

Adriana Marques de Oliveira¹ 

Jair Lício Ferreira Santos² 

Simone Aparecida Capellini¹ 

Descritores

Leitura
Avaliação
Ensino Fundamental
Aprendizagem
Escolaridade

Keywords

Reading
Assessment
Education, Primary
Learning
Education Status

Endereço para correspondência:

Adriana Marques de Oliveira
Departamento de Fonoaudiologia,
Universidade Estadual Paulista “Júlio de Mesquita Filho”
Av. Hygino Muzzi Filho, 737, Marília
(SP), Brasil, CEP: 17.525-900.
E-mail: adriana.oliveira@unesp.br

Recebido em: Junho 05, 2019

Aceito em: Junho 08, 2020

Banco de palavras para leitura de escolares do Ensino Fundamental ciclo I, E-LEITURA I

Words database for reading by students from Basic Education I, E-READING I

RESUMO

Objetivo: Apresentar o processo de elaboração de um banco de palavras adequadas ao nível de proficiência de leitura de escolares do Ensino Fundamental I. **Método:** Selecionaram-se palavras de livros didáticos de Língua Portuguesa da rede pública de ensino de São Paulo, Paraná, Rio de Janeiro e Minas Gerais. Optou-se pelos substantivos e adjetivos. Excluíram-se as palavras homófonas, escritas em outros idiomas, com grafia errada, compostas por justaposição, abreviações, advérbios, locuções adverbiais, locuções prepositivas, meses do ano, numerais, palavras no aumentativo ou diminutivo, nomes próprios e gírias. As palavras foram categorizadas segundo frequência de ocorrência nos livros. Para tanto, foram utilizados os tercis da distribuição, a frequência média e o ponto de corte dos tercis. Para detectar possíveis falhas na seleção das palavras, foram selecionados 50 escolares do 1º ao 5º ano (10 por ano escolar) para leitura individual, com duração de 20 minutos, do banco de palavras. **Resultados:** Foram digitadas 286.290 palavras. Após análise dos critérios de inclusão/exclusão e categorização por frequência de ocorrência, o banco ficou constituído por 4.195 palavras. Após leitura pelos escolares, foram excluídas palavras que contemplavam os critérios de exclusão e que geravam desconforto por parte dos alunos. O banco ficou constituído por 4190 palavras, divididas em frequência: baixa (n= 3735, 88,59%), média (n= 374, 8,93%) e alta (n= 81, 1,93%), denominado E-LEITURA I. **Conclusão:** a elaboração de um banco de palavras de baixa, média e alta frequência de ocorrência para servir de estímulo linguístico foi adequadamente alcançado e disponibilizado para a prática clínica e pedagógica.

ABSTRACT

Purpose: To present the process of elaborating a word bank appropriate for the reading proficiency level of elementary school students. **Methods:** Words from Portuguese language textbooks used in the public school system of São Paulo, Paraná, Rio de Janeiro and Minas Gerais states of Brazil were selected. We opted for those belonging to the class of nouns and adjectives. Were excluded: homophones; other languages; abbreviations; adverbs; adverbial phrases; prepositional phrases; months of the year; numerals; diminutive or augmentative forms; proper names; misspellings; slang; and words composed by juxtaposition. The words were then categorized according to frequency of occurrence in the textbooks. For this purpose, the tertiles of the distribution, the mean frequency and cutoff point of the tertiles were used. To detect possible mistakes in the selection of words, 50 students from the 1st to 5th year, with 10 per school year were selected for individual reading, for 20 minutes from the word bank. **Results:** A total of 286,290 words were typed. After analyzing the inclusion / exclusion criteria and categorizing by frequency of occurrence, the bank amounted to 4,195 words. After reading by students, the E-READING I comprised 4,190 words classified according to frequency: low (n = 3735), medium (n = 374) and high (n = 81). **Conclusion:** The development of a low, medium and high frequency word bank, to serve as a linguistic stimulus, was achieved and made available for clinical and pedagogical practice.

Trabalho realizado no Departamento de Fonoaudiologia, Faculdade de Filosofia e Ciências, Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP - Marília (SP), Brasil.

¹ Faculdade de Filosofia e Ciências, Universidade Estadual Paulista “Júlio de Mesquita Filho” - Marília (SP), Brasil.

² Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo – USP - Ribeirão Preto (SP), Brasil.

Fontes de financiamento: Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, processo nº150687/2017-6 – Pós-Doutorado Júnior – PDJ.

Conflito de interesses: nada a declarar.



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

INTRODUÇÃO

A Avaliação Nacional da Alfabetização (ANA), realizada pelo Ministério da Educação - MEC, avalia, segundo níveis de proficiência, as habilidades básicas de leitura, escrita e matemática até o final do 3º ano do Ensino Fundamental. Os últimos resultados para os escolares do 3º ano (na época, considerado o último ano do ciclo de alfabetização) foram publicados em 2016: o número de escolares considerados proficientes em leitura era de 45,3%, o que significa que 65,3% dos escolares do 3º ano escolar não haviam atingido as competências mínimas de leitura para completar o ciclo de alfabetização⁽¹⁾.

O Saeb[†] (Sistema de Avaliação da Educação Básica), outro sistema de avaliação periódica, conduzido pelo Instituto Nacional de Estudos e Pesquisas Educacionais (Inep) em larga escala, teve como público alvo estudantes de 5º e 9º ano do ensino fundamental e do 3º Médio em 2017. Na prova de língua portuguesa, 39,3% dos escolares do 5º ano do Ensino Fundamental (último ano deste ciclo) estão abaixo do nível médio nacional (214,5 – nível 4 da escola de proficiência). Isso significa que esses escolares não estão hábeis, por exemplo, para localizar as informações explícitas nos textos, identificar o assunto principal ou os personagens, e inferir o sentido da palavra, por exemplo⁽²⁾.

Com os dados apresentados, fica evidente que a defasagem da educação brasileira se inicia no ensino fundamental, especialmente na fase mais importante - a alfabetização. A incapacidade de ler com acurácia e fluência afeta diretamente a compreensão da leitura e traz consequências negativas que vão além do sucesso na educação formal, pois tem o condão de prejudicar a vida profissional e até mesmo social do indivíduo, que não será capaz de compreender um simples texto de rede social⁽³⁻⁵⁾.

Pesquisadores⁽⁶⁻¹²⁾ buscam incessantemente a elaboração de instrumentos de avaliação e intervenção para essa faixa escolar tanto para promover e facilitar o processo de alfabetização como para identificar precocemente se algo em seu processo não está evoluindo de maneira esperada para que se possa intervir o quanto antes⁽⁴⁾. No entanto, no Brasil, muitas vezes os profissionais se deparam com dificuldades na elaboração desses procedimentos devido à inexistência de bancos de palavras para a seleção de estímulos linguísticos adequados para avaliar e intervir, especialmente, quando é necessário classificá-las por sua frequência de ocorrência.

Essa classificação, no entanto, pode envolver dificuldades em reconhecer que a frequência de ocorrência da palavra no material escrito geralmente é diferente da frequência da mesma palavra observada na comunicação oral. Alguns profissionais partem do que para eles é mais comum para classificar as palavras como de alta ou de baixa frequência. Como ressaltado por Pinheiro⁽⁶⁾, a exposição à palavra auditivamente afeta o seu reconhecimento

auditivo e não o reconhecimento visual. O mesmo se dá para a leitura, isto é, ao ler uma palavra ocorre o reconhecimento visual e não o auditivo. O sistema de reconhecimento visual e o sistema de reconhecimento auditivo são diferentes.

Em outros países, os pesquisadores disponibilizam bancos ou dicionários de frequências, nos quais as palavras são classificadas por frequência e extensão, a partir dos quais os profissionais podem selecionar as que melhor atendem a seus objetivos⁽¹³⁻¹⁵⁾. Na Espanha, por exemplo, a Real Academia Espanhola disponibiliza on-line lista das palavras de acordo com sua frequência.

No Brasil, o *Corpus Brasileiro*⁽¹⁶⁾ é uma coletânea de textos falados e escritos, composto por aproximadamente um bilhão de palavras com informações sobre a frequência de ocorrência e categorias. Nesse banco todas as categorias gramaticais de palavras estão incluídas, desde verbos, estejam eles no infinitivo, participio ou em sua forma conjugada, até advérbios, preposições, pronomes, siglas, nomes próprios, numerais, entre outros. O *Corpus*, por procurar abranger todas as variedades linguísticas da Língua Portuguesa, não permite a classificação das palavras por nível de escolaridade.

Em 2007, pesquisadores realizaram análise quantitativa da frequência dos fonemas e estruturas silábicas da língua portuguesa. Para tanto, fizeram o *corpus* da língua escrita com utilização do dicionário *Houaiss* da Língua Portuguesa. Nesse *corpus* não incluíram abreviações, siglas, estrangeirismos, homônimos, palavras com hífens e aquelas formadas por justaposição, bem como não consideraram a frequência de uso das palavras⁽¹⁷⁾.

No Brasil, existem listas de palavras, classificadas por frequência, regularidade ortográfica e extensão. A mais conhecida e amplamente utilizada é a Lista de Palavras e Pseudopalavras de Pinheiro⁽⁶⁾, formada por 96 palavras reais e 96 não-palavras.

Mediante o exposto, constata-se a necessidade de construir e disponibilizar um banco de palavras, classificado segundo a frequência de ocorrência com a qual aparece no material escrito, para que o profissional possa elaborar sua própria lista de palavras ou utilizá-las em instrumentos de intervenção.

A seleção de palavras classificadas em frequência e de acordo com a escolaridade também é importante para a elaboração de procedimentos que avaliam e intervêm em habilidades metalinguísticas, vocabulário e escrita^(6,7,18-20). Mesmo que as palavras não sejam utilizadas em sua forma escrita, os autores partem da seleção para a elaboração dos estímulos, seja a representação pela forma oral, escrita ou pictórica.

Por isso, esse estudo tem por objetivo apresentar o processo de elaboração de um banco de palavras adequadas para o nível de proficiência de leitura de escolares do Ensino Fundamental I, categorizado por frequência de ocorrência na apresentação do material escrito.

MÉTODO

Pesquisa aplicada para elaboração de um banco de palavras adequadas para o nível de proficiência de leitura de escolares do Ensino Fundamental I, a ser denominado E-LEITURA I.

[†] Desde 2019, a Avaliação Nacional de Alfabetização (ANA), a Avaliação Nacional da Educação Básica (Aneb) e a Avaliação Nacional do Rendimento Escolar (Anresc, também conhecida como Prova Brasil) deixaram de existir com esse nome. Agora todas as avaliações são denominadas SAEB e são diferenciadas pelo ano escolar avaliado a fim de realizar um diagnóstico do ensino fundamental e médio brasileiro e de fatores que podem interferir no desempenho do escolar. O SAEB foi reestruturado para se adequar à Base Nacional Comum Curricular (BNCC) que se tornou a referência na formulação das questões das provas em língua portuguesa, matemática, ciências da natureza e ciências humanas. Sua aplicação foi em 2019 e os resultados não foram publicados até o momento⁽⁹⁾.

A pesquisa aplicada objetiva gerar conhecimentos para aplicação prática visando à solução de problemas identificados⁽²¹⁾.

Procedimentos éticos

O estudo foi registrado na Plataforma Brasil (CAAE: 74853317.3.0000.5406) e aprovado pelo Comitê de Ética em Pesquisa da Faculdade de Filosofia e Ciências, Universidade Estadual Paulista “Júlio de Mesquita Filho” - UNESP (parecer nº 2.375.716).

Elaboração do Banco de Palavras para escolares do Ensino Fundamental I

Seleção do material didático

Para elaboração do banco, as palavras foram retiradas do material didático utilizado pelas redes de ensino municipal e estadual referente ao ensino da Língua da Portuguesa no Ensino Fundamental, ciclo I, nos estados de São Paulo, Paraná, Rio de Janeiro e Minas Gerais.

A seleção do material didático abrangeu três escolas do interior do estado de São Paulo, sendo duas municipais e uma estadual.

Para os demais estados, selecionou-se apenas uma escola municipal da capital de cada estado. Foram consultadas sete coleções de materiais didáticos distribuídos entre os estados (Tabela 1 e 2).

Com exceção dos livros da coleção *Ler e Escrever*, elaborados pela Secretaria da Educação do Estado de São Paulo, os demais materiais didáticos são organizados por editoras e aprovados pelo MEC, integrando o Plano Nacional do Livro Didático - PNLD 2016, 2017 e 2018.

A partir da relação dos livros didáticos autorizados pelo MEC, as escolas são livres para eleger aqueles que melhor atendam a seus objetivos. Nas escolas municipais consultadas do interior do Estado de São Paulo, os coordenadores informaram que, além dos livros didáticos autorizados pelo MEC e eleitos para serem trabalhados, há outro material utilizado na organização das aulas que é fornecido pela Secretaria da Educação do Estado de São Paulo, ou seja, os livros da coleção *Ler e Escrever*.

Na escola estadual consultada sobre o material didático, a coleção *Ler e Escrever* é adotada. Essa obra possui cadernos específicos de atividades e de recuperação para cada ano escolar, além de um caderno único apenas de textos. Para este banco não foram digitadas as palavras dos cadernos de recuperação, considerando que não são todos os alunos que realizam essas atividades.

Tabela 1 - Distribuição do material didático de Língua Portuguesa para o Ensino Fundamental, ciclo I, Estado de São Paulo

Escola 1	Referências
Ler e escrever: coletânea de atividades – 1º ao 5º ano	São Paulo (Estado) Secretaria da Educação. <i>Ler e escrever: coletânea de atividades – 1º ano</i> . Secretaria da Educação. 4. Ed. Rev. e atual. São Paulo: FDE; 2014. São Paulo (Estado) Secretaria da Educação. <i>Ler e escrever: coletânea de atividades – 2º ano</i> . Secretaria da Educação. 7. Ed. Rev. e atual. São Paulo: FDE; 2014. São Paulo (Estado) Secretaria da Educação. <i>Ler e escrever: coletânea de atividades – 3º ano</i> . Secretaria da Educação. 7. Ed. Rev. e atual. São Paulo: FDE; 2014. São Paulo (Estado) Secretaria da Educação. <i>Ler e escrever: coletânea de atividades – 4º ano</i> . Secretaria da Educação. 6ª Ed. Rev. e atual. São Paulo: FDE; 2015. São Paulo (Estado) Secretaria da Educação. <i>Ler e escrever: coletânea de atividades – 5º ano</i> . Secretaria da Educação. 6ª Ed. Rev. e atual. São Paulo: FDE; 2015.
Ler e escrever: livro de textos do aluno – 1º ao 5º ano	São Paulo (Estado) Secretaria da Educação. <i>Ler e escrever: livro de textos do aluno</i> . Secretaria da Educação. 7ª Ed. São Paulo: FDE; 2013.
Escola 2	
Coleção Porta Aberta – Edição renovada – Letramento e alfabetização, 1º ao 5º ano	Carpaneda IPM. <i>Porta Aberta – língua portuguesa, 1º ano: ensino fundamental: anos iniciais</i> . 1. Ed. São Paulo: FTD; 2014. Carpaneda IPM. <i>Porta Aberta – língua portuguesa, 2º ano: ensino fundamental: anos iniciais</i> . 1. Ed. São Paulo: FTD; 2014. Carpaneda IPM. <i>Porta Aberta – língua portuguesa, 3º ano: ensino fundamental: anos iniciais</i> . 1. Ed. São Paulo: FTD; 2014. Carpaneda IPM. <i>Porta Aberta – língua portuguesa, 4º ano: ensino fundamental: anos iniciais</i> . 1. Ed. São Paulo: FTD; 2014. Carpaneda IPM. <i>Porta Aberta – língua portuguesa, 5º ano: ensino fundamental: anos iniciais</i> . 1. Ed. São Paulo: FTD; 2014.
Escola 3	
Aprender e criar – Letramento e Alfabetização 1º ao 3º ano	Neves AAA, Carvalho A, Bevilacqua E, Grilo M. <i>Aprender e criar: letramento e alfabetização</i> , 1. 2ª Ed. São Paulo: Escala Educacional; 2014. Neves AAA, Carvalho A, Bevilacqua E, Grilo M. <i>Aprender e criar: letramento e alfabetização</i> , 2. 2ª Ed. São Paulo: Escala Educacional; 2014. Neves AAA, Carvalho A, Bevilacqua E, Grilo M. <i>Aprender e criar: letramento e alfabetização</i> , 3. 2ª Ed. São Paulo: Escala Educacional; 2014.
Ápis: Língua Portuguesa – 4º ao 5º ano	Borgatto AMT, Bertin TCH, Marchezi VLC. <i>Ápis: Língua Portuguesa 4º ano</i> . 2ª Ed. São Paulo: Ática; 2014. Borgatto AMT, Bertin TCH, Marchezi VLC. <i>Ápis: Língua Portuguesa 5º ano</i> . 2ª Ed. São Paulo: Ática; 2014.

Tabela 2. Distribuição do material didático de Língua Portuguesa para o Ensino Fundamental, ciclo I, dos estados do Minas Gerais, Paraná e Rio de Janeiro

Minas Gerais	Referências
Coleção Quatro Cantos – Português – Letramento e Alfabetização, 1º ao 3º ano	Porto A, Antoniol V. Coleção Quatro Cantos: Português – Letramento e Alfabetização – 1º ano: livro do aluno. Belo Horizonte: Dimensão; 2013. Porto A, Antoniol V. Coleção Quatro Cantos: Português – Letramento e Alfabetização – 2º ano: livro do aluno. Belo Horizonte: Dimensão; 2013. Porto A, Antoniol V. Coleção Quatro Cantos: Português – Letramento e Alfabetização – 3º ano: livro do aluno. Belo Horizonte: Dimensão; 2013.
Projeto Buriti: português: Ensino Fundamental: anos iniciais, 4º ao 5º ano	Editora Moderna. Projeto Buriti: português: ensino fundamental anos iniciais – 4º ano. 3ª Ed. São Paulo: Moderna; 2014. Editora Moderna. Projeto Buriti: português: ensino fundamental anos iniciais – 5º ano. 3ª Ed. São Paulo: Moderna; 2014.
Paraná	
Aprender juntos – Letramento e Alfabetização – 1º ao 5º ano	Vasconcelos A. Aprender Juntos: letramento e alfabetização, 1º ano: ensino fundamental: anos iniciais. 4ª Ed. São Paulo: Edições SM; 2014. Vasconcelos A. Aprender Juntos: letramento e alfabetização, 2º ano: ensino fundamental: anos iniciais. 4ª Ed. São Paulo: Edições SM; 2014. Vasconcelos A. Aprender Juntos: letramento e alfabetização, 3º ano: ensino fundamental: anos iniciais. 4ª Ed. São Paulo: Edições SM; 2014. Vasconcelos A. Aprender Juntos: português, 4º ano: ensino fundamental: anos iniciais. 4ª Ed. São Paulo: Edições SM; 2014. Vasconcelos A. Aprender Juntos: português, 5º ano: ensino fundamental: anos iniciais. 4ª Ed. São Paulo: Edições SM; 2014.
Rio de Janeiro	
Aprender juntos – Letramento e Alfabetização – 1º ao 3º ano	Vasconcelos A. Aprender Juntos: letramento e alfabetização, 1º ano: ensino fundamental: anos iniciais. 4ª Ed. São Paulo: Edições SM; 2014. Vasconcelos A. Aprender Juntos: letramento e alfabetização, 2º ano: ensino fundamental: anos iniciais. 4ª Ed. São Paulo: Edições SM; 2014. Vasconcelos A. Aprender Juntos: letramento e alfabetização, 3º ano: ensino fundamental: anos iniciais. 4ª Ed. São Paulo: Edições SM; 2014.
Projeto Buriti: português: Ensino Fundamental: anos iniciais, 4º ao 5º ano	Editora Moderna. Projeto Buriti: português: ensino fundamental anos iniciais – 4º ano. 3ª Ed. São Paulo: Moderna; 2014. Editora Moderna. Projeto Buriti: português: ensino fundamental anos iniciais – 5º ano. 3ª Ed. São Paulo: Moderna; 2014.

Elaboração do E-LEITURA I

Os aspectos metodológicos adotados para a elaboração deste estudo são semelhantes aos descritos por Oliveira e Capellini^(19,20).

Todas as palavras dos textos integrantes dos materiais didáticos foram digitadas em planilha *Excel*. Após a digitação, foram selecionadas aquelas que estavam de acordo com os critérios de inclusão do E-LEITURA I.

Para a seleção das palavras, optou-se por digitar aquelas dos textos contidos nos livros de Português. Foram inseridas as que pertencem, simultaneamente, à classe dos substantivos e adjetivos, de acordo com o contexto em que são usadas, ou seja, as palavras que poderiam ser classificadas como substantivos e/ou adjetivos (classificação essa constante no dicionário de língua portuguesa *Michaelis online*).

Ressalta-se que não bastava o adjetivo ser substantivado; a classificação adjetivo e substantivo precisava estar presente no dicionário de língua portuguesa. Por exemplo: no caso das palavras como “coruja”, classificadas nos dicionários como substantivo e adjetivo, como no exemplo: “Ela é uma mãe coruja” (adjetivo) e “A coruja não costuma caçar durante o dia” (substantivo). Já a palavra “abacate” nos dicionários é apenas classificada como substantivo, razão pela qual essas palavras foram incluídas no banco de palavras.

Os substantivos foram selecionados por serem uma classe frequente em qualquer texto, visto que exercem funções sintáticas importantes na frase. Optou-se por se ater a essa classe de palavras por ser o núcleo do sintagma nominal^(19,20).

Os substantivos podem ser reconhecidos por alguns critérios, assim como aqueles apresentados por Cegalla⁽²²⁾:

1. Representam os nomes dos seres;
2. São sempre núcleos do sintagma nominal;
3. Aceitam, em geral, o artigo;
4. Sintaticamente, podem exercer várias funções: sujeito; complemento verbal; agente da passiva; complemento nominal; predicativo; aposto;
5. Flexionam-se em gênero, número e grau.

Os substantivos e adjetivos (adjetivos substantivados) que flexionam-se em gênero, número e grau e que possuem essa classificação nos dicionários foram inseridos no banco de palavras.

Foram excluídas do banco palavras homófonas que pudessem oferecer ambiguidade dependendo do contexto por se tratarem de palavras isoladas, ou seja, é necessário a oração para que seja possível recuperar o seu significado e a sua pronúncia.

São as palavras homônimas homógrafas (escritas do mesmo jeito, decodificadas diferente – por exemplo: cor “a *cor* do seu vestido” e “essa música eu sei de *cor*”); palavras homônimas perfeitas (escritas da mesma forma e também pronunciadas igualmente – como na palavra *cedo* “eu *cedo* meu lugar para você” e “está muito *cedo* para dormir”).

Foram excluídas do banco de palavras, ainda, as palavras escritas em outros idiomas (mesmo as já incorporadas no dicionário de português), as abreviações, os advérbios, as locuções adverbiais e prepositivas, os adjetivos, os meses do ano, os numerais e as palavras no aumentativo ou diminutivo, além de gírias e palavras compostas por justaposição. Também não foram considerados nomes próprios, palavras com grafia registrada no material de forma errada e/ou para representar a pronúncia popular, palavras com ortografia divergente e arcaísmos, ou seja, palavras que não são mais utilizadas atualmente.

Como no Português Brasileiro o gênero dominante é o masculino, as palavras encontradas no gênero feminino eram passadas para o masculino. Foram mantidas as palavras no feminino se houvesse outro termo para o masculino ou se o sufixo fosse modificado.

Quanto ao número, usou-se sempre o singular. Se a palavra, ao ser passada do plural para o singular assumisse forma homônima homógrafa ou homônima perfeita ou, ainda, oferecesse qualquer tipo de ambiguidade, era, portanto, excluída do banco.

De acordo com a reforma ortográfica, o trema deixou de ser usado (por exemplo, bilíngue, pinguim, antiguidade) e os ditongos abertos (ei, oi, eu) só são acentuados no final da palavra. Quando a palavra aparecia escrita na regra anterior à reforma ortográfica, ela era adaptada.

Na elaboração do E-LEITURA II e III^(19,20), foram retiradas algumas palavras que geraram desconforto ou piadas por parte dos escolares, outras que geraram confusão no momento da leitura como a palavra *face* [f’a.si] que era pronunciada como em inglês *face* [f’ej.si] - oriundo de *facebook*. Em vista disso, também foram excluídas da relação algumas palavras, visto que elas podem despertar comportamentos indesejados nos escolares do Ensino Fundamental I como, por exemplo, *face*, *sexo*, *sexual*, *calcinha*, *cueca*, *virgem* e *capeta*.

Após esses critérios de seleção, foram contabilizadas todas as palavras e a quantidade de vezes que apareceram no material, a fim de levantar a frequência de ocorrência em cada ano escolar. As palavras foram organizadas em um banco único e enviadas ao estatístico para selecionar quais eram as de baixa, média e alta frequência comuns a todos os anos.

A partir dos procedimentos supracitados, foi criado um banco único de palavras para o Ensino Fundamental I (1º ao 5º anos), nomeado de E-LEITURA I.

As palavras selecionadas foram classificadas segundo a frequência de ocorrência no material didático consultado e, para isso, foram utilizados os tercis da distribuição, a frequência média e o ponto de corte dos tercis devido às frequências que se encontram próximas ao centro - para classificá-las em baixa, média e alta frequência.

Participantes

Com a finalidade de detectar possíveis falhas na seleção das palavras (que não atenderam aos critérios de inclusão e exclusão,

digitadas erroneamente ou que geraram comportamento de recusa para leitura), participaram deste estudo 50 escolares do ensino público estadual de uma cidade do centro-oeste paulista.

Os pais ou responsáveis pelos escolares assinaram duas vias do Termo de Consentimento Livre e Esclarecido (TCLE), conforme resolução do Conselho Nacional de Saúde CNS 196/96. Os 50 participantes foram agrupados da seguinte forma:

- GI: 10 escolares do 1º ano do Ensino Fundamental ciclo I;
- GII: 10 escolares do 2º ano do Ensino Fundamental ciclo I;
- GIII: 10 escolares do 3º ano do Ensino Fundamental ciclo I;
- GIV: 10 escolares do 4º ano do Ensino Fundamental ciclo I;
- GV: 10 escolares do 5º ano do Ensino Fundamental ciclo I.

Os escolares participantes foram selecionados de acordo com os critérios de inclusão e exclusão. As informações foram observadas no prontuário escolar e/ou obtidas a partir dos professores.

Adotaram-se como critérios de inclusão para seleção dos participantes: 1) estar regularmente matriculado no Ensino Fundamental ciclo I; 2) assinatura do Termo de Consentimento Livre e Esclarecido pelos pais ou responsáveis; 3) assinatura do Termo de Assentimento. Foram excluídos: 1) escolares que se recusaram a participar, embora os pais ou responsáveis tivessem assinado o termo de consentimento; 2) escolares com diagnóstico interdisciplinar de transtorno de aprendizagem, dislexia e transtorno de *déficit* de atenção e hiperatividade; 3) queixa de aprendizagem; 4) alteração de linguagem ou fala; 5) prejuízo na acuidade visual e auditiva; 6) diagnóstico de síndromes genéticas ou neurológicas; 7) histórico de repetência; 8) rebaixamento intelectual.

Procedimentos

As palavras de baixa, média e alta frequência do E-LEITURA I foram apresentadas aos escolares em folha sulfite, papel A4, digitadas em *ARIAL*, tamanho 14, letras maiúsculas, espaçamento duplo, separadas em três colunas. Cada folha apresentava, em média, 51 palavras, sendo lidas individualmente em voz alta, uma por vez, pelo escolar.

Cada escolar era solicitado a ler todas as palavras que conseguisse por 20 minutos. Ao completar o tempo, esperava-se terminar a página onde o aluno se encontrava para encerrar a avaliação. O escolar seguinte começava a leitura do ponto onde o último havia parado.

Esclarece-se que o procedimento inicialmente delineado solicitava que os escolares lessem todas as palavras do banco em sessões de 15 minutos. Estimou-se que seriam necessárias em média 15 sessões. Contudo, esse procedimento teve que ser repensado no primeiro dia da coleta dos dados, uma vez que os escolares mostraram cansaço. Diante do ocorrido, a coleta dos dados foi reduzida para no máximo duas sessões com duração de 20 minutos cada.

Os alunos do 2º e 3º anos participaram duas vezes da leitura do Banco de Palavras, completando 20 minutos, em cada vez, com distância de duas semanas entre elas. Isso ocorreu de

Tabela 3. Descrição do número e frequência das palavras digitadas para a elaboração do E-LEITURA I

COLEÇÃO	Número de Palavras digitadas	Frequência das palavras
Ler e escrever: coletânea de atividades 1º ao 5º ano	20.446	104.199
Coleção Porta Aberta – Letramento e alfabetização 1º ao 5º ano	8.992	43.335
Aprender juntos – Letramento e Alfabetização – 1º ao 5º ano	10.555	45.596
Aprender e criar – Letramento e Alfabetização 1º ao 3º ano	3.573	13.219
Coleção Quatro Cantos – Português – Letramento e Alfabetização, 1º ao 3º ano	3.651	11.724
Ápis: Língua Portuguesa – 4º ao 5º ano	9.830	49.834
Projeto Burity: português: Ensino Fundamental: anos iniciais – 4º ao 5º ano	4.869	18.383
TOTAL	61.916	286.290

Tabela 4. Distribuição a partir do ponto de corte dos tercis para determinação da frequência de ocorrência das palavras, total das frequências e frequência média, quantidade de palavras por frequência

	E-LEITURA I
Total das frequências	48.185
Frequência média	11,48
Ponto de corte	
1º tercil	22/23
2º tercil	100/101
Número de repetições por frequência	
Baixa	1-22
Média	23-100
Alta	101-590
Quantidade de palavras por frequência	
Baixa	3.738
Média	375
Alta	82
Total de palavras E-LEITURA I	4.195

forma aleatória. Ao terminar a leitura dos 10 alunos, iniciou-se novamente com o primeiro aluno até acabar a leitura do banco. Isso resultou na participação de todos os alunos novamente. Já os escolares do 4º e 5º anos participaram apenas uma vez. Os alunos foram retirados da sala de aula de acordo com a disponibilidade da escola e autorização do professor e da direção.

Os escolares do 4º e 5º anos realizaram a leitura das palavras nos meses de agosto e setembro; os do 2º e 3º anos nos meses de outubro e novembro e os do 1º ano, no mês de novembro. A leitura feita pelos escolares do 1º ano foi a última a ser iniciada, a fim de poderem ter maior tempo de alfabetização e, portanto, condições de receber instrução quanto à atividade de decodificação.

Análise dos resultados

As informações coletadas foram registradas em um banco de dados do programa *Microsoft Excel*. Para classificar as palavras como de baixa, média e alta frequência, foram utilizados os tercis da distribuição e também a frequência média e o ponto de corte dos tercis (devido às frequências que se encontram próximas ao centro). O tercil divide o intervalo de uma distribuição de frequência em três classes de igual número (33,33%). No caso de uma distribuição simétrica para encontrar os valores dos tercis, verifica-se quais se localizam no intervalo entre 33,33% e entre 66,66%.

RESULTADOS

Descrição dos resultados da seleção de palavras

Foram digitadas 286.290 palavras (contadas em frequência de ocorrência, ou seja, uma mesma palavra aparece mais de uma vez) entre artigos, preposições, adjetivos, verbos, substantivos e outras. Ao retirar as palavras repetidas, ficamos com 61.916 palavras, como apresentado na Tabela 3, por coleção de livros utilizados para a elaboração do E-LEITURA I.

Todas as 61.916 palavras foram analisadas e selecionadas de acordo com os critérios estabelecidos. Após análise e seleção, o E-LEITURA I consta de 4.195 palavras que foram enviadas para a classificação da frequência de ocorrência (baixa, média ou alta).

No E-LEITURA I, utilizaram-se os valores da distribuição acumulada (frequência acumulada) por se tratar de um banco de palavras com muitas repetições. A distribuição dos dados, portanto, é assimétrica. Como se trata de uma variável nominal, a distribuição dos tercis no E-LEITURA I não contém exatamente 33,33% do total.

Nesse caso, o primeiro tercil inicia-se na repetição 22, porcentagem acumulada 33,30% até a porcentagem 34,07% e termina na repetição 23, porcentagem acumulada 34,12%. O segundo tercil inicia-se na porcentagem acumulada 66,56% (repetição 100) e termina na porcentagem acumulada 66,77% (repetição 101).

A partir dos valores dos tercis, foi possível classificar as palavras como de baixa, média e alta frequência, conforme a Tabela 4.

No E-LEITURA I, as sílabas canônicas (consoante “C” e vogal “V” - CV) representam 44,25% do total de palavras do E-LEITURA I, com 1844 delas, seguida das não canônicas CVC (n= 855, 20,52%, exemplo: “*bactéria*”), VC (n=458, 10,99%, como em “*ar*”), V (n=400, 9,60%, como em “*idade*”) e CCV (n=321, 7,70%, como em “*brejo*”). Abaixo de cinco por cento encontram-se as seguintes: CVV (n = 136, 3,26%, como em “*goiaba*”), CCVC (n= 63, 1,51%, como em “*clássico*”), VV (n= 25, 0,60%, como em “*autor*”), CCVV (n= 15, 0,36%, como em “*chão*”), CVCC (n= 13, 0,31%, como em “*monstro*”), CVVC (n= 12, 0,29%, como em “*questão*”), VVC (n= 1, 0,26%, “*austriaco*”), CCVCC (n= 10, 0,24%, como em “*translação*”), CVVC (n= 2, 0,05%, como em “*cais*”), VVC (N=1, 0,02%) e CCVVC (n= 1, 0,02%, como em “*braille*”). Ao todo foram encontradas 16 estruturas silábicas.

Quanto à extensão, as trissilábicas são as de maior ocorrência, representando o total de 35,23% com 1468 palavras, seguidas pelas polissilábicas com quatro sílabas, com 1159 palavras (27,81%). As palavras dissilábicas e polissilábicas com cinco sílabas vêm a seguir com 87 palavras (19,85%) e 509 (12,22%), respectivamente. As polissilábicas com mais de seis sílabas representam 3,62% do total das palavras do E-LEITURA I, sendo com seis sílabas n= 115 (2,76%), com sete n= 31 (0,74%), com oito sílabas n= 4 (0,10%) e com dez sílabas n= 1 (0,02%). As monossilábicas representam 1,27% do banco, com 53 palavras.

A distribuição das palavras de baixa, média e alta frequência por extensão e complexidade silábica encontra-se na Tabela 5.

Descrição dos resultados da leitura do E-LEITURA I

Quanto aos escolares do 1º ano do Ensino Fundamental I, não foi possível avaliá-los. Ao iniciar a leitura com esses escolares, ainda nas palavras de alta frequência, observou-se que apenas liam as monossilábicas e dissilábicas canônicas (CV). Por exemplo, se a palavra tivesse a letra x, como em bruxa, ou a letra r, no meio da palavra, como em “porta”, ou, no final, como em “amor”, eles já avisavam que não conseguiam lê-la. Foram chamadas duas alunas do 1º ano para ler as palavras de alta frequência e, após a leitura, optou-se por excluir o 1º ano do estudo, visto

que o banco apresenta, em relação a essas crianças alvo, alta complexidade.

Após a leitura das palavras pelos escolares do 2º ao 5º ano, anotou-se que algumas palavras presentes contemplavam o critério de exclusão e outras geravam desconforto por parte dos alunos. Portanto, as palavras seguintes foram retiradas do banco: alta frequência: cor (que pode ser lida com a vogal tônica aberta ou fechada / o/ ou /o/); média frequência: ministro (do verbo ministrar) e as de baixa frequência: arrear (verbo), palhava (não consta no dicionário *Michaelis e no VOLP – Vocabulário Ortográfico da Língua Portuguesa – da Academia Brasileira de Letras*) e a palavra inferno (recusa de leitura).

Após a exclusão dessas palavras, o E-LEITURA I ficou constituído por 4190 palavras, divididas em frequência: alta - 81 palavras, o que corresponde a 1,93% do banco, média - 374 (8,93%) e baixa - 3735 (88,59%) palavras.

O banco de palavras E-LEITURA I é apresentado nos Tabela Suplementar 1, 2 e 3, divididos em partes - Tabela Suplementar 1, palavras de alta frequência; Tabela Suplementar 2, palavras de média frequência e Tabela Suplementar 3, palavras de baixa frequência. As palavras são apresentadas em ordem alfabética, indicando-se a quantidade de sílabas (extensão silábica), complexidade da palavra segundo a estrutura da sílaba inicial (como as consoantes e vogais são organizadas na sílaba, as vogais

Tabela 5. Distribuição das palavras de alta, média e baixa frequência segundo extensão silábica e complexidade da sílaba inicial

E-LEITURA I					
Alta frequência					
Extensão silábica		Complexidade da Sílabla Inicial			
Monossílabo	13 (16,05%)	V	7 (8,64%)	CVC	12 (14,81%)
Dissílabo	42 (51,85%)	VC	5 (6,17%)	CVV	8 (9,88%)
Trissílabo	23 (28,40%)	CV	40 (49,38%)	CCVC	4 (4,94%)
Polissílabo (quatro sílabas)	3 (3,70%)	CCV	4 (4,94%)	CCVV	1 (1,23%)
Total palavras	81 (100%)				
Média Frequência					
Extensão silábica		Complexidade da Sílabla Inicial			
Monossílabo	18 (4,81%)	V	31 (8,29%)	VCC	1 (0,27%)
Dissílabo	145 (38,77%)	VC	32 (8,56%)	CCVC	8 (2,14%)
Trissílabo	140 (37,43%)	CV	180 (48,13%)	CCVV	2 (0,53%)
Polissílabo (quatro sílabas)	62 (16,58%)	CCV	21 (5,61%)	CVVC	3 (0,80%)
Polissílabo (cinco sílabas)	7 (1,87%)	CVC	79 (21,12%)	CVCC	1 (0,27%)
Polissílabo (seis sílabas)	2 (0,53%)	CVV	16 (4,28%)		
Total palavras	374 (100%)				
Baixa Frequência					
Extensão silábica		Complexidade da Sílabla Inicial			
Monossílabo	22 (0,59%)	V	362 (9,75%)	CVV	2 (0,05%)
Dissílabo	642 (17,24%)	VC	421 (11,34%)	CVCC	12 (0,32%)
Trissílabo	1307 (35,16%)	CV	1626 (43,75%)	CCVC	51 (1,37%)
Polissílabo (quatro sílabas)	1094 (29,47%)	VV	26 (0,67%)	CCVV	12 (0,32%)
Polissílabo (cinco sílabas)	502 (13,52%)	VVC	1 (0,03%)	CCVCC	10 (0,27)
Polissílabo (seis sílabas)	113 (3,04%)	VCC	10 (0,27%)	CCVVC	1 (0,03%)
Polissílabo (sete sílabas)	31 (0,84%)	CCV	296 (7,97%)		
Polissílabo (oito sílabas)	4 (0,11%)	CVC	764 (20,58%)		
Polissílabo (nove sílabas)	0 (0,00%)	CVV	112 (3,02%)		
Polissílabo (dez sílabas)	1 (0,03%)	CVVC	9 (0,24%)		
Total palavras	3716 (100%)				
Total de palavras E-LEITURA I	4190				

estão representadas pela letra “V” e as consoantes pela letra “C”), bem como o número de repetições de cada palavra no banco.

DISCUSSÃO

A elaboração do banco de palavras E-LEITURA I partiu da necessidade de estímulos linguísticos para a elaboração de instrumentos de avaliação e intervenção em escolares do Ensino Fundamental I. Ao utilizar palavras para avaliação da leitura, por exemplo, a lista de palavras isoladas é um dos materiais mais eficazes. No entanto, a seleção das palavras para composição de uma lista deve seguir alguns critérios de observação de características psicolinguísticas como regularidade, comprimento e frequência de ocorrência^(6,18,23). No entanto, ao falar de frequência de ocorrência não se refere à modalidade oral que o escolar possui com a palavra e, sim, à modalidade escrita, à experiência de leitura, da quantidade de vezes que ele visualizou essa palavra, isto é, a representação visual da palavra⁽⁶⁾. Esse cuidado na seleção por frequência e nível escolar não é importante apenas para atividades de leitura, mas também para elaboração de instrumentos que avaliem e intervenham em habilidades metalinguísticas, vocabulário e escrita^(6,7,19,20).

Para o E-LEITURA I, optou-se, por utilizar o material elaborado pela Secretaria do Estado de São Paulo e o indicado pelo MEC (para todo o Brasil) eleitos para serem trabalhados no ano de 2017, tanto em São Paulo, como em outros estados (Paraná, Minas Gerais e Rio de Janeiro) para escolas públicas e municipais. Não há livros específicos para cada região do país, portanto não se privilegiam os regionalismos de cada uma.

Não contemplar os regionalismos é a principal crítica recebida aos poucos bancos de palavras já existentes^(19,20). Neste estudo, observou-se que, apesar do maior número de livros utilizados, não foi observado um aumento significativo de palavras em relação a bancos conhecidos^(19,20), mas sim da quantidade de vezes que cada palavra aparece.

No E-LEITURA I, a estrutura CV é predominante, seguida da CVC. Esse achado vem ao encontro de estudos que levantaram o perfil das palavras do português brasileiro e também identificaram maior ocorrência da estrutura CV seguida da CVC^(17,24). A estrutura CV (canônica) é a mais frequente na maioria das línguas, inclusive na portuguesa, seguida das não canônicas V, VC e CVC. A estrutura CV, por ser a predominante, é uma das primeiras a surgir na aquisição do sistema linguístico e os escolares tendem a aprender primeiro essa estrutura silábica^(17,24,25).

De acordo com as diretrizes da Base Nacional Curricular Comum - BNCC⁽²⁶⁾, as estruturas CV, V, CVC e CCV devem ser adquiridas ao longo do 2º e 3º ano, quando os escolares devem ler e escrever corretamente as palavras com tais estruturas em todas as sílabas. A estrutura VC deve ser dominada no 3º ano escolar e as estruturas VV e CVV ao longo do 3º e 4º ano escolar.

Ao classificar quais estruturas ocorrem mais frequentemente no banco (CV, CVC, VC, V, CCV, CVV, CCVC e VV), observa-se que seguem o padrão de desenvolvimento das diretrizes da BNCC⁽²⁶⁾, o que mostra que o E-LEITURA está em consonância com o estipulado para o Ensino Fundamental I. A estrutura mais complexa CCVC encontra-se presente na sequência, provavelmente

porque no E-LEITURA é constituído pelo material didático do 1º ao 5º do Ensino Fundamental I.

Em relação à extensão da palavra, os achados deste estudo também vêm ao encontro da literatura que afirma que o português é uma língua predominantemente trissilábica, com poucos monossílabos^(17,25). No entanto, ao contrário do que ocorre na análise de Marques⁽²⁵⁾, na qual em primeiro aparece um maior número de palavras trissilábicas, seguida por dissilábicas, neste estudo encontrou-se uma maior número de palavras trissilábicas, seguida de polissilábicas com quatro sílabas e depois as dissilábicas.

Neste estudo o menor número de sílaba encontrado foi um e o maior 10, sendo a palavra de maior número “*otorrinolaringologista*”. Em estudo realizado no Brasil⁽¹⁷⁾, de levantamento do *Corpus* da língua escrita, do dicionário *Houaiss*, o menor número de sílabas é um e o maior é de 20 sílabas (“*pneumoultramicroscopicossilicovulcanoconiótico*”).

Quanto aos escolares do 1º ano, não foi possível a leitura das palavras do banco, porque eles apenas liam as palavras monossilábicas e dissilábicas canônicas (CV). Palavras com estruturas CVC ou CVV, já na primeira sílaba não eram lidas por eles. Tal comportamento vem ao encontro do que está previsto na Base Nacional Curricular Comum - BNCC⁽²⁶⁾, que dispõe que a decodificação e a fluência de leitura, que, para o documento, são descritas como a leitura de palavras novas e de palavras de uso frequente, serão lidas globalmente por memorização. Essa precisão, para o documento da Base Nacional Curricular Comum, só é encontrada em alunos que já compreendem o sistema de escrita, fato que, segundo seus organizadores, pode ocorrer até o final do 2º ano.

O comportamento dos escolares de parar a leitura, relatando cansaço e a decisão de limitar o tempo, deve-se ao fato de que, ao ler, dependendo da rota de leitura a ser utilizada, exige-se mais da memória operacional fonológica ou visuoespacial, o que possivelmente gera sobrecarga⁽²⁷⁻³⁰⁾. É consenso, na literatura, que essas palavras exigem um tempo maior de leitura do que as palavras de alta e média frequência e que o número de erros também costuma aumentar^(6,7,18-20,27-29). Por serem de baixa frequência, o leitor não tem sua representação no léxico visual de entrada, o que o leva a realizar a leitura pela via fonológica, solicitando com mais frequência seus processos de atenção.

O principal objetivo do banco de palavras é oferecer uma ferramenta na qual os profissionais possam obter os estímulos linguísticos necessários para a seleção de palavras de acordo com a frequência de ocorrência no material escrito.

CONSIDERAÇÕES FINAIS

O objetivo de elaborar um banco de palavras de baixa, média e alta frequência do material escrito para servir de estímulo linguístico foi adequadamente alcançado.

Do total de palavras do E-LEITURA I, as de baixa frequência correspondem a 88,59% do banco. Quanto à estrutura da sílaba inicial, as sílabas canônicas são as mais frequentes. No que concerne à extensão, os trissílabos são as de maior ocorrência, seguida dos polissílabos com quatro sílabas.

O banco de palavras, apesar de nomeado E-LEITURA, foi realizado a partir da frequência de ocorrência das palavras,

portanto, tais estímulos também podem ser utilizados em atividades de escrita e ortografia. A frequência das palavras não se altera se o objeto de análise é decodificação ou codificação.

O acesso ao banco de palavras e as possibilidades de ferramentas que poderão ser desenvolvidas a partir dos estímulos linguísticos nele apresentados poderão auxiliar o profissional a identificar e a intervir precocemente nas dificuldades de leitura.

AGRADECIMENTOS

À Gabriela Franco dos Santos Liporaci, Irene B. Marques de Oliveira e Larissa Sellin pelo auxílio na digitação das palavras para a Elaboração do E-LEITURA I. À Cristiane Tomazinho Fumagalli, Edmilton Oseias da Cunha, Fernanda Boatto Belgamasco, Juliana Mendes Alves, Luciana Cássia Pereira Capel, Luciana Cordeiro Felipetto e Patrícia Hiraoka pelo fornecimento e indicação do material didático.

REFERÊNCIAS

1. INEP: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. MEC: Ministério da Educação. Sistema de avaliação da Educação Básica – avaliação nacional da alfabetização – Edição 2016 [Internet]. Brasília: INEP/MEC; 2017 [citado em 2017 Out 9]. Disponível em: <http://portal.mec.gov.br/docman/outubro-2017-pdf/75181-resultados-ana-2016-pdf/file>
2. INEP: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Relatório SAEB. Brasília: INEP/MEC; 2019.
3. Lervåg A, Hulme C, Melby-Lervåg M. Unpicking the developmental relationship between oral language skills and reading comprehension: it's simple, but complex. *Child Dev.* 2018;89(5):1821-38. <http://dx.doi.org/10.1111/cdev.12861>. PMID:28605008.
4. Colenbrander D, Ricketts J, Bredmore HL. Early identification of Dyslexia: understanding the issues. *Lang Speech Hear Serv Sch.* 2018;49(4):817-28. http://dx.doi.org/10.1044/2018_LSHSS-DYSLC-18-0007. PMID:30458543.
5. Oakhill J, Cain K, Elbro C. Compreensão de leitura: teoria e prática. São Paulo: Hogrefe CETEPP; 2017.
6. Pinheiro AMV. Leitura e escrita: uma abordagem cognitiva. 2. ed. Campinas: Livro Pleno; 2006.
7. Capellini SA, Oliveira AM, Cuetos F. PROLEC: provas de avaliação dos processos de leitura. 3. ed. São Paulo: Casa do Psicólogo; 2014.
8. Salles JF, Parente MAPP. Avaliação da leitura e escrita de palavras em crianças de 2ª série: abordagem neuropsicológica cognitiva. *Psicol Reflex Crit.* 2007;20(2):220-8. <http://dx.doi.org/10.1590/S0102-79722007000200007>.
9. Lamprecht R, Santos RM, Freitas GM, Brodacz R, Siqueira M, Costa AC, et al. Confias: consciência fonológica instrumento de avaliação sequência. São Paulo: Pearson; 2015.
10. Santos B, Capellini SA. PRONAR-LE - Programa de remediação com nomeação automática rápida e leitura. Ribeirão Preto: Booktoy; 2018.
11. Seabra AG, Capovilla FC. TCLPP - Teste de competência de leitura de palavras e pseudopalavras. São Paulo (SP): Memnon; 2010.
12. Silva C, Capellini SA. Programa de Intervenção fonológica para escolares em fase inicial de alfabetização: manual e caderno do aplicador. Ribeirão Preto: Booktoy; 2019.
13. Balota DA, Yap MJ, Cortese MJ, Hutchison KA, Kessler B, Loftis B, et al. The English Lexicon Project. *Behav Res Methods.* 2007;39(3):445-59. <http://dx.doi.org/10.3758/BF03193014>. PMID:17958156.
14. Martínez JA, García E. Diccionario. Frecuencias del castellano escrito en niños de 6 a 12 años. Salamanca: Publicaciones de la Universidad Pontificia, 2004.
15. López MXB. O Galego fundamental: dicionário de frequências. Santiago de Compostela: A Coruña Fundación Pedro Barrié de la Maza; 2007.
16. Corpus Brasileiro. Projeto AC/DC: corpo Corpus Brasileiro [Internet]. 2018 [citado em 2018 Nov 11]. Disponível em: <http://www.linguateca.pt/acesso/corpus.php?corpus=CBRAS>
17. Viaro ME, Guimarães-Filho ZO. Análise quantitativa da frequência dos fonemas e estruturas silábicas portuguesas. *Estudos Linguísticos.* 2007;26(1):27-36.
18. Cuetos F. *Psicología de la lectura.* 8. ed. Las Rozas, Madrid: Wolters Kluwer España; 2010.
19. Oliveira AM, Capellini SA. E-LEITURA II: banco de palavras para leitura de escolares do Ensino Fundamental II. *CoDAS.* 2016;28(6):778-817. <http://dx.doi.org/10.1590/2317-1782/20162016049>. PMID:27982255.
20. Oliveira AM, Capellini SA. Banco de palavras para leitura de escolares do ensino médio: E-LEITURA III. *Rev CEFAC.* 2016;18(6):1404-46. <http://dx.doi.org/10.1590/1982-0216201618610516>.
21. Silva EL, Menezes EM. Metodologia da pesquisa e elaboração de dissertação. 4. ed. Florianópolis: UFSC; 2005.
22. Cegalla DP. *Novíssima Gramática da Língua Portuguesa.* 48. ed. São Paulo: Companhia Editora Nacional; 2008.
23. Pinheiro AMV, Rothe-Neves R. Avaliação cognitiva de leitura e escrita: as tarefas de leitura em voz alta e ditado. *Psicol Refl Crit (Lond).* 2001;14(2):399-408. <http://dx.doi.org/10.1590/S0102-79722001000200014>.
24. Matzenauer CLB. Bases para o entendimento da aquisição fonológica. In: Lamprecht RR. (Ed.), *Aquisição Fonológica do Português: perfil de desenvolvimento e subsídios para a terapia.* Porto Alegre: Artmed; 2004. p. 33-58.
25. Marques LF. Estruturas silábicas do português do Brasil: uma análise tipológica [dissertação]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo; 2008. <http://dx.doi.org/10.11606/D.8.2008.tde-06082009-163245>.
26. Brasil. Ministério da Educação. Secretaria da Educação Básica. Base nacional comum curricular. Brasília, DF: Ministério da Educação; 2016. [citado em 2018 Dez 11]. Disponível em: <http://basenacionalcomum.mec.gov.br/#/site/inicio>
27. Baddeley A. Memória de trabalho. In: Baddeley A, Anderson MC, Eysenck MW. (Eds.), *Memória.* Porto Alegre: Artmed; 2011. p. 54-82.
28. Gonçalves HA, Viapiana VF, Sartori MS, Giacomoni CH, Stein LM, Fonseca RP. Funções executivas predizem o processamento de habilidades básicas de leitura, escrita e matemática? *Neuropsicol Lat AM.* 2017;9(3):42-54. <http://dx.doi.org/10.5579/rml.2016.0393>.
29. Zhao J, Liu M, Liu H, Huang C. Increased deficit of visual attention span with development in Chinese children with developmental dyslexia. *Sci Rep.* 2018;8(1):3153. <http://dx.doi.org/10.1038/s41598-018-21578-5>. PMID:29453430.
30. Awadh FHR, Phénix T, Antzaka A, Lallier M, Carreiras M, Valdois S. Cross-Language Modulation of visual attention span: an Arabic-French-Spanish in skilled adult readers. *Front Psychol.* 2018;7:307. <http://dx.doi.org/10.3389/fpsyg.2016.00307>. PMID:27014125.

Contribuição dos autores

AMO: participou da concepção, do planejamento do projeto de pesquisa, coleta dos dados, análise e interpretação dos dados, redação e revisão crítica do manuscrito; JLFS: participou do planejamento do projeto de pesquisa, da análise (responsável pela utilização das frequências observadas – divididas em tercís – para a classificação das palavras) e interpretação dos dados e revisão crítica do manuscrito; SAC: participou da concepção, planejamento e orientação do projeto de pesquisa, da redação e revisão crítica do manuscrito.

MATERIAL SUPLEMENTAR

Este artigo acompanha material suplementar.

Tabela Suplementar 1. Banco de palavras para leitura de escolares do Ensino Fundamental I - E-LEITURA I – palavras de alta frequência

Tabela Suplementar 2. Banco de palavras para leitura de escolares do Ensino Fundamental I - E-LEITURA I – palavras de média frequência

Tabela Suplementar 3. Banco de palavras para leitura de escolares do Ensino Fundamental I - E-LEITURA I – palavras de baixa frequência

Este material está disponível como parte da versão online do artigo na página: <https://www.scielo.br/j/codas/>