

O PAPEL DA AQUISIÇÃO DA ESCRITA NA REPRESENTAÇÃO DA CODA SILÁBICA DO PORTUGUÊS BRASILEIRO: DA FONOLOGIA INFANTIL À FONOLOGIA ADULTA

Lissa Pachalski*

Ana Ruth Moresco Miranda**

- **RESUMO:** A aquisição da escrita tem sido considerada uma importante fonte de mudança na mente/cérebro do aprendiz, especialmente na maneira como a linguagem é processada. Vários aspectos desse tipo de mudança têm sido estudados, entretanto, poucos trabalhos abordam mudanças especificamente na representação fonológica. Este artigo busca promover esta discussão, com foco na representação da coda silábica do Português Brasileiro. Para tanto, são analisados dados de escrita de 166 crianças brasileiras monolíngues, estudantes dos quatro primeiros anos do Ensino Fundamental. Os resultados mostram que as crianças produzem mais erros na escrita das consoantes nasal e lateral em relação à fricativa e à rótica nesta posição silábica. Este resultado é atribuído principalmente a uma assimetria entre a fonologia infantil e a fonologia adulta, que é refletida no sistema ortográfico. Em razão desse conflito, argumenta-se que a representação da coda pode ser alterada no sistema fonológico das crianças brasileiras, tendo como resultado o seu alinhamento com a gramática alvo.
- **PALAVRAS-CHAVE:** Coda; Sílabas; Português Brasileiro; Aquisição da escrita; Aquisição fonológica.

Introdução

É amplamente reconhecido que os sistemas de escrita alfabética, mesmo aqueles com profundidade ortográfica, demandam do aprendiz um acesso consciente ao conhecimento fonológico, uma vez que o princípio operacional desses sistemas se baseia nas relações entre fonemas e grafemas (Coulmas, 1989, 2003; Landsmann, 1993; Treiman; Kessler, 2014). O aprendiz precisa entender que sua língua é formada por unidades não-significativas, que podem ser analisadas em partes cada vez menores e em

* Universidade Federal de Pelotas (UFPeL), Centro de Letras e Comunicação, Pelotas, RS, Brasil. Doutoranda do Programa de Pós-Graduação em Letras (PPGL-UFPeL). pachalskil@gmail.com. ORCID: <https://orcid.org/0000-0002-7288-8067>.

** Universidade Federal de Pelotas (UFPeL), Faculdade de Educação, Pelotas, RS, Brasil. Professora titular da Faculdade de Educação (FaE-UFPeL). anaruthmmiranda@gmail.com. ORCID: <https://orcid.org/0000-0002-1380-5751>.

diferentes camadas de organização (Lieberman *et al.*, 1974; Morais *et al.*, 1979; Gombert, 1992). A aquisição da escrita, nesse sentido, pode ser considerada uma retomada do conhecimento fonológico internalizado. Além disso, cria condições para a atualização¹ da língua em uma nova substância, alternativa à fala (Miranda, 2017; Lyons, 1968). É possível, então, esperar algum grau de mudança representacional no conhecimento fonológico ao longo do processo de aquisição da escrita, especialmente quando a fonologia infantil não é bem refletida no sistema ortográfico da mesma maneira como é a gramática-alvo (Miranda, 2017; Veloso, 2006, 2019; Karmiloff-Smith, 1992). É por isso também que os dados de desenvolvimento da fala e da escrita podem ser tomados como uma base empírica relevante para discussões não apenas sobre o conhecimento fonológico infantil, mas também sobre o sistema fonológico-alvo. Esta é a perspectiva adotada neste artigo, de forma que se possa argumentar a favor de um tipo particular de mudança na representação fonológica de crianças brasileiras que estão adquirindo a escrita, a saber, a representação da coda silábica do Português Brasileiro (PB).

Não há consenso sobre o estatuto fonológico dessa estrutura silábica em termos da sua composição segmental. Os autores divergem quanto ao número e ao tipo de segmentos que podem compor a coda no PB, especialmente sobre a presença ou não de uma consoante nasal e de uma lateral. A proposta mais difundida é a que atribui quatro segmentos consonantais a essa posição silábica, a saber, /S, N, L, R/ (Bisol, 1999). Deve-se notar, entretanto, que essa discussão diz respeito à gramática-alvo, isto é, à gramática adulta, letrada. Na aquisição fonológica do PB, os fatos podem ser um pouco diferentes. Embora alguns autores, como Mezzomo (2004), assumam que as crianças brasileiras adquirem os quatro segmentos consonantais propostos por Bisol (1999), outros questionam essa suposição. Miranda (2009, 2019), por exemplo, sugere que, na verdade, as crianças podem não adquirir as codas nasal e lateral – pelo menos não até adquirirem a escrita, um processo capaz de ajustar a gramática em desenvolvimento à gramática-alvo.

O presente artigo segue a posição de Miranda, com o objetivo de contribuir com a discussão sobre o estatuto representacional da coda silábica no PB e argumentar em favor do potencial dos dados de aquisição da escrita nesse tipo de discussão. Para tanto, as seções que seguem detalham a discussão sobre esse constituinte silábico no Português e revisam os estudos existentes sobre o tema, abrangendo a gramática adulta, a aquisição fonológica e a aquisição da escrita. Também são feitos alguns comentários sobre a questão da mudança representacional. Em seguida, os procedimentos metodológicos são esclarecidos, seguidos pelos resultados e sua discussão, que inclui algumas implicações dos achados para os três campos abordados neste artigo, quais sejam, a teoria fonológica, a aquisição fonológica e a aquisição da escrita.

¹ O termo “atualização” corresponde à sua concepção linguística formal, que se refere à ideia de realização de unidades linguísticas abstratas (por exemplo, fonemas) em uma substância física – primária ou secundária, som ou escrita (Crystal, 1988; Lyons, 1968).

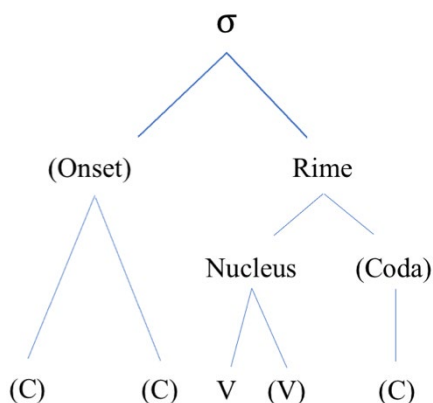
Estrutura silábica no PB: fonologia e ortografia

O PB é uma língua românica com uma organização segmental e silábica bastante previsível. Segue claramente o Princípio de Sequência de Sonoridade (SSP), evitando platôs e preferindo mudanças suaves de sonoridade dentro das sílabas (Bisol, 1999; Collischonn; Wetzels, 2016). O PB também apresenta uma tendência clássica para o padrão CV, que está presente em mais da metade do léxico (Viaro; Guimarães-Filho, 2007). Apesar dessa previsibilidade, há muitos tópicos controversos sobre a representação da fonologia do Português. Um deles é a coda, quando se trata da estrutura silábica.

Há três análises principais para a estrutura silábica e para a determinação dos princípios de silabificação no PB dentro de um quadro gerativo: Collischonn (1997), que segue a proposta em *templates* de Ito (1986); Bisol (1999), cuja abordagem é baseada na fonologia métrica; e Lee (1999), que segue a Teoria da Otimidade (TO). Apesar de suas diferenças, geralmente é assumido por esses autores que a sílaba é composta pelas subunidades Ataque e Rima, conforme Selkirk (1982): o primeiro é opcional e a última tem um Núcleo obrigatório e uma Coda opcional (Figura 1). Neste artigo, mencionaremos apenas as diferenças e/ou semelhanças em relação à composição segmental da coda, deixando de lado os detalhes sobre o processo de silabificação.

Figura 1 – Representação da sílaba (σ) no PB.

Parêntesis indicam elementos opcionais



Fonte: Elaboração dos autores com base em Selkirk (1982)

Collischonn (1997) e Bisol (1999) argumentam que a coda é composta por quatro segmentos consonantais, a saber, /R/, /L/, /N/, /S/, como em ‘carta’ [‘kar.tə] /kaRta/, ‘calma’ [‘kaw.mə] /kaLma/, ‘banco’ [‘bãᵛ.ko] /baNko/ e ‘pasta’ [‘pas.tə] /paSta/. Os

glides [j, w] também são assumidos como coda pelas autoras. No PB, os glides são combinados com vogais para formar ditongos, como em ‘leite’ [‘lej.tʃi] /leite/ e ‘pauta’ [‘paw.tə] /pauta/. Fonologicamente, segundo as autoras, seriam derivados de vogais altas no processo de silabificação e mapeados como codas silábicas.

Por outro lado, Lee (1999) propõe que os glides não são mapeados como codas, mas como núcleo complexo. Além disso, /l/ não ocupa a coda, apenas o ataque. Essa posição sobre /l/ considera o fato de que esse segmento é produzido majoritariamente como [w] no PB, por exemplo, ‘calma’ [‘kaw.mə] /kaLma/. Poucos dialetos no Brasil produzem /l/ como uma líquida velarizada [ɫ] em posição pós-vocálica, e os que produzem são influenciados por línguas minoritárias de herança, principalmente presentes no sul do Brasil (Callou; Moraes; Leite, 2013 [1999]).

Em resumo, Collischonn (1997), Bisol (1999) e Lee (1999) concordam em alguns aspectos quanto à composição segmental da coda medial do PB, com exceção dos casos de /l/ e glides. É importante notar que esses autores assumem uma abordagem bifonêmica para o caso da nasalidade. Essa visão não é unânime entre outros autores, como Couto (1978), Callou e Leite (1990) e Cristófaros-Silva (1999), que entendem que a nasalidade é representada de forma monofonêmica, ou seja, com vogais nasais, como em ‘banco’ [‘bã.ko] /bãku/. Assim, não se teria uma consoante nasal na posição de coda, e o inventário vocálico se tornaria mais extenso no PB.

A visão bifonêmica é apoiada por Collischonn (1997), Bisol (1999) e Lee (1999), com base em alguns fatos linguísticos, a maioria deles observados primeiramente pelo pioneiro linguista Joaquim Camara Jr. (1969), que trabalhou em uma abordagem estruturalista e também defendeu uma estrutura bifonêmica para a nasalidade no PB:

- (i) no PB, sílabas travadas são seguidas pelo r-forte /χ/, como em ‘Israel’ [iz.χa.‘ew] /iSχael/, ‘guelra’ [‘gew.χa] /gelχa/ e ‘honra’ [‘õŋ.χə] /oNχa/, enquanto sílabas abertas são seguidas pelo r-fraco /r/, como em ‘aurora’ [aw.‘rõ.rə] /aurõra/;
- (ii) nas derivações de formas terminadas em uma vogal nasal tônica, o segmento nasal emerge, como em ‘irmã’ [ir.mã] /irmaN/ → ‘irmanar’;
- (iii) o prefixo ‘in-’ perde o segmento nasal antes de laterais e do r-forte por assimilação, emergindo como uma vogal oral, como em ‘ilegal’ [i.le.‘gaw] /ilegal/ e ‘irregular’ [i.χe.gu.‘lar] /iχegulaR/, em contraste com ‘infeliz’ e ‘inabalável’, casos em que a nasal é mantida e produzida;
- (iv) o processo de sândi é bloqueado entre uma vogal nasalizada e uma vogal oral, como em ‘sótão antigo’ vs. *sótântigo/sótontigo, diferente de outras instâncias similares no PB onde o sândi ocorre – fronteira de palavra, sílabas átonas, p. ex. ‘casa amarela’ → [kazama’re]la].

Vale mencionar que a coda medial no PB é, como em muitas línguas, um *locus* de variabilidade. Se considerarmos a proposta de Bisol (1999), cada um dos quatro segmentos consonantais que ocupam essa posição tem, pelo menos, duas possibilidades

diferentes de realização fonética, através das quais podemos observar processos de neutralização ou de alofonia. Por isso, essas consoantes são frequentemente representadas como arquifonemas.² Assim, dependendo do dialeto, /S/ pode ser realizado como [s], [z], [ʃ] ou [ʒ]; /R/ como [r], [h], [ɦ] e muitos outros; /L/ como [w] ou [ɭ], e /N/ como [Ń], [Ńn], [Ńm] e assim por diante (‘V’ significando ‘vogal’).

Na ortografia do PB, a escrita dessas quatro consoantes em posição de coda é bastante previsível, considerando que se está lidando com um nível mais abstrato de representação: as relações são entre *fonemas* e *grafemas*, não entre sons e letras. Note-se que a posição silábica por si só já é um tipo de condicionador ortográfico para a grafia de determinado segmento – *position conditioning* nos termos de Treiman e Kessler (2014). Assim, em outra palavra e/ou posição silábica, a ortografia das mesmas consoantes pode ser mais simples ou mais complexa, como é o caso da rótica e da fricativa no PB.

Na coda, a rótica apresenta uma relação de *um-para-um* (biunívoca), onde /R/ é representado apenas pelo grafema <r>, como em /pɔRta/ → <porta>. O mesmo é válido para a lateral, onde /L/ é sempre grafada com <l>, como em /baLde/ → <balde>.³ Para a nasal, há uma relação de *um-para-muitos* (múltipla), condicionada por um vínculo de *onset-para-coda* que ajuda a prever o grafema correto: /N/ é representado por <m> apenas quando precede uma consoante oclusiva labial, como em /kaNpu/ → <campo>; em qualquer outro caso, é grafado com <n>, como em /poNte/ → <ponte>. Finalmente, a fricativa tem uma relação *um-para-muitos* (múltipla), com alguma ajuda do condicionamento contextual, mas não de forma suficiente para dar total previsibilidade à escrita. Após qualquer vogal, exceto /e/ e /ɛ/, /S/ é escrito como <s>, como em /paSta/ → <pasta>. Se houver um /e/ ou /ɛ/, permanece uma competição arbitrária entre <s> e <x>, como em /eStudo/ → <estudo> e /eSplode/ → <explode>.

O que ocorre na aquisição da fonologia?

As crianças brasileiras tendem a seguir a seguinte ordem de aquisição em relação à estrutura silábica (Lamprecht *et al.*, 2004; Ferreira-Gonçalves; Freitas, 2016):

$$(1) CV \sim V < CVG < CVC(C) < CCV(C)(C)$$

A coda, então, é a penúltima estrutura adquirida, seguida apenas pelo *onset* complexo. As idades de aquisição podem variar de acordo com o segmento específico que ocupa cada estrutura silábica. De modo geral, as estratégias de reparo mais recorrentes relatadas no percurso de aquisição da coda são a omissão da consoante, como

² Embora o arquifonema reflita uma abordagem estruturalista quando a neutralização ocorre no sistema, ele ainda é utilizado como uma forma de representar segmentos na gramática (léxico) que não estão especificados até que haja um processo fonológico na derivação para definir sua forma de saída.

³ Embora a realização fonética de /l/ como [w] no BP possa causar ambiguidade, devido à sua semelhança resultante com os ditongos decrescentes, como em /kalda/ [‘kawda] ‘calda’ e /kauda/ [‘kawda] ‘cauda’.

em [‘siŋ.ku] → [‘si.ku] ‘cinco’; a coalescência, como em [‘kaw.sə] → [‘kɔ.sə] ‘calça’; a metátese, como em [bor.bo.‘le.tə] → [bo.bor.‘le.tə] ‘borboleta’, e a semivocalização da consoante, como em [‘pɔr.ta] → [‘pɔj.tə] ‘porta’. A ordem de aquisição dos diferentes segmentos em corresponderia ao seguinte:

(2) nasal (2:00) < lateral (2:2) < fricativa (2:8) < rótica (3:10)

Esta ordem considera uma média entre a presença desses segmentos na posição medial e final da palavra (Lamprecht *et al.*, 2004; Mezzomo, 2004; Ceron *et al.*, 2022 – amostras transversais; Mezzomo *et al.*, 2010 – comparação entre amostras transversais e longitudinais). Há alternâncias relatadas para nasal e lateral dependendo do estudo e da posição na palavra considerados, mas a tendência observada é que ambos sejam adquiridos mais cedo do que fricativa e rótica. A rótica é unânime em ser a última a ser adquirida.

Também é importante notar que essa ordem reflete a visão de que as crianças brasileiras adquirem quatro codas consonantais, incluindo nasal e lateral, e excluindo glides. Hernandorena (1990) e Bonilha (2005) são autoras que questionam se ocorre a aquisição de uma coda lateral, que parece ser interpretada pelas crianças como um glide que forma um ditongo decrescente com a vogal precedente. Essa hipótese é sugerida devido a três fatos principais:

- (i) a precocidade observada na aquisição do suposto segmento lateral em comparação com as outras consoantes da classe das líquidas nas diferentes posições que podem ocupar na estrutura silábica do PB;
- (ii) a precocidade observada na aquisição da lateral em relação às outras consoantes que compõem a coda, especialmente /S/ e /R/;
- (iii) e a coincidência entre a idade de aquisição do lateral, em média, com a idade de aquisição observada para os outros ditongos decrescentes na língua, que são compostos por glides e são adquiridos mais cedo (Bonilha, 2005; Ferreira-Gonçalves; Freitas, 2016; ver a ordem em (1)).

Além disso, Miranda (2009, 2019) e Pachalski (2020) chamam a atenção para o fato de que o *input* fonético das crianças é um glide [w] na maioria dos dialetos brasileiros, como em [‘paw.ku] /paLko/ ‘palco’ (ver seção seguinte para mais detalhes). Segundo as autoras, o contato com a forma consonantal geralmente ocorre por meio do processo de alfabetização, que proporcionará um *input* visual conflitante. Esses fatos tornam ainda mais difícil assumir que há uma representação lateral no sistema das crianças. Assim, o que Hernandorena (1990), Bonilha (2005), Miranda (2009, 2019) e Pachalski (2020) propõem como representação fonológica seria a de um glide em um núcleo ramificado, alinhando-se com outras análises para o sistema adulto, como as de Camara Jr. (1969) e Lee (1999), mencionadas na seção anterior.

Miranda (2009, 2019) e Pachalski (2020) fazem uma alegação semelhante em relação à representação da nasalidade. As autoras sustentam que as crianças brasileiras não adquirem uma consoante nasal na posição de coda; em vez disso, adquirem vogais nasais. Nesse sentido, apontam para fatos análogos ao caso da lateral:

- (i) a precocidade observada na aquisição do suposto segmento nasal em relação às outras consoantes que compõem a coda, especialmente /S/ e /R/;
- (ii) a proximidade da aquisição da nasal à conclusão do inventário de vogais (orais) (Rangel, 2002; Lamprecht *et al.*, 2004; Ferreira-Gonçalves; Freitas, 2016).

Neste caso, haveria uma assimetria entre a fonologia infantil e o que é descrito para a fonologia adulta no PB, especialmente sob a análise de Bisol (1999). Se as consoantes laterais e nasais não fazem parte do paradigma de coda das crianças, como podem fazer parte do sistema adulto? Uma possível solução para essa questão é considerar o que ocorre na aquisição da escrita, um caminho seguido por Miranda e Pachalski, e o tópico da próxima seção.

Mudança representacional por meio da aquisição da escrita

Miranda (2019) analisa a grafia do que seriam as quatro consoantes da coda medial no BP, de acordo com Bisol (1999). Seus dados foram extraídos de 204 textos espontâneos produzidos por crianças dos primeiros quatro anos de uma escola pública e uma escola particular e foram analisados descritivamente. Em resumo, a autora encontrou mais erros na grafia dos segmentos lateral e nasal em comparação com a fricativa e rótica, tanto na escola pública quanto na particular. Miranda (2019) também aponta uma diminuição nos erros à medida que as séries escolares avançam, uma mudança que ocorre de forma mais tímida na escola pública.

Pachalski (2020) encontrou resultados muito semelhantes em uma análise estatística de dados extraídos de textos espontâneos produzidos por 280 crianças nos primeiros cinco anos do Ensino Fundamental de uma escola pública no Rio Grande do Sul. Seus resultados mostraram haver significativamente mais erros na grafia de lateral e nasal do que na grafia de rótica e fricativa. Não houve diferença significativa entre a frequência de erros ao se comparar, entre si, os pares lateral e nasal, e rótica e fricativa. Outro resultado interessante foi que a omissão da consoante se apresentou como o tipo de erro mais frequente na grafia da nasal, e a substituição de consoante foi mais frequente na grafia da lateral, coincidindo novamente com os resultados de Miranda, que também encontrou uma grande incidência de omissão da nasal. Com relação à variável ano escolar, os erros apresentaram a tendência de diminuir à medida que as séries avançam, mas sem significância estatística. Pachalski demonstrou, no entanto, que havia uma correlação negativa entre escolaridade e erros ortográficos *se* a natureza dos erros fosse considerada, caso em que seguiu a proposta de Miranda (2019). Pachalski, então,

encontrou que os erros de natureza *fonológica* foram mais frequentes nos primeiros três anos escolares (o chamado “ciclo de alfabetização”) e reduziram significativamente nos dois últimos anos (4º e 5º), dando lugar aos erros de natureza *ortográfica* (para uma explicação das categorias de erro, ver seção 4).

O primeiro ponto a notar é que os dados analisados por Miranda e Pachalski revelam um comportamento oposto ao que é observado na aquisição fonológica do PB, uma vez que se supõe que as codas nasal e lateral sejam adquiridas precocemente em relação à fricativa e à rótica. Se a hipótese de senso comum for assumida, ou seja, que a aquisição da escrita segue o mesmo roteiro da aquisição fonológica, o que se esperaria na escrita é que as crianças apresentassem mais erros na representação dos segmentos róticos e fricativos, porque esses segmentos são adquiridos mais tarde do que se infere, portanto, maior complexidade associada a eles.

De fato, Miranda e Pachalski interpretam essa assimetria como uma indicação de que laterais e nasais não são considerados como codas pelas crianças no início do desenvolvimento da escrita – consequentemente, também não o são durante a aquisição fonológica. No entanto, as autoras sugerem que, ao longo do curso da escolaridade e do desenvolvimento da ortografia pelas crianças, a representação da nasalidade e do glide em núcleo silábico pode ser reprocessada e, portanto, reestruturada como consoantes nasal e lateral em coda, respectivamente.

As autoras sustentam a posição de que esse tipo de *mudança representacional* é possível com base em dois pressupostos, também assumidos neste artigo:

- (i) a aquisição da linguagem é um processo de construção gradual de estruturas linguísticas (Kiparsky; Menn, 1977; Vihman; Croft, 2013). Como a aquisição da escrita é vista como uma parte integrante desse processo, embora tenha suas especificidades, é plausível supor que não há exatamente a criação de novas estruturas, mas sim algum tipo de remodelagem daquelas que já existem;
- (ii) o caráter flexível das representações que caracterizam os diferentes domínios cognitivos específicos da mente, razão pela qual é possível que os seres humanos tomem seu próprio conhecimento como objeto de atenção cognitiva, conforme sugerido por Karmiloff-Smith (1992).

De acordo com Karmiloff-Smith (1992), o desenvolvimento metacognitivo resulta de um processo cíclico e progressivo de redescrição representacional, que transforma informações implícitas *na* mente em conhecimento explícito *para* a mente. Assim, diferentes níveis de representação do conhecimento tomam lugar na cognição à medida que o processo de redescrição acontece, motivado tanto por aspectos endógenos quanto exógenos do indivíduo.⁴

⁴ Considerando os objetivos da nossa discussão, não descreveremos o Modelo de Redescrição Representacional (MRR) de Karmiloff-Smith em detalhes. É apenas necessário ter em mente pressupostos gerais do modelo: a possibilidade de redescrição representacional na mente e a gradualidade desse processo, que pode ocorrer em qualquer ponto do

Conforme aponta Frith (1998), o fato de se descobrir que a língua opera com unidades mínimas e abstratas, como os fonemas, modifica substancialmente o próprio processamento da linguagem, uma vez que o que antes era processado como uma unidade indivisível agora é processado de maneira analisada. Nesse sentido, a mudança representacional pode operar tanto sobre as estruturas linguísticas em si quanto sobre o grau de consciência que se tem em relação a essas estruturas.

Em resumo, quando Miranda e Pachalski levam em consideração os dados de escrita para a discussão, fazem isso a partir de um ponto de vista que considera a aquisição da escrita como parte integrante do processo de aquisição da linguagem. Quando as crianças compreendem o princípio alfabético, seu conhecimento fonológico é demandado, ou seja, elas precisam retomar sua representação linguística já estabelecida e utilizá-la de uma maneira nova. Em linha com a visão clássica de Saussure (2012 [1912]) e Lyons (1968), a escrita torna-se uma forma alternativa de atualização da língua. Assim, uma nova janela é aberta para analisar a linguagem, além da fala somente. A ideia de que a escrita pode influenciar a (re)estruturação das representações fonológicas é, portanto, crucial para o argumento desenvolvido neste artigo.⁵

Método

Sujeitos e amostra

Analisamos um total de 843 dados extraídos de palavras escritas em um ditado, produzidas por 166 crianças de 1º a 4º ano do Ensino Fundamental de uma escola pública da cidade de Pelotas-RS, com idades entre 6 e 10 anos.⁶ Todas eram falantes nativas do PB e apresentavam desenvolvimento típico de linguagem. Os dados foram coletados entre 2014 e 2015 e fazem parte do 7º estrato do BATALE⁷, um banco de dados criado e mantido pelo GEALE⁸.

desenvolvimento (meta)cognitivo. Em Miranda e Cunha (2013), e Pachalski e Miranda (2021), há uma discussão mais detalhada sobre o uso do MRR para explicar os dados de aquisição da escrita na abordagem aqui adotada.

⁵ Ver Veloso (2019) para uma discussão com base nessa perspectiva usando dados do português europeu, e Hamman e Colombo (2017) para uma proposta de formalização da relação fonologia-ortografia com base no modelo BiPhon-OT.

⁶ Os dados analisados neste estudo foram coletados entre 2014 e 2015, em contexto escolar, quando ainda não era prática consolidada nas áreas de Educação e Letras a submissão de pesquisas com produções escritas de crianças a Comitês de Ética em Pesquisa. Ainda assim, todos os procedimentos seguiram princípios éticos básicos: a participação foi autorizada pela escola e pelos responsáveis, garantiu-se o anonimato das crianças e não houve qualquer prejuízo a elas em decorrência da pesquisa.

⁷ Acrônimo para Banco de Textos de Aquisição da Linguagem Escrita.

⁸ Acrônimo para Grupo de Estudos sobre Aquisição da Linguagem Escrita.

Instrumento/estímulos

Os dados foram obtidos pela aplicação de um instrumento de coleta de dados chamado *ditado balanceado de palavras*, conforme proposto por Moojen (2009). Este instrumento foi construído pela autora com o objetivo de avaliar o conhecimento geral de ortografia das crianças brasileiras. A autora selecionou palavras em uso corrente que contemplassem as letras mais frequentes do português, além de submeter a lista à avaliação de um juiz linguístico. Originalmente, o ditado contém 50 palavras. Para evitar fadiga dos participantes, especialmente das crianças mais novas, decidiu-se reduzir o ditado para 25 palavras para sua aplicação no 1º e no 2º ano. Para a aplicação no 3º e no 4º ano, as 50 palavras originais foram mantidas, mas coletadas em duas seções separadas, com 25 palavras em cada. Não se sucederam mais de 2 semanas entre as duas seções. As palavras foram ditadas oralmente por membros do GEALE, os quais faziam parte do mesmo grupo dialetal das crianças. A aplicação foi feita nas salas de aula dos alunos, seguindo a divisão das turmas escolares. As palavras consideradas neste estudo foram: ‘cambalhota’, ‘fazenda’, ‘horror’, ‘explosão’, ‘desfile’ e ‘alguém’. Essas palavras foram selecionadas porque têm uma sílaba com coda consonantal, representando cada um dos segmentos possíveis de ocupar essa posição segundo Bisol (1999), e estão presentes em ambas as listas de palavras ditadas para 1º e 2º anos e para 3º e 4º anos.

Cabe mencionar uma limitação do instrumento utilizado. A lista original do ditado de Moojen (2009) contém um item com coda rótica medial (“exército”), mas este não foi aplicado às crianças do 1º e 2º ano. Para assegurar a comparabilidade entre os anos escolares, optou-se pela inclusão da palavra “horror” como representante da rótica, ainda que em posição final. Reconhecemos que essa decisão constitui uma restrição metodológica, decorrente do uso de um instrumento padronizado originalmente concebido para outra finalidade. No entanto, entende-se que é possível manter o item na análise, uma vez que a rótica apresenta comportamento fonológico estável em nomes (não verbos) independentemente da posição silábica, ao contrário do que se observa, por exemplo, para a nasalidade, cujas realizações e análises fonológicas variam consideravelmente conforme a posição na palavra. Além disso, note-se que foram incluídos nas análises dois itens lexicais com coda nasal e dois itens com coda fricativa, de modo a contemplar as diferentes opções ortográficas existentes para esses segmentos no sistema (<m>, <n>; <s>, <x>, respectivamente).

Organização dos dados e categorias de análise

Uma vez extraídas dos ditados as grafias das codas, os dados foram organizados em planilhas *xlsx* e classificados em duas grandes categorias: acertos e erros. Estes últimos foram classificados e analisados de acordo com sua natureza, definida nos

termos da proposta de Miranda (2020).⁹ A autora propõe serem três as principais fontes de conhecimento que motivam os erros ortográficos, a saber: fonologia, ortografia e fonografia. Este artigo foca nos erros de natureza fonológica, uma vez que os interesses de discussão estão relacionados aos efeitos da aquisição da escrita na representação fonológica. Portanto, apenas dados classificados nesta categoria serão abordados na seção de resultados, sendo as exceções explicitamente mencionadas. O Quadro 1 apresenta uma descrição e exemplos de erros das três principais categorias propostas por Miranda (2020).

Quadro 1 – Quadro descritivo das categorias de erros ortográficos utilizados neste estudo

Natureza do erro	<i>Fonológica</i>	<i>Ortográfica</i>	<i>Fonográfica</i>
Descrição	Relacionada à complexidade fonológica própria a determinadas unidades segmentais e prosódicas, e/ou a fenômenos da fala	Relacionada à complexidade ortográfica gerada pelas regras contextuais e arbitrárias do sistema ortográfico	Relacionada ao processamento fonema-grafema e/ou aos aspectos motor e mecânico da escrita (livre de complexidade fonológica e/ou ortográfica)
Principais tipos de erro	Determinados tipos de omissão, substituição, inserção e transposição de grafemas, atestados em processos fonológicos observados na aquisição ou variação da língua	Uso indevido de grafema que compete com outro(s) no mesmo contexto em regras contextuais e/ou arbitrárias	Uso indevido de grafema devido à semelhança ou maior complexidade do traçado da letra; sequenciamento, omissão e inserção de letras e sílabas não atestados em processos fonológicos observados na aquisição ou variação da língua
Exemplos de erros	horror → ‘orro’; alguém → ‘augei’; desfile → ‘disifile’; explosão → ‘sitrosão’	cambalhota → ‘canbalhota’; explosão → ‘esplosão’	horror → ‘orrete’; fazenda → ‘feda’

Fonte: Adaptado de Miranda (2020) e de Pachalski (2020)

É importante notar que a categoria fonológica considera a complexidade representacional de certas unidades melódicas e prosódicas da gramática que podem levantar dúvidas quanto à sua representação gráfica, embora as relações fonografêmicas

⁹ Estamos cientes de que existem várias propostas na literatura para a classificação de erros ortográficos, como aquelas relatadas e resumidas por Treiman e Kessler (2014). Como sua exploração excederia o limite de espaço que temos neste artigo, não discutiremos o tópico com mais detalhes.

sejam de um-para-um na maioria dos casos. Este é o caso de *onsets* e rimas ramificados, que têm uma representação fonológica complexa no nível da sílaba, mas que, com exceção de /N/ e /S/ em coda, possuem uma representação ortográfica simples. Normalmente, os tipos de erros incluídos nesta categoria assemelham-se a processos fonológicos relatados na aquisição fonológica ou mesmo na variação (cf. seções anteriores).

Também convém salientar que, quando mais de um erro na grafia da coda foi identificado na mesma palavra, cada um dos erros foi contado separadamente. Por exemplo, em *cambalhota* → ‘cã bada’, três diferentes erros na grafia da nasalidade foram computados: o uso do diacrítico para marcar a nasalidade, a omissão da consoante nasal e a hipersegmentação (uso de espaço em branco em um lugar não previsto pela norma ortográfica; para uma discussão sobre esse fenômeno no PB, cf. Cunha e Miranda, 2009). Neste caso, deve-se notar que consideramos erros que afetam não apenas a consoante, mas também a vogal, especialmente no registro de nasalidade. Esta abordagem é respaldada pelo pressuposto de que vogais (Núcleo) e consoantes finais (Coda) têm um comportamento solidário dentro da estrutura silábica, formando uma unidade coesa chamada Rima. Além disso, a nasalidade no PB apresenta uma característica de espriamento, geralmente afetando segmentos vizinhos, sobretudo vogais.

Perguntas e hipóteses

Três questões principais guiaram a análise deste artigo, acompanhadas de suas hipóteses:

- I. O tipo de segmento influencia a ocorrência de erros de natureza fonológica na posição de coda?
 - a. Sim, o tipo de segmento é um preditor para erros de natureza fonológica na posição de coda. Especificamente, /L/ e /N/ aumentam a probabilidade de erros em comparação com /R/ e /S/.
- II. O ano escolar influencia a ocorrência de erros de natureza fonológica na posição de coda?
 - a. Sim, o ano escolar é um preditor para erros de natureza fonológica na posição de coda. Quanto mais avançam os anos, menos se observam erros de natureza fonológica.
- III. O tipo de segmento influencia a qualidade dos erros de natureza fonológica na posição de coda?
 - a. Sim, o tipo de segmento é um preditor para a qualidade dos erros de natureza fonológica na posição de coda. Especificamente, a substituição é mais provável de ocorrer em /L/ e a omissão é mais provável de ocorrer em /N/.

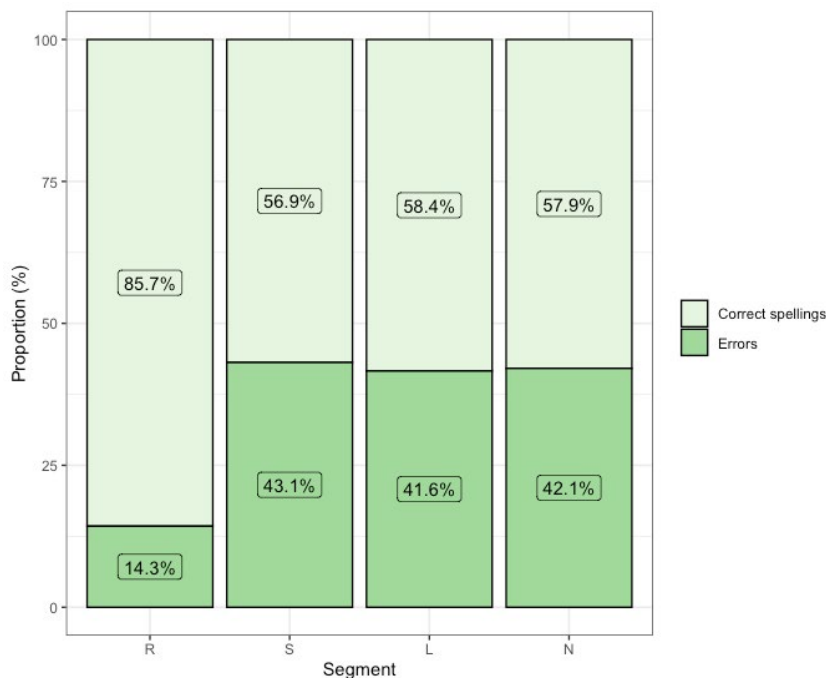
A seção de resultados que segue é organizada a partir dessas perguntas.

Resultados

Questões I – O tipo de segmento influencia a ocorrência de erros de natureza fonológica em posição de coda? – e II – O ano escolar influencia a ocorrência de erros de natureza fonológica em posição de coda?

A Figura 2 mostra a distribuição das grafias conforme acertos e erros encontrados na amostra, por tipo de segmento na coda. Note que a distribuição de erros inclui todas as categorias de erro: fonológico, ortográfico e fonográfico, conforme a proposta de Miranda (2020).

Figura 2 – Gráfico de barras da distribuição de acertos e erros nas grafias por tipo de segmento na coda (todas as categorias de erro)

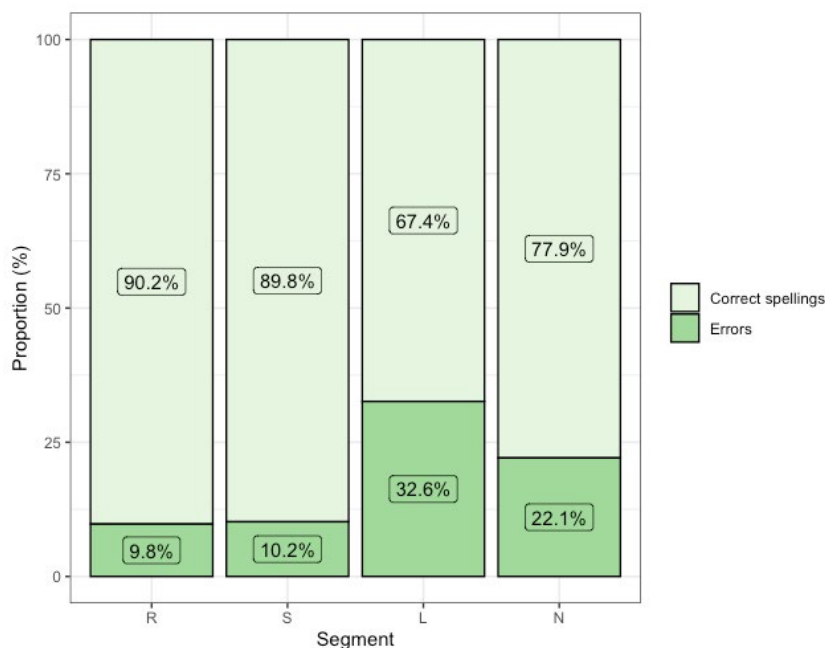


Fonte: Elaboração das autoras, com o pacote *ggplot2* no *RStudio* (Wickham, 2016).

Com um rápido olhar para o gráfico de barras, podemos perceber que a consoante rótica não parece ser tanto um foco de dúvida gráfica como parecem ser os outros segmentos em coda. No entanto, de acordo com a literatura relatada neste artigo, o resultado esperado era o de que as frequências mais altas de erro estivessem concentradas apenas em /L/ e /N/, não em /S/ também. Isso pode se dever a dois fatores. Primeiro, nossa amostra diz respeito a dados controlados. Esse tipo de dado força as crianças a escrever certos tipos de palavras que provavelmente evitariam em uma situação

espontânea. Isso porque tais itens apresentam um grau maior de complexidade em algum sentido – este é o caso, por exemplo, da escrita arbitrária de /S/ com <x> em ‘explosão’. Segundo, todas as categorias de erro estão sendo consideradas, o que torna a comparação mais opaca. De fato, se excluirmos os erros ortográficos e fonográficos e mantivermos apenas os fonológicos, o que faz sentido considerando os interesses de discussão, temos algumas diferenças importantes e uma distribuição mais similar ao que é encontrado na literatura:

Figura 3 – Gráfico de barras da distribuição de acertos e erros nas grafias por tipo de segmento em coda (apenas erros de natureza fonológica)



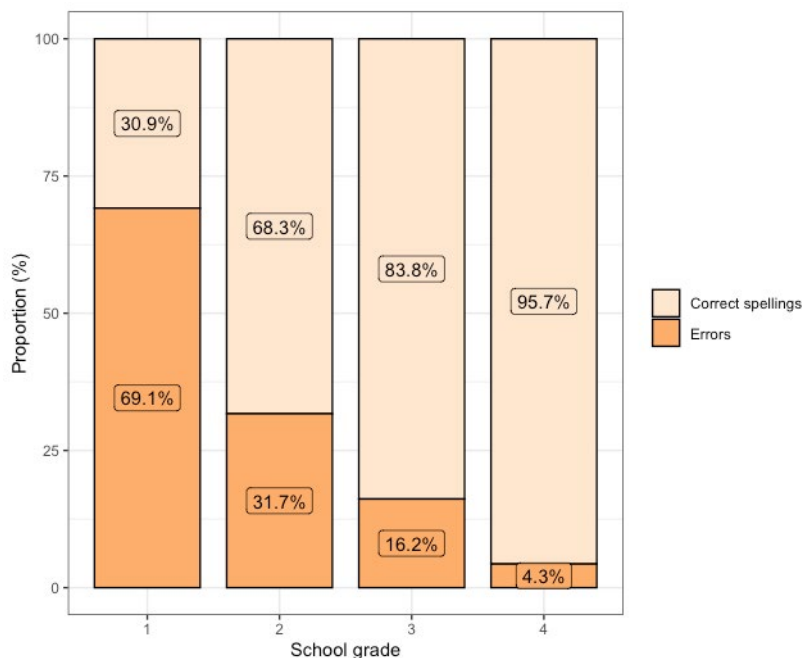
Fonte: Elaboração dos autores, com o pacote *ggplot2* no *RStudio* (Wickham, 2016)

Comparando apenas a complexidade fonológica envolvendo a grafia da coda, é possível observar outro comportamento dos dados: /R/ e /S/ concentram menos erros, formando um grupo de escrita menos complexo em comparação com /L/ e /N/. Isso agora está de acordo com o que Miranda (2019) e Pachalski (2020) encontraram em textos espontâneos e constitui a base da primeira hipótese para a modelagem estatística.

Uma segunda hipótese está relacionada ao papel do ano escolar. Conforme apresentado em seções anteriores, essa variável extralinguística foi testada e provou ser importante para a análise dos dados de escrita inicial, não apenas no que diz respeito às sílabas complexas, mas a várias outras estruturas do sistema do PB. A previsão é que os erros de escrita diminuam conforme as séries avançam, assumindo-se o papel

positivo que deve ter a escolarização no processo de aprendizagem. A Figura 4 exhibe a distribuição das grafias entre acertos e erros na posição de coda por ano escolar:

Figura 4 – Gráfico de barras da distribuição de acertos e erros nas grafias da coda por ano escolar (apenas erros de natureza fonológica)



Fonte: Elaboração dos autores, com o pacote *ggplot2* no *RStudio* (Wickham, 2016)

Uma análise dos dados apresentados nas Figuras 2 e 3 foi realizada com um modelo de efeitos mistos, tendo como efeitos aleatórios os participantes. Esta e outras análises de modelos mistos foram feitas no *software R* versão 4.2.3 (R Core Team, 2023) utilizando o pacote *lme4* (Bates *et al.*, 2013). Como a variável resposta é binária (acertos/erros), foi utilizada a distribuição binomial com função de *link logit*. Os fatores fixos foram segmento (R, S, L, N) e ano escolar (1º a 4º anos), ambos tratados como variáveis categóricas e não ordenadas, e sem interação. Três modelos alternativos também foram testados: um com interação entre segmento e ano escolar, um com segmento como único preditor e um com série escolar como único preditor. A Tabela 1 apresenta os resultados daquele que se mostrou como o melhor modelo misto testado, ou seja, aquele com segmento e ano escolar como fatores fixos, mas sem interação entre si:

Tabela 1 – Resultados da análise com o modelo de efeitos mistos: preditores ‘segmento’ + ‘ano escolar’ + efeitos aleatórios para ‘participantes’ para a VD ‘acerto/ erro nas grafias’ → VD ~ SEGMENTO + ANO.ESCOLAR (1 | PARTICIPANTE)

Preditores	<i>Odds Ratios</i>	Intervalo de confiança	p-valor	Acertos n (%)	Erros n (%)	Ocorrências totais n
Segmento /R/	–	–	–	114 (87%)	17 (13%)	131
Segmento /S/	0.73	0.31 – 1.74	0.477	161 (85.6%)	27 (14.4%)	188
Segmento /L/	12.08***	5.06 – 28.83	<0.001	80 (58.4%)	57 (41.6%)	137
Segmento /N/	2.64*	1.21 – 5.76	0.015	168 (71.5%)	67 (28.5%)	235
				523 (75.7%)	168 (24.3%)	691
Ano escolar 1	–	–	–	29 (30.9%)	65 (69.1%)	94
Ano escolar 2	0.06***	0.02 – 0.18	<0.001	125 (68.3%)	58 (31.7%)	183
Ano escolar 3	0.02***	0.01 – 0.07	<0.001	192 (83.8%)	37 (16.2%)	229
Ano escolar 4	0.00***	0.00 – 0.01	<0.001	177 (95.7%)	8 (4.32%)	185
				523 (75.7%)	168 (24.3%)	691
Efeitos aleatórios						
σ^2	3.29					
τ_{00}	3.00					
ICC	0.48					
N	138					
Observações	691					
R ² Marginal /	0.406 /					
R ² Condicional	0.689					
AIC	533.391					
* p<0.05 ** p<0.01 *** p<0.001						

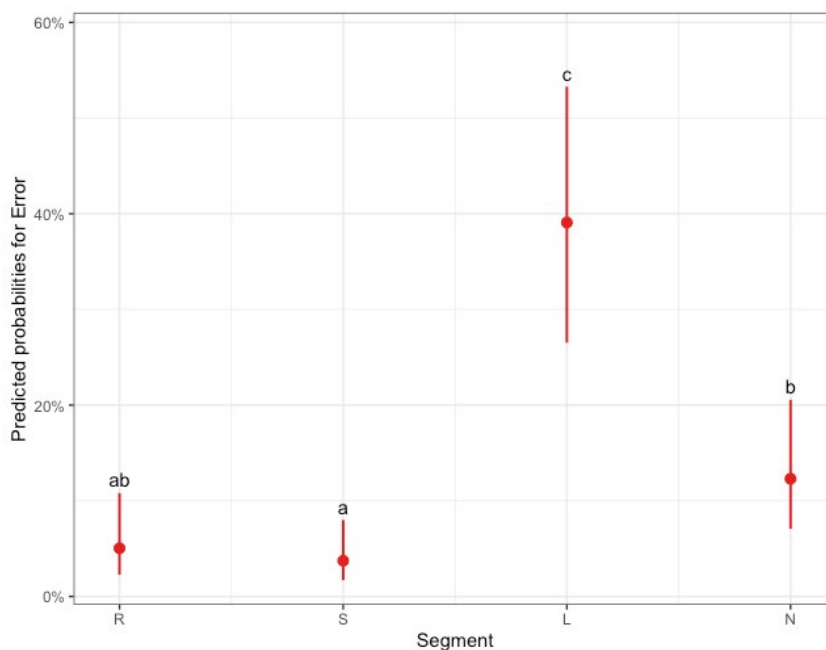
Fonte: Elaboração dos autores

Este modelo se mostrou mais explicativo em relação aos outros, pois seu valor de AIC é menor e seu valor de R² Condicional é maior. É interessante notar que o valor do ICC é menor no modelo apresentado também, revelando uma diminuição no poder dos efeitos aleatórios para explicar a variância nos dados quando ambos os preditores são considerados e, consequentemente, reforçando o poder dos efeitos fixos.

Considerando o efeito do segmento, é possível observar que /S/ não aumenta a chance de erros em relação à categoria de referência /R/. Por outro lado, /L/ e /N/ sim, com destaque para /L/, que aumenta em 12,08 vezes a chance de ocorrerem erros na grafia da coda, em relação a /R/. Quanto ao ano escolar, o modelo mostra que os 2º, 3º e 4º anos reduzem a chance de erros em comparação com o 1º ano.

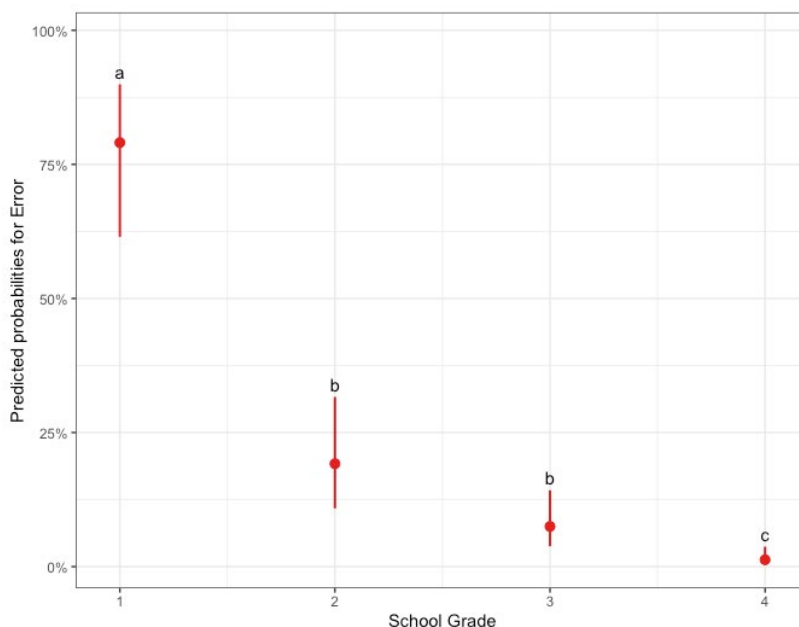
Comparações pareadas das médias marginais estimadas, com correção de *Tukey*, também foram realizadas para verificar as diferenças entre cada segmento e cada ano escolar. As diferenças podem ser visualizadas nas Figuras 5 e 6:

Figura 5 – Gráfico das probabilidades previstas para erros de natureza fonológica em posição de coda, por tipo de segmento. As diferenças podem ser verificadas pelas letras acima das barras. As linhas verticais correspondem ao intervalo de confiança de 95% para as probabilidades previstas. Letras diferentes indicam diferenças estatisticamente significativas entre os grupos de acordo com comparações pareadas das médias marginais estimadas com ajuste de *Tukey*



Fonte: Elaboração das autoras, com o pacote *sjPlot* no *RStudio* (Lüdeck, 2023)

Figura 6 – Gráfico das probabilidades previstas para erros de natureza fonológicas em posição de coda por série escolar. As diferenças podem ser verificadas pelas letras acima das barras. As linhas verticais correspondem ao intervalo de confiança de 95% para as probabilidades previstas. Letras diferentes indicam diferenças estatisticamente significativas entre os grupos de acordo com comparações pareadas das médias marginais estimadas com ajuste de *Tukey*



Fonte: Elaboração das autoras, com o pacote *sjPlot* no *RStudio* (Lüdeck, 2023)

Na Figura 5, é possível observar uma maior probabilidade de erro de /L/ em relação a qualquer outro segmento. /N/ ocupa o segundo lugar, compartilhando diferenças com /L/ e /S/, mas não com /R/. A nasal apresenta menor probabilidade de erros em relação a /L/, mas maiores probabilidades em relação a /S/. Por fim, /R/ e /S/ não apresentam diferenças entre si.

A Figura 6, por sua vez, mostra que o 1º ano apresenta maior probabilidade de erros do que todos os outros anos escolares, fato já apontado pela Tabela 1. O único par que não exibe diferença nas probabilidades previstas de erros é a comparação entre o 2º e o 3º ano. Todos os outros anos mostram diferenças entre si.

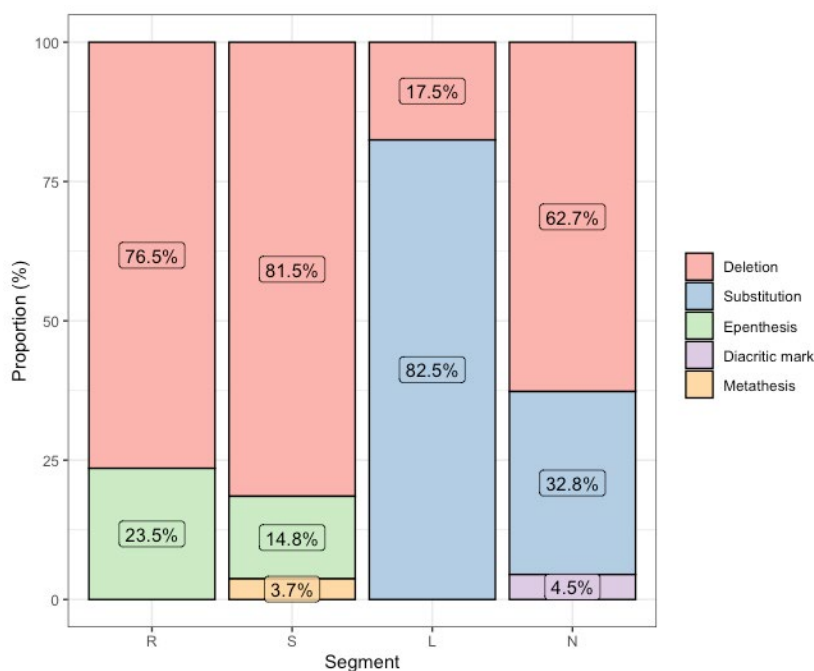
Com esses resultados, nossas duas primeiras questões são respondidas, corroborando as hipóteses iniciais: o tipo de segmento influencia os erros de natureza fonológica em posição de coda, onde /L/ e /N/ aumentam a probabilidade de erros em comparação a /R/ e /S/, e o ano escolar influencia a ocorrência de erros de natureza fonológica em posição de coda, sendo que quanto mais avançam as séries, menos erros ocorrem. É

importante ressaltar que, no modelo de efeitos mistos apresentado na Tabela 1, há uma diferença significativa verificada entre /R/ e /N/. Ao realizar as comparações pareadas entre os segmentos, essa diferença desaparece. Atribuímos isso a alguma característica específica de nossa amostra, já que esse tipo de diferença é apontado no modelo inicial e já foi verificado em trabalhos anteriores mencionados neste artigo.

Questão III – O tipo de segmento influencia a qualidade dos erros de natureza fonológica na posição de coda?

A Figura 7 exibe a distribuição dos erros de natureza fonológica em posição de coda por tipo de segmento. As análises com modelos de efeitos mistos não foram bem-sucedidas com esses dados, pois os modelos falharam em convergir e as alternativas se mostraram pobres em termos explicativos. No entanto, esse problema de convergência é esperado, dado o perfil dos dados: nem todos os segmentos apresentam todas as categorias de erro, e algumas categorias têm baixa frequência. Para essa questão, então, a análise será apenas descritiva.

Figura 7 – Gráfico de barras da distribuição dos diferentes tipos de erros de natureza fonológica na posição de coda



Fonte: Elaboração das autoras, com o pacote *ggplot2* no *RStudio* (Wickham, 2016)

O gráfico mostra que a omissão de consoantes é a estratégia predominante em todos os segmentos, exceto para /L/, segmento para o qual a substituição é notavelmente mais frequente (83,1%), como no dado ‘alguém’ → ‘agei’. A substituição também é observada na escrita de /N/ em uma quantidade considerável (33,8%), por exemplo, ‘cambalhota’ → ‘cãobaliota’. A epêntese figura em terceiro lugar, na escrita de /R/ e /S/, como em ‘horror’ → ‘ororo’ e ‘desfile’ → ‘disifile’. O uso de um diacrítico para representar a nasalidade aparece de forma episódica para /N/ – ‘cambalhota’ → ‘cã bada’ – assim como a metátese, que foi verificada apenas uma vez na escrita de /S/ – ‘explosão’ → ‘sitrosão’.

O Quadro 2 apresenta alguns exemplos das escritas extraídas diretamente dos ditados, que ilustram cada categoria de erro presente no gráfico da Figura 6.

Quadro 2 – Exemplos de erros de natureza fonológica presentes na amostra analisada

	Omissão	Substituição	Epêntese	Uso de diacrítico	Metátese
/R/	<u>ORRO</u>	–	<u>ORORO</u>	–	–
/S/	<u>DI F I L I</u>	–	<u>DISIFILI</u>	–	<u>SITROSÃO</u>
/L/	<u>AGAI</u>	<u>auguem</u>	–	–	–
/N/	<u>FAZEDA</u>	<u>CÃO BALIOTA</u>	–	<u>CA BADA</u>	–

Fonte: Elaboração das autoras com dados da pesquisa

É interessante comparar a qualidade da estratégia de substituição entre os dois contextos fonológicos em que ocorre, ou seja, entre as consoantes lateral e nasal. O *input* auditivo está possivelmente na base do índice observado. Neste caso, está em uma relação um-para-um com a organização da fonologia infantil, uma vez que a produção é semivocalizada – [w] – e a forma fonológica correspondente seria uma vogal alta que resulta em uma semivogal – /u/ → [w]. No entanto, embora haja menos casos de substituição para a grafia da nasalidade, há mais heterogeneidade na forma como a substituição pode se manifestar nesse contexto: seja como registro de ditongo nasal ou oral (por exemplo, ‘fazeida’, ‘cãobaliota’, conforme apresentado na tabela) ou como uma mudança na qualidade da vogal envolvida na sequência VN (por exemplo, ‘fazanda’).

Para Miranda (2009), a mudança na qualidade da vogal pode estar ligada à sendesibilidade da criança às distintas características fonéticas que as vogais assumem nesse contexto: sua produção nasalizada, que por si só já as distingue das vogais orais, também as aproxima perceptualmente de alguns segmentos vocálicos, especialmente /a/ e /e/ (Berti; Chacon; Santos, 2008), o que explicaria, por exemplo, a escrita de

<a> em vez de <e> em ‘fazanda’ para ‘fazenda’. A autora também estabeleceu uma comparação entre a incidência de erros na escrita da nasalidade de acordo com as diferentes qualidades das vogais envolvidas, através da qual constatou que as vogais /a/ e /e/ são as que mais favorecem erros, concentrando, juntas, 58% das ocorrências totais.

Embora a análise neste caso se baseie apenas em dados descritivos, o tipo de segmento parece, de fato, influenciar o tipo de erro de natureza fonológica encontrado na grafia da coda, pelo menos nesta amostra, em conformidade com outros estudos relatados (Miranda, 2009; Pachalski, 2020; Pachalski; Miranda, 2021). A tendência parece ser de uma predominância da omissão em todos os segmentos, com exceção de /L/. Este segmento mostra uma tendência para erros de substituição, que também estão presentes na escrita de /N/, embora em menor proporção.

Acreditamos que esses dados reforçam o comportamento distinto das crianças em relação à escrita dessas estruturas. Consoantes laterais e nasais não apenas concentram mais erros em comparação com consoantes róticas e fricativas, de forma geral, mas também diferem destas em termos da qualidade dos erros de natureza fonológica produzidos.

Discussão

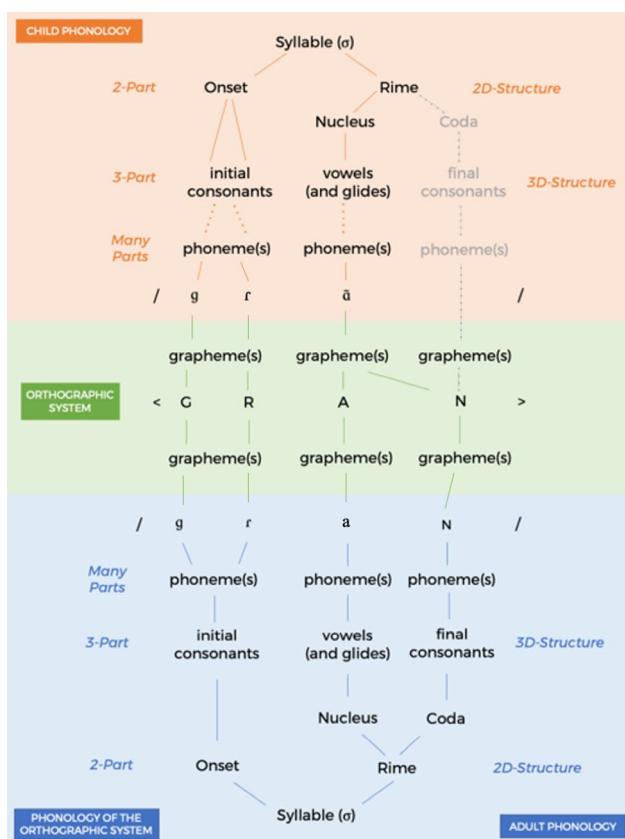
Os resultados apresentados poderiam ser interpretados a partir do efeito de outras variáveis, como *nome de letra*, à luz da análise de Treiman (1992) sobre a grafia da rima silábica por crianças falantes de inglês. Treiman mostra que os alunos de primeiro ano usam o conhecimento do nome das letras para grafar a rima *quando* a sequência VC de uma rima corresponde a um nome de letra. Assim, em ‘farm’ (*fazenda*), as crianças às vezes escreviam ‘frm’, ou em ‘helped’ (*ajudou*) o resultado era ‘hlpt’. No PB, no entanto, a rima em sequências CVC não corresponde aos nomes das letras, especialmente por causa das consoantes. A pronúncia das consoantes termina em vogais e segue um modelo V.CV – por exemplo, <s> [‘e.sɪ], <r> [‘e.χɪ], <l> [‘e.li], <n> [‘ẽ.nɪ]. Além disso, esta variável pode não explicar as diferenças encontradas nos erros ortográficos entre cada segmento. Por que as líquidas laterais e as nasais apresentam mais erros do que as róticas e fricativas se possuem o mesmo *template* fonológico para os respectivos nomes de letra? E por que a qualidade dos erros difere entre os segmentos?

Entendemos que essas questões são mais bem respondidas de um ponto de vista representacional, isto é, se seguirmos a perspectiva de que as crianças brasileiras não interpretam a líquida lateral e a consoante nasal como componentes da coda, e que essa interpretação adiciona um desafio particular às crianças quando iniciam a aquisição da escrita (Miranda, 2019; Pachalski, 2020). Isso ocorre porque o sistema ortográfico brasileiro parece representar esses segmentos refletindo o que é proposto para a fonologia adulta (escolarizada), considerando as análises de Bisol (1999) e Collischonn e Wetzels (2016). Para uma consoante nasal, há <n> ou <m>, como em ‘tampa’ /taNpa/ (não um diacrítico ‘~’ sobre a vogal como há para outros casos de nasalidade, como /

masaN/ ‘maçã’). Para uma consoante lateral, há o <l>, como em ‘palco’ /paLko/ (não um <u> como temos para casos de ditongos verdadeiros, como ‘caule’ /kaule/). Dado esse conflito entre a forma como as crianças representam esses segmentos específicos e a forma como o sistema ortográfico o faz, mais erros de natureza fonológica são registrados na escrita de <l> e de <m, n> quando comparados aos outros na mesma posição – <s, x> para a fricativa e <r> para a rótica, que não apresentam esse tipo de assimetria.

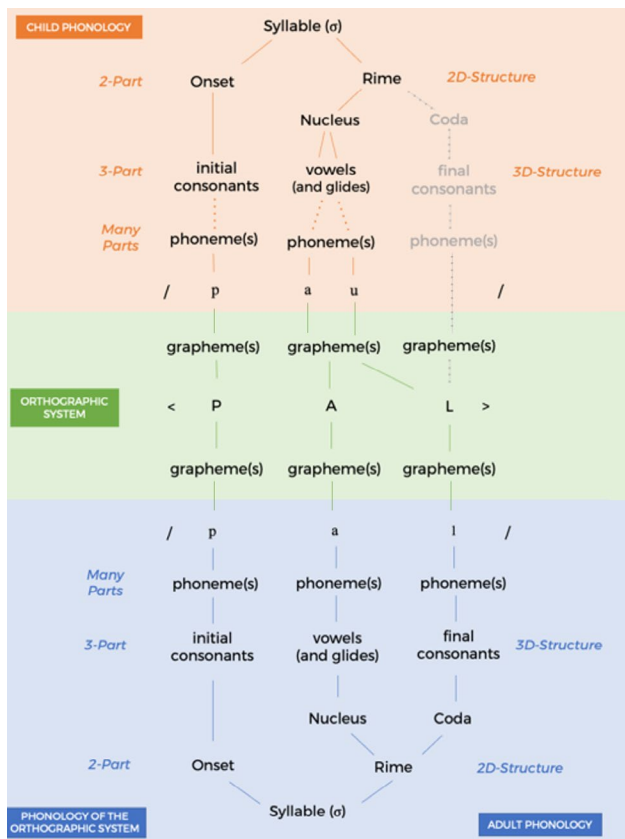
As Figuras 8 e 9 a seguir têm como objetivo formalizar e, conseqüentemente, ilustrar o conflito a respeito do qual argumentamos, a primeira para o caso da nasalidade e a segunda para o caso da lateral/(semi)vogal.

Figura 8 – Formalização do conflito entre a fonologia das crianças e a “fonologia do sistema ortográfico” (correspondente à fonologia adulta escolarizada), que resulta no processo de mudança representacional da nasalidade medial (exemplo com as sílabas /grã/ vs. /graN/ da palavra ‘grande’).



Fonte: A representação da estrutura silábica é adaptada de Seymour (1997). Figura original de Pachalski (2020), adaptado com permissão

Figura 9 – Formalização do conflito entre a fonologia das crianças e a “fonologia do sistema ortográfico” (correspondente à fonologia adulta escolarizada), que resulta no processo de mudança representacional da semivogal, derivada de uma vogal alta, em um núcleo ramificado (exemplo com as sílabas /pau/ vs. /pal/ da palavra ‘palco’)



Fonte: A representação da estrutura silábica é adaptada de Seymour (1997). Figura original de Pachalski (2020), adaptado com permissão

As Figuras 8 e 9 estão estruturadas da mesma forma: no topo, em laranja, está o que sugerimos ser a representação da sílaba na fonologia infantil, dominando uma determinada sequência de segmentos. No meio, em verde, está a representação alfabética/ ortográfica da sequência de fonemas. Abaixo, em azul, está a representação da sílaba na fonologia adulta, que corresponde ao que chamamos de “fonologia do sistema ortográfico”. Com essa expressão, queremos nos referir ao fato de que todo sistema ortográfico se baseia em um sistema fonológico, já que estabelece representações entre fonemas e grafemas. E, geralmente, o sistema fonológico representado por um sistema ortográfico corresponde às análises propostas para as fonologias adultas.

Na Figura 8, consideramos o caso da nasalidade. Em nossa visão, o que é representado na fonologia infantil é uma vogal nasal, que ocupa o núcleo silábico, como já afirmado por Miranda (2019) e Pachalski (2020). Não há coda (ainda), e é por isso que essa posição está colorida em cinza na figura. Essa representação conflita com a informação que a criança recebe do sistema ortográfico, que representa este caso particular de nasalidade com uma sequência de dois grafemas – <a, e, i, o, u> seguido de <n, m> – indicando que está lidando com uma representação bifonêmica formada por uma vogal oral + consoante nasal (VN), a mesma da fonologia adulta.

Isso explicaria por que a omissão de <n, m> é o tipo de erro mais frequente registrado na escrita da coda nasal. Neste caso, a dúvida da criança seria do tipo “*existe algo na posição pós-vocálica?*”. Essa proposição está ligada à forma como os erros ortográficos são interpretados, especialmente os de natureza fonológica: segue uma visão clássica da criança como “solucionadora de problemas” (*task solver*) ao passar pelo processo de aquisição da escrita ou mesmo da fonologia (Kiparsky; Menn, 1977; Ferreiro; Teberosky, 1984; Vihman; Croft, 2013). Nesse sentido, a criança é ativa no processo de desenvolvimento e formula hipóteses para resolver um problema. No caso discutido aqui, trata-se de um problema específico, que tomaria a forma da dúvida sugerida anteriormente. Em outras palavras, trata-se de um problema representacional.

A Figura 9 apresenta o caso da líquida lateral. O que é representado na fonologia infantil é um ditongo formado por duas vogais, como é afirmado por Hernandorena (1990), Bonilha (2005), Miranda (2019) e Pachalski (2020). Essas vogais ocupam um núcleo silábico ramificado e uma delas se tornará uma semivogal no processo de silabificação. Aqui, também não há coda (ainda), mas a diferença importante é que *há algo na posição pós-vocálica*. Essa representação conflita com a informação que a criança recebe do sistema ortográfico, o qual representa essa sequência com dois grafemas – <a, e, i, o, u> seguido de <l> – indicando que está categorizando o segmento pós-vocálico como consoante lateral (a mesma da fonologia adulta escolarizada). Em nossa interpretação, é por isso que a substituição de <l> por <u> é o tipo de erro mais frequente registrado na escrita da lateral. Nesse caso, a dúvida da criança poderia ser formulada como “*o que é esse algo na posição pós-vocálica?*”. Novamente, a proposição de que há tipo de dúvida tem amparo na forma como abordamos os erros ortográficos neste artigo, especialmente os de natureza fonológica, seguindo a visão da criança como solucionadora de problemas e ativa no processo de desenvolvimento.

Esse tipo de conflito cria um contexto para que as crianças trabalhem as suas próprias representações fonológicas. Treiman (1992) já mostrou que unidades intrassilábicas desempenham um papel significativo na leitura e na escrita do inglês americano. Como essas unidades funcionam de forma coesa, as crianças têm mais dificuldade em analisar *onsets* e rimas em seus componentes fonêmicos, o que aumenta a probabilidade de erros na grafia dessas unidades. À medida que esse conhecimento é explicitado, os erros vão diminuindo. Entendemos que esse fato também subjaz aos nossos dados. Contudo, além disso, o que propomos aqui é que o conhecimento fonológico não é apenas explicitado

via aquisição da escrita, mas também pode ser reestruturado, no sentido do conceito de redescrição representacional de Karmiloff-Smith (1992).

Como revisado anteriormente neste artigo, esse processo cognitivo refere-se a uma mudança no formato das representações ao longo do desenvolvimento, à medida que elas exigem atenção cognitiva ou acesso consciente. No caso que discutimos, à medida que a criança tem acesso a novas formas de representar os segmentos na rima ramificada, através da aquisição do sistema alfabético-ortográfico, com um grafema consonantal correspondente a um fonema consonantal, é exigida atenção às estruturas fonêmicas e silábicas envolvidas, já que há um conflito entre sua estrutura representacional e aquela proposta pelo sistema ortográfico. Essa demanda por atenção pode, assim, desencadear um processo de redescrição representacional, fazendo com que a representação da nasalidade e do ditongo no sistema fonológico infantil seja redescrita de acordo com o modelo apresentado, isto é, de acordo com a forma de categorização fonológica do sistema ortográfico, que corresponde, em grande medida, à fonologia adulta.

Conclusão

Os resultados que reportamos neste estudo reforçam aqueles apresentados por Abaurre (1988), Miranda (2019) e Pachalski (2020). Acreditamos que eles lançam luz à discussão sobre a estrutura silábica do PB e a sua representação, não apenas na fonologia infantil, mas também na fonologia adulta. Nossa contribuição empírica diz respeito a trazer mais robustez aos resultados disponíveis na literatura através da modelagem estatística proposta. Também oferecemos uma forma de representar e formalizar esse fenômeno de reestruturação representacional especificamente do paradigma da coda silábica no desenvolvimento fonológico das crianças brasileiras. Além disso, os dados são de natureza controlada, diferente dos dados espontâneos analisados nos trabalhos das autoras referenciadas.

Por outro lado, nossos dados provêm de um instrumento que não foi construído especificamente para estudar a grafia da coda, o que pode limitar nossos resultados ou torná-los menos precisos. Pesquisas futuras devem prestar atenção a este ponto, fornecendo dados de um instrumento específico que enfoque critérios relevantes para analisar a grafia da coda no PB. Seria interessante também se os dados de escrita inicial pudessem ser comparados com dados de fala, não apenas de crianças, mas também de adultos, para obter mais material empírico referente à discussão sobre as assimetrias entre a fonologia infantil e adulta.

À luz de nossa discussão, acreditamos que, em uma sociedade grafocêntrica como a que vivemos atualmente, faz sentido ponderar que pode haver diferença entre aqueles que adquiriram a escrita e aqueles que não adquiriram, do ponto de vista da forma como o conhecimento linguístico é representado. Em nossa visão, é necessário considerar a aquisição da escrita como parte do processo mais amplo de aquisição da linguagem, bem como levar os dados de escrita em maior consideração nas discussões

sobre a fonologia da língua, para, assim, sustentar um caminho intermediário entre o observado binarismo de descontinuidade vs. continuidade entre gramáticas infantis e adultas (Vihman; Croft, 2013).

Agradecimentos

Esta pesquisa contou com o apoio das agências nacionais de fomento à pesquisa, (CAPES: Processos n. 88887.721461/2022-00, n. 88881.933716/2024-01 e n. 88887.937773/2024-00; CNPq: Processo n. 312387/2020-2). Agradecemos, também, a Lia Pachalski, Fernanda Peres, Lorenzo Richetti, Nathalia Reinehr e Yasmin de Campos pela colaboração nas diferentes etapas da elaboração deste manuscrito.

PACHALSKI, Lissa; MIRANDA, Ana Ruth Moresco. The role of writing acquisition in the representation of Brazilian Portuguese syllabic coda: from child to adult phonology. *Alfa*, São Paulo, v. 70, 2025.

- *ABSTRACT: The acquisition of writing has been considered an important source of change in the learner's mind/brain, especially in the way language is processed. There are several aspects of this kind of change that have been studied; however, few works have addressed changes specifically in phonological representation. This article seeks to provide that discussion, focusing on the representation of the syllabic coda in Brazilian Portuguese. To this end, we analyze spelling data from 166 Brazilian monolingual children, who study in the first four years of Elementary School. The results show that children have trouble when spelling the nasal and lateral consonants more than the fricative and the rhotic in this syllabic position. This is mainly attributed to an asymmetry between child and adult phonology, which is reflected in the orthographic system. Due to this conflict, we argue that the representation of the coda may have changed in the phonological system of Brazilian children, resulting in its alignment of the coda with the target grammar.*
- *KEYWORDS: Coda; Syllable; Brazilian Portuguese; Writing acquisition; Phonological acquisition.*

Contribuição dos autores (conforme taxonomia CRediT)

Lissa Pachalski: Conceituação; Curadoria de dados; Análise formal; Aquisição de financiamento; Investigação; Redação – rascunho original; Visualização; Administração do projeto; Metodologia.

Ana Ruth Moresco Miranda: Conceituação; Recursos; Supervisão; Redação – revisão e edição; Aquisição de financiamento.

Declaração de disponibilidade de dados

Todo o conjunto de dados que dá suporte aos resultados deste estudo está disponível em arquivo pessoal a ser solicitado ao(s) autor(es).

Referências

ABAURRE, M. B. M. A relação entre escrita espontânea e representações linguísticas subjacentes. **Verba Volant**, v. 2, n. 1, p. 167-200, 2011 [1988].

BATES, D.; MAECHLER, M.; BOLKER, B.; WALKER, S. **lme4**: Linear mixed-effects models using Eigen and S4. 2013. Disponível em: <http://cran.r-project.org/package=lme4>. Acesso em: 12 nov. 2025.

BERTI, L. C.; CHACON, L.; SANTOS, A. P. A escrita de /an/ por pré-escolares: pistas acústico-auditivas. **Cadernos de Educação**, v. 35, 2010. Disponível em: <https://periodicos.ufpel.edu.br/index.php/caduc/article/view/1620>. Acesso em: 12 nov. 2025.

BISOL, L. A sílaba e seus constituintes. In: NEVES, M. H. M. (org.). **Gramática do português falado**. v. 7. São Paulo: Humanitas/FFLCH/USP; Campinas: Editora da Unicamp, 1999.

BONILHA, G. F. G. **Aquisição fonológica do português brasileiro**: uma abordagem conexcionista da teoria da otimidade. 2005. Tese (Doutorado em Letras). Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2005. Disponível em: http://www.leffa.pro.br/tela4/Textos/Textos/Teses/giovana_bonilha.pdf. Acesso em: 12 nov. 2025.

CALLOU, D.; LEITE, Y. **Iniciação à Fonética e à Fonologia**. 5. ed. Rio de Janeiro: Jorge Zahar, 1990.

CALLOU, D.; MORAES, J. A.; LEITE, Y. Consoantes em coda silábica: /s, r, l/. In: ABAURRE, M. B. M. (ed.). **Gramática do português culto falado no Brasil**: a construção fonológica da palavra. São Paulo: Contexto, 2013. p. 167-194.

CAMARA JR., J. M. **Estrutura da língua portuguesa**. 23. ed. Petrópolis: Vozes, 1995 [1970].

CERON, M. I. *et al.* Aquisição segmental do Português Brasileiro: *onset* simples, complexo e coda. **CoDAS**, v. 34, n. 3, 2022. DOI: 10.1590/2317-1782/20212020439.

COLLISCHONN, G. **Análise prosódica da sílaba em Português**. 1997. Tese (Doutorado em Linguística) – Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1997.

COLLISCHONN, G.; WETZELS, W. L. Syllable Structure. *In*: WETZELS, W. L.; COSTA, J.; MENUZZI, S. (ed.). **The Handbook of Portuguese Linguistics**. Wiley Blackwell, 2016. p. 86-106.

COULMAS, F. **The writing systems of the world**. Oxford, UK: Blackwell, Repr, 1989.

COULMAS, F. **Writing systems**: an introduction to their linguistic analysis. Cambridge, U.K.; New York, NY: Cambridge University Press, 2003.

COUTO, H. H. **Vowel nasality**: case of Portuguese and Guarani. The Fifth Lacus Forum, Columbia, South Carolina, 1978.

CRISTÓFARO-SILVA, T. **Fonética e fonologia do português**: roteiro de estudos e guia de exercícios. São Paulo: Contexto, 1999.

CUNHA, A. P. N.; MIRANDA, A. R. M. A hipo e a hipersegmentação nos dados de aquisição de escrita: a influência da prosódia. **ALFA: Revista de Linguística**, São Paulo, v. 53, n. 1, 2009. Disponível em: <https://periodicos.flcar.unesp.br/alfa/article/view/1681>. Acesso em: 12 nov. 2025.

FERREIRA-GONÇALVES, G.; FREITAS, M. J. Acquisition of phonology. *In*: WETZELS, W. L.; COSTA, J.; MENUZZI, S. (ed.). **The handbook of Portuguese Linguistics**. Wiley Blackwell, 2016. p. 545-561.

FERREIRO, E.; TEBEROSKY, A. **Psicogênese da língua escrita**. Porto Alegre: Artes Médicas, 1999 [1984].

FRITH, U. Literally changing the brain. **Brain** **121**, v. 6, n. 6, p. 1011-1012, 1998. Disponível em: https://www.researchgate.net/publication/13635314_Literally_changing_the_brain. Acesso em: 12 nov. 2025.

GOMBERT, J. E. **Metalinguistic development**. Chicago: University of Chicago Press, 1992.

HAMANN, S.; COLOMBO, I. E. A formal account of the interaction of orthography and perception: English intervocalic consonants borrowed into Italian. **Nat Lang Linguist Theory**, v. 35, 2017, p. 683-714. Disponível em: <https://doi.org/10.1007/s11049-017-9362-3>. Acesso em: 12 nov. 2025.

HERNANDORENA, C. L. B. M. **Aquisição da fonologia do português**: estabelecimento de padrões com base em traços distintivos. 1990. Tese (Doutorado em Linguística) – Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1990.

ITO, J. **Syllable Theory in Prosodic Phonology**. Ph.D. Dissertation – University of Massachusetts, Amherst: [s.n.].

KARMILOFF-SMITH, A. **Más allá de la modularidad**: la ciencia cognitiva desde la perspectiva del desarrollo. Madrid: Alianza Editorial, S.A., 1992.

KIPARSKY, P.; MENN, L. On the acquisition of phonology. In: MACNAMARA, J. (org.). **Language learning and thought**. New York: Academic Press, 1977. p. 47-78.

LAMPRECHT, R. R. *et al.* (org.). **Aquisição fonológica do português**: perfil de desenvolvimento e subsídios para terapia. Porto Alegre: ArtMed, 2004.

LANDSMANN, L. T. **Aprendizagem da linguagem escrita**: processos evolutivos e implicações didáticas. São Paulo: Ática, 1995.

LIBERMAN, I. Y.; SHANKWEILER, D.; FISCHER, F. W.; CARTER, B. J. Explicit syllable and phoneme segmentation in the young child. **Journal of Experimental Child Psychology**, v. 18, p. 201-212, 1974. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.602.5825&rep=rep1&type=pdf>. Acesso em: 12 nov. 2025.

LYONS, J. **Introduction to Theoretical Linguistics**. London: Cambridge University Press, 1968.

MEZZOMO, C. L. **Aquisição da coda no português brasileiro**: uma análise via teoria de Princípios e Parâmetros. 2004. Tese (Doutorado em Letras) – Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2004.

MEZZOMO, C. L.; QUINTAS, V.G.; SAVOLDI, A.; BRUNO, L.B. Aquisição da coda: um estudo comparativo entre dados transversais e longitudinais. **Rev Soc Bras Fonoaudiol.**, 2010, v. 15, n. 3, pp. 401-407. DOI: <https://doi.org/10.1590/S1516-80342010000300015>.

MIRANDA, A. R. M. Aquisição da escrita: as pesquisas do GEALE. In: MIRANDA, A. R. M.; CUNHA, A. P. N.; DONICHT, G. (org.). **Estudos sobre Aquisição da Linguagem Escrita**. Pelotas: Editora UFPel, 2017. p. 15-50. Disponível em: <https://repositorio.ufpel.edu.br/handle/prefix/4391?show=full&locale-attribute=es>. Acesso em: 12 nov. 2025.

MIRANDA, A. R. M. Os dados de aquisição oral e escrita e o estatuto das codas mediais do português. In: FERREIRA-GONÇALVES, G.; KESKE-SOARES, M.; BRUM-DE-PAULA, M. R. **Estudos em aquisição fonológica**. 1. ed., v. 2. Santa Maria: Sociedade Vicente Pallotti, 2009. p. 111-130.

MIRANDA, A. R. M. As sílabas complexas: fonologia e aquisição da linguagem oral e escrita. **Fórum linguístico**, Florianópolis, v. 16, n. 2, p. 3825-3848, 2019. Disponível em: <https://periodicos.ufsc.br/index.php/forum/article/view/1984-8412.2019v16n2p3825/40613>. Acesso em: 12 nov. 2025.

MIRANDA, A. R. M. Um estudo sobre a natureza dos erros (orto)gráficos produzidos por crianças dos anos iniciais. **Educ. rev.** [online]. 2020, v. 36, e221615. Epub Jan 31, 2020. Disponível em: https://wp.ufpel.edu.br/geale/?page_id=1097. Acesso em: 12 nov. 2025.

MIRANDA, A. R. M.; CUNHA, A. P. N. Índícios de reestruturação do conhecimento fonológico da criança em dados de reparo na escrita inicial. **Letras de Hoje**, Porto Alegre, v. 48, n. 3, p. 343-354, jul./set. 2013. Disponível em: <https://revistaseletronicas.pucrs.br/fale/article/view/12627>. Acesso em: 12 nov. 2025.

MOOJEN, S. M. P. **Ditado Balanceado – Avaliação da escrita alfabético-ortográfica na clínica e na escola.** [S.l.] Editora Hogrefe, 2009.

MORAIS, J.; CARY, L.; ALEGRIA, J.; BERTELSON, P. Does awareness of speech as a sequence of phones arise spontaneously? **Cognition**, v. 7, p. 323-331, 1979. DOI: [https://doi.org/10.1016/0010-0277\(79\)90020-9](https://doi.org/10.1016/0010-0277(79)90020-9).

PACHALSKI, L. **A grafia de sílabas complexas na aquisição da escrita: relações entre fonologia e ortografia.** 2020. 195f. Dissertação (Mestrado em Letras) – Programa de Pós-Graduação em Letras, Centro de Letras e Comunicação, Universidade Federal de Pelotas, Pelotas, 2020. Disponível em: <https://guaiaca.ufpel.edu.br/handle/prefix/7458?show=full>. Acesso em: 12 nov. 2025.

PACHALSKI, L.; MIRANDA, A. R. M. Sílabas complexas na escrita de crianças dos anos iniciais: indícios sobre o acesso às estruturas intrassilábicas. **Linguagem & Ensino**, Pelotas, v. 24, n. 4, p. 868-892, out./dez. 2021. Disponível em: <https://periodicos.ufpel.edu.br/ojs2/index.php/rle/index>. Acesso em: 12 nov. 2025.

R CORE TEAM. **R: A language and environment for statistical computing.** R Foundation for Statistical Computing, Vienna, Austria, 2024. Disponível em: <https://www.R-project.org/>. Acesso em: 12 nov. 2025.

RANGEL, G. A. **Aquisição do sistema vocálico no português brasileiro.** 2002. Tese (Doutorado em Linguística) – Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2002.

SELKIRK, E. The syllable. In: HULST, H.; SMITH, N. (org.). **The structure of phonological representations.** v. 2. Dordrecht: Foris, 1982. p. 337-379.

SEYMOUR, P. H. K. Foundations of orthographic development. In: PERFETTI, C. A.; RIEBEN, L.; FAYOL, M. (org.). **Learning to spell.** New Jersey: Lawrence Erlbaum Associates, 1997. p. 319-337.

TREIMAN, R. The Role of Intrasyllabic Units in Learning to Read and Spell. *In*: GOUGH, P. B.; EHRI, L. C.; TREIMAN, R. **Reading Acquisition**. New Jersey, London: Lawrence Erlbaum Associates Publishers, 1992. p. 65-106. Disponível em: https://www.researchgate.net/publication/280301296_The_role_of_intrasyllabic_units_in_learning_to_read. Acesso em: 12 nov. 2025.

TREIMAN, R.; KESSLER, B. **How children learn to write words**. New York: Oxford University Press, 2014.

VELOSO, J. Reavaliando o estatuto silábico das seqüências Obstruinte+Lateral em Português europeu. **D.E.L.T.A.**, v. 22, n. 1, p. 127-158, 2006. Disponível em: https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502006000100005. Acesso em: 12 nov. 2025.

VELOSO, J. Phonology and writing: can we look at written productions to “see the unseeable” in phonology? **Loquens**, v. 6, n. 1, p. 2386-2637, 2019. DOI: <https://doi.org/10.3989/loquens.2019.059>.

VIARO, M. E.; GUIMARÃES-FILHO, Z. Análise quantitativa da frequência dos fonemas e estruturas silábicas portuguesas. **Estudos Linguísticos**, v. 36, n. 1, p. 27-36, 2007. Disponível em: <http://www.usp.br/gmhp/publ/Via32.pdf>. Acesso em: 12 nov. 2025.

VIHMAN, M. M.; CROFT, W. Phonological development: toward a “radical” templatic phonology. *In*: VIHMAN, M. M.; KEREN-PORTNOY, T. (ed.). **The emergence of phonology**: whole-word approaches and cross-linguistic evidence. London: Cambridge University Press, 2013. p. 17-57.

Recebido em 15 de julho de 2024

Aprovado em 16 de dezembro de 2024