

Sílabas complexas na escrita de crianças dos anos iniciais: indícios sobre o acesso às estruturas intrassilábicas

Lissa Pachalski¹

Universidade Federal de Pelotas, Pelotas, RS, Brasil

Ana Ruth Moresco Miranda²

Universidade Federal de Pelotas, Pelotas, RS, Brasil

Resumo: Este artigo trata do que seria a primeira e mais abrangente motivação para a ocorrência de erros (orto)gráficos em sílabas complexas na escrita inicial segundo o estudo de Pachalski (2020), cujo objetivo foi de descrever e analisar a grafia de *onset* e de rima ramificados em textos de crianças brasileiras do 1º ao 5º ano do Ensino Fundamental. A hipótese defendida é a de que a maioria dos erros observáveis em sílabas complexas seria decorrente da *explicitação incompleta da camada correspondente às ramificações silábicas*. O estudo parte da perspectiva de que a aquisição da escrita e da fala integram o processo mais amplo de aquisição da linguagem, de forma que relações simétricas e assimétricas podem ser observadas entre os dois processos. A discussão dos resultados baseia-se em uma análise quanti-qualitativa dos dados extraídos dos textos produzidos por crianças dos anos iniciais de escolarização.

Palavras-chave: *Onset* complexo; Coda medial; Aquisição da escrita; Fonologia e ortografia; Erro (orto)gráfico.

Title: Complex syllables on children's spelling: clues about access to intrasyllabic units

Abstract: This article discusses what would be the first and broader motivation for the occurrence of (ortho)graphic errors in complex syllables in the early writing, according to Pachalski (2020), whose objective was to describe and analyze the spelling of both complex onset and rhyme in texts written by Brazilian children from the 1st to the 5th grade of Elementary School. The hypothesis defended is that the majority of observable errors in complex syllables would be due to an *incomplete explicitation of the layer corresponding to the syllabic branches*. The study assumes that writing and speech acquisition are both part of the broader process of language acquisition; thus, symmetrical and asymmetrical relationships can be observed between the two processes. The results discussion is based on a quanti-qualitative analysis of the data extracted from texts produced by Elementary School children.

Keywords: Complex onset; Medial coda; Writing acquisition; Phonology and orthography; (Ortho)graphic error.

¹ Mestra em Letras pela Universidade Federal de Pelotas (UFPEL). Licenciada em Pedagogia pela UFPEL. Integrante do Grupo de Estudos sobre Aquisição da Linguagem Escrita (GEALE/CNPq/UFPEL). ORCID: <https://orcid.org/0000-0002-7288-8067>. E-mail: pachalskil@gmail.com.

² Doutora em Linguística e Letras pela Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS). Professora associada da Universidade Federal de Pelotas (UFPEL), atua nos Programas de Pós-Graduação em Educação e em Letras. Bolsista de Produtividade em Pesquisa do CNPq. ORCID: <https://orcid.org/0000-0002-1380-5751>. E-mail: anaruthmiranda@gmail.com.

Introdução

Quem acompanha crianças em fase de aquisição da escrita, provavelmente já se deparou com grafias tais como ‘tabem’ para ‘também’, ‘pecisa’ para ‘precisa’ ou ‘espolzão’ para ‘explosão’³, em dado momento desse percurso. Em um período que antecede a aquisição da escrita, o desenvolvimento da gramática fonológica, também é comum observar formas faladas como [bi’ka] para ‘brincar’ ou [se’veza] para ‘cerveja’⁴. Tais dados, tanto os referentes à aquisição da fala quanto à da escrita, compartilham uma característica: a alteração das estruturas silábicas.

Contudo, uma diferença importante decorre do fato de a gramática fonológica ser gradualmente *construída* pelas crianças ao longo do desenvolvimento linguístico, de modo que formas cada vez mais complexas vão sendo incorporadas e sua representação vai sendo definida por meio de movimentos de composição e de síntese (MACKEN, 1992, 1996; KIPARSKY; MENN, 1977). Na aquisição da escrita alfabética, o que se observa, no entanto, é a *retomada* desse conhecimento fonológico já internalizado e alinhado à gramática adulta – o movimento predominante, neste período, é de decomposição e de análise das unidades prosódicas e melódicas da língua construídas até então (MIRANDA, 2017). Por essa razão, é raro observar na fala de crianças em fase inicial de escolarização, por exemplo, formas como as apresentadas anteriormente – [bi’ka] ou [se’veza]. Na verdade, nessa etapa tais formas passam a ser consideradas indicativo de atraso no desenvolvimento fonológico (LAMPRECHT et al., 2004). Com base em tais considerações, surge a pergunta: se as crianças já adquiriram fonologicamente estruturas de sílaba com *onset* e rima ramificados, o que as motiva a omitir, inserir, substituir ou alterar a ordem dos segmentos dessas estruturas na escrita?

Neste artigo⁵ pretende-se chegar a uma resposta possível para a pergunta; mais especificamente, busca-se aquela que seria a primeira e mais abrangente motivação para a ocorrência de erros (orto)gráficos⁶ em sílabas complexas na escrita inicial. A pesquisa de Pachalski (2020), que focaliza a grafia de *onset* e de rima ramificados em textos de crianças brasileiras em fase de aquisição da escrita (1º a 5º ano do Ensino Fundamental), está na base deste estudo.

A hipótese defendida é a de que a maioria dos erros observáveis em sílabas complexas seria decorrente do que Pachalski (2020) denominou *explicitação incompleta da camada correspondente às ramificações silábicas* – o que significa dizer que erros em sílabas complexas não ocorrem por haver estruturas que precisam ainda ser incorporadas ao

³ Dados extraídos do Banco de Textos de Aquisição da Linguagem Escrita – BATALE (MIRANDA, 2001).

⁴ Dados extraídos de Matzenauer e Miranda (2012, p. 97-98).

⁵ Pesquisa desenvolvida com apoio das agências de fomento nacionais (processos 88882.424116/2019-01 da CAPES; 312387/2020-2 e 446882/2013-3 do CNPq).

⁶ O uso de parênteses tem como objetivo demarcar a diferença existente entre erros relacionados às regras do sistema ortográfico propriamente dito, os quais envolvem as relações múltiplas entre fonemas e grafemas, definidas contextual ou arbitrariamente, e aqueles produzidos na fase inicial do desenvolvimento da escrita, muitas vezes motivados por questões representacionais, isto é, referentes à fonologia da língua.

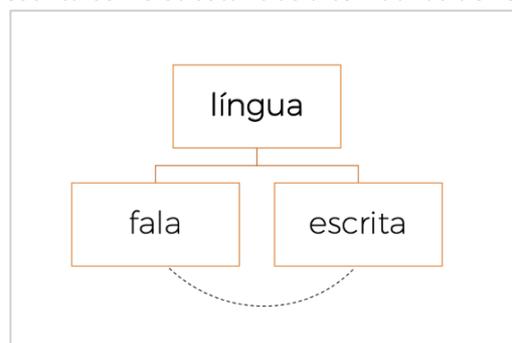
inventário da criança, pois isso já ocorreu durante o período de aquisição fonológica. Antes, ocorrem devido à incompletude existente em processos cognitivos de conversão da informação implícita para conhecimento explícito, isto é, um acesso consciente parcial à forma como se constituem internamente as estruturas silábicas (KARMILOFF-SMITH, 1994; TREIMAN, 1993; SEYMOUR, 1997). Nesse sentido, à medida que conhecimentos implícitos venham a ser explicitados, uma atividade epilinguística dá lugar a uma atividade metalinguística (GOMBERT, 1992).

O texto que segue está estruturado em 7 seções além desta introdução: na primeira, são definidas as concepções adotadas acerca da aquisição da fala e da escrita, bem como do conhecimento linguístico e de seu processamento; em seguida, considerações são feitas acerca da teoria e da aquisição da sílaba, com especial atenção para sua configuração no Português Brasileiro (PB); depois, são apresentados alguns estudos que versam sobre o desenvolvimento metacognitivo e o acesso consciente à estrutura fonológica; na seção 5, é revisada a configuração do sistema ortográfico do PB e descrita a maneira como os erros (orto)gráficos são neste estudo classificados, segundo sua natureza; adiante, os procedimentos metodológicos são descritos; na seção 7, os resultados são apresentados e discutidos, e, por fim, seguem-se as considerações finais.

Fala e escrita na aquisição da linguagem

Neste artigo, os processos de aquisição da fala e da escrita são compreendidos como partes integrantes do processo mais amplo de aquisição da linguagem (ABAURRE, [1988] 2011; MIRANDA, 2017). Essa concepção está ligada a uma noção anterior acerca de fala e de escrita como *realizações* distintas de um mesmo conhecimento abstrato, a língua:

Figura 1 – Fala e escrita como substâncias alternativas de realização da língua



Fonte: adaptado de Miranda (2020).

A ideia expressa pela Figura 1 é a de que a língua, um conhecimento abstrato e internalizado na mente humana, pode ser atualizada⁷ via substâncias distintas: os sons da fala e os grafismos da escrita. A primeira é a substância prioritária (LYONS, 1968; SAUSSURE,

⁷ O termo *atualização* faz referência à sua aceção linguística formal, ligada à ideia de *realização* ou *manifestação* de unidades linguísticas abstratas (fonemas, por exemplo) em uma substância física primária ou secundária (som ou escrita) (CRYSTAL, 1988; LYONS, 1968).

[1916] 2012), um sistema adquirido natural e espontaneamente pelo ser humano; a segunda é a substância secundária, um artefato cultural que exige ensino explícito para que seja adquirido. A similaridade entre as duas substâncias, expressa pelas linhas pontilhadas da figura, reside no fato de que elas realizam fisicamente um mesmo conhecimento.

A escrita alfabética, especificamente, compõe o grupo dos sistemas *fonográficos*, que operam a partir de unidades de segunda articulação da língua, não-significativas (MARTINET, 1960) – como o são as unidades constitutivas da camada sonora das línguas (traços, fonemas e sílabas). O nível fonêmico é o domínio de referência específico do sistema alfabético de escrita, cujo princípio de funcionamento são as relações entre fonemas e grafemas (GELB, 1952).

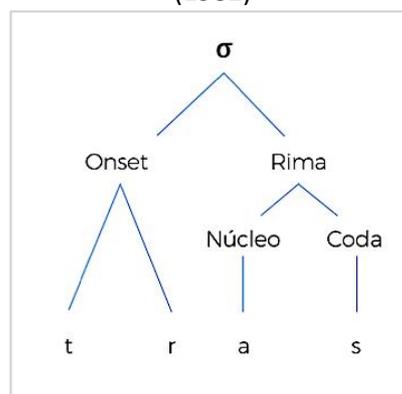
Assim, o sujeito que adquire um sistema alfabético tem a oportunidade de retomar o conhecimento fonológico internalizado, adquirido espontaneamente nos primeiros anos de vida, e de atualizar esse conhecimento em uma nova substância, razão pela qual se pode conceber a aquisição da escrita como parte do processo mais amplo de aquisição da linguagem. Em sendo retomado por via de uma substância distinta daquela que até então era a habitual, é como se o conhecimento linguístico passasse a ser compreendido a partir de um novo ponto de vista pela criança: não mais como unidades de primeira articulação apenas, mas também pelo viés da segunda articulação, o nível da estrutura fônica. Tal fato cria condições para que *mudanças representacionais* possam ocorrer (KARMILOFF-SMITH, 1994), tanto no que se refere às estruturas linguísticas em si quanto à consciência que se tem de tais estruturas. Como constata Frith (1998), o fato de descobrir que a linguagem opera com unidades mínimas não-significativas e abstratas tais como os fonemas modifica substancialmente o próprio processamento da linguagem, pois o que outrora era processado como unidade indivisível agora é processado de forma analítica, tornando-se disponível ao acesso e à manipulação conscientes.

Partindo dessa perspectiva, Miranda (2017) expressa a ideia de que o sujeito, a partir de princípios gerais constitutivos da Gramática Universal (CHOMSKY, 1988) em interação com o *input* recebido da sua comunidade linguística, constrói e organiza progressivamente seu conhecimento linguístico internalizado – mais especificamente, no caso que se está tratando neste texto, o conhecimento fonológico, que inclui todo tipo de informações melódicas e prosódicas, realizadas via fala, a substância primária (KIPARSKY; MENN, 1977; MACKEN, 1996; SAUSSURE, [1912] (2012); LYONS, 1968). Esse conhecimento é retomado no momento em que a criança compreende o domínio de referência estrito da escrita alfabética, a saber, a camada fonêmica da língua (LANDSMANN, 1995; GELB, 1952), criando condições para que: (i) a língua seja atualizada em uma substância alternativa, gráfica, e (ii) esse conhecimento seja disponibilizado *para a mente*, de modo que possa ser manipulado e, eventualmente, experimentar mudança representacional (KARMILOFF-SMITH, 1994).

Sílaba na teoria e na aquisição fonológica

Uma noção fundamental que guiou diferentes abordagens sobre a sílaba na teoria fonológica é a ideia de que a organização interna dessa unidade se dá em torno de um pico de soância, que, na maioria das línguas, corresponde a uma vogal; às margens desse pico ficam situados os segmentos menos soantes. Em uma abordagem métrica da sílaba, como a de Selkirk (1982), essa noção é transformada em uma hierarquia de constituintes, isto é, a camada melódica da língua projeta constituintes sub-silábicos, organizados de forma hierárquica e inter-relacionada, como mostra a Figura 2:

Figura 2 – Representação da estrutura hierárquica interna da sílaba /tras/, com base em Selkirk (1982)



Fonte: Pachalski (2020).

Assim, tem-se que a Sílaba (σ), nas línguas em geral, pode ser constituída de um *Onset* (O), não-obrigatório e opcionalmente ramificado; e de uma *Rima* (R), constituinte que pode se ramificar tendo uma *Coda* (Co), mas que tem o *Núcleo* (Nu) como obrigatório. O número de segmentos licenciados em uma sílaba e em seus constituintes mínimos fica a critério de cada língua em particular.

No PB, o licenciamento de segmentos na posição de *Coda* é tema controverso. Segundo a proposta de Bisol (1999), que segue Câmara Jr. [1970] (1995) especialmente no que se refere à nasalidade fonológica, os segmentos consonantais que podem ocupar a *Coda* medial são quatro: /S/, /N/, /R/ e /l/, os quais são realizados, majoritariamente e respectivamente, pelas contrapartes fonéticas [fricativa coronal], [nasal], [rótica] e [glide labiodorsal]. Já pela proposta de Costa e Freitas (2001), que analisam o Português Europeu (PE), não há uma consoante nasal em *Coda*, mas sim vogal nasal. Para Mateus e Andrade (2000), cuja análise também se volta ao PE, a nasalidade pode ser formalizada por meio de um autosegmento nasal flutuante que se liga ao *Núcleo*. Além disso, para os autores supracitados, os glides, que são derivados de vogais altas, ocupam a segunda posição do *Núcleo*, proposta que também é adotada por analistas do PB como Câmara Jr. [1970] (1995) e Lee (1999). Para Bisol (1989,

1999), os glides em ditongo decrescente têm estatuto de consoante e, portanto, ocupam a coda da sílaba.⁸

Já o licenciamento dos segmentos para o *onset* no PB é um ponto relativamente pacífico na literatura fonológica⁹. Na posição de *onset* simples, praticamente todas as consoantes são licenciadas, com algumas restrições no que se refere à posição na palavra, como é o caso das soantes palatais /ʎ/ e /ø/, que ocupam apenas a posição medial, como em ‘calha’ e ‘sonho’. Relativamente ao *onset* complexo, as possibilidades são reduzidas. Na primeira posição, são licenciadas as oclusivas e fricativas labiodentais, cujas presenças podem ser restringidas a depender da líquida – lateral ou não-lateral – que ocupará a segunda posição do *onset* (*dl e *vl), e também da posição do *onset* na palavra – inicial ou medial (*#tl e *#vr).

No tocante à aquisição da estrutura silábica, Lamprecht et al. (2004) apresentam um panorama descritivo útil para os fins da presente discussão. Em relação à rima ramificada, as autoras admitem que as crianças brasileiras adquirem os mesmos quatro segmentos que Bisol (1999) define para essa posição silábica no sistema adulto, na seguinte ordem: /N/, /l/, /S/ e /R/, sendo a idade média de estabilização da produção da rima ramificada aos 3 anos e 10 meses. A aquisição do *onset* complexo é mais tardia em relação à da rima, com média de estabilização em torno dos 5 anos; no entanto, não se observa diferença no tempo de aquisição entre os grupos consonantais, isto é, entre o *onset* formado por líquida lateral e o *onset* formado por líquida não-lateral. Diferentes tipos de estratégias são registrados para a produção de cada segmento na sua respectiva posição silábica; porém, a estratégia predominante na produção dos alvos com ambas as estruturas silábicas, em praticamente todos os segmentos, é a omissão, como em [‘kasa] para ‘calça’ e [‘tSigi] para ‘tigre’.

Acesso consciente¹⁰ à estrutura fonológica da língua

Conforme observado na seção 2, a aquisição da escrita demanda que as crianças acessem a estrutura fonológica da língua de forma relativamente consciente. Inúmeros estudos já foram realizados a respeito dessa relação, isto é, entre a aquisição da escrita e o

⁸ Neste trabalho, o foco das análises em rima ramificada incide sobre os segmentos que ocupam a posição medial na palavra, e glides em ditongos decrescentes não são analisados. Além disso, o leitor irá notar o uso deliberado do termo *rima ramificada*, opção feita devido às controvérsias existentes em torno do estatuto da coda no PB e, também, como é discutido no estudo completo de Pachalski (2020), às assimetrias existentes entre a fonologia infantil e a fonologia adulta em relação ao mesmo tema. Entende-se que o uso do termo *rima ramificada* permite transitar mais livremente sobre o assunto, sem a necessidade de se comprometer *a priori* com uma ou outra proposta específica para a composição da rima no PB.

⁹ Alguns autores sugerem que as sequências [kw] e [gw] formariam *onsets* complexos no PB (cf. HERNANDORENA, 1990; LAMPRECHT, 1990; LAMPRECHT; BONILHA, 2003; BONILHA, 2005). Neste estudo, tais sequências são consideradas como segmentos complexos, conforme propõe Bisol (1999), e, por isso, não são analisados.

¹⁰ É importante considerar que se está abordando a *consciência* de um ponto de vista não-dicotômico. Isto é, segue-se a ideia de Karmiloff-Smith (1994) de que o acesso consciente à informação implícita se dá de forma progressiva e não pode ser descrita apenas em termos da dicotomia *inconsciente* vs. *consciente*. Assim, na referência ao termo ‘consciência’, neste estudo, sempre está se levando em consideração a ideia de que o acesso consciente pode apresentar diferentes níveis ou gradações.

acesso consciente à fonologia. Um deles, considerado clássico e pioneiro sobre esse tema, é o de Liberman et al. (1974), que compara o grau de dificuldade na segmentação (i. e., abstração) de fonemas e de sílabas por crianças em fase inicial de escolarização. Os autores têm por hipótese, corroborada por estudo experimental, que a segmentação em fonemas é mais custosa e menos natural para as crianças do que a segmentação em sílabas.

Treiman (1992), inspirada no trabalho de Liberman et al. (op. cit.), analisa o papel das unidades intrassilábicas no desenvolvimento da consciência fonológica e, conseqüentemente, na aquisição da leitura e da escrita. A autora argumenta que, entre a consciência silábica e a consciência fonêmica, existem níveis intermediários de consciência fonológica correspondentes às unidades intrassilábicas. Sua argumentação sustenta-se em uma série de pesquisas por ela realizadas à luz de teorias não-lineares para o enfoque da sílaba. Segundo seu argumento, antes de as crianças analisarem as palavras em fonemas, elas analisariam as palavras em constituintes silábicos, *onset* e rima, de modo a estarem cientes de que, por exemplo, uma sílaba como /blo/, de 'bloqueio', começa com /bl/ e termina com /o/. No entanto, nesse nível, as crianças "não são capazes de analisar /bl/ como o fonema /b/ seguido do fonema /l/" (TREIMAN, 1992, p. 71).

De forma geral, os resultados apresentados por Treiman (1992) apontam que a habilidade de dividir sílabas faladas em *onset* e rima emerge antes de as crianças ingressarem em classes de alfabetização e que a estrutura silábica afeta o desempenho nas tarefas de segmentação de palavras em unidades menores, o que significa dizer que crianças pré-escolares têm dificuldade de analisar tanto a estrutura interna de um *onset* quanto de uma rima, quando ambos os constituintes são ramificados. Assim, indo ao encontro da conclusão mais abrangente de Liberman et al. (1974), Treiman (1992, p. 77, interpolações nossas) pontua que, "apesar de fonemas desempenharem um papel tácito na produção e na percepção da fala, eles não são acessíveis à consciência assim como são as unidades maiores [sílaba, *onset*, rima]".

No tocante à ideia mais ampla de desenvolvimento (meta)cognitivo, segue-se, neste artigo, a proposta de Karmiloff-Smith (1994) de que ele seja decorrente de um processo cíclico e progressivo de redescrição representacional, que vai transformando informações implícitas – *na* mente – em explícitas – *para* a mente. Assim, diferentes níveis de representação do conhecimento vão tendo lugar na cognição, à medida que o processo de redescrição vai acontecendo, motivado tanto por aspectos endógenos quanto exógenos ao indivíduo: "a redescrição é uma operação interna que preserva certas características e elimina outras em benefício de um nível diferente de acessibilidade" (LANDSMANN, p. 193, 1995).

No Modelo de Redescrição Representacional (MRR) proposto pela autora, quatro níveis são sugeridos, conforme o quadro a seguir:

Quadro 1 – Quadro descritivo do Modelo de Representação Representacional de Karmiloff-Smith (1994)

Nível	Descrição	Exemplo de conduta resultante
Implícito (I)	Representações inflexíveis e armazenadas de forma independente <i>na</i> mente (codificadas “entre parênteses”, ilegíveis para o sistema cognitivo); o conhecimento é procedimental e não se pode acessar suas partes componentes.	Conhecer e utilizar uma palavra como ‘ouvi’ (pret. perf. ind.), mas não ser capaz de decompor em ‘ouv- + -i’.
Explícito 1 (E1)	Redescrição do conhecimento procedimental; representações mais flexíveis e disponíveis a vínculos entre si dentro de um domínio cognitivo ¹¹ ; o conhecimento começa a se abrir <i>para a</i> mente, mas ainda não se pode falar em acesso consciente.	Generalizar o morfema ‘-i’ para outros contextos de pret. perf. ind. que não seguem a mesma tendência de afixação, como ‘sabi’ (para ‘soube’).
Explícito 2 (E2)	Redescrição do nível E1; representações de um domínio cognitivo disponíveis a vínculos com as de outros domínios cognitivos; acesso consciente ao conhecimento, mas ainda não verbalizável.	Voltar a utilizar os morfemas de verbos regulares e irregulares adequadamente, mas sem ser capaz de verbalizar no <i>quê</i> e por <i>quê</i> essas formas são semelhantes e diferentes, de modo a também justificar a mudança no comportamento.
Explícito 3 (E3)	Redescrição do nível E2; acesso consciente ao conhecimento e suas partes componentes, expresso por meio de verbalização – o conhecimento é “traduzível”.	Ser capaz de verbalizar que, para se referir ao passado, se utiliza geralmente, mas nem sempre, formas com ‘-i’ final.

Fonte: elaboração própria; exemplos adaptados de Lorandi e Karmiloff-Smith (2012).

O primeiro nível, Implícito (nível I), codifica o conhecimento em um formato que capacita o sujeito a executar determinados procedimentos, mas não de forma que se consiga acessar o conhecimento em termos de centração cognitiva; o segundo nível, Explícito 1 (nível E1), já comporta representações um pouco mais flexíveis, disponíveis ao sistema cognitivo, mas ainda não ao sujeito propriamente dito; já os níveis Explícito 2 (nível E2) e Explícito 3 (nível E3), são caracterizados pelo acesso consciente às representações, sendo o que os diferencia mormente é a capacidade de explicitar verbalmente os motivos que justificam a realização de determinadas condutas em detrimento de outras, característica do nível E3.

A natureza dos erros (orto)gráficos

No que diz respeito à representação ortográfica das consoantes do PB, existem casos nos quais as relações entre fonemas e grafemas são de um para um, isto é, um fonema será sempre representado pelo mesmo grafema. Em proporção um pouco maior, existem casos em que as relações entre fonemas e grafemas são de um para muitos: um mesmo fonema

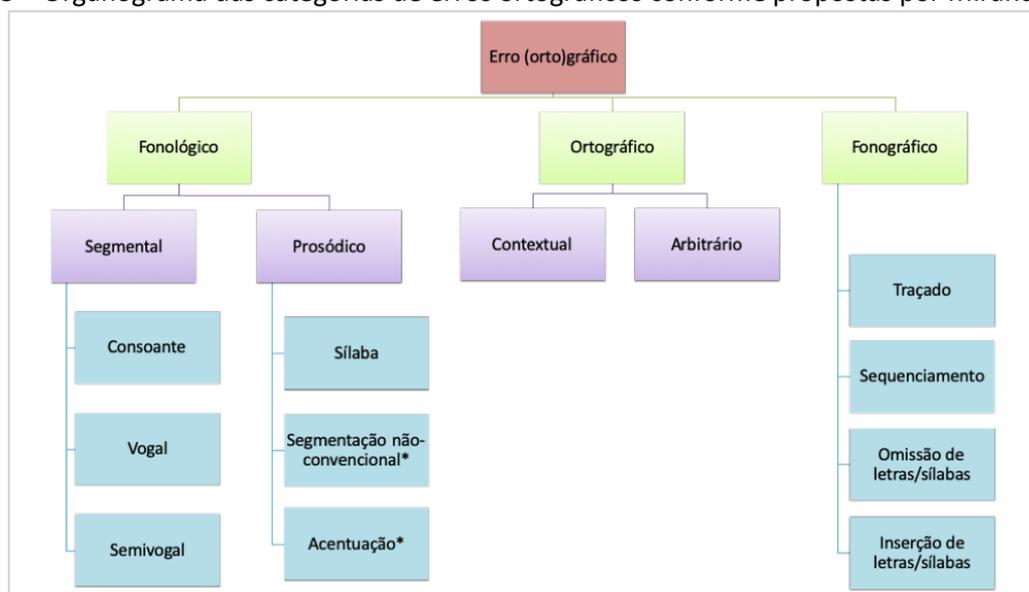
¹¹ “Da perspectiva da mente da criança, um ‘domínio’ é um conjunto de representações que sustenta uma área específica do conhecimento: linguagem, número, espaço, e assim por diante. Da mesma forma, existem microdomínios, como a aquisição pronominal, que pode ser pensada como um subconjunto dentro do domínio particular da linguagem. Por seu turno, um módulo consiste em uma unidade de processamento de informação que encapsula nele o conhecimento e a computação. Nesse sentido, considerar o desenvolvimento de domínio específico não significa modularidade, pois armazenar e processar informação pode ser de domínio específico sem ser encapsulado” (LORANDI; KARMILOFF-SMITH, 2012, p. 7, tradução nossa).

pode ser representado por mais de um grafema em diferentes contextos. Ao sistematizar as relações fonográfêmicas do PB, Lemle (1987) chama os primeiros casos de relações *biunívocas* e os segundos de relações *múltiplas*.

Considerando essa configuração da ortografia do PB, é possível classificá-la pelo critério da *profundidade ortográfica* (SEYMOUR; ARO; ERSKINE, 2003), definido pela maior ou menor complexidade das relações entre fonemas e grafemas, isto é, pela maior ou menor presença de relações múltiplas. Um sistema profundo ortograficamente é um sistema que apresenta *opacidade* na relação entre fonemas e grafemas ou, em outras palavras, as relações múltiplas prevalecem. Um sistema raso apresenta *transparência* nas relações fonográfêmicas, ou seja, mais relações biunívocas que múltiplas. Assim, em conformidade com Veloso (2005), que analisa a ortografia do PE em relação ao critério de profundidade ortográfica, é possível estipular o sistema do PB como sendo *relativamente transparente*, uma vez que apresenta relações biunívocas e múltiplas de forma mais ou menos equilibrada (SOARES, 2016).

Tendo em vista essa caracterização do sistema ortográfico, juntamente à concepção teórica mais abrangente apresentada na seção 2, é possível compreender propostas que visam à definição da natureza dos conhecimentos envolvidos no percurso de desenvolvimento ortográfico. Uma delas é a de Miranda (2020), cuja configuração atual pode ser visualizada na Figura 3 e por meio da qual é possível capturar os principais tipos de conhecimentos mobilizados pelas crianças nessa etapa de desenvolvimento, ao menos em um sistema ortográfico como o do PB, considerando seu grau de profundidade relativamente transparente:

Figura 3 – Organograma das categorias de erros ortográficos conforme propostas por Miranda (2020)



Fonte: acervo do Grupo de Estudos sobre Aquisição da Linguagem Escrita (GEALE).

O organograma procura representar a ideia de que os erros (orto)gráficos produzidos pelas crianças em processo de aquisição da escrita podem ser motivados por conhecimentos ligados a três fontes distintas: fonologia, ortografia e fonografia (MIRANDA, 2020). A primeira

categoria (*Fonológico*) diz respeito à complexidade de determinadas unidades melódicas e prosódicas da gramática fonológica, que pode provocar dúvida relativamente à sua representação gráfica, ainda que a relação fonografêmica seja biunívoca, na maioria dos casos. A segunda categoria (*Ortográfico*) relaciona-se à aprendizagem de regras ortográficas, definidas contextual ou arbitrariamente, casos em que as relações entre fonemas e grafemas são múltiplas e que, portanto, não apresentam motivação primária no sistema fonológico. A terceira (*Fonográfico*), por fim, se refere ao processamento das relações fonema/grafema, do ponto de vista do funcionamento da memória, da atenção, da motricidade e/ou da velocidade da codificação ou decodificação da informação.

Este artigo tem como foco a natureza fonológica dos erros, uma vez que enfatiza uma motivação para a ocorrência de erros em sílabas complexas que envolve o acesso consciente às estruturas fonológicas. *Onset* e rima ramificados têm representação fonológica complexa no nível da sílaba, e, à exceção de /N/ e /S/ em coda, têm representação ortográfica simples¹². Para ambas as estruturas silábicas analisadas, foram estabelecidos 4 tipos de erro de base fonológica que os dados poderiam apresentar, com base naquilo que estudos anteriores apontam sobre a aquisição das sílabas complexas na escrita¹³ e nos principais tipos de fenômenos verificáveis na fonologia infantil quando se trata da aquisição de sílabas complexas (ver seção seguinte): omissão, metátese, epêntese e substituição.

Procedimentos metodológicos

Neste texto, são apresentados e discutidos dados do estudo de Pachalski (2020), dissertação que teve como objetivo descrever e analisar a grafia de *onset* e de rima ramificados em textos de crianças brasileiras em fase de aquisição da escrita (1º a 5º ano do Ensino Fundamental). As grafias foram extraídas de 280 textos espontâneos¹⁴ componentes do sétimo estrato do Banco de Textos de Aquisição da Linguagem Escrita (BATALE)¹⁵ e produzidos por crianças de 1º a 5º anos do Ensino Fundamental de uma escola pública municipal da cidade de Pelotas/RS/Brasil.

Os textos do BATALE são, em sua maioria, coletados por integrantes do GEALE em turmas dos anos iniciais de escolas parceiras, seguindo uma metodologia padrão, constituída de três etapas principais: *motivação*, uma atividade coletiva de leitura de imagens e de

¹² Para uma discussão sobre erros de natureza ortográfica e fonológica na rima ramificada, ver Pachalski e Miranda (2021).

¹³ Cf. Pachalski e Miranda (2017, 2018), Miranda (2009a, 2009b, 2012, 2019), Abaurre (1991, 2001) e Santos (2013).

¹⁴ Para uma discussão sobre o conceito de texto espontâneo e a sua pertinência para o tipo de análise proposta em estudos como este, conferir Abaurre [1988] (2011) e Nörnberg et al. (2018). Estas últimas também trazem uma descrição mais detalhada da metodologia utilizada nas coletas de texto realizadas pelo GEALE. Além disso, importa observar que, embora seja 280 o número total de textos analisados, este número não equivale ao total de sujeitos, pois a maioria dos alunos produziu ao menos dois textos cada um.

¹⁵ O BATALE foi criado em 2001, como resultado das atividades de pesquisa do Grupo de Estudos sobre Aquisição da Linguagem Escrita (GEALE), seguindo ampliado e mantido pelo grupo de pesquisa. O banco hoje dispõe de mais de 7423 textos espontâneos produzidos por crianças brasileiras, portuguesas e moçambicanas, e pode ser acessado virtualmente, mediante solicitação, através do endereço eletrônico: <https://sistemavestigios.org>.

diálogos entre pesquisador e crianças, que visa a estimular e a oferecer subsídios para a escrita do texto; *produção de texto*, a atividade individual de escrita propriamente dita, e *socialização*, uma atividade de leitura em voz alta para a turma e pesquisador das produções textuais realizadas individualmente.

Uma vez extraídas dos textos as grafias de *onset* e rima ramificados, os dados foram separados em duas grandes categorias: acertos e erros. Estes últimos foram classificados e analisados segundo a sua natureza, definida nos termos apresentados na seção 5 deste texto. Como mencionado, este artigo tem como foco erros de natureza fonológica e, por isso, são abordados apenas os dados classificados nessa categoria. A seguir, um quadro exhibe uma descrição e exemplificação dos erros considerados de natureza fonológica no contexto das sílabas complexas:

Quadro 2 – Quadro descritivo da categoria de erros *Fonológica*
(OC = *onset* complexo; RR = rima ramificada)

Categoria/natureza		<i>Fonológica</i>
Descrição		Relacionada a algum tipo de complexidade representacional fonológica ou, ainda, à fala – no caso em foco, a complexidade está relacionada à ramificação silábica
Tipos de erros	OC	Omissão, metátese, epêntese e substituição (exemplificados respectivamente nas linhas abaixo para cada um dos constituintes)
	RR	
Exemplos de erros	OC	‘binçar’ – ‘brincar’; ‘garde’ – ‘grande’; ‘espoloção’ – ‘explosão’; ‘refrequiço’ – ‘reflexo’
	RR	‘decobriram’ – ‘descobriram’; ‘enprotante’ – ‘importante’; ‘geniti’ – ‘gente’; ‘autura’ – ‘altura’

Fonte: adaptado de Miranda (2020) e de Pachalski (2020).

Por meio da classificação dos dados nas categorias mais abrangentes *acertos* e *erros* e também nas categorias que definem a *natureza* e os *tipos de erro*, foram também definidas as variáveis do estudo, sendo as que interessam neste caso específico: a) erro (considerando *onset* e rima separadamente) como variável *dependente*; e b) tipo de estrutura silábica e tipo de erro (orto)gráfico, como variáveis *independentes*.

Após essas definições, foi feito um levantamento das *taxas* de erro e de acerto, de modo que as variáveis se configuraram como numéricas. Seguiu-se, então, para as análises estatísticas descritiva e inferencial, realizadas através do *software R* (versão 4.0.2) e orientadas por profissional da área de estatística. Na tabela que segue, podem ser conferidos os testes estatísticos utilizados na análise de Pachalski (2020), referente à amostra aqui considerada:

Tabela 1 – Testes de diferença utilizados na análise dos dados

<i>Design</i>	Variável dependente	Distribuição das variáveis	Teste de diferença	Teste <i>Post Hoc</i> ¹⁶	Valor alfa
Intrassujeitos (3 ou mais momentos)	Intervalar/numérica	Anormal	<i>Friedman</i>	<i>Conover</i>	0,05

Fonte: adaptado de Pachalski (2020).

Como se observa por meio da Tabela 1, a análise exploratória dos dados revelou anormalidade na distribuição, motivo pelo qual os testes utilizados foram não-paramétricos. Na seção seguinte, os resultados obtidos com os testes são apresentados, descritos e discutidos, a fim de que se compreenda a motivação para a ocorrência dos erros (orto)gráficos em sílabas complexas, conforme defendida por Pachalski (2020).

Resultados e discussão

Antes de explorar a questão principal e mais específica deste texto, pode ser útil observar alguns dados mais gerais analisados por Pachalski (2020), pois dão ao leitor uma compreensão mais ampla das características da amostra e, assim, fornecem um enquadramento para as questões discutidas posteriormente. As Tabelas 2 e 3, adiante, mostram como estão distribuídos os dados de acordo com as categorias erro e acerto em cada um dos constituintes silábicos considerados:

Tabela 2 – Acertos e erros na grafia de *onset* complexo

	Taxas de acerto (n = 145) ¹⁷	Taxas de erro (n = 145)
Média (%)	94,60	5,40
Desvio padrão	19,72	19,72
Frequência	536/562	26/562

Fonte: adaptado de Pachalski (2020).

Tabela 3 – Acertos e erros na grafia da rima ramificada

	Taxas de acerto (n = 184)	Taxas de erro (n = 184)
Média (%)	89,11	10,90
Desvio padrão	23,80	23,80
Frequência	1428/1536	108/1536

Fonte: adaptado de Pachalski (2020).

¹⁶ Com valores de *p* corrigidos pela *correção de Bonferroni*.

¹⁷ O valor de *n* equivale ao número de sujeitos que propiciaram dados para computação. Isso significa que, no caso do *onset* complexo, 145 sujeitos produziram pelo menos uma vez essa estrutura em algum dos textos que escreveram (certa e/ou errada); no caso da rima ramificada, 184 sujeitos produziram pelo menos uma vez essa estrutura em algum dos textos que escreveram (certa e/ou errada). Assim, embora o número total de textos analisado tenha sido 280, não significa que todos eles continham as estruturas silábicas enfocadas.

As Tabelas 2 e 3 são bastante informativas acerca da amostra analisada. Primeiro, observa-se que a quantidade de alunos que grafou o *onset* complexo (n = 145) é menor que a quantidade de alunos que grafou a rima ramificada (n = 184), fato que se reflete na frequência das grafias registrada para cada estrutura: 562 para *onset* e 1536 para rima. Isso possivelmente também está ligado à frequência que tais estruturas silábicas têm no léxico do PB, uma em relação a outra: enquanto rima ramificada apresenta em torno de 22% de ocorrência no léxico, o *onset* complexo apresenta apenas 5% (VIARO; GUIMARÃES-FILHO, 2007).

As estruturas apresentam semelhanças, no entanto, em relação às diferenças entre as médias computadas para as taxas de acerto e as taxas de erro: respectivamente, o *onset* conta com médias de 94,60% e 5,40%, e a rima com médias 89,11% e 10,90%. Os testes estatísticos realizados por Pachalski (2020), para comparar as duas médias em cada constituinte, sinalizaram a significância das diferenças entre os valores, de modo que as taxas de acerto são significativamente maiores que as taxas de erro. Os valores de desvio padrão também são semelhantes ao serem comparadas as duas estruturas – não apenas isso, mas são igualmente altos, revelando a anormalidade e heterogeneidade na distribuição da amostra, isto é, a alta variabilidade entre as produções dos sujeitos. Esta, aliás, é uma característica que se estende em todos os diferentes levantamentos feitos por Pachalski (2020) e reflete um fato comum nos estudos que lidam com a linguagem em desenvolvimento, como é o caso da aquisição da escrita.

Por fim, um último aspecto ainda merece comentário: a diferença entre os constituintes silábicos em relação às médias das taxas de erro – 10,90% para rima ramificada e 5,40% para *onset* complexo. Embora pareça pequena, trata-se de uma diferença que se mostrou significativa em testes *Wilcoxon* realizados por Pachalski (2020), fato que desencadeou importantes discussões na pesquisa da autora, com destaque para o estatuto das codas mediais na fonologia infantil e na fonologia adulta.

Dentro desse enquadramento mais amplo está, então, inserida a questão específica da pesquisa de Pachalski (2020) enfatizada neste texto: existe diferença entre as médias dos diferentes tipos de erro de natureza fonológica em sílabas complexas? A hipótese da autora é a de que a omissão de consoantes, em ambos os constituintes analisados, é significativamente maior do que os demais tipos de erro de natureza fonológica estabelecidos (metátese, epêntese e substituição). Essa hipótese tem por base estudos exploratórios anteriores realizados a respeito da grafia de sílabas complexas, como os de Miranda (2009a, 2009b, 2012, 2018, 2019) e de Pachalski e Miranda (2017, 2018), que indicam maior proeminência da omissão tanto em *onset* quanto em rima ramificados.

As Tabelas 4 e 5 apresentam as médias obtidas por Pachalski (2020) para cada um dos tipos de erro de natureza fonológica, considerando cada constituinte silábico separadamente¹⁸:

¹⁸ Na análise aqui apresentada, a variável ano escolar não foi analisada devido ao caráter mais amplo do fenômeno a ser discutido no artigo. A opção foi pela amalgamação dos dados de diferentes anos escolares, uma

Tabela 4 – Médias dos erros de natureza fonológica em *onset* complexo

	Omissão (n = 8)	Metátese (n = 8)	Epêntese (n = 8)	Substituição (n = 8)
Média (%)	93,30	6,70	0,00	0,00
Desvio padrão (%)	12,84	12,84	0,00	0,00
Frequência	16/18	2/18	0/18	0/18
$\chi^2(3)$	22,00*			

* $p < 0,05$

Fonte: adaptado de Pachalski (2020).

Tabela 5 – Erros de natureza fonológica na rima ramificada

	Omissão (n = 41)	Metátese (n = 41)	Epêntese (n = 41)	Substituição (n = 41)
Média (%)	56,91	3,66	0,00	37,00
Desvio padrão (%)	46,98	17,28	0,00	46,32
Frequência	34/58	2/58	0/58	22/58
$\chi^2(3)$	43,82***			

*** $p < 0,001$

Fonte: adaptado de Pachalski (2020).

Ao comparar ambas as tabelas, em primeiro lugar, vê-se o reflexo daquilo que foi observado nas Tabelas 2 e 3: o menor número de *n* para *onset* complexo em relação à ramificada. Tem-se, então, que, dos 145 sujeitos que grafaram pelo menos um *onset* complexo nos textos, apenas 8 efetivamente produziram erros (de natureza fonológica); da mesma forma, dos 184 sujeitos que grafaram rimas ramificadas nos textos, 41 deles produziram erros envolvendo essa estrutura. Igualmente, constatam-se elevados valores de desvio padrão, especialmente na rima, refletindo a já apontada característica assimétrica da amostra em termos da distribuição dos dados.

No que diz respeito aos tipos de erro, observa-se que a omissão apresentou médias mais altas em relação às médias dos demais tipos de erro em ambas as estruturas silábicas, com 93,30% para *onset* e 56,91% para rima; a metátese exibiu uma média levemente mais alta em *onset*, com 6,70% em relação aos 3,66% da rima; casos de epêntese não foram registrados na grafia de ambos os constituintes, e a substituição foi bastante registrada na rima, com 37% de média, se comparada ao zero de registro em *onset*.

Nas figuras a seguir, são expostos excertos de erros encontrados nos textos analisados por Pachalski (2020), que exemplificam os diferentes tipos de erro em discussão:

Figura 4 – Exemplos de grafias com omissão em *onset* complexo e em rima ramificada, por alunos do 4º e do 2º ano, respectivamente: ‘binquedos’ para ‘brinquedos’ e ‘tabem’ para ‘também’

vez que, juntos, constituem uma unidade de aprendizagem do Ensino Fundamental (Anos Iniciais ou 1º Ciclo). Além disso, o tamanho exíguo da amostra estudada limitou análises estatísticas mais robustas. A comparação geral dos erros entre os diferentes anos escolares, no entanto, pode ser conferida em Pachalski (2020).



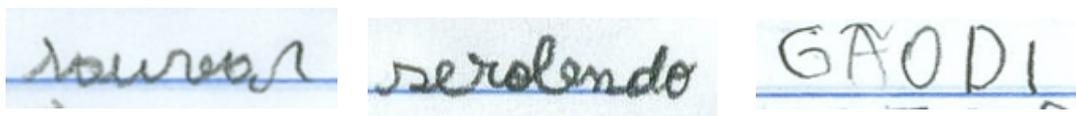
Fonte: BATALE – 7º estrato (MIRANDA, 2001) via Pachalski (2020).

Figura 5 – Exemplos de grafias com metátese em *onset* complexo e em rima ramificada, por alunos do 4º ano, respectivamente: ‘garde’ para ‘grande’ e ‘enprotante’ para ‘importante’



Fonte: BATALE – 7º estrato (MIRANDA, 2001) via Pachalski (2020).

Figura 6 – Exemplos de grafias com substituição em rima ramificada, por alunos do 5º ano (‘sauvar’ para ‘salvar’ e ‘serolendo’ para ‘se rolando’) e do 1º ano (‘gãodi’ para ‘grande’)



Fonte: BATALE – 7º estrato (MIRANDA, 2001) via Pachalski (2020).

Omissão e metátese se mostram bastante transparentes quanto àquilo que denotam seus próprios nomes, em ambas as estruturas: a omissão envolve a falta da consoante em posição silábica ramificada e a metátese envolve o deslocamento de consoante para outra posição silábica na palavra. A substituição é que apresenta características mais heterogêneas: pode ser o uso de um grafema no lugar de outro na última posição da rima, como em ‘sauvar’ para ‘salvar’, seu sentido mais comum; mas também pode ser a troca de grafemas que representam a nasalidade pós-vocálica medial por grafemas que representam um ditongo nasal, como em ‘gãodi’ para ‘grande’, ou, ainda, o uso de um grafema no lugar de outro na primeira posição da rima, como em ‘serolendo’ para ‘se rolando’. Quando aqui se refere à substituição, portanto, está se referindo a uma alteração que pode ocorrer na rima como um todo¹⁹.

Quanto aos resultados da estatística inferencial, os testes *Friedman* revelaram haver diferenças significativas entre as médias das taxas de ocorrência de cada tipo de erro nas duas estruturas silábicas. Em *onset* complexo, os valores obtidos foram $\chi^2(3) = 22,00$ e $p = 0,043$, e em rima ramificada $\chi^2(3) = 43,82$ e $p = 0,000$, permitindo, assim, reter as respectivas hipóteses alternativas. Os testes *post hoc* *Conover* ($p < 0,0083$) permitiram verificar entre quais médias, especificamente, essas diferenças se manifestam: em *onset*, a média de omissão se mostrou significativamente maior que a de metátese, epêntese e substituição, da

¹⁹ Casos de substituição em *onset* complexo, não registrados na amostra analisada, são mais simples, e envolvem o uso do grafema que representa uma líquida no lugar do grafema que representa outra líquida (<l> por <r> ou <r> por <l>). Miranda (2019) exhibe como exemplo desse tipo de erro em *onset* complexo a grafia ‘bruzéu’ para ‘blusão’.

mesma forma que em rima, com a diferença de que, nesta estrutura, a média de substituição também se mostrou significativamente maior que a de metátese e epêntese.

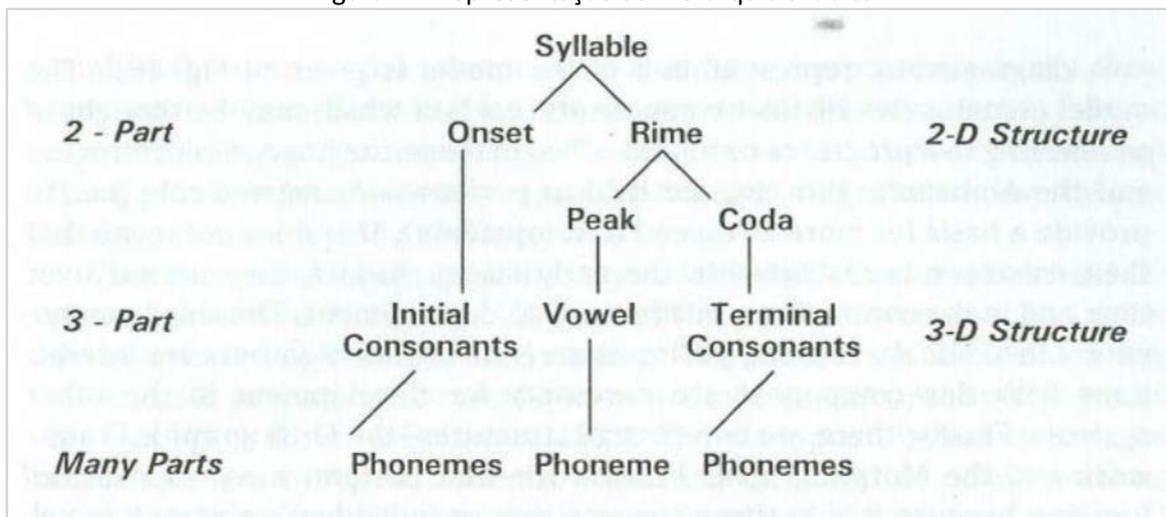
Dessa forma, Pachalski (2020) constata que, sim, existe diferença significativa entre as médias dos diferentes tipos de erro de natureza fonológica em sílabas complexas. Naturalmente, a principal questão que se coloca é como interpretar esse resultado: por que seria a omissão o tipo de erro mais frequente dentre os demais? O que isso pode indicar acerca dos conhecimentos mobilizados no processo de aquisição da escrita, particularmente em relação às estruturas silábicas complexas?

Para a autora, a omissão parece indicar que representar graficamente um fonema ligado a um constituinte silábico ramificado é um problema²⁰, e isso de forma mais prototípica do que os outros tipos de erro. Diferente dos outros fenômenos – metátese, epêntese e substituição –, na omissão, o segmento e a própria ramificação à qual ele está ligado, aparentam passar despercebidos durante o processamento da informação. Pachalski (2020) sugere, como uma das principais motivações para esse fato, haver uma *explicitação incompleta* das unidades intrassilábicas na representação do sujeito em processo de aquisição da escrita. Ou seja, perante uma parcial acessibilidade à organização prosódica a que está submetida uma sequência segmental, o segmento em posição ramificada em alguns momentos pode simplesmente ser ignorado, o que torna mais fácil e natural, em termos de esforço cognitivo, omiti-lo ao invés de produzi-lo em posição ou qualidade alterada – ações que corresponderiam aos casos de metátese, epêntese e substituição. Em outras palavras, identificar um segmento ou, uma vez identificando-o, decidir onde posicioná-lo na sequência dada, demandaria mais acesso consciente à constituição interna da sílaba do que não o produzir.

Essa hipótese interpretativa encontra amparo em bases teóricas mais diretas. A começar por Seymour (1997), autor que, ao propor um modelo de desenvolvimento ortográfico, discute o papel que teria a consciência fonológica nesse processo, e mais particularmente o acesso às unidades silábica e intrassilábicas. Para tanto, ele retoma estudos brevemente revisados na seção 4 deste texto, diante da seguinte figura:

²⁰ O termo *problema* faz remissão à sua concepção científica, isto é, de questão a ser investigada, questão essa à qual estão associadas hipóteses. Assim, as crianças têm, na grafia de fonemas em posição silábica ramificada, um problema no sentido que essa tarefa lhes impõe uma questão a ser investigada. Procurando responder a essa questão, formulam hipóteses interpretativas: provisórias, testadas por meio da escrita e reveladas através dos erros (orto)gráficos. Essa perspectiva está diretamente associada à concepção piagetiana de *erro construtivo*, paradigma basilar das pesquisas desenvolvidas no GEALE.

Figura 7 – Representação da hierarquia silábica

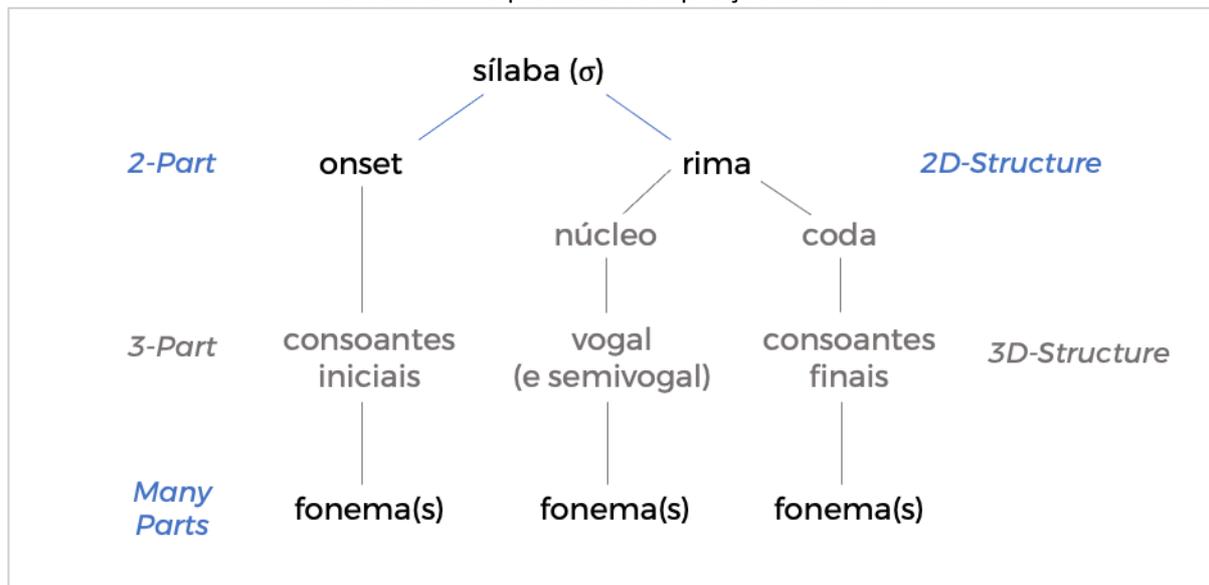


Fonte: SEYMOUR, 1997, p. 323.

Considerando o que apontam pesquisas como as de Liberman et al. (1974) e de Treiman (1992), Seymour (1997) ressalta que a capacidade de acessar conscientemente unidades fonológicas tende a se desenvolver em um processamento do tipo “de cima para baixo” (*top-down*): antes mesmo de as crianças ingressarem em classes formais de alfabetização, elas já demonstram a capacidade de segmentar a língua primeiro em unidades maiores, como a sílaba inteira, e em seguida nos constituintes imediatos *onset* e rima, os quais correspondem à camada *2-Part* ou *2D Structure* da Figura 7. Treiman (1992), porém, assinala que o acesso consciente de crianças pré-escolares para por aí, isto é, crianças não-alfabetizadas ainda não demonstram ser capazes de segmentar a língua em unidades menores que *onset e rima*, o que inclui não somente a camada fonêmica (*Many Parts*), como já é de praxe esperar, mas também a camada que corresponde às ramificações desses constituintes silábicos, representada como *3-Part* ou *3D Structure* na Figura 7. Considerando também o que outros célebres estudos da área de consciência fonológica demonstram (cf. MORAIS et al., 1979; READ et al., 1986), o acesso à camada fonêmica será possibilitado como um efeito da própria aquisição da escrita alfabética, uma vez que o sistema demanda a sensibilidade à existência de fonemas, em razão de seu princípio operacional eleger essas unidades da língua como foco da representação gráfica.

Tendo em conta esse quadro, Seymour (1997) supõe que, a partir da aquisição do princípio alfabético do sistema de escrita, a tendência de o processamento continuar a se dar no sentido *top-down* é descontinuada, pois, com a aquisição da escrita, o acesso à camada fonêmica da língua (*Many Parts*) precederia o acesso à camada que corresponde à ramificação silábica (*3-Part*). Assim, um processamento do tipo *bottom-up* (de baixo para cima) se interpõe ao *top-down*. Em outras palavras, dá-se um salto da camada *2-Part* à camada *Many Parts*, em termos de acesso consciente, deixando a camada *3-Part* suspensa. Uma representação visual desse processo é proposta adiante na Figura 8, em uma adaptação da Figura 7, na qual a explicitação incompleta da camada *3-Part* é representada com a cor cinza:

Figura 8 – Representação da *explicitação incompleta* das unidades intrassilábicas – camada *3-Part/3D Structure* – no processo de aquisição da escrita



Fonte: adaptação de Pachalski (2020).

Pachalski (2020) interpreta que esse fato implica em uma noção inicial dos fonemas como unidades isoladas do ponto de vista das crianças, isto é, unidades que não necessariamente obedecem a uma fonotática. É preciso lembrar, a esta altura, que a sílaba tem como principal papel, na gramática fonológica, organizar a distribuição dos segmentos da língua, de modo que define, para isso, restrições fonotáticas. Uma vez que não se tem acesso claro a essa informação, fornecida de forma direta pela camada *3-Part* à camada *Many Parts*, os segmentos em posição silábica ramificada tornar-se-iam, então, suscetíveis a erros na sua representação gráfica.

Portanto, a ideia da expressão *explicitação incompleta* se refere, de modo mais específico, a um acesso insuficiente à camada que corresponde às ramificações silábicas, que fica “em suspenso”, por assim dizer, em meio às outras camadas já acessadas. É importante observar aqui um aspecto fundamental desse argumento: a dúvida envolvendo a representação das sílabas complexas na escrita não diz respeito a uma *total* inacessibilidade à organização da estrutura silábica, mas sim a um acesso insuficiente, ou seja, *incompleto*. Se é incompleto, subentende-se tanto que se trata de um conhecimento ainda em construção quanto que o acesso a alguma informação existe, ainda que não por inteiro. Do contrário, seria observado um número muito maior de erros em relação ao que se registra efetivamente, como visto nas Tabelas 2 e 3, e não seria possível observar textos nos quais a mesma criança simultaneamente produz erros e acertos em sílabas complexas, como o que segue adiante na Figura 9:

Figura 9 – Texto produzido por aluno do 2º do Ensino Fundamental²¹

Taba
Um animal muito forte
muito rápido come fruta
e salva animais e vive
na floresta e cava e inteligente.
porque ele salva a mimais
e vai ajuda floresta e vai
salva os ovos de passaro
nunca vai para de ajuda
todos.

Fonte: BATALE – 7º estrato (MIRANDA, 2001) via Pachalski (2020).

No texto da Figura 9, observa-se de um lado a coexistência de erros e acertos em sílabas complexas, pois palavras como ‘salva’ e ‘inteligente’ são grafadas conforme a norma, enquanto em ‘nuca vai’ para ‘nunca vai’, a nasal pós-vocálica é omitida. De outro lado, também se vê a coexistência de diferentes tipos de erros nas duas estruturas silábicas analisadas: em ‘floresta’ para ‘foresta’ e ‘frote’ para ‘forte’, por exemplo, tem-se omissão e metátese, respectivamente. Repare-se ainda que em ‘foresta’ o erro em *onset* é seguido de um acerto na coda fricativa, e em ‘frote’ a metátese na rima ramificada gera como resultado um *onset* complexo. A ideia aqui é que se fosse o caso de haver uma falta completa de acesso à informação sobre como se estruturam as sílabas internamente, os erros seriam sistemáticos e predominantes em relação aos acertos, e, conjuntamente, a omissão seria a única possibilidade tipológica de erro, uma vez que a metátese demandaria ao menos um nível inicial de acesso consciente às camadas intrassilábicas, pois está ligada a uma dúvida sobre onde posicionar, na estrutura da sílaba, um segmento que já foi identificado.²²

²¹ Sugestão de leitura: “Taba [título] Um animal muito forte, muito rápido, come fruta. Ele salva animais e vive na floresta e cava e [é] inteligente. Porque ele salva animais e vai ajudar [a] floresta e vai salvar os ovos de pássaro. Nunca vai parar de ajudar todos”.

²² Esta é uma importante diferença existente entre a aquisição da escrita e a aquisição fonológica. Enquanto a estrutura não está disponível na fonologia, os “erros” ou a não-produção da estrutura são sistemáticos, e produções conforme o alvo são raras e esporádicas. Na aquisição da escrita, porém, o conhecimento fonológico já construído é “somente” retomado e analisado em suas partes componentes, fazendo com que, nesta etapa, acertos sejam mais sistemáticos e erros mais esporádicos.

Do ponto de vista da sustentação teórica da ideia de explicitação incompleta, além do que pontua Seymour (1997) de modo mais específico, é fundamental recobrar o Modelo de Redescritção Representacional (MRR) de Karmiloff-Smith (1994), apresentado em linhas gerais na seção 4 deste texto. O MRR ampara a ideia de que podem ocorrer explicitações incompletas ao propor que o desenvolvimento (meta)cognitivo acontece por meio de um processo cíclico que transforma progressivamente representações implícitas na mente em representações explícitas para a mente. O acesso consciente, assim, não ocorre de forma automática e em duas etapas apenas – do inconsciente ao consciente – mas de forma gradativa, passando por diferentes níveis. Além disso, o modelo presume que o conhecimento de um domínio cognitivo, como o linguístico, pode estar em um diferente nível de explicitação em relação a outro tipo de conhecimento que faz parte do mesmo domínio.

Primeiro, então, há de se ter em conta que o conhecimento fonológico, principal insumo das crianças para o desenvolvimento da escrita alfabética, estaria inicial e predominantemente em um nível I, conforme a terminologia da autora, isto é, codificado em um formato que inviabiliza a sua manipulação, a sua análise e a interface entre domínios cognitivos. Em outras palavras, permitindo às crianças um bem-sucedido uso da linguagem atualizada em substância fônica, mas impossibilitando um exercício metalinguístico, por exemplo.

Como demanda para a aquisição da escrita alfabética, esse conhecimento precisa ser acessado ou explicitado, o que ocorreria de forma progressiva, de acordo com o que propõe o MRR sobre o desenvolvimento metacognitivo. Assim, uma estrutura que carece de explicitação, ou seja, uma estrutura cuja explicitação é incompleta, é uma estrutura que estaria em um nível E1 em termos de abertura ao acesso consciente, o nível seguinte ao Implícito no MRR. Uma das características que revela a existência desse nível segundo Karmiloff-Smith (1994) é a ocorrência de erros de execução em uma determinada tarefa que um sujeito costumava cumprir com *maestria procedimental*, isto é, com um uso tácito bem-sucedido. Nesse sentido, poder-se-ia pensar que a coexistência de erros e de acertos em um texto produzido pela mesma criança, ou a coexistência de erros de diferentes tipologias nesse mesmo texto, estaria incluída na caracterização feita pela autora: a criança costuma produzir conforme o alvo as sílabas complexas – tanto na fala quanto na escrita, pois o conhecimento que serve como matéria-prima para a produção é o mesmo – mas, por vezes e em dado período de seu desenvolvimento, erra a sua grafia na escrita, uma vez que esta modalidade da língua exige o acesso consciente ao conhecimento fonológico, que ainda não está suficientemente explícito no caso que se está tratando.

Outra característica importante do nível E1 no MRR para os fins da presente discussão é o formato das representações mentais, que já seriam mais flexíveis em relação ao nível anterior e, por isso, *capazes de criar vínculos* com outras representações. Se a camada correspondente às ramificações silábicas não está suficientemente explicitada, como se viu, o vínculo com as demais camadas que a envolvem conseqüentemente também não está. É um momento do percurso de desenvolvimento em que determinadas estruturas não estão nem completamente abertas ao acesso consciente, tampouco inacessíveis, como no nível I. Como

conclui Pachalski (2020), a tarefa cognitiva do sujeito a partir daí seria, então, a de relacionar as representações, de criar vínculos entre elas, uma vez que já estão em um formato disponível para tanto, o que fará com que avance o processo de explicitação das estruturas silábicas complexas. Com o decorrer da aquisição da escrita e a consequente estabilização das grafias, as crianças poderiam chegar a um nível E2 de explicitação – hipótese que testes de consciência fonológica, voltados à segmentação de unidades intrassilábicas, poderiam ajudar a corroborar, pois, por meio deles, as crianças poderiam revelar mais claramente a que camada da sílaba chega a sua capacidade de segmentação. Alcançar o nível E3 de explicitação, por fim, não seria o esperado, uma vez que se trata de um conhecimento bastante abstrato para ser verbalizável, e que, muito provavelmente, apenas adultos com formação linguística teriam capacidade para fazê-lo.

Considerações finais

Este artigo procurou explorar o que o estudo de Pachalski (2020) sugere ser a principal motivação para a ocorrência de erros (orto)gráficos em sílabas complexas na escrita inicial. A hipótese defendida com sustentação teórica e empírica é a de que a maioria dos erros observáveis em sílabas complexas seria decorrente da *explicitação incompleta da camada correspondente às ramificações silábicas*, isto é, erros em sílabas complexas ocorrem não porque existem estruturas que ainda precisam ser incorporadas ao inventário da criança, uma vez que isso já ocorreu durante o período de aquisição da fala, mas porque essas mesmas estruturas precisam ser acessadas conscientemente pela criança.

O que sustenta a ideia de haver uma explicitação incompleta é que, em primeiro lugar, isso teria correspondência com a tendência natural do curso de desenvolvimento (meta)cognitivo mais abrangente conforme proposto por Karmiloff-Smith (1994), o qual prevê a existência de níveis cada vez mais sofisticados de acesso consciente à informação que o sistema cognitivo vai alcançando, passando de um nível implícito, formato com o qual todo conhecimento inicia, até diferentes níveis de explicitação dos conhecimentos, momento em que atividades de caráter metalinguístico podem passar a ser realizadas.

Depois, de forma mais específica, estaria a interposição de um processamento *bottom-up* naquilo que seguia uma tendência *top-down*, isto é, relativamente ao acesso consciente às unidades silábica e intrassilábicas. Crianças pré-escolares (não-alfabetizadas) já demonstram capacidade de segmentar a língua tanto em sílabas inteiras quanto nas unidades intrassilábicas *onset* e rima (LIBERMAN et al., 1974; TREIMAN, 1992). No entanto, crianças não-alfabetizadas ainda não seriam capazes de segmentar em unidades menores que *onset* e rima (TREIMAN, 1992), o que inclui não somente a camada fonêmica, *mas também as ramificações desses constituintes silábicos*. O acesso à camada fonêmica será possibilitado como um efeito da própria aquisição da escrita alfabética (cf. MORAIS et al., 1979; READ et al., 1986), uma vez que o sistema demanda a sensibilidade à existência de fonemas, em razão de seu princípio operacional eleger essas unidades da língua como foco da representação gráfica.

Assim, Seymour (1997) supõe que, a partir da aquisição do princípio alfabético do sistema de escrita, a tendência de o processamento continuar a se dar no sentido *top-down* é descontinuada, pois, com a descoberta do princípio alfabético, o acesso à camada fonêmica da língua precederia o acesso à camada que corresponde à ramificação silábica. É assim que um processamento do tipo *bottom-up* (de baixo para cima) se interporia ao *top-down*. Em outras palavras, considerando-se o exposto nas Figuras 7 e 8 deste texto, dar-se-ia um salto da camada silábica *2-Part* à camada *Many Parts*, em termos de acesso consciente, deixando a camada *3-Part* suspensa.

Na terminologia do MRR de Karmiloff-Smith (1994), esta camada estaria em um nível E1 de explicitação – nem inacessada, nem completamente acessada; com um formato de representações do conhecimento já mais flexíveis e capazes de criar vínculos com outras representações – exatamente aquilo que é necessário neste ponto do desenvolvimento ortográfico, particularmente no que toca às sílabas complexas, pois a camada fonêmica (*Many Parts*) precisa associar-se à camada correspondente às ramificações silábicas (*3-Part*). Diante desse quadro, a tarefa do sujeito, então, seria relacionar as representações, criar vínculos entre elas, processo que acarretará a explicitação das mesmas representações.

Do ponto de vista empírico, a predominância da omissão em relação aos demais tipos de erro de natureza fonológica, tanto em *onset* quanto em rima ramificada, fortalece a sustentação teórica ao mesmo tempo em que é por ela interpretada. Se a organização interna da sílaba não está suficientemente explícita, tampouco está a informação sobre como devem os segmentos se organizar em termos de restrições fonotáticas. Assim, torna-se mais custoso identificá-los e registrá-los, e, diante disso, a solução mais simples em termos de esforço cognitivo terminaria por ser a omissão do segmento ligado ao constituinte ramificado, cuja informação seria parcial ao sujeito, dada a explicitação incompleta.

Portanto, conforme pontua Pachalski (2020), o que se defende com o argumento exposto é que a motivação majoritária dos erros (orto)gráficos em sílabas complexas *não está ligada a uma análise fonêmica insuficiente*, como intuitivamente se pressupõe muitas vezes, mas sim a uma análise insuficiente da associação existente entre a camada fonêmica, já explicitada devido à demanda exigida pelo sistema de escrita alfabética, e a última camada da sílaba, formada pelas ramificações de *onset* e rima.

É importante salientar, por fim, que a sustentação dessa hipótese não se dá sem a pressuposição, defendida por Miranda (2017), de que o principal conhecimento a ser acessado pela criança especialmente no início do curso de aquisição da escrita alfabética é a *gramática fonológica* já construída durante a aquisição da fala, e que, nesse processo, essa gramática pode ser reestruturada – razão por que se concebe a aquisição da escrita como integrante do processo mais amplo de aquisição da linguagem e por que, conseqüentemente, os estudos de aquisição da escrita podem lançar luz sobre a configuração da gramática infantil.

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].

- ABAURRE, M. B. M. A relevância dos critérios prosódicos e semânticos na elaboração de hipóteses sobre segmentação na escrita inicial. *Boletim da ABRALIN*, n. 11, p. 203-217, 1991b.
- ABAURRE, M. B. M. Dados da escrita inicial: indícios da construção da hierarquia de constituintes silábicos? In: MATZENAUER, C. L. B. (org.). *Aquisição de língua materna e de língua estrangeira: aspectos fonético-fonológicos*. Pelotas: EDUCAT, 2001. p. 63-86.
- 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.
- BISOL, L. O ditongo na perspectiva da fonologia atual. *D.E.L.T.A.*, São Paulo, v. 5, n. 2, p. 185-168, 1989.
- BONILHA, G. F. G. Aquisição fonológica do português brasileiro: uma abordagem conexionista da teoria da otimidade. 2005. 388 f. Tese (Doutorado em Letras) – Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2005.
- CÂMARA JR., J. M. *Estrutura da língua portuguesa*. 23. ed. Petrópolis: Vozes, [1970] 1995.
- CHOMSKY, N. *Language and problems of knowledge: the Managua Lectures*. Cambridge: MIT Press, 1988.
- COSTA, J.; FREITAS, M. J. Sobre a representação das vogais nasais. In: MATZENAUER, C. L. (org.). *Aquisição de língua materna e de língua estrangeira: aspectos fonético-fonológicos*. 1ª ed. Pelotas: ALAB/EDUCAT, 2001. p. 87-109.
- FRITH, U. Literally changing the brain. *Brain* 121, v. 6, n. 6, p. 1011-1012, 1998. <https://doi.org/10.1093/brain/121.6.1011>
- GELB, I. J. *A study of writing*. Chicago: University of Chicago Press, 1952.
- GOMBERT, J. E. *Metalinguistic development*. Chicago: University of Chicago Press, 1992.
- 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.
- KARMILOFF-SMITH, A. *Más allá de la modularidad: la ciencia cognitiva desde la perspectiva del desarrollo*. Madrid: Alianza Editorial, S.A., 1994.
- 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.; BONILHA, G. F. G. A aquisição de [kw] e [gw] sob o enfoque de restrições. VI ENAL, 2003, Porto Alegre, *Anais* [...]. Porto Alegre: PUCRS, 2003.
- LAMPRECHT, R. R. et al. (orgs.). *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.
- LEE, S. Sílabas no português brasileiro na visão da Teoria da Optimalidade. In: II Congresso Internacional da Abralín, 1999, Florianópolis, *Anais*... Florianópolis: UFSC, 1999.
- LEMLE, M. *Guia teórico do alfabetizador*. São Paulo: Ática, 1987.

LIBERMAN, I. Y. *et al.* Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, v. 18, n. 2, p. 201-212, 1974. [https://doi.org/10.1016/0022-0965\(74\)90101-5](https://doi.org/10.1016/0022-0965(74)90101-5)

LORANDI, A.; KARMILOFF-SMITH. From sensitivity to awareness: morphological knowledge and the Representational Redescription model. *Letras de Hoje*, v. 47, n. 1, p. 6-16, 2012.

LYONS, J. *Introduction to Theoretical Linguistics*. London: Cambridge University Press, 1968. <https://doi.org/10.1017/CBO9781139165570>

MACKEN, M. A. Phonological acquisition. In: GOLDSMITH, J. A. (org.). *The Handbook of Phonological Theory*. Oxford: Blackwell Publishing, 1996. p. 1-17.

MACKEN, M. A. Where is phonology? In: FERGUSON, C; MENN, L; STOEL-GAMMON, C. (orgs.). *Phonological development: models, research, implications*. The Hague: Holland Academic Graphies, 1992. p. 249-269.

MARTINET, A. *Elementos de linguística geral*. 2. ed. São Paulo: Editora Livraria Sá da Costa, 1970 [1960].

MATEUS, M. H.; D'ANDRADE, E. *The phonology of portuguese*. Oxford: Oxford University Press, 2000.

MATZENAUER, C. L. B.; MIRANDA, A. R. M. A construção do conhecimento fonológico na aquisição da linguagem. *Rev. Est. Ling.*, v. 20, n. 2, p. 91-124, 2012. <https://doi.org/10.17851/2237-2083.20.2.91-124>

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

MIRANDA, A. R. M. A grafia de estruturas silábicas complexas na escrita de crianças das séries iniciais. In: PINHO, S. Z. (org.). *Formação de Educadores: o papel do educador e sua formação*. São Paulo: Unesp, 2009a. p. 409-426.

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. <https://doi.org/10.5007/1984-8412.2019v16n2p3825>

MIRANDA, A. R. M. Aquisição da escrita: as pesquisas do GEALE. In: MIRANDA, A. R. M.; CUNHA, A. P. N.; DONICHT, G. (orgs.). *Estudos sobre Aquisição da Linguagem Escrita*. Pelotas: Editora UFPel, 2017. p. 15-50.

MIRANDA, A. R. M. *BATALE: Banco de Textos de Aquisição da Linguagem Escrita*. Pelotas: Faculdade de Educação, Universidade Federal de Pelotas, 2001.

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, 2009b. p. 111-130.

MIRANDA, A. R. M. Reflexões sobre a fonologia e a aquisição da linguagem oral e escrita. *Veredas online*, v. 16, p. 118-135, 2012.

MIRANDA, A. R. M. Um estudo sobre a natureza dos erros (orto)gráficos produzidos por crianças dos anos iniciais. *Educ. rev. [online]*, vol. 36, 2020. <https://doi.org/10.1590/0102-4698221615>

PACHALSKI, L. A grafia de sílabas complexas na aquisição da escrita: relações entre fonologia e ortografia. 2020. 197 f. Dissertação (Mestrado em Letras) – Programa de Pós-Graduação em Letras, Centro de Letras e Comunicação, Universidade Federal de Pelotas, Pelotas, 2020.

PACHALSKI, L.; MIRANDA, A. R. M. A grafia da rima ramificada em textos dos anos iniciais do ensino fundamental: conhecimento fonológico e ortográfico. In: CAVALCANTE, M. C. B.; BARROS, I. B. R. *Linguagem: aquisição da fala e da escrita*, 1. ed. vol. 3. Pontes Editores, 2021. p. 138-161.

PACHALSKI, L.; MIRANDA, A. R. M. A grafia de sílabas complexas na escrita inicial: um estudo comparativo entre dados de escrita espontânea e controlada. In: XX Encontro de Pós-Graduação, 2018a, Pelotas, *Anais eletrônicos...* Pelotas, UFPel: 2018a. Disponível em: <http://cti.ufpel.edu.br/siepe/arquivos/2018/LA_02952.pdf>.

PACHALSKI, L.; MIRANDA, A. R. M. Da grafia do onset complexo à grafia da metátese na aquisição da escrita: um trajeto investigativo na iniciação científica. In: XXVI Congresso de Iniciação Científica UFPel, 2017, Pelotas. *Anais eletrônicos...* Pelotas: UFPel, 2017b. Disponível em: <http://cti.ufpel.edu.br/siepe/arquivos/2017/LA_03800.pdf>.

READ, C. *et al.* The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition*, v. 24, p. 31-44, 1986. [https://doi.org/10.1016/0010-0277\(86\)90003-X](https://doi.org/10.1016/0010-0277(86)90003-X)

SANTOS, R. N. Aquisição de grupos consonânticos e seu impacto nos desempenhos escritos no 1º Ciclo do Ensino Básico. 2013. 154 f. Dissertação (Mestrado em Letras) – Faculdade de Letras, Universidade de Lisboa, 2013.

SAUSSURE, F. *Curso de linguística geral*. In: BALLY, C.; SECHEHAYE, A. (orgs.). 28ª ed. São Paulo: Cultrix, 2012 [1916].

SELKIRK, E. The syllable. In: HULST, H.; SMITH, N. (orgs.). *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. (orgs.). *Learning to spell*. New Jersey: Lawrence Erlbaum Associates, 1997. p. 319-337.

SEYMOUR, P. H. K.; ARO, M.; ERSKINE, J. M. Foundation literacy acquisition in european orthographies. *British Journal of Psychology*, v. 94, p. 143-174, 2003. <https://doi.org/10.1348/000712603321661859>

SOARES, M. *Alfabetização: a questão dos métodos*. São Paulo: Contexto, 2016.

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. <https://doi.org/10.4324/9781351236904-4>

VELOSO, J. Língua na escrita e a escrita da língua: algumas considerações gerais sobre transparência e opacidade fonémicas na escrita do português e outras questões. *INVEP*, Lisboa, v. VI, n. 1, p. 49-69, 2005.

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.

Recebido em: 15/06/2021.

Aceito em: 30/09/2021.