1993), a questão é saber (a) se o tipo de informação que chega a um estágio (ou dele sai), (b) se a forma de transformação que ocorre em um determinado estágio e (c) se a informação que é transmitida para o próximo estágio; ocorre de uma maneira discreta ou contínua. No segundo caso, o

importante é determinar se o estágio pode manipular apenas um item de cada vez ou múltiplos itens simultaneamente.

As características acima expostas apresentam repercussões teóricas importantes. Primeiro, é necessário distinguir entre a abordagem do processamento de informações e as pressuposições feitas dentro de aplicações particulares desta abordagem. Assim, a transmissão discreta da informação é aplicável à metodologia dos fatores aditivos de Sternberg (1969), enquanto que o modelo de dupla-rota em cascata (o referencial teórico para a presente pesquisa) pressupõe uma transmissão do tipo contínua da informação. Ambos os modelos teóricos fazem parte da metateoria do processamento de informação (respectivamente, para o estudo da memória e da linguagem), mas apresentam pressuposições distintas de como esse processamento ocorre em determinado contexto.

A segunda repercussão refere-se à avaliação dos modelos. Uma das vantagens da abordagem do processamento de informação é a possibilidade de se identificar, dentre vários modelos concorrentes, o modelo mais adequado utilizando-se, para tanto, manipulações experimentais específicas e comparando-se os resultados gerados com as previsões que são feitas por esses modelos. Isso implica que diferentes visões acerca da forma de processamento do estímulo podem coexistir e explicar de maneira eficiente um mesmo fenômeno, havendo a necessidade de se refinar as análises e os procedimentos metodológicos a fim de se produzir resultados que nos faça decidir por uma visão ou por outra. Desse modo, a questão não é se o processamento é discreto ou contínuo, ou se é serial ou paralelo, mas quando que ele se manifestará de uma forma ou de outra. Massaro e Cowan (1993) chamam, ainda, a atenção para o fato de que pode haver uma alteração na forma de processamento do estímulo em função de mudanças na demanda da tarefa (efeito das estratégias) – ou seja, um tipo informação tipicamente transmitida

de maneira contínua pode ser transmitida de forma discreta em função das demandas da tarefa e vice-versa.

No estudo da cognição humana, os psicólogos cognitivos utilizam modelos para representar as diversas teorias existentes sobre os processos mentais. Um modelo é “uma representação física de uma teoria que relaciona dada teoria a uma situação em particular”, permitindo que predições bem detalhadas sejam feitas (Eysenck & Keane, 1994, p. 33). Desta forma, os modelos de processamento de informação podem ser definidos como descrições teóricas de uma seqüência de estágios através dos quais o processamento cognitivo ocorre (Massaro & Cowan, 1993). Há dois tipos de modelagem utilizados pela pesquisa psicológica: os modelos verbais e os modelos computacionais. Os modelos verbais são descrições do funcionamento mental baseadas no desempenho de indivíduos normais e com lesões cerebrais no desempenho de tarefas experimentais. Com base nesses achados, essas descrições são transformadas em modelos de processamento da informação que são pictoricamente representadas pela notação de caixas e setas. Nas caixas, são representados os diferentes estágios do processamento e, nas setas, o fluxo da informação. Os estágios (ou caixas) são arranjados em uma ordem temporal e a informação pode fluir nas direções apontadas pelas pontas das setas, que podem ser unidirecionais ou bidirecionais, estas últimas indicando que os estágios anteriores podem ser influenciados por informações posteriormente processadas. Conforme vimos, no início da criação do modelo de estágios acreditava-se que a informação só passaria para o estágio seguinte quando o processamento em um estágio anterior findasse, constituindo os modelos baseados em limiar de ativação. Atualmente, admite-se que informações parciais dos estágios anteriores podem ser passadas para os estágios seguintes antes que o processo finalize o que é conhecido como modelo em cascata (Coltheart, Rastle, Perry, Langdon & Ziegler,

2001). A Figura 1 apresenta um modelo de caixa-e-setas com estágios comumente presentes nos modelos de processamento da informação.

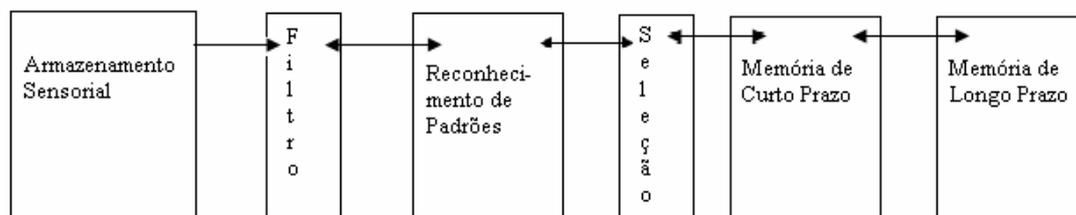


Figura 1. Estágios geralmente presentes em um modelo de processamento de informações (reproduzido de Reed (1982) com tradução).

A outra forma de modelagem utilizada pelos pesquisadores da cognição constitui os modelos computacionais, que correspondem a simulações computadorizadas de funções e de processos cognitivos humanos. Os modelos computacionais representam formas importantes de expressão de teorias sobre a cognição e apresentam a vantagem de poderem ser utilizados para testar, para aprimorar ou para corrigir as teorias psicológicas propostas pelos modelos verbais (Coltheart et al., 2001; Lopes, Lopes & Teixeira, 2004). Além destas, os modelos computacionais apresentam a vantagem de permitir a operacionalização de muitos termos mentalistas utilizados pela psicologia cognitiva (Crowter-Heyck, 1999).

Os modelos teóricos propostos pela abordagem do processamento de informação têm recebido constantes modificações desde seu surgimento com a “revolução cognitiva” nos anos 50 (Baars, 1986, citado por Lopes et al., 2004; Eysenck & Keane, 1994; Gardner, 2003; Miller, 2003), mas o aparecimento de outras abordagens nos últimos anos levou Lopes et al. (2004) a concluir que atualmente o paradigma está

enfrentando uma crise, particularmente no que se refere ao trabalho desenvolvido pela psicologia cognitiva experimental. Para os autores, a abordagem do processamento de informação se encontra em um estágio que necessita de revisões face ao desenvolvimento de três abordagens alternativas (ou complementares), a saber, a abordagem computacional da mente (representada pelos modelos conexionistas), a psicologia ecológica e as neurociências. A principal dificuldade da abordagem em se consolidar como um sistema psicológico forte seria de ordem metodológica, pois, ao mesmo tempo em que surgiu como uma proposta de ser uma ciência da mente, em substituição ao behaviorismo, manteve a experimentação como seu principal método de investigação, o mesmo utilizado pela teoria rival. Assim, essas abordagens alternativas seriam capazes de fornecer elementos necessários para se compreender a cognição tanto por meio da introdução de um novo método, quanto pelo esforço em compreender o papel do ambiente na construção do comportamento. No entanto, embora os autores afirmem que tais abordagens coloquem desafios importantes para o paradigma do processamento de informação, eles reconhecem que algumas delas não constituem seus substitutos potenciais, mas sim trazem contribuições complementares e que possivelmente preencherão algumas das lacunas deixadas por ele.

A crítica de Lopes e colaboradores quanto ao uso exclusivo do método experimental como forma de explicar a cognição humana é pertinente, mas talvez pouco adequada, uma vez que quase nenhum pesquisador na atualidade utiliza uma abordagem exclusiva para o estudo da cognição. Isso fica evidente quando Eysenck e Keane (1994), no início dos anos 90, especulam se os psicólogos cognitivos experimentais seriam “uma espécie ameaçada de extinção” (p. 15). Quanto à crítica ao processamento da informação, o problema é a confusão que os autores fazem entre a abordagem de um modo geral com método que os psicólogos experimentais utilizam dentro dessa

abordagem. Assim, uma crítica que deveria recair sobre o uso exclusivo do método experimental passa a se referir à abordagem do processamento de informação como um todo. Além disso, um dos principais pontos de argumentação dos autores é que o surgimento de abordagens alternativas coloca novos desafios à abordagem. No entanto, conforme afirmam Massaro e Cowan (1993), o paradigma do processamento de informação é uma metateoria geral para a investigação psicológica e as outras metateorias (representadas pelo sistema de símbolos físicos, pelo conexionismo, pela modularidade e pelo realismo ecológico) constituem mais variações dentro dessa arquitetura do que propriamente alternativas, uma vez que elas permanecem compartilhando características importantes com a abordagem geral. Por exemplo, as três primeiras metateorias supracitadas apresentam a característica comum da busca de explicação para os processos e representações subjacentes ao comportamento, enquanto que a abordagem ecológica, como a dos neo-Gibsonianos, apesar de se interessar pela influência das propriedades invariantes do ambiente nas respostas perceptuais, ainda não conseguem explicar o processo inverso – ou seja, a influência do sistema perceptivo sobre a situação ambiental – sem apelar para a descrição desses mesmos processos mentais internos.

Por fim, as neurociências, representadas pelo conjunto de conhecimentos sobre o cérebro do ponto de vista bioquímico e pelo conjunto de técnicas emergentes de mapeamento cerebral, apesar de constituírem um avanço para o estudo da cognição, não são incompatíveis com a abordagem do processamento de informação e não constituem uma revolução metodológica. Por exemplo, a interpretação da área ativada pelas técnicas de neuroimagem funcional tem por base o método da subtração que se baseia fortemente na idéia de que o processamento cognitivo ocorre em estágios seqüenciais e independentes, conforme vimos. Assim, para que a interpretação dos resultados gerados

por tais técnicas não esteja comprometido, é necessário que se tenha uma conceituação clara dos componentes envolvidos na execução da tarefa, o que é possibilitado pela presença de um modelo cognitivo adequado para explicá-la (Kristensen, Almeida, & Gomes, 2001).

A abordagem adotada no presente estudo tem por base o método da psicologia experimental, os pressupostos teóricos da psicologia cognitiva e reconhece a influência dos achados da neuropsicologia cognitiva. Assim, a psicologia experimental pressupõe que o uso de tarefas cujos estímulos são manipulados em termos de determinadas características é capaz de fornecer inferências acerca dos processos cognitivos que são utilizados para se executar tais tarefas e, por meio dessas inferências, os psicólogos cognitivos criam modelos sobre processamento mental. A neuropsicologia cognitiva, definida como “o estudo das relações entre a cognição e o comportamento humano e as funções mentais preservadas ou alteradas” (Kristensen et al., 2001, p. 267), testa os modelos então criados a partir do desempenho em tarefas cognitivas de indivíduos normais, daqueles que não se desenvolveram normalmente ou de indivíduos com déficits específicos ocasionados por lesões cerebrais. A influência entre a psicologia cognitiva e a neuropsicologia é, desta forma, bidirecional, sendo que a psicologia cognitiva oferece explicações coerentes sobre a cognição humana e a neuropsicologia promove provas empíricas para essas explicações, reformulando-as quando necessário. Além da importância teórica da ligação entre as disciplinas, a aplicabilidade clínica das tarefas cognitivas para o diagnóstico neuropsicológico é também evidente.

Na próxima seção, será descrito o modelo no qual se baseia a construção da tarefa proposta nesse trabalho, o modelo de dupla-rota da leitura. Este surgiu como um modelo verbal nos primeiros trabalhos de John Morton na década de 60 na Inglaterra e a

partir dele foram desenvolvidos modelos computacionais, como o proposto pelo grupo de Max Coltheart na Austrália.

2.2 O modelo de dupla-rota de leitura de palavras isoladas

Uma área da avaliação cognitiva da leitura que tem recebido bastante respaldo teórico e empírico nos mais diversos idiomas é a área de reconhecimento de palavras, também conhecida como leitura em voz alta de palavras isoladas. Trata-se da leitura de palavras fora de contexto, que permite a avaliação de modelos de processamento cognitivo da leitura por meio da manipulação de determinadas características psicolinguísticas das palavras. Essa manipulação propicia o surgimento de diversos tipos de efeitos na leitura feita por leitores competentes, iniciantes e por aqueles que, devido a lesões cerebrais, perderam com maior ou menor gravidade as habilidades anteriormente adquiridas. Os desempenhos desses indivíduos nas tarefas de leitura servem tanto para confirmar quanto para revisar as teorias de processamento propostas por esses modelos verbais.

Um dos modelos da psicologia cognitiva que tem se mostrado bastante útil na área de avaliação do reconhecimento de palavras, e que tem recebido sucessivas confirmações tanto de ordem teórica¹ quanto prática, é o Modelo de dupla-rota. Este modelo foi originalmente proposto por Morton em 1979 e desde então tem recebido diversas reformulações, sendo a mais atual delas o Modelo de dupla-rota em Cascata – DRC (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). Segundo Rastle e Coltheart (1999), o termo “dupla-rota” refere-se a uma classe particular de teorias sobre a leitura cuja característica definidora é a postulação da existência de dois procedimentos

¹ Representada pelos modelos computacionais.

distintos utilizados para se converter a grafia em fala: um procedimento lexical (ou rota lexical), em que a pronúncia e o significado da palavra são acessados diretamente, por meio da memória visual; e um procedimento fonológico (ou rota fonológica), que constrói a pronúncia da palavra por meio do uso de regras de conversão entre a ortografia e pronúncia e cujo significado deve ser acessado indiretamente, por meio desta pronúncia então gerada (memória acústica). O modelo de dupla-rota da leitura é reproduzido na Figura 2.

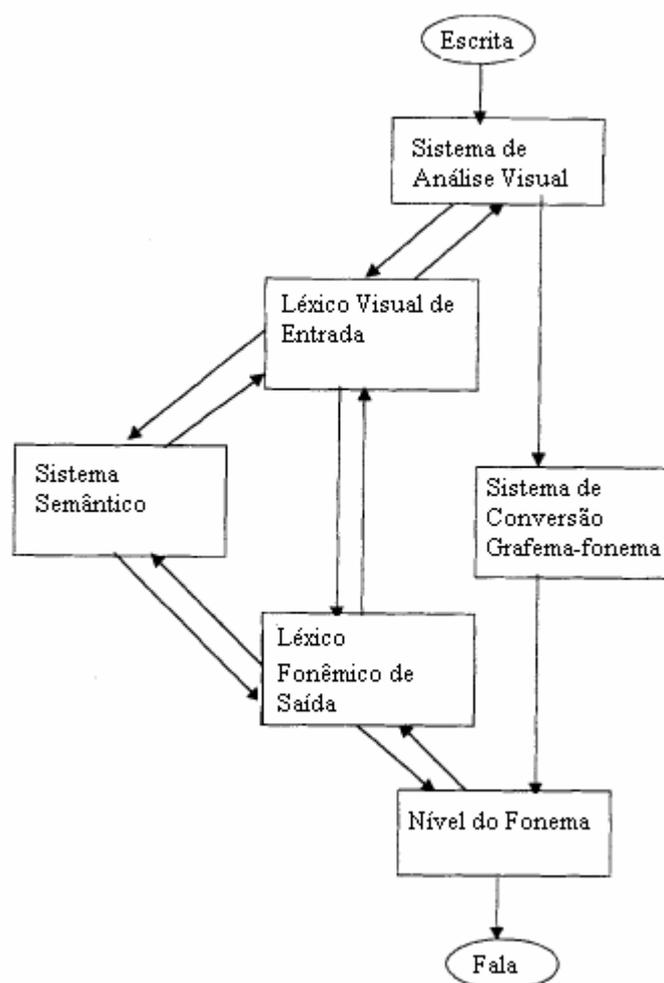


Figura 2. Estrutura básica do modelo de dupla-rota em cascata para o reconhecimento de palavras e a leitura em voz alta (reproduzido de Coltheart et al. (2001) com tradução).

A rota fonológica opera de acordo com a aplicação de regras de correspondência entre grafemas e fonemas. Os fonemas são as menores unidades significativas da fala (ou seja, unidades capazes de alterar o significado das palavras). Por exemplo, /b/ e /k/ são fonemas distintos da língua portuguesa porque geram significados distintos quando colocados diante da mesma terminação /alə/ – por produzirem respectivamente as pronúncias das palavras *bala* e *cala*. No entanto, variações regionais nas pronúncias das palavras, sem a correspondente mudança de significado, não caracterizam fonemas, e sim alofones (como as pronúncias /^hdiə/ e /^hdʒiə/ da palavra *dia*). A cada fonema corresponde-se um único grafema, que constitui uma unidade da escrita. Em uma palavra, nem sempre o número de grafemas coincide com o número de letras, como, por exemplo, o caso da palavra *ninho*, que apresenta cinco letras e quatro fonemas (pois o grafema <nh>, composto de duas letras, representa o único fonema /ŋ/). No português, diferentemente de línguas mais irregulares, como o inglês, os grafemas são formados por uma ou duas letras, este último caso representando os dígrafos (Parente, Silveira e Lecours, 1997). Alguns autores defendem que os grafemas (e não as letras) são processados como unidades perceptuais nos sistemas alfabéticos de escritas como o inglês e o francês, sendo as unidades básicas da leitura (Henderson, 1985; Rey, Ziegler, e Jacobs, 2000), mas ainda não foram conduzidos estudos em língua portuguesa que tratassem diretamente desta questão. No entanto, o estudo de Sucena e Castro (2005) pode fornecer *insights* preliminares sobre a aquisição da relação grafema-fonema no português. Os autores avaliaram crianças portuguesas cursando da 1ª à 4ª série na leitura de não-palavras com ortografia simples e complexa e observaram a ausência de diferenças entre esses estímulos nas séries mais avançadas. Tais resultados sugerem

uma mudança lenta do uso do processo de conversão letra-som para o processo grafema-fonema em nosso idioma.

No modelo DRC, as regras de conversão entre grafemas e fonemas são efetuadas segundo um critério estatístico, ou seja, cada grafema está associado ao fonema que mais o representa na língua. A rota fonológica opera de uma maneira serial, da esquerda para a direita, convertendo, letra por letra, uma seqüência de letras em uma seqüência de fonemas, respeitando-se a posição em que a letra se encontra na palavra². É em função disso que o número de letras que compõe as palavras afeta a leitura feita pela rota fonológica, ou seja, as palavras de maior comprimento tendem a ser lidas de uma maneira mais lenta e mais susceptível a erros do que as palavras menores, o que é conhecido na literatura como *efeito de extensão*. O comprimento das palavras não afeta a leitura lexical³ que, ao contrário da leitura pela rota fonológica, é feita de forma paralela. Outra evidência que aponta para o processamento do tipo serial da rota fonológica relaciona-se à diminuição do efeito de regularidade⁴ entre as palavras de baixa frequência à medida que a posição da irregularidade da palavra se afasta da primeira sílaba (Coltheart & Rastle, 1994; Rastle & Coltheart, 1999). Assim, como a rota fonológica seleciona os fonemas das palavras serialmente, quando a irregularidade se encontra na primeira sílaba, o fonema incorreto pode ser ativado antes que a pronúncia correta da palavra seja ativada pela rota lexical, o que aumenta a

² Se a palavra contiver um grafema formado por letras compostas, a leitura pela rota fonológica é afetada pelo chamado efeito “whammy”, que foi descrito pela primeira vez por Rastle & Coltheart (1998), utilizando simulações computacionais e resultados com seres humanos. Esse efeito constitui o aumento das latências na leitura fonológica de palavras que apresentam grafemas compostos. Por exemplo, na leitura da palavra *ganho*, a rota fonológica irá gerar o fonema /n/ para a letra <n> na terceira posição da palavra. No entanto, quando a próxima letra for processada (no caso, <h>), a rota irá produzir o fonema /ŋ/ (correspondente ao grafema <nh>) na terceira posição. Como dois fonemas distintos foram produzidos para a mesma posição, o fonema correto (no caso, /ŋ/) irá sofrer uma competição inibitória, gerando, portanto, um aumento no tempo de reação.

³ Particularmente no que se refere aos processos que ocorrem no léxico visual de entrada, no léxico fonêmico de saída e no sistema semântico.

⁴ Que se trata, conforme veremos, da vantagem em termos de precisão ou tempo de processamento de estímulos regulares em relação a irregulares.

probabilidade de erros ou de incremento do tempo de reação na leitura. Em contrapartida, quando a irregularidade se encontra nas sílabas finais, a chance de que a rota lexical ative a pronúncia da palavra antes que a rota fonológica termine conversão do fonema irregular aumenta, diminuindo, assim, a probabilidade de que ocorra um conflito entre as pronúncias geradas pelas rotas lexical e fonológica.

Geralmente, a utilização da rota lexical ou fonológica varia de acordo com os tipos de estímulos que são apresentados. Desta forma, a leitura de palavras reais (principalmente a de palavras muito comuns, ou de alta frequência) é feita pela rota lexical, enquanto que a leitura de estímulos novos para o leitor e de não-palavras⁵ é feita pela rota fonológica. No entanto, as não-palavras não são lidas inteiramente pelo processo fonológico, havendo a possibilidade de influência lexical na leitura desses estímulos (Coltheart et al., 2001; Rastle & Coltheart, 1999). Isso ocorre porque as não-palavras são capazes de ativar as palavras reais que lhes são ortograficamente semelhantes no léxico mental, o qual, por sua vez, ativa a representação fonológica das palavras reais e de seus fonemas, facilitando, assim, a leitura das não-palavras. Desse modo, ambos os processos aqui descritos (o lexical e o fonológico), apesar de distintos, são simultâneos, o que torna possível essa influência lexical na leitura de não-palavras (e também de palavras pouco frequentes), o que é conhecido como efeito de vizinhança (para maiores detalhes, ver Andrews 1989, 1992; ver também Sears, Hino, & Lupker, 1995; Grainger, O'Regan, Jacobs, & Segui, 1992).

De uma maneira geral, o uso prioritário da rota fonológica pode causar problemas na leitura de palavras irregulares, cuja pronúncia é definida de maneira arbitrária, ou seja, estas palavras podem não ser lidas com sucesso por meio da aplicação de regras de conversão entre grafemas e fonemas, pelo menos em grande parte

⁵ As não-palavras, também referidas na literatura como pseudopalavras, são seqüências de letras que respeitam a estrutura ortográfica de uma língua, mas que não apresentam significados aceitos pela comunidade lingüística. São, portanto, palavras sem significado.

das vezes⁶. Assim, a leitura fonológica de palavras irregulares pode estar sujeita a erros de pronúncia, o que não ocorre no caso da leitura das palavras regulares, em que a aplicação eficiente das regras de correspondência entre grafema e fonema é capaz de fornecer uma pronúncia correta. No português, as palavras consideradas regulares são aquelas que podem ser lidas corretamente por meio do uso de regras que são independentes de contexto, ou seja, que apresentam uma relação unívoca entre a grafia e o fonema (como a palavra *faca*), ou que são dependentes de contexto que governam a pronúncia (como é o caso, no português, dos grafemas <r> e <s> intervocálicos, que representam, respectivamente, os fonemas /r/ e /z/). Por outro lado, existem apenas dois casos de irregularidade no português, descritos por Parente, Silveira e Lecours (1997) como irregularidade do tipo 1A e do tipo 1B. No primeiro caso, os autores consideram as palavras que são simultaneamente irregulares tanto para a leitura quanto para a escrita, sendo o caso do <x> no contexto intervocálico. Na irregularidade do tipo B agrupam-se os casos dos grafemas <e> e <o> em sílaba tônica que, em posição paroxítona, apresentam pronúncia arbitrária aberta /ɛ/ ou /ɔ/ ou fechada /o/ ou /e/. No primeiro caso, a leitura pela rota fonológica gera principalmente erros de regularização (substituição da pronúncia irregular do <x> por sua pronúncia mais freqüente, o fonema /ʃ/) e, no segundo caso, a leitura está sujeita a erros de troca de qualidade da vogal (leitura de uma vogal aberta por uma fechada e vice-versa).

⁶ Muitas vezes se pensa que as palavras irregulares não podem ser pronunciadas corretamente por meio do uso da rota fonológica, o que é uma interpretação inadequada. Conforme dito anteriormente, a rota fonológica escolhe os fonemas por meio de um critério estatístico. Assim, ao utilizar o processo fonológico, a chance que o leitor tem de escolher uma pronúncia incorreta para um grafema irregular aumenta se a ocorrência desse grafema é pouco freqüente na língua, mas há chances de se escolher o fonema correto, principalmente se o fonema irregular for muito freqüente, como é o caso, por exemplo, do fonema /ʃ/, que é a pronúncia mais freqüente do grafema <x> no português, ocorrendo 75% das vezes (Silva & Rothe-Neves, 2002). Desta forma, há uma grande probabilidade que um leitor, mesmo lendo pela rota fonológica, acerte a pronúncia de uma palavra que contenha o <x> intervocálico com este som, por exemplo, na leitura da palavra *luxo*.

Ao contrário da leitura pela rota fonológica, a rota lexical produz uma leitura eficiente independentemente do nível de regularidade das palavras. Desta forma, é possível fazer um teste do uso dos processos lexical ou fonológico na leitura por meio da manipulação do nível de regularidade das palavras. Assim, se as palavras irregulares forem lidas de maneira significativamente menos precisa do que as palavras regulares, isso implica que a leitura está sendo feita pelo processo fonológico, e não pelo processo lexical. À maior correção na leitura das palavras regulares em relação às palavras irregulares dá-se o nome de *efeito de regularidade na precisão* da leitura. O efeito de regularidade ocorre principalmente entre leitores iniciantes, que confiam no processo fonológico por ainda não apresentarem um amplo vocabulário lexical, o qual apresenta as representações ortográficas, semânticas e fonológicas das palavras conhecidas.

Conforme dito anteriormente, a leitura de palavras é feita simultaneamente pelas rotas lexical e fonológica. Além disso, no caso da leitura de palavras regulares, tanto o processo fonológico quanto o lexical são capazes de gerar a mesma pronúncia, enquanto que a leitura de palavras irregulares pode estar sujeita a erros de pronúncia quando estabelecida pelo processo fonológico. Neste caso, haverá um desacordo entre as duas pronúncias produzidas, sendo necessária uma nova consulta ao léxico para uma confirmação. Isto gera um atraso na produção da resposta e caracteriza o *efeito de regularidade no tempo de processamento* ou tempo de reação (TR). O efeito de regularidade no tempo de processamento pode, portanto, ser definido como uma função da relação temporal que se estabelece entre as duas rotas (Kinoshita, Lupker, & Rastle, 2004). Assim, a leitura fonológica de palavras irregulares pode estar tanto sujeita a erros de pronúncia, que caracteriza o efeito de regularidade na precisão, quanto a um atraso na produção da resposta, que corresponde ao efeito de regularidade no tempo de reação.

A regularidade das palavras afeta a leitura de maneira diferenciada de acordo com seu nível de frequência de ocorrência. As palavras de alta frequência (como havia antecipado) são aquelas que o leitor entra constantemente em contato, sendo, portanto, muito ativadas no léxico mental. Por esse motivo, o reconhecimento dessas palavras ocorre de uma maneira muito mais rápida do que o das palavras de baixa frequência, o que caracteriza o *efeito de frequência*. Assim, se uma palavra irregular for muito frequente, a tendência é que seu reconhecimento se efetue de maneira tão rápida e eficaz quanto ao de uma palavra regular, de tal forma que o efeito de regularidade não atinge a leitura das palavras muito comuns. Isso implica que o efeito de regularidade geralmente se restringe à leitura das palavras pouco frequentes na língua, como é o caso do português, conforme demonstrado por Pinheiro (1994, 1995, no prelo). (Voltaremos a esse ponto mais adiante, na seção seguinte). A presença do efeito de frequência é um indício do uso de processo lexical na leitura, mas um efeito exagerado de frequência, principalmente entre leitores mais experientes, ou a ausência deste, pode indicar falhas no processo lexical (Pinheiro, 1999; Pinheiro & Rothe-Neves, 2001), ou problemas na formação das representações das palavras no léxico.

Em português, o desenvolvimento dos processos lexicais e fonológicos ocorre paralelamente ao longo da alfabetização (Godoy, 2005; Pinheiro, 1995; Salles & Parente, 2005) e isso pode ser demonstrado por uma diminuição dos efeitos de regularidade, extensão e de frequência com o decorrer da escolarização. A redução de tais efeitos indica que cada vez mais palavras (curtas e longas, regulares e irregulares) são lidas pelo processo lexical, havendo um desprendimento do uso prioritário do processo fonológico com o desenvolvimento da habilidade de leitura. Tal fato apresenta desdobramentos de ordem clínica. Assim, se a manifestação dos efeitos supracitados persiste de uma forma intensa entre leitores mais experientes, isso pode fornecer

indícios sobre dificuldades na capacidade de reconhecer palavras, o que, por sua vez, pode gerar repercussões para a aquisição dos processos superiores da leitura como, por exemplo, a compreensão de texto. Neste caso específico, a contribuição de um reconhecimento eficiente de palavras se efetuará por meio do automatismo (caracterizado pela rapidez e precisão na leitura), uma vez que, quanto mais automática a leitura é realizada, menos recursos da memória de trabalho são gastos na decodificação, os quais podem ser alocados para a compreensão (Morais, 1986; Perfetti, 1985). No entanto, alguns autores questionam o impacto isolado da velocidade de processamento para a compreensão e acreditam que a principal contribuição do reconhecimento de palavras neste processo é o conhecimento ortográfico (para uma revisão, ver Goff, Chris Pratt, & Ben Ong, 2005).

Em suma, o desempenho dos indivíduos na tarefa de leitura em voz alta de palavras isoladas é capaz de fornecer indícios sobre os processos (lexicais ou fonológicos) utilizados por eles na leitura. Isso é medido pelos efeitos que ocorrem devido à manipulação das características psicolinguísticas das palavras. Assim, se o contraste entre palavras regulares e irregulares gerar diferenças significativas, na forma de um efeito de regularidade, pode-se concluir que a leitura está sendo mediada por um mecanismo fonológico, pois as palavras irregulares são lidas com sucesso se o processo lexical for utilizado. O mesmo pode ser concluído se o contraste entre palavras grandes e curtas gerar um forte efeito de extensão, uma vez que o processo fonológico, por ser serial, gera a pronúncia mais lentamente para as palavras grandes do que para as de menor comprimento. Se as palavras muito comuns apresentarem uma vantagem em relação às palavras pouco frequentes, dentro do esperado pelo nível de escolaridade do indivíduo, temos indícios da existência de um mecanismo lexical em formação⁷.

⁷ Um quarto efeito que o modelo de dupla-rota é demonstrado na tarefa de leitura em voz alta é o efeito lexical (ou efeito de lexicalidade), que corresponde à leitura mais rápida e eficiente das palavras reais em

Em países de língua inglesa, em que o modelo de dupla-rota foi desenvolvido, a presença dos efeitos referidos anteriormente é constantemente relatada na literatura (para o efeito de frequência, ver Foster & Chambers, 1973; para o efeito de regularidade, ver Seidenberg, Waters, Barnes, & Tanenhaus, 1984 e Taraban & McClelland, 1987 (todos citados por Coltheart et al., 2001); para o efeito de extensão, ver Rastle & Coltheart, 1999). Na seção seguinte, descrevemos como esses efeitos têm aparecido em estudos na língua portuguesa.

2.3 Estudos sobre a tarefa de leitura em voz na língua portuguesa

Em muitos países desenvolvidos, vários testes de reconhecimento de palavras são publicados e apresentam normas para as populações a que se referem, sendo também capazes de fornecer medidas de idade de leitura. Como exemplo, podemos citar, na Inglaterra, “*The British Ability Scales – BAS*” (Elliott, 1996) e *The Schonell Word Reading* (Schonell & Goodacre, 1971; Schonell & Schonell, 1960). Nos Estados Unidos, citamos os subtestes de leitura do “*The Wide Range Achievement Test – WRAT3*” (Wilkinson, 1993) e do “*Woodcock Reading Mastery Tests-Revised*” (Woodcock, 1998). Na Nova Zelândia, temos o “*Burt Word Reading Test, New Zealand Revision*” (Gilmore, Croft, & Reid, 1981). Na França, uma bateria computadorizada de avaliação da leitura e de habilidades relacionadas, o EVALEC, recebeu recentemente dados normativos para crianças francesas da 1ª à 4ª série do Ensino Fundamental e,

relação às não-palavras. Essa vantagem das palavras reais ocorre porque elas apresentam representações arquivadas no léxico mental, enquanto que a pronúncia das não-palavras é construída pela rota fonológica. Desta forma, é esperado que o efeito lexical seja maior para o contraste com as palavras de alta frequência do que com as palavras de baixa frequência, as quais, conforme vimos, estão menos ativadas no léxico mental. Aqui não tratamos mais extensamente desse efeito por envolver a leitura de não-palavras.

dentre os vários subtestes, a bateria apresenta um subteste de leitura de palavras isoladas (Sprenger-Charolles, Colé, Béchennec, & Kipffer-Piquard, 2005).

Ao contrário dos países supracitados, no Brasil a avaliação cognitiva da leitura tem sido feita principalmente por meio da utilização de listas de palavras que variam em termos das características psicolinguísticas contrastantes de regularidade, extensão e frequência a que se referiu anteriormente e que, quase em sua totalidade, não passaram pelos processos psicométricos de validação, de precisão e de criação de normas. A única exceção⁸ constitui o subteste de leitura do Teste de Desempenho Escolar (Stein, 1994) e, por esse motivo, analisaremos essa tarefa em uma seção à parte. Nesta seção, descreveremos os principais resultados encontrados por estudos em língua portuguesa que utilizaram a tarefa de leitura em voz alta realizada com crianças em fase de desenvolvimento da habilidade de leitura, com ênfase na tarefa de leitura de palavras reais e na descrição dos efeitos que, de certa forma, confirmam os achados do modelo de dupla-rota. Referências a outras tarefas serão feitas somente quando necessário.

O primeiro estudo conduzido no Brasil com o objetivo (dentre outros) de investigar a aplicabilidade do modelo de dupla-rota para o contexto da leitura em nossa língua foi o de Pinheiro (1994, 1995, no prelo), que investigou as habilidades de leitura e escrita de palavras reais e não-palavras isoladas entre crianças de 1^a à 4^a série de uma escola particular. Em relação à leitura de palavras reais, os resultados mostraram que, tanto em termos de precisão quanto de tempo de processamento, houve um efeito da

⁸ Capovilla e Capovilla (1998) estabeleceram normas preliminares para uma tarefa de leitura em voz alta de crianças cursando da pré-escola à 2^a série do Ensino Fundamental. O trabalho apresenta como problemas a amostra reduzida (76 participantes) e homogênea (foi utilizada apenas uma escola da rede particular). A mesma tarefa passou por uma validação de construto (Capovilla, Capovilla e Macedo, 1998), mas não foram relatadas análises de itens, de fidedignidade ou de outros índices de validade. Além disso, Capovilla, Varanda e Capovilla (2006) publicaram recentemente a normatização e a validação de uma bateria de testes de competência na leitura de palavras reais e de não-palavras. Tal bateria difere-se da tarefa de reconhecimento de palavras que estamos lidando aqui por ser constituída de tarefas de leitura silenciosa em que, adicionalmente aos estímulos escritos, são também apresentados estímulos pictóricos. Por essas razões, o subteste de leitura do TDE pode ser considerado como o único teste de leitura em voz alta do Brasil com as características psicométricas supracitadas.

frequência e da escolarização na leitura das crianças, tendo ocorrido também uma interação entre estes fatores. Tais resultados indicam, respectivamente, que houve uma influência lexical na leitura, que houve um ganho em proficiência e rapidez na leitura com a escolarização e que a frequência das palavras afetou de forma diferente a leitura das crianças das séries iniciais e das finais. Em relação ao efeito de regularidade, este se restringiu às séries iniciais e ao tempo de processamento das palavras pouco frequentes, o que foi interpretado como uma influência da regularidade somente entre esses estímulos. O comprimento das palavras também só afetou o tempo de reação, e em maior magnitude entre as crianças das séries iniciais. Como conclusão do estudo, a autora aponta a identificação da existência dois processos distintos que se desenvolviam de maneira paralela, um lexical e outro fonológico, confirmando, assim, que o modelo de dupla-rota poderia ser estendido para a leitura do português brasileiro.

Em relação ao estudo de Pinheiro (1994, 1995, no prelo), duas considerações devem ser feitas. Primeiro, tanto para a avaliação da leitura, quanto da escrita, a classificação de regularidade das palavras utilizada no estudo considerou a relação fonema-grafema, ou seja, a direção da escrita. Para Pinheiro e Rothe-Neves (2001), tal fato pode ter induzido à ausência do efeito de regularidade nos erros, pois, conforme lembram os autores, muitas palavras irregulares para a escrita são regulares para a leitura (como *cidade*), ou seja, podem não gerar erros de pronúncia mesmo se lidas pela rota fonológica. Outro ponto que merece destaque é a metodologia utilizada no estudo, no qual foram considerados estatisticamente válidos apenas os resultados que se mostraram significativos tanto na análise de itens quanto na análise de sujeitos. Tal medida pode ter levado a divergências de resultados com outros trabalhos. Por exemplo, Capovilla, Capovilla e Macedo (1998), utilizando a mesma lista de palavras de Pinheiro (1995), com a diferença de apenas três itens, e ao contrário desta autora, encontraram o

efeito de regularidade nos erros cometidos na leitura. No entanto, como ressaltam Pinheiro e Rothe-Neves (2001), os referidos autores utilizaram apenas análises de sujeitos para reportar o efeito, o que contraria as recomendações de Clark (1973) – assim como de outros autores, como, por exemplo, Raaijmakers (2003), Raaijmakers, Schrijnemakers e Gremmen (1999, 2006). Segundo Clark (1973), para que os resultados sejam generalizáveis tanto para a população de sujeitos a qual a amostra representa, quanto para a população de itens do qual o conjunto de itens experimentais foi derivado, é necessário que se encontre resultados significativos em ambos os tipos de análise, de sujeito e de itens. Isso implica que os resultados de Capovilla e colegas só podem ser generalizados com segurança para a população de sujeitos, não se estendendo para o corpus de palavras que dispomos na língua. Com exceção dos estudos de Pinheiro, todos os resultados a serem relatados nesta revisão foram obtidos apenas com a análise de sujeitos, sendo, portanto as suas generalizações de alcance limitado.

Em um estudo com crianças portuguesas, Sucena e Castro (2005) avaliaram a leitura de palavras reais e de não-palavras feita por crianças cursando da 1ª à 4ª série (N = 99) com o objetivo de investigar o efeito da consistência ortográfica na precisão e no tempo de processamento na leitura e determinar as estratégias preferenciais (lexicais ou fonológicas) utilizadas pelos participantes. As palavras do estudo foram divididas em quatro condições ortográficas, podendo ser simples (relação letra-som unívoca), complexas (formadas por grafemas compostos por mais de uma letra), regidas por regras contextuais e irregulares. Entre as não-palavras, só não havia a condição irregular. Entre outras conclusões, o estudo mostrou a presença dos efeitos de regularidade (palavras regulares *versus* irregulares), de complexidade ortográfica (grafemas simples *versus* compostos) e de lexicalidade (palavras *versus* não-palavras) na leitura, tanto em termos de tempo de reação quanto de precisão. Além disso, o efeito

de extensão apareceu a partir da 2ª série no tempo de processamento de palavras e de não-palavras, tendo sido maior entre estas últimas. Os autores concluíram que as crianças utilizavam paralelamente tanto as estratégias lexicais (efeito lexical) quanto e fonológicas (efeitos de regularidade, extensão e complexidade) na leitura, o que aponta para um léxico ortográfico em desenvolvimento mesmo entre as crianças da 4ª série. Os resultados desse estudo estão em consonância com os resultados obtidos com crianças brasileiras (Pinheiro, 1994, 1995; Salles & Parente, 2002).

Com o objetivo avaliar as rotas preferenciais (lexical ou fonológica) utilizadas por crianças de 2ª e 3ª série de escolas particulares, Salles e Parente (2002) investigaram a emergência dos efeitos de regularidade, frequência, extensão e lexicalidade na leitura de palavras isoladas utilizando uma classificação que considera a regularidade para a leitura das palavras. Os resultados mostraram que a maioria das crianças da 3ª série utilizava ambas as rotas com proficiência, enquanto que as crianças da 2ª série eram preferencialmente leitores fonológicos, o que é condizente com a idéia de que, com o desenvolvimento, a leitura fonológica evolui para uma leitura lexical, mas que um bom leitor é proficiente no uso de ambas as rotas. Em relação aos efeitos na leitura, o maior efeito encontrado foi o de regularidade, seguido do de extensão (na leitura de não-palavras) e de frequência, enquanto que o efeito de lexicalidade foi quase inexistente. Resultados semelhantes foram obtidos em Salles (2005), que estudou uma amostra de crianças cursando a 2ª série, desta vez oriunda de escolas públicas. A autora relatou os efeitos de regularidade, frequência e extensão (na leitura de não-palavras), mas novamente não encontrou o efeito lexical.

A presença do efeito de regularidade no trabalho de Salles e Parente (2002) e de Salles (2005) reforçam a idéia de que este efeito pode ser obtido considerando-se a regularidade grafema-fonema das palavras, mas os resultados destes estudos devem ser

considerados com cautela, uma vez que em ambos os trabalhos somente foram relatados os tamanhos dos efeitos em termos de diferenças de médias, não tendo sido conduzidas análises estatísticas para se verificar se tais diferenças foram significativas. Mesmo assim, pode-se afirmar que o tamanho do efeito encontrado aponta para tal direção.

Um fato que chama a atenção nos trabalhos de Salles e Parente (2002) e de Salles (2005) é a ausência do efeito lexical, constantemente relatado tanto na língua inglesa (p. ex., Kinoshita, et al., 2004; McCann & Besner, 1987; Rastle & Coltheart, 1999) quanto no português (Capovilla et al., 1998; Godoy, 2005; Pinheiro, 1994; 1995; Sucena & Castro, 2005). Duas breves considerações serão feitas para tentar explicar esses resultados. Primeiramente, o efeito lexical nesses trabalhos pode ter sido anulado em função de as não-palavras terem sido contrastadas com as palavras reais de ambas as frequências juntas, e não separadamente. Os estudos em português apontam para o fato de que a frequência de ocorrência da palavra interfere na produção desse efeito, pelo menos em uma lista com a classificação de regularidade do ponto de vista da leitura⁹. Assim, Godoy (2005) demonstrou uma superioridade lexical na precisão e na velocidade de processamento da leitura de palavras de alta frequência em relação às não-palavras, mas o mesmo efeito não apareceu entre as palavras de baixa frequência (de fato, houve um efeito marginalmente significativo na precisão a favor das não-palavras, com $p = 0,65$. Esse resultado é esperado considerando a fase de desenvolvimento da leitura das crianças testadas (alunos da pré-escola e da 1ª série) e o tipo de classificação de regularidade de palavras utilizadas (palavras regulares e irregulares para leitura)). Sucena e Castro (2005), também utilizando uma lista de palavras classificadas quanto à regularidade do ponto de vista da leitura, obtiveram o

⁹ Pinheiro (2006) encontrou o efeito geral de lexicalidade no tempo de reação e na precisão de palavras de alta e de baixa frequência, mas conforme visto anteriormente este estudo utilizou uma classificação de palavra diferente dos estudos que estamos tratando.

mesmo efeito em todas as séries avaliadas (1ª à 4ª série) utilizando palavras de frequência média. É possível que o nível de frequência das palavras utilizado (palavras de média frequência ao invés de palavras de baixa frequência) explique o efeito encontrado nas crianças de 1ª série.

Outra hipótese para os resultados refere-se ao método utilizado nos estudos. Nos trabalhos de Salles e Parente (2002) e Salles (2005), as palavras reais e as não-palavras foram apresentadas conjuntamente em uma mesma lista e isso pode ter induzido à ausência do efeito de lexicalidade. A explicação está no efeito das estratégias utilizadas na leitura (Coltheart, 1978, citado por Coltheart et al., 2001; Kinoshita et al., 2004). Assim, Coltheart (1978) demonstrou que a estratégia – lexical ou fonológica – escolhida pelo leitor pode ser afetada pelo tipo de estímulo – palavra real ou não-palavra – que está sendo lido. Nesse estudo, os participantes tinham que ler uma lista de não-palavras pronunciáveis, com exceção do último item, que era uma palavra real altamente irregular (*wolf*). Vários dos participantes regularizaram esta palavra para rimar com a palavra *golf*, sendo que muitos deles sequer perceberam que havia uma palavra real na lista. Coltheart (1978) reporta esse achado como a influência do efeito da estratégia que é utilizada na leitura. Ele interpretou o resultado como sendo o efeito do “desligamento” da rota lexical pelos participantes, que utilizaram de forma prioritária a rota fonológica por esta gerar sempre a resposta correta na leitura das não-palavras. Em outros termos, a previsibilidade de que o próximo estímulo seria também uma não-palavra forçou a leitura dos estímulos da lista por meio da rota fonológica.

Desta forma, fato semelhante pode ter acontecido entre os participantes das pesquisas citadas anteriormente. Como as palavras reais e as não-palavras foram reportadas em uma mesma lista, os participantes podem ter sido induzidos a utilizar a mesma estratégia para ler ambos os tipos de estímulos, a saber, a estratégia fonológica,

dato que esta poderia gerar pronúncias plausíveis para a maioria dos estímulos. No trabalho de Salles (2005), além da ausência do efeito de lexicalidade, o efeito das estratégias ficou evidente com o melhor desempenho das crianças na leitura de não-palavras em relação às palavras irregulares. Esse fato ilustra muito bem como a metodologia utilizada pelo pesquisador pode interferir na produção de efeitos que são utilizados para testar o modelo de processamento cognitivo da leitura e, conseqüentemente, na interpretação dos resultados.

Em um estudo longitudinal, Godoy (2005) avaliou a influência do método de ensino no desenvolvimento de estratégias de processamento na leitura e na escrita de crianças cursando da pré-escola à 1ª série em escolas particulares de Florianópolis. Na Fase 3 do estudo, em que as crianças cursavam a 1ª série, e utilizando uma lista experimental de palavras contrastadas quanto à frequência, regularidade para a leitura e extensão¹⁰, a autora descobriu, além do efeito lexical supracitado, a presença de um efeito de extensão e de frequência na precisão e na velocidade de leitura de ambos os grupos de crianças. O efeito de regularidade se restringiu à precisão na leitura, mas pôde ser obtido para o tempo de reação quando valores extremos (*outliers*) foram retirados da análise. Tais resultados confirmam o uso do processo fonológico por ambos os grupos de crianças, mas também apontam para a formação paralela de um processo lexical. Apesar de nesta fase do estudo as crianças expostas ao método fônico terem apresentado um melhor resultado do que as expostas pelo método global, não foi encontrado um efeito geral do método de alfabetização na leitura. Este resultado sugere que, independentemente do método de ensino, o processo de decodificação pode ser alcançado com eficiência em uma ortografia transparente, como é o caso do português. No entanto, não se descarta a hipótese de que o contraste entre os métodos de

¹⁰ Essa lista foi criada por Pinheiro (2003) especialmente para o estudo.

alfabetização feitos pelo estudo não tenha sido suficientemente forte para demonstrar o efeito do método na leitura (Godoy, 2005; Godoy, Defior & Pinheiro, 2007).

Com o objetivo de avaliar o impacto da relação grafema-fonema e fonema-grafema na emergência do efeito de regularidade na leitura, Pinheiro, Lúcio e Silva (2008) conduziram um estudo piloto no qual o nível de regularidade das palavras foi controlado tanto do ponto de vista da leitura quanto da escrita. Assim, um grupo de crianças da 1ª à 3ª série de uma escola particular leu o mesmo corpo de palavras que está sendo testada no presente estudo. Conforme esperado a partir das recomendações de Parente, Hosogi e Lecours (1997) e de Pinheiro e Rothe-Neves (2001), controlando-se o nível de regularidade das palavras para a leitura, Pinheiro e colaboradores encontraram um efeito de regularidade na precisão. No entanto, e ao contrário dos achados anteriores de Pinheiro (1994, 1995, no prelo), em que as palavras da lista testada foram classificadas na direção fonema-grafema, o efeito não se estendeu ao tempo de reação. O efeito no tempo de processamento só foi observado quando se considerou a classificação de regularidade das palavras para a escrita, o que está em consonância com os resultados de Pinheiro (1994, 1995, no prelo), mas diferentemente destes, o efeito também se estendeu aos erros produzidos. Como explicação, supôs-se que a classificação de palavras utilizada no estudo de Pinheiro e colaboradores (que considerou apenas dois níveis de regularidade, tanto para a leitura quanto para a escrita) possa ter interferido nos resultados. Outra possibilidade seria a de que o efeito de regularidade na precisão tenha ocorrido em função da grande diferença em termos de número de palavras utilizadas nos dois estudos. Enquanto a lista de palavras da pesquisa original era composta de 96 palavras, a da pesquisa sob consideração era formada de 323 palavras, o que pode ter induzido as crianças a errarem mais como resultado de fadiga (possibilidade não testada).

Os resultados encontrados para o efeito de regularidade em todos os estudos citados acima são resumidos na Tabela 1. Pela análise desta tabela, podemos observar que é plausível a hipótese levantada por Pinheiro et al. (2008) de que a emergência do efeito de regularidade na leitura possa variar de acordo com a classificação de regularidade das palavras (direção da leitura ou da escrita). Assim, pela análise desta tabela, podemos observar que, independentemente do número de categorias de regularidade utilizadas no estudo, *quando se considera a regularidade para a leitura das palavras, o efeito de regularidade está sempre presente na precisão*. De maneira semelhante, e novamente independente da quantidade de categorias, *utilizando-se a classificação de regularidade para a escrita das palavras, o efeito está sempre presente no tempo de processamento na leitura*. Isso conduz à hipótese de que o efeito de regularidade na leitura aparecerá na precisão, sempre que o nível de regularidade das palavras for classificado quanto à direção da leitura, enquanto que, para o tempo de processamento, esse efeito sempre aparecerá quando a classificação para a escrita é considerada.

No entanto, um fato complicador é que o surgimento do efeito para a precisão e para o TR parece variar em relação ao *número de categorias* utilizadas no estudo quando se considera, respectivamente, a classificação das palavras para a escrita ou para a leitura. Desta forma, nos estudos de Godoy (2005) e de Pinheiro et al. (2008), em que se consideraram apenas duas categorias, o efeito *não* apareceu no tempo de reação controlando-se o nível de regularidade das palavras para a leitura, mas o contrário ocorreu no trabalho de Sucena e Castro (2005), que utilizaram quatro categorias diferentes. Em relação à precisão, Pinheiro (1994, 1995, no prelo), utilizando a classificação para a escrita e três categorias de regularidade, encontrou o efeito de regularidade apenas para o grupo de sujeitos, mas não para os itens (os resultados foram

replicados por Capovilla, Capovilla e Macedo (1998) também na análise de sujeitos). No entanto, Pinheiro e colaboradores encontraram o efeito tanto para os itens quanto para os sujeitos, utilizando apenas duas categorias de regularidade.

Tabela 1

Resultados para o efeito de regularidade na leitura em diversos estudos com a língua portuguesa, em função da direção da classificação de regularidade das palavras, e estatísticas por eles utilizadas

Estudo	Direção da Classificação de regularidade	Número de categorias	Efeito de regularidade (precisão)	Efeito de regularidade (TR)	Estatística utilizada
Capovilla, Capovilla e Macedo (1998)	Escrita	Três	Presente	Presente	ANOVA (F1)
Godoy (2005)	Leitura	Duas	Presente	Ausente	ANOVA (F1)
Pinheiro (1994, 1995, no prelo)	Escrita	Três	Ausente	Presente	ANOVA (F1 e F2)
Pinheiro, Lúcio e Silva (2008)	Leitura	Duas	Presente	Ausente	<i>minF'</i>
Pinheiro, Lúcio e Silva (2008)	Escrita	Duas	Presente	Presente	<i>minF'</i>
Salles (2005)	Leitura	Duas	Presente	—	—
Salles e Parente (2002)	Leitura	Duas	Presente	—	—
Sucena e Castro (2005)	Leitura	Quatro	Presente	Presente	ANOVA (F1)

Nota. TR = Tempo de reação

Neste estudo, buscar-se-á esclarecer essas questões teóricas a partir da condução de análises estatísticas separadas considerando duas ou quatro categorias de regularidade grafofonêmicas e fonografêmicas. A descrição das categorias de regularidade utilizadas nesse trabalho será feita na seção 4.

Tendo em vista o que foi exposto a respeito do estudo da avaliação cognitiva de leitura no Brasil, particularmente no que diz respeito à tarefa de leitura em voz alta de palavras reais isoladas, podemos concluir três pontos importantes: (1) os diversos estudos em língua portuguesa avaliados confirmam, de uma maneira geral, os postulados do modelo de dupla-rota, segundo o qual a leitura pode ocorrer por meio de dois processos distintos e simultâneos, um lexical e outro fonológico; (2) as crianças brasileiras parecem utilizar preferencialmente a estratégia fonológica no início da aprendizagem da leitura, a qual vai sendo gradualmente substituída pelo uso prioritário estratégia lexical, que é mais eficaz do ponto de vista do desenvolvimento. O apego à estratégia fonológica nesse período parece ser devido ao caráter quase regular da língua portuguesa, que apresenta apenas dois casos de irregularidade para a leitura. No entanto, a influência do método de ensino utilizado nas escolas, que geralmente é calcado no ensino formal das relações letra-som aos leitores iniciantes, ainda não pode ser descartada; (3) questões metodológicas interferem nos resultados encontrados pelos diferentes pesquisadores, e essas questões esbarram em dois pontos principais: (3.1) a maioria dos pesquisadores que trabalham com a avaliação da leitura no Brasil tem desconsiderado que, em estudos que utilizam materiais lingüísticos, é importante examinar a generalização dos resultados para o corpo de itens, além dos grupos de sujeitos, o que é obtido respectivamente pela presença de resultados significativos na análise de itens e na análise de sujeitos¹¹; (3.2) os estudos iniciais sobre o efeito de

¹¹ Ou pela combinação das duas medidas em uma terceira denominada $\min F'$, que constitui o valor mínimo da estatística F da ANOVA para que um resultado seja significativo simultaneamente nas

regularidade consideravam o nível de regularidade das palavras para a escrita, ao invés de para a leitura, o que pode interferir nos resultados dos estudos até então conduzidos; (3.3) a influência do nível de regularidade para a leitura e para a escrita das palavras na precisão e no tempo de reação na leitura ainda não está claramente compreendida.

Passaremos agora para uma investigação dos resultados das pesquisas que utilizaram o único teste de leitura em voz alta de palavras isoladas que apresenta normas para a população brasileira e estudos de validade e fidedignidade, o subteste de leitura do TDE.

2.4 Avaliação cognitiva com o subteste de leitura do teste de desempenho escolar

O Teste de Desempenho escolar (TDE) foi criado com o objetivo de tentar fornecer uma avaliação objetiva das capacidades básicas necessárias para o desempenho escolar nas áreas de leitura, escrita e aritmética. O teste apresenta normas para a população de crianças cursando da 1ª à 6ª série das redes pública (municipal e estadual) e particular do Ensino Fundamental do município de Porto Alegre – RS. Cada série apresenta três diferentes grupos de proficiência (grupos inferior, superior e médio) criados a partir das faixas de percentis 25 e 75 – com exceção da 1ª série, que adicionalmente apresenta as classificações “médio inferior” e “médio superior” (Stein, 1994).

No que se refere ao procedimento de construção do teste, tanto o procedimento de seleção dos itens, quanto o de validação e do cálculo do índice de fidedignidade foram estabelecidos por meio da análise da consistência interna. Assim, no

análises de itens e de sujeitos (Clark, 1973; Raaijmakers, 2003). Para um exemplo da aplicação do $\min F'$ em um estudo brasileiro de tarefa de decisão lexical, ver Justi (2004) e Justi & Pinheiro (2006).

procedimento de seleção dos itens, foram excluídos aqueles que apresentaram baixas correlações item-total. O procedimento de validação limitou-se à análise do conteúdo do teste e do critério interno, não havendo outros estudos sobre a validação do teste (como a de critério externo, a convergente, a divergente ou a preditiva). Por fim, o procedimento de fidedignidade foi dado por meio do Coeficiente Alfa de Cronbach.

Segundo constam informações do manual, diferenças significativas entre as séries foram encontradas para os três subtestes utilizando-se a técnica de análise multivariada de variância (MANOVA), mas não foi relatada a localização do efeito (ou seja, não foram investigadas diferenças entre séries consecutivas). Isso seria essencial para avaliar a proporção em que o teste avalia as diferenças entre as séries.

Voltando-nos especificamente ao subteste de leitura, e observando a Tabela 1 do manual (reproduzida aqui na Tabela 2), notamos que a partir da quarta série as médias e os desvios padrão de acertos são muito semelhantes uns aos outros. Portanto, não se pode afirmar com segurança que o teste discrimine as habilidades investigadas entre essas séries sem o estabelecimento de uma análise de *post hoc*. Além disso, o desvio padrão da primeira série apresenta um valor extremamente elevado em relação à média de acerto neste subteste, o que talvez prejudique a avaliação das diferenças individuais dentro desta série. Pela análise dessas médias, a única diferença que parece ser realmente grande é entre a primeira série e as demais séries. Desta forma, o fornecimento de dados a respeito da localização do efeito tornaria possível avaliar se os escores médios da primeira série estariam carregando o efeito de série no subteste de leitura.

Tabela 2

Valores das Médias e Desvios Padrões (D.P.) por série para o subteste de leitura do TDE (dados da padronização. Reproduzido da Tabela 1 de Stein, 1994)

Série	N	Média	D.P.
1 ^a	90	34,67	27,18
2 ^a	91	58,83	13,43
3 ^a	88	64,57	4,67
4 ^a	89	66,67	2,56
5 ^a	90	67,97	2,71
6 ^a	90	68,72	1,82

Um fato que talvez corrobore com o que foi exposto é demonstrado no trabalho de Dias, Enumo e Turini (2006). Os autores conduziram um estudo longitudinal no qual avaliaram, por meio do TDE, a evolução das habilidades de leitura, escrita e aritmética de 172 alunos da 2^a à 5^a série do Ensino Fundamental de Vitória (ES), os quais foram submetidos a duas avaliações nos anos de 2001-2002. Foram conduzidas ANOVAS com análises *a posteriori* do Teste de Turkey para localizar os efeitos entre as séries subseqüentes, separadamente para ambas as avaliações. Desta forma, notou-se que o padrão das diferenças interseriais para o subteste de leitura diferiu da primeira para a segunda avaliação. Em 2001 houve diferenças significativas apenas entre a 3^a e a 4^a série, enquanto que em 2002 essas diferenças ocorreram entre a 3^a e a 6^a série¹² e entre a 4^a e a 6^a série. Desta forma, os autores concluíram que, de uma maneira geral, foi necessário um intervalo de pelo menos dois anos para que houvesse diferenças interseriais no desempenho dos participantes no subteste. No entanto, os resultados

¹² Note que os alunos avaliados em 2001 geralmente cursavam a série posterior em 2002.

desse estudo devem ser interpretados com cautela, uma vez que a idade dos participantes variava entre 8 e 19 anos, não havendo diferenciação por idade no relato dos resultados. Assim, a heterogeneidade da amostra, que foi composta tanto por estudantes regulares, quanto por adolescentes e adultos cursando a educação continuada, prejudica a avaliação dos resultados obtidos no estudo. A separação da amostra de acordo com a idade forneceria resultados mais conclusivos a respeito da capacidade do TDE em discriminar as habilidades de alunos cursando diferentes séries.

Outro estudo relacionou o desempenho no TDE de crianças cursando da 1ª à 5ª série a habilidades básicas de leitura – avaliadas por meio da precisão e do tempo de reação no reconhecimento de palavras reais e de não-palavras (Tonelotto et al., 2005). Os resultados mostraram que o subteste de leitura do TDE não foi capaz de discriminar os desempenhos dos grupos Superior, Médio e Inferior da amostra. Além disso, houve uma associação entre o desempenho dos três grupos de crianças, tanto em termos de tempo de reação quanto de precisão, entre a tarefa de reconhecimento de palavras e o escore total e os escores dos subtestes de leitura e de escrita do TDE. Assim, quanto melhor foi o desempenho no TDE, maior foi a precisão na leitura de palavras e não-palavras e mais rápido foi o tempo de processamento desses itens. A presença de relação entre a tarefa de leitura feita pelos autores e os subtestes de leitura e escrita do TDE constitui uma evidência a favor da validade convergente do teste. No entanto, um fato que chama a atenção é que essa associação também foi encontrada para as relações entre o reconhecimento de palavras e de não-palavras e o subteste de aritmética¹³. Isto constitui uma incongruência com a validade divergente do teste, dado que as

¹³ Com exceção do tempo de processamento na leitura de palavras reais.

habilidades básicas de aritmética e de reconhecimento de palavras pertencem a construtos diferentes e, portanto, não deveriam estar relacionadas¹⁴.

Com o objetivo de investigar o poder em discriminar o desempenho dos indivíduos de diferentes níveis de aprendizagem que o subteste apresenta, foi conduzida uma revisão de literatura na qual se procurou sintetizar os principais resultados de trabalhos que utilizaram o subteste de leitura do TDE. Ao todo, foram encontrados 34 trabalhos, utilizando-se as bases *PsychInfo*, *Science Direct*, *Scielo* e *Google Acadêmico* no período de maio de 2007. Dos 34 estudos encontrados, apenas 16 (47%) realizaram análises com o subteste de leitura do TDE. A Tabela 3 resume os principais resultados encontrados por esses artigos científicos que utilizaram o subteste de leitura do TDE.

Dos 16 trabalhos que utilizaram o subteste, seis (37,5%) tiveram como objetivo investigar o efeito de variáveis independentes diversas (como a interação pai-filho, ou qualidade do ambiente escolar) no desempenho acadêmico de participantes, quatro trabalhos (25%) procuraram avaliar o efeito de treinamento em habilidades nesse desempenho e dois (12,5%) fizeram uma análise descritiva do desempenho de participantes no teste. Outros objetivos somam 25%, sendo quatro temas distintos: uma comparação do desempenho de participantes pela avaliação formal e pela avaliação feita pela professora; uma análise descritiva do desempenho no teste de uma população com dificuldades específicas (deficiência auditiva); uma comparação do desempenho escolar de crianças com e sem epilepsia e a relação entre o desempenho no teste e o desempenho em tarefa de leitura de palavras reais e de não-palavras isoladas. Desta forma, dos dois estudos que apresentaram análises descritivas do subteste, um

¹⁴ É importante ressaltar que a mesma ausência de relação não se aplicaria quando se considera as habilidades de resolução de problemas complexos de aritmética e de compreensão textual – pois a resolução desse tipo de problema exige a compreensão do enunciado, ou seja, requer a representação de problemas textuais que devem ser coordenados com as operações matemáticas a serem utilizadas (Kintsch & Greeno, 1985).

representa um caso clínico e o outro demonstrou uma piora significativa do desempenho dos participantes no subteste com a escolarização.

Tabela 3

Resumo das conclusões de trabalhos que relataram os desempenhos de participantes no subteste de leitura do TDE.

Estudo	Objetivo do estudo	Resultado (somente subteste de leitura)
Brancalhone et al. (2004)	Comparar o desempenho acadêmico de crianças expostas à situação de violência conjugal com o de crianças não-expostas a esta situação.	Não houve diferenças significativas no subteste entre as crianças do grupo exposto à violência e as crianças expostas à violência.
Capellini et al. (2004)	Comparar o desempenho de escolares com e sem dificuldades de aprendizagem em relação à avaliação formal e ao julgamento das professoras.	O subteste não discriminou o desempenho na leitura do grupo de crianças com dificuldades de aprendizagem em relação ao grupo sem dificuldades.
Capellini et al. (2003)	Avaliar a inclusão de alunos com deficiência auditiva a partir da comparação de seu desempenho com o de crianças normais.	Mais de 90% das crianças com dificuldade auditiva apresentaram desempenho abaixo do esperado pela série. O teste mostrou-se muito difícil para a amostra, devido ao alto ponto de corte para a classificação inferior, fazendo com que alunos com um percentual de acertos muito baixo ficassem no mesmo grupo de outros que tiveram resultados mais altos.
Cia (2005)*	Comparar o desempenho escolar aos hábitos e ambiente de estudo da criança, às condições de trabalho do pai, a fatores familiares e pessoais, relacionamento entre pai e filho e autoconceito da criança.	Não houve correlações entre a leitura e os fatores investigados (ambiente e hábitos de estudo, condições de trabalho dos pais e fatores familiares e pessoais). Todos os índices de relacionamento pai-filho utilizados (p. ex., nº de horas que o pai passa com o filho) se correlacionaram positivamente com o subteste.
Cia et al. (2004)	Estudar o impacto da qualidade da relação pai-filho sobre o desempenho acadêmico de escolares.	Não houve efeito de variáveis sociodemográficas (idade e sexo da criança, idade e anos de escolaridade do pai e renda familiar) nas correlações entre envolvimento paterno e o desempenho na leitura dos participantes. As três medidas de interação pai-filho utilizadas no estudo se correlacionaram com a leitura.

Tabela 3 (Continuação).

Estudo	Objetivo do estudo	Resultado (somente subteste de leitura)
D'Affonseca (2005)*	Investigar a relação entre a frequência das interações da mãe com os filhos com o desempenho acadêmico de crianças.	Não houve diferenças significativas entre as crianças cujas mães eram donas de casa e aquelas cujas mães trabalhavam fora. Quanto aos índices de interação mãe-filho, quanto maior a participação da mãe dona-de-casa nas atividades escolares, culturais e de lazer do filho, <i>pior</i> foi o desempenho na leitura.
Dias e Enumo (2004)	Investigar o efeito de um programa de intervenção em habilidades cognitivas e criativas em crianças com dificuldades de aprendizagem.	Não houve diferenças entre os grupos (experimental e de controle) na habilidade de leitura nem no pré-teste nem no pós-teste (após o treinamento).
Dias e Enumo (2006)	Avaliar os efeitos de um programa de promoção da criatividade no desempenho acadêmico, cognitivo e na criatividade de alunos com dificuldades de aprendizagem, divididos em dois grupos: G1 – que recebeu o treinamento em criatividade e G2 – que não recebeu o treinamento.	1) G1 X G2: Não houve diferenças entre os grupos nem no pré-teste nem no pós-teste; 2) pré-teste X pós-teste: em ambos os grupos, não houve diferenças no desempenho na leitura antes e após o programa de treinamento.
Dias et al. (2004)	Investigar os efeitos de um programa de criatividade sobre o desempenho acadêmico e cognitivo de alunos com dificuldade de aprendizagem, divididos em dois grupos: experimental (GE) e de controle (GC).	Tanto no GE quanto no GC, houve um aumento da classificação <i>inferior</i> em leitura no pós-teste em relação ao pré-teste. A classificação <i>superior</i> deixou de existir no pós-teste no GE e teve um aumento sutil no GC.
Dias et al. (2006)	Investigar o desempenho acadêmico de uma mostra de escolares no intervalo de um ano.	Houve uma piora significativa do desempenho na leitura nas duas avaliações (amostra geral). Para cada série separadamente, as diferenças entre a primeira e a segunda avaliação restringiram-se ao grupo de crianças mais velhas (houve uma piora significativa no desempenho). Não houve diferenças de gênero em nenhuma avaliação. Não houve diferenças significativas de desempenho entre as mesmas séries (3 ^a , 4 ^a e 5 ^a) para as duas avaliações.

Tabela 3 (Continuação).

Estudo	Objetivo do estudo	Resultado (somente subteste de leitura)
Elias e Marturano (2005)	Investigar os efeitos de uma intervenção baseada em princípios da aprendizagem mediada sobre o desempenho acadêmico e o comportamento de crianças com baixo desempenho escolar e problemas de comportamento.	Houve uma melhoria significativa no desempenho em leitura no pós-teste em relação ao pré-teste. O estudo não apresentou grupo de controle.
Fonseca et al. (2004)	Comparar o desempenho escolar de crianças com epilepsia benigna da infância com pontas centrotemporais (EBICT) com o de crianças saudáveis. Estudaram-se as relações entre o TDE e a lateralidade do foco e o número de descargas no eletrencefalograma.	As crianças com EBICT tiveram significativamente desempenho inferior em leitura. As crianças com desempenho inferior em leitura apresentaram maior número de descargas do que aquelas com desempenho médio e superior. O uso de medicação antiepiléptica não interferiu no desempenho. Conclui-se que o número de descargas, ao interferir com a função cerebral, pode ser um fator a explicar o desempenho mais baixo na leitura das crianças com EBICT.
Oliveira et al. (2005)	Descrever as alterações de linguagem em quatro membros de uma família com Síndrome Perisylviana, e relacioná-las a exames de neuroimagem.	Análise descritiva: a criança obteve desempenho inferior em leitura.
Pereira (2006)*	Caracterizar o desempenho escolar de crianças vitimizadas atendidas no Fórum Judicial.	Não houve diferenças significativas na leitura das crianças vitimizadas em relação ao grupo de controle.
Tonelotto et al. (2005)	Avaliar o desempenho escolar e de leitura (de palavras reais e de pseudopalavras isoladas) entre estudantes do Ensino Fundamental.	Houve uma associação entre o melhor desempenho em leitura do TDE e uma maior porcentagem de acertos e menor tempo de reação no reconhecimento de palavras e pseudopalavras.
Zanoti-Jeronymo e Carvalho (2005)	Comparar o desempenho escolar, o autoconceito e o comportamento de crianças filhas de pais alcoólicos (FA) com crianças cujos pais não são alcoólicos (FNA).	Os FA apresentaram um pior desempenho escolar em leitura quando comparados aos FNA.

Nota. Nos estudos marcados com o sinal * houve outros objetivos além da investigação do desempenho acadêmico dos participantes.

Esta revisão demonstrou que os estudos que dispomos com o subteste de leitura do TDE são pouco esclarecedores em relação à capacidade do teste em discriminar grupos, pois apenas dois dos 34 estudos encontrados investigaram diretamente as habilidades investigadas pelo TDE, os quais também não deixam de apresentar problemas que dificultam possíveis conclusões. Assim, conforme exposto anteriormente, o trabalho de Dias et al. (2006) é pouco conclusivo por utilizar uma amostra que apresenta uma grande amplitude de idade (amostra que, inclusive, ultrapassa a idade das crianças do estudo de normatização) enquanto que o de Tonelotto et al. (2005) tem como principal limitação o fato de os autores terem analisado os resultados sem considerar a série do participante, focando-se apenas na classificação descritiva de toda a amostra.

O resultado da busca tornou evidente a importância da condução de estudos que buscassem investigar de uma maneira cuidadosa o poder que o subteste de leitura do TDE apresenta para discriminar o desempenho dos indivíduos de diferentes níveis de aprendizagem. Tendo em vista este objetivo, Lúcio, Pinheiro e Nascimento (2007) investigaram o desempenho de 306 crianças de 1ª à 4ª série da Rede de Ensino de Belo Horizonte e descobriram que o critério de acerto apresentado no manual (em que devem ser consideradas corretas as respostas de silabação e de correção espontânea) reduz a variabilidade dos escores e aumenta o efeito de teto. Além disso, as autoras demonstraram que esse problema é também em grande parte devido ao desequilibrado nível de dificuldade das palavras, sendo que o teste apresenta uma quantidade excessiva de palavras fáceis e poucas palavras com um nível maior de dificuldade. Tal fato ficou demonstrado pela curva acentuadamente assimétrica à direita e da elevada proporção de acerto necessária para se atingir a classificação inferior. A grande quantidade de

palavras fáceis no teste ficou demonstrada pela alta proporção de palavras regulares para a leitura em relação às palavras irregulares.

A partir do que foi exposto, podemos notar que o subteste de leitura do TDE, por ser um instrumento que, como a própria autora define, fornece uma medida das habilidades básicas necessárias para o desempenho na leitura, talvez fosse mais bem descrito como um teste de prontidão. Assim, o subteste parece ser mais útil na identificação daqueles indivíduos que adquiriram a habilidade de decodificação na leitura do que para informar se o indivíduo adquiriu as habilidades de leitura exigidas em uma determinada série, tampouco para discriminar o quanto que essas habilidades foram de fato aprendidas. Nesse sentido, iniciativas para se criar um novo teste de leitura de palavras isoladas, como a que aqui se propõe a fazer, mostram-se tanto válidas quanto necessárias.

3 FASES DE CONSTRUÇÃO DE UM TESTE PSICOLÓGICO

3.1 Etapas na construção de testes

A construção de testes psicológicos constitui uma atividade complexa que exige a execução de uma série de atividades inter-relacionadas em que as decisões tomadas nas etapas iniciais interferem no curso dos procedimentos a serem adotados posteriormente (Adánez, 1999). Além disso, uma construção adequada de testes requer do pesquisador a aquisição de conhecimentos específicos da área de interesse, tanto de ordem teórica quanto prática, indo desde a definição clara do construto a ser avaliado, até o domínio de técnicas estéticas e da psicometria (Noronha, Freitas & Ottati, 2002).

Segundo Adánez (1999), o planejamento cuidadoso das etapas de construção de um teste é um pré-requisito fundamental para o sucesso de uma pesquisa. Nesse processo, o autor ressalta ainda a importância das etapas iniciais, como a construção dos itens ou a redação das instruções, as quais seriam frequentemente (e de uma maneira injustificada) negligenciadas por muitos pesquisadores, o que comprometeria de uma forma irremediável o produto final ou a qualidade das pesquisas.

Apesar de a construção de um teste não seguir etapas rigidamente pré-estabelecidas, muitos pesquisadores (p. ex., Adánez, 1999; Oakland, 1999; Pasquali, 1996; Urbina, 2007) têm se dedicado à descrição das fases envolvidas nesse processo. Adánez (1999), por exemplo, expõe 13 etapas a serem consideradas e as divide em dois estágios distintos. Desta forma, as sete primeiras etapas seriam partes de um processo de planificação, elaboração e correção do material básico de um teste, cujo critério norteador seria a teoria. Em contrapartida, nas etapas seguintes teriam peso os critérios empíricos, referindo-se à obtenção de dados da prova experimental com a finalidade de

selecionar os itens que mais bem se ajustem às características planificadas. As etapas descritas por Adánez são resumidas na Tabela 4:

Tabela 4

Etapas de construção de um teste descritas por Adánez (1999)

Etapa	Descrição
1. Definição dos objetivos do teste	Constitui a identificação dos usos ou a natureza das inferências que se pretende fazer com a pontuação dos testes. São geralmente empregados três objetivos de testagem, sendo que dois colocam ênfase na característica pragmático-utilitária dos testes (avaliação do rendimento e predição) e um na relevância teórica (medição de um traço psicológico). Esta fase teria repercussões na identificação das características que os itens deverão ter (etapa 4) e no tipo de instruções a serem dadas na tarefa (etapa 5).
2. Especificação do contexto	Observação das características da população a que o teste se dirige e das condições de aplicação (se individual ou coletiva, usos do tempo).
3. Eleição do modelo matemático	A escolha pela Teoria Clássica dos Testes (TCT) ou pela Teoria de Resposta ao Item (TRI), que apresenta repercussões na construção e nos critérios de seleção dos itens, tendo também impacto no tamanho da amostra a ser avaliada.
4. Definição do domínio	Etapa em que se define o construto que se pretende avaliar, delineando-se de maneira precisa as características dos indicadores que serão utilizados para representar o domínio (conteúdo, padrões característicos, etc.), de tal maneira que seja possível amostrá-lo adequadamente por meio dos itens do teste.
5. Construção dos itens e das instruções	A partir do que foi definido pelas etapas anteriores, constrói-se os itens e, posteriormente, as instruções. Estas últimas devem ser formuladas de tal forma a garantir que cada participante entenda claramente o que deve fazer, além de motivá-los a conseguir seu máximo nível de execução, ou responda com sinceridade às questões.
6. Revisão do teste inicial por especialistas	É conveniente consultar especialistas para que julguem aos itens e às instruções em aspectos tais como a precisão dos enunciados, a adequação do vocabulário empregado, a pertinência do material a respeito do domínio previamente definido, etc.
7. Estudo piloto	A aplicação piloto do teste a uma pequena amostra da população em estudo pode ser útil para verificar aspectos pragmáticos como: se os itens e a forma de responder aos mesmos são bem compreendidos, o tempo de aplicação do teste, o cálculo do tamanho amostral, etc.
8. Seleção de amostras e aplicação do teste	É necessário garantir que o teste será administrado a sujeitos característicos de todos os níveis do traço medido e a distintos setores da população com diferentes níveis em tais atributos, garantindo-se, assim, a representatividade da amostra.

Tabela 4 (Continuação).

Etapa	Descrição
9. Análise e seleção empírica dos itens	Seleção de itens baseada na TCT ou na TRI. No primeiro caso, a análise se baseia no índice de dificuldade e discriminação dos itens. Comumente, emprega-se uma análise da consistência interna, que fornece um índice de fidedignidade dos itens e também fornece subsídios para a validação de construto. A TRI se efetua por meio do ajuste dos itens ao modelo, seleção dos itens adequados e calibração do banco de itens.
10. Avaliação da fidedignidade do teste	As formas mais comuns de se obter o índice de fidedignidade são por meio de formas paralelas, teste-reteste, duas metades e o Coeficiente de fidedignidade Cronbach.
11. Avaliação da validade do teste	Atualmente, as pesquisas com testes psicológicos têm se concentrado na validação de construto, que é o grau em que um teste mede um traço teórico. Nesse sentido, a validade não pode se resumir a uma só medida, mas deve basear-se em evidências múltiplas, progressivas e convergentes. Apesar desse reconhecimento crescente, as técnicas empregadas são comumente divididas em três categorias: correlacionais, diferenciais e experimentais.
12. Elaboração das normas	A maior parte dos testes TCT usa <i>normas de grupo</i> , que se baseiam em uma comparação da pontuação de um sujeito com a distribuição de pontuações do grupo a que o sujeito pertence. A maior parte das normas é apresentada mediante a transformação das pontuações diretas em escalas como os percentis, escores <i>z</i> e pontuações típicas normalizadas derivadas.
13. Redação do manual de uso	Devem ser especificadas todas as características do teste: fundamentos teóricos, objetivos, possíveis usos, população a que se dirige, recursos necessários para a aplicação, procedimento de aplicação e correção, normas, validade, fidedignidade, referências bibliográficas e qualquer outra recomendação necessária de uso correto.

Obviamente, a divisão proposta por Adánez (1999) possui um caráter mais didático do que propriamente um caminho a ser seguido para a construção de testes psicológicos. Mesmo assim, serão resumidos a seguir os procedimentos de pesquisa utilizados anteriormente ao estágio atual da construção da prova computadorizada de reconhecimento de palavras que comporá a bateria AVACLE considerando-se as etapas descritas por Adánez. Posteriormente, serão apresentadas as etapas a serem cumpridas ao longo dessa dissertação, também se tendo como referência as fases propostas por este autor.

3.2 Procedimentos anteriores

O conjunto de provas (em processo de desenvolvimento) sob a denominação “AVACLE” consiste de uma bateria de avaliação cognitiva composta de tarefas de leitura (de aplicação individual) e de escrita (de aplicação coletiva) para crianças. No que diz respeito especificamente à avaliação da leitura, a bateria é composta por tarefas de decisão (lexical e semântica) e de leitura em voz alta de palavras reais e de não-palavras. Esta dissertação trata dos estágios iniciais da investigação psicométrica da tarefa de leitura em voz alta de palavras reais isoladas da bateria AVACLE. Esta tarefa, doravante denominada prova computadorizada de reconhecimento de palavras para crianças, tem como foco principal a descrição da seleção de itens candidatos, que serão posteriormente aplicados a uma nova amostra a fim de se selecionar os itens que farão parte da versão final da referida tarefa. Isso porque, de acordo com muitos autores, a validação de um instrumento ou medida se coroa quando existe pelo menos uma segunda tentativa (estudo) que se ajuste às descobertas da primeira pesquisa.

De acordo com as duas primeiras etapas de construção de um teste proposta por Adánéz (1999), a seleção de itens a compor a prova computadorizada de leitura foi planejada para atender tanto a propósitos de ordem teórica quanto de ordem prática. No primeiro caso, a prova tem como objetivo testar a rota lexical do modelo de dupla-rota (Coltheart, Rastle, Perry, Langdon & Ziegler, 2001). No segundo, o objetivo é conduzir uma análise psicométrica de um banco de itens que será uma etapa em direção da construção de prova computadorizada de reconhecimento de palavras para crianças que será uma medida que fornecerá as normas de desempenho em leitura de crianças cursando as séries iniciais do Ensino Fundamental. Isso tornará possível o diagnóstico das capacidades atuais de reconhecimento de palavras dessa população.

Para alcançar o objetivo de avaliar os componentes do modelo, é necessário investigar a emergência dos efeitos de regularidade, extensão e frequência na amostra em estudo (ver seção 2.2). Essa investigação somente é possível por meio da manipulação das características psicolinguísticas das palavras, as quais são: o nível de regularidade, o tamanho e a frequência de ocorrência das palavras.

Para investigar o efeito de regularidade, Pinheiro (2004, 2007b) classificou as palavras quanto ao nível de regularidade grafema-fonema e fonema-grafema, o que será descrito na próxima seção. Para investigar o efeito de extensão, foram escolhidas palavras de 4 a 8 letras. Para que a investigação do efeito de frequência fosse possível, Pinheiro (1996) conduziu uma contagem de frequência de ocorrência de palavras expostas a crianças cursando da 1ª à 4ª série do Ensino Fundamental, criando um banco de dados capaz de fornecer as frequências relativa e absoluta das palavras e, também, de classificar essas frequências como alta, média ou baixa para a população em questão.

No presente estudo, contrariando as recomendações de Parente, Hosogi e Lecours (1997), foram selecionadas no corpus de Pinheiro (2004, 2007b) apenas palavras de baixa frequência. No entanto, tal escolha se justifica por duas razões, uma de ordem empírica e outra de ordem prática. Do ponto de vista empírico, ressaltamos anteriormente (seção 2.3) que o nível de regularidade não afetou a leitura das palavras de alta frequência no estudo de Pinheiro (1995), nem mesmo entre os leitores iniciantes, havendo apenas uma presença restrita do efeito de regularidade nas palavras de baixa frequência de ocorrência. Sendo assim, o conhecimento prévio do efeito da frequência de ocorrência das palavras na produção de erros e no tempo de reação relatados naquele trabalho norteou a seleção de apenas as palavras de baixa frequência no presente estudo. Do ponto de vista prático, Pinheiro e Rothe-Neves (2001) chamam a atenção para um problema de ordem metodológica na construção de listas pareadas de palavras de baixa

e de alta frequência para a avaliação do efeito de regularidade na leitura. Isso ocorre porque o universo de palavras de alta frequência é muito pequeno, tornando inviável o pareamento dessas palavras com as de baixa frequência em termos de regularidade, número de grafemas e estrutura silábica. Por esses motivos, neste estudo decidiu-se utilizar apenas palavras de baixa frequência e, após a seleção daquelas palavras que se mostrarem mais adequadas em termos de dificuldade e discriminação, será finalmente feito o pareamento com as palavras de alta frequência que constam do corpo de itens original de Pinheiro (1996).

Com o objetivo de selecionar os itens que farão parte da tarefa, planejou-se inicialmente avaliar 400 crianças de 1ª à 4ª série de escolas públicas (estaduais e municipais) e particulares. Em função do número limitado de participantes em relação ao total de palavras presentes no banco de itens, optou-se por uma abordagem estatística clássica de tratamento de dados em detrimento da análise via teoria de resposta ao item. Nesse último caso, seria necessário avaliar pelo menos 10 participantes para cada item do banco (resultando em uma amostra em torno de 3.000 indivíduos), o que obviamente tornaria inviável a coleta de dados¹⁵.

Em síntese, podemos observar que todas as sete etapas iniciais propostas por Adánéz (1999) (com exceção da etapa 6, que não se aplica a este estudo) já foram cumpridas e são aqui sintetizadas: 1) definição dos objetivos da testagem – testar o modelo de dupla-rota (objetivo teórico) e criar um instrumento capaz de servir como auxílio para o diagnóstico de problemas de aprendizagem da leitura (objetivo de ordem prática); 2) especificação do contexto – criar uma prova computadorizada de aplicação individual dirigida a crianças de 1ª à 4ª série do Ensino Fundamental; 3) especificação do modelo matemático – em função do tamanho da amostra, escolha pela TCT; 4)

¹⁵ Por exemplo, Nunes & Primi (2005) mostraram que em um teste contendo 17 itens, o tamanho mínimo da amostra para uma estimativa confiável da calibração dos itens e das habilidades dos testandos foi de 200 participantes, o que confere uma média de 11 participantes por item

definição do domínio – as palavras isoladas devem variar em termos regularidade e extensão para atender aos objetivos propostos, sendo que pelas razões empíricas e práticas optou-se pelo uso exclusivo de palavras de baixa frequência; 5) construção dos itens e das instruções – as palavras foram classificadas em termos da relação grafema-fonema e fonema-grafema e número de letras e as instruções foram redigidas; 7) aplicação do banco de itens em um estudo piloto – um estudo preliminar foi conduzido no qual se constatou a adequação das instruções, do equipamento, do procedimento de aplicação e também tornou possível a familiarização com a linguagem do *software* utilizado para aplicação da tarefa.

3.3 Proposta de trabalho

Das fases de construção de testes descritas por Adánez (1999), a presente dissertação propõe relatar os procedimentos referentes à etapa oito (seleção de amostras e aplicação do banco de itens) e nove (análise e seleção empírica dos itens). Conforme dito acima, será feita uma seleção dos itens candidatos a compor a tarefa idealizada. No entanto, a título de exemplo, propõe-se apresentar uma simulação de duas amostras reduzidas e paralelas de itens, a partir das quais se descreverá as etapas 10 (fidedignidade) e 11 (avaliação da validade). Os procedimentos de seleção da amostra e de aplicação da tarefa serão descritos em detalhes no método (seção 5).

A análise de itens é um termo genérico utilizado para se referir a um conjunto de procedimentos estatísticos empregados para averiguar a qualidade dos itens de um teste, de tal maneira que os escores dos sujeitos nos itens variem de acordo com os propósitos da avaliação (Anastasi & Urbina, 2000). Três procedimentos clássicos foram aqui escolhidos para tal fim, a saber, a dificuldade, a discriminação e a consistência interna

dos itens. A seleção dos itens terá por base a melhor adequação a estas três medidas, as quais são descritas brevemente a seguir.

A dificuldade do item, definida como a proporção de indivíduos da amostra que o acertam, depende tanto de suas características intrínsecas, quanto do nível de habilidade de quem está sendo avaliado (Urbina, 2007). No processo de construção de um teste, a escolha de um nível de dificuldade adequado é essencial para a avaliação das diferenças individuais, uma vez que a manutenção de itens com dificuldades extremamente elevadas ou baixas é pouco ou nada informativa sobre aquilo que os indivíduos são capazes de fazer. Nesse sentido, um dos objetivos da análise dos itens é a exclusão de itens que quase ninguém acerta ou erra, itens estes que não afetam a variabilidade dos escores (Anastasi & Urbina, 2000; Martin & Pratt, 2001).

A discriminação constitui a extensão em que um item diferencia adequadamente os testandos no domínio de comportamento que o teste propõe medir (Pasquali, 1996). Em outros termos, esse índice informa o grau em que um determinado item serve para distinguir os indivíduos que obtêm pontuações altas e os que obtêm pontuações baixas no teste (Bisquerra et al., 2004). O poder de discriminação de um item pode ser obtido por meio do uso de critérios internos ou externos ao teste. No primeiro caso, uma medida bastante utilizada é a consistência interna, que considera a variância comum entre os itens de um teste (Pasquali, 2003). Além de fornecer dados sobre a precisão do item, esta técnica é capaz de dizer da validade dos mesmos, uma vez que oferece um índice da unidimensionalidade do construto que está sendo avaliado. No caso do uso de critérios externos (definidos como medidas independentes dos escores de testes e que estes pretendem avaliar ou prever), são utilizados os mesmos índices de critérios destinados a validar o teste como um todo, como a pertença a determinado grupo diagnóstico ou faixa etária (Urbina, 2007).

A fidedignidade (ou precisão) refere-se à estabilidade dos escores dos indivíduos testados em diferentes ocasiões em um mesmo teste ou em um conjunto equivalente de itens. Ela representa o grau em que podemos confiar que as diferenças apresentadas nos escores de um teste entre uma avaliação e outra devem ser atribuídas a erros casuais ou aleatórios, ou seja, ela permite estimar uma margem confiável de variação dos escores dos indivíduos em um teste. Existem diversas formas de se avaliar a fidedignidade de um instrumento psicológico, a qual pode fornecer uma medida da variância de erro que é atribuída ao tempo, à amostragem de conteúdo e à concordância entre avaliadores (Anastasi & Urbina, 2000).

Nesse trabalho, são apresentadas evidências de dois tipos de fidedignidade para as amostras reduzidas de itens que serão simuladas: a concordância entre avaliadores e a fidedignidade de forma alternativa com intervalo. As diferenças entre os avaliadores constituem fontes de erros nos escores do teste que se devem a elementos de subjetividade na avaliação (Urbina, 2007). No que se refere à tarefa de leitura de palavras, a maioria das respostas dos participantes não representa fonte de ambigüidades, uma vez que elas apresentam pronúncias pré-determinadas. Entretanto, em alguns casos, a conferência de escores às respostas depende do julgamento do avaliador. Isto é o que ocorre com as respostas de leitura silabada (na qual o participante lê com pausas entre as letras ou as sílabas que compõem as palavras) que foram consideradas incorretas nesse estudo e as respostas que representam variações regionais ou de grupos sociais, as quais devem ser consideradas corretas. Desta forma, depende do avaliador julgar se a pronúncia de uma resposta é suficientemente contínua a ponto de não ser enquadrada como silabação e, portanto, como um erro, e também decidir se uma resposta denota uma variação lingüística (por exemplo, ler a palavra *flor* como “frô”) ou um erro de pronúncia propriamente dito (por exemplo, ler *filtro* como

“friuto”). Uma grande concordância entre diferentes avaliadores a respeito dos escores dos participantes indicará que os erros de medida então obtidos provavelmente não se devem a esta fonte.

A fidedignidade de forma alternativa com intervalo consiste na aplicação de formas alternativas de um mesmo teste com uma lacuna de tempo entre as duas aplicações (Urbina, 2007). O banco de palavras aplicado às crianças, por ser bastante grande, foi dividido em duas listas. Estas foram pareadas, na medida do possível, quanto às características que se supõe interferir na dificuldade dos itens (como o comprimento, a estrutura silábica, etc.), o que faz com que constituam, pelo menos em tese, formas alternativas. Como as listas foram aplicadas em dias diferentes, medidas preliminares de estabilidade temporal podem ser obtidas por meio da técnica acima descrita, se confirmada a hipótese de que as listas são, de fato, formas alternadas (o que também fornece indícios a respeito da consistência interna dos itens sob investigação).

O cálculo da fidedignidade por meio da técnica de forma alternativa com intervalo está sujeito a fontes de erros devidos a efeitos de prática. A redução das correlações entre um momento e outro da avaliação é capaz de revelar a influência desse efeito nos escores. Outro fator importante a ser considerado é o efeito de prática *dentro de uma mesma aplicação da tarefa*. Em uma mesma aplicação as crianças leram uma grande quantidade de itens, o que pode ter gerado um efeito de prática dos itens iniciais aos finais. A contrapartida desse efeito é a fadiga, ou seja, uma redução no desempenho com o decorrer da avaliação. A característica da avaliação aqui conduzida permite a possibilidade de ambos os tipos de fontes de erros, sendo, portanto, alvo de nossa investigação.

Por fim, o presente trabalho pretende apresentar um procedimento capaz fornecer indícios para a validade do banco de itens sob investigação. Segundo a

perspectiva atual, a validade se refere ao grau em que as evidências empíricas e a teoria dão suporte às interpretações dos escores de um teste. Isto é, o processo de validação está relacionado aos propósitos ou aos usos da avaliação psicológica em um determinado contexto, não se referindo a uma qualidade que os testes apresentam de uma maneira ampla (AERA, APA, NCME, 1999). Isso implica que as informações ou tipos de evidências a serem coletadas a favor da validade irá depender das interpretações que se propõe fazer em relação aos escores do teste.

Conforme visto na seção 2.2, o modelo de dupla-rota pressupõe que a leitura pode ser feita por meio de dois processos distintos, sendo que as respostas dos indivíduos podem variar em função de qual mecanismo é utilizado para ler. Isso implica que a presença de efeitos na leitura constituem fontes de validação da teoria que pressupõe o uso desses mecanismos. Desta forma, a validade do banco de itens aplicado na amostra será investigada por meio da emergência de efeitos na leitura, efeitos estes que se devem tanto às características do estímulo quanto dos participantes (seu nível de desenvolvimento, padrão econômico, entre outras características que se tem relacionado ao sucesso na leitura). Para esta finalidade, será utilizado o modelo marginal produzido pelo método de Estimção de Equação Generalizada (GEE) para determinar o peso relativo de cada covariável selecionada na proficiência na leitura (entende-se por proficiência na leitura, índices satisfatórios ou médios de precisão e tempos de reação). Modelos diferenciados serão reportados para a precisão (modelo GEE logístico) e para o tempo de reação (modelo GEE linear). O método GEE fornece modelos semelhantes ao da regressão hierárquica, com a diferença que não apresenta suposições com relação à distribuição dos efeitos (ou seja, é adequado a dados não paramétricos). A escolha dessa técnica se deve a dois fatores: (1) os modelos GEE são capazes de considerar as múltiplas observações de um mesmo sujeito (no nosso caso, as palavras) como

pertencentes a um grupo único, o qual é independente das observações dos demais grupos de sujeitos; (2) modelos hierárquicos são utilizados para determinar os melhores preditores de um critério, ou seja, servem como métodos de validação de modelos teóricos cujas relações entre variáveis já são conhecidas em termos de magnitude e direção (Abbad e Torres, 2002; Hu FB, Hedeker, Flay, & Pentz, 1998; Twist, 2004).

Um último ponto merece ser mencionado aqui. Apesar de os procedimentos de análise de itens, fidedignidade e validação terem sido relatados em momentos separados, a visão atual a respeito do processo de construção de instrumentos de avaliação psicológica é a de que todas as fontes de evidências coletadas a favor da interpretação de um escore acrescentam indícios à validade do mesmo (AERA, APA, NCME, 1999). Isso implica que toda a investigação psicométrica conduzida nesse trabalho poderá servir como fonte de acúmulo para a validade das inferências que estão sendo feitas aqui. Por exemplo, é sabido que a análise de itens contribui para a validade das interpretações que podem ser feitas a partir dos escores dos testes (Adánez, 1999; Cascio & Aguinis, 2005; Mitchell & Klimoski, 1986). De maneira análoga, Urbina (2007) destaca que a fidedignidade dos escores de um teste pode contribuir para a validade desses escores de várias maneiras:

Se (...) o teste for delineado para avaliar um construto unidimensional (...), altos índices de consistência interna iriam confirmar a alegação de unidimensionalidade. Da mesma forma, se for obtida uma consistência de escores entre diferentes avaliadores, pode-se supor que todos eles estão empregando os mesmos critérios e, assim, provavelmente avaliando as mesmas características. Se o construto a ser avaliado for estável (...) uma alta fidedignidade teste-reteste nos escores seria um pré-requisito essencial para evidências de validade. (p. 172).

Nesse sentido, e por razões didáticas, optou-se por preservar a divisão clássica entre os procedimentos psicométricos comumente utilizados no processo de construção de testes psicológicos. Ademais, apesar de neste trabalho se propor coletar evidências a respeito da validade de construto da tarefa, isso não implica que não seja necessária a condução de investigações de outras fontes de evidência de validade da mesma. Por exemplo, os *Standards* (AERA, APA, NCME, 1999) descrevem cinco fontes de evidência de validade, a saber, as evidências baseadas no conteúdo do teste, nos processos de resposta, na estrutura interna do instrumento, nas conseqüências da avaliação e nas relações entre o teste e outras variáveis. Estas últimas fontes de evidência ainda são subdivididas em vários mecanismos de exploração da validade, os quais são: a validade convergente e a discriminante, as relações entre o critério e o teste (o que é conhecido como validade de critério) e a generalização da validade. Conforme se vê, o presente trabalho não pretende, e nem poderia, esgotar a questão a respeito da validade da tarefa cuja investigação psicométrica está sob investigação.

4 CLASSIFICAÇÃO DAS PALAVRAS QUANTO AO NÍVEL DE REGULARIDADE E HIPÓTESES GERADAS A PARTIR DESSA CLASSIFICAÇÃO

O objetivo dessa seção é descrever a classificação de regularidade bidirecional de palavras que foi introduzida por Pinheiro (2004, 2007b). Essa classificação consistiu na identificação da relação grafofonêmica (direção da leitura) e fonografêmica (direção da escrita) de todas as consoantes e vogais das palavras portuguesas e se baseou, entre outros, nos trabalhos de Lemle (1981), Souza (1978) e de Parente, Silveira e Lecours (1997). Para todos esses autores, as relações grafofonêmicas e fonografêmicas podem ser regulares (independentes ou dependentes de contexto) ou irregulares (também referidas na literatura como ambíguas). A classificação de palavras sob consideração estabelece a distinção entre as pronúncias que são dependentes de contexto daquelas que são independentes de contexto. Inova por classificar cada palavra tanto do ponto de vista de sua regularidade para a leitura quanto para a escrita, por considerar nas palavras irregulares a posição da correspondência irregular e por identificar uma nova categoria que é formada de palavras em que se encontram uma combinação de uma (ou mais) correspondência irregular com uma (ou mais) correspondência regida por regras.

Desta forma, para cada grafema ou fonema das palavras do corpo de itens considerado, Pinheiro (2004, 2007b) atribuiu um nível de regularidade e, ao final, cada palavra foi classificada como: (a) regular – todos os grafemas ou fonemas da palavra são independentes de contexto; (b) regra – pelo menos um dos grafemas ou fonemas da palavra dependem de regras contextuais que governam a pronúncia; (c) irregular – pelo menos um dos fonemas ou grafemas da palavra é ambíguo e (d) irregular-regra – pelo menos um dos fonemas ou grafemas da palavra é irregular e, simultaneamente, pelo

menos um dos fonemas ou grafemas da palavra apresenta pronúncia dependente de contexto.

Cabe ressaltar que a regularidade da palavra para a leitura é *independente* de sua classificação para a escrita e vice-versa, isto é, uma palavra regular para a leitura pode ser regular, regra, irregular ou irregular-regra para a escrita, e assim por diante. A Tabela 5 mostra todas as variações possíveis da classificação das palavras para a leitura (relação grafema-fonema) e para a escrita (relação fonema-grafema) por meio de exemplos de palavras utilizadas no estudo. Note que as correspondências fonema-grafema irregulares ou regidas por regras contextuais estão sublinhadas e seus equivalentes, na leitura, acham-se destacadas em negrito.

O uso de quatro categorias da maneira como foram descritas difere de todos os estudos anteriormente conduzidos e esse procedimento se justifica pela possibilidade de que apenas alguns estímulos sejam suficientemente difíceis a ponto de produzirem o efeito de regularidade. Esta é uma hipótese plausível em função da variação encontrada no efeito de regularidade de acordo com o número de categorias que a pesquisa utiliza (ver seção 2.3). Assim, por meio da referida classificação, espera-se identificar os estímulos responsáveis pela emergência do efeito de regularidade na leitura. Além disso, a comparação do efeito de regularidade obtido com o uso de duas ou quatro categorias de regularidade permitirá verificar a validade empírica da hipótese de que a emergência desse efeito sofre influência do número de categorias utilizadas na classificação de palavras.

Tabela 5

Exemplos de palavras em cada categoria de classificação para a leitura, em função da classificação para a escrita.

Classificação para a escrita	Classificação para a leitura			
	R	IR	IRRG	RG
R	bala	moda	amora	morada
RG	fum <u>o</u>	bote <u>e</u>	soro <u>o</u>	caro <u>o</u>
IR	raça	<u>c</u> ela	<u>c</u> era	ra <u>s</u> a
IRRG	recurso <u>o</u>	<u>c</u> esto	aces <u>o</u>	espiral <u>i</u>

Nota. R = regular; RG = regra; IR = irregular; IRRG = irregular-regra. Em negrito: irregularidades e regras contextuais para a leitura. Sublinhado: irregularidades e regras contextuais para a escrita.

O presente estudo adota classificação de Pinheiro (2004, 2007b) para as palavras que compõem o banco de itens sob investigação, com apenas algumas alterações. A principal mudança diz respeito à classificação das vogais, que no presente estudo está de acordo com aquela apresentada por Silveira (1986)¹⁶. As demais modificações se referem a casos particulares, e não a um contexto geral, como a opção pela classificação de Silveira (1986) aqui apontada. (Para maiores detalhes sobre os casos específicos, ver nota de rodapé n.º 23).

Em função do que foi exposto, conclui-se que os itens presentes na tarefa de leitura que está em construção foram escolhidos segundo uma estratégia racional

¹⁶ Em relação ao número de fonemas vocálicos no português brasileiro, existe uma discordância entre os lingüistas, sendo que um grupo considera a existência de 26 fonemas (grupo de Matoso Câmara, 1985) enquanto outro considera a existência de 33 fonemas (grupo de Silveira, 1986). Segundo Parente, Silveira e Lecours (1997), a abordagem de Silveira (1986) não modifica as irregularidades nas conversões, sendo que apenas aumenta o número de regras contextuais. No entanto, conforme dito acima, no presente trabalho, a classificação de palavras diferencia entre as pronúncias dependentes de contexto das independentes de contexto, o que nos fez optar pela classificação de Silveira (1986).

(Cascio & Aguinis, 2005; Mitchell & Klimoski, 1986), ou seja, partiu-se do pressuposto de que a classificação de regularidade das palavras é determinante para o nível de dificuldade que o item apresenta para o leitor. Assim sendo, acredita-se que as palavras regulares e regra para a leitura serão fáceis de serem lidas, enquanto que as palavras irregulares e irregulares-regra serão mais difíceis. No entanto, considerando-se o caráter altamente previsível de nossa língua, a despeito de as palavras irregulares imporem maiores dificuldades à leitura, é possível que as palavras difíceis não imponham um nível de dificuldade suficientemente grande a ponto de gerarem efeito de solo entre esses itens.

5 OBJETIVOS

5.1 Objetivo geral:

- Conduzir a investigação psicométrica de um banco de itens que oferecerá os subsídios para a construção de versões paralelas da prova computadorizada de reconhecimento de palavras para crianças que comporá a bateria AVACLE.

5.2 Objetivos específicos:

Os objetivos específicos serão reportados em função da amostra de itens que será trabalhada. Os objetivos descritos em 5.2.1 e 5.2.2 se referem às análises feitas com todo o banco de itens lido pela amostra de crianças. Em 5.2.3 são apresentados os objetivos que serão realizados com a simulação da amostra reduzida de itens, a qual foi gerada a partir dos itens-candidatos obtidos em 5.2.1.

5.2.1 Análise de itens

- Selecionar, dentre um amplo corpo de itens, as palavras que apresentarem níveis de discriminação e de dificuldade satisfatórios para a avaliação da habilidade de reconhecimento de palavras isoladas. Esta seleção inicial buscará identificar as palavras-candidatas a fazerem parte de um teste de leitura em voz alta – a prova computadorizada de reconhecimento de palavras para crianças –, cujo desenvolvimento se dará em um estudo posterior.

5.2.2 Efeito de fadiga

- Investigar o efeito de fadiga ou de treino no tempo de reação e na precisão da leitura do banco de itens aplicado à amostra.

5.2.3 Efeitos na leitura

- Avaliar o impacto do nível de regularidade grafema-fonema e fonema-grafema na emergência do efeito de regularidade nos erros e no tempo de reação da amostra de crianças.
- Investigar a influência relativa do número de categorias de regularidade utilizadas na classificação das palavras para a emergência do efeito de regularidade na leitura.

5.2.4 Investigação da fidedignidade a partir de simulação de amostra de itens-candidatos

- Simular uma seleção de itens baseando-se na amostra de itens-candidatos.
- Avaliar a consistência interna e temporal das amostras simuladas de itens obtidas;
- Avaliar a precisão na correção dos protocolos de avaliadores da mesma amostra de itens.

6 MÉTODO

Este trabalho é parte de um projeto maior, intitulado “Avaliação de Competências Psicológicas da População Infanto-Juvenil de Belo Horizonte: Inteligência e Habilidades de Leitura e Escrita” (Pinheiro, 2007a). Este foi financiado pela FAPEMIG e em processo de desenvolvimento pelo grupo de pesquisa da professora Ângela Pinheiro do Departamento de Psicologia da UFMG. O projeto foi aprovado pelo Comitê de Ética em Pesquisa com Seres Humanos (COEP) da UFMG em 09/03/2005 (processo n.º ETIC 347/04). No Apêndice A encontra-se a cópia da aprovação do projeto pelo comitê.

6.1 Características do estudo

Este é um estudo transversal, do tipo descritivo, com delineamento exploratório e comparativo.

6.2 Amostra

A seleção das escolas foi feita com base nos dados fornecidos pela Secretaria de Educação do Estado de Minas Gerais. Tomando-se a cidade de Belo Horizonte, foi feito um levantamento de todas as escolas (estaduais, municipais e particulares) existentes nessa cidade. A escolha da quantidade de escolas participantes por tipo (estadual, municipal e particular) seguiu o critério de representar proporcionalmente a quantidade de alunos existentes em cada um desses tipos.

O processo de seleção das escolas seguiu dois critérios: (1) escolher escolas pertencentes às zonas de conveniência dos aplicadores dos instrumentos da pesquisa, a saber, escolas situadas nas áreas da Pampulha, do Centro e Nordeste da cidade¹⁷ e (2) concordância da escola selecionada em participar da pesquisa por meio do termo de consentimento (Apêndice B).

Considerando as áreas escolhidas, oito escolas foram contatadas, dessas apenas duas se recusaram a participar. Sendo assim, a amostra de escolas foi formada por 6 (seis) escolas, sendo quatro estaduais, duas municipais e uma particular. A Tabela 6 apresenta os principais dados de todas as escolas que participaram do estudo.

A seleção das crianças foi feita a partir de critérios pré-estabelecidos, que foram passados para as professoras. Cada professor recebeu uma carta (Apêndice C) em que esses critérios eram explicitados, de forma a proceder à seleção dos alunos que poderiam participar da pesquisa.

Os alunos que satisfizeram os critérios estabelecidos receberam os termos de consentimento livre e esclarecido (Apêndice D), que lhes foram entregues pelos dirigentes de suas respectivas escolas. Em cada série, foram distribuídos de 15 a 20 termos de consentimento. Os alunos foram instruídos a entregar aos pais a autorização e a devolvê-las assinadas. As escolas que apresentavam mais de uma turma de uma mesma série geralmente distribuíram os termos entre as diferentes turmas. Somente foram avaliadas as crianças que devolveram os termos de consentimento assinados por um dos pais ou responsáveis.

¹⁷ Em função da falta de recursos para subsidiar o deslocamento dos membros da equipe responsável pela coleta de dados, a escolha das escolas que compuseram a amostra deste estudo tomou como referência a sua proximidade com a UFMG (onde os aplicadores dos testes estudam) e a facilidade de deslocamento entre a residência dos aplicadores e o local de coleta de dados. Portanto, a amostra do presente estudo, apesar de heterogênea, é de conveniência.

Tabela 6

Perfil socioeconômico das escolas participantes do estudo.

Escola	Tipo	Zona	Renda^{18, a}	Mínimo^{a,b}	Máximo^{a,b}	Moda¹⁹
Escola 1	Estadual	Centro-sul	1326,20	927	1699	C
Escola 2	Estadual	Centro-sul	916,32	424	4648	C e D
Escola 3	Estadual	Centro-sul	1611,68	424	4648	C
Escola 4	Estadual	Pampulha	1724,15	927	4648	B2
Escola 5	Municipal	Nordeste	2698,54	927	7793	B1
Escola 6	Municipal	Pampulha	635,32	424	927	D
Escola 7	Particular	Pampulha	4265,90	927	7793	A2

^a Valores em reais. ^b Renda familiar mínima e máxima encontrada em cada escola.

Ao todo, 352 crianças devolveram os termos de consentimento assinados. Destas, 340 foram avaliadas. Desses participantes, sete foram excluídos (por falhas técnicas apresentadas pelo equipamento ou pela falta de aplicação de uma das listas). Sendo assim, a amostra do presente estudo conta com um total de 333 participantes (média de idade = 8,77, D.P. = 1,14, 52% do sexo feminino), distribuídos por série e escola conforme mostra a Tabela 7.

¹⁸ A renda média do público-alvo da escola foi calculada por meio da média dos rendimentos das famílias que responderam ao questionário socioeconômico. Esse valor foi obtido por meio da tabela de cálculo fornecida pelo Critério Brasil que, além classe social, fornece o rendimento médio da família em função dessa classe. As classes sociais variam de A1 (nível mais alto) a E (nível mais baixo). Desta forma, para cada família, foi obtido um rendimento médio baseado na classificação socioeconômica gerada pelo questionário. Em seguida, calculou-se a média dos rendimentos de todas as famílias que responderam ao questionário, o que pode ser visualizado na coluna quatro da Tabela 6.

¹⁹ Classificação socioeconômica mais freqüente das famílias.

Tabela 7

Quantidade de alunos por série e escola na prova que leram o banco de itens testado.

Escola	Série				Total
	1 ^a	2 ^a	3 ^a	4 ^a	
Escola 1	10	10	12	09	41
Escola 2	11	07	10	10	38
Escola 3	—	09	11	10	30
Escola 4	16	15	20	15	66
Escola 5	17	18	20	09	64
Escola 6	06	08	09	09	32
Escola 7	17	14	18	13	62
Total	77	81	100	75	333

6.3 Instrumentos

Todos os participantes leram em voz alta o corpo de itens contendo 323 palavras de baixa frequência de ocorrência, variando em níveis de regularidade (descritas na seção 4), extensão (4 a 8 letras) e estrutura silábica. Esse conjunto de itens foi apresentado por meio de três computadores *lap tops*, nos quais foi instalado o programa *Cognitive Workshop (Cogwork)*, desenvolvido pela equipe do Prof. Philip Seymour na Universidade de Dundee, na Escócia, especialmente para apresentar estímulos e medir o tempo de reação. O *Cogwork* é rodado no *Windows 98* e constitui um programa do *Acess 5.0* que possui uma interface com o *Soundforge 4.5*; este último grava a pronúncia das crianças, assim como as palavras na ordem em que são lidas. A Figura 3

apresenta um modelo de *output* obtido pelo programa *Soundforge* para um dos participantes da pesquisa. A palavra *start* na figura sinaliza o momento em que o item aparece na tela do computador. Onde está escrito *begin* indica o começo da pronúncia da criança, assim como *end* indica o término da resposta. O tempo de reação é registrado pela diferença entre os registros dados em *start* e em *begin* no computador.

Para fins de aplicação, a referida lista de palavras foi dividida em duas listas: a Lista 1 e a Lista 2. As palavras apareciam no centro da tela do computador em letra *times new roman* tamanho 40, escritas na cor branca com fundo preto. As respostas e o tempo de reação das crianças (medido em milésimos de segundo) foram automaticamente gravados pelo próprio computador.

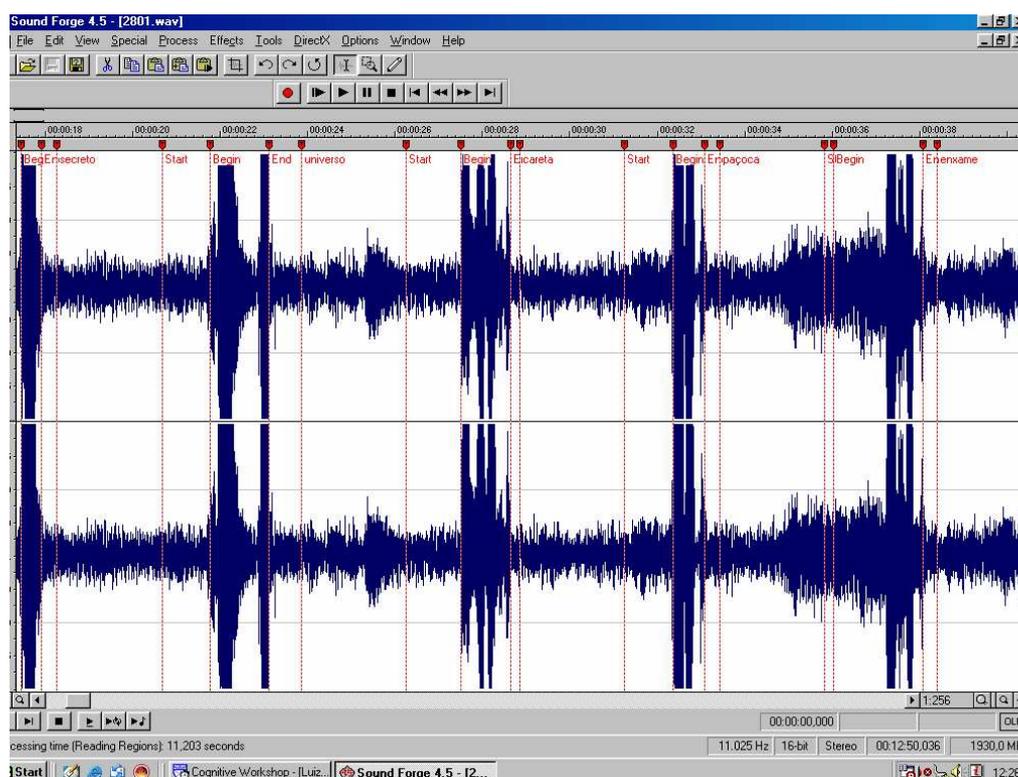


Figura 3. Exemplo de *output* do programa *Soundforge* 4.5, utilizado para a gravação da leitura das crianças.

Para a correção do conjunto de itens considerou-se como correta uma pronúncia que é aceita pela comunidade lingüística, emitida de forma contínua e sem pausas entre as sílabas ou letras da palavra. Desta forma, as respostas emitidas com pausas na pronúncia, as respostas de correção espontânea e as respostas que geraram não-palavras (ou mesmo palavras reais diferentes dos estímulos conferidos), foram consideradas incorretas. Entretanto, as pronúncias coloquiais da palavra, as quais representam variações lingüísticas de um subgrupo de falantes, foram consideradas corretas (por exemplo, ler *catorze* como “quatorze”), desde que não houvesse pausas durante e execução da resposta. De maneira semelhante, considerou-se acerto a emissão de uma pronúncia formal, geralmente aprendida na escola, mas que não corresponde à maneira usual de se falar no nosso Estado (por exemplo, ler *medo* como /¹medo/ e não como /¹medu/).²⁰

Além da leitura, as crianças foram avaliadas quanto à inteligência verbal (subtestes verbais do testes WISC-III) e não-verbal (escala colorida do RAVEN) e quanto ao desempenho escolar pelo subteste de leitura do Teste de Desempenho Escolar (TDE). Além disso, as crianças responderam oralmente um questionário de auto-avaliação (Apêndice E) no qual fizeram um julgamento a respeito de suas próprias habilidades nas áreas de leitura e matemática. Às professoras foi solicitado que respondessem à Escala de Avaliação de Competência em Leitura pelo Professor desenvolvida por Pinheiro e Costa (2005) e em fase de validação (Apêndice F).²¹

²⁰ Não apresentaremos aqui uma análise qualitativa dos erros. No entanto, para detalhes sobre como essa análise deve ser conduzida, ver Pinheiro, Lúcio e Cunha (2007).

²¹ A inclusão dos instrumentos de avaliação da inteligência, do desempenho escolar e da escala competência da criança na leitura segundo julgamento da professora foi necessária em função do processo de validação das duas listas que compõe o corpo de itens. A descrição dos resultados destas avaliações encontra-se em Pinheiro (2007a), mas não será apresentada aqui por fugir ao escopo do presente projeto de dissertação.

6.4 Procedimento de coleta de dados

A coleta de dados foi feita no período de agosto a dezembro de 2006 e de fevereiro a abril de 2007. As provas foram aplicadas por uma equipe de estagiários incluindo dois bolsistas PROBIC/FAPEMIG, 14 alunos de graduação voluntários e a própria mestranda. A equipe de bolsistas e voluntários que realizou a coleta de dados foi treinada pela mestranda em, no mínimo, três sessões. Uma equipe de alunos ficou encarregada da aplicação dos testes de inteligência e desempenho escolar, enquanto outra equipe dedicou-se exclusivamente à aplicação da prova computadorizada de leitura. Findada a coleta de dados, parte da equipe permaneceu no projeto e recebeu treinamento para a correção dos testes.

Cada criança foi avaliada individualmente em uma sala especialmente oferecida para a pesquisa pela escola em três sessões diferentes. As crianças foram submetidas à prova computadorizada de leitura em duas sessões e foram avaliadas quanto à inteligência (WISC-III e RAVEN) e desempenho escolar (TDE) e responderam à auto-avaliação em uma terceira seção. As sessões ocorreram em dias diferentes e a ordem de aplicação foi aleatória em todos os casos.

Na aplicação do banco de itens, inicialmente a criança era exposta a um treinamento, em que uma lista de 10 palavras de alta frequência era apresentada. Após esse treinamento, introduzia-se a primeira lista experimental (Lista 1 ou Lista 2), escolhida aleatoriamente. Na seção seguinte, era apresentada a outra lista. Os itens das duas listas eram randomizados pelo próprio computador e durante a sua aplicação o programa introduzia duas pausas. Para todos os participantes, foi dada a seguinte instrução: “Agora nós vamos fazer uma atividade de leitura. Olhe bem para o centro da tela do computador. Nesta tela irão aparecer várias palavras, uma de cada vez, que você

deve ler o mais rápido que puder e procurando não errar. Se você não souber ler, pode dizer que não sabe, mas o importante é que você tente ler. Podemos começar?”.

Para os testes de inteligência e o de desempenho escolar, foram seguidas as instruções apresentadas em seus respectivos manuais (para as instruções do WISC-III, ver Wechsler, 1991; para as instruções do Raven, ver Raven, Raven, & Court, 1993; para as instruções do subteste de leitura do TDE, ver Stein, 1994).

6.5 Procedimento de análise estatística

Para cada análise estatística efetuada, a primeira seção (denominada “procedimento”) descreve os procedimentos estatísticos utilizados dentro de cada objetivo proposto. Em todas as análises, as variáveis medidas (ou variáveis dependentes) foram a precisão e o tempo de reação dos participantes, com exceção da primeira e da última seção de resultados, em que foi medida apenas a precisão. Para todas as comparações, resultados que obtiveram um p-valor $<0,05$ foram considerados significativos. Para todas as análises, foi utilizado o pacote estatístico SPSS versão 16.0 *for windows*, com exceção da análise de regressão, para a qual foi utilizada o programa STATA versão 8.0.

7 RESULTADOS

7.1 Análise de itens

7.1.1 Procedimento

Três medidas serão aqui relatadas para esta análise, a saber, a dificuldade (p), o poder de discriminação (D) e a consistência interna dos itens ($I-T$). A dificuldade constitui a proporção de participantes que acertam o item. Para a composição dos grupos superior (S) e inferior (I) na análise da discriminação dos itens, foram escolhidos os participantes que obtiveram escores que ficaram posicionados aos 27% das extremidades superior e inferior, respectivamente. Por fim, o cálculo da consistência interna foi feito por meio da correlação item-total. Para as duas últimas medidas, os valores acima de 0,35 foram considerados adequados. Desta forma, os itens-candidatos foram escolhidos com base na sua adequação a essas três medidas.

7.1.2 Resultados

O Apêndice G²² lista todos os índices de dificuldade, discriminação e de correlação item total para as palavras das Listas 1 e 2 em cada série e no total²³. (Devido

²² Com o objetivo de melhorar a visualização do Apêndice G, foi utilizada a notação de ponto ao invés de vírgula nos resultados.

²³ Como se pode observar, algumas palavras mudaram de classificação depois que os dados foram coletados. Manteve-se nesta Tabela a classificação feita na época da coleta de dados (a qual foi baseada em Pinheiro (2004, 2007b)) pelo fato de as análises de consistência interna e de estabilidade temporal do instrumento depender dessa classificação. No entanto, para a análise de regressão hierárquica, será utilizada a classificação modificada. Assim, todas as palavras contendo <ce>, <ci>, <ge> e <gi> foram classificadas como “regra” para a leitura e os grafemas <c> e <g> nos demais contextos foram classificados como regulares. Além disso, as palavras contendo vogais na sílaba tônica, seguidas de <m>, <n> ou de <nh> foram classificadas como “regra”, pois essas vogais são sempre nasaladas. (Observe que algumas dessas palavras – como *trena* e *nono* – estavam classificadas como irregulares, mas este não é o

a limitações de espaço, as palavras de cada lista foram divididas em termos do nível de regularidade para a leitura e para a escrita e do número de letras, apesar de outros critérios terem sido utilizados para o pareamento, conforme se verá na próxima seção). Os valores de alfa de Cronbach atingiram níveis altamente satisfatórios, variando de 0,97 na 3ª série a 0,99 na 1ª série e no geral (todas as séries juntas). Os índices de consistência interna obtidos atestam para a homogeneidade da tarefa e confirmam a unidimensionalidade do construto avaliado, isto é, a habilidade de reconhecimento de palavras isoladas. Os dados apresentados no Apêndice G são resumidos na Tabela 8, que mostra a média dos índices p , D e $I-T$ para cada série e no total²⁴.

caso de irregularidade tipo 1B, uma vez que as pronúncias do <e> e do <o>, em ambos os casos, são previsíveis pela regra de nasalação). Todas as palavras contendo o grafema <x> estavam classificadas como irregulares, mas consideramos como irregulares apenas o <x> no contexto intervocálico e como “regra” nos demais contextos. A porcentagem de palavras assim alteradas foi de 12%. No entanto, é importante ressaltar que, a despeito das modificações introduzidas, algumas palavras mativeram a sua classificação original. É o caso, por exemplo, da palavra *aceso*, classificada como IRRG por conter o grafema <e> em sílaba tônica em posição paroxítona e o <s> intervocálico. Assim, ter considerado a pronúncia do <c> nessa palavra como “regra” não alterou a classificação final da mesma. Por estas razões, a classificação final das palavras utilizada no presente trabalho se difere da introduzida por Pinheiro (2004, 2007b) e pode ser vista no Apêndice H (ver p. 76, na seção 4 acima).

²⁴ Mesmo para a correlação item-total, optou-se por utilizar os valores de média e desvio-padrão ao invés da transformação desses valores em escores z pelo fato de a distribuição ter sido aproximadamente normal para todas as séries e no total (o valor modular mínimo e máximo para a assimetria e para a curtose foi de 0,40 e 0,65 e de 0,15 e 0,53, respectivamente).

Tabela 8

Valores médios da dificuldade, discriminação e correlação item-total nas Listas 1 e 2 e no total para cada série separadamente e toda a amostra.

Amostra	Lista 1			Lista 2			Total		
	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T
1ª série	0,73	0,45	0,49	0,72	0,41	0,45	0,72	0,43	0,47
2ª série	0,81	0,32	0,45	0,78	0,32	0,42	0,80	0,32	0,43
3ª série	0,87	0,22	0,31	0,85	0,23	0,29	0,86	0,23	0,30
4ª série	0,88	0,27	0,48	0,86	0,28	0,49	0,87	0,27	0,48
Total	0,82	0,34	0,47	0,81	0,33	0,44	0,81	0,34	0,45

Nota. *p* = proporção de acertos; D = coeficiente de discriminação; I-T = correlação item-total.

A Tabela 9 mostra a proporção de palavras das Listas 1 e 2 e do total que se distribuem ao longo dos diferentes níveis de dificuldade adotados, bem como dos níveis de discriminação e da consistência interna. Por ela, podemos observar que a maioria das palavras das listas é fácil ($0,60 \leq p < 0,80$), ou muito fácil ($p \geq 0,80$), e que apenas uma proporção irrisória de itens apresenta um nível extremo de dificuldade ($p < 0,20$). Apesar disso, a maioria dos itens produziu níveis adequados de discriminação ($D \geq 0,30$), o que indica o conjunto de itens como um todo foi capaz de segregar os indivíduos com competência na leitura de palavras isoladas daqueles indivíduos cuja competência é reduzida. Por fim, observou-se que a maioria dos itens também produziu níveis adequados de correlação item-total ($I-T \geq 0,30$), o que mais uma vez atesta para a homogeneidade da tarefa.

Tabela 9

Distribuição da proporção de palavras (%) em cada nível de dificuldade, discriminação e correlação item-total.

Análise estatística	Lista 1	Lista 2	Total
Dificuldade			
<0,20	0,62	0	0,31
≥ 0,20 e < 0,40	1,23	1,24	1,24
≥ 0,40 e < 0,60	3,09	8,70	5,88
≥ 0,60 e < 0,80	24,07	22,98	23,53
≥ 0,80	70,99	67,08	69,04
Discriminação			
<0,30	38,27	45,34	41,80
≥ 0,30	61,73	55,66	58,20
Item-total			
<0,30	3,09	8,07	5,57
≥ 0,30	96,91	91,93	94,93

Devido à pequena quantidade de itens que apresentam índices de dificuldade de alta ($p < 0,20$) a média ($0,40 \leq p < 0,60$), todos os itens de ambas as Listas 1 e 2 com estes níveis de dificuldade são reportados na Tabela 10. O índice de discriminação desses itens variou de 0,13 a 0,61 na Lista 1 e de 0,21 a 0,70 na Lista 2. Para as palavras com nível de dificuldade baixa ($0,60 \leq p < 0,80$), foram selecionados os itens com nível de discriminação maior do que 0,50. Com isso, foram selecionados 17 itens, tanto da Lista 1 quanto da Lista 2. Mantiveram-se, assim, proporções em torno de 45% dos itens

originais em ambas as listas para o nível de dificuldade baixo, todos eles apresentando valores de correlação item-total maiores do que 0,40, o que é bastante satisfatório.

Tabela 10

Itens selecionados por meio da análise, separados em função do nível de dificuldade e da lista de aplicação.

Dificuldade	R-1 (Lista 1)	R-2 (Lista 2)
<0,20	ampola	-
$\geq 0,20$ e $< 0,40$	aceso, exceto	tola, boxe
$\geq 0,40$ e $< 0,60$	serelepe, sarjeta, foco, cera, humilde	palhoça, forca, catorze, roxa, saxofone, zebu, despesa, monarca, novelo, guizo, bodoque, rasa, maroto, careta
$\geq 0,60$ e $< 0,80$	excesso, exposto, gengiva, freguesa, bondosa, caroço, atleta, colo, manhoso, bezerro, caçula, cigana, celeste, marreco, nublado, vasilha, cadelã.	externo, colete, rabanete, cometa, cevada, cesto, enxoval, fofo, delicada, serrote, deserto, doutora, luminosa, reflexo, mariposa, defesa, bexiga

Tabela 10 (*Continuação*).

Dificuldade	R-1 (Lista 1)	R-2 (Lista 2)
$\geq 0,80$	cereja, cartela, espiral, veste, cantora, nervosa, caxumba, covarde, esquerdo, farda, retalho, jeitoso, cabana, recurso, teimoso, repleto, universo, famoso, moderna, diretora, aflito, faxina, morena, favela, preguiça, adulta, brasa, dengoso, capela, feminina, frota, miolo, cartola, trevo, carroça, negra, pepino, peteca, honesto, tombo, bandeja, tapete, gemada, coberta, uniforme, espanto, horta, filtro. combate, fiapo terço, sapeca, selva	

Nota. R-1 (palavras obtidas pela análise de itens da Lista 1); R-2 (palavras obtidas pela análise de itens da Lista 2).

Dos itens com nível de dificuldade muito baixo ($p \geq 0,80$), foram selecionados aqueles que apresentaram dificuldade maior ou igual a 0,35 e correlação item-total maior ou igual a 0,40. Com isso, foram selecionados 30 itens da Lista 1 (mantidos 26% dos itens) e 24 da Lista 2 (22% dos itens). No total, foram selecionados 112 itens candidatos a compor a prova computadorizada de leitura. Destes, 69% são irregulares (IR e IRRG). A Tabela 10 resume os resultados. Em R-1²⁵ desta Tabela estão as palavras obtidas pela análise dos itens na Lista 1 e em R-2 as palavras oriundas da Lista 2.

A Tabela 11 mostra a configuração final da distribuição dos itens selecionados em termos das quatro classificações de regularidade para a leitura e do nível de dificuldade, para as Listas 1 e 2 e para o conjunto total de itens. Por esta tabela, pode-se

²⁵ R de “forma reduzida”.

observar que em ambas as Listas 1 e 2 as palavras de dificuldade muito alta ($0,20 \leq p < 0,40$) e muito alta ($p < 0,20$) se concentram entre os itens irregulares para a leitura (IR e IRRG) e que esses itens também são predominantes entre os níveis de dificuldade média. As palavras regulares (R e RG) só aparecem entre os itens com este último índice de dificuldade, mas mesmo assim nunca são prevalentes em nenhum dos níveis de dificuldade aqui especificados.

Tabela 11

Porcentagens de palavras em cada classificação de regularidade para a leitura presentes para os níveis de dificuldade especificados.

Dificuldade/Lista	R	RG	IR	IRRG
R-1 (Lista 1)				
<0,20	-	-	100	-
$\geq 0,20$ e $< 0,40$	-	-	50	50
$\geq 0,40$ e $< 0,60$	20	-	40	40
$\geq 0,60$ e $< 0,80$	12	18	41	29
$\geq 0,80$	37	7	43	13
R-2 (Lista 2)				
<0,20	-	-	-	-
$\geq 0,20$ e $< 0,40$	-	-	100	-
$\geq 0,40$ e $< 0,60$	14	14	50	21
$\geq 0,60$ e $< 0,80$	6	12	35	47
$\geq 0,80$	33	4	46	17

Tabela 11 (*Continuação*).

Dificuldade/Lista	R	RG	IR	IRRG
Total				
<0,20	-	-	100	-
≥ 0,20 e < 0,40	-	-	0,75	0,25
≥ 0,40 e < 0,60	16	11	47	26
≥ 0,60 e < 0,80	9	15	44	32
≥ 0,80	35	6	44	15

Nota. R = regular; RG = regra; IR = irregular; IRRG = irregular-regra. (regularidade para a leitura).

7.2 Efeito de treino e fadiga

7.2.1 Procedimento

Conforme dito na seção “método”, o banco de itens foi dividido em duas partes aplicadas em dias diferentes: a Lista 1, que contém 162 palavras, e a Lista 2, com 161 palavras. Em cada uma dessas partes, havia duas pausas automaticamente programadas pelo computador, pausas estas que se efetuavam a cada 55 palavras lidas. Portanto, cada aplicação ficou subdividida em três partes, cada uma delas contendo mais ou menos a mesma quantidade de palavras, com exceção de que, na última parte da Lista 1, foram apresentadas 52 palavras, enquanto que na última parte da Lista 2 foram apresentadas 51 palavras. O aplicador era instruído a questionar à criança se havia necessidade de descanso nessas pausas programadas (o que geralmente não ocorria) ou se poderia dar seguimento à seção.

O objetivo desta análise foi o de verificar se existem diferenças de desempenho entre cada um dos terços de prova, que demonstraria um efeito de fadiga (caso o desempenho no primeiro terço da prova for superior ao dos demais) ou de treino (caso contrário). Como as palavras são apresentadas de forma aleatória, espera-se que os terços de prova sejam mais ou menos homogêneos em termos de dificuldade (isso é importante, pois poderiam surgir diferenças entre as partes não por efeito da fadiga ou do treino, mas pela presença de palavras mais ou menos difíceis em cada uma dessas partes).

7.2.2 Resultados

Análises univariadas de variância foram conduzidas para verificar a existência de diferenças significativas entre os três terços de cada uma das listas. As comparações foram feitas no geral (com todas as séries juntas) e para cada série separadamente. Análises de *post hoc* foram feitas pelo teste Bonferroni. A Tabela 12 mostra os valores de média e DP para a proporção de acertos (%) e para o TR (mseg.) em cada condição.

Para a precisão, as análises indicaram um efeito geral de treino na Lista 1 ($F=4,649$, $p<0,01$; $\beta = 0,79$) e o *post hoc* de Bonferroni indicou que essas diferenças se devem às comparações entre o primeiro terço da lista e o segundo ($p<0,012$). Além disso, na terceira série o efeito também esteve presente ($F = 5,811$, $p<0,03$; $\beta = 0,87$) e as diferenças ocorreram entre o primeiro terço da lista e os demais ($p<0,034$ e $p<0,004$ para as comparações com o segundo e o terceiro terço, respectivamente). As outras comparações não produziram diferenças significativas.

Tabela 12

Valores de média (M) e desvio padrão (D.P.) para a proporção de acertos (%) e para o tempo de reação (mseg) nas três partes de cada uma das listas de palavras e para todo o corpo de itens aplicados às crianças da amostra para cada série separadamente e no total.

Precisão	1ª série		2ª série		3ª série		4ª série		Total	
	M	D.P	M	D.P	M	D.P	M	D.P	M	D.P
L1/1º terço	72	45	80	40	85	35	87	33	82	39
L1/2º terço	74	44	80	40	87	34	88	32	83	38
L1/3º terço	72	45	81	39	88	33	88	33	83	38
L2/1º terço	70	46	78	42	85	36	87	34	80	40
L2/2º terço	71	45	78	42	85	36	87	34	80	40
L2/3º terço	72	45	79	41	85	36	86	44	81	39
Total*	72	45	79	45	86	35	87	33	81	40
TR	M	D.P	M	D.P	M	D.P	M	D.P	M	D.P
L1/1º terço	1382	970	1124	745	931	429	833	386	1057	691
L1/2º terço	1488	980	1159	727	949	440	834	353	1087	696
L1/3º terço	1440	1004	1176	781	968	468	855	413	1098	725
L2/1º terço	1338	912	1147	811	933	448	822	385	1050	693
L2/2º terço	1394	981	1165	806	972	471	863	482	1087	729
L2/3º terço	1393	952	1186	823	994	490	863	459	1100	727
Total*	1399	967	1156	781	957	458	885	414	1079	710

*Média de TR para todo o corpo de itens (Lista 1 e Lista 2 juntas).

A análise do tamanho do efeito (ou a magnitude da diferença entre as condições) mostrou que o valor de eta (η^2) é praticamente nulo tanto para as comparações gerais na Lista 1 ($\eta^2 = 0,002$) quanto na 3ª série ($\eta^2 = 0,001$). Os valores para a estatística F de tamanhos de efeito pequenos, médios e grandes são de 0,1, 0,25 e 0,4, respectivamente (Cohen, 1992; Fox e Mathers, 1997). Isso implica que a diferença estatisticamente significativa aqui obtida pode não ter, de fato, significância prática, havendo a possibilidade ter ocorrido um erro do tipo II (ou seja, a rejeição da hipótese nula quando ela deveria ser aceita).

Para o tratamento do tempo de reação, foi considerado um ponto de corte (*cutoff*) de todos os valores abaixo de 400 msec²⁶ (Ratcliff, 1993). Todas as comparações foram significativas no geral e em todas as séries, para ambas as listas, na direção do efeito de fadiga. A exceção ficou por conta da Lista 2 na segunda série, cujo efeito não foi significativo. No entanto, para todas essas comparações o tamanho do efeito atingiu valores próximos a zero. Assim, a título do que ocorreu para os casos específicos relatados na precisão, apenas as comparações de *post hoc* apontam para o efeito de fadiga, podendo-se, possivelmente, descartar a hipótese de um efeito geral. A Tabela 13 resume os resultados e apresenta as comparações que foram significativas no teste de Bonferroni.

²⁶ Ademais este critério, o TR de algumas palavras foi anulado nos casos de registro acidental em função de um barulho ambiental, ou mesmo um ruído emitido pelo próprio participante que não correspondia à sua resposta. Na Figura 3, por exemplo, o TR da palavra *paçoca* foi anulado para o participante em questão. Outro ponto a ser mencionado é que o TR de uma das escolas (Escola 4) foi anulado para todas as crianças, pois houve falha no equipamento e os valores não puderam ser salvos.

Tabela 13

Níveis de significância para as comparações entre os três terços das Listas 1 e 2, valores do tamanho do efeito (η^2), do poder observado (β) e de p para as comparações significativas encontradas no teste de Bonferroni, para cada série e no geral.

Comparação	Lista	Significância	η^2	β	Bonferroni
Geral	L1	F = 12,491	0,001	1,00	1° vs 2°, p <0,001
		p<0,001			1° vs 3°, p<0,001
	L2	F = 17,181	0,001	1,00	1° vs 2°, p <0,001
		p<0,001			1° vs 3°, p<0,001
1ª série	L1	F = 3,991	0,001	0,72	1° vs 2°, p <0,02
		p<0,01			
	L2	F = 3,244	0,001	0,62	ns
		p<0,04			
2ª série	L1	F = 4,481	0,001	0,77	1° vs 3°, p<0,01
		p<0,01			
	L2	Ns	0,000	0,47	ns
3ª série	L1	F = 6,632	0,001	0,91	1° vs 3°, p<0,001
		p<0,001			
	L2	F = 16,233	0,003	1,00	1° vs 2°, p <0,001
		p< 0,000			1° vs 3°, p<0,001
4ª série	L1	F = 3,150	0,001	0,64	ns
		p<0,04			
	L2	F = 9,154	0,002	0,98	1° vs 2°, p <0,001
		p<0,000			1° vs 3°, p<0,001

Nota. L1 = Lista 1; L2 = Lista 2.

7.3 Efeitos na leitura

7.3.1 Procedimento

Para a análise de regressão hierárquica, optou-se por utilizar todo o banco de itens ao invés de apenas os itens que foram selecionados por uma razão: conforme se observou acima (seção 7.1.2, página 86; ver Tabela 11, p. 92), na amostra de itens-candidatos selecionada, a maioria dos itens regulares para a leitura se concentra entre os níveis de dificuldade média a muito baixa, enquanto que os itens irregulares, além de prevalentes em todos os níveis de dificuldade, são exclusivos entre as faixas de maior dificuldade. Apesar de este resultado já apontar para o efeito de regularidade na leitura, utilizar um modelo confirmatório para este efeito com uma amostra de itens com as características mencionadas poderia produzir um modelo explicativo circular, uma vez que os itens de dificuldade alta foram todos irregulares.

Tanto para os modelos de precisão quanto de tempo de reação, foram considerados os seguintes preditores: (a) características do estímulo: regularidade da palavra para a leitura, para a escrita e número de letras da palavra; (b) características do participante: sexo, série, idade de aquisição da leitura (heteroavaliação feita pelos pais), pré-escola, QI verbal (QIV), compreensão verbal (QICV) e resistência à distração (QIRD) (todos avaliados pelo WISC-III), percentil no RAVEN (normas da escala geral); dificuldade com a língua portuguesa (medida pela auto-avaliação da criança (Apense E) e por uma heteroavaliação feita pelos pais (Apense D)); (c) características sociais: nível socioeconômico (dado pela renda familiar), escolaridade do pai e escolaridade da mãe. Para ambos os modelos, e com o objetivo de investigar a hipótese anterior de que os efeitos de treino e fadiga encontrados não apresentam de fato

significância prática, a variável “ordem de apresentação do estímulo” também foi incluída no primeiro modelo. Além disso, cabe ressaltar que, para a característica de regularidade da palavra, serão feitos dois modelos distintos: um com dois níveis de regularidade para a leitura e para a escrita (R e IR) e outro utilizando quatro níveis (R, RG, IR e IRRG). Conforme afirmado acima (seção 4), este procedimento tem por objetivo investigar a validade da hipótese de que o número de categorias utilizado pode interferir na expressão do efeito de regularidade na leitura.

Das características da palavra, espera-se que a regularidade seja o fator que mais contribua para as diferenças de desempenho das crianças, tanto em termos de precisão quanto de tempo de reação. No entanto, no primeiro caso, a regularidade da palavra para a leitura deve afetar a precisão na amostra, enquanto que a regularidade para a escrita deve afetar os tempos de reação. Assim, o modelo final para a precisão deve conter apenas a regularidade para a leitura, enquanto que o modelo para o tempo de reação deve conter apenas a regularidade para a escrita. A direção do efeito esperada é a de melhores resultados nas palavras regulares em relação às irregulares. Em relação ao número de letras das palavras, é esperado um efeito de extensão favorável às palavras mais curtas (Godoy, 2005; Pinheiro, 1995; Salles, 2005; Sucena & Castro, 2005).

Os demais fatores devem afetar ambos os modelos na mesma direção. Desta forma, das características do participante, é esperado um efeito positivo da série (Godoy, 2005; Pinheiro, 1995; Salles, 2005; Salles & Parente, 2005) e da resistência à distração (por ter relação com a memória de trabalho (Sawson & Ashbaker, 2000; Sawson, Howard, & Saéz, 2006)) e um efeito contrário (isto é, uma melhoria da proficiência com uma redução dos valores das variáveis) da idade de aquisição da leitura (Zevin & Seidenberg, 2002). A avaliação da inteligência feita pelo Raven não deve se relacionar com a leitura, pois, por se tratar de uma medida de inteligência fluida,

não deverá se correlacionar com uma habilidade verbal (validade divergente); assim, esta variável não deverá participar do modelo final. O QIV e a compreensão verbal entram na primeira equação como forma de averiguar sua importância relativa, podendo não haver relação entre essas medidas da inteligência e o desempenho na leitura (Maia & Fonseca, 2002). Todas as características sociais devem apresentar relações positivas com a precisão e negativas com o tempo de reação. Por fim, deve haver uma maior proficiência na leitura das meninas em relação aos meninos, uma vez que a literatura tem relatado que as meninas se sobressaem mais em tarefas de leitura (p. ex., Chiu & McBride-Chang, 2006; Hedges & Nowell, 1995; Rosén, 2001) e de desempenho acadêmico (p. ex., Carvalho, 2004a e b) do que os meninos.

Para as variáveis não métricas, foram escolhidos indicadores, os quais são: as palavras regulares, o sexo masculino, a primeira série, ausência de dificuldade com a língua portuguesa, não ter realizado pré-escola e o menor nível de escolaridade dos pais (analfabeto/primário incompleto).

Foi utilizado um modelo GEE logístico para a variável precisão e um modelo linear para o tempo de reação, ambos com uma estrutura homogênea. Para a seleção de variáveis dos modelos linear e logístico, foi utilizado o procedimento de *stepwise backward*. Este método ajusta um modelo completo e, em seguida, retira a variável não significativa de menor peso no modelo. Deste modo, o procedimento é repetido até que sobrem apenas variáveis estatisticamente significativas.

7.3.2 Resultados

7.3.2.1 Modelo de tempo de reação

Na Tabela 12 são mostradas as médias e os desvios padrão para o tempo de reação das crianças em função da série e no total. A figura 4 mostra a distribuição dos tempos de reação apresentados pela amostra (como se pode observar, o padrão da distribuição se aproxima muito de uma curva de qui-quadrado teórica).

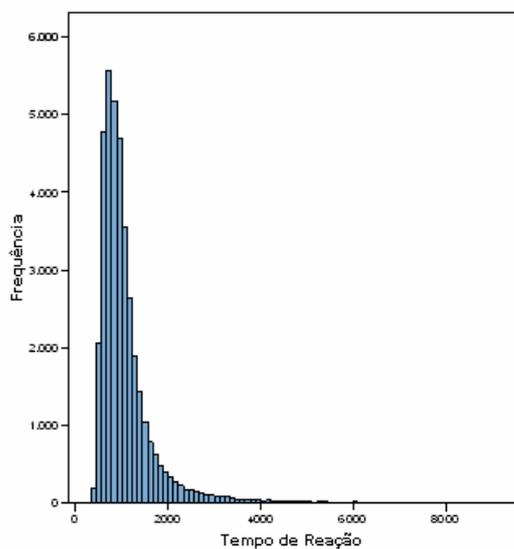


Figura 4. Distribuição do tempo de reação da amostra de participantes

A Tabela 14 apresenta a equação de estimação do tempo de reação, com todas as co-variáveis que se mostraram significativas.

Tabela 14
Equação de estimação para o tempo de reação dos alunos

Modelo para o tempo de reação	Coef.	DP	Z	P-valor	IC(95%)	
Ordem	0.50	0.10	4.74	<0.0001	0.29	0.70
Idade leitura (<i>anos</i>)	114.02	48.36	2.36	0.0180	19.24	208.79
Série (<i>1ª série</i>)				<0.0001(conj.)		
2ª série	-342.77	149.70	-2.29	0.0220	-636.17	-49.37
3ª série	-474.32	126.88	-3.74	<0.0001	-722.99	-225.65
4ª série	-685.09	140.64	-4.87	<0.0001	-960.74	-409.44
QICV (<i>escore</i>)	-7.60	2.24	-3.49	<0.0001	-12.16	-3.41
Letras (<i>total</i>)	62.64	6.97	10.49	<0.0001	50.94	74.34
Classificação escrita (<i>Regular</i>)				<0.0001(conj.)		
regra	6.27	3.96	1.58	0.1140	-1.50	14.04
Irregular	43.76	6.89	7.43	<0.0001	32.20	55.30
Irregular regra	52.49	6.97	7.53	<0.0001	38.83	66.15

Confirmando as expectativas levantadas, foram consideradas preditoras para o tempo de reação (TR) as covariáveis idade de aquisição da leitura, série, número de letras da palavra e a classificação de regularidade da palavra para a escrita. Todas estas covariáveis entraram no modelo de regressão na direção esperada.

No que diz respeito às características do participante, observou-se que cada ano de atraso na aquisição da leitura corresponde a um aumento de cerca de 114mseg no TR da criança. Isso indica que, quanto mais cedo a criança aprende a ler, maior a sua rapidez em termos de tempo de processamento. Além disso, foi observado um forte efeito de escolaridade: passando da primeira para a segunda série, há uma redução de quase 340mseg no TR. Uma redução de 474mseg no TR é observada da primeira para a terceira série, enquanto que nas comparações com a quarta série a redução chegou aos 685mseg. Todas as comparações da primeira série com as demais foram significativas.

Contrariando as expectativas, o fator de resistência à distração não foi relevante para explicar os resultados. Dos fatores de inteligência, apenas a compreensão verbal mostrou-se relevante para o modelo, havendo uma redução de 8mseg para cada unidade de aumento no QICV. Nenhuma das variáveis sociais contribuiu para a equação, inclusive o nível socioeconômico da criança. Não houve efeito de gênero.

Para as características da palavra, houve um forte efeito de extensão, sendo que para cada unidade de aumento da palavra (em número de letras) houve um acréscimo de 63mseg no TR. Corroborando nossa hipótese, apenas a regularidade para a escrita da palavra foi relevante para as variações no TR. Além disso, o modelo com quatro categorias de regularidade foi o que mais bem se adequou aos dados, tornando evidente a influência da quantidade de categorias adotadas no trabalho na produção de efeitos na leitura. As palavras regulares foram lidas com mais rapidez do que as palavras irregulares (IR e IRRG), mas as comparações com as palavras da categoria regra não foram significativas. O aumento do TR das demais classificações comparação com as palavras regulares variou de seis a 52mseg.

Ao contrário do que foi esperado, houve uma forte influência da ordem na direção de um efeito de fadiga no TR. A cada palavra lida, houve um aumento médio de 0,50mseg no TR.

7.3.2.2 Modelo de precisão

A Tabela 12 apresenta a proporção geral de acertos no banco de itens para cada série e no total. Para o modelo de precisão, mais variáveis foram consideradas preditoras do que o modelo de TR. Em comum, a idade de aquisição da leitura, a série, o QICV, o número de letras da palavra e a regularidade para a escrita da palavra (a entrada desta última variável contrariou a hipótese levantada acima (página 99) de que apenas a regularidade para a leitura seria importante para a emergência desse efeito)²⁷. Adicionalmente, mostraram-se adequadas para o modelo o fato de a criança ter feito pré-escola, a escolaridade da mãe, a auto-avaliação da criança sobre suas próprias

²⁷ Também de maneira análoga ao modelo de TR, não houve efeito de gênero na precisão da leitura.

dificuldades com a língua portuguesa e a classificação da palavra para a leitura. A ordem não foi relevante para a explicação dos resultados em termos de precisão. A Tabela 15 apresenta a equação de estimação da precisão, com todas as co-variáveis que se mostraram significativas.

Tabela 15
Equação de estimação da precisão dos alunos

Modelo para a variável precisão	RC	DP	Z	P-valor	IC(95%)	
Pré-escola (<i>0 não; 1 sim</i>)	1.77	0.39	2.61	0.0090	1.15	2.73
Idade leitura (<i>anos</i>)	0.76	0.05	-3.79	<0.0001	0.66	0.88
Dif. port. auto-avaliação (<i>0 não; 1 sim</i>)	0.76	0.10	-2.10	0.0360	0.59	0.98
Série (<i>1ª série</i>)				<0.0001 (conj.)		
2ª série	1.53	0.26	2.53	0.0110	1.10	2.13
3ª série	2.08	0.37	4.13	<0.0001	1.47	2.96
4ª série	3.06	0.70	4.90	<0.0001	1.96	4.80
Escolaridade Mãe (<i>Analfabeto/Primário incompleto</i>)				0.0290 (conj.)		
Primário completo/Ginasial incompleto	1.02	0.30	0.07	0.9470	0.57	1.83
Ginasial completo/Colegial incompleto	0.85	0.22	-0.64	0.5220	0.52	1.40
Colegial completo/Superior incompleto	0.99	0.24	-0.04	0.9710	0.62	1.59
Superior completo	0.50	0.14	-2.40	0.0170	0.28	0.88
QICV (<i>escore</i>)	1.02	0.00	5.40	<0.0001	1.01	1.03
Letras (<i>total</i>)	0.91	0.01	-6.67	<0.0001	0.89	0.94
Classificação leitura (<i>Regular</i>)				<0.0001 (conj.)		
regra	0.86	0.03	-4.52	<0.0001	0.80	0.92
Irregular	0.58	0.02	-14.33	<0.0001	0.54	0.62
Irregular regra	0.51	0.02	-14.06	<0.0001	0.47	0.56
Classificação escrita (<i>Regular</i>)				<0.0001 (conj.)		
regra	0.80	0.02	-7.89	<0.0001	0.75	0.84
Irregular	0.78	0.03	-7.69	<0.0001	0.73	0.83
Irregular regra	0.64	0.02	-12.14	<0.0001	0.59	0.68

O modelo GEE logístico apresenta a estimativa do efeito em termos de razão de chance (RC na Tabela 15). Primeiramente, serão descritas as influências das características do sujeito na precisão. A cada ano de atraso na aprendizagem da leitura, as chances de acerto são 24% menores. Isso significa que aumentando a idade de aprendizagem da leitura, ocorre uma diminuição na probabilidade de acerto. Ter feito pré-escola aumentam em 77% as chances de acerto. O reconhecimento pela criança de dificuldades com a aprendizagem do português reduz em 35% suas chances de acerto.

Houve um forte efeito de série na precisão: passando da primeira para a segunda série, as chances de acerto aumentam em 53%. Da primeira para a terceira série, o aumento é de 100%, enquanto que em relação à quarta este aumento da chance de acerto é de 300%. Por fim, a cada unidade de aumento do QICV há um aumento de 2% na chance de acerto.

Das características sociais, conforme dito acima, apenas a escolaridade materna mostrou-se importante e a direção de sua influência ocorreu de uma maneira inusitada. Com exceção das relações entre os dois níveis mais inferiores de escolaridade, a relação entre escolaridade da mãe e proficiência na leitura é inversa, ou seja, quanto maior a escolaridade da mãe, pior o desempenho da criança. A diferença do conjunto foi significativa ao nível de $p < 0,03$. Além disso, a única diferença significativa com as comparações com a variável indicadora foi com o nível mais superior de escolaridade da mãe, sendo que a chance de acerto neste nível é 50% menor em relação à variável indicadora.

A cada aumento no número de letras da palavra, há uma redução de 9% na chance de acerto e isso demonstra um forte efeito de extensão. Em relação à classificação da palavra para a leitura e para a escrita, passando da categoria regular para todas as demais categorias, há uma redução significativa da razão de chance de acerto. Diferentemente do modelo de TR, as comparações entre as palavras regulares e as regras também foram significativas. No geral, a maior redução da razão de chance ocorreu nas comparações com as palavras irregulares-regra tanto para a regularidade na leitura quanto para a escrita, sendo a chance de acerto nesses itens respectivamente 49% e 36% menor em relação à variável indicadora. A título do que ocorreu com o TR, o modelo com quatro categorias de regularidade foi o que melhor explicou os resultados da amostra.

7.4 Simulação de seleção de amostra final de itens a partir da amostra de itens-candidatos: procedimentos psicométricos

O objetivo desta seção é descrever a seleção de uma amostra final de itens a partir dos itens-candidatos cuja seleção foi apresentada na seção 7.1. Tal procedimento tem um caráter tanto didático quanto exploratório. Assim, além de uma simulação de seleção de itens, serão conduzidos alguns procedimentos psicométricos para se verificar a fidedignidade (temporal e de concordância entre juízes) desta amostra final de itens selecionada. É importante ressaltar que as amostras de itens apresentadas a título de exemplo nessa seção não devem ser tomadas como formas definitivas da prova computadorizada de reconhecimento de palavras para crianças, tampouco as investigações psicométricas aqui conduzidas como evidências decisivas, o que faz com que o produto gerado não permita, portanto, a extração de inferências a respeito dos processamentos utilizados na leitura pelos participantes, conforme seria esperado se as palavras tivessem sido pareadas em termos de regularidade e extensão (e de frequência, posteriormente).

7.4.1 Análise de itens

7.4.1.1 Procedimentos e resultados

Para a redução dos itens-candidatos a uma amostra menor de itens, seguimos o critério de Pasquali (1996) para a seleção da quantidade de palavras que devem estar presentes quanto ao nível de dificuldade. Este critério considera que a distribuição dos itens em termos de dificuldade deve ser aproximada à da curva normal.

Conforme se pôde observar pela Tabela 9, a maioria das palavras do banco de itens apresentou níveis de dificuldade baixa ou muito baixa. Ademais, a análise dos itens preservou a relação observada da escassez de itens difíceis ou muito difíceis (Tabela 10), o que nos mostra que as opções de seleção de itens são de fato reduzidas para estes últimos níveis de dificuldade. Diante dessa configuração de resultados, optou-se por criar duas listas de palavras contendo 20 itens cada, metade dos quais extraídos da lista de palavras-candidatas gerada a partir da Lista 1 e a outra metade daquela gerada a partir da Lista 2 (chamadas aqui de Listas Reduzidas 1 e 2, respectivamente ou LR-1 e LR-2). As duas listas paralelas então geradas serão doravante referidas como Exemplo-1 e Exemplo-2.

Desta forma, de acordo com o critério de Pasquali (1996), para cada forma paralela, o ideal seria selecionar 2 itens muito difíceis, 4 difíceis, 8 com nível médio de dificuldade, 4 fáceis e 2 muito fáceis. No entanto, conforme vimos, em todo o banco de itens apenas uma palavra foi considerada muito difícil e quatro foram consideradas difíceis, o que nos fez decidir pela criação de itens fixos que estariam presentes em ambas as LR-1 e LR-2, alterando-se apenas os itens com dificuldade média, baixa ou muito baixa. Isso implica que os critérios de discriminação e dificuldade não se aplicam

aqui para os itens mais difíceis. De fato, esses itens não apresentam índices de discriminação satisfatórios (pois tanto as crianças do grupo S quanto as do I têm proporções de acertos baixas nesses itens), além de baixa correlação item-total (pois, os itens são muito fáceis, qualquer dificuldade imposta pelo item faz com que ele se difira dos demais). Entretanto, a seleção desses itens se justifica por duas razões: (1) o valor do alfa de Cronbach do banco de itens praticamente não se altera com a exclusão desses itens considerados difíceis (passa de 0,989 para 0,987); (2) estes constituem os únicos itens capazes de gerar alguma dificuldade para as crianças das séries finais.

Em função de não haver uma proporção suficiente de itens difíceis para a composição das duas Formas de Exemplo, os itens com índices $p < 0,40$ foram tratados como “itens âncora”, selecionados para ambas as formas. O mesmo ocorreu com três itens de dificuldade média na LR-1, a qual apresentou apenas cinco itens com esse nível de dificuldade. Assim, na LR-1, a palavra *cera* foi selecionada para o Exemplo-2, a fim de se equilibrar o número de palavras de tamanho pequeno (até 5 letras) em ambas as formas, permanecendo a palavra *humilde* no Exemplo-1. A Tabela 16 apresenta todos os itens selecionados, e aqueles destacados na cor cinza constituem os itens-âncora para os Exemplos 1 e 2.

Para o nível de dificuldade média, foram selecionados cinco itens na LR-2, pois, conforme vimos, não apareceram itens com nível de dificuldade extrema nessa lista reduzida. O primeiro critério de entrada foi a apresentação de um alto índice de discriminação, tendo sido selecionadas as palavras *bodoque* e *guizo* no Exemplo-1 e *saxofone* e *monarca* no Exemplo-2. Note que as palavras foram pareadas em ambas as formas em termos de nível de discriminação e de regularidade, havendo uma palavra regular e outra irregular em cada forma da lista. Em seguida, considerou-se o nível de dificuldade (uma vez que LR-2 não apresenta palavras muito difíceis, esse é um critério

importante, mesmo entre palavras cujo nível de dificuldade é semelhante). Foram selecionadas as palavras *palhoça* (Exemplo-1) e *forca* (Exemplo-2). Observa-se que a palavra *forca* apresenta coeficientes de discriminação e correlação item-total menor do que 0,30. Apesar desse fato, a imposição de dificuldade para as crianças que são boas leitoras por esta palavra direcionou a escolha. As demais palavras (*catorze*, *maroto*, *zebu* e *novelo*) foram escolhidas por apresentarem simultaneamente bons índices D e de correlação item-total.

Tabela 16

Palavras selecionadas a partir da simulação de análise dos itens-candidatos distribuídos em termos de dificuldade pelo critério de Pasquali (1996)

Nível de dificuldade	Palavra/R-1 Exemplo-1	Palavra/R-2 Exemplo-1	Palavra/R-1 Exemplo-2	Palavra/R-2 Exemplo-2
<.20	ampola	-	ampola	-
≥.20 e <.40	aceso	tola	aceso	tola
	exceto	boxe	exceto	boxe
≥.40 e <.60	humilde	catorze	cera	saxofone
	serelepe	guizo	serelepe	forca
	foco	maroto	foco	monarca
	sarjeta	palhoça	sarjeta	novelo
	-	bodoque	-	zebu
≥.60 e <.80	gingiva	reflexo	exposto	enxoval
	freguesa	cometa	caroço	doutora
≥.80	esquerdo	preguiça	teimoso	adulta

Como existem muitas palavras com dificuldade baixa, como critério para a seleção dos itens foi adotado um valor maior do que 0,60 no índice D e um valor maior que 0,40 na correlação item-total. Isto acarretou na seleção de quatro palavras da R-1 e cinco da LR-2. As palavras da R-1 foram então pareadas em termos de valores de índice D e de nível de regularidade (uma palavra IRRG para cada forma – *freguesa* e *caroço* – e a palavra regular *gingiva* no Exemplo-1 e a irregular *exposto* no Exemplo-2 para equiparar o número de palavras regulares em cada uma das formas). Para a LR-2, excluiu-se a palavra *colete* por apresentar o índice D menos adequado (0,60). As palavras *reflexo* e *enxoval* foram colocadas em formas separadas por conterem o grafema <x> e por apresentarem os maiores índices D.

Por fim, para a seleção das palavras muito fáceis o valor mínimo de D para as palavras da R-1 foi de 0,40 e de 0,50 para as palavras da R-2. Com isso, selecionaram-se quatro palavras em cada lista. As palavras *esquerdo* e *teimoso* foram selecionadas na R-1 por apresentarem os melhores índices D. As palavras *preguiça* e *adulta* foram selecionadas na Lista 2, a despeito de terem índices D menores do que os outros dois itens que ainda restavam, devido ao fato de serem regulares. Tanto na versão final dos exemplos aqui demonstrados, quanto na seleção de entrada dos itens-candidatos, houve poucos itens regulares. Assim, nas Formas de Exemplo, dos 32 itens apenas nove são regulares (cinco no Exemplo-1 e quatro no Exemplo-2). De forma semelhante, dos 41 itens com nível de dificuldade e discriminação dentro das faixas de critérios aqui adotados (reduzidos dos 113 itens-candidatos totais), apenas 11 eram regulares (25%). Assim, preservou-se com essa seleção uma proporção aproximada de 25% de palavras regulares com níveis adequados de dificuldade e discriminação²⁸, a despeito de as palavras regulares se concentrarem entre os itens de dificuldade de média a alta, o que já

²⁸ 25% dos itens no Exemplo-1 são regulares e 20% no Exemplo-2.

era esperado a partir das hipóteses apresentadas sobre a dificuldade das palavras apresentada na seção 4.

Em suma, esta análise mostrou como podem ser selecionados itens a partir de critérios previamente adotados. Assim, as versões finais de exemplos que foram obtidas seguiram o critério de apresentação de bons níveis de dificuldade e discriminação pelas palavras, assim como a manutenção de proporções em cada nível de dificuldade esperada de acordo com a curva normal. Com esse exemplo, pôde-se ver ainda como a teoria pode participar da seleção dos itens, uma vez que algumas palavras, a despeito de terem apresentado piores índices psicométricos, foram selecionadas por pertencerem a determinada classificação de regularidade ou número de letras.

7.4.2 Consistência interna e temporal

7.4.2.1 Procedimento

Com o objetivo de maximizar a semelhança entre as Listas 1 e 2 que compõem o banco de itens, estas foram pareadas em termos da classificação da regularidade da palavra (tanto para a leitura quanto para a escrita), do número de letras, da estrutura silábica e da semelhança estrutural (por exemplo, a presença de um dígrafo ou de uma letra inicial em comum). Esse procedimento objetivou proporcionar uma equiparidade entre as listas. Como as Formas de Exemplo 1 e 2 são compostas por palavras de ambas as listas (ou de suas formas reduzidas, LR-1 e LR-2), se comprovada a consistência entre as palavras advindas dessas listas é possível determinar a estabilidade temporal das Formas de Exemplo como um todo (uma vez que as Listas 1 e 2 foram aplicadas em dias diferentes). Desse modo, o objetivo dessa seção é descrever a consistência interna

das Formas de Exemplo 1 e 2 utilizando método das metades (*split half*) e, confirmada, a consistência, passar para a determinação da estabilidade temporal do instrumento, fazendo-se uso, portanto, da técnica de fidedignidade de forma alternativa com intervalo.

7.4.2.2 Resultados

7.4.2.2.1 Consistência interna

Conforme vimos, as Formas de Exemplo obtidas descritas na seção anterior são compostas por 10 palavras retiradas da Lista 1 e 10 palavras da Lista 2²⁹. Assim, na divisão pelo método das metades (*split half*), a primeira metade das Formas de Exemplo 1 e 2 ficou composta de palavras da Lista 1, enquanto que a segunda metade foi constituída de palavras da Lista 2. A Tabela 17 mostra os valores de alfa de Cronbach, de correlação e do coeficiente de Spearman-Brown para as Formas de Exemplo 1 e 2 e de suas respectivas metades.

Os valores dos coeficientes mostrados na Tabela 17 nos revelam que a consistência interna das Formas de Exemplo 1 e 2 do instrumento é bastante satisfatória. Houve uma tendência de a metade 2 de ambas as formas apresentar valores levemente mais elevados de consistência interna do que a metade 1, mas nos dois casos os níveis atingidos de alfa de Cronbach foram aceitáveis. A correlação entre as metades das duas formas demonstrou uma relação moderada, enquanto que a correção pela fórmula de Spearman-Brown elevou os coeficientes de correlação a níveis considerados altos.

²⁹ Uma vez que a lista reduzida 1 (LR-1) foi produzida a partir da Lista 1, enquanto que a forma reduzida 2 (LR-2) constitui-se de palavras da Lista 2.

A partir do que foi exposto, pode-se concluir que as Formas de Exemplo 1 e 2 apresentaram níveis adequados de consistência interna. Além disso, a análise pelo método das metades mostrou que existe consistência entre as palavras retiradas das Listas 1 e 2, o que indica que essas listas podem ser consideradas formas alternativas. Com esse resultado, torna possível avaliar a estabilidade temporal das palavras retiradas das Listas 1 e 2.

Tabela 17

Valores dos coeficientes alfa de Cronbach para as Formas de Exemplo 1 e 2 e valores de correlação e de Spearman-Brown para as metades de cada uma dessas formas

Forma	Coefficientes	Valores
Exemplo-1	Alfa de Cronbach	0,82
	Alfa Metade 1	0,67
	Alfa Metade 2	0,73
	Correlação	0,69
	Spearman-Brown	0,82
Exemplo-2	Alfa de Cronbach	0,82
	Alfa Metade 1	0,67
	Alfa Metade 2	0,72
	Correlação	0,67
	Spearman-Brown	0,80

7.4.2.2.2 Estabilidade temporal

Esta seção tem como objetivo avaliar em que medida os escores dos participantes são afetados por amostragem de tempo. A lacuna de tempo entre a aplicação das Listas 1 e 2 variou de zero a 141 dias. Três faixas de tempo entre a aplicação das listas foram então selecionadas: até sete dias ($N = 262$), entre oito e 30 dias ($N = 40$) e mais de 31 dias ($N = 29$). Também foram reportadas as análises para todo o período ($N = 331$).

A Tabela 18 mostra as correlações de Pearson feitas entre os escores obtidos pelas crianças nas Listas 1 e 2, as quais estão distribuídas conforme as faixas de tempo supracitadas e para todo o período. As palavras retiradas da Lista 1 podem ser vistas nas colunas 2 e 4 da Tabela 16 e as palavras retiradas da Lista 2 estão nas colunas 3 e 5 da mesma tabela.

Tabela 18

Valores de correlação de Pearson para a precisão das respostas dos participantes nas palavras selecionadas das Listas 1 e 2 em cada intervalo de tempo selecionado.

Período	Correlação
Até 7 dias	0,77
Entre 8 e 30 dias	0,73
Mais de 30 dias	0,83
Total	0,78

No geral, houve uma relação moderada entre as palavras selecionadas das Listas 1 e 2 em todos os intervalos de tempo utilizados, com exceção do período superior a 30

dias, em que a relação pode ser considerada forte. Todas as correlações encontradas foram significativas ao nível de $p < 0,001$. Os resultados são condizentes com a interpretação de que os escores obtidos pelos participantes apresentam estabilidade temporal, dentro das faixas de tempo aqui estipuladas.

7.4.3 Concordância entre avaliadores

7.4.3.1 Procedimento

Participaram do processo de correção da leitura das crianças cinco avaliadores devidamente treinados (quatro estudantes de graduação do sétimo período de psicologia – que já completaram, portanto, todas as disciplinas de avaliação psicológica – e uma lingüista), além da própria mestrandia. Os critérios para se considerar uma resposta como acerto foram descritos no método. A análise da concordância entre os avaliadores se baseou nas correlações entre as correções protocolos dos participantes estabelecidas por três dos cinco avaliadores.

No processo de correção, dois dos avaliadores corrigiram os mesmos protocolos de nove crianças (Lista 1 ou Lista 2 de cada criança). Sendo assim, esses dados foram aproveitados para a análise da fidedignidade entre avaliadores. Um terceiro avaliador foi então sorteado e corrigiu os protocolos novamente. A Tabela 19 mostra a distribuição dos participantes desta análise por série, sexo e escola.

A concordância entre os avaliadores foi feita separadamente para as Formas de Exemplo 1 e 2 por meio do coeficiente W de Kendall (que fornece valores que podem oscilar entre 0 e 1 e testa a hipótese nula de que as variáveis – no caso, as avaliações dos juízes – são independentes). Quanto mais o coeficiente W se aproxima de 1, p tenderá a

zero e, portanto, se rejeitará a hipótese nula, o que indica concordância entre os juízes (Bisquerra et al., 2004).

Tabela 19

Descrição das características dos participantes que fizeram parte da análise de concordância entre avaliadores.

Série	Sexo	Escola
1	M	1
1	M	7
1	M	7
2	F	4
2	M	1
2	M	4
3	F	5
3	F	6
4	M	2
4	M	2

7.4.3.2 Resultados

Para ambas as Formas de Exemplo, houve uma forte concordância entre os avaliadores, apesar de ter havido uma tendência maior de concordância no Exemplo-2 ($W = 0,92$, $p < 0,005$ para o Exemplo-1 e $W = 0,99$, $p < 0,003$ para o Exemplo-2).

8 CONSIDERAÇÕES FINAIS

A presente pesquisa reportou as etapas intermediárias conduzidas para a investigação psicométrica de uma tarefa de leitura em voz alta de palavras isoladas – prova computadorizada de reconhecimento de palavras para crianças –, a qual será destinada à avaliação de alunos cursando da 1ª à 4ª série do Ensino Fundamental da cidade de Belo Horizonte. Especificamente, reportou-se uma análise de itens a partir da qual se selecionou os itens-candidatos a fazer parte da referida prova de leitura. Além disso, foram descritos efeitos na leitura (treino, fadiga, regularidade, extensão e série) e foi simulada uma seleção de itens a partir da amostra de itens-candidatos. Por meio das Formas de Exemplo geradas por esse procedimento, foi possível demonstrar a consistência interna e temporal dos itens selecionados, assim como a consistência nas correções dos avaliadores. A análise destes resultados mostrou que estes procedimentos, em conjunto, forneceram indícios a respeito da validade do construto investigado (neste caso, a habilidade de reconhecer palavras fora de contexto) por terem sido capazes de produzir informações que corroboram com as hipóteses levantadas e com as interpretações esperadas a respeito dos escores nas listas de palavras (tanto a original e quanto as reduzidas).

A análise de itens confirmou a hipótese de que, do ponto de vista de acurácia, o banco de itens testado não apresentaria muitas dificuldades para os leitores de uma maneira geral. A grande maioria dos itens mostrou-se fácil ou muito fácil para toda a amostra e apenas uma pequena proporção de itens apresentou algum nível de dificuldade maior para os leitores. Isso corrobora com questões levantadas por pesquisadores a respeito do caráter regular ou previsível de nossa língua (p. ex., Parente, Silveira & Lecours, 1997), que impõe pouca ou nenhuma dificuldade mesmo para os

leitores iniciantes. Este fato se torna evidente quando observamos que apenas cerca de 1,5% do total de itens apresentou um índice $p < 0,40$ (ver Tabela 9). Considerando que mais de 50% das palavras foram classificadas como irregulares para a leitura, a porcentagem de itens que impuseram dificuldades para a amostra nos faz refletir sobre a classificação de regularidade utilizada no português brasileiro. Diante desse fato, pelo menos tomando a direção da leitura, talvez fosse mais apropriado se falar em ambigüidade ou em inconsistência do que propriamente em “irregularidade” entre as palavras do português que apresentam pronúncias que não são regulares ou não são previsíveis por regras na leitura. De fato, Silva e Rothe-Neves (2002) demonstraram que o grafema <x>, independentemente da posição que se encontra na palavra, em 75% das vezes apresenta uma pronúncia consistente (lido como /ʃ/), ou seja, a criança que desconhece qualquer outra pronúncia do <x> tem 75% de acertar na leitura de uma palavra contendo esse grafema³⁰. Isso não caracteriza uma irregularidade, tal como se observa em outros idiomas, como o inglês.

A despeito desse fato, a onipresença de itens irregulares entre as palavras mais difíceis (índice $p < 0,40$) e a quantidade marcante de itens com esse tipo de regularidade (69%) que apresentou os níveis satisfatórios de discriminação e de dificuldade atentam para o poder do efeito de regularidade no português brasileiro. A presença deste efeito na seleção dos itens demonstra que a dificuldade que a palavra apresenta ao leitor iniciante está intimamente relacionada ao processo de leitura que é utilizado por ele no momento da realização da tarefa. Por este motivo, tal fato corrobora de forma indireta com a aplicabilidade do modelo de dupla-rota no nosso idioma, uma vez que os itens

³⁰ Pelo menos entre as palavras de alta frequência, de onde foi extraída proporção feita por Silva e Rothe-Neves (2002) da representação de fonemas por grafemas da nossa língua. No entanto, é possível que essa proporção se mantenha mesmo entre as palavras de alta frequência, uma vez que, das 19 palavras que apresentam o grafema <x> no banco de itens utilizado no presente estudo, mais de 68% apresentam a pronúncia consistente do grafema como /ʃ/.

irregulares, quando lidos pela rota fonológica, podem estar sujeitos a erros de pronúncia, o que justamente caracteriza a dificuldade dos itens com os quais estamos lidando.

Esta hipótese é reforçada pela configuração da seleção simulada de itens que foi aqui desenvolvida. Por meio de tal simulação, não foi possível sequer criar duas formas provisórias de listas de palavras que se diferem em termos dos níveis de dificuldade mais elevada, tendo sido necessário se criar itens-âncora para esta finalidade. Para os 32 itens selecionados nesta simulação, 75% eram irregulares para a leitura. Como o critério de entrada dos itens nas Formas de Exemplo simuladas diferiu-se da análise de itens principalmente em termos dos níveis de discriminação dos itens (que foram mais exigentes), pode-se afirmar que os itens irregulares são também mais discriminativos do que os regulares (uma vez que a proporção desses itens subiu de 69% da análise de itens para 75% da seleção simulada de itens).

Apesar de ser plausível a hipótese de que o efeito de regularidade está relacionado ao nível de dificuldade da palavra, outra possibilidade deve ser levada em consideração. É possível que outras características das palavras além da regularidade (como o número de palavras vizinhas ou a presença de um grafema muito incomum) que não foram estudadas possam ter interferido nos resultados apresentados. Isso fica mais evidente a partir dos índices de correlação item-total que as palavras de dificuldade extremamente alta apresentaram, os quais ficaram abaixo de 0,30 na maioria dos casos (60%). Este fato demonstra que habilidades outras que não a de reconhecer palavras³¹ possa ter interferido no sucesso na leitura desses itens, uma vez que os índices de consistência interna que aqui foram obtidos (dados pelo alfa de Cronbach) apontaram para uma forte unidimensionalidade do construto investigado.

³¹ O conhecimento da pronúncia de grafemas infreqüentes pode ser um desses fatores. O grafema <xc>, por exemplo, apesar de apresentar a pronúncia regular /s/, por ser extremamente infreqüente (Silva & Rothe-Neves, 2002), pode não ser lido corretamente por leitores iniciantes ou pouco habilidosos.

As questões relativas à dificuldade das palavras no português, aliadas à configuração final dos itens-candidatos obtidos pela análise³² nos faz refletir sobre a aplicação em nosso idioma das recomendações de Pinheiro de Rothe-neves (2001) em relação às características das listas de palavras utilizadas para uma avaliação cognitiva da leitura. Para esses autores, com base na teoria de reconhecimento de palavras, é essencial que as palavras de uma lista construída para fins de permitir uma avaliação dos processos de leitura sejam pareadas em termos de regularidade, frequência e em número de letras a fim de se extrair conclusões a respeito do processamento utilizado pela criança na leitura. No entanto, conforme foi demonstrado, em uma lista de centenas de itens de baixa frequência de ocorrência, apenas 7% apresentaram índices de dificuldade mais elevada ($p < 0,40$) para distinguir as habilidades de leitura de crianças em fase de alfabetização. Isso mostra que as recomendações dos autores parecem se aplicar mais aos estudos que visam identificar as estratégias utilizadas pelos leitores durante a leitura do que a um trabalho em que a metodologia se liga aos procedimentos psicométricos de construção de um teste, cujo objetivo é discriminar leitores competentes dos não-competentes. Os resultados aqui apresentados mostram que o pareamento de palavras que se diferem em termos de frequência³³ e regularidade pode não ter relevância do ponto de vista normativo, sendo talvez importantes apenas os índices psicométricos de dificuldade de discriminação. Obviamente, o referido pareamento, assim como de outros tipos discutidos na revisão de literatura aqui apresentada, permanece relevante no momento de se investigar a validação do construto da tarefa, conforme explicitado nos parágrafos seguintes, assim como no momento de se

³² Que, como vimos, apresentam uma alta proporção de palavras irregulares e diferentes distribuições em termos de número de letras.

³³ Isso pode ser afirmado por inferência, uma vez que, como dito no início desse parágrafo, apenas uma minoria dos itens presentes no corpus original apresentou alguma dificuldade para a amostra. Considerando que as palavras de baixa frequência são mais difíceis de serem lidas do que as palavras comuns, uma vez que sua representação no léxico mental é mais fraca, é esperado que as palavras de alta frequência imponham uma dificuldade menor ainda para os leitores.

conduzir uma avaliação dos processos de leitura que mudam de acordo com a competência do leitor e de seu estágio de desenvolvimento nessa habilidade. Assim, uma hipótese é de que a inclusão de palavras de alta frequência de ocorrência entre os itens-candidatos que foram aqui obtidos apenas acrescentaria itens redundantes e que nada contribuem para a avaliação da habilidade geral de leitura requerida por um teste.

A questão aqui levantada acerca da aplicabilidade do pareamento de itens em termos de regularidade e frequência para a avaliação normativa da habilidade de reconhecimento de palavras pode ser esclarecida pela diferenciação feita por Embretson (1983) a respeito das pesquisas de validação de construto, nomeadas por ele como pesquisas de representação do construto e pesquisas de intervalo nomotético. No primeiro caso, o objetivo é identificar os mecanismos teóricos que estão envolvidos no desempenho de uma tarefa, ou seja, há uma busca pela decomposição da tarefa em etapas irreduzíveis. Desta forma, no paradigma do processamento de informação a representação do construto se relaciona à dependência que o desempenho nas tarefas possui em relação aos processos, estratégias e conhecimento anteriores do sujeito. Apesar de esta dependência poder ser um aspecto essencial para o desempenho na tarefa, isso não implica que as pessoas irão necessariamente variar em termos deste desempenho. Em outras palavras, a pesquisa de representação do construto está interessada com a variabilidade das tarefas de um teste e não necessariamente com as diferenças individuais produzidas nas mesmas. No caso da presente pesquisa, por exemplo, em um banco de itens com centenas de palavras consideradas mais difíceis para os leitores por serem pouco familiares, ocorreu uma proporção geral de 81% de acertos em uma amostra composta por leitores competentes e menos habilidosos. Desta forma, por mais que supostamente os processos de leitura utilizados por ambos os tipos de leitores sejam diferentes (ou seja, o uso de processamento lexical no primeiro grupo

e um apego à estratégia fonológica no segundo), esta diferença não refletiu em uma grande variabilidade nos escores dos participantes, o que é representado pela assimetria na distribuição dos escores do banco de itens.

A pesquisa de intervalo nomotético, por outro lado, se refere à rede de inter-relações que se pode estabelecer entre um teste e outras medidas (que meçam o mesmo construto ou mesmo construtos diferentes). Nesse sentido, este tipo de pesquisa se preocupa com diferenças individuais e se interessa pela variabilidade dos escores dos participantes nas tarefas.

Pode-se dizer que os estudos experimentais que objetivam observar os efeitos diversos na leitura de palavras (e, portanto, validar os modelos que distinguem os processos e as etapas utilizados durante o ato de ler) constituem pesquisas de representação do construto. Enquanto isso, as pesquisas voltadas para a construção de testes e o estabelecimento de normas que são utilizadas para classificar os sujeitos em diferentes níveis de habilidade são caracterizadas como pesquisas de intervalo nomotético. Desta forma, ambos os tipos de pesquisa (de representação do construto e de intervalo nomotético) irão buscar fontes distintas de validação que são independentes umas das outras. Aplicando essa distinção para as questões levantadas acima, pode-se dizer que o pareamento das palavras em termos de características psicolinguísticas é relevante para a validação da pesquisa experimental na área da leitura de palavras isoladas, enquanto que a sua separação em termos de dificuldade e discriminação é importante para as comparações intragrupo que são feitas na mesma área.

A partir do que foi exposto, restam aqui duas possibilidades. A configuração da regularidade das palavras numa análise de itens feita em uma segunda amostra de participantes, a partir dos itens-candidatos aqui gerados, pode ou não ser mantida. Isto é, em uma segunda amostra, os itens podem ou não manter o comportamento que foi aqui

exibido nesta seleção simulada de itens, no que diz respeito à proporção de itens regulares e irregulares³⁴. No caso de se manter uma grande proporção de itens irregulares, a versão final do teste poderia ser útil apenas para uma análise das diferenças entre os testandos em termos do que é esperado em seu desempenho em função da série. Isso implica que esta versão não permitiria a extração de inferências a respeito dos processamentos utilizados na leitura pelos participantes, conforme seria esperado se as palavras fossem pareadas em termos de regularidade e extensão (e, posteriormente, de frequência). Se isto se comprovar, o alcance de um objetivo não seria obtido sem o sacrifício do outro. Dito em outros termos, a criação de um teste de palavras isoladas capaz de apresentar variabilidade suficiente a ponto de distinguir as habilidades de indivíduos em níveis diferentes de desenvolvimento poderá exigir a desconsideração do pareamento de palavras que é previsto pela teoria. De maneira análoga, a criação de um instrumento capaz de testar os componentes do modelo de dupla-rota poderia ser alcançada à custa da redução da variabilidade das respostas dos participantes (representada pela entrada no teste de itens mais fáceis – regulares e pequenos – passíveis de serem lidos sem dificuldade pelos leitores menos competentes). Somente as investigações posteriores nos dirão se as hipóteses aqui levantadas serão ou não confirmadas.

Uma possível de solução para esse impasse pode ser pensada se fossem criados artifícios que aumentassem a dificuldade dos itens regulares para a leitura entre os leitores menos competentes, o que acarretaria em um impacto na discriminação desses itens, forçando sua entrada na seleção de itens-candidatos. Isso seria possível apenas por meio da alteração do critério de acerto utilizado na pesquisa. Segue-se uma sugestão, que se relaciona ao uso do tempo de reação para a análise dos erros. Conforme visto, o

³⁴ E também grandes e pequenos, uma vez que apenas 25 palavras (22% da amostra total de itens-candidatos) apresentam 4 ou 5 letras.

presente trabalho considerou as respostas de silabação como erro, assim como foi feito em outros trabalhos (p. ex., Pinheiro, Lúcio, & Silva, 2008; Pinheiro, Lúcio, & Cunha, 2008). Também dissemos que estas respostas estão sujeitas à interferência da avaliação do juiz. Assim, apesar dos bons índices de concordância entre os juízes obtidos na simulação aqui reportada, um ponto fundamental para a conferência dos escores seria a eliminação da necessidade de julgamento nesse tipo de erro. Isso seria possível por meio do uso da medida da duração da resposta, ou seja, a soma do tempo de reação com a duração locucional (Capovilla, Capovilla, & Macedo, 2001). Para um dado item, a média e o desvio padrão dessa medida poderia ser utilizada como critério de acerto na palavra. As respostas que se desviassem muito dessa medida (por exemplo, dois desvios-padrão) seriam consideradas incorretas. Isso impediria, por exemplo, que fossem consideradas corretas tanto as respostas de silabação explícita (por ser a leitura feita com pausas, a duração locucional é maior) quanto implícita (ou seja, quando o participante silaba para si a palavra e, em seguida, emite uma resposta, que pode estar correta, principalmente se a palavra lida for regular. Isso aumenta o tempo de reação³⁵). Com essas medidas, é possível que o número de erros cresça entre os leitores menos competentes, que fazem uso do processo fonológico na leitura, aumentando, assim, a discriminação e a dificuldade mesmo dos itens regulares. Isso proporcionaria uma maior variabilidade dos itens dentro de cada espectro de nível de dificuldade, melhorando as possibilidades de escolha de itens.

Em síntese, da mesma forma que o critério de Pasquali (1996) para a seleção de itens propicia a variabilidade intergrupo ao forçar uma distribuição normal dos escores, uma maior rigidez nos critérios de acerto poderia levar ao aumento da oferta de itens regulares e pequenos entre aqueles que apresentam índices adequados de dificuldade e

³⁵ É importante notar que nenhum controle sobre a silabação implícita foi feito no presente estudo.

discriminação. Com isso, seria possível conciliar em um mesmo teste tanto a manifestação de diferenças entre os testandos quanto dos processos de leitura por eles utilizados. Para esse fim, a análise de itens aqui reportada deveria ser refeita seguindo-se esses critérios.

No que diz respeito ao efeito de regularidade, uma hipótese que foi lançada nesse estudo é que a emergência desse efeito na precisão e no tempo de processamento na leitura poderia variar em função do número e do tipo de categorias que se utiliza na pesquisa. De fato, os resultados obtidos pelo modelo de Equação de Estimação Generalizada (EEG) mostraram que o uso de quatro categorias de regularidade introduzidas por Pinheiro (2004) e testadas no presente trabalho, é mais eficaz para explicar o desempenho das crianças do que o uso de duas categorias apenas. As comparações entre as diversas categorias mostraram que o efeito maior ocorre nas comparações entre as palavras regulares e as irregulares-regra. Isto sugere que, considerando as palavras irregulares, faz sentido a distinção introduzida por Pinheiro (2004) entre palavras que apresentam apenas correspondências irregulares daquelas que, adicionalmente a uma correspondência irregular, apresentam também correspondências dependentes de regras contextuais. Nesse sentido, parece que o primeiro grupo de itens impõe maior demanda de processamento do que o segundo.

Esses resultados, aliados à presença de diferenças significativas na leitura das palavras regulares em relação às palavras regras em termos de precisão, sugerem que o domínio de regras contextuais na leitura pode não se estabelecer muito cedo durante a aprendizagem da língua, como sugerem alguns autores (p. ex., Pinheiro e Rothe-Neves, 2001). Pelo menos no que se referem aos erros, essas palavras continuam impondo alguma dificuldade aos leitores, dificuldade esta que supera à que é imposta pelos itens cuja pronúncia é independente de contexto. Entretanto, um fato importante a ser

mencionado é que a classificação das palavras adotada nesse estudo gerou um maior número de palavras dependentes de regras contextuais do que outros estudos que não encontraram diferenças significativas entre as palavras regulares dependentes e independentes de contexto (p. ex., nos estudos de Pinheiro (1995) e de Pinheiro, Lúcio, & Silva, 2008). Este fato reforça a hipótese de que, dependendo da classificação utilizada pelo estudo, os efeitos entre as categorias de palavras podem ser observados ou suprimidos. Além disso, uma análise detalhada dos tipos de erros cometidos em cada palavra poderá confirmar se o efeito aqui obtido se deve ou não a descondições de regras contextuais.

A hipótese de que a regularidade para a escrita das palavras é relevante para o tempo de reação recebeu suporte empírico e confirma o efeito de regularidade encontrado por Pinheiro (1994, 1995, no prelo) em uma lista de palavras cuja classificação tomou como referência a direção fonema-grafema. Ainda não está claro porque esse efeito ocorre, mas é possível que o fato de a aquisição das regras na escrita se efetuar de forma mais lenta e laboriosa do que as regras para a leitura tenha alguma interferência (Parente, Silveira, & Lecours, 1997). Por exemplo, Pinheiro e Rothe-Neves (2001) levantam a hipótese de que mesmo palavras regulares para a escrita podem causar dificuldades na leitura das crianças mais novas devido ao seu domínio mais fraco de conhecimento das relações fonema-grafema. Nesse sentido, pode-se supor que a falta de domínio das regras na escrita de um modo geral possa causar uma lentidão no tempo de processamento das crianças (e mesmo na precisão, como o que ocorreu aqui).

De maneira análoga ao que foi encontrado no trabalho de Pinheiro et al. (2008), o efeito de regularidade na precisão ocorreu tanto quando se considera a classificação de regularidade das palavras para a leitura quanto da escrita. Esse achado facilita o estudo

do efeito de regularidade no português, uma vez que é possível encontrar esse efeito simultaneamente nos erros e no TR controlando-se a classificação para a escrita das palavras. O efeito na precisão enfraquece com a adoção desta classificação em comparação com a classificação para a leitura, mas ainda mostra-se presente e, para a pesquisa, esta é uma boa notícia, uma vez que reduz o número de variáveis a serem controladas pelo pesquisador.

Os efeitos de série e de extensão apareceram na direção esperada em ambos os modelos de precisão e de TR. É interessante notar que, em relação ao efeito de extensão, o aumento do TR para cada acréscimo de letra na palavra encontrado neste trabalho (63mseg) se assemelhou ao que foi encontrado por Pinheiro (1999). A autora, utilizando uma medida de amplitude de variação, descobriu que, para cada acréscimo de letra na palavra, houve um aumento de 1 a 70mseg no TR.

O fato de a criança fazer pré-escola foi um fator muito importante encontrado na precisão. As crianças com pré-escola têm, em média, um ano a mais de escolarização, e isso parece ter um efeito benéfico para o desempenho na tarefa em que estamos lidando. O fato de termos na amostra crianças que fizeram ou não pré-escola não prejudica os dados aqui reportados por duas razões: (1) as crianças que não fazem pré-escola existem na população geral e (2) apenas 3% das crianças da amostra estão nessa situação.

Em consonância com os achados de Zevin e Seidenberg (2002), a idade de aquisição da leitura foi um fator de grande importância para a proficiência na leitura. Apesar de ter sido utilizado uma avaliação subjetiva e aproximada desse fator (ou seja, a heteroavaliação dos pais), este se mostrou importante para a explicação do modelo.

Contrariando as expectativas, não houve efeito de gênero na leitura. O efeito de gênero tem sido constantemente relatado em tarefas que medem habilidades verbais, sendo que os indivíduos do sexo feminino geralmente apresentam melhores

desempenhos do que os do sexo masculino (p. ex., Chiu & McBride-Chang, 2006; Hedges & Nowell, 1995; Rosén, 2001). Os resultados aqui apresentados estão em consonância com o ponto de vista de outros autores (p. ex., White, 2007) de que as diferenças de gênero têm sido superestimadas, ou seja, de que as diferenças de gênero, quando aparecem, podem não ter uma significância prática, uma vez que os tamanhos dos efeitos são geralmente pequenos.

Nenhuma das medidas de inteligência utilizadas foram importantes para o desempenho na leitura, o que está de acordo com o esperado (p. ex., Carraher & Rego, 1984; Nunes, Buarque, & Bryant, 1992; Siegle, 1992; Toth & Siegle, 1994). A exceção constituiu a compreensão verbal (medida em QI), que se mostrou um fator relevante tanto para a precisão quanto para o TR. Esses resultados estão em desacordo com o que foi obtido por Maia e Fonseca (2002) com crianças brasileiras. No entanto, uma diferença fundamental entre o presente trabalho e o conduzido pelas autoras é o fato de termos utilizado aqui o WISC-III, que apresenta adaptação e normas para nossa população, e as autoras terem utilizado o WISC. Existe a possibilidade de que a compreensão verbal, medida pelo WISC-III, tenha captado o conhecimento verbal geral das crianças as quais são importantes para o desempenho na leitura. Por exemplo, em estudo conduzido com 73 crianças com risco genético para dislexia, foi demonstrado que essas crianças não se diferiram do grupo de controle em termos de inteligência não-verbal – assim como no presente estudo – mas tiveram um desempenho mais fraco em tarefas de fala e de processamento da linguagem, particularmente a compreensão do vocabulário e a nomeação (Gallagher, Frith & Snowling, 2000). De fato, vários estudos têm apontado que crianças com dificuldade na aprendizagem da leitura e da escrita apresentam problemas com tarefas que medem habilidades lingüístico-verbais e fonológicas (Carroll & Snowling, 2004; Siegel & Linder, 1984; Snowling, 2004).

O único fator social que interferiu na leitura da amostra foi a escolaridade da mãe, que mostrou uma relação negativa com a precisão (ou seja, os filhos de mulheres altamente escolarizadas apresentaram pior desempenho na leitura do que daquelas que apresentam baixo nível de escolaridade). Uma hipótese plausível seria a de que as mães com escolaridade alta, por trabalharem fora, não dedicam tempo suficiente para a educação dos filhos. Entretanto, D´Affonseca (2005) mostrou que não existem diferenças no desempenho na leitura de crianças cujas mães trabalham fora daquelas em que as mães são donas de casa. Ademais disso, os resultados deste trabalho ainda apontam para uma relação inesperada entre as características da mãe e o desempenho da criança, tal como a que aqui obtivemos. Assim, foi encontrado que quanto maior a participação de mães donas-de-casa nas atividades escolares, culturais e de lazer do filho, pior o desempenho na leitura da criança. Desta forma, a interpretação desses resultados merece um estudo à parte, que avalie de maneira mais ampla o contexto social da criança e sua relação com o desempenho acadêmico.

Um fato interessante foi que o nível socioeconômico da família da criança não interferiu na proficiência da amostra (nem em termos de TR, nem de precisão). Isto é um resultado bastante positivo para a educação no Brasil, uma vez que estamos acostumados a ver tanto na literatura nacional quanto na internacional dados que apontam para diferenças socioeconômicas importantes em várias áreas de domínio cognitivo (p. ex. Duncan & Seymour, 2000; Sisto, 2006; Soares, 2004). No entanto, os autores geralmente relatam diferenças socioeconômicas grupais, enquanto que no presente estudo consideramos diferenças individuais (representadas pela renda da família e não, por exemplo, o tipo de escola, que seria um fator grupal). Existe a possibilidade de que os fatores grupais ainda sejam importantes para a manifestação de diferenças em habilidades cognitivas, devido ao efeito do grupo de pares (Hanushek,

Kain, Markman, & Rivkin, 2003). Uma vez que um maior nível socioeconômico propicia o acesso a vários tipos de recursos culturais pela criança, e que as escolas particulares concentram estudantes com este nível, poderia haver um efeito do nível socioeconômico na proficiência do grupo de alunos, em contraposição a um efeito anulado ou inexistente do nível socioeconômico quando se considera o indivíduo de maneira isolada (Soares, 2006).

A análise da fidedignidade pelo método da forma alternativa com intervalo mostrou que ambas as Formas de Exemplo 1 e 2 apresentam uma estabilidade temporal satisfatória em todas as faixas de tempo investigadas. A aplicação desse método só foi possível pelo fato de ter se demonstrado uma consistência entre as Listas 1 e 2, as quais foram aplicadas na amostra em dias diferentes. A presença de consistência entre as palavras das Listas 1 e 2 das Formas de Exemplo constitui mais uma evidência de que a dificuldade dos itens está relacionada às características psicolinguísticas da palavra (como a regularidade e a extensão), pois elas foram pareadas justamente em termos dessas características. Outro ponto importante a ser mencionado representa os índices adequados de consistência interna obtidos nas Formas de Exemplo 1 e 2 (0,82 para as duas formas).

A análise de fidedignidade entre avaliadores mostrou que houve uma concordância quanto aos escores conferidos aos participantes nas Formas de Exemplo 1 e 2. Este fato é importante porque indica que os critérios de acerto explicitados para os avaliadores foram eficientes para nortear o processo de correção do teste e que, portanto, asseguraram a validade dos escores obtidos pelos participantes. O resultado também demonstra a relevância da objetividade no momento de se estabelecer os parâmetros para a correção de testes que podem estar propensos a algum tipo de viés de subjetividade por parte do avaliador (Urbina, 2007).

Enfim, a análise da magnitude dos efeitos (η^2) de treino (na precisão da Lista 1) e de fadiga (no tempo de reação de ambas as listas) na leitura das crianças nos fez suspeitar que não houvesse uma significância prática dos mesmos, a despeito da significância estatística encontrada. Nesse caso, uma hipótese seria de que o efeito estatístico tivesse sido alcançado em função do tamanho da amostra de participantes e de itens. Como foram feitas muitas comparações em uma amostra de tamanho considerável, os resultados podem ter sido induzidos ao erro Tipo II. Com o objetivo de investigar essa hipótese, colocamos no modelo de regressão a variável “ordem de apresentação da palavra” e demonstrou-se que, pelo menos para o TR, a hipótese de erro Tipo II não parece plausível. Assim, para a precisão, não se confirmou o efeito de treino, mas para o TR é possível que tenha ocorrido um efeito de fadiga. Apesar disso, as outras variáveis do modelo de TR, conforme vimos, não sofreram alterações em termos de seu poder explicativo sobre os resultados.

Em função do que foi exposto, conclui-se que todos os procedimentos conduzidos nesse trabalho para se efetuar a investigação psicométrica de um banco de palavras de baixa frequência, assim como de suas formas reduzidas (respeitando-se, nesse último caso, as ressalvas feitas), trouxeram informações que corroboram com a validade da tarefa em questão. Os itens-candidatos obtidos por meio da análise de itens apresentaram índices altamente satisfatórios de discriminação e de dificuldade. Além disso, as Formas de Exemplo obtidas por meio da seleção simulada de itens mostrou níveis satisfatórios de consistência interna e de estabilidade temporal. A análise de regressão pelo modelo GEE mostrou que todos os fatores tradicionalmente considerados relevantes para a proficiência na leitura contribuíram, de fato, de maneira significativa para os modelos criados e apontou novas fontes de influência (como a compreensão verbal e a escolaridade da mãe). Além disso, o trabalho contribuiu com questões

teóricas importantes em relação à emergência do efeito de regularidade no português brasileiro ao demonstrar que o número de categorias utilizadas no trabalho pode afetar a emergência e a extensão do efeito de regularidade e que o controle da regularidade do ponto de vista da escrita pode ser suficiente para a obtenção do efeito de regularidade que, como vimos, é de fundamental importância para a identificação do processo de leitura empregado pelo leitor. Por tudo isso, este estudo pode ser considerado de grande relevância para a pesquisa em avaliação psicológica no Brasil e certamente indicará o caminho para a construção de um instrumento diagnóstico para as dificuldades específicas de aprendizagem que acometem as crianças em fase de desenvolvimento da leitura.

9 REFERÊNCIAS

- Abbad, G. & Torres, C. V. (2002). Regressão múltipla *stepwise* e hierárquica em psicologia organizacional: aplicações, problemas e soluções. *Estudos em Psicologia*, 7 (número especial), 19-29.
- Adánez, G. P. (1999). Procedimientos de construcción y análisis de test psicométricos. In S. M. Wechsler & R. S. L. Guzzo (Eds.), *Avaliação Psicológica: Perspectiva Internacional* (pp. 57-100). São Paulo: Casa do Psicólogo.
- AERA, APA, & NCME (1999). *Standards for educational and psychological testing*. Washington, D.C.: Author.
- Anastasi, A. & Urbina, S. (2000). *Testagem Psicológica* (Maria Adriana Verissimo Veronese, Trad.). Porto Alegre: Artes Médicas.
- Andrews, S. (1989). Frequency and Neighborhood Effects on Lexical Access: Activation or Search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(5), 802-814.
- Andrews, S. (1992). Frequency and Neighborhood Effects on Lexical Access: Lexical Similarity or Orthographic Redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(2), 234-254.
- Bisquerra, R., Sarriera, J. C., & Martinez, F. (2004). Introdução à estatística: enfoque informático com o pacote estatístico SPSS. Porto Alegre: Artmed.
- Brancahhone, P. G., Fogo, J. C., & Williamseg, L. C. A. (2004). Crianças expostas à violência conjugal: Avaliação do desempenho acadêmico. *Psicologia: teoria e pesquisa*, 20(2), 113-117.

- Capellini, S. A., Tonelotto, J. M. F., & Ciasca, S. M. (2004). Medidas de desempenho escolar: avaliação formal e opinião de professores. *Revista Estudos de Psicologia, Puc-Campinas*, 21(2), 79-90.
- Capellini, V. L. M., Mendes, E. G., & Salgado, H. M. (2003). Avaliando o rendimento escolar de alunos com deficiência auditiva em escolas regulares. *Revista espaço*, 18-19, dezembro de 2002 a julho 2003.
- Capovilla, A. G. S. & Capovilla, F. C. (1998). Desenvolvimento de leitura e escrita de pré 3 à 2ª série: tabelas preliminares de normatização de uma lista de itens psicolinguísticos. *Ciência Cognitiva: Teoria, Pesquisa e Aplicação*, 2(4), 821-840.
- Capovilla, A. G. S., Capovilla, F. C., & Macedo, E. C. (1998). Validação do *software* CronosFonos para a análise de tempo de reação, duração e frequência de segmentação locucionais na leitura em voz alta de itens isolados. *Ciência Cognitiva: Teoria, Pesquisa e Aplicação*, 2(3), 253-340.
- Capovilla, A. G. S., Capovilla, F. C., & Macedo, E. C. (2001). Rota pré-lexical na leitura em voz alta: tempo de reação, duração e segmentação da pronúncia. *Psicologia: Reflexão e Crítica*, 14(2), 409-427.
- Capovilla, F. C., Varanda, C., & Capovilla, A. G. S. (2006). Teste de Competência de Leitura de Palavras e Pseudopalavras: normatização e validação. *Revista de Psicologia da Vetor Editora*, 7(2), 47-59.
- Carraher, T. N. & Rego, L. L. B. (1984). Desenvolvimento cognitivo e alfabetização. *Revista Brasileira de Estudos Pedagógicos*, 65, 38-55.
- Carroll, J. M. & Snowling, M. J. (2004). Language and phonological skills in children at high risk of reading difficulties. *Journal of Child Psychology and Psychiatry*, 45(3), 631-640.

- Carvalho, M. P. (2004a). Quem são os meninos que fracassam na escola? *Cadernos de pesquisa*, 34(1421), 11-40.
- Carvalho, M. P. (2004b). O fracasso escolar de meninos e meninas: articulações entre gênero e cor/raça. *Cadernos Pagu*, 22, 247-290.
- Cascio, W. F. & Aguinis, H. (2005). Test development and use: new twists on old questions. *Human Resource Management*, 44(3), 219-235.
- Chiu, M. M. & McBride-Chang, C. (2006). Gender, Context, and Reading: a Comparison of Students in 43 Countries. *Scientific Studies of Reading*, 10(4), 331-362.
- Cia, F. (2005). O impacto do turno de trabalho do pai no desempenho acadêmico e no autoconceito de crianças escolares. Tese de doutorado não publicada, Universidade Federal de São Carlos, São Carlos, Brasil.
- Cia, F., D´Affonseca, S. M., & Barham, E. J. (2004). A relação entre o envolvimento paterno e o desempenho acadêmico dos filhos. *Cadernos de Psicologia e Educação Paidéia*, 14(29), 277-286.
- Clark, H. (1973). The language-as-fixed-effect fallacy: a critique of language statistics in psychological research. *Journal of verbal learning and verbal behaviour*, 12, 335-359.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Coltheart, M. & Rastle, K. (1994). Serial processing in reading aloud: evidence for dual models of reading. *Journal of Experimental Psychology: Human Perception and Performance*, 20(6), 1197-1211.
- Coltheart, M., Rastle, K. Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204-256.

- Conselho Federal de Psicologia [CFP] (2003). *Resolução 02/2003*. Acesso em 5 de julho, 2007, em <http://www.pol.org.br>
- Crowter-Heyck, H. (1999). George A. Miller, Language, and the computer metaphor of mind. *History of Psychology*, 2(1), 37-64.
- D’Affonseca, S. M. (2005). Prevenindo o fracasso escolar: comparando o autoconceito e o desempenho acadêmico de filhos de mães que trabalham fora e donas de casa. Tese de doutorado não publicada, Universidade Federal de São Carlos, São Carlos, Brasil.
- Dias, T. L. & Enumo, S. R. F. (2004, novembro). *Avaliação dinâmica: Uma proposta alternativa e complementar de avaliação cognitiva em crianças com indicação de dificuldade de aprendizagem*. Trabalho apresentado na 27ª Reunião Anual da Associação Nacional de Pós-Graduação em Educação, Caxambu, MG. Acesso em 24 de outubro, 2007, em <http://www.anped.org.br/reunioes/27/textosgt20.htm>
- Dias, T. L. & Enumo, S. R. F. (2006). Criatividade e dificuldade de aprendizagem: Avaliação com procedimentos tradicional e assistido. *Psicologia: Teoria e Pesquisa*, 22(1), 69-78.
- Dias, T. L., Enumo, S. R. F., & Junior, R. R. A. (2004). Influências de um programa de criatividade no desempenho cognitivo e acadêmico de alunos com dificuldade de aprendizagem. *Psicologia em Estudo, Maringá*, 9(3), 429-437.
- Dias, T. L., Enumo, S. R. F., & Turini, F. A. (2006). Avaliação do desempenho acadêmico de alunos do ensino fundamental em Vitória, Espírito Santo. *Estudos em Psicologia, Campinas*, 24(3), 381-390.
- Duncan, L. G. & Seymour, P. H. K. (2000). Socio-economic differences in foundation-level literacy. *British Journal of Psychology*, 91, 145-166.

- Elias, L. C. S. & Maturano, E. M. (2005). Oficinas de linguagem: proposta de atendimento psicopedagógico para crianças com queixas escolares. *Estudos de Psicologia, 10*(1), 53-61.
- Elliott, C. D. (1996). The British Ability Scales II. Windsor, Berkshire: NFER-NELSON Publishing Company.
- Embretson, S. (1983). Construct Validity: construct representation versus nomothetic Span. *Psychological Bulletin, 92*(1), 179-197.
- Eysenck, M. W. & Keane, M. T. (1994). *Psicologia cognitiva: um manual introdutório* (Wagner Gesser e Maria Helena Fenalti Gesser, trad.). Porto Alegre: Artes Médicas.
- Fonseca, L. C., Tedrus, G. M. A. S., Tonelotto, J. M. F., Antunes, T. A., & Chiodi, M. G. (2004). Desempenho escolar em crianças com epilepsia da infância com pontas centrotemporais. *Arquivos em Neuropsiquiatria, 62*(2-B), 459-462.
- Fox, N. & Mathers, N. (1997). Empowering research: statistical power in general practice research. *Family Practice, 14*, 324-329.
- Gallagher, A., Frith, U., & Snowling, M. (2000). Precursors of literacy delay among children with genetic risk of dyslexia. *Journal of Child Psychology and Psychiatry, 41*(2), 203-213.
- Gardner, H. (2003). *A nova ciência da mente: uma história da revolução cognitiva* (Claudia Malbergier, Trad.). 3. ed., São Paulo: Edusp.
- Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: the neighbourhood frequency effect. *Perception & Psychophysics, 45*, 189-195.
- Godoy, D. M. A. (2005). Aprendizagem inicial da leitura e da escrita no português do Brasil: influência da consciência fonológica e do método de alfabetização. Tese de

- doutorado não publicada, Universidade Federal de Santa Catarina, Florianópolis, Brasil.
- Godoy, D., Defior, S., & Pinheiro, A. M. V. (2007). Impacto do método de alfabetização sobre o desenvolvimento da consciência fonêmica, da leitura e da escrita no português do Brasil. Manuscrito submetido para publicação.
- Goff, D. A., Pratt, C., & Ong, B. (2005). The relations between children's reading comprehension, working memory, language skills and components of reading decoding in a normal sample. *Reading and Writing: an interdisciplinary journal*, 18, 583-616.
- Gilmore, A., Croft, C., & Reid, N.A. (1981). *Burt Word Reading Test New Zealand Revision: Teachers Manual*. Wellington: New Zealand Council for Educational Research.
- Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, 269, 41-45.
- Henderson, L. (1985). On the use of the term 'grapheme'. *Language and Cognition Process*, 1(2), 135-148.
- Hu FB, G. J., Hedeker D., Flay BR, P. M. A. (1998). Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes. *American Journal of Epidemiology*, 147(7), 694-703.
- Justi, F. (2004). O efeito de vizinhança ortográfica no português do Brasil: estudo com a tarefa de decisão lexical. Dissertação de mestrado não-publicada, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
- Justi, F. & Pinheiro, A. M. V. (2006). O efeito de vizinhança ortográfica no português do Brasil: acesso lexical ou processamento estratégico? *Revista Interamericana de Psicología*, 40(3), 257-280.

- Kinoshita, S., Lupker, S. J., & Rastle, K. (2004). Modulation of regularity and lexicality effects in reading aloud. *Memory & Cognition*, 32(8), 1255-1264.
- Kintsch, W. & Greeno, J. G. (1985). Understanding and solving arithmetic problemseg.. *Psychological Review*, 92(2), pp. 109-129.
- Kristensen, C. H., Almeida, R. M. M., & Gomes, W. B. (2001). Desenvolvimento histórico e fundamentos metodológicos da neuropsicologia cognitiva. *Psicologia: Reflexão e Crítica*, 14(2), 259-274.
- Lemle, M. (1981). A tarefa de alfabetização: etapas e problemas no português. *Letras de Hoje*, 15(4), 41-60.
- Lopes, E. J., Lopes, R., & Teixeira, J. F. (2004). A Psicologia Cognitiva experimental cinquenta anos depois: a crise do paradigma do processamento de informação. *Paidéia: Cadernos de Psicologia e Educação – Ribeirão Preto*, 14(27), 17-27.
- Lúcio, P. S., Pinheiro, A. M. V., & Nascimento, E. (2007). O impacto da mudança no critério de acerto na distribuição dos escores do subtteste de leitura do Teste de Desempenho Escolar. Manuscrito submetido para publicação.
- Maia, A. C. B. & Fonseca, M. L. (2002). Quociente de Inteligência e Aquisição de Leitura: Um Estudo Correlacional. *Psicologia: Reflexão e Crítica*, 15(2), 261-270.
- Massaro, D. W. & Cowan, N. (1993). Information processing models: microscopes of the mind. *Annual Review of Psychology*, 44, 383-425.
- Martin, F., & Pratt, C. (2001). The Martin and Pratt nonword reading test. Melbourne: Australian Council for Educational Research Press.
- Matoso Câmara Jr., J. (1985). Estrutura da língua portuguesa. Petrópolis: Vozes.
- McCann, R. & Besner, D. (1987). Reading Pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in

- naming. *Journal of Experimental Psychology: Human Perception and Performance*, 13(1), 14-24.
- Miller, J. (1982). Discrete versus continuous stage models of human information processing: in search of partial output. *Journal of Experimental Psychology: Human Perception and Performance*, 8(2), 273-296.
- Miller, J. (2003). The cognitive revolution: a historical perspective. *TRENDS in Cognitive Science*, 7(3), 141-144.
- Mitchell, T. W. & Klimoski, R. J. (1986). Estimating the Validity of Cross-Validity Estimation. *Journal of Applied Psychology*, 71(2), 311-317.
- Morais, J. (1996). A arte de ler. São Paulo: Editora da Universidade Estadual Paulista.
- Neisser, U. (1967). *Cognitive Psychology*. Englewood Cliffs, N.J.: Prentice-Hall, Inc.
- Noronha, A. P. P., Freitas, F. A., & Ottati, F. (2002). Parâmetros psicométricos de testes psicológicos de inteligência. *Interação em Psicologia*, 6(2), 195-201.
- Nunes, T., Buarque, L. & Bryant, P. (1992). Dificuldades na aprendizagem da leitura: Teoria e prática (Vol. 47). São Paulo: Cortez.
- Nunes, C. H. S. S. & Primi, R. (2005). Impacto do tamanho da amostra na calibração de itens e estimativa de escores por teoria de resposta ao item. *Avaliação Psicológica*, 4(2), 141-153.
- Oakland, T. (1999). Developing Standardized Tests. In S. M. Wechsler & R. S. L. Guzzo (Eds.), *Avaliação Psicológica: Perspectiva Internacional* (pp. 101-118). São Paulo: Casa do Psicólogo.
- Oliveira, E. P., Guerreiro, M. M., Guimarães, C. A., Brandão-Almeida, I. L., Montenegro, M. A., Cendes, F., & Hage, S. R. V. (2005). Caracterização das manifestações lingüísticas de uma família com síndrome perisylviana. *Pró-Fono Revista de Atualização Científica*, 17(3), 393-402.

- Parente, M. A. M. P., Hosogi, M. L., & Lecours, A. R. (1997). Conduta clínica. In A. R. Lecours & M. A. M. Parente (Eds.), *Dislexia: Implicações do Sistema de Escrita do Português* (pp. 86-105). Porto Alegre: Artes Médicas.
- Parente, M. A. M. P., Silveira, A., & Lecours, A. R. (1997). As palavras do português escrito. In A. R. Lecours & M. A. M. Parente (Eds.), *Dislexia: Implicações do Sistema de Escrita do Português* (pp. 41-55). Porto Alegre: Artes Médicas.
- Pasquali, L. (1996). *Teorias e métodos de medida em ciências do comportamento*. Brasília: Laboratório de medida em Ciências do Comportamento. Instituto de Psicologia da Universidade de Brasília [INEP].
- Pasquali, L. (2003). *Psicometria: Teoria dos testes na Psicologia e na Educação*. Petrópolis, RJ: Vozes.
- Pereira, P. C. (2005). Violência doméstica e desempenho escolar: desafios para o judiciário e para a educação. Tese de doutorado não publicada, Universidade Federal de São Carlos, São Carlos, Brasil.
- Perfetti, C. A. (1985). *Reading Ability*. New York Oxford University Press.
- Pinheiro, A. M. V. (1994). *Leitura e Escrita: uma abordagem cognitiva*. 1. ed., Editorial Psy. Campinas.
- Pinheiro, A. M. V. (1995). Reading and spelling development in Brazilian Portuguese. *Reading and Writing: an interdisciplinary journal*, 7, 111-138.
- Pinheiro, A. M. V. (1996). Contagem de frequência de ocorrência de palavras expostas a crianças da 1ª à 4ª série do Ensino Fundamental. São Paulo: Associação Brasileira de Dislexia.
- Pinheiro, A. M. V. (1999). Cognitive assessment of competent and impaired reading in Scottish and Brazilian children. *Reading and Writing: an interdisciplinary journal*, 11, 175-211.

- Pinheiro, A. M. V. (2003). Classificação de regularidade de palavras. Relatório técnico-científico de atividades do CNPq (Ref.: COCHS/DPH 168/03). Belo Horizonte: Universidade Federal de Minas Gerais, Departamento de Psicologia.
- Pinheiro, A. M. V. (2004). Banco de palavras de baixa frequência de ocorrência, para crianças brasileiras da 1^a à 4^a série do Ensino Fundamental, classificadas em termos de estrutura silábica, número de letras e regularidade para leitura e para escrita. Relatório técnico-científico de atividades da FAPEMIG (Ref.: DC/SOT 1202/2004). Belo Horizonte: Universidade Federal de Minas Gerais, Departamento de Psicologia.
- Pinheiro, A. M. V. (2007a). Avaliação de competências psicológicas da população infanto-juvenil de Belo Horizonte: inteligência e habilidades de leitura e escrita. Relatório técnico-científico de atividades da FAPEMIG (Ref.: DC/SOT 1806/2007). Belo Horizonte: Universidade Federal de Minas Gerais, Departamento de Psicologia.
- Pinheiro, Ângela M. V. (2007b). Banco de palavras de baixa frequência de ocorrência, para crianças brasileiras da 1^a à 4^a série do Ensino Fundamental, classificadas em termos de estrutura silábica, número de letras e regularidade para leitura e para escrita. In Sim-Sim, Inês & Vianna, Fernanda L. (2007). Ministério da Educação-Gabinete de Estatística e Planejamento da Educação. Portugal, Lisboa, 2007.
- Pinheiro, A. M. V. (no prelo). *Leitura e Escrita: uma abordagem cognitiva*. 2. ed., Editora Livro Pleno: Campinas.
- Pinheiro, A. M. V. & Costa, A. E. B. (2005, março). Escala de avaliação de competência em leitura pelo professor. Pôster apresentado no VII Encontro Mineiro de Avaliação Psicológica, Belo Horizonte, MG.

- Pinheiro, A. M. V., Lúcio, P. S., & Cunha, C. R. (2008). Tarefa de leitura de palavras em voz alta: uma proposta de análise dos erros. Manuscrito submetido para publicação.
- Pinheiro, A. M. V., Lúcio, P. S., & Silva, D. M. R. (2008). O efeito de regularidade grafema-fonema e fonema-grafema na leitura em voz alta de palavras isoladas. Manuscrito submetido para publicação.
- Pinheiro, A. M. V. & Rothe-Neves. (2001). Avaliação cognitiva da leitura: as tarefas de leitura em voz alta e ditado. *Psicologia: Reflexão e Crítica*, 14(2), 399-408.
- Raaijmakers, J. G. W. (2003) A Further Look at the “Language-as-Fixed-Effect Fallacy”. *Canadian Journal of Experimental Psychology*, 57(3), 141-151.
- Raaijmakers, J. G., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to Deal with “The language-as-Fixed-Effect Fallacy”: Common Misconceptions and Alternative Solutions. *Journal of Memory and Language*, 41, 416–426
- Raaijmakers, J. G., Schrijnemakers, J. M. C., & Gremmen, F. (2006). Corrigendum “How to Deal with “The language-as-Fixed-Effect Fallacy”: Common Misconceptions and Alternative Solutions” [*Journal of Memory and Language*, 41, 416–426]. *Journal of Memory and Language*, 54, 634.
- Rastle, K. & Coltheart, M. (1998). Whammies and double whammies: The effect of length on nonword reading. *Psychonomic Bulletin & Review*, 5(2), 277-282.
- Rastle, K. & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 461-481.
- Ratcliff, R. Methods for Dealing With Reaction Time Outliers. *Psychological Bulletin*, 114(3), 510-532.
- Raven, J., Raven, J. C., & Court, J. H. (1993). Manual das Matrizes Progressivas de Raven: escala especial. Tradução e Adaptação: L. A. Angelini, I. C. B. Alves, E. M.

- Custódio, W. F. Duarte, & J. L. M. Duarte. São Paulo: Centro Editor de Testes e Pesquisas em Psicologia, 1999.
- Reed, S. K. (1982). *Cognition: Theory and applications*. Brooks/ Cole Publishing Company: Monterey, California.
- Rey, A., Ziegler, J. C., & Jacobs, A. M. (2000). Graphemes are Perceptual Reading Units. *Cognition*, 75, B1-B12.
- Rodrigues, C. M. L. & Lopes, E. J. (2002). A psicologia cognitiva no Brasil: Um panorama dos anos 90. *Horizonte Científico – Revista On-line da PROPP-UFU, Uberlândia, 1*, 13p.
- Rosén, M. (2001). Gender differences in reading performance on documents across countries. *Reading and Writing: an interdisciplinary journal*, 14, 1-38.
- Salles, J. F. (2005). Habilidades e dificuldades de leitura e escrita em crianças de 2ª série: abordagem neuropsicológica cognitiva. Tese de doutorado não publicada, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil.
- Salles, J. F. & Parente, M. A. M. P. (2002). Processos Cognitivos na leitura de palavras em crianças: relações com compreensão e tempo de leitura. *Psicologia: Reflexão e Crítica*, 15(2), 321-331.
- Sawson H. L., & Ashbaker, M. H. (2000). Working memory, short-term memory, speech rate, word recognition and reading comprehension in learning disabled readers: does the executive system have a role? *Intelligence*, 28(1), 1-30.
- Sawson, H. L., Howard, C. B., & Saéz, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities? *Journal of Learning Disabilities*, 39(3), May/June, 252-269.
- Schonell, F. & Goodacre, E. (1971). *The Psychology and Teaching of Reading*. London and Edinburgh: Oliver and Boyd.

- Schonell, F. J. & Schonell, P. E. (1960). *Diagnostic and attainment testing*. Edinburgh: Oliver and Boyd.
- Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 876-900.
- Siegel, L. (1992). Na evaluation of the discrepancy definition of dyslexia. *Journal of learning disabilities*, 25(10), 618-629.
- Siegel, L. & Linder, B. (1984). Short-term memory processes in children with reading and arithmetic disabilities. *Developmental Psychology*, 20, 200-207.
- Silva, D. M. R. & Rothe-Neves, R. (2002, novembro). O conceito de foneticidade e a relação grafema-fonema no português brasileiro. Pôster apresentado na X Semana de Iniciação Científica da UFMG, Belo Horizonte, 336-336.
- Silveira R. C. P. (1986). Estudos de fonologia portuguesa. São Paulo: Cortez.
- Sisto, F. F. (2006). O funcionamento diferencial dos itens. *Psico-USF*, 11(1), 35-43.
- Snowling, M. (2004). Dislexia desenvolvimental: uma introdução e visão teórica geral. In M. Snowling & J. Stackhouse (Eds.). *Dislexia, Fala e Linguagem: um manual profissional*. (Tradução Magda França Lopes). Porto Alegre: Artmed. Pp. 11-21.
- Soares, J. F. (2004). O efeito da escola no desempenho cognitivo de seus alunos. *REICE – Revista Eletrônica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 2(2), 83-104.
- Soares, J. F. & Collares, A. C. M. (2006). Recursos familiares e o desempenho cognitivo dos alunos do ensino básico brasileiro. *DADOS – Revista de Ciências Sociais*, 49(3), 615-650.

- Souza, V. P. (1978). Lingüística aplicada à alfabetização. *Ensaio de Lingüística: Cadernos de Lingüística e Teoria da Literatura*, Faculdades de Letras da UFMG, Centro de Extensão, 7, 96-166.
- Sprenger-Charolles, L., Colé, P., Béchenec, D., & Kipffer-Piquard, A. (2005). French normative data on reading and related skills from EVALEC, a new computerized battery of tests (end Grade 1, Grade 2, Grade 3, and Grade 4). *Revue européenne de psychologie appliquée*, 55, 157-186
- Stein, L. M. (1994). TDE: teste de desempenho escolar: manual para aplicação e interpretação. São Paulo: Casa do Psicólogo.
- Sternberg, S. (1969). The discovery of processing stages: extensions of Donders' Method. *Acta Psychol*, 30, 276-315.
- Sucena, A. & Castro, S. L. (2005) Estratégias fonológicas e ortográficas na aprendizagem da leitura do português europeu. *Anales de la Revista de Psicología General y Aplicada*, 10(3). Publicação eletrônica. Acessado em 25 de julho, 2007. Disponível na página <http://www.fedap.es/IberPsicologia/iberpsi10/indiceip10-3.htm>
- Tonelotto, J. M. F., Fonseca, L. C., Tedrus, G. M. S. A., Martins, S. M. V., Gibert, M. A. P., Antunes, T. A., & Pensa, N. A. S. (2005). Avaliação do desempenho escolar e habilidades básicas de leitura em escolares do ensino fundamental. *Avaliação Psicológica*, 4(1), 33-43.
- Toth, G., & Siegel, L. S. (1994). A critical evaluation of the IQ-achievement discrepancy based definition of dyslexia. In K. P. van den Bos e L. S. Siegel (Eds.), *Current directions in dyslexia research*, (pp. 45–70). Lisse, Netherlands: Swets & Zeitlinger.

- Twisk, J. (2004). Longitudinal data analysis: a comparison between generalized estimating equations and random coefficient analysis. *European Journal of Epidemiology*, 19, 769–776
- Urbina, S. (2007). Fundamentos da testagem psicológica. (Cláudia Dornelles, Trad.). Porto Alegre: Artmed.
- Wechsler, D. (1991). WISC-III: Escala de Inteligência Wechsler para Crianças: Manual (3. ed.). Adaptação e Padronização: V. L. M. de Figueiredo. São Paulo: Casa do Psicólogo, 2002.
- White, B. (2007). Are girls better readers than boys? Which boys? Which girls? *Canadian Journal of Education* 30(2), 554-581.
- Wilkinson, G. S. (1993). WRAT 3: Wide Range Achievement Test. Version 3, published by the Psychological Corporation.
- Woodcock, R. W. (1998). Woodcock reading mastery tests (revised). Allen, TX: DLM Teaching Resources.
- Zanoti-Jerônimo, D. V. & Carvalho, A. M. P. (2005). Self-concept, academic performance and behavioral evaluation of the children of alcoholic parents. *Revista Brasileira de Psiquiatria*, 27(3), 233-236.
- Zevin, J. D. & Seidenberg, M. S. (2002) Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language* 47, 1-29.

Apêndice A – Cópia da aprovação do projeto pelo Comitê de Ética em Pesquisas com Seres Humanos (COEP) da UFMG.

Universidade Federal de Minas Gerais
Comitê de Ética em Pesquisa da UFMG - COEP

Parecer nº. ETIC 347/04

Interessada: Profa. Dra. Ângela Maria Vieira Pinheiro
Depto. De Psicologia - FAFICH/UFMG

DECISÃO

O Comitê de Ética em Pesquisa da UFMG – COEP, aprovou no dia 09 de março de 2005, após cumprimento das solicitações de diligência, o projeto de pesquisa intitulado « **Avaliação de competências psicológicas da população infanto-juvenil de Belo Horizonte: inteligência e habilidades de leitura e escrita** » bem como o Termo de Consentimento Livre e Esclarecido do referido projeto.

O relatório final ou parcial deverá ser encaminhado ao COEP um ano após o início do projeto.


Profa. Dra. Maria Elena de Lima Perez Garcia
Presidente do COEP/UFMG

Apêndice B – Modelo de termo de consentimento enviado às escolas



Universidade Federal de Minas Gerais
Faculdade de Filosofia e Ciências Humanas
Departamento de Psicologia

**PEDIDO DE AUTORIZAÇÃO PARA PARTICIPAÇÃO DE ESCOLAS
EM PESQUISA**

Prezado(a) Diretor(a),

A equipe da profa. Ângela Maria Vieira Pinheiro do Departamento de Psicologia da UFMG está conduzindo uma pesquisa intitulada “Avaliação de Competências Psicológicas da População Infanto-Juvenil de Belo Horizonte: Inteligência e Habilidades de Leitura e de Escrita” que tem como principal objetivo a criação de um instrumento de avaliação de leitura e de escrita, para crianças da 1ª à 4ª série do Ensino Fundamental que, além de estabelecer os parâmetros de desenvolvimento normal dessas habilidades, possibilite o diagnóstico das dificuldades de leitura e/ou escrita surgidas dentro de cada faixa etária estudada.

Para conduzir tal pesquisa, a prof. Ângela necessita aplicar o instrumento que desenvolveu em uma amostra representativa de crianças e, para isso, solicitamos sua autorização para que os alunos desta escola possam participar da referida pesquisa. A coleta de dados será feita a partir de agosto de 2006 e consistirá da aplicação de um teste de leitura (em duas sessões, cada uma com duração média de 20 minutos) de inteligência e de desempenho escolar. Cada criança será testada individualmente em seu horário de aula, em ambiente fornecido pela escola.

A professora Ângela se dispõe a dar maiores esclarecimentos sobre eventuais dúvidas que possam surgir a respeito da pesquisa por meio dos telefones (31) 3499-6268 e (31) 9637-8085 ou pelos emails: amvpinheiro@superig.com.br e pslucio@gmail.com. Contamos com a sua colaboração e agradecemos-lhes pela atenção,

Eu, _____, na qualidade de diretor(a) da Escola _____, autorizo a participação de alunos na pesquisa “Avaliação de Competências Psicológicas da População Infanto-Juvenil de Belo Horizonte: Inteligência e Habilidades de Leitura e de Escrita”.

Belo Horizonte, ____ de _____ de 2006.

Assinatura do(a) diretor(a) da escola

Apêndice C – Critérios de participação na pesquisa

Caro professor,

Seguem abaixo os critérios para a participação das crianças de sua série na pesquisa. Eventuais dúvidas serão esclarecidas por email (pslucio@gmail.com) ou pessoalmente. Obrigada,

Patrícia.

1. Todas as crianças participantes devem estar lendo, inclusive as da primeira série;
2. Não é exigido que a criança leia sempre bem, desde que esteja lendo. Assim, entram na pesquisa crianças que lêem bem, mal ou razoavelmente.
3. Não podem participar da pesquisa crianças com suspeita de retardo mental, qualquer tipo de lesão cerebral ou déficit neurológico, ou crianças com problemas comportamentais e de conduta;
4. Para participar da pesquisa, as crianças devem ter entre 7 e 11 anos de idade;
5. As crianças devem estar em idade regular (ou seja, dentro da faixa de idade esperada para a série cursada). Assim, as crianças da primeira série devem ter entre 7 e 8 anos, as da segunda entre 8 e 9 anos, as da terceira entre 9 e 10 anos e as da quarta série entre 10 e 11 anos. Idades fracionadas são aceitas (por exemplo, uma criança da 1ª série que tenha 8 anos e 4 meses pode participar, mas não uma que tenha 9 anos completos).
6. As crianças que caírem dentro desses critérios deverão ser sorteadas a fim de participarem da pesquisa.

Apêndice D – Modelo de termo de consentimento livre e esclarecido entregue aos pais
dos participantes (frente e verso)



Universidade Federal de Minas Gerais
Faculdade de Filosofia e Ciências Humanas
Departamento de Psicologia

PEDIDO DE AUTORIZAÇÃO PARA PARTICIPAÇÃO DE ALUNOS EM PESQUISA

Prezados Pais,

A equipe de pesquisa da professora Ângela Maria Vieira Pinheiro, do Departamento de Psicologia da UFMG, está conduzindo uma pesquisa intitulada “Avaliação de Competências Psicológicas da População Infanto-Juvenil de Belo Horizonte: Inteligência e Habilidades de Leitura e de Escrita” que tem como principal objetivo criar um instrumento de avaliação de leitura para crianças da 1ª à 4ª série do Ensino Fundamental, que, além de estabelecer os parâmetros de desenvolvimento normal dessa habilidade, possibilite o diagnóstico das dificuldades de leitura surgidas dentro de cada faixa etária estudada.

Para conduzir tal pesquisa, a equipe da profa. Ângela necessita aplicar o instrumento que desenvolveu em uma amostra representativa de crianças e, para isso, solicita sua autorização para que seu filho participe da referida pesquisa. A coleta de dados será feita a partir de agosto de 2006 e consistirá na avaliação, em três seções diferentes, da leitura, do desempenho escolar e da inteligência da criança. Cada seção tem uma duração média de 15 minutos e a criança será testada individualmente em seu horário de aula e na sua escola.

À criança é garantido o direito de não participar da pesquisa, se assim o desejar. Além disso, a criança poderá desistir participar da pesquisa em qualquer das seções de teste, sem que isso lhe acarrete prejuízo de qualquer natureza. Não será cobrada nenhuma taxa ou qualquer outro valor aos responsáveis por aqueles que se dispuserem a participar da pesquisa. Os resultados da pesquisa poderão ser publicados em artigos e eventos científicos. Entretanto, os participantes não serão identificados por seus nomes ou quaisquer outros tipos de informações que quebrem sua condição de anônimos. No final da pesquisa, os resultados do desempenho dos participantes serão entregues à escola para consulta dos pais interessados. Os resultados estarão disponíveis a partir do segundo semestre de 2007.

A professora Ângela se dispõe a dar maiores esclarecimentos sobre eventuais dúvidas que possam surgir a respeito da pesquisa por meio dos telefones (31) 3499-6268 – terças e quintas feiras, de 14:30 às 17:00 – e (31) 9637-8085 (falar com Patrícia) ou pelos emails: amvpinheiro@superig.com.br e pslucio@gmail.com. Contamos com a sua colaboração e agradecemos-lhes pela atenção. No entanto, caso você não consinta que seu filho participe desta pesquisa, por favor, devolva este documento em branco para a professora por meio de seu filho.

Autorização:

Eu, _____ autorizo meu
(minha) filho(a) _____ a participar da pesquisa
“Avaliação de Competências Psicológicas da População Infanto-Juvenil de Belo Horizonte:
Inteligência e Habilidades de Leitura e de Escrita”.

Belo Horizonte, ____ de _____ de 2006.

Assinatura (pode ser da mãe, do pai ou responsável)

Para melhor conhecermos o seu filho ou sua filha, por favor, responda o questionário abaixo:

Nome completo da criança: _____ sexo: F M
 Data de Nascimento: __/__/____ Idade: ____ anos Escreve com a mão: direita esquerda
 Já foi reprovado(a) em alguma disciplina? Sim Não
 Se sim: (a) quantas vezes? _____ (b) em qual(is) série(s)? _____ (c) em qual(is) disciplina(s)? _____
 Tem dificuldades com a disciplina “Língua Portuguesa”? Sim Não
 Tem dificuldades com a disciplina “Matemática”? Sim Não
 Seu filho(a) fez pré-escola? Sim Não Com que idade seu filho(a) aprendeu a ler? _____
 Qual a cor de seu filho? Negro Branco Pardo outro

Para que possamos avaliar se o desempenho da criança no teste é semelhante ao desempenho em sala de aula, necessitamos da nota de seu filho nas disciplinas “Língua Portuguesa” e “Matemática”. Marque abaixo a opção “sim”, se você autoriza que a escola forneça a nota média de seu filho nessas disciplinas, e marque a opção “não”, caso contrário: Sim Não

O seguinte questionário será utilizado para verificar se a leitura é influenciada pelo nível sócio-econômico das crianças (fonte: Critério Brasil IBGE (2000), retirado do *site* da ABEP – Associação Brasileira de Empresas de Pesquisa).

Posse de itens: marque sobre o número que representa a quantidade de itens presentes no lar.

Item	Quantidade				
	0	1	2	3	4 ou +
Televisão em cores	0	1	2	3	4 ou +
Banheiro	0	1	2	3	4 ou +
Empregada mensalista	0	1	2	3	4 ou +
Máquina de lavar	0	1	2	3	4 ou +
Rádio	0	1	2	3	4 ou +
Automóvel	0	1	2	3	4 ou +
Aspirador de pó	0	1	2	3	4 ou +
Vídeo cassete e/ou DVD	0	1	2	3	4 ou +
Geladeira	0	1	2	3	4 ou +
Freezer	0	1	2	3	4 ou +

Quem é o chefe da família? Mãe Pai

Marque o grau de instrução do Pai (se houver)

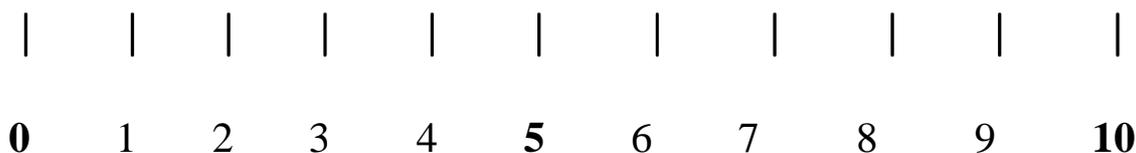
Analfabeto/Primário incompleto:
 Primário completo/Ginasial incompleto:
 Ginásial completo/Colegial incompleto:
 Colegial completo/Superior incompleto:
 Superior completo:

Marque o grau de instrução da mãe (se houver)

Analfabeto/Primário incompleto:
 Primário completo/Ginasial incompleto:
 Ginásial completo/Colegial incompleto:
 Colegial completo/Superior incompleto:
 Superior completo:

Apêndice E – Questionário de auto-avaliação (Modelo de entrevista com as crianças)

1. Você acha que você tem dificuldades com a disciplina “português?”. E “matemática”?
2. Olhe para esta régua. Nela, temos números que vão de zero a 10. Ela contém o valor que representa a nota que você se daria em algumas matérias que você vê na escola. Se você fosse dar uma nota de zero a dez para o seu desempenho em português neste ano, qual seria essa nota? Uma nota zero significa que você acha que você está se saindo muito mal, uma nota dez significa que você está excelente, uma nota cinco significa que você não está nem bem nem mal. Escolha qualquer número entre zero e 10 que mais bem representa a sua nota. Pode ser um, dois, três, quatro, cinco, seis, sete, oito, nove e dez. Quanto mais perto de 10, ou seja, notas maiores do que cinco, melhor o seu desempenho neste ano. Quanto mais perto de zero, ou seja, notas menores do que cinco, pior é o seu desempenho. Seja sincero ao dar sua resposta, pois quero saber como você acha que você está em português este ano.



3. Faça a mesma coisa com seu desempenho em matemática.

Qual você acha que é a cor que você acha que você tem? Você é negro, branco ou pardo?

Apêndice F – Escala de Avaliação de Competência em Leitura pelo Professor para
Crianças de 1ª série (Apêndice F-1) e de 2ª à 4ª séries (Apêndice F-2)

Apêndice F-1: Escala de Avaliação de Competência em Leitura pelo Professor para
Crianças de 1ª série.

Nome do aluno: _____
Série: 1ª idade: ____ anos ____ meses ____ data: ____/____/____
Professora: _____
Escola: _____

Situação de leitura em voz alta

Nº	Critério	SIM	NÃO
1	Demonstra ter entendido o que leu quando indagado sobre o texto lido.		
2	Lê com entonação compatível com a pontuação, expressando emoções e sentimentos de acordo com o texto lido, por exemplo: dá entonação de questionamento em toda a sentença, quando há sinal de interrogação no texto; dá entonação de alegria ou de surpresa em toda a sentença, quando há sinal de exclamação.		
3	Lê muito devagar, sem ritmo, soletrando cada sílaba, não observa a pontuação.		
4	Às vezes, lê e não sabe contar o que leu.		
5	Lê rapidamente as palavras “conhecidas”.		
6	Lê, mas não sabe contar o que leu, nem quando é estimulado com questões.		
7	Lê muito devagar ou muito depressa.		
8	Às vezes comete erros ao ler palavras “novas”.		
9	Coloca entonação do sinal de interrogação e/ou o de exclamação somente na palavra que precede o sinal.		
10	Lê as palavras corretamente.		
11	Não observa a entonação compatível com os sinais de pontuação, fazendo uma leitura num só tom.		
12	Lê com ritmo, nem tão devagar, nem tão rápido.		
13	Freqüentemente comete erros ao ler palavras novas.		

Situação de Leitura Silenciosa

Nº	Critério	SIM	NÃO
01	Nem sempre é capaz de identificar o assunto a partir do título e vice-versa.		
02	Identifica personagens, lugares, mas tem alguma dificuldade para identificar idéias principais.		
03	É capaz de identificar personagens, lugares e idéias principais do texto.		
04	Não identifica o assunto a partir do título e vice-versa.		
05	É capaz de escolher um título para as passagens apresentadas sem título ou mesmo um título alternativo para passagens com título.		
06	Não identifica personagens, lugares ou idéias principais expressos no texto.		
07	Não é capaz de resumir o que leu, seja oralmente ou por escrito.		
08	É capaz de identificar o assunto a partir do título e vice-versa.		
09	É capaz de resumir oralmente o texto lido.		
10	Apresenta certa dificuldade em resumir oralmente o que leu.		

Apêndice F-2: Escala de Avaliação de Competência em Leitura pelo Professor para
Crianças de 2ª à 4ª série.

Nome do aluno: _____

Série: ___ idade: ___ anos ___ meses ___ data: ___/___/___

Professora: _____

Escola: _____

Situação de leitura em voz alta

Nº	Critério	SIM	NÃO
01	Lê, mas não sabe contar o que leu, nem quando estimulado com questões.		
02	Às vezes comete erros ao ler palavras “novas”.		
03	Lê com entonação compatível com a pontuação, expressando emoções e sentimentos de acordo com o texto lido, por exemplo: dá entonação de questionamento, em toda a sentença, quando há sinal de interrogação no texto. Dá entonação de alegria ou de surpresa, em toda a sentença, quando há sinal de exclamação.		
04	Não observa a entonação compatível com os sinais de pontuação, fazendo uma leitura num só tom.		
05	Diz “não sei” quando encontra palavra nova.		
06	Às vezes, lê e não sabe contar o que leu.		
07	Lê rapidamente as palavras “novas” e palavras inventadas.		
08	Lê muito devagar, sem ritmo, soletrando cada sílaba, não observa a pontuação.		
09	Lê soletrando tanto palavras “novas” quanto palavras “conhecidas”.		
10	Coloca a entonação do sinal de interrogação e/ou o de exclamação somente na palavra que precede o sinal de pontuação.		
11	Demora iniciar a leitura quando encontra palavras “novas”, necessitando soletrá-las.		
12	Lê rapidamente palavras “conhecidas” e palavras “pouco conhecidas”.		
13	Freqüentemente comete erros ao ler palavras “novas”.		
14	Demonstra ter entendido o que leu quando indagado sobre o texto lido.		
15	Lê com ritmo, nem tão devagar, nem tão rápido.		
16	Lê muito devagar ou muito depressa.		
17	Lê palavras corretamente.		

Situação de leitura silenciosa

Nº	Critério	SIM	NÃO
01	Identifica personagens, lugares, mas tem alguma dificuldade para identificar idéias principais sem uma segunda leitura.		
02	Não identifica o assunto a partir do título e vice-versa.		
03	É capaz de resumir oralmente o texto lido.		
04	Não identifica personagens, lugares ou idéias principais expressos no texto.		
05	É capaz de escolher um título para passagens apresentadas sem título ou mesmo um título alternativo para passagens com título.		
06	É capaz de identificar o assunto a partir do título e vice-versa.		
07	Apresenta certa dificuldade em resumir oralmente o que leu.		
08	É capaz de identificar personagens, lugares e idéias principais do texto, após a primeira leitura.		
09	Nem sempre é capaz de identificar o assunto a partir do título e vice-versa.		
10	Não é capaz de resumir o que leu, seja oralmente ou por escrito.		

Observação: para todas as séries, ao final da escala, era solicitado à professora a avaliar seu aluno com uma nota de 0 a 10 nas disciplinas “língua portuguesa” e “matemática”, conforme especificado abaixo:

Agora, gostaria que você pensasse no desempenho nas disciplinas “língua Portuguesa” e “matemática” do aluno que você acabou de avaliar. Escreva abaixo uma nota de zero a 10 que melhor represente seu desempenho e suas capacidades nessas duas disciplinas durante o corrente ano letivo. Uma nota de valor 10 significa que o aluno apresenta todas as habilidades e demonstra ter todas as competências referidas na(s) disciplina(s) acima listada(s), em função da sua série ou idade. Uma nota zero implica que o aluno apresenta grandes dificuldades com a(s) disciplina(s), não tendo atingido as competências necessárias para um desempenho esperado de acordo com sua série ou idade. Uma nota cinco representa o aluno mediano (aquele que apresenta algumas dificuldades, mas que não são extremas, ou seja, aquele aluno que não se desempenha nem tão bem, nem tão mal). Qualquer nota entre zero e dez pode ser fornecida, mas não utilize decimal. Obrigada pela atenção.

Nota em português: _____ Nota em matemática: _____

Apêndice G – Valores dos índices de dificuldade, discriminação e correlação item-total apresentados pelas séries e pela amostra geral nas palavras do teste (distribuídas em função da classificação de regularidade para a leitura e para a escrita, do número de letras e da Lista de aplicação)

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T												
RL/RE															
L1/4L															
bala	.91	.20	.39	.94	.14	.17	.96	.15	.32	.95	.15	.20	.94	.15	.29
cuca	.90	.35	.63	.94	.14	.30	.97	.04	.35	.92	.20	.58	.93	.17	.49
mina	.79	.36	.43	.84	.28	.20	.89	.15	.24	.89	.20	.46	.86	.22	.35
tapa	.84	.45	.61	.93	.09	.33	.94	.15	.29	.91	.16	.48	.91	.20	.46
L2/4L															
cava	.90	.25	.46	.93	.17	.55	.99	.00	.02	.97	.00	.32	.95	.15	.44
luta	.94	.20	.64	.96	.09	.36	.98	.08	.09	1.0	.00	.00	.97	.08	.41
suja	.91	.30	.52	.83	.44	.47	.96	.07	.17	.96	.10	.14	.92	.23	.38
vala	.82	.26	.36	.85	.18	.33	.89	.77	.30	.92	.30	.66	.87	.28	.40
L1/5L															
pista	.91	.35	.71	.99	.04	.47	.97	.11	.34	.97	.00	.36	.96	.12	.52
farda	.79	.41	.48	.72	.44	.44	.88	.30	.43	.84	.35	.50	.81	.37	.46
brava	.91	.30	.62	.98	.09	.32	.98	.01	.29	.96	.10	.66	.96	.11	.52
ladra	.86	.36	.60	.86	.35	.45	.91	.19	.37	.93	.20	.62	.89	.28	.50
prata	.87	.26	.53	.90	.30	.51	.98	.07	.20	.93	.25	.67	.92	.22	.53
viúva	.82	.36	.45	.93	.22	.59	.93	.11	.33	.97	.10	.49	.91	.25	.48
L2/5L															
farta	.84	.35	.45	.80	.26	.20	.94	.15	.39	.95	.20	.70	.89	.26	.42
briga	.94	.16	.46	.90	.09	.23	.98	.07	.22	.97	.05	.37	.95	.11	.37
dupla	.86	.45	.69	.91	.26	.61	.95	.07	.11	.97	.10	.45	.92	.22	.54

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T
RL/RE															
L1/6L															
espiga	.81	.41	.52	.86	.30	.43	.88	.30	.19	.85	.35	.49	.85	.29	.42
cocada	.84	.26	.41	.95	.30	.51	.94	.22	.49	.93	.20	.66	.89	.28	.50
maluca	.90	.35	.51	.93	.17	.44	.94	.15	.44	.91	.30	.51	.92	.23	.46
salada	.92	.20	.44	.98	.09	.56	.99	.00	-.06	.95	.15	.56	.96	.08	.44
picada	.83	.40	.49	.91	.22	.35	.96	.04	.02	.91	.35	.69	.91	.27	.45
L2/6L															
antiga	.82	.31	.41	.90	.22	.25	.96	.11	.44	.96	.15	.76	.91	.22	.46
adulta	.73	.61	.56	.77	.65	.65	.88	.33	.39	.91	.30	.62	.82	.51	.58
cabana	.77	.51	.50	.78	.57	.49	.85	.41	.41	.83	.50	.52	.81	.43	.47
colina	.83	.50	.57	.94	.09	.19	.90	.30	.32	.96	.05	.60	.91	.20	.43
danada	.66	.46	.47	.84	.35	.49	.90	.08	.23	.91	.20	.52	.83	.35	.49
sabida	.79	.31	.52	.91	.18	.29	.91	.04	.00	.92	.15	.52	.89	.20	.40
estima	.79	.41	.48	.90	.30	.53	.93	.11	.27	.96	.15	.73	.89	.27	.50
L1/7L															
bengala	.71	.41	.52	.72	.44	.42	.80	.37	.36	.83	.40	.46	.77	.38	.41
revista	.88	.45	.70	.93	.26	.68	.97	.07	.35	.96	.10	.09	.94	.20	.54
batalha*	.75	.46	.42	.85	.30	.37	.92	.26	.55	.96	.15	.68	.87	.32	.49
L2/7L															
ervilha*	.81	.35	.39	.91	.22	.54	.93	.26	.65	.93	.25	.67	.90	.31	.53
redonda	.79	.46	.59	.89	.22	.49	.96	.15	.31	.89	.25	.57	.89	.28	.54
monarca	.45	.39	.27	.46	.62	.41	.53	.60	.42	.69	.55	.47	.53	.58	.39
L1/8L															
feminina	.77	.36	.53	.83	.39	.49	.85	.44	.43	.91	.25	.55	.84	.37	.50

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T												
RL/RE															
L2/8L															
delicada	.61	.51	.40	.68	.36	.25	.83	.30	.25	.84	.45	.60	.74	.51	.41
limonada	.77	.16	.20	.81	.43	.42	.84	.30	.42	.88	.35	.64	.82	.36	.38
RL/RGE															
L1/4L															
cano	.84	.40	.58	.94	.23	.48	.94	.15	.14	.92	.25	.62	.91	.26	.49
fumo	.87	.45	.72	.96	.14	.50	.96	.04	.08	.99	.05	.59	.95	.17	.53
mico	.91	.30	.54	.99	.05	.38	.99	.04	.29	.96	.15	.38	.96	.11	.44
nabo	.82	.55	.61	.89	.26	.41	.90	.11	.32	.85	.45	.47	.87	.31	.48
ramo	.61	.38	.43	.70	.40	.49	.88	.37	.48	.77	.45	.35	.75	.46	.47
L2/4L															
bife	.86	.45	.64	.94	.13	.34	.96	.04	.08	.95	.10	.34	.93	.18	.45
gago	.78	.36	.39	.77	.19	.38	.88	.04	.42	.84	.35	.50	.82	.29	.41
mudo	.81	.50	.59	.83	.35	.43	.91	.19	.23	.88	.20	.29	.86	.32	.43
nave	.83	.21	.35	.95	.09	.26	.93	.11	.21	.92	.25	.64	.91	.17	.37
tubo	.86	.26	.31	.90	.22	.58	.93	.11	.11	.89	.30	.58	.90	.26	.40
L1/5L															
bomba	.88	.21	.35	.90	.22	.51	.97	.07	.35	.92	.15	.54	.92	.17	.44
surdo	.84	.45	.62	.89	.22	.54	.93	.15	.27	.92	.25	.56	.90	.26	.52
tonto	.82	.60	.73	.91	.30	.62	.94	.15	.47	.91	.25	.59	.90	.32	.63
visto	.86	.50	.65	.96	.09	.27	.99	.04	.15	.99	.05	.12	.95	.16	.47
frade	.69	.56	.52	.74	.61	.55	.86	.19	.28	.87	.40	.61	.79	.48	.52
grilo	.88	.30	.45	.93	.13	.36	.96	.07	.19	.91	.20	.54	.92	.17	.41
fiapo	.75	.46	.53	.83	.44	.51	.92	.26	.32	.92	.20	.42	.86	.39	.50

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T
RL/RGE															
L2/5L															
apito	.92	.25	.57	.91	.22	.29	.90	.00	.03	.89	.20	.44	.91	.10	.30
mente	.87	.40	.68	.86	.30	.44	.94	.15	.47	.95	.15	.64	.91	.27	.55
risco	.84	.36	.47	.93	.22	.53	.96	.11	.49	.95	.15	.76	.92	.22	.55
turno	.92	.20	.57	.95	.17	.58	.99	.04	.49	.97	.10	.49	.96	.13	.55
tombo	.82	.41	.58	.86	.44	.55	.88	.26	.30	.85	.45	.62	.86	.37	.51
grade	.58	-.01	.13	.63	.36	.27	.72	.52	.47	.72	.55	.47	.67	.44	.32
trave	.88	.26	.59	.95	.17	.73	.95	.19	.42	.93	.20	.44	.93	.18	.54
fiado	.87	.26	.38	.94	.05	.18	.94	.04	.10	.93	.15	.56	.92	.14	.33
trigo	.86	.45	.61	.91	.05	.17	.96	.00	.02	.93	.15	.58	.92	.18	.42
L1/6L															
afrito	.69	.51	.45	.79	.44	.43	.89	.15	.16	.92	.30	.52	.83	.37	.45
limite	.86	.50	.72	.86	.39	.60	.95	.19	.44	.95	.15	.65	.91	.29	.62
pepino	.74	.66	.61	.83	.30	.43	.89	.23	.41	.92	.20	.53	.85	.37	.53
tucano	.92	.25	.60	.96	.13	.60	.99	.04	.49	.96	.15	.72	.96	.14	.61
ovinho	.81	.55	.63	.80	.44	.38	.85	.79	.18	.88	.30	.53	.83	.32	.44
L2/6L															
cabide	.81	.36	.34	.81	.35	.53	.69	.45	.25	.84	.45	.57	.78	.31	.34
peludo	.81	.41	.51	.88	.30	.47	.98	.74	.18	.95	.15	.61	.91	.28	.52
relato	.78	.41	.56	.81	.48	.60	.93	.11	.05	.93	.25	.71	.87	.33	.53
veludo	.82	.45	.46	.86	.26	.47	.92	.22	.39	.96	.15	.68	.89	.32	.49

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T	<i>p</i>	D	I-T									
RL/RGE															
L1/7L															
estante	.79	.50	.55	.90	.30	.62	.96	.75	.39	.96	.10	.40	.91	.28	.55
combate	.78	.36	.55	.84	.35	.58	.89	.30	.43	.92	.20	.54	.86	.37	.54
covarde	.65	.66	.57	.79	.48	.59	.88	.26	.36	.84	.36	.58	.80	.46	.57
nublado	.58	.81	.57	.81	.39	.50	.83	.37	.42	.84	.45	.56	.77	.51	.55
L2/7L															
espanto	.78	.70	.60	.85	.39	.56	.91	.26	.49	.93	.20	.63	.87	.38	.62
campina	.68	.22	.27	.70	.27	.35	.72	.34	.29	.80	.50	.48	.72	.32	.33
destino	.83	.31	.41	.90	.30	.47	.93	.15	.41	.97	.10	.80	.91	.25	.49
ditongo	.57	.72	.52	.57	.45	.41	.83	.26	.35	.88	.16	.42	.72	.46	.48
retalho*	.74	.60	.55	.77	.44	.52	.84	.19	.21	.89	.25	.54	.81	.38	.48
L1/8L															
mosquito	.90	.35	.65	.96	.13	.73	.99	.04	.29	.96	.10	.62	.95	.15	.63
RL/IRE															
L1/4L															
raça	.88	.25	.46	.90	.30	.50	.96	.11	.39	.97	.05	.054	.93	.20	.41
L1/5L															
salsa	.70	.42	.43	.85	.30	.48	.86	.41	.42	.93	.25	.34	.84	.35	.45
L2/5L															
talco	.84	.46	.62	.90	.18	.39	.95	.15	.39	.95	.15	.38	.91	.21	.49
L1/6L															
buzina	.88	.40	.58	.94	.17	.47	.96	.11	.34	.96	.10	.66	.94	.19	.53
caçula	.62	.66	.52	.75	.57	.61	.79	.37	.47	.80	.60	.60	.74	.58	.55
cigana*	.66	.36	.43	.73	.52	.52	.77	.41	.35	.83	.50	.58	.75	.52	.57
gemada	.73	.46	.41	.80	.48	.57	.94	.19	.44	.92	.25	.65	.85	.42	.54

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T
RL/IRE															
L2/6L															
caçada	.66	.27	.35	.75	.23	.30	.77	.23	.20	.88	.30	.55	.77	.32	.36
cevada*	.56	.61	.46	.72	.44	.44	.74	.37	.32	.72	.60	.52	.69	.50	.45
cilada	.82	.45	.46	.83	.26	.35	.94	.04	.20	.88	.30	.50	.87	.25	.41
concha	.66	.51	.54	.77	.48	.53	.84	.26	.23	.77	.40	.44	.77	.44	.47
L1/7L															
gengiva	.49	.67	.52	.54	.61	.43	.75	.45	.41	.77	.45	.54	.65	.67	.51
L2/7L															
mochila	.79	.36	.29	.86	.39	.67	.95	.15	.24	.97	.10	.50	.90	.28	.46
RL/IRRGE															
L1/4L															
laje	.79	.31	.38	.93	.13	.11	.93	.15	.25	.93	.20	.60	.90	.21	.37
maço	.78	.41	.44	.83	.18	.31	.88	.23	.35	.89	.40	.64	.85	.35	.43
L2/4L															
gaze	.81	.55	.63	.78	.31	.48	.88	.15	.29	.89	.35	.61	.84	.34	.51
sino	.99	.01	.39	.98	.04	.06	1.0	.00	.00	1.0	.00	.00	.99	.00	.18
L1/5L															
haste	.65	.33	.42	.77	.35	.42	.80	.30	.44	.77	.31	.28	.75	.38	.40
lenço	.84	.45	.57	.94	.13	.46	.99	.00	-.06	.95	.15	.36	.93	.19	.50
traje	.81	.41	.48	.81	.39	.44	.97	.04	.13	.89	.40	.64	.88	.32	.48
L2/5L															
balde	.81	.32	.54	.91	.22	.35	.94	.15	.40	.92	.30	.65	.90	.29	.51
salto	.88	.40	.65	.93	.17	.57	1.0	.00	.00	.97	.10	.80	.95	.18	.62
ciclo	.83	.36	.51	.86	.39	.50	.97	.07	.20	.92	.20	.51	.90	.26	.49
guizo	.40	.49	.30	.49	.83	.51	.60	.34	.26	.71	.53	.42	.55	.62	.40

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T												
RL/IRGE															
L1/6L															
gozado	.78	.60	.57	.83	.52	.62	.91	.15	.24	.91	.30	.59	.86	.35	.54
ensino	.83	.36	.54	.93	.09	.47	.96	.07	.17	.95	.20	.71	.92	.20	.52
L2/6L															
regime	.79	.41	.47	.90	.30	.52	.94	.19	.28	.96	.10	.65	.90	.32	.51
filtro	.71	.66	.64	.88	.39	.69	.91	.30	.55	.96	.15	.75	.87	.41	.66
L1/7L															
recurso	.65	.76	.64	.79	.65	.70	.85	.23	.32	.95	.10	.57	.81	.48	.60
humilde	.45	.72	.50	.63	.27	.31	.54	.27	.25	.75	.55	.43	.59	.46	.39
L2/7L															
cimento	.90	.35	.65	.90	.26	.51	.95	.08	.33	.99	.05	.10	.93	.21	.49
saudade	.51	.52	.40	.65	.27	.41	.75	.41	.38	.88	.30	.30	.70	.46	.43
L2/8L															
preguiça	.66	.61	.58	.84	.55	.61	.91	.22	.43	.84	.45	.53	.82	.51	.58
IRL/RE															
L1/4L															
boba	.60	.88	.53	.52	.49	.34	.70	.31	.27	.61	.21	.23	.61	.35	.35
cota	.77	.36	.36	.80	.39	.45	.87	.30	.43	.92	.10	.24	.84	.33	.39
gola	.81	.41	.42	.90	.13	.37	.94	.19	.28	.95	.20	.43	.90	.23	.42
soja	.82	.26	.38	.90	.13	.20	.99	.00	.03	.93	.10	.04	.92	.15	.29
sopa	.79	.41	.41	.84	.44	.63	.93	.11	.10	.95	.20	.49	.88	.32	.46
vela	.88	.40	.49	.90	.13	.36	.97	.04	.14	.96	.15	.67	.93	.18	.45

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T
IRL/RE															
L2/4L															
copa	.91	.25	.54	.93	.09	.29	.93	.00	.03	.96	.15	.71	.93	.15	.38
cova	.65	.28	.33	.72	.22	.25	.88	.11	.03	.91	.25	.55	.79	.31	.35
moda	.91	.16	.50	.96	.13	.57	.96	.07	.05	.95	.20	.67	.95	.14	.46
sola	.92	.20	.40	.90	.22	.25	.98	.00	.00	.95	.20	.43	.94	.16	.32
tela	.90	.26	.49	.94	.17	.57	.97	.04	.15	.97	.10	.19	.95	.15	.43
tola	.17	.27	.22	.19	-.05	.07	.18	.28	.23	.27	.47	.26	.20	.21	.19
L1/5L															
besta	.57	.76	.51	.95	.32	.37	.85	.30	.44	.89	.15	.45	.75	.49	.50
morta	.79	.55	.54	.85	.44	.67	.96	.04	.12	.96	.10	.66	.89	.31	.57
testa	.78	.46	.49	.79	.22	.28	.85	.15	.25	.91	.10	.51	.83	.27	.39
trena*	.78	.51	.60	.93	.22	.46	.87	.22	.31	.92	.20	.55	.87	.28	.49
negra	.71	.70	.60	.83	.39	.40	.91	.11	.23	.88	.20	.47	.84	.40	.49
moeda	.91	.35	.72	.94	.13	.44	.92	.22	.41	.95	.20	.65	.93	.20	.54
zebra	.91	.30	.41	.95	.09	.24	.98	.07	.38	.97	.10	.80	.95	.14	.44
L2/5L															
forca	.26	.12	.16	.43	.45	.38	.45	.17	.01	.49	.32	.21	.41	.29	.24
resta	.78	.41	.57	.88	.26	.45	.88	.19	.40	.88	.26	.33	.86	.31	.46
torta	.86	.26	.39	.90	.35	.40	.93	.22	.27	.93	.15	.55	.91	.23	.41
frota	.71	.46	.46	.90	.39	.46	.93	.19	.14	.88	.30	.50	.84	.38	.47
sobra	.78	.12	.24	.81	.27	.29	.80	.33	.35	.76	-.04	.01	.75	.14	.21
viola	.86	.31	.51	.65	.17	.45	.93	.15	.24	.96	.15	.49	.92	.21	.44

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T
IRL/RE															
L1/6L															
cadela	.77	.41	.52	.78	.19	.46	.79	.59	.51	.83	.40	.32	.79	.58	.43
canela	.71	.46	.48	.77	.65	.59	.85	.19	.18	.81	.55	.60	.79	.45	.49
capela	.75	.40	.48	.80	.44	.60	.87	.33	.50	.87	.40	.57	.83	.43	.53
gaveta	.83	.40	.61	.81	.27	.40	.92	.19	.41	.93	.20	.63	.88	.31	.51
maleta	.70	.56	.48	.68	.48	.48	.87	.19	.22	.80	.40	.47	.77	.39	.45
peteca	.75	.56	.52	.80	.52	.64	.92	.15	.22	.91	.10	.55	.85	.39	.54
sacola	.92	.30	.61	.94	.09	.27	1.0	.00	.00	.97	.00	.80	.96	.12	.51
atleta	.62	.81	.59	.74	.52	.44	.76	.26	.36	.81	.35	.35	.74	.50	.46
sapeca	.78	.36	.46	.83	.39	.55	.93	.26	.46	.93	.25	.70	.87	.36	.55
mestra	.66	.31	.38	.65	.31	.29	.82	.37	.38	.83	.35	.44	.74	.41	.39
cebola*	.88	.40	.56	.90	.30	.67	1.0	.00	.00	.97	.10	.22	.94	.19	.52
L2/6L															
caneca	.51	.62	.41	.60	.31	.30	.68	.41	.31	.73	.40	.39	.63	.49	.38
caneta	.82	.35	.38	.80	.39	.38	.89	.15	.13	.87	.25	.34	.85	.29	.32
cometa	.52	.62	.45	.65	.53	.38	.79	.41	.37	.71	.75	.56	.68	.63	.47
favela	.69	.75	.63	.75	.57	.57	.90	.30	.53	.93	.25	.49	.82	.49	.60
novela	.90	.35	.65	.94	.22	.67	.99	.00	-.04	.95	.00	.21	.95	.13	.49
panela	.90	.11	.32	.90	.13	.22	.96	.04	.25	.92	.15	.38	.92	.12	.30
pipoca	.97	.10	.38	.98	.09	.48	.99	-.04	-.11	.95	.15	.63	.97	.07	.37
aposta	.73	.56	.61	.90	.26	.47	.89	-.03	.01	.92	.25	.64	.86	.31	.49
L1/7L															
cartela	.68	.56	.45	.79	.39	.42	.86	.26	.36	.88	.35	.39	.80	.42	.45
coberta	.70	.66	.61	.89	.22	.41	.93	.19	.45	.88	.30	.49	.86	.36	.54

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T												
IRL/RE															
L2/7L															
bandeja	.77	.36	.31	.86	.44	.71	.90	.19	.39	.89	.40	.71	.86	.37	.51
cartola	.73	.61	.55	.83	.44	.55	.90	.15	.40	.91	.20	.49	.84	.38	.53
moderna	.69	.52	.49	.83	.44	.54	.84	.52	.49	.93	.25	.67	.82	.43	.54
carreta	.58	.57	.43	.63	.23	.27	.74	.12	.15	.75	.26	.31	.68	.32	.33
IRL/RGE															
L1/4L															
colo	.61	.46	.42	.64	.57	.44	.84	.33	.28	.83	.40	.52	.74	.51	.46
foco	.39	.30	.30	.47	.15	.14	.57	.35	.29	.57	.27	.31	.50	.36	.28
mole	.81	.55	.62	.93	.13	.53	.96	.11	.12	.95	.15	.40	.91	.26	.52
nono*	.77	.27	.41	.79	.48	.48	.89	.11	.24	.88	.25	.50	.83	.32	.43
teto	.70	.27	.17	.79	.35	.49	.81	.33	.23	.84	.30	.43	.79	.36	.33
L2/4L															
bote	.74	.42	.44	.88	.05	.15	.84	.30	.47	.91	.20	.52	.84	.29	.40
fofo	.56	.42	.36	.60	.66	.54	.81	.44	.43	.79	.45	.47	.70	.52	.48
moto	.86	.01	.31	.81	.27	.26	.91	.11	.29	.83	.11	.17	.86	.15	.24
neto	.83	.26	.31	.90	.13	.41	.95	.15	.32	.93	.20	.60	.91	.22	.42
voto	.69	.41	.40	.78	.18	.21	.91	.15	.17	.84	.25	.28	.81	.31	.33

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T
IRL/RGE															
L1/5L															
sorte	.87	.25	.52	.95	.17	.66	.97	.04	.32	.97	.10	.48	.94	.17	.53
teste	.71	.51	.56	.86	.26	.42	.90	.26	.47	.91	.25	.30	.85	.33	.49
bloco	.78	.46	.42	.85	.18	.40	.96	.07	.15	.97	.10	.53	.89	.23	.49
prego	.78	.31	.52	.90	.26	.55	.83	.34	.40	.88	.35	.40	.85	.31	.42
veste	.62	.52	.52	.79	.57	.39	.89	.22	.31	.89	.25	.45	.80	.45	.49
afeto	.60	.42	.39	.69	.27	.27	.87	.30	.34	.85	.25	.37	.76	.41	.40
queda	.71	.36	.38	.89	.22	.49	.90	.22	.37	.89	.25	.58	.85	.32	.47
L2/5L															
forno	.81	.46	.51	.86	.26	.51	.91	.26	.39	.93	.25	.43	.88	.27	.48
poste	.68	.32	.36	.79	.26	.36	.91	.08	.18	.84	.30	.43	.81	.29	.39
brejo	.56	.19	.19	.59	.40	.30	.71	.04	.30	.68	.41	.36	.64	.41	.29
globo	.84	.36	.54	.95	.09	.27	.96	.11	.16	.96	.15	.49	.93	.19	.44
trevo	.69	.71	.54	.85	.31	.26	.84	.15	.27	.91	.25	.39	.85	.39	.46
febre	.88	.26	.44	.89	.13	.20	.87	.19	.17	.93	.25	.42	.89	.22	.28
miolo	.79	.55	.54	.79	.52	.61	.87	.30	.37	.91	.35	.52	.84	.38	.51
leque	.78	.41	.46	.90	.26	.44	.87	.15	.26	.89	.25	.39	.86	.32	.40
L1/6L															
ampola	.16	-.02	.01	.16	.25	.21	.14	.03	.04	.25	.28	.18	.17	.13	.11
bigode	.82	.55	.67	.89	.26	.58	.91	.26	.42	.88	.40	.63	.88	.32	.58
camelo	.81	.26	.36	.91	.09	.58	.94	.08	.13	.87	.30	.31	.89	.16	.31
tijolo	.88	.26	.33	.91	.22	.29	.99	.04	.06	.95	.10	.40	.94	.18	.42

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T												
IRL/RGE															
L2/6L															
estojo	.83	.50	.62	.84	.31	.56	.86	.11	.27	.93	.20	.63	.89	.28	.57
colete	.45	.67	.36	.57	.49	.41	.67	.48	.25	.69	.50	.45	.60	.60	.40
novelo	.47	.44	.35	.60	.49	.40	.54	.56	.39	.57	.66	.50	.55	.55	.39
tapete	.74	.55	.53	.86	.35	.46	.95	.19	.56	.91	.30	.58	.87	.38	.56
cabelo	.87	.35	.51	.98	.04	.08	.97	.11	.45	.93	.15	.54	.94	.17	.43
L1/7L															
martelo	.71	.51	.52	.93	.13	.42	.92	.26	.50	.88	.35	.46	.86	.32	.51
materno	.69	.51	.51	.70	.36	.41	.74	.30	.27	.83	.30	.47	.74	.40	.42
repolho*	.82	.46	.51	.90	.30	.64	.94	.19	.51	.93	.25	.61	.80	.30	.58
marreco	.65	.56	.49	.68	.57	.46	.83	.52	.57	.87	.45	.67	.76	.57	.55
repleto	.77	.65	.56	.75	.35	.37	.85	.22	.30	.91	.25	.57	.82	.37	.46
L2/7L															
catorze	.43	.49	.42	.47	.66	.51	.47	.38	.19	.55	.51	.35	.48	.52	.37
socorro	.78	.65	.62	.90	.35	.69	.96	.15	.45	.96	.15	.75	.90	.34	.65
bodoque	.42	.63	.47	.60	.49	.43	.56	.60	.35	.65	.61	.45	.56	.63	.44
L1/8L															
sabonete	.84	.36	.54	.90	.26	.55	.94	.15	.26	.95	.15	.40	.91	.27	.49
uniforme	.77	.65	.64	.78	.57	.58	.94	.15	.27	.92	.05	.25	.86	.37	.52
esquerdo	.64	.71	.51	.80	.61	.67	.85	.33	.38	.88	.40	.61	.80	.51	.57
gabinete	.48	.44	.36	.68	.44	.52	.73	.60	.54	.85	.35	.50	.69	.49	.50

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			2ª série			3ª série			Total		
	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T
IRL/RGE															
L2/8L															
rabanete	.52	.61	.46	.65	.66	.45	.69	.41	.33	.73	.40	.37	.65	.52	.43
basquete	.81	.41	.49	.91	.30	.62	.97	.04	.33	.95	.20	.71	.91	.29	.57
IRL/IRE															
L1/4L															
coxa	.77	.46	.42	.88	.26	.43	.88	.19	.25	.92	.15	.31	.86	.31	.39
gema*	.88	.21	.40	.91	.17	.40	.97	.04	.086	.97	.10	.43	.94	.14	.38
roça	.75	.41	.43	.88	.26	.34	.88	.30	.41	.91	.25	.53	.86	.30	.44
L2/4L															
roxa	.39	.29	.21	.57	.32	.27	.51	.24	.13	.52	.61	.41	.50	.36	.26
cela*	.82	.46	.44	.85	.13	.17	.95	.11	.15	.92	.20	.21	.89	.16	.31
seda	.74	.66	.57	.84	.31	.36	.92	.22	.26	.97	.05	.40	.87	.32	.48
L1/5L															
selva	.74	.60	.49	.86	.35	.60	.93	.15	.42	.92	.25	.62	.87	.39	.56
sexta	.79	.45	.35	.77	.44	.40	.91	.19	.29	.88	.20	.32	.84	.32	.36
graxa	.79	.35	.38	.85	.39	.44	.90	.26	.38	.84	.45	.34	.85	.38	.39
L2/5L															
relva	.56	.67	.49	.74	.27	.20	.75	.30	.31	.73	.40	.33	.70	.37	.37
bolsa	.73	.41	.51	.96	.26	.59	.97	.11	.25	.96	.10	.41	.90	.28	.53
magra	.79	.26	.25	.90	.26	.27	.97	.07	.32	.93	.15	.31	.90	.19	.32
horta	.77	.70	.72	.88	.30	.54	.91	.26	.40	.92	.25	.37	.87	.39	.57
L1/6L															
tigela	.88	.30	.54	.94	.13	.30	.94	.11	.27	.95	.20	.683	.93	.18	.45
faxina	.81	.40	.42	.84	.35	.42	.83	.52	.64	.85	.30	.34	.83	.39	.42
ameixa	.90	.40	.66	.89	.26	.51	.98	.00	.05	.93	.10	.50	.93	.17	.51

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			2ª série			3ª série			Total		
	<i>p</i>	D	I-T												
IRL/IRE															
L2/6L															
paçoca	.69	.27	.34	.77	.27	.36	.75	.30	.38	.84	.40	.51	.76	.34	.38
bexiga	.62	.71	.57	.73	.57	.51	.89	.30	.44	.89	.35	.67	.79	.53	.59
L1/7L															
cerveja*	.84	.45	.59	.88	.30	.48	.94	.11	.17	.93	.10	.50	.90	.24	.48
sarjeta	.32	.53	.30	.42	.52	.35	.54	.45	.33	.67	.65	.42	.49	.57	.38
carroça	.79	.40	.42	.81	.31	.47	.91	.26	.50	.81	.50	.57	.84	.38	.47
L2/7L															
gorjeta	.61	.46	.41	.62	.61	.44	.73	.34	.22	.83	.35	.47	.70	.44	.41
certeza*	.70	.56	.59	.88	.13	.41	.95	.19	.42	.96	.10	.63	.88	.28	.56
palhoça*	.29	.40	.32	.35	.54	.41	.45	.45	.29	.51	.61	.41	.40	.55	.38
IRL/IRRGE															
L1/4L															
xale*	.62	.37	.40	.73	.23	.15	.75	.48	.39	.77	.16	.33	.72	.32	.33
cego*	.68	.36	.36	.77	.39	.45	.86	.08	.02	.85	.40	.61	.79	.41	.40
L2/4L															
luxo	.86	.26	.30	.80	.01	.12	.86	.11	.14	.89	.20	.43	.85	.13	.23
moço	.79	.40	.53	.88	.26	.51	.93	.19	.45	.93	.25	.62	.89	.29	.55
boxe	.13	.18	.14	.20	.05	.35	.28	.61	.45	.39	.27	.24	.25	.48	.31
L1/5L															
polvo	.82	.41	.58	.89	.39	.62	.95	.08	.16	.87	.40	.60	.89	.28	.53
terço	.73	.65	.59	.80	.48	.56	.95	.11	.41	.92	.25	.64	.86	.42	.60
L2/5L															
cesto*	.55	.66	.51	.68	.52	.54	.77	.56	.52	.76	.36	.36	.69	.57	.51
treze	.65	.42	.37	.68	.35	.26	.72	.45	.40	.81	.55	.54	.71	.46	.38

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			2ª série			3ª série			Total		
	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T
IRL/IRRGE															
L1/6L															
enxada*	.75	.51	.51	.75	.39	.49	.83	.22	.29	.83	.45	.50	.79	.40	.45
colcha	.55	.76	.51	.77	.40	.47	.83	.08	.15	.81	.30	.42	.74	.48	.46
exceto	.19	.13	.14	.21	.07	.10	.41	.57	.39	.72	.55	.53	.38	.52	.34
L2/6L															
enxame*	.64	.09	.27	.62	-.07	.00	.66	.12	.18	.67	.22	.25	.65	.11	.17
azeite	.81	.50	.59	.84	.30	.46	.97	.07	.25	.97	.10	.80	.90	.28	.55
choque	.82	.45	.53	.90	.22	.46	.97	.04	.08	.93	.10	.29	.91	.22	.44
L1/7L															
exposto	.40	.68	.45	.59	.53	.39	.75	.41	.43	.81	.36	.47	.65	.64	.49
excesso	.51	.67	.47	.57	.61	.43	.69	.41	.40	.73	.70	.56	.63	.59	.48
celeste*	.62	.66	.55	.69	.57	.57	.82	.30	.30	.82	.45	.60	.75	.54	.54
bezerro	.60	.76	.54	.69	.57	.53	.82	.48	.59	.87	.35	.41	.74	.56	.54
L2/7L															
caxumba	.65	.66	.59	.81	.44	.42	.86	.37	.43	.71	.45	.65	.80	.53	.55
externo	.38	.25	.24	.53	.44	.32	.74	.67	.58	.71	.65	.48	.60	.58	.43
enxoval*	.52	.52	.37	.67	.70	.42	.76	.59	.53	.81	.65	.63	.69	.69	.50
honesto	.77	.55	.60	.85	.26	.35	.90	.22	.25	.91	.35	.69	.86	.36	.51
serrote	.57	.71	.54	.73	.44	.55	.78	.41	.33	.88	.40	.56	.74	.54	.53
secreto	.68	.32	.44	.75	.48	.51	.80	.44	.39	.83	.45	.54	.77	.46	.47
reflexo	.52	.855	.54	.69	.70	.62	.86	.37	.46	.93	.25	.65	.76	.67	.61
L2/7L															
universo	.70	.51	.48	.77	.52	.59	.90	.26	.34	.85	.45	.45	.81	.46	.50
saxofone	.26	.54	.38	.40	.66	.48	.64	.52	.42	.68	.60	.44	.50	.70	.48

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T
RGL/RE															
L1/4L															
caju	.94	.25	.63	.99	.029	-.01	.97	.07	.17	.95	.15	.33	.96	.09	.34
vara	.83	.16	.35	.93	.09	.25	.95	.19	.45	.93	.25	.41	.91	.21	.38
L2/4L															
dura	.92	.20	.55	.94	.09	.38	.98	.07	.25	1.0	.00	.00	.96	.11	.41
zebu	.39	.57	.27	.46	.45	.33	.59	.41	.24	.57	.51	.41	.51	.53	.33
L1/6L															
coruja	.84	.35	.51	.90	.17	.49	.96	.11	.35	.97	.10	.53	.92	.22	.50
morena*	.68	.46	.39	.83	.30	.54	.89	.22	.21	.89	.35	.42	.83	.41	.45
parada	.84	.41	.50	.93	.17	.57	.91	.33	.48	.96	.05	.14	.91	.23	.45
segura	.78	.60	.54	.90	.30	.61	.95	.00	.06	.92	.20	.60	.89	.29	.53
L2/6L															
barata	.90	.21	.35	.94	.09	.30	.95	.07	.24	.92	.20	.56	.93	.15	.36
morada	.79	.46	.51	.85	.22	.36	.84	.22	.31	.95	.20	.62	.86	.28	.44
marido	.87	.31	.46	.90	.26	.34	.98	.00	-.06	.97	.10	.80	.93	.18	.44
camisa	.94	.10	.23	.95	.17	.60	1.0	.00	.00	.97	.05	.39	.97	.01	.38
RGL/RGE															
L1/4L															
caro	.52	.52	.38	.62	.53	.45	.68	.26	.21	.80	.11	.14	.65	.45	.35
L2/4L															
rasa	.44	.44	.26	.56	.32	.23	.65	.38	.31	.63	.46	.37	.57	.44	.32
L1/5L															
caqui	.71	.46	.43	.86	.30	.54	.95	.19	.43	.96	.10	.44	.88	.33	.51

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T												
RGL/RGE															
L2/7L															
risonho*	.74	.50	.49	.83	.39	.49	.92	.11	.25	.96	.05	.27	.86	.31	.46
RG//IRE															
L2/5L															
brasa	.73	.36	.37	.77	.57	.56	.92	.11	.08	.91	.25	.58	.83	.37	.45
L1/6L															
chuchu	.81	.40	.45	.89	.18	.52	.95	.07	.06	.93	.15	.30	.90	.20	.42
L1/7L															
espiral	.74	.51	.45	.78	.44	.53	.81	.33	.43	.88	.30	.42	.80	.45	.46
vasilha*	.62	.66	.55	.74	.35	.48	.87	.33	.58	.85	.40	.57	.78	.54	.57
IRRGL/RE															
L2/4L															
fera	.70	.22	.26	.84	.13	.13	.85	.26	.26	.96	.15	.16	.84	.25	.26
L2/5L															
amora	.83	.26	.52	.86	.18	.34	.88	.33	.44	.92	.15	.31	.87	.28	.41
L1/6L															
espera	.86	.41	.60	.89	.26	.54	.94	.15	.40	.93	.25	.72	.91	.26	.57
sonora	.53	.29	.32	.73	.52	.40	.77	.37	.33	.77	.55	.56	.71	.49	.42
L2/6L															
enxuto*	.49	.34	.28	.62	.40	.28	.67	.26	.20	.79	.16	.26	.64	.33	.30
orelha*	.78	.40	.37	.93	.26	.47	.93	.22	.35	.93	.25	.64	.89	.31	.46
direta	.57	.28	.26	.68	.40	.33	.74	.38	.39	.79	.31	.36	.70	.38	.35
L2/7L															
cantora	.68	.27	.39	.79	.39	.49	.85	.26	.33	.85	.40	.50	.80	.36	.45

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T												
IRRGL/RE															
L1/8L															
diretora	.69	.61	.52	.84	.39	.49	.84	.41	.53	.92	.25	.60	.82	.39	.54
IRRGL/RGE															
L1/4L															
soro	.49	.43	.33	.69	.23	.21	.79	.37	.31	.72	.16	.27	.68	.35	.33
L2/6L															
maroto	.45	.48	.33	.53	.45	.39	.64	.41	.27	.65	.65	.50	.57	.54	.39
IRRGL/IRE															
L1/4L															
cera*	.39	.53	.29	.53	.36	.24	.60	.52	.41	.71	.60	.36	.56	.61	.35
L1/5L															
presa	.58	.38	.35	.63	.31	.26	.64	.45	.25	.77	.40	.37	.65	.43	.32
L1/6L															
cereja*	.66	.66	.49	.84	.44	.64	.81	.30	.35	.88	.20	.23	.80	.39	.46
garota	.79	.46	.52	.89	.26	.39	.95	.11	.31	.88	.30	.50	.88	.26	.47
L2/6L															
defesa	.62	.56	.45	.75	.52	.50	.86	.33	.45	.87	.45	.64	.78	.50	.53
careta	.55	.33	.34	.48	.19	.23	.63	.34	.27	.72	.16	.27	.59	.29	.29
L1/7L															
bondosa	.52	.44	.40	.69	.52	.54	.84	.26	.30	.77	.50	.47	.71	.50	.50
formosa	.60	.33	.41	.69	.23	.22	.82	.26	.29	.77	.31	.37	.73	.36	.36
medrosa	.78	.35	.48	.89	.35	.65	.94	.11	.40	.89	.20	.54	.88	.26	.54
vaidosa	.75	.55	.52	.84	.39	.60	.89	.30	.48	.92	.25	.63	.85	.39	.55

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T												
IRRGL/IRE															
L2/7L															
despesa	.39	.25	.27	.48	.53	.40	.89	.38	.26	.65	.51	.35	.52	.47	.34
nervosa	.66	.61	.54	.73	.57	.56	.55	.30	.56	.91	.25	.59	.80	.52	.59
doutora	.68	.76	.56	.62	.52	.58	.81	.48	.48	.88	.40	.64	.75	.62	.56
L1/8L															
japonesa	.75	.46	.56	.88	.30	.40	.94	.22	.41	.95	.20	.68	.88	.33	.54
L2/8L															
luminosa	.62	.51	.51	.70	.48	.41	.81	.52	.44	.88	.25	.49	.76	.51	.49
mariposa	.62	.62	.46	.78	.52	.60	.82	.41	.47	.84	.60	.68	.77	.56	.56
IRRGL/IRRGE															
L1/5L															
aceso	.21	.41	.29	.26	.25	.14	.18	.36	.33	.33	.18	.11	.24	.28	.21
L1/6L															
caroço	.53	.76	.51	.70	.52	.52	.76	.45	.35	.81	.45	.55	.71	.60	.51
L2/6L															
famoso	.69	.51	.50	.77	.35	.50	.92	.22	.34	.85	.40	.59	.81	.46	.52
L1/7L															
teimoso	.64	.71	.58	.88	.39	.54	.86	.37	.57	.85	.35	.51	.81	.49	.57
dengoso	.75	.75	.67	.85	.22	.38	.83	.33	.30	.88	.40	.59	.83	.41	.50
manhoso	.61	.71	.53	.69	.53	.52	.83	.26	.37	.80	.40	.49	.74	.54	.52
L2/7L															
jeitoso	.68	.61	.48	.79	.35	.47	.89	.22	.42	.85	.45	.60	.81	.41	.52
deserto	.68	.55	.42	.69	.44	.38	.80	.59	.57	.81	.45	.57	.75	.56	.47

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice G (Continuação).

Palavra/lista	1ª série			2ª série			3ª série			4ª série			Total		
	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T	<i>p</i>	D	I-T
IRRGL/IRE															
L1/8L															
freguesa	.47	.68	.47	.54	.74	.52	.80	.23	.37	.80	.55	.55	.66	.66	.53
serelepe	.31	.40	.331	.35	.50	.36	.45	.35	.24	.63	.61	.44	.43	.58	.37

Nota. L1 = Lista 1 (cinza); L2 = Lista 2 (cinza claro); 1L a 8L = número de letras da palavra; *p* = Dificuldade; D = Discriminação; I-T = Correlação item-total.

Apêndice H – Classificação das palavras adotada no presente estudo.

RL/RE

bala cuca tapa pista farda brava ladra prata viúva espiga cocada maluca salada picada bengala revista cava luta suja vala farta briga dupla magra antiga adulta³⁶ danada sabida estima monarca delicada limonada

RL/RGE

mico nabo surdo visto frade grilo fiapo aflito limite combate covarde nublado mosquito bife gago mudo nave tubo apito risco turno grade trave fiado trigo cabide peludo relato veludo

RL/IRE

raça salsa caçula batalha caçada ervilha mochila

RL/IRRGE

laje maço haste traje gozado recurso humilde gaze talco balde salto guizo filtro retalho saudade preguiça

RGL/RE

mina caju vara trena coruja morena parada feminina dura zebu cabana colina barata morada redonda

RGL/RGE

cano fumo ramo nono caro bomba tonto caqui enxada pepino tucano ovinho estante mente tombo enxame marido enxuto espanto campina destino ditongo

RGL/IRE

gema segura buzina cigana faxina gemada chuchu espiral vasilha gengiva rasa brasa camisa cevada cilada concha

RGL/IRRGE

xale lenço ensino sino ciclo regime enxoval cimento risonho

IRL/RE

boba cota gola soja sopa vela besta morta testa negra moeda zebra cadela canela capela gaveta maleta peteca sacola atleta sapeca mestra cartela coberta copa cova moda sola tela tola forca resta torta frota sopra viola caneca caneta cometa favela novela panela pipoca aposta bandeja cartola moderna

IRL/RGE

colo foco mole teto sorte teste bloco prego veste afeto ampola queda bigode camelo tijolo martelo materno marreco repleto sabonete uniforme esquerdo gabinete bote fofa moto neto voto forno poste brejo globo trevo febre miolo estojo leque colete novelo tapete cabelo catorze socorro carreta bodoque rabanete basquete

IRL/IRE

coxa seda selva graxa tigela colcha sarjeta roxa roça relva bolsa horta paçoca bexiga gorjeta palhoça

IRL/IRRGE

polvo terço excesso repolho carroça bezerro exceto luxo moço boxe treze azeite choque caxumba honesto serrote secreto reflexo universo

IRRGL/RE

garota espera sonora diretora fera amora careta direta cantora

³⁶ Note que a palavra *adulta* foi incorretamente classificada no presente estudo. A classificação adequada seria RL/IRE. Manteve-se aqui a classificação que foi utilizada nas análises estatísticas.

Apêndice H (*Continuação*).

IRRGL/RG

soro maroto

IRRGL/IRE

cera sexta presa cereja ameixa cebola bondosa formosa cerveja medrosa vaidosa
japonesa cela defesa orelha despesa nervosa certeza doutora luminosa mariposa

IRRGL/IRIGE

cego aceso caroço exposto celeste teimoso dengoso manhoso freguesa serelepe cesto
famoso externo jeitoso deserto saxofone

Nota. R = regular; RG = regra; IR = irregular; IRRG = irregular-regra; L = classificação para a
leitura; E = classificação para a escrita.