

UNIVERSIDADE FEDERAL DE PELOTAS
Centro de Letras e Comunicação
Programa de Pós-Graduação em Letras



Dissertação

A grafia de sílabas complexas na aquisição da escrita:
relações entre fonologia e ortografia

Lissa Pachalski

Pelotas, 2020

Lissa Pachalski

A grafia de sílabas complexas na aquisição da escrita:
relações entre fonologia e ortografia

Dissertação apresentada ao Programa de Pós-Graduação em Letras da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Mestre em Letras.

Orientadora: Ana Ruth Moresco Miranda

Pelotas, 2020

Universidade Federal de Pelotas / Sistema de Bibliotecas
Catalogação na Publicação

P116g Pachalski, Lissa

A grafia de sílabas complexas na aquisição da escrita :
relações entre fonologia e ortografia / Lissa Pachalski ; Ana
Ruth Moresco Miranda, orientadora. — Pelotas, 2020.

197 f.

Dissertação (Mestrado) — Programa de Pós-Graduação
em Letras, Centro de Letras e Comunicação, Universidade
Federal de Pelotas, 2020.

1. Sílabas complexas. 2. Aquisição da escrita. 3.
Fonologia e ortografia. 4. Erro (orto)gráfico. I. Miranda, Ana
Ruth Moresco, orient. II. Título.

CDD : 410

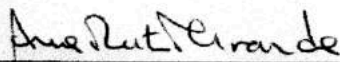
Lissa Pachalski

**A GRAFIA DE SÍLABAS COMPLEXAS NA AQUISIÇÃO DA ESCRITA:
RELAÇÕES ENTRE FONOLOGIA E ORTOGRAFIA**

Dissertação aprovada, como requisito parcial, para obtenção do grau de Mestre em Letras, Área de concentração Linguagem, Texto e Imagem, do Programa de Pós-Graduação em Letras da Universidade Federal de Pelotas.

Pelotas, 31 de julho de 2020

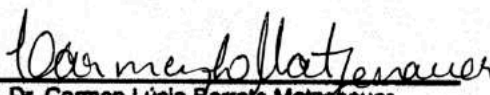
Banca examinadora:



Prof. Dr. Ana Ruth Moresco Miranda
Orientadora/Presidente da Banca
Universidade Federal de Pelotas



Prof. Dr. João Manuel Pires da Silva e Almeida Veloso
Membro da Banca
Universidade do Porto – Portugal



Prof. Dr. Carmen Lúcia Barreto Matzenauer
Membro da Banca
Universidade Federal de Pelotas



Prof. Dr. Giovana Ferreira Gonçalves
Membro da Banca
Universidade Federal de Pelotas

Agradecimentos

À Ana Ruth, minha mestra, orientadora, amiga e inspiração acadêmica, responsável direta por me apresentar à Linguística e por estabelecer bases sólidas e confiáveis para a minha formação na área. Obrigada pela confiança que demonstras ter em mim, por esses anos todos de trabalho intenso e produtivo no GEALE, pela tua generosidade e pelo teu desapego a vaidades acadêmicas, o que me permitiu ir muito além do previsto para um “mero aluno”;

às professoras presentes nas bancas de qualificação e de defesa, Carmen e Giovana, pelas contribuições precisas que deram para que meu trabalho tivesse mais qualidade. Para mim, é uma honra tê-las como avaliadoras da minha pesquisa. As admiro muito pelas suas trajetórias profissionais e pelo rigor científico que demonstram ter em cada uma de suas produções acadêmicas;

à professora Dina, presente na banca de qualificação, pela leitura cuidadosa e gentil do projeto de dissertação, que me rendeu ajustes importantes no trabalho. Obrigada também pelo empenho em tentar participar da banca de defesa, mesmo em meio a um difícil quadro de saúde;

ao professor João Veloso, pela prontidão em aceitar compor a banca de defesa da dissertação, trazendo a esse momento de avaliação ainda mais rigor e distinção. Também agradeço pela leitura generosa, interessada e entusiástica que fizeste do trabalho, proporcionando-me um momento singular e honrado de diálogo acadêmico;

à Marta, minha “coorientadora”. Obrigada pela confiança, pelas oportunidades de trabalho que tu me deste, pela parceria produtiva tanto no curso de Pedagogia quanto no GEALE, e por dividir desinteressadamente o vigor, a intensidade, o conhecimento e a perspicácia que tens, os quais potencializaram minhas ideias e ações no âmbito do trabalho acadêmico;

à CAPES, pela bolsa concedida durante o período de estudos no mestrado, indispensável para que eu pudesse dedicar-me exclusivamente à pesquisa e,

assim, conseguir entregar um produto acadêmico de maior qualidade; da mesma forma, agradeço à agência pelas bolsas concedidas durante o período da iniciação científica, fundamentais para que eu pudesse desenvolver as bases do que veio a se constituir esta dissertação, motivo pelo qual agradeço igualmente ao CNPq, que me concedeu financiamento nessa etapa de formação também.

ao GEALE, grupo de pesquisa no qual venho construindo minha identidade profissional e onde encontrei parcerias de trabalho, oportunidade de crescimento e espaço para a produção de conhecimento séria, paciente e honesta;

ao Nilton, pela generosa assistência prestada com a análise estatística da dissertação;

aos bolsistas de iniciação científica do GEALE, Diandra, Lorenzo e Nathalia, que me prestaram ajuda valiosa no processo de análise dos dados da dissertação, tornando-o mais rápido e eficiente;

aos colegas do Programa de Pós-Graduação em Letras da UFPel, pelas parcerias de pesquisa, de trabalhos, de conversas, de cafés, de *insights*, de organização de eventos, que tornaram minha passagem pelo PPGL mais intensa, produtiva e, ao mesmo tempo, leve;

à minha mãe, por ser minha maior incentivadora desde sempre e por ter me dado as melhores condições para poder fazer do estudo e da pesquisa o meu trabalho. E também por, “sem querer”, tornar o universo das Letras e da Linguagem tão *familiar* para mim, que, depois de algumas voltas, foi nele que eu acabei me instalando;

ao Tinho, que acredita no que eu faço e que, sempre com paciência e compreensão, me dá a segurança necessária para fazer as escolhas que julgo corretas na minha trajetória profissional;

à Luiza, à Jaque e à Bel, minhas colegas de profissão e grandes amigas, que estiveram presentes de forma fundamental na gênese do que veio a se tornar esta dissertação – e ainda depois continuaram sempre presentes, me apoiando e incentivando a seguir em frente (ou, quando necessário, a parar também);

por fim, ao único Deus, três em um, quem me possibilita a dádiva de experimentar tudo aquilo que acima citei; criador do mistério da Palavra, que move a minha curiosidade e a minha vontade; com quem está as chaves da sabedoria e do conhecimento; por meio de quem posso me mover, viver e existir; razão da minha dedicação; quem dá sentido e esperança, em primeira e última instância, a tudo que sou e faço, em cada detalhe da minha vida, incluindo, obviamente, esta dissertação, que me leva a contemplar: “então foi assim que Tu fizeste...”.

The scientist must not dally in realizing practical aims. He no doubt will obtain them, but must obtain them in addition. He never must forget that the special object he is studying is only a part of this big whole, which must be the sole motive of his activity. [...] He does not study nature because it is useful; he studies it because he delights in it, and he delights in it because it is beautiful. [...] Of course I do not speak here of that beauty that strikes the senses, the beauty of quality and appearances; [...] I mean that deeper beauty coming from the harmonious order of the parts, and that a pure intelligence can grasp.
(Henri Poincaré)

*“Grandes são as obras do SENHOR;
dignas de estudo para quem as ama”
(Salmos 111:2 – Bíblia Sagrada)*

Resumo

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

Esta pesquisa buscou descrever e analisar a grafia de *onset* e de rima ramificados em textos de crianças em fase de aquisição da escrita, estabelecendo relações com a aquisição fonológica dessas estruturas e com a teoria fonológica, a fim de compreender a motivação dos erros (orto)gráficos que as crianças produzem nas estruturas silábicas complexas. Foram analisados, quanti e qualitativamente, dados extraídos de 280 textos espontâneos produzidos por crianças estudantes do 1º Ciclo do Ensino Fundamental (1º a 5º anos) de uma escola pública da rede municipal da cidade de Pelotas/RS. Os textos foram coletados em 2014 e em 2015, e integram o Banco de Textos de Aquisição da Linguagem Escrita (BATALE). No que se refere às estruturas silábicas estudadas, foram considerados: em *onset* complexo, os grupos consonantais formados por lateral e rótica, nos moldes CCV e CCVC; em rima ramificada, os segmentos lateral, nasal, fricativo e rótico em posição medial na palavra, nos moldes VC, CVC e CCVC. O efeito das variáveis independentes (i) *tipo de estrutura silábica*, (ii) *tipo de segmento*, (iii) *tipo de erro (orto)gráfico* e (iv) *ano escolar* foi testado sobre a frequência de ocorrência dos erros (orto)gráficos encontrados nos textos. Em relação a (i), verificou-se que os erros (orto)gráficos são significativamente mais frequentes em rima ramificada do que em *onset* complexo; em relação a (ii), verificou-se que os erros (orto)gráficos são influenciados pelo tipo de segmento presente na rima ramificada, sendo que a grafia das consoantes lateral e nasal concentra significativamente mais erros do que a grafia das consoantes fricativa e rótica; em relação a (iii), verificou-se, para a rima ramificada, que no primeiro período do 1º Ciclo do EF (1º a 3º anos), os erros de natureza fonológica são significativamente mais frequentes do que os erros de natureza ortográfica; já no segundo período do 1º Ciclo do EF (4º e 5º anos), os erros de natureza ortográfica são significativamente mais frequentes que os erros de natureza fonológica. Também aferiu-se que a omissão é o tipo de erro de natureza fonológica mais frequente em ambas as estruturas silábicas analisadas, sendo que, na rima ramificada, a substituição também tem importância significativa, embora secundária em relação à omissão; em relação a (iv), por fim, verificou-se que a frequência de erros (orto)gráficos em *onset* complexo apresenta diferenças significativas entre o 1º ano e o 3º ano do EF. A partir de tais resultados, concluiu-se que os erros ortográficos em sílabas complexas produzidos por crianças em fase de aquisição da escrita podem apresentar três tipos de motivação não mutuamente excludentes: (i) a explicitação incompleta *da relação* entre os constituintes silábicos *onset* complexo e rima ramificada e a camada fonêmica da língua; (ii) nos casos da grafia de nasal e de lateral em rima ramificada apenas, o conflito entre a fonologia infantil e a fonologia do sistema ortográfico, sendo esta última equivalente à fonologia de adultos alfabetizados, as quais preveem estatutos distintos para as estruturas com nasal e lateral; (iii) nos casos da grafia de nasal e de fricativa em rima ramificada apenas, os conhecimentos relacionados a regras ortográficas contextuais e arbitrárias.

Palavras-chave: Sílabas complexas. Aquisição da escrita. Fonologia e ortografia. Erro (orto)gráfico.

Abstract

PACHALSKI, Lissa. **The spelling of complex syllables on writing acquisition: relations between phonology and orthography.** 2020. 197s. Dissertation (Masters in Linguistics) – Programa de Pós-Graduação em Letras, Centro de Letras e Comunicação, Universidade Federal de Pelotas, Pelotas, 2020.

This research aimed to describe and analyze the spelling of both complex onset and rime in texts of children in writing acquisition phase, establishing relationships with the phonological acquisition of these structures and with the phonological theory, in order to understand the motivation of (ortho)graphic errors that children produce in complex syllabic structures. Data extracted from 280 spontaneous texts produced by children from the first cycle of elementary school (1st to 5th grades) of a public school in the city of Pelotas/RS/Brazil were analyzed, quantitatively and qualitatively. The texts were collected in 2014 and 2015, and are part of the Written Language Acquisition Text Database (BATALE). With regards to the syllabic structures focused on the study, the following were considered: in complex onset, consonant groups formed by lateral and rotic, in the CCV and CCVC frames; in complex rime, the lateral, nasal, fricative and rotic segments in medial position in the word, in the VC, CVC and CCVC frames. The effect of the independent variables (i) *type of syllabic structure*, (ii) *type of segment*, (iii) *type of (ortho)graphic error*, and (iv) *school grade* was tested on the frequency of occurrence of the (ortho)graphic errors found in the texts. In relation to (i), it was found that (ortho)graphic errors are significantly more frequent in complex rimes than in complex onset; in relation to (ii), it was found that the (ortho)graphic errors are influenced by the type of segment present in the complex rime, with the spelling of the lateral and nasal consonants concentrating significantly more errors than the spelling of the fricative and rotic consonants; in relation to (iii), it was found, for the complex rime, that in the 1st period of the 1st cycle of elementary school (1st to 3rd grades), errors of phonological nature are significantly more frequent than errors of orthographic nature; in the 2nd period of the 1st cycle of elementary school (4th and 5th grades), errors of orthographic nature are significantly more frequent than errors of phonological nature. It was also verified that the omission is the most frequent type of phonological error in both the syllabic structures analyzed, and, in the complex rime, the substitution is also significant, although secondary in relation to the omission; in relation to (iv), finally, it was found that the frequency of (ortho)graphic errors in complex onset presents significant differences between the 1st grade and the 3rd grade of elementary school. From these results, it is possible to conclude that the spelling errors in complex syllables produced by children in the writing acquisition phase can present three types of motivation that are not mutually exclusive: (i) the incomplete explicitation of *the relationship* between the syllabic constituents complex onset and rime and the phonemic layer of the language; (ii) in the cases of nasal and lateral spelling in complex rhyme only, the conflict between children's phonology and the orthographic system's phonology, the latter being equivalent to the phonology of literate adults, which provide different statutes for structures with nasal and lateral; (iii) in the cases of nasal and fricative spelling in complex rime only, the knowledge related to contextual and arbitrary orthographic rules.

Keywords: Complex syllables. Writing acquisition. Phonology and orthography. (Ortho)graphic error.

Lista de abreviaturas e siglas

AQUIFONO	Banco de Dados de Aquisição Fonológica
ASL	Língua Americana de Sinais
BATALE	Banco de Textos de Aquisição da Linguagem Escrita
BiPhon	Modelo Bidirecional de Fonologia e Fonética
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CEAAL	Centro de Estudos sobre Aquisição e Aprendizagem da Linguagem
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
EF	Ensino Fundamental
GEALE	Grupo de Estudos sobre Aquisição da Linguagem Escrita
GPEL	Grupo de Pesquisa Estudos sobre a Linguagem
INIFONO	Banco de Dados de Aquisição Fonológica Inicial
LDB	Lei de Diretrizes e Bases da Educação
MRR	Modelo de Redescrição Representacional
PB	Português Brasileiro
PE	Português Europeu
PNE	Plano Nacional de Educação
PPGL	Programa de Pós-Graduação em Letras
PUCRS	Pontifícia Universidade Católica do Rio Grande do Sul
UP	Universidade do Porto
UFPeI	Universidade Federal de Pelotas

Lista de figuras

Figura 1	Modelo Biphon	36
Figura 2	Diagrama conceitual: o domínio teórico das pesquisas do GEALE	48
Figura 3	Quadro das relações entre fonemas consonantais e grafemas que caracterizam o sistema ortográfico do Português Brasileiro	54
Figura 4	Organograma das categorias de erros ortográficos do GEALE	56
Figura 5	Fluxograma do modelo de desenvolvimento ortográfico de Seymour (1997)	58
Figura 6	Representação da hierarquia silábica com base em Seymour (1997)	60
Figura 7	Representação da sílaba /mar/ conforme Kahn (1976)	62
Figura 8	Representação silábica da palavra /kaRo/ conforme Clements e Keyser (1983)	63
Figura 9	Representação da estrutura hierárquica interna da sílaba /tras/, com base em Selkirk (1982)	63
Figura 10	Escala de Soância conforme Clements e Hume (1995)	64
Figura 11	Escala de Soância conforme Selkirk (1984)	64
Figura 12	Organograma dos moldes silábicos licenciados para o Português Brasileiro	65
Figura 13	Quadro de segmentos licenciados para o Português conforme o constituinte silábico	67
Figura 14	Quadro de frequência de estruturas silábicas do Português Brasileiro de acordo com a tonicidade	68
Figura 15	Quadro-síntese do percurso de aquisição fonológica das estruturas silábicas complexas no Português Brasileiro	77
Figura 16	Cronologia de aquisição dos segmentos do PB conforme posição ocupada no constituinte silábico	78
Figura 17	Quadro com a configuração do Banco de Textos de Aquisição da Linguagem Escrita (BATALE)	96
Figura 18	Imagens do livro “Animais Fantásticos”	98

Figura 19	Imagens do livro “Animais Fantásticos” em tamanho menor, manipuladas pelas crianças participantes da coleta.	99
Figura 20	Exemplo dos processos de tratamento de textos do BATALE: digitalização, codificação e digitação	100
Figura 21	Quadro com as relações fonografêmicas das sílabas complexas no PB consideradas neste estudo	104
Figura 22	Quadro das categorias de erros (orto)gráficos utilizadas neste estudo	105
Figura 23	Quadro descritivo da categoria de erros residuais <i>outros</i>	106
Figura 24	Captura de tela com a imagem de parte de uma das planilhas de <i>Excel</i> utilizadas no processo de contagem, tabulação e organização dos dados da pesquisa (variáveis relativas a acerto e erro)	107
Figura 25	Captura de tela com a imagem de parte de uma das planilhas de <i>Excel</i> utilizadas no processo de contagem, tabulação e organização dos dados da pesquisa (variáveis relativas a tipo de erro)	107
Figura 26	Captura de tela com a imagem de parte de uma das planilhas de SPSS (aba <i>data view</i>) utilizadas no processo de levantamento estatístico descritivo e inferencial (variáveis relativas a acerto e erro)	108
Figura 27	Captura de tela com a imagem de parte de uma das planilhas de SPSS (aba <i>variable view</i>) utilizadas no processo de levantamento estatístico descritivo e inferencial (variáveis relativas a acerto e erro)	109
Figura 28	Gráfico linear das médias de frequência de erros (orto)gráficos em <i>onset</i> complexo conforme cada ano escolar	132
Figura 29	Gráfico linear das médias de frequência de erros (orto)gráficos em rima ramificada conforme cada ano escolar	135
Figura 30	Exemplos de grafias com erros de natureza ortográfica em rima ramificada, feitas por alunos do 4º e do 5º ano: ‘emgraçado’ para ‘engraçado’, ‘tanben’ para ‘também’ e ‘estinção’ para ‘extinção’	137

Figura 31	Exemplos de grafias com omissão de C2 em <i>onset</i> complexo, por alunos do 2º e do 4º ano, respectivamente: ‘floresta’ para ‘floresta’ e ‘binquedos’ para ‘brinquedos’	143
Figura 32	Exemplo de grafias com metátese em <i>onset</i> complexo, por alunos do 2º e do 4º ano, respectivamente: ‘furta’ para ‘fruta’ e ‘garde’ para ‘grande’	143
Figura 33	Exemplos de grafias com omissão em rima ramificada, por alunos do 2º e do 4º ano, respectivamente: ‘tabem’ para ‘também’ e ‘imãsinha’ para ‘irmãzinha’	145
Figura 34	Exemplos de grafias com substituição em rima ramificada, por alunos do 5º, do 1º ano e do 2º ano, respectivamente: ‘autura’ para ‘altura’, ‘imãsinha’ para ‘irmãzinha’ e ‘bastente’ para ‘bastante’	145
Figura 35	Exemplo de grafia com metátese em rima ramificada, por alunos do 2º e do 4º ano, respectivamente: ‘frote’ para ‘forte’ e ‘enprotante’ para ‘importante’	146
Figura 36	Exemplo de grafia com hipersegmentação em contexto de nasalidade medial, por alunos do 2º e do 4º ano, respectivamente: ‘gra de’ para ‘grande’ e ‘en graçado’ para ‘engraçado’	150
Figura 37	Exemplo de textos produzidos por alunos do 3º ano e do 2º ano do Ensino Fundamental, com erros e acertos em sílabas complexas	154
Figura 38	Representação da hierarquia silábica adaptada à ideia de <i>explicitação incompleta</i>	157
Figura 39	Formalização do conflito entre a fonologia infantil e a fonologia do sistema ortográfico (correspondente à fonologia adulta alfabetizada), que resulta no processo de mudança representacional da nasalidade medial (exemplo com as sílabas /grã/ vs. /graN/ da palavra ‘grande’)	161
Figura 40	Formalização do conflito entre a fonologia infantil e a fonologia do sistema ortográfico (correspondente à fonologia adulta alfabetizada), que resulta no processo de mudança	162

representacional do glide, derivado de vogal alta, em núcleo ramificado (exemplo com as sílabas /pau/ vs. /pal/ da palavra 'palco')

Figura 41 Síntese visual do argumento da pesquisa

164

Lista de tabelas

Tabela 1	Caracterização geral do Estrato 7	101
Tabela 2	Testes de diferença utilizados na análise dos dados	109
Tabela 3	Testes de correlação utilizados na análise dos dados	110
Tabela 4	Principais questões e hipóteses de pesquisa	116
Tabela 5	Acertos e erros na grafia de <i>onset</i> complexo	121
Tabela 6	Acertos e erros na grafia da rima ramificada	119
Tabela 7	Erros (orto)gráficos por tipo de estrutura silábica	121
Tabela 8	Erros (orto)gráficos por tipo de segmento (rima ramificada)	123
Tabela 9	Significância dos testes <i>Conover</i> para verificação das diferenças entre as médias de frequência de erros envolvendo cada segmento em rima ramificada	123
Tabela 10	Erros (orto)gráficos na rima ramificada considerando os diferentes estatutos de coda	125
Tabela 11	Erros (orto)gráficos por tipo de estrutura silábica (somente codas fricativa e rótica)	126
Tabela 12	Erros (orto)gráficos por tipo de segmento (<i>onset</i> complexo)	128
Tabela 13	Erros (orto)gráficos na grafia de <i>onset</i> complexo conforme o ano escolar	130
Tabela 14	Significância dos testes <i>Mann-Whitney</i> para verificação das diferenças entre as médias de frequência de erros em onset complexo em cada ano escolar	131
Tabela 15	Erros (orto)gráficos na grafia da rima ramificada conforme o ano escolar	134
Tabela 16	As diferentes naturezas de erro em rima ramificada	139
Tabela 17	As diferentes naturezas de erro em rima ramificada no Ciclo de Alfabetização (1º a 3º anos)	140
Tabela 18	As diferentes naturezas de erro em rima ramificada nos 4º e 5º anos	140
Tabela 19	Erros de natureza fonológica em <i>onset</i> complexo	143

Tabela 20	Significância do teste <i>Conover</i> para verificação das diferenças entre as médias de frequência de cada tipo de erro de natureza fonológica em <i>onset</i> complexo	144
Tabela 21	Erros de natureza fonológica na rima ramificada	145
Tabela 22	Significância do teste <i>Conover</i> para verificação das diferenças entre as médias de frequência de cada tipo de erro de natureza fonológica na rima ramificada	146
Tabela 23	Erros de substituição em rima ramificada por tipo de coda	148
Tabela 24	Erros de omissão em rima ramificada conforme tipo de coda	159

Sumário

1 Introdução	19
2 Fundamentação teórica	25
2.1 Aquisição e processamento da linguagem	25
2.1.1 A aquisição fonológica	34
2.1.2 A aquisição da escrita	39
2.1.2.1 Aquisição da escrita e consciência fonológica	49
2.1.3 Ortografia: funcionamento e aprendizagem	53
2.2 Teoria da Sílabas	61
2.3 A aquisição de estruturas silábicas complexas	69
2.3.1 Fonologia	69
2.3.1.1 Rima ramificada	70
2.3.1.2 <i>Onset</i> complexo	74
2.3.1.3 Síntese e aspectos para discussão	76
2.3.2 Escrita	80
2.3.2.1 Rima ramificada	81
2.3.2.2 <i>Onset</i> complexo	85
2.3.2.3 Síntese e aspectos para discussão	88
3 Metodologia	94
3.1 Caracterização da pesquisa	94
3.2 Os dados, os instrumentos de coleta e os sujeitos	95
3.3 Os procedimentos de análise	102
3.4 As variáveis	110
3.4.1 Variáveis dependentes	110
3.4.2 Variáveis independentes	111
3.5 Perguntas e hipóteses de pesquisa	114
4 Descrição, análise e discussão dos resultados	118
4.1 As diferenças gerais entre acertos e erros (orto)gráficos envolvendo a grafia de sílabas complexas: questão 1	118
4.2 O papel da estrutura silábica e do tipo de segmento na ocorrência de erros (orto)gráficos em sílabas complexas: questões 2 e 3	121

4.3 O papel da escolarização na ocorrência de erros (orto)gráficos em sílabas complexas: questão 4	129
4.4 A natureza dos erros (orto)gráficos em sílabas complexas: questões 5, 6 e 7	136
4.5 Síntese dos resultados: respondendo ao problema geral da pesquisa e sistematizando as conclusões	163
5 Considerações finais	178
Referências	184

1 Introdução

Para aqueles que acompanham o processo de aquisição da escrita de uma criança, é relativamente comum, em algum momento desse percurso, deparar-se com grafias tais como ‘eseplo’ para ‘exemplo’, ‘buxa’ para ‘bruxa’ ou ‘estar’ para ‘extra’¹. De forma semelhante, também é comum observar formas faladas como [so’vetʃi] para ‘sorvete’ ou [i^m’pesta] para ‘empresta’², produzidas por crianças durante o desenvolvimento da gramática fonológica, um período que antecede a aquisição do sistema de escrita.

Esses dados, tanto os referentes à aquisição da fala quanto à aquisição da escrita, compartilham a característica de alterar a estrutura da sílaba. Uma importante diferença, no entanto, é que, durante o desenvolvimento linguístico, as crianças gradualmente *constroem* a gramática fonológica, incorporando formas cada vez mais complexas e definindo sua representação por meio de um movimento de composição e síntese (MACKEN, 1992, 1996; KIPARSKY; MENN, 1977). Na aquisição do sistema de escrita alfabética, o que ocorre é uma *retomada* desse conhecimento fonológico internalizado que, em grande medida, já está estabilizado e bastante alinhado à gramática adulta: o movimento é, pois, de análise e decomposição das unidades prosódicas e melódicas (MIRANDA, 2017). Torna-se difícil observar na fala de crianças em fase inicial de escolarização, por exemplo, formas como as apresentadas anteriormente – [so’vetʃi] ou [i^m’pesta] –, as quais, na verdade, já são consideradas, nesta etapa, indícios de atraso ou desvio no desenvolvimento fonológico (LAMPRECHT et al., 2004).

¹ Dados extraídos do Banco de Textos de Aquisição da Linguagem Escrita – BATALE (MIRANDA, 2001).

² Dados extraídos de Matzenauer e Miranda (2012, p. 112).

Assim, 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 trabalho, opera-se com a hipótese de que erros de tal natureza ocorrem não porque existem estruturas que ainda precisam ser incorporadas ao inventário da criança, pois isso já ocorreu durante o período de aquisição da fala. Antes, o que majoritariamente está na base desses erros seriam processos cognitivos de conversão de uma informação *implícita* em conhecimento *explícito*, isto é, devido à passagem por níveis de acesso à consciência necessários à descoberta, pela criança, da existência e da forma como se constituem internamente as estruturas silábicas (KARMILOFF-SMITH, 1994; SEYMOUR, 1997; MIRANDA, 2017). Na medida em que conhecimentos implícitos precisam ser explicitados, pode-se dizer, também, que há a passagem de uma atividade *epilinguística* a uma atividade *metalinguística* (GOMBERT, 1992).³

Por essa perspectiva, portanto, a criança constrói ao longo do desenvolvimento fonológico um conhecimento tácito dos diferentes componentes melódicos e prosódicos, dentre eles a sílaba, que a habilita a organizar os segmentos de sua língua em unidades que, combinadas em outras maiores, formam a gramática fonológica. Antes de iniciar o processo de alfabetização, o grau máximo de acesso consciente que as crianças podem ter em relação às suas representações fonológicas é referente às unidades intrassilábicas *onset* e rima (TREIMAN, 1992). Em outras palavras, as crianças não-alfabetizadas ainda não são capazes de segmentar a linguagem em unidades menores que *onset* e rima, o que inclui não somente a camada fonêmica, *mas também as ramificações desses constituintes silábicos*. A consciência⁴ fonêmica é desenvolvida quando as crianças iniciam o processo de aquisição da escrita

³ Uma atividade *epilinguística* é aquela que evidencia uma “sensibilidade funcional do sujeito às regras de organização e de uso da língua” (GOMBERT, 1992, p. 13), podendo ser observada desde os primeiros anos de vida. O conhecimento que possibilita esse tipo de atividade é caracterizado como tácito e implícito, pois não é acessível à consciência, tampouco é manipulável; ele apenas habilita o sujeito ao uso eficaz de um procedimento. Na palavras de Karmiloff-Smith (1994, p. 71), é “informação na mente e, ainda não, conhecimento para a mente”. Uma atividade *metalinguística*, por sua vez, é aquela na qual o sujeito evidencia algum grau de consciência e de controle deliberados sobre a própria linguagem (GOMBERT, 1992), geralmente decorrentes de demandas externas, que forcem a explicitação – a “abertura” – de conhecimentos antes implícitos na mente.

⁴ Para Karmiloff-Smith (1994), cujo modelo é revisado no segundo capítulo deste trabalho, 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.

alfabética, uma vez que a escrita alfabética demanda sensibilidade à existência de fonemas, em razão do seu princípio operacional, que estabelece relações entre essas unidades fonológicas e os grafemas (MORAIS et al., 1979; READ et al., 1986).

Se bem observado, a consciência fonológica segue uma tendência natural de desenvolver-se a partir de unidades maiores a unidades menores – processamento *top-down* – até que se inicie a aquisição da escrita alfabética que, impondo sensibilidade aos fonemas, interrompe o fluxo de cima para baixo, visto que ainda resta ser acessada a camada correspondente à ramificação silábica (SEYMOUR, 1997). Assim, um processamento do tipo *bottom-up* se interpõe ao *top-down*. Este fato acaba implicando em uma percepção dos fonemas como unidades isoladas, que não necessariamente se relacionam em termos de restrições fonotáticas, uma vez que não se tem acesso claro à informação de que estão *ligados* ou *associados* a constituintes mais altos, e, por isso, não estão submetidos a uma organização estrutural explícita – papel por excelência desempenhado pelas sílabas dentro da gramática fonológica. Sendo assim, tais elementos acabam sendo mais suscetíveis à omissão, à substituição ou a alterações de ordem ao serem representados graficamente.

As bases para a elaboração do presente trabalho foram erigidas a partir dos estudos já evocados e, de forma especial, a partir de estudos exploratórios realizados em âmbito de iniciação científica e de pós-graduação no Grupo de Estudos sobre Aquisição da Linguagem Escrita (GEALE)⁵, entre o período de 2014 a 2019 (PACHALSKI et al., 2014, 2015; PACHALSKI; MIRANDA, 2016, 2017, 2018, 2019a), os quais estiveram todos orientados pela temática homônima a desta pesquisa, a saber, a grafia de sílabas complexas na aquisição da escrita. Eles estão inseridos, portanto, naquela que é a principal linha de investigação do GEALE, caracterizada pelo objetivo de descrever e analisar as relações simétricas e/ou assimétricas entre o conhecimento fonológico infantil e as grafias alfabéticas de crianças em período de aquisição da escrita, especialmente aquelas grafias em que são verificados erros (orto)gráficos⁶.

⁵ O GEALE é um Grupo de Pesquisa vinculado ao CNPq, criado em 2001 e ligado ao programa de Pós-Graduação em Educação da Faculdade de Educação da UFPel.

⁶ O uso de parênteses, isolando o elemento de composição “orto-”, é uma constante nos estudos do GEALE, e tem a finalidade de demarcar a diferença existente entre erros relacionados às normas do sistema ortográfico propriamente dito (*orto*) – os quais envolvem relações múltiplas entre fonemas e grafemas, definidas contextual ou arbitrariamente – e erros produzidos na fase inicial do desenvolvimento da escrita, muitas vezes motivados por questões representacionais (*gráficas*) ou por

Esses estudos procuraram dialogar com investigações produzidas na área de aquisição e de teoria fonológica formal, especificamente aquelas que tratam sobre a sílaba com constituintes ramificados no Português Brasileiro (PB), dentre as quais destacam-se as de Ribas (2002), que versam sobre a aquisição de *onset* complexo em crianças com desenvolvimento fonológico típico e atípico; as de Zitzke (1998, 2001), Lamprecht (2002) e Redmer (2007), sobre a ocorrência da metátese na aquisição fonológica típica; a de Mezzomo (2004), sobre a aquisição das codas em crianças com desenvolvimento fonológico típico, e a de Bisol (1999), que propõe uma caracterização da sílaba, de seus constituintes e de seu molde para o PB, com base no modelo não-linear de Selkirk (1982).

Os mesmos estudos também dialogaram com pesquisas situadas na área de aquisição da escrita, especialmente aquelas que têm como objeto a grafia de sílabas complexas: Miranda (2009a, 2009b, 2012, 2019), Miranda e Cunha (2013) e Abaurre (1991) foram referências para tratar das discussões em torno da grafia da rima ramificada; Coelho (2016) auxiliou a pensar o fenômeno da metátese⁷ na escrita inicial; Abaurre (2001) apresentou pistas para analisar quaisquer fenômenos que tenham a sílaba como domínio principal, e Santos (2013) foi a referência disponível para tratar da grafia do *onset* complexo, embora seu estudo tenha como substrato dados de crianças portuguesas.

Considerando-se o exposto até aqui, o presente estudo tem por objetivo geral descrever e analisar a grafia de *onset* e de rima ramificados em textos de crianças em fase de aquisição da escrita, estabelecendo relações com a aquisição fonológica dessas estruturas e com a teoria fonológica formal, a fim de responder ao problema de pesquisa formulado anteriormente, qual seja, *o que motiva as crianças a omitir, inserir, substituir ou alterar a ordem dos segmentos de estruturas com onset e rima ramificados na escrita, se elas já adquiriram fonologicamente tais estruturas?*

alguma influência da fala, isto é, referentes à fonologia da língua. Para uma discussão sobre as diferenças entre sistema gráfico e ortográfico, ver Ferreira [2001] (2013), e para uma discussão sobre a proposta de classificação dos erros conforme essas duas naturezas, ver Miranda (2014, 2020).

⁷ Fenômeno no qual se verifica um reordenamento em uma sequência de sons, como em [‘pros.tɛ.tu] para ‘protesto’ (dado extraído do BATALE).

A esse intento mais abrangente, associam-se os seguintes objetivos específicos⁸:

- a) descrever e analisar, quanti e qualitativamente, erros (orto)gráficos nos quais se verificam a alteração de estruturas de sílaba com *onset* e rima ramificados em textos espontâneos produzidos por alunos de 1º a 5º ano do Ensino Fundamental;
- b) caracterizar os erros (orto)gráficos partindo da proposta de categorização do GEALE⁹, a fim de delimitar a sua natureza (fonológica e/ou ortográfica);
- c) verificar e comparar, quanti e qualitativamente, o efeito das variáveis *tipo de estrutura silábica*, *tipo de segmento*, *tipo de erro (orto)gráfico* e *ano escolar* sobre os erros ortográficos;
- d) discutir sobre a natureza dos erros (orto)gráficos relacionados às sílabas complexas;
- e) discutir sobre simetrias e assimetrias existentes entre a aquisição da fala e da escrita das estruturas silábicas complexas;
- f) discutir sobre simetrias e assimetrias existentes entre a fonologia infantil e a fonologia adulta;
- g) propor uma formalização para as operações envolvidas na conversão das informações fonológicas *implícitas* em conhecimento *explícito*, especificamente no que se refere à representação subjacente da sílaba em crianças em fase de aquisição da escrita.

A justificativa para a investigação proposta está amparada, sobretudo, por se tratar de uma pesquisa com potencial de preenchimento de uma lacuna existente no contexto científico brasileiro no que diz respeito ao tema da grafia das sílabas complexas na aquisição da escrita. Existem alguns (e já referidos) estudos que

⁸ No capítulo 3, que se ocupa da metodologia deste trabalho, os objetivos específicos são retomados e explanados de forma mais detalhada especialmente em relação às características dos sujeitos da pesquisa, às categorias de análise e às variáveis controladas. Além disso, considerando que para cada variável monitorada em uma pesquisa há pelo menos uma pergunta subjacente (MARTINS, 2011) – do tipo “x influencia y?” – no capítulo 3 também são apresentadas o que se poderia chamar (sub)questões de investigação, acompanhadas de suas respectivas hipóteses, que decorrem das variáveis controladas e que, tomadas em conjunto, auxiliam na formulação de uma resposta à pergunta geral de pesquisa e na verificação da hipótese correspondente a ela, ambas apresentadas nesta introdução. Vale ressaltar que cada (sub)questão de investigação contempla um ou mais objetivos específicos, conforme explicitado na metodologia.

⁹ Cf. Miranda (2017, 2020).

versam sobre a temática (cf. ABAURRE, 1991a, 2001, [1988] 2011; MIRANDA, 2009a, 2009b, 2012, 2019a; MIRANDA; CUNHA, 2013; PACHALSKI et al., 2014, 2015; PACHALSKI; MIRANDA, 2016, 2017, 2018, 2019), mas, como observa Soares (2016), a maioria tem caráter exploratório ou descritivo, focalizando um tipo de estrutura específica e/ou algum fenômeno relacionado a elas. Além disso, é dada maior ênfase à grafia da rima ramificada. Falta, portanto, uma investigação mais abrangente, que descreva, analise e indique motivações para a ocorrência de erros que envolvem a grafia das estruturas com *onset* e rima ramificados, aprofundando e sistematizando as pistas, hipóteses e discussões já levantadas por outras pesquisas. É nesse sentido, também, que a execução deste projeto oportuniza a desejável continuidade de uma trajetória investigativa originada na iniciação científica.

Desta feita, o trabalho está estruturado, no texto que segue, da seguinte forma, além desta introdução:

- a) no capítulo 2, encontra-se a fundamentação teórica, que contempla as principais teorizações das diferentes áreas abordadas no trabalho, desenhando o quadro teórico no qual ele está situado. São tecidas considerações sobre a aquisição da linguagem falada e escrita, a consciência epilinguística e a metalinguística, o sistema ortográfico, a teoria da sílaba e a aquisição fonológica e escrita de estruturas silábicas;
- b) no capítulo 3, é descrita a metodologia da pesquisa, caracterizando-a quanto à sua natureza e esclarecendo sobre seus principais aspectos: os sujeitos, os dados, as variáveis e os procedimentos de análise;
- c) no capítulo 4, são descritos e discutidos os resultados obtidos a partir dos procedimentos de análise de dados adotados, conforme expostos no capítulo 3. A descrição e discussão dos resultados está organizada em 5 subseções, tendo como roteiro as questões de investigação apresentadas na Tabela 2 da metodologia;
- d) no capítulo 5, são feitas as considerações finais, orientadas pela retomada das questões de investigação e da maneira como elas foram respondidas no desenvolvimento da pesquisa. Também são expostos tanto os limites que o trabalho apresenta quanto as possibilidades investigativas que ele desperta.

2 Fundamentação teórica

Neste capítulo, são expostas as principais questões teóricas concernentes às diferentes áreas de estudo nas quais esta pesquisa se situa, de forma a delimitar o seu domínio teórico e a definir bases conceituais que sustentem a descrição, a análise e a discussão dos dados.

Inicialmente, são feitas considerações sobre a aquisição e o processamento da linguagem humana, aspectos que suscitam históricos debates no âmbito das Ciências Cognitivas e da Linguística. Neste tópico abrangente, estão incluídas as discussões entre inatismo e empirismo na aquisição da linguagem, o formato que assume o conhecimento linguístico na mente humana e a natureza do conhecimento linguístico em termos de atenção cognitiva. Em seguida, a aquisição fonológica e a aquisição da escrita recebem um tratamento diferenciado, por meio de que se busca evidenciar as especificidades envolvidas em cada um desses processos que integram a aquisição da linguagem de modo mais amplo, bem como a relação que pode se estabelecer entre ambos. Nesse contexto, considerações específicas também são feitas relativamente às relações entre aquisição da escrita e a consciência fonológica, e, por último, ao funcionamento e ao desenvolvimento do sistema ortográfico do Português Brasileiro. Depois, uma seção é dedicada à revisão das principais teorias existentes sobre a representação silábica na Fonologia, preparando o caminho para a revisão dos estudos que se ocuparam de descrever e analisar, primeiro, a aquisição fonológica de estruturas silábicas complexas, e, por último, a sua aquisição na escrita.

2.1 Aquisição e processamento da linguagem

A constituição do campo da Aquisição da Linguagem no âmbito da ciência moderna tem como um de seus mais reconhecidos impulsos a reviravolta epistemológica ocasionada pelo paradigma gerativista, cujo estabelecimento se deve

de forma estratégica à publicação da resenha crítica de Noam Chomsky (1959) em resposta à obra *Verbal Behavior*, de B. F. Skinner (1957) (SLOBIN, 1980; LYONS, 1973; OTHERO; KENEDY, 2019).

Skinner (1957) propunha uma definição de linguagem como *comportamento verbal* (o que seria uma tradução literal do título da obra para o português), cuja aprendizagem e uso ocorreriam por meio de mecanismos de estímulo-resposta, sem o suporte de qualquer recurso inato. Com essa proposição, Skinner expressava de forma clara o que se configurava como *mainstream* epistemológico tanto na Linguística quanto na Psicologia até meados da década de 1950: o empirismo comportamentalista. Por essa perspectiva, a mente humana seria, em seu estado inicial, uma *tabula rasa*, preenchida progressivamente com os dados da experiência sensível (COPLESTON, 1994). A pesquisa científica deveria ser, portanto, de base indutiva, partindo dos dados empíricos para chegar a conclusões de ordem teórica (MARCONI; LAKATOS, 2017).

A corrente principal da Linguística desenvolvida nesse período, conhecida pelo rótulo de *estruturalismo norte-americano*, fundava-se no empirismo, tendo como marca o desenvolvimento de um método que tornasse rápida e eficiente a descrição das línguas. Franz Boas, Edward Sapir, Leonard Bloomfield e Zellig Harris (este último, professor de Chomsky) são nomes importantes desse movimento que evitava o apelo à intuição dos falantes e à criação de hipóteses *a priori*, deixando que, da rigorosa descrição linguística, emanassem as categorias e o modo de funcionamento que compunham a gramática de cada língua particular (LYONS, 1970; ILARI, [2004] 2011).

A partir de sua crítica à obra de Skinner (1957) – e, em conjunto, ao estruturalismo norte-americano – Chomsky (1959, 1965, 1988) elabora uma adaptação da paráfrase moderna ao *Problema de Platão* criada por Bertrand Russell, que tomou emprestado do pensador grego o questionamento de como pode o ser humano saber tanto com um número tão pequeno de evidências e com tão pouco tempo de vida¹⁰. Chomsky, com base nisso, cunha o que ficou conhecido como o

¹⁰ “Platão ilustrou o problema com o primeiro experimento psicológico do qual se tem registro (pelo menos, um ‘experimento mental’). Em *Mênon*, Sócrates demonstra que um menino escravo sem instrução sabia os princípios da geometria ao conduzi-lo, por meio de uma série de perguntas, à descoberta de teoremas da geometria. Este experimento levanta o problema que ainda está conosco: como pôde o menino escravo estar apto a encontrar as verdades da geometria sem instrução ou informação?” (CHOMSKY, 1988, p. 4, tradução nossa).

“problema lógico da aquisição”: como o ser humano pode aprender um mecanismo tão complexo como a linguagem com tão poucas evidências disponíveis e de forma tão rápida? Dessa forma, coloca as questões ligadas à natureza da linguagem e, conseqüentemente, à sua aquisição, no centro dos debates das ciências cognitivas, chamando a atenção para aspectos que se tornaram peças-chave na sustentação do *gerativismo*, o programa investigativo que ele passa a liderar e que, simultaneamente, passa a ser o paradigma dominante na Linguística, cuja influência, inclusive, se estende para além dessa área do conhecimento (LYONS, 1970).

Dentre os vários aspectos que se podem considerar importantes a respeito da teoria gerativa, aqueles que de modo especial concernem à aquisição da linguagem podem ser resumidos nos seguintes:

- a) parte-se da hipótese da *pobreza de estímulos* (CHOMSKY, 1965, 1988), isto é, da ideia de que o *input* linguístico disponível ao sujeito é fragmentado, pouco transparente e, portanto, insuficiente para guiar o processo de aquisição da linguagem, este objeto que consiste em uma complexa engenharia mental cujo domínio inconsciente as crianças manifestam rapidamente. O conhecimento internalizado da língua é, nesse contexto, chamado de *competência linguística*;
- b) é plausível supor, considerando que os estímulos externos não dão conta da tarefa de aquisição, a existência de um mecanismo interno e inato aos seres humanos que os habilitam a adquirir a linguagem, o qual é acionado pelo *input* linguístico (falado ou sinalizado) disponível no ambiente externo em que vivem os seres humanos. Tal mecanismo, que ficou conhecido como *Dispositivo de Aquisição de Linguagem (LAD)*, na sigla em inglês), é modular, não possui interface necessária com outros módulos cognitivos e pode ser comparado a um órgão do corpo, tal como o sistema de visão ou de digestão (CHOMSKY, 1965, 1988; BERWICK; CHOMSKY, 2016). No LAD, tem-se especificada uma *Gramática Universal (GU)*, que consiste em um conjunto de propriedades gerais que regem a construção das gramáticas das línguas particulares;
- c) existe um *período crítico* para que a aquisição ocorra, isto é, o órgão da linguagem, como qualquer outro órgão corporal, tem um período maturacional que depende da sua correta estimulação para que se

desenvolva adequadamente. Caso contrário, a função que o órgão exerce estará comprometida (CHOMSKY, 1988; BERWICK; CHOMSKY, 2016).

Nas palavras do próprio Chomsky (1988, p. 17, tradução e interpolações nossas):

[...] o problema [de Platão] é resolvido em termos de certas propriedades da mente/cérebro e certos traços do ambiente linguístico. [...] A interação desses fatores produz um sistema de conhecimento que é representado na mente/cérebro, como o estado maduro da faculdade da linguagem. Esse sistema de conhecimento provê a interpretação de expressões linguísticas, incluindo novas [expressões] que a criança em fase de aquisição da linguagem não ouviu ainda.

As proposições teóricas divergentes ao gerativismo procuraram elaborar suas críticas especialmente em torno dessas teses, recorrendo a uma postura que começava a se tornar tendência na dinâmica mais ampla da Academia, de acordo com Koch e Cunha-Lima (2004): a retomada dos aspectos socioculturais, interacionais e externos aos indivíduos, em reação a filosofias racionalistas e imanentistas que, com a consolidação da Linguística moderna nas bases chomskianas, completaram sua hegemonia no cenário científico ocidental.

Tal fato suscita e consolida uma das controvérsias mais características da área de Aquisição da Linguagem até os dias atuais: o debate “*nature versus nurture*” (cf. SLOBIN, 1980; FERRARI; FRANÇA; MAIA, 2016), ou seja, o debate sobre quais os aspectos de maior proeminência quando se fala em natureza e aquisição da linguagem: processos endógenos ou exógenos aos indivíduos? Estruturas inatas ou adquiridas/aprendidas? Aspectos biológicos ou sociais? Notavelmente, a teoria gerativista assume o lado *nature* no debate, enquanto, de forma geral, teorias díspares assumem o lado *nurture*, as quais incluem versões remodeladas do empirismo clássico. Na verdade, segundo Slobin (1980), considerando a data de seus escritos até os dias atuais, o que se pode aferir também em Ferrari, França e Maia (2016), “o debate passou, até certo ponto, de uma discussão entre nativismo [*nature*] e empirismo [*nurture*] para uma discussão acerca da *qualidade de nativismo* que é adequada aos assuntos de aquisição da linguagem” (SLOBIN, 1980, p. 142, interpolações e grifos nossos).

É nessa direção, por exemplo, que acontece o célebre debate organizado por Piattelli-Palmarini [1979] (1983) entre Noam Chomsky e Jean Piaget. Ambos os pesquisadores são críticos do empirismo como teoria da aprendizagem humana (e da linguagem, conseqüentemente), e concordam no sentido de que nenhum

conhecimento é possível sem que uma espécie de organização cognitiva esteja presente desde o início do desenvolvimento humano e sem que haja algum tipo de interação entre indivíduo e meio ambiente (PIATELLI-PALMARINI, [1979] 1983). A diferença, entretanto, que os coloca em lados distintos no debate *nature* vs. *nurture*, é relativamente a que espécie de estrutura cognitiva, afinal, está disponível ao sujeito para que ele adquira a linguagem.

Para Chomsky, como se aferiu anteriormente, o estado inicial da linguagem é caracterizado pela existência de uma Gramática Universal: o conjunto de princípios, regras, restrições ou operações que regulam a aquisição das gramáticas particulares de cada língua e orientam as crianças à extração das informações relevantes do seu ambiente linguístico. O papel do ambiente é minimizado (embora não eliminado), restando a ele a função de desencadear ou *amadurecer* as estruturas da GU. Trata-se, portanto, de uma visão modularista, ou de domínio específico, do desenvolvimento cognitivo.

Para Piaget, por outro lado, não há organização cognitiva específica para a linguagem (tampouco para qualquer outra área específica de conhecimento), apenas mecanismos gerais, funcionais e sensoriais de inteligência que permitem à cognição gerar estruturas de conhecimento de forma progressiva. O indivíduo, assim, age sobre os dados linguísticos por meio de esquemas de ação, os quais progressivamente geram, modificam e complexificam essa estrutura cognitiva. A visão piagetiana insere-se na perspectiva não-modularista, ou de domínio geral, a respeito do desenvolvimento cognitivo.

Não haveria, contudo, alguma maneira de conciliar tais perspectivas sobre o desenvolvimento da linguagem humana ou a ciência estaria destinada a sempre se colocar em polos antagônicos no debate a respeito dos objetos de investigação? Procurando responder a essa pergunta e minimizar as diferenças aparentemente incompatíveis entre as perspectivas debatidas em Piatelli-Palmarini [1979] (1983), Karmiloff-Smith (1994) propõe um modelo da organização e do desenvolvimento cognitivo que, neste contexto, também se torna útil para compreender a maneira como a informação linguística pode ser processada, armazenada e representada na mente humana.

Com o Modelo de Redescrição Representacional (MRR), Karmiloff-Smith (1994) sustenta que a linguagem se desenvolve no ser humano sob a condição de haver algum tipo de informação inata disponível, embora em graus mais moderados

e menos especificados. São necessárias restrições iniciais de domínio específico, que orientem a extração de informação relevante do *input*. No entanto, para que haja flexibilidade representacional, é necessário assumir também que existem processos de domínio geral, tal como sustentam perspectivas de base piagetiana. Além disso, para a autora, mais importante que o feito de se tornar um usuário eficiente da língua em um curto período de tempo – fato central para a teoria gerativa –, é “a potencialidade de tomar as próprias representações como objeto de atenção cognitiva” (KARMILOFF-SMITH, 1994, p. 52), isto é, de analisar o próprio conhecimento adquirido.

Para tanto, o MRR sugere que o desenvolvimento cognitivo ocorre por meio de um processo iterativo (cíclico) de *redescrição representacional*, o qual progressivamente transforma representações que estão implícitas *na* mente em representações explícitas, disponíveis *para* a mente. Segundo Karmiloff-Smith (1994, p. 40), “as redescrições são abstrações em uma linguagem de nível superior”, que se sucedem por motivações endógenas e flexibilizam as representações mentais, colocando-as à disposição de outros domínios cognitivos. No MRR, o que se sugere como sendo de domínio geral é justamente a *redescrição representacional*, que ocorre em todos os diferentes domínios específicos da mente da mesma maneira, recodificando informações a fim de torná-las acessíveis à consciência. Nas palavras de Landsmann (1995, p. 193),

A *redescrição* é uma operação interna que preserva certas características e elimina outras em benefício de um nível diferente de acessibilidade. O sistema [...] cria uma *redescrição* para poder trabalhá-la, trabalhar suas partes componentes.

Para Karmiloff-Smith (1994), existem diferentes níveis representacionais que correspondem aos diferentes formatos que o conhecimento pode assumir na mente.

O primeiro nível é nomeado pela autora **Implícito (nível I)**. Neste momento, o foco do sujeito está no dado externo, na maneira como esse dado será computado, internalizado. Não há condições de centrar a atenção nas representações mentais, de manipulá-las de forma consciente e de haver interface entre o dado externo, que está sendo computado, e as representações que já estão na mente. Ou seja, neste estado inicial, há independência entre as representações inter e intra domínios, as quais apresentam pouca flexibilidade. As representações estão codificadas *entre parênteses*, para utilizar a metáfora da autora. Esse tipo de representação, de nível I, possibilita o desenvolvimento de habilidades procedimentais, que geram a execução

bem-sucedida de determinadas condutas/procedimentos (por exemplo, a produção de formas linguísticas sintáticas, fonológicas, morfológicas etc., ou o encaixamento de blocos de formas geométricas no espaço que corresponde às formas), levando àquilo que Karmiloff-Smith (1994) denomina de *maestria procedimental*, o ponto culminante do nível I.

A partir desse momento de estabilidade, as representações de nível I estão prontas para serem redescritas em um formato de nível **Explícito 1 (nível E1)**, o qual é caracterizado por comportar representações um pouco mais flexíveis, que não estão mais entre parênteses, capazes de criar vínculos com outras representações tanto intra quanto inter domínios. Contudo, tal disponibilidade ou explicitude é relativa, pois, segundo a autora, no nível E1 as representações tornam-se disponíveis ou explícitas apenas ao sistema cognitivo e não ao sujeito. Por esse motivo, embora ainda não se possa falar em consciência propriamente dita, já é possível constatar um funcionamento metaprocedural. Uma característica que, para Karmiloff-Smith (1994), sustenta a existência deste nível é a ocorrência de “erros tardios”. Depois de um período em que um sujeito apresenta maestria procedimental em relação a determinado conhecimento, é possível observar que ele passa a cometer alguns erros que antes não cometia, dando origem à *curva em U* da aprendizagem (STRAUSS; STAVY, 1982).

Para exemplificar, a autora relata um experimento feito com crianças cuja língua nativa é a ASL (Língua Americana de Sinais) e cujos pais, também usuários de ASL, aprenderam a língua tardiamente. As crianças, depois de um período de tempo apresentando maestria na execução de um conjunto de sinais, passaram a cometer erros na sua produção, os quais consistiam na realização fragmentada de sinais cuja realização deveria se dar em um movimento único e contínuo, a fim de separar marcadores morfológicos. Como ilustra Karmiloff-Smith (1994), seria semelhante a segmentarmos na fala uma palavra observando as fronteiras morfológicas, como ‘pint-a-r’. No entanto, ao serem indagadas sobre o motivo que as levava a fragmentar os sinais, as crianças não foram capazes de responder, isto é, simplesmente executaram as novas formas.

Esse tipo de conduta demonstraria justamente que o *sistema cognitivo* está lidando com as representações de forma distinta, manipulando-as e analisando-as em suas partes componentes, o que gera novas condutas – “erradas”, do ponto de vista da forma-alvo esperada, mas precisas se for levado em consideração o tipo de

conhecimento e de processamento que está sendo mobilizado. Por outro lado, o fato de as crianças não conseguirem explicitar verbalmente as razões pelas quais realizam tais condutas, demonstra, segundo Karmiloff-Smith (1994), que se trata de um conhecimento ainda não acessível ao sujeito, apenas disponível para o sistema. Nos termos do MRR, o que acontece é que representações antes implícitas, isto é, não disponíveis à manipulação tanto pelo sistema cognitivo quanto pelo próprio sujeito, que geravam condutas corretas (como na execução dos sinais de forma não-segmentada), são redescritas em um formato que passa a ter um grau inicial de explicitação, por meio do qual o sistema cognitivo passa a conseguir analisá-las e relacioná-las com outras representações.

Os últimos níveis do MRR são os **Explícitos 2 e 3 (nível E3 e E3)**, caracterizados pelo acesso consciente às representações e pela capacidade de explicitar verbalmente os motivos que justificam a realização de determinadas condutas em detrimento de outras. Karmiloff-Smith (1994) apresenta-os de forma conjugada, visto que a sua distinção é muito sutil e que não existem estudos que se dedicam a analisar especificamente o nível E2. A grande marca que evidenciaria o nível E2/E3 são as “teorias verbalmente expressáveis”, ou seja, as explicações dadas pelos sujeitos para justificar a motivação de determinado tipo de conduta. Para exemplificar, é possível retomar o exemplo anterior, das crianças usuárias da ASL. O que evidenciaria que suas representações estão codificadas em formato E2/E3, acessíveis à consciência, seria uma resposta explicativa ao motivo pelo qual realizam a segmentação dos sinais a partir dos marcadores morfológicos, uma resposta que demonstraria uma capacidade metalinguística da criança, a elaboração de uma espécie de teoria sobre aquilo que ela está executando.

É importante observar que a redescrição representacional, por se tratar de um processo cíclico, não implica na existência de *estágios* de desenvolvimento. É possível, portanto, que um sujeito esteja em um nível I para certos conhecimentos relacionados à sintaxe, por exemplo, e em um nível E1 ou E2 para outros aspectos associados ao mesmo módulo linguístico.

Com o MRR, além de procurar superar ou mesmo conciliar as dicotomias *modular vs. não-modular*, *nature vs. nurture* e *racionalismo vs. empirismo*, Karmiloff-Smith (1994) também busca problematizar a dicotomia *inconsciente vs. consciente*. Para a autora, o acesso consciente à informação antes implícita se dá de forma progressiva e não-automática, motivo pelo qual são propostos os 3 níveis distintos

correspondentes à explicitação da informação. Essa progressiva passagem da informação linguística a formatos distintos de representação mental, responsáveis por sua abertura à consciência dos sujeitos, também é descrita na literatura psicolinguística como desenvolvimento metalinguístico (GOMBERT, 1992), o qual envolve a passagem de atividades *epilinguísticas* a atividades *metalinguísticas*, caracterizadas, respectivamente, pelos conhecimentos de natureza implícita e explícita, na terminologia de Karmiloff-Smith (1994).

Uma atividade epilinguística é aquela que evidencia uma “sensibilidade funcional do sujeito às regras de organização e de uso da língua” (GOMBERT, 1992, p. 13), podendo ser observada desde os primeiros anos de vida. Um exemplo são os juízos de gramaticalidade emitidos sobre uma sentença como “sentada vermelho a estava chapeuzinho”, sendo, por isso, facilmente rejeitada por falantes do Português, inclusive por crianças em idade precoce (cf. MALUF; GOMBERT, 2008). O conhecimento que possibilita esse tipo de atividade é caracterizado como tácito e implícito, pois não é acessível à consciência, tampouco é manipulável; ele apenas habilita o sujeito ao uso eficaz de um procedimento. Na palavras de Karmiloff-Smith (1994, p. 71), é “informação *na* mente e, ainda não, conhecimento *para* a mente”.

Uma atividade metalinguística, por sua vez, é aquela na qual o sujeito evidencia algum grau de consciência e de controle deliberados sobre a própria linguagem (GOMBERT, 1992), o que pode se manifestar em condutas que envolvem a manipulação de sílabas, por exemplo, comum em testes de consciência fonológica e brincadeiras infantis. Neles, as crianças recebem um comando para a omissão de alguma sílaba de uma palavra, e a sua tarefa é identificar a palavra gerada a partir de tal omissão: “se tirarmos a sílaba ‘pi’ de ‘piolho’, qual palavra fica?”.

Para Gombert (1992), o desenvolvimento de habilidades metalinguísticas é decorrente de demandas externas, que forçam a explicitação – a “abertura” – de conhecimentos antes implícitos na mente. A aprendizagem da escrita e da leitura é amplamente considerada como a principal desencadeadora desse processo (cf. MORAIS et al., 1979; ALEGRIA et al., 1982; READ et al., 1986 especialmente sobre os efeitos recíprocos entre habilidade metafonêmica e aprendizagem da escrita alfabética), uma vez que torna visíveis e acessíveis formas que, para as crianças, são bastante abstratas e para as quais sua atenção não estava voltada.

Apesar disso, não recai sobre a aprendizagem do sistema de escrita a exclusividade de desencadeamento de atividades metalinguísticas. Conforme

Gombert (1992, p. 190, tradução nossa), com o que é possível fazer equivaler pressupostos do modelo de Karmiloff-Smith (1994), antes de qualquer demanda externa, “o pré-requisito absoluto para essa consciência é o controle epilinguístico. Apenas aquilo que já foi dominado em um nível funcional também o será em um nível consciente”.

2.1.1 A aquisição fonológica

Na subseção anterior (2.1), foi realizada uma breve revisão sobre a gênese e o desenvolvimento do campo de estudos da Aquisição da Linguagem. Foram assinalados os pontos e os conceitos que, para a construção do quadro teórico desta dissertação, são os mais relevantes. Um deles diz respeito aos embates das distintas epistemologias que sustentam as teorias linguísticas, os quais emergem de modo particular quando estas procuram responder à maneira pela qual esse sistema de conhecimento que é a língua origina-se e desenvolve-se na mente dos seres humanos. O debate *nature/nurture* sintetiza as divergências que estão em jogo na explicação sobre como ocorre a aquisição da linguagem, alternando-se historicamente em termos da proeminência que assumem no contexto científico.

Para Macken (1996), não é diferente na área específica de aquisição fonológica, uma vez que ainda há uma polarização entre perspectivas empiristas (*conexionistas*, atualmente), geralmente associadas ao trabalho de foneticistas, e racionalistas, cuja tendência é de remeter ao trabalho de fonólogos formalistas ou cognitivistas. Nesse sentido, ainda conforme a autora, um dos maiores desafios para os estudiosos da área trata-se de

delimitar adequadamente os domínios entre a fonologia e a fonética, determinar a interação entre aprendizado e restrições inatas, e separar o aprendizado geral do processo e da estrutura linguística de domínio específico (MACKEN, 1996, p. 2, tradução nossa).

A aquisição fonológica, partindo dessa perspectiva, pode ser tida como um processo pelo qual passam as crianças e que delas demanda pelo menos a resolução de duas tarefas primordiais segundo Kiparsky e Menn (1977, p. 47, tradução nossa):

Primeiramente, a criança precisa dominar o repertório fonético da língua. Subsequentemente, a criança aprende as regras fonológicas que representam as regularidades que governam a variação de suas palavras e morfemas. Cada realização envolve habilidades perceptuais e de produção.

À despeito da discussão envolvendo o estatuto de *regras fonológicas* e do funcionamento das gramáticas, a ideia central é a de que, no curso da aquisição fonológica, as crianças precisam aprender como a língua da comunidade de falantes da qual participam discretiza o contínuo fonético, isto é, como a língua organiza o som em unidades discretas que vão compor o seu inventário fonológico e se articular com outros componentes da gramática, de modo a permitir à criança a interpretação de quaisquer expressões de sua língua materna, de forma especial aquelas que nunca ouviu (CHOMSKY, 1988). Matzenauer (2015) trata como caso de *fonologização* o processo de aprender a interpretar unidades físicas (contínuo fonético) como categorias gramaticais (fonológicas).

Além disso, para Macken (1992), assim como para Kiparsky e Menn (1977), adquirir a fonologia de uma língua é uma tarefa complexa, caracterizada por um dinamismo que, ao mesmo tempo, é atravessado por regularidades. Tendo isso em vista, os autores entendem ser necessário conceber a criança em período de aquisição fonológica como um sujeito ativo, ainda que não consciente de suas ações, que está constantemente solucionando problemas por meio da elaboração, testagem e reformulação de hipóteses sobre o funcionamento da gramática fonológica de sua língua materna.

Para explicar como ocorre a fonologização do som, Matzenauer (2015) se vale do Modelo Bidirecional de Fonologia e Fonética – BiPhon, proposto por Boersma (2007). Trata-se de um modelo de processamento de língua materna e também de gramática, o qual pode oferecer um caminho para a elaboração de uma promissora resposta ao desafio proposto por Macken (1996). Nele, leva-se em consideração percepção e produção linguísticas (por esse motivo identificado como bidirecional) bem como a interface entre fonética e fonologia. A Figura 1 sintetiza os principais aspectos que envolvem o processamento fonológico conforme preconizados pelo Modelo BiPhon:



Figura 1 – Modelo BiPhon.

Fonte: adaptação de Boersma e Hamann (2009).

A Figura 1 exhibe que o Modelo BiPhon opera a partir de dois módulos inter-relacionados, um responsável pela compreensão (ou percepção) e outro responsável pela produção da informação fonológica. Ambos são formados por três níveis distintos de representação: a forma fonética, que integra informações de ordem acústica e articulatória; a forma de superfície, que integra a estrutura fonológica em seus níveis de organização melódico e prosódico (traços, segmentos, sílabas, pés...¹¹), e a forma subjacente, que constitui a camada mais abstrata, relacionando unidades discretas da fonologia com o léxico e a morfologia. Os níveis de representação são articulados por meio das restrições de Pista, de Estrutura e de Fidelidade. Uma das virtudes do Modelo BiPhon é captar as ideias de construção e de gradatividade características da aquisição fonológica (e de qualquer processo de aquisição), tal como assinaladas por Macken (1992) e Kiparsky e Menn (1977). Há um constante processo de ajuste e estruturação da informação, capturado por meio da interação tanto entre os diferentes níveis de representação, através das restrições, quanto entre os módulos de Compreensão e de Produção.

Macken (1979, 1992) propõe que a palavra fonológica possui um papel de destaque especialmente nos estágios iniciais de aquisição, período no qual as

¹¹ Em Nespor e Vogel [1986] (1994), por meio da Hierarquia Prosódica, há uma proposta de formalização das unidades que compõem a camada prosódica das línguas. A unidade que ocupa a base da hierarquia é a sílaba (σ), seguida do pé métrico (Σ), da palavra fonológica (w), do grupo clítico (C), da frase fonológica (ϕ), da frase entonacional (I) e do enunciado (U). Alguns princípios gerais regem a relação entre os constituintes. Dentre eles, destaca-se o da proeminência relativa, que estabelece relações de dominância entre os constituintes de cada camada, aos quais é atribuído um valor forte e um (ou mais) valor(es) fraco(s).

crianças manifestam a formação de *templates* prosódicos, caracterizados por certos padrões que constituem o substrato linguístico para a generalização de regras de atribuição de acento e da fonotática da língua. Segundo a autora, as crianças formulam hipóteses sobre a composição da gramática sonora com base em similaridades globais que esses *templates* apresentam, incorporando novas palavras tendo-os como referência. À medida que avançam no desenvolvimento fonológico, esses *templates* passam a ser analisados em unidades menores, isto é, em termos de segmento e de traços, componentes fundamentais para a elaboração de uma gramática fonológica.

Miranda (2009b) e Matzenauer e Miranda (2012), por exemplo, ancoram-se em Macken (1992), a fim de explicar o comportamento inesperado de uma criança pequena em relação à produção da fricativa coronal em coda final. Chamou a atenção das autoras, ao analisarem os dados da criança, a produção correta do segmento em posição de coda final ([‘dojs] para ‘dois’) em uma idade considerada muito precoce, tendo em vista o que é relatado na literatura da área (ver subseção 2.4.1) para dar a fricativa como adquirida na referida posição. Para resolver essa assimetria, as autoras se valem da proposta de Macken (1992), interpretando a “produção precoce de palavras com [s] final não como a emergência de codas no sistema da criança, mas como resultado de uma operação que envolve uma unidade prosodicamente mais alta que a sílaba” (MIRANDA, 2009b, p. 120), ou seja, a palavra fonológica. Desta forma, o segmento estaria ligado diretamente à palavra fonológica e não à estrutura silábica, compondo o *template* lexical ainda não efetivamente analisado em termos de suas estruturas mínimas, tanto no que diz respeito aos traços constitutivos do segmento quanto ao constituinte silábico específico que ele ocupa. A proposta de Macken (1979, 1992) pressupõe que o processamento da informação ocorra, ao menos nos estágios mais iniciais de aquisição, na modalidade *top-down*, por meio da qual o conhecimento é (re)construído partindo de unidades maiores até chegar no que seriam as menores.

A ideia de que esse tipo de processamento seja predominante no período inicial de aquisição fonológica também recebe amparo de estudos como os de Morgan e Demuth (1996), Christophe, Guasti e Nespor (1997) e Gout e Christophe (2006), os quais argumentam em favor de que a informação prosódica, especialmente aquela ligada aos constituintes mais altos da hierarquia, desencadeia e orienta não somente a organização e construção das representações fonológicas pela criança, mas

também as representações de outros níveis linguísticos, especialmente a sintaxe e o léxico. É a chamada teoria do *bootstrapping*¹² prosódico. Os autores relatam uma série de experimentos que evidenciam que diferentes tipos de informação prosódica, como acento e fronteiras de frase fonológica e/ou de frase entonacional, são utilizadas pelos bebês de até 12 meses como pistas decisivas para identificar parâmetros de sua língua materna como o de direcionalidade do núcleo sintático (verbo-complemento ou complemento-verbo) bem como identificar palavras no fluxo contínuo da fala, as quais servem à formação do léxico.

Teorizações semelhantes que partem da análise de dados do Português podem ser aferidas em Scarpa (1997, 1999) e Santos (1995, 2001, 2017), para o PB, e Frota e Vigário (2008), para o PE. As autoras, com base em análises de dados naturalísticos transversais e longitudinais, defendem que o processo de aquisição da prosódia se dê de forma *top-down*, utilizando a frase entonacional como referência para o estabelecimento dos padrões que caracterizam as unidades mais baixas da hierarquia. Em uma faixa etária precoce, as crianças demonstram lidar com esse domínio por meio da demarcação de contornos entonacionais com sons preenchedores (*filler sounds*), os quais muitas vezes não correspondem a segmentos existentes no inventário fonológico do Português, do que decorre que a sequência segmental não é necessariamente analisada pela criança e que o acento da palavra coincide, na verdade, com o acento entonacional. É por esse motivo, também, que o padrão acentual inicialmente produzido pelas crianças seria iâmbico (acento com proeminência final), segundo Santos (2001, 2017) e Frota e Vigário (2008). Ainda, as autoras portuguesas, à luz de tais proposições, entendem que as crianças podem coincidir, em uma só produção, uma sílaba, uma palavra prosódica e uma frase, sendo que este último constituinte guia a construção dos demais.

Scarpa (1997, 1999), Santos (1995, 2001, 2017) e Frota e Vigário (2008) também chamam a atenção para o fato de que não é porque a prosódia guia e/ou desencadeia o processo de aquisição, que ela mesma, enquanto hierarquia de unidades da gramática fonológica, não vá sendo gradativamente construída. Em outras palavras, isso significa dizer que a informação prosódica não está pré-especificada na mente. O ponto central, no entanto, trata-se da fixação precoce de

¹² *Bootstrapping* é um termo em inglês cuja tradução para o português pode se dar por meio das palavras *inicialização* ou *desencadeamento*. Por ser amplamente difundido, neste trabalho será utilizado o termo em inglês.

padrões prosódicos pela criança e da utilização desses padrões como condutores da aquisição da linguagem.

Colaboram, ainda, para fortalecer a ideia do *bootstrapping* prosódico, ou ao menos a ideia de que a informação prosódica oferece pistas essenciais para guiar o processo de aquisição da linguagem, trabalhos clássicos de percepção que se valem do método de sucção não-nutritiva para demonstrar a sensibilidade do bebê recém-nascido a regularidades rítmicas e melódicas da língua. Mehler et al. (1978) demonstram que os bebês utilizam tais regularidades no reconhecimento da voz materna. Nazzi, Bertoncini e Mehler (1998) mostram que bebês de 2 meses já são capazes de diferenciar o padrão acentual da sua língua materna quando contrastada com o padrão de outra língua. O mesmo padrão acentual parece ser utilizado por crianças de 6 a 9 meses para a delimitação das fronteiras de palavra, segundo Juckzyk, Cutler e Redanz (1993).

Depois dessa rápida e profícua etapa inicial do desenvolvimento fonológico, na qual os padrões prosódicos são fundamentais para a formação de um léxico primitivo que assume a configuração de *templates*, “a aquisição será por ele dirigida” (MATZENAUER, 2015, p. 74), em um movimento bidirecional de extração e cruzamento de informação tanto do *input* quanto da própria gramática em construção. Com um léxico formado, mesmo que ainda em forma de *templates* prosódicos, hipotetiza-se, conforme Matzenauer (2015), que a criança passe a utilizá-lo como fonte de informação relevante para começar a compor o inventário segmental, ou, em outras palavras, para iniciar a fonologização de traços e segmentos¹³.

2.1.2 A aquisição da escrita

Até este ponto da fundamentação teórica, se discorreu sobre aspectos que dizem respeito à aquisição da linguagem na modalidade falada. Apesar de assumir, neste trabalho, uma perspectiva que considera a aquisição da escrita como parte integrante do processo mais abrangente de aquisição da linguagem (noção que será explorada adiante, na presente subseção, com base em Miranda (2017)), devem ser

¹³ A discussão sobre a direcionalidade do processamento e sobre as unidades fonológicas condicionadoras da aquisição é retomada na subseção 2.4.1, na qual são apresentados os percursos de aquisição fonológica de rima ramificada e de onset complexo, e por meio dos quais são oferecidas outras evidências em favor de um processamento predominantemente *top-down*, o qual interage, em momentos posteriores, com o processamento *bottom-up*.

consideradas as importantes diferenças que existem entre fala e escrita enquanto substâncias alternativas de atualização da língua (SAUSSURE, [1916] 2012; LYONS, 1968).

A mais basilar das distinções refere-se ao modo natural e espontâneo como se dá a aquisição da fala, em se comparando ao processo de ensino sistemático que a escrita alfabética exige para que seja adquirida. Por esse motivo é que encontramos adultos pouco ou não escolarizados que não sabem escrever, mas que são usuários competentes da língua na sua modalidade falada ou sinalizada. Também é pelo mesmo motivo que muitas civilizações das quais se têm notícia na história humana (algumas até hoje) não desenvolveram sistemas de notação escrita, mantendo uma rica tradição oral (GELB, 1952; LANDSMANN, 1995).

No entanto, apesar de a escrita não ser adquirida de forma natural e espontânea, ela deriva de uma capacidade, esta sim, exclusiva da espécie humana, “a capacidade *notacional*; isto é, a capacidade de utilizar ferramentas para deixar marcas permanentes de atos intencionais” (LANDSMANN, 1995, p. 111, grifos da autora). Conforme aponta Landsmann (1995), baseada em Goodman (1968), a capacidade notacional se manifesta em diferentes tipos de registro *simbólico*, que podem ser icônicos ou não, e que podem, também, constituir *sistemas* notacionais, cujas condições para sua constituição estão relacionadas diretamente àquilo que remonta a uma concepção saussureana de sistema, o qual é definido de forma imanente, ou seja, por meio da própria rede de relações opositivas que se estabelecem entre as unidades que o compõem.

Essa característica possibilita aos usuários do sistema a identificação de informações, permitindo sua reiterada reprodução, ideia que possibilita a assunção do sistema de escrita alfabética como um sistema notacional, já que ele “inclui uma série finita de caracteres”, os quais, por serem finitos, permitem “identificar se alguma marca ou inscrição produzida por alguém pertence ou não ao sistema” (LANDSMANN, 1995, p. 116). Mas o que o sistema de escrita alfabética *representa*? Seus símbolos fazem referência a que tipo de domínio? Essa espécie de indagação e a resposta dada a ela é crucial, como se verá em seguida, para compreender o processo por que passam as crianças na aquisição da escrita.

De forma geral, “o domínio de referência da escrita alfabética é a linguagem” (LANDSMANN, 1995, 119). Contudo, essa é uma caracterização por demais abrangente, uma vez que reúne em seu escopo outros tipos de sistemas de escrita,

tais como os *logográficos* (a escrita chinesa, por exemplo), cujo funcionamento opera preponderantemente com unidades de primeira articulação da língua (MARTINET, 1960), isto é, os grafismos utilizados representam unidades significativas (morfemas). A escrita alfabética compõe o grupo dos sistemas *fonográficos*, os quais operam a partir de unidades de segunda articulação, não-significativas – o caso das 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, o qual tem como princípio de funcionamento, por essa razão, as relações entre fonemas e grafemas (GELB, 1952).

Como as crianças chegam à compreensão de um sistema de escrita e quais são os processos cognitivos envolvidos em sua aquisição, desde o período pré-escolar até (e incluindo) o início de práticas de instrução intencionais, são perguntas que demoraram a mobilizar investigações, uma vez que se supunha que aprender a escrever seria uma consequência da aprendizagem da leitura, a qual era compreendida pela via associacionista de respostas sonoras a estímulos gráficos (FERREIRO; TEBEROSKY, [1984] 1999; LANDSMANN, 1995).

Com exceção dos estudos vanguardistas de Luria e Vygotsky [1929] (1988)¹⁴, é apenas na segunda metade do século XX, especialmente a partir da década de 1970, que se observa o desenvolvimento de pesquisas que buscavam sistematicamente responder a tais questões, acompanhando o *boom* dos estudos psicolinguísticos cuja impulsão está diretamente ligada ao surgimento da teoria da gramática gerativa (BALIEIRO JR., 2002).

Dentre as pesquisas que se destacam, a unanimidade recai sobre a de Ferreiro e Teberosky [1984] (1999)¹⁵, cujo objetivo esteve em oferecer respostas a perguntas

¹⁴ Segundo Landsmann (1995), Luria e Vygotsky foram os primeiros a conceber a escrita considerando haver uma *pré-história* dessa manifestação da língua do ponto de vista ontogenético, no sentido de reconhecer que, antes mesmo da instrução formal, as crianças já se aproximam da escrita com uma “predisposição pedagógica” (PREMACK, 1984). Para os pesquisadores, a escrita é essencialmente um *instrumento*, isto é, um sistema de representação que serve à transmissão de conceitos e ideias por meio do uso funcional dos seus mecanismos formais (letras, traços, pontos etc.). Dessa forma, Luria procurou principalmente observar se as crianças fazem um uso funcional da escrita (por exemplo, utilizar o registro escrito como forma de se lembrar de frases recém enunciadas oralmente), a fim de avaliar a maneira como dela se apropriam. Por meio de alguns experimentos, o autor chegou à proposição de que as crianças passam por quatro etapas em direção à compreensão da escrita nos termos de uma técnica instrumental, aprendendo a fazer progressivamente uso diferenciado e instrumental dos signos. Para uma revisão sobre as diferenças entre a teoria psicogenética e as proposições dos autores russos, conferir Ferreiro (1994), Landsmann (1995) e Soares (2016).

¹⁵ É imprescindível referir os estudos que, de forma paralela ou mesmo anterior à publicação de Ferreiro e Teberosky [1984] (1999), passaram a emergir procurando abordar a aprendizagem da escrita pelas crianças. São eles: Carol Chomsky (1970, 1971), Read (1971, 1975) e Bissex (1980), cujo ponto em

tais como o que a criança sabe sobre a escrita e quais são suas hipóteses sobre o sistema. Além disso, as autoras questionaram a maneira reducionista por meio da qual eram explicados – e consequentemente didatizados – os processos de aprendizagem de leitura e de escrita pelas crianças. Para tanto, se serviram da concepção piagetiana sobre o desenvolvimento cognitivo do ser humano e da visão chomskiana sobre a aquisição da linguagem – e aqui é necessário lembrar que, apesar de Chomsky e Piaget discordarem no que diz respeito à aquisição da linguagem, ambos se unem, por vias distintas, na crítica ao empirismo (cf. PIATELLI-PALMARINI, [1979] 1983). É este aspecto em especial que serve aos interesses de Ferreiro e Teberosky [1984] (1999), uma vez que elas se propõem a explicar os processos de compreensão da leitura e da escrita *do ponto de vista do próprio sujeito que aprende*, o que somente é possível se este sujeito for biologicamente potente. Nesse sentido, segundo as autoras, dois aspectos, salientados pelas teorias chomskiana e piagetiana respectivamente, não podem ser descuidados: a competência linguística da criança e as suas capacidades cognoscitivas (FERREIRO; TEBEROSKY, [1984] 1999).

A *competência linguística* diz respeito ao conhecimento internalizado e inconsciente que qualquer falante constrói naturalmente sobre a estrutura e o funcionamento de sua língua materna, por meio da interação de princípios gerais da Gramática Universal com o *input* linguístico recebido da comunidade de fala (CHOMSKY, 1988; LYONS, 1970; 1987). Já a ideia de *sujeito cognoscente*, refere-se à capacidade do ser humano de buscar ativamente compreender o ambiente que o rodeia, formulando hipóteses provisórias (porque sempre podem ser reformuladas) que respondam às suas indagações sobre os objetos que constituem esse ambiente. Dessa forma, mais do que meramente responder a estímulos, o sujeito *os interpreta*, transformando-os via mecanismos de *assimilação* e incorporando-os a sua estrutura cognitiva através dos mecanismos de *acomodação* (PIAGET, [1979] 1983). Nesse sentido, é importante notar que o sujeito não acumula conhecimentos, “empilhando-

comum reside no uso do conceito de *escrita inventada*, por meio do qual se referem às escritas iniciais das crianças como formas autênticas, consistentes e reveladoras das relações que estabelecem entre seu conhecimento fonológico internalizado e o funcionamento do sistema de escrita alfabética; e Gentry (1982), Frith (1985) e Ehri (1997), cuja semelhança está em tratar da aprendizagem da escrita propondo modelos de estágios ou fases. A diferença desses estudos relativamente à abordagem psicogenética reside, sobretudo, em que a última procura explicitamente vincular a compreensão sobre a aquisição da escrita a uma teoria do desenvolvimento cognitivo, enquanto as demais focam no princípio alfabético da escrita e em como as crianças vão se aproximando das correspondências fonográfêmicas adequadas ao sistema em aquisição (FERREIRO; TEBEROSKY, [1984] 1999; SOARES, 2016).

os”, por assim dizer, um por um em sua cognição; antes, o sujeito constrói uma estrutura cognitiva à medida que interage com o ambiente, lançando mão daquilo que já sabe, ou seja, daquilo que está já incorporado à sua cognição para oferecer esquemas interpretativos aos objetos que o cercam.

Partindo dessa perspectiva, uma criança que é iniciada formalmente no processo de alfabetização já carrega consigo um elaborado conhecimento de sua língua materna bem como sobre a própria linguagem escrita, visto que, participando de uma sociedade grafocêntrica, procura ativamente interpretar este objeto com o qual frequentemente se depara. Logo, conduzir as crianças à apreensão do princípio de funcionamento do sistema de escrita, como ressaltam Ferreiro e Teberosky [1984] (1999, p. 27),

não se trata de transmitir um conhecimento que o sujeito não teria fora desse ato de transmissão, mas sim de fazer-lhe cobrar a consciência de um conhecimento que o mesmo já possui, mas sem ser consciente de possuí-lo.

O caminho cognitivo percorrido pelo sujeito na compreensão dos objetos que o cercam, particularmente do sistema de escrita alfabética, é, pois, precoce e, segundo Ferreiro e Teberosky [1984] (1999), obedece a um certo padrão de sucessivas reestruturações cognitivas globais, as quais correspondem às diferentes hipóteses que elaboram as crianças sobre o valor, as características e a função da escrita. Em outras palavras, as crianças elaboram *conceitos* sobre a escrita por meio da construção de respostas que são constantemente reformuladas, a respeito *do que* é a escrita, *como* ela funciona e *por que* ela é assim, tomando como base o conhecimento já incorporado às suas estruturas cognitivas, sobretudo aquele que diz respeito à competência linguística.

As investigações que as pesquisadoras conduziram até chegar à proposição de tal percurso evolutivo foram orientadas pelo método clínico piagetiano, cuja premissa de partida é a de que o sujeito entrevistado já tem hipóteses sobre o objeto de conhecimento que está sendo investigado (cf. CARRAHER, 1989). Assim, mais importante que medir acertos e erros em relação à concepção do sujeito, é captar o processo que gera as respostas dadas, pois ele poderá evidenciar o tipo de conhecimentos e os arranjos cognitivos que estão em jogo na direção da plena compreensão do objeto.

Ferreiro e Teberosky [1984] (1999) propuseram às crianças que entrevistaram a realização de algumas tarefas, como a escrita do próprio nome, a escrita de palavras

familiares (mamãe, papai, menino) e não-familiares (sapo, formiga, mapa), e a escrita de uma oração que contivesse alguma das palavras anteriores (“o menino brinca com o sapo”). Em momento posterior à resolução das tarefas, as pesquisadores indagaram as crianças a respeito das respostas dadas, pedindo-lhes que lessem aquilo que escreveram, apontando a leitura com o dedo indicador.

Com base nesse tipo de situação experimental, as autoras chegaram à proposição do percurso evolutivo pelo qual passam as crianças na compreensão do sistema de escrita alfabética, constituído, segundo elas, de 5 níveis distintos de conceitualização que, em Ferreiro (1985), são distribuídos em 3 níveis mais abrangentes. A distinção em 3 níveis toma como critério fundamental a interpretação que dá a criança a respeito do domínio de referência da escrita, ou seja, como ela responde à pergunta “se a escrita representa alguma coisa, o que ela representa?”.

A característica predominante do **primeiro nível** é a busca pela distinção entre formas icônicas (desenhos) e não-icônicas (letras), sem a diferenciação, no entanto, das cadeias de letras em termos de sua legibilidade (não importa se há letras repetidas e sequências de letras que não formam palavras reais da língua, por exemplo). Embora se faça a distinção entre desenho e escrita, a criança não necessariamente atribui um valor notacional ao não-icônico, isto é, não associa as cadeias de letras à intenção de representação de um conceito ou objeto externo. A escrita não é, pois, realização da linguagem.

No **segundo nível**, observam-se características que estão associadas ao que se convencionou chamar como períodos pré-silábico indiferenciado e pré-silábico diferenciado, especialmente a construção de critérios para a legibilidade de uma cadeia de letras e para a diferenciação entre formas escritas, o que ocorre nos eixos quanti e qualitativos do sistema. O eixo quantitativo se refere à variedade e ao número de caracteres que são utilizados, e o qualitativo refere-se ao tipo e à ordem de caracteres utilizados. Assim, segundo o ponto de vista da criança, uma sequência gráfica, para ser legível e legitimamente considerada como escrita, deve possuir um número mínimo de marcas (duas letras, por exemplo), variar quanto ao tipo dessas marcas (não repetir letras), e variar quanto à ordem que as marcas estão postas na sequência. Deve-se observar que não necessariamente a criança utiliza letras para registrar as sequências escritas – é o caso do período pré-silábico indiferenciado, no qual é verificável o uso de linhas contínuas, traços circulares, retilíneos, curvilíneos, quadrados, no que seria o lugar das letras do alfabeto.

Neste momento, como começa a se estabelecer algum tipo de referência entre a linguagem e a escrita – isto é, a criança passa a entender a escrita como representação de significado – é comum observar o fenômeno do *realismo nominal* (REGO, 1986), caracterizado pela associação não-arbitrária entre significante e significado, aferível quando as crianças sistematicamente escrevem palavras mais extensas para conceitos que remetem a entes fisicamente maiores. Um exemplo clássico é a escrita da palavra ‘boi’ em contraponto à da palavra ‘formiga’. Uma criança que opera segundo a lógica do realismo nominal irá escrever ‘boi’ com maior número de caracteres ou com maior traçado, enquanto a palavra ‘formiga’ será grafada com menos caracteres ou menores marcas gráficas.

O **terceiro nível**, por fim, compreende as chamadas hipóteses silábica, silábica-alfabética e alfabética, as quais têm como característica primordial a compreensão da escrita como sistema de representação da camada fônica da língua, ou seja, a concepção de que a escrita alfabética tem como eixo de funcionamento as propriedades formais e não-significativas da língua.

Inicialmente, correspondendo ao período silábico, as crianças tomam a sílaba como sendo a unidade fonológica a que as marcas gráficas se referem, de modo que associam uma marca para cada unidade silábica da palavra-alvo. Neste momento, também se observa uma operação diferenciada entre os eixos quantitativo e qualitativo, para além daqueles observados para os períodos pré-silábicos: uma criança pode utilizar quaisquer tipos de grafismo (traços circulares, retilíneos, curvilíneos, quadrados etc.) para demarcar a quantidade de sílabas (‘///’ pode denotar uma palavra trissílaba, por exemplo) ou pode deliberadamente escolher as letras que correspondem a um dos fonemas constituintes de cada sílaba (‘R T’ ou ‘A O’ pode denotar a palavra ‘rato’, por exemplo), atendendo não somente à quantidade de unidades mas também à sua natureza. Em geral, ao operar com letras, as crianças elegem entre a classe das consoantes ou a classe das vogais, mas existem casos em que é verificado um uso alternado entre as classes. É importante notar que no período silábico os sujeitos ainda podem operar com as hipóteses de número mínimo e de variedade de caracteres, razão pela qual determinadas sequências gráficas que eles mesmos produzem passam a causar conflitos de ordem cognitiva: se para uma criança é indispensável a presença de um mínimo de duas letras em uma cadeia gráfica, como lidar com palavras monossílabas, as quais, segundo a hipótese silábica, deveriam ser grafadas com uma letra só? E se para outra criança a exigência é a de

que os caracteres devem variar, como aceitar sequências como ‘A A A’ ou ‘A R R’ para ‘arara’, por exemplo, que obedecem perfeitamente ao princípio silábico¹⁶?

A busca pela resolução desse tipo de impasse é um dos fatores que leva à reformulação das hipóteses infantis, passando pelo período de transição chamado silábico-alfabético até chegar ao período alfabético, momento em que se alcança o refinado reconhecimento do princípio regulador do sistema de escrita alfabética: a ideia de que um grafema representa um fonema. Este é o ponto culminante do percurso evolutivo pelo qual passam as crianças em direção à compreensão do sistema de escrita. Nele estando,

[...] a criança já franqueou a ‘barreira do código’; compreendeu que cada um dos caracteres da escrita corresponde a valores sonoros menores que a sílaba e realiza sistematicamente uma análise sonora dos fonemas das palavras que vai escrever. Isto não quer dizer que todas as dificuldades tenham sido superadas: *a partir desse momento, a criança se defrontará com as dificuldades próprias da ortografia, mas não terá problemas de escrita, no sentido estrito* (FERREIRO; TEBEROSKY, [1984] 1999, p. 219, grifos das autoras).

São as grafias produzidas por crianças situadas neste ponto do percurso evolutivo que interessam à presente pesquisa, mais precisamente as grafias em que são produzidas escritas alfabéticas consolidadas a partir das quais são observados os chamados erros ortográficos.

A concepção adotada neste trabalho sobre o erro ortográfico está diretamente associada à concepção piagetiana de *erro construtivo*, subjacente, também, à pesquisa de Ferreiro e Teberosky [1984] (1999). Não se trata de desconsiderar a existência de erro; em relação ao conhecimento *final*, que é objeto da aquisição, uma escrita como ‘A A A’ para ‘arara’, por exemplo, de fato está errada. Porém, classificar esta representação meramente como não-acerto desconsidera uma série de conhecimentos importantes que a criança está construindo sobre o objeto, sobretudo o de que a escrita alfabética opera com a camada sonora da língua, um avanço conceitual considerável em relação à concepção de escrita que caracterizava o período anterior. Assim, o erro é construtivo na medida em que sinaliza para a existência de um processo subjacente de construção ativa em torno da compreensão de determinado objeto.

Se considerado por essa perspectiva positiva, o erro ortográfico oferece pistas importantes sobre as hipóteses com as quais estão operando as crianças alfabéticas

¹⁶ Para uma discussão recente e mais detalhada sobre a hipótese silábica, ver Ferreiro [2009] (2013).

a respeito dos conhecimentos envolvidos no percurso de desenvolvimento ortográfico (MIRANDA, 2010), os quais podem ter naturezas mais ou menos variáveis a depender do grau de transparência do sistema, como se verá adiante: se a transparência for de alto grau, os erros podem elucidar especialmente questões concernentes ao conhecimento fonológico; se houver mais opacidade, outros aspectos entram em concorrência com o componente fonológico, como a morfologia (SEYMOUR, 1997, 2005).

É em torno dessa potencialidade reveladora do erro ortográfico que se desenha o domínio teórico dos estudos desenvolvidos no âmbito do GEALE, no qual esta pesquisa está inclusa, tendo por base, também, as demais considerações até aqui realizadas. Nas palavras de Miranda (2017, p. 21),

[...] considerando a premissa de que experiências prévias são comumente ativadas para a resolução de novos desafios cognitivos, constante nos estudos de base piagetiana, procuramos vislumbrar, nas escritas heteróclitas produzidas nos períodos iniciais, a forma como o conhecimento já construído sobre sua língua materna, especialmente aquele referente à fonologia da língua, é retomado pela criança.

Por essa perspectiva, a linguagem escrita – e, particularmente, a escrita inicial – é considerada como uma substância tão potente de atualização da língua quanto é considerada a linguagem falada, a ponto de ser possível, pela sua análise, a elucidação de aspectos concernentes especialmente à aquisição e ao funcionamento do sistema fonológico – abordagem também adotada por Veloso (2003, 2006, 2019), que argumenta em favor de que a escrita, não somente a fala, pode revelar importantes informações sobre a natureza das representações fonológicas. O termo *atualização* corresponde à sua acepção linguística formal, que remete à ideia de realização das unidades linguísticas abstratas (fonemas, por exemplo) em uma substância física primária ou secundária (som ou escrita) (CRYSTAL, 1988; LYONS, 1968). Dessa forma, a “retomada do conhecimento fonológico” (MIRANDA, 2017, p. 21) pelas crianças a partir da compreensão do princípio alfabético possibilita a atualização desse conhecimento em uma substância alternativa, a escrita, razão pela qual se pode conceber aquisição da escrita como parte do processo mais amplo de aquisição da linguagem.

Em sendo retomado via substância distinta daquela que é a habitual, é como se o conhecimento linguístico passasse a ser compreendido a partir de um novo ponto de vista pela criança: não mais apenas do ponto de vista da primeira articulação, por via do qual a linguagem é constituída somente de significado, mas também pelo viés

da segunda articulação, por meio do qual se observa que a língua, para veicular significado, possui uma estrutura formal. Tal fato cria condições para que *mudanças representacionais* possam ocorrer (MIRANDA, 2017; KARMILOFF-SMITH, 1994), tanto no que se refere às estruturas em si quanto à consciência que se tem das estruturas. Como observa Frith (1998), o fato de se descobrir que a linguagem opera com unidades mínimas e abstratas tais como os fonemas modifica substancialmente o próprio processamento da linguagem, pois o que antes era processado como unidade indivisível agora é processado de forma analisada, um importante aspecto da aquisição da escrita que é explorado com mais detalhe na subseção seguinte. A ideia de que a escrita pode, por assim dizer, influenciar a estruturação das representações fonológicas também está presente na argumentação de Veloso (2003, 2006, 2019).

Miranda (2017) sintetiza essa concepção que define o domínio teórico do GEALE em um diagrama tal como o que se observa na Figura 2:

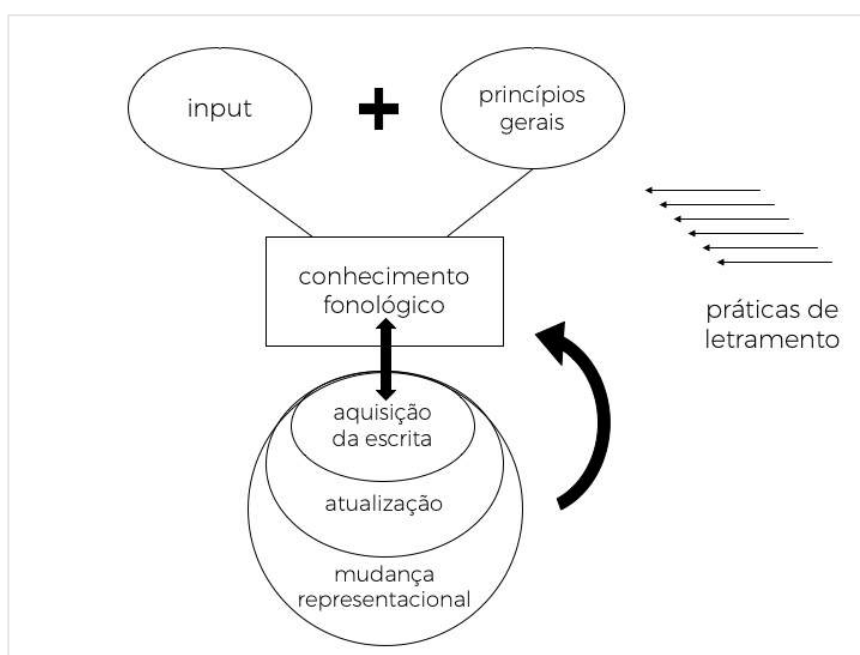


Figura 2 – Diagrama conceitual: o domínio teórico das pesquisas do GEALE.
Fonte: MIRANDA, 2017, p. 22.

Em síntese, o diagrama expressa a ideia de que o sujeito, a partir de princípios gerais constitutivos da Gramática Universal (CHOMSKY, 1988) que interagem com o *input* recebido da sua comunidade linguística, constrói e organiza progressivamente seu conhecimento linguístico internalizado – mais especificamente, neste caso, o conhecimento fonológico, o que inclui toda sorte de informações melódicas e

prosódicas, realizadas via fala, a substância primária (KIPARSKY; MENN, 1977; MACKEN, 1996; SAUSSURE, [1912] (2012); LYONS, 1968). Este conhecimento é retomado no momento em que a criança compreende o domínio de referência estrito da escrita, a saber, a camada fonêmica da língua (LANDSMANN, 1995; GELB, 1952), criando condições para: (i) atualizar a língua em substância alternativa, gráfica, e (ii) disponibilizar esse conhecimento *para a mente*, acarretando na possibilidade de sua manipulação e consequente mudança representacional (KARMILOFF-SMITH, 1994).

2.1.2.1 Aquisição da escrita e consciência fonológica

Conforme observado na subseção anterior, a aquisição da escrita passa, em boa medida, pela consciência cada vez mais refinada, por parte das crianças, a respeito da estrutura fonológica da língua. Nesse sentido, faz-se necessário apreciar com mais detalhe as relações que podem se estabelecer entre a consciência fonológica e a aquisição da escrita.

Em *Explicit Syllable and Phoneme Segmentation in the Young Child* (1974), I. Y. Liberman, Shankweiler, Fischer e Carter lançam as bases de problemáticas que se tornaram clássicas nas investigações envolvendo tais relações. Os autores tratam de dois aspectos em especial: primeiro, da desafiadora tarefa cognitiva imposta à criança que está aprendendo a escrever, qual seja, abstrair da linguagem falada os segmentos adequados à sua representação no sistema de escrita, e, segundo, de definir se essa tarefa torna-se cognitivamente mais desafiadora à medida que a unidade de segmentação do sistema de escrita é tanto menor quanto menos significativa do ponto de vista da estruturação do sistema linguístico.

I. Y. Liberman e seus colaboradores (1974) comparam, nesse sentido, o grau de dificuldade que têm crianças em fase inicial de escolarização na segmentação (i. e., abstração) de fonemas e de sílabas. Os autores têm por hipótese que a segmentação em fonemas é mais custosa e menos natural para as crianças do que a segmentação em sílabas, baseando-se em dois argumentos principais: o primeiro relaciona-se à maneira como foram surgindo e se desenvolvendo os diferentes sistemas de escrita ao longo da história humana (GELB, 1952), e o segundo assenta-se na maneira como a fala é percebida e codificada pelo sistema linguístico humano (A. M. LIBERMAN; COOPER; SHANKWEILER; STUDDERT-KENNEDY, 1967).

I. Y. Liberman et al. (1974) retomam a concepção clássica de que os sistemas de escrita desenvolvidos ao longo da história fundaram-se na intuição humana de que a língua está duplamente estruturada em uma camada significativa e uma camada não-significativa (ver subseção 2.1.2). É central para os autores observar que, “no desenvolvimento histórico da escrita, sistemas que utilizaram unidades significativas vieram primeiro [...]. Escrever com unidades não-significativas é um desenvolvimento mais recente” (LIBERMAN et al., 1974, p. 201-202). Entre os sistemas fonográficos, ainda, há de se considerar a existência mais tardia da escrita alfabética, cuja unidade de representação é o fonema, em relação à silábica.

Por essa razão, I. Y. Liberman et al. (1974, p. 202) supõem que

o desenvolvimento histórico da escrita pode refletir a facilidade (ou dificuldade) com que a segmentação explícita pode ocorrer. [...] podemos concluir que em ordem crescente de dificuldade está a palavra, a sílaba, e [...] o fonema. [...] podemos supor que para a criança deve haver a mesma ordem de dificuldade e, de forma correspondente, a mesma ordem de emergência no desenvolvimento¹⁷.

Somado a esse fato – e, talvez, como a própria causa dele – está o segundo argumento, que diz respeito à percepção humana dos sons linguísticos. Os autores lembram que até mesmo bebês demonstram a capacidade de identificar e discriminar contrastes fônicos em pares como [mar] e [par]¹⁸, e o fazem como os adultos, isto é, de forma categórica. Porém, conforme notam I. Y. Liberman et al. (1974), isso não significa que as crianças consigam *analisar* essa diferença de forma correspondente à estrutura fonêmica que lhe é subjacente, ou seja, demonstrar um conhecimento *explícito* de que os pares são constituídos, cada um, por três segmentos e que a informação concernente à distinção entre os pares concentra-se no primeiro segmento.

Isso porque, de acordo com o que os estudos sobre percepção da fala apresentam (A. M. LIBERMAN et al., 1967), a percepção do sinal acústico não corresponde diretamente à segmentação fonêmica. O ouvido humano não percebe a palavra [mar], por exemplo, em relação biunívoca com a estrutura fonêmica da língua. Pelo contrário, “os segmentos fonêmicos são codificados no nível acústico em unidades maiores aproximadamente de tamanho silábico” (I. Y. LIBERMAN et al,

¹⁷ Essa intuição parece estar de acordo, por exemplo, com os estudos de Ferreiro e Teberosky [1984] (1999), apresentados na subseção 2.1.2 deste trabalho, cuja teoria prevê 3 níveis globais de desenvolvimento da escrita alfabética que parecem acompanhar a ordem crescente do desafio cognitivo que é segmentar a linguagem em unidades não-significativas.

¹⁸ Pares análogos aos exemplos originais do texto *bad* e *bat* (*mau* e *morcego*, em inglês).

1974, p. 203). A informação acústica a respeito das consoantes adjacentes à vogal em [mar] fica contida, por assim dizer, na própria vogal. Assim, três segmentos fonológicos equivalem a um segmento acústico, demandando regras linguísticas que realizem a conversão do sinal acústico à informação fonológica, e vice-versa (A. M. LIBERMAN et al., 1967).

É por esse motivo que a segmentação explícita da linguagem em sílabas deveria ser mais simples que a segmentação em fonemas. Se a informação acústica das consoantes está “embrulhada” na vogal, e se, no sistema fonológico, o constituinte obrigatório para a formação de sílabas – o Núcleo – é projetado por uma vogal, decorre que a identificação do *número* de sílabas seria mais acessível, ainda que as pistas para identificar precisamente as *fronteiras* silábicas estejam encobertas (I. Y. LIBERMAN et al, 1974).

Como forma de testar a hipótese de que a análise explícita em sílabas é mais acessível do que a análise em fonemas, I. Y. Liberman et al. (1974) desenvolveram um experimento que procurou verificar o “quão bem crianças de berçário, jardim de infância e primeiro ano conseguem identificar o número de segmentos fonêmicos em sentenças faladas e o quanto isso se compara com a sua habilidade de lidar similarmente com sílabas” (I. Y. LIBERMAN et al, 1974, p. 204).

Os sujeitos eram 135 crianças de classe média de uma escola pública de educação infantil na região suburbana de Manchester, na Inglaterra, e de uma escola pública de ensino fundamental da cidade vizinha de Andover. As crianças foram distribuídas em dois grupos experimentais: um grupo realizava o teste de segmentação fonêmica e o outro realizava o teste de segmentação silábica. Para ambos os grupos, o procedimento era o mesmo: as crianças deveriam repetir a palavra ou o som falado pelo pesquisador e, em seguida, indicar o número de segmentos (fonemas ou sílabas, a depender do grupo) contidos na palavra ou no som, apertando em um botão tantas vezes quantos segmentos elas julgassem haver no estímulo.

Os resultados demonstraram que em todos os níveis escolares as crianças obtiveram mais sucesso no teste segmentação silábica em relação ao teste de segmentação fonêmica. No berçário, por exemplo, nenhuma das crianças foi bem-sucedida na segmentação fonêmica. Nos níveis escolares seguintes, mesmo as crianças que obtiveram sucesso nesse teste, levaram mais tempo e maior número de

tentativas até atingir o número mínimo de acertos exigidos (6 acertos consecutivos) em relação ao outro teste.

Apesar de os resultados corroborarem a hipótese da pesquisa, I. Y. Liberman et al. (1974, p. 210) assinalam que “seria um equívoco supor que a estrutura silábica é completamente transparente às crianças”. Mesmo no teste de segmentação silábica observa-se um aumento gradual na taxa de acerto conforme o ano escolar – de 46% no berçário a 90% no primeiro ano – o que indica que a capacidade de analisar, segmentar e abstrair a linguagem, mesmo que em suas camadas mais elementares, requer práticas de instrução intencional (I. Y. LIBERMAN et al, 1974).

Uma interessante discussão decorrente do trabalho de I. Y. Liberman et al. (1974) é explorada por Treiman (1992) e concerne ao papel das unidades intrassilábicas no desenvolvimento da consciência fonológica e, conseqüentemente, na aquisição da leitura e da escrita. Para a autora, estudos como os de Isabelle Liberman et al. (1974) não consideram a estrutura silábica como uma variável relevante, pois, à época de sua realização, os modelos fonológicos consideravam a sílaba como uma sequência linear e não-hierárquica de fonemas (ver subseção 2.3).

Treiman (1992) argumenta, apresentando resultados de uma série de pesquisas por ela realizadas, que entre a consciência silábica e a consciência fonêmica existem níveis intermediários de consciência fonológica correspondentes às unidades intrassilábicas¹⁹. Portanto, antes de as crianças analisarem as palavras em fonemas, elas analisariam as palavras em *onset* e rima, de modo a estarem cientes de que, por exemplo, uma sílaba como /blo/ começa com /b/ e termina com /o/. No entanto, nesse nível, as crianças “não são capazes de analisar /b/ como o fonema /b/ seguido do fonema /l/” (TREIMAN, 1992, p. 71). A autora assinala que somente nas sílabas CV a análise em *onset* e rima equivale biunivocamente à análise em fonemas, uma vez que a consoante projeta o *onset* e a vogal, a rima. Dessa forma, é necessário averiguar se as crianças, quando aparentemente demonstram capacidade de segmentar as palavras em fonemas, não estão, na verdade, apenas lidando com palavras de sílaba CV.

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 formais de alfabetização e que a estrutura silábica afeta a

¹⁹ Considerando-se o modelo não-linear de Selkirk (1982), explorado em detalhe na subseção 2.3.

performance 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 constituintes são ramificados. Assim, indo ao encontro da conclusão mais abrangente de I. Y. Liberman et al. (1974), Treiman (1992, p. 77, interpolações minhas) acorda 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 [sílabas, *onset*, rima]”.

2.1.3 Ortografia: funcionamento e aprendizagem

Esta subseção encerra as principais discussões acerca da aquisição e do processamento da linguagem, tratando de noções específicas concernentes ao funcionamento e à aprendizagem do sistema ortográfico.

Antes de compreender a natureza dos conhecimentos mobilizados pela criança no curso do desenvolvimento da ortografia, momento que se inicia a partir da compreensão, pelo aprendiz, do princípio alfabético, é necessário ter em vista como se organiza e funciona o sistema ortográfico, o que implica em observar quais são as relações que se estabelecem entre fonemas e grafemas, quais as regras que regem essas relações e como as mesmas relações definem a *profundidade* do sistema. No caso deste estudo, interessa particularmente observar como se organiza o sistema ortográfico do Português Brasileiro, que, do ponto de vista das consoantes, está sintetizado na Figura 3, conforme segue:

Quadro de consoantes do Português						
ponto modo	bilabial	labiodental	alveolar	palato- alveolar	palatal	velar
plosivas	/p/ <p>		/t/ <t>			/k/ <c> <qu>
	/b/ 		/d/ <d>			/g/ <g> <gu>
fricativas		/f/ <f>	/s/ <s> <z> <c/> <x> <sc/sç> <ss> <xs> <xc>	/ʃ/ <x> <ch>		/r/ <r> <rr>
		/v/ <v>	/z/ <z> <s> <x>	/ʒ/ <j> <g>		
africadas				[tʃ] 't' [dʒ] 'd'		
nasais	/m/ <m>		/n/ <n>		/ɲ/ <nh>	[ŋ] 'n'
laterais			/l/ <l>		/ʎ/ <lh> 	
róticas			/r/ <r>			
glides	[w]				[j] 'i'	[w] 'u' 'i'

34 grafemas para 19 fonemas
5 alofones

Figura 3 – Quadro das relações entre fonemas consonantais e grafemas que caracterizam o sistema ortográfico do Português Brasileiro.
Fonte: MIRANDA, 2020, p. 5.

O quadro mostra que, no que se refere às consoantes do PB, existem alguns casos em que as relações entre fonemas e grafemas são de um para um, isto é, um fonema será sempre representado, na escrita, pelo mesmo grafema. Em proporção um pouco maior, existem casos em que as relações entre fonemas são de um para muitos, ou seja, um mesmo fonema pode ser representado por mais de um grafema em diferentes contextos. Lemle (1987), ao sistematizar as relações fonografêmicas, chama os primeiros casos de relações *biunívocas* e os segundos de relações *múltiplas*.

Diversos autores, inspirados em Lemle (1987), apontaram para o fato de que as relações múltiplas são regidas por regras que podem ser definidas contextual ou arbitrariamente (cf. CAGLIARI, 1987; CARRAHER, 2001; MORAIS, 2002; MIRANDA,

2013, 2020). Assim, por exemplo, as duas possibilidades existentes para a grafia do fonema /r/ são regidas por regras contextuais, uma vez que o grafema <r> é utilizado em início de palavra (como em ‘rato’) e após consoantes em final de sílaba (como em ‘Israel’ e ‘honra’), sendo que o <rr> é somente utilizado em posição intervocálica. Regras arbitrárias regem relações como a do fonema /ʒ/ com seus correlatos gráficos <j> e <g>, sobretudo diante das vogais /i/ e /e/, como em ‘jiboia’, ‘girafa’, ‘jeito’ e ‘geladeira’. Nesses casos, regulados pela etimologia, não é possível extrair um princípio gerador, restando ao aprendiz o apelo à estratégia mnemônica.

No tocante à representação gráfica do sistema vocálico²⁰, a única relação biunívoca que se estabelece, segundo Lemle (1987), é entre o fonema /a/ e o grafema <a>²¹; para os demais casos, relações múltiplas são observadas.

Considerando as relações fonografêmicas que se estabelecem na ortografia do PB, é possível classificá-la segundo o que Seymour, Aro e Erskine (2003) chamam de critério da *profundidade ortográfica*, 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 fonografêmicas, ou seja, mais relações biunívocas que múltiplas. Assim, considerando-se a Figura 3 e em conformidade com Veloso (2005), que analisa a ortografia do PE em relação ao critério de profundidade ortográfica, é possível entender o sistema do PB como *relativamente transparente*, uma vez que apresenta relações biunívocas e múltiplas de forma mais ou menos equilibrada.

Tendo em vista essa caracterização do sistema ortográfico, juntamente à concepção teórica mais abrangente apresentada no final da subseção 2.1.2, é possível compreender propostas que visam definir a natureza dos conhecimentos envolvidos no percurso de desenvolvimento ortográfico. Uma delas decorre de inúmeras pesquisas realizadas no âmbito do GEALE, por meio das quais se chegou

²⁰ É necessário considerar que o sistema vocálico do PB organiza-se em relação direta com a tonicidade: na posição tônica, 7 são as vogais contrastivas - /a e ε i o ɔ u/; na posição pretônica, o número de segmentos funcionais cai para 5, com a perda da distintividade entre as médias-baixas e as médias-altas - /a e i o u/; na posição pós-tônica não-final, o sistema é reduzido a 4 vogais, com a perda de contraste entre a média-alta dorsal /o/ e a vogal alta /u/ - /a e i u/; e, por fim, na posição pós-tônica final, a última vogal média restante perde o contraste e 3 são as vogais funcionais - /a i u/ (cf. CÂMARA JR., [1970] 1995).

²¹ Considerando-se a existência apenas de vogais orais no sistema do PB.

a um sistema de categorização de erros ortográficos, cuja configuração atual pode ser visualizada na Figura 4²² 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 Português Brasileiro, considerando seu grau de profundidade relativamente transparente:

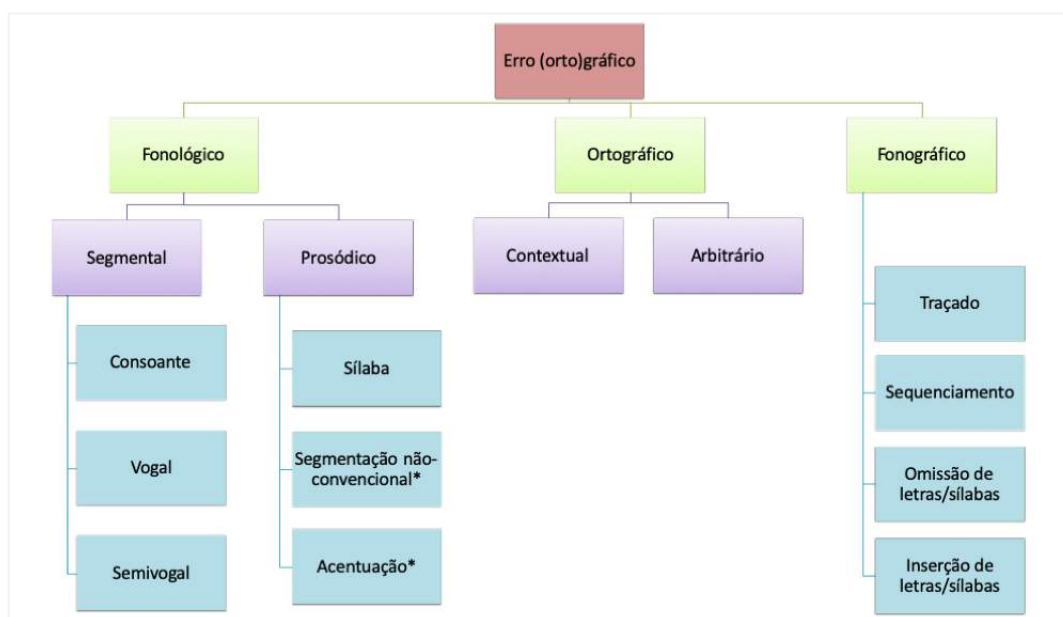


Figura 4 – Organograma das categorias de erros ortográficos do GEALE.
Fonte: acervo do GEALE.

O organograma evidencia que os erros (orto)gráficos produzidos por aqueles que estão em processo de aquisição da escrita advêm de conhecimentos referentes a três fontes distintas: fonologia, ortografia e fonografia (MIRANDA, 2020). A primeira categoria diz respeito à complexidade representacional 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 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 no sistema fonológico. A terceira, 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

²² Uma versão mais detalhada das categorias utilizadas neste estudo pode ser visualizada na Figura 22 (capítulo 3).

informação²³. Cada uma das categorias Cada uma delas possui desdobramentos que possibilitam uma descrição mais precisa do dado analisado. Por exemplo, em uma grafia como ‘pri car’ para ‘brincar’, observa-se a interveniência de diferentes aspectos relacionados à grafia tanto do *onset* quanto da rima, todos vinculados à fonologia em uma perspectiva mais abrangente: a dessonorização de /b/, o que envolve o primeiro *segmento* do *onset* complexo; a omissão da consoante nasal em coda, o que envolve a estrutura da *sílaba* propriamente dita, e a hipersegmentação da palavra (a inserção do espaço em branco em lugar não previsto pela norma) exatamente na fronteira silábica onde há apagamento da coda nasal, aspecto ligado à *segmentação não-convencional*.

Outra proposta que procura evidenciar a natureza dos aspectos envolvidos no processamento e no percurso de compreensão do sistema ortográfico pelas crianças que chegam à compreensão do princípio alfabético é a de Seymour (1997), consistindo não propriamente em uma classificação de erros ortográficos mas em uma definição de módulos cognitivos que estariam envolvidos no processamento da ortografia. A proposta do autor é feita levando-se em consideração especificidades da ortografia do inglês que, apesar de constituir um sistema fonográfico, apresenta relações bastante opacas entre fonemas e grafemas, o que leva os pesquisadores a considerarem-no mais precisamente como um sistema morfofonêmico (SEYMOUR, 1997; C. CHOMSKY, 1971; BRYANT; DEACON; NUNES, 2006). O modelo é, nas palavras do autor, de “duplo-fundamento”, pois leva em conta a relação da ortografia inglesa com componentes morfológicos e lexicais, além do alfabético. Na Figura 5, a seguir, é possível visualizar a proposta de Seymour (1997):

²³ A categoria de erros *fonográficos* não é considerada neste estudo em virtude de ainda ser uma proposta embrionária que necessita de testes para melhor compreensão de sua natureza e de sua validade para erros do tipo que envolvem estruturas como as sílabas complexas. No capítulo 3, que descreve a metodologia da pesquisa, as categorias consideradas são apresentadas e melhor explicadas.

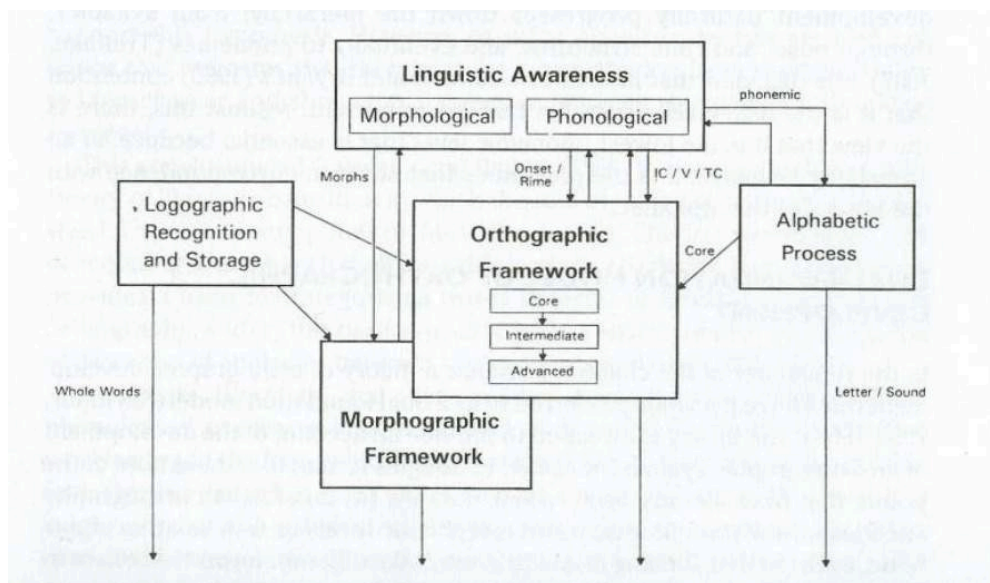


Figura 5 – Fluxograma do modelo de desenvolvimento ortográfico de Seymour (1997).
 Fonte: SEYMOUR, 1997, p. 324.

Da figura, depreende-se que este modelo de desenvolvimento da ortografia é constituído de 5 componentes distintos que interagem entre si: dois compõem o fundamento, funcionando como pré-requisito – o Processo Alfabético e o Processo Logográfico; um torna o sistema operacional e funcional – o módulo da Consciência Linguística –, e outros dois possuem função representacional, onde o conhecimento abstrato do sistema está armazenado, constituindo-se como as estruturas centrais – os módulos da Estrutura Ortográfica e da Estrutura Morfográfica.

O Processo Alfabético está fundado na consciência sobre a realidade fônica da língua e de sua atualização em substância gráfica, o que possibilita o estabelecimento das relações iniciais entre fonemas e grafemas. Dentro do modelo de Seymour (1997), a principal contribuição do processo alfabético é a formação do núcleo estrutural do sistema ortográfico (*core*).

O Processo Logográfico se refere ao reconhecimento e armazenamento direto de palavras, que são tratadas como unidades inteiras. Neste módulo,

[...] O ponto principal é que representações de *palavras* são codificadas, normalmente consistindo em um informação parcial ou completa da identidade das letras. Esse processo é considerado vital para o desenvolvimento ortográfico porque é o caminho pelo qual os exemplares, por meio dos quais a informação ortográfica é extraída, são internalizados (SEYMOUR, 1997, p. 325).

Não há relações intrínsecas entre o armazenamento logográfico e o conhecimento fonológico, porém ele pode fornecer uma sensibilidade a restrições grafotáticas. Assim, as crianças podem discriminar bem combinações de grafemas

licenciadas no sistema, ainda que relacionem a formas fonológicas incorretas (por exemplo, saber reconhecer que dígrafos com a letra ‘h’ – <ch>, <lh>, <nh> – fazem parte do sistema ortográfico, mas, no momento de escrever, atribui-los aos fonemas incorretos, como em ‘ovenha’ para ‘ovelha’). Em geral, o processo logográfico é mobilizado especialmente na leitura ou escrita de uma palavra familiar, enquanto o alfabético é mobilizado na leitura ou na escrita de palavras inventadas ou não-familiares.

Ao módulo de Consciência Linguística Seymour (1997) dedica mais atenção, detalhando uma discussão que, em vista do objetivo da presente pesquisa, é importante evocar. Retomando as proposições de I. Y. Liberman et al. (1974) e de Treiman (1992), exploradas na subseção 2.1.2.1 deste trabalho, o autor salienta que é necessário realizar uma distinção entre os tipos de consciência fonológica, aspecto ao qual teóricos como Gombert (1992) se empenharam em fazer (ver subseção 2.1). Um tipo emerge de forma natural, e particularmente no tocante às sílabas, pode ser evidenciado por uma sensibilidade à realidade de constituintes silábicos maiores, como rima e *onset* (ver subseção 2.3), ainda que de forma não analisada, isto é, não decomposta nas unidades menores que os compõem. A habilidade de manipular rimas e aliteraões geralmente está associada a esse tipo de sensibilidade.

Seymour (1997) observa que, por essa perspectiva, o curso natural do desenvolvimento da consciência fonológica partiria das unidades maiores até chegar às menores. No entanto, esse processo natural é atravessado por uma demanda externa, a aquisição da escrita alfabética, fazendo com que a capacidade de segmentação da camada fônica em suas menores unidades seja alcançada antes mesmo da percepção de que há um nível anterior de organização ao qual essas unidades devem respeitar. Isso significa que os fonemas, no momento inicial de aquisição da escrita alfabética, seriam percebidos como unidades isoladas, que não necessariamente se relacionam em termos de restrições fonotáticas (e também, posteriormente, em termos de restrições grafotáticas que regulam a representação gráfica dos fonemas). A Figura 6 auxilia a visualizar as unidades de representação silábica com as quais a criança lida no curso de desenvolvimento da consciência fonológica:

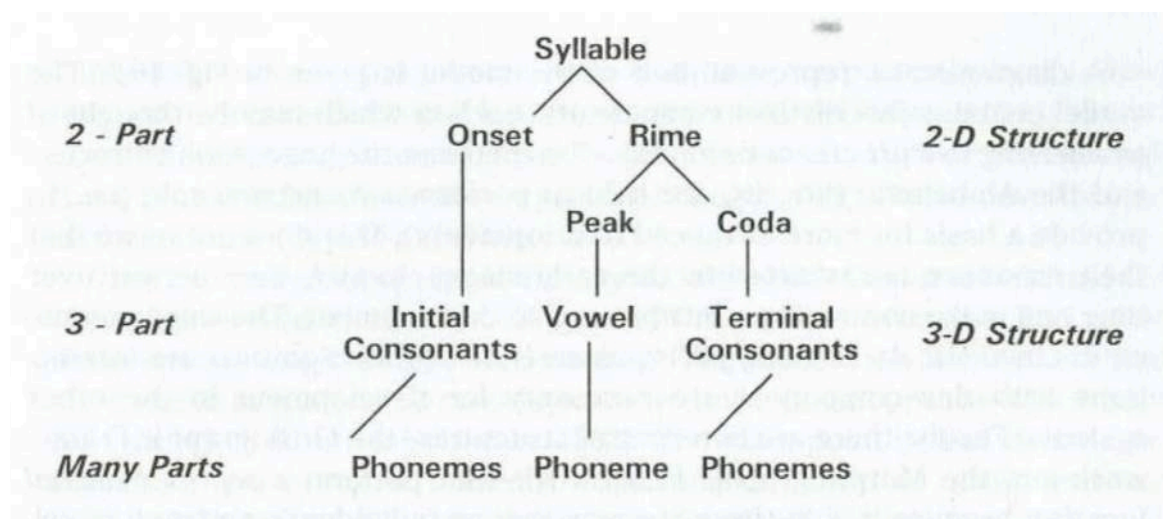


Figura 6 – Representação da hierarquia silábica com base em Seymour (1997).
 Fonte: SEYMOUR, 1997, p. 323.

Segundo as proposições do autor, portanto, a consciência fonológica teria a tendência natural de se desenvolver partindo da sílaba como unidade indivisível e posteriormente tornando o indivíduo sensível ao que seria a *estrutura 2D*, conforme apresenta a Figura 6. Esse processo de *cima para baixo* é atravessado pela aquisição de um SEA, fazendo com que a sensibilidade à *estruturação 3D* da sílaba seja precedida pela percepção da segmentação fonêmica da língua. Seymour (1997, p. 327, tradução nossa) salienta que

Isso não significa que estruturas de ordem superior não sejam importantes no desenvolvimento da alfabetização, apenas que a sua importância emerge posteriormente ao invés de inicialmente. A razão para isso é que estruturas de ordem superior não se tornam relevantes até o estágio em que a estrutura ortográfica passa a se formar.

Dessa maneira, para o autor, no decorrer do processo de desenvolvimento ortográfico, unidades silábicas de ordem superior à camada fonêmica, particularmente a camada da estrutura 3D, passam a ser tornar objeto de atenção, devido à introdução do aprendiz às complexidades ortográficas e fonológicas que caracterizam o sistema em aquisição (relações múltiplas, moldes silábicos da língua, restrições fonotáticas etc.). Assim, as hipóteses sobre as relações fonografêmicas são reorganizadas constantemente à luz da hierarquia interna da sílaba, com estipulações das possíveis combinações entre grafemas que são licenciadas para a ocupação das devidas posições silábicas.

Esse processo é fundamental para a construção da Estrutura Ortográfica, o eixo central do modelo de Seymour (1997), a qual se desenvolve progressivamente, primeiro com estruturas mais simples até chegar em formas mais complexas, com as

relações múltiplas entre fonemas e grafemas definidas. Para cada relação ortográfica há um conjunto de palavras que oferece a referência correta da relação entre fonema e grafema, o que equivaleria a uma espécie de *template* lexical gráfico. Esse léxico é diferente em termos de sua extensão daquele armazenado no Processo Logográfico, embora extraia seu conteúdo deste último. A diferença entre as entradas logográficas e ortográficas está em que aquelas são repositório e estas são lugar de construção de representações que visam a formação de um sistema, podendo ser redescritas sempre que houver a entrada de uma nova informação.

O modelo tem como pressuposto a ideia de que o desenvolvimento da estrutura básica da ortografia (*orthographic framework*) se dá pela passagem progressiva por níveis representacionais (*core* → *intermediate* → *advanced*). A afirmação do primeiro estágio, que dá início ao desenvolvimento da Estrutura Ortográfica, ocorre a partir do estabelecimento do Processo Alfabético e das principais relações entre fonemas e grafemas, especialmente aquelas que não envolvem complexidades ortográficas, ou seja, as relações biunívocas. Com isso, passa a ocorrer uma reorganização constante dessas relações à luz da hierarquia interna da sílaba, com estipulações das possíveis combinações entre letras que são licenciadas para a ocupação das posições silábicas. O avanço dos estágios se dá, no inglês, depois de já estabelecido o estágio nuclear (*core*), com o desenvolvimento da consciência morfológica e com a formação de um módulo morfográfico, acionado em leitura e escrita de palavras e combinações de letras irregulares.

Apesar das diferenças consideráveis existentes entre os sistemas ortográficos do Português e do Inglês, o modelo de Seymour (1997) apresenta pistas importantes que auxiliam a elucidar com maior profundidade a natureza dos conhecimentos mobilizados pelo aprendiz de uma ortografia que, em seu fundamento, é de base alfabética. Assim, parece que a proposta do autor pode dialogar com a proposta de categorização de erros ortográficos do GEALE (apresentada na Figura 5), ajudando a definir, com maior precisão, o que está em jogo no processamento da informação (orto)gráfica. Esse aspecto é recuperado na discussão dos resultados desta pesquisa.

2.2 Teoria da Sílaba

A sílaba é um dos conceitos teóricos que provavelmente goza de maior estabilidade e consenso entre os estudiosos de Fonologia. Independente das

diferenças de tratamento que cada teoria dê à configuração estrutural da sílaba, há acordo em considerá-la uma unidade da gramática fonológica, a menor da camada prosódica (NESPOR; VOGEL, [1986] 1994).

Conforme observa Freitas (2017), a sílaba, enquanto tema, já estava presente em gramáticas antigas, provavelmente em virtude de ser a unidade de identificação mais intuitiva por parte dos falantes das línguas em geral, e, por isso, também comumente tomada como referência para a segmentação da fala em partes menores, quando existe tal demanda (cf. A. M. LIBERMAN et al., 1968).

É, no entanto, na segunda metade do século XX que a sílaba é incluída como peça-chave na teoria fonológica, especialmente com o advento dos modelos não-lineares. Mesmo no célebre *The Sound Pattern of English* (SPE), de Chomsky e Halle (1968), estudo que forneceu as bases para o desenvolvimento da Fonologia Gerativa, a sílaba ainda não possuía um estatuto próprio, enquanto unidade gramatical (CLEMENTS; KEYSER, 1983).

Collischonn [1996] (2014), ao revisar diferentes teorias sobre a sílaba desenvolvidas a partir da fonologia não-linear, destaca dois tipos de caracterizações propostas pelos estudiosos gerativistas como possibilidades de representação da estrutura silábica:

- a) uma autossegmental, pela qual os segmentos ou unidades CV estão diretamente ligados ao nó silábico, sem hierarquia entre si, concepção defendida por Kahn (1976) e Clements e Keyser (1983), cuja formalização pode ser observada nas Figuras 7 e 8:

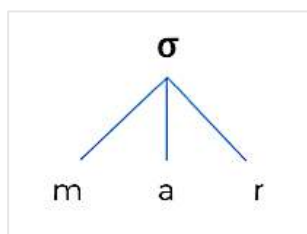


Figura 7 – Representação da sílaba /mar/ conforme Kahn (1976).
Fonte: elaboração própria.

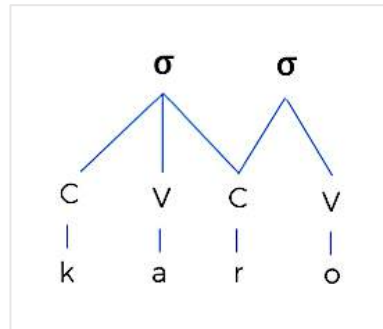


Figura 8 – Representação silábica da palavra /kaRo/ conforme Clements e Keyser (1983).
Fonte: elaboração própria.

- b) uma arbórea (ou métrica), expressa pelo modelo de Selkirk (1982) que, aproveitando noções de pesquisadores anteriores (cf. KAHN, 1976; PIKE; PIKE, 1947; FUDGE, 1969), propõe uma hierarquia de constituintes sub-silábicos que são projetados pela camada melódica, procurando resolver melhor a inter-relação entre picos e margens da sílaba e entre os próprios segmentos. A Figura 9 ilustra essa proposição:

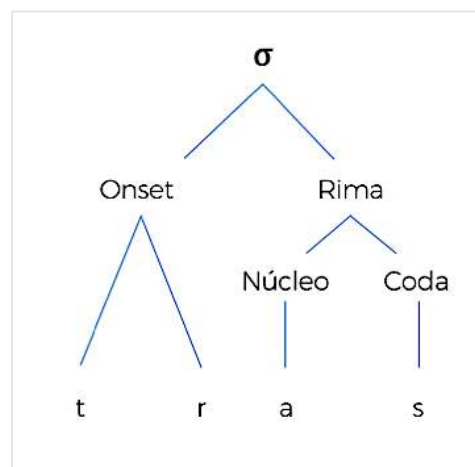


Figura 9 – Representação da estrutura hierárquica interna da sílaba /tras/, com base em Selkirk (1982).
Fonte: elaboração própria.

No presente trabalho, é adotada a perspectiva (b), por meio da qual se entende que a sílaba (σ) é constituída de *Onset* (O) e *Rima* (R) que, por sua vez, é projetada por um *Núcleo* (N) e, opcionalmente, ramifica-se em *Coda* (Co).

Uma noção fundamental que guiou tanto Selkirk (1982) quanto seus colegas de área na proposição de uma teoria da sílaba é a ideia de que a organização interna dessa unidade se dá em torno de um pico de soância, que majoritariamente é realizado por uma vogal – daí a obrigatoriedade da projeção do *Núcleo*. Às margens

deste pico de soância ficam situados os constituintes não-obrigatórios *Onset* e *Coda*, os quais devem apresentar uma sonoridade respectivamente crescente e decrescente em direção ao Núcleo. Assim se estabelece uma das mais importantes *condições universais de silabação*, concernente a uma *sequência de sonoridade* que as línguas devem respeitar na organização silábica que elegem realizar. Esse princípio demanda a existência de uma *escala de soância*, a qual, enquanto construto teórico, tem recebido diversas versões a depender do autor (BISOL, 1999; COLLISCHONN, [1996] 2014; COLLISCHONN; WETZELS, 2016). Nas Figuras 10 e 11, adiante, seguem dois exemplos diferentes de escala de soância, propostas por Clements e Hume (1995) e Selkirk (1984), respectivamente:

obstruintes	<	nasais	<	líquidas	<	vogais
0		1		2		3

Figura 10 – Escala de Soância conforme Clements e Hume (1995).
Fonte: elaboração própria.

som	sonoridade
p, t, k	0,5
b, d, g	1
f, θ	2
v, z, ʁ, ð	3
s	4
m, n	5
l	6
r	7
i, u	8
e, o	9
a	10

Figura 11 – Escala de Soância conforme Selkirk (1984).
Fonte: elaboração própria.

Na escala proposta por Clements e Hume (1995), há maior generalização, e cada grande classe natural tem um índice de soância. A escala de Selkirk (1984), por sua vez, é mais específica, esmiuçando as grandes classes naturais em grupos menores de segmentos e trazendo, assim, mais gradiência para a escala. Independentemente das diferenças, ambas podem sustentar os princípios de

silabação. Com elas, por exemplo, sabe-se que uma sequência como /pl/ pode ocupar o *Onset* e não a Coda, pela sua sonoridade crescente; tampouco essa sequência pode ocupar o Núcleo, uma vez que não tem sonoridade suficiente para projetar esse constituinte. A sequência de sonoridade trata-se de um princípio universal que as línguas devem obedecer e que é fundamental para a definição da sua fonotática, isto é, para o estabelecimento das sequências de sons licenciadas em cada língua.

Além disso, as línguas devem definir o número de segmentos licenciados em uma sílaba e em seus constituintes mínimos, ou seja, estabelecer o *molde silábico*. Os moldes silábicos do Português Brasileiro (PB)²⁴ podem ser visualizados na Figura 12:

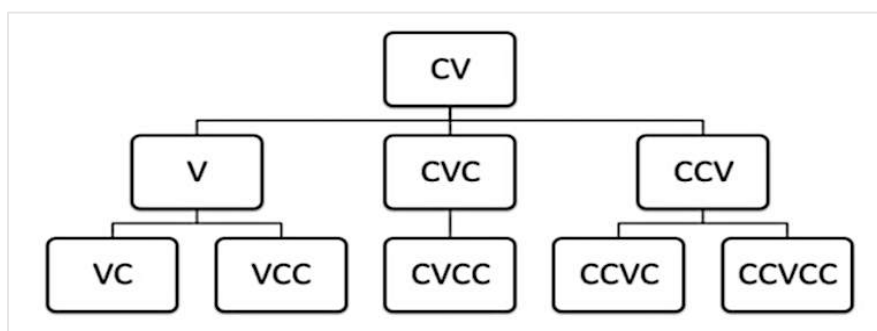


Figura 12 – Organograma dos moldes silábicos licenciados para o Português Brasileiro.
Fonte: MIRANDA; MATZENAUER, 2010, p. 369.

Pela figura, observa-se que, da sílaba canônica universal, derivam os demais *templates* silábicos, os quais, com exceção de V, possuem *Onset* e/ou Rima ramificados (ou complexos²⁵). Com relação a estes últimos, foco deste trabalho, poucos são os segmentos licenciados para ocupá-los conforme a fonotática definida para o sistema do Português – e muitas são as divergências teóricas envolvendo a definição de tal licenciamento, sobretudo na posição de Coda.

Pela proposta de Bisol (1999), que segue a proposição de Câmara Jr. [1970] (1995) especialmente no que se refere à nasalidade fonológica, os segmentos consonantais que podem ocupar a Coda medial são /S/, /N/, /r/ e /l/, os quais são realizados, majoritariamente e respectivamente, pelas contrapartes fonéticas [fricativa coronal], [nasal], [rótica] e [glide lábio-dorsal]. Já pela proposta de Costa e Freitas

²⁴ Neste estudo, são considerados para análise os moldes VC, CVC, CCV e CCVC.

²⁵ Crystal [1985] (1988, p. 55), no *Dicionário de Linguística e Fonética*, define da seguinte forma o termo *complexidade*: “O sentido geral [...] é encontrado na Linguística para indicar tanto a estruturação interna formal de unidades linguísticas quanto a dificuldade em usá-las e aprendê-las”.

(2001), que analisam o Português Europeu (PE), não há uma consoante nasal em Coda, mas sim vogais nasais. Para Mateus e Andrade (2000), cuja análise também se volta ao PE, a nasalidade pode ser formalizada por meio de um autossegmento que não ocupa a Coda. 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. Em se assumindo o modelo de Bisol (1999), é possível sustentar a existência de padrões como CVCC, VCC e CCVCC (ver Figura 12), como nos itens '**con**strução', '**ins**petor' e '**trans**formação'. Nesses casos, consoantes nasais projetariam a Coda, e a fricativa /S/ estaria ligada diretamente à Rima.²⁶

O estabelecimento dos segmentos licenciados para o *onset* no PB trata-se de um ponto mais pacífico na literatura fonológica. Na posição de *onset* simples, praticamente todos os fonemas são licenciados, com algumas restrições no que se refere à posição na palavra, como é o caso das soantes palatais /ɲ/ e /j/, que ocupam apenas a posição medial. No que se refere ao *onset* complexo, as possibilidades são reduzidas. Na primeira posição, são licenciadas algumas oclusivas e fricativas, cuja presença pode ser restringida a depender da líquida, lateral ou não-lateral, que ocupará a segunda posição do *Onset*.²⁷

É possível sintetizar o exposto sobre a fonotática do Português Brasileiro conforme cada constituinte silábico com a Figura 13:

²⁶ É importante esclarecer que, neste trabalho, o foco das análises, no caso da rima ramificada, reside sobre os segmentos que ocupam a *posição medial* na palavra. Glides em ditongos decrescentes também não são analisados. Contemplar esses aspectos – coda final e glides – excederia a capacidade de alcance da pesquisa, devido à complexidade e às controvérsias específicas também existentes em torno de seus estatutos fonológicos, muitas das quais envolvem interface direta com a morfologia. Para uma breve revisão a esse respeito, conferir Collischonn [1996] (2014) e Collischonn e Wetzels (2016). Estudos feitos no GEALE que relacionam fonologia e ortografia nesses casos também podem ser conferidos em Adamoli (2006, 2012), Miranda (2018), Ávila (2019) e Pachalski e Miranda (2019b). Além disso, opta-se, neste trabalho, pelo uso deliberado do termo *rima ramificada*, devido às controvérsias existentes em torno dos segmentos que teriam estatuto de coda no PB, e, também, como é discutido nas subseções seguintes, às assimetrias existentes entre a fonologia infantil e a fonologia adulta no que se refere às mesmas questões que geram os embates teóricos revisados. Entende-se que o uso do termo rima ramificada permite transitar mais livremente sobre assunto.

²⁷ 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).

Constituinte silábico	Ramificação	Posição na palavra	Segmentos licenciados		Exemplos
Onset	Simple	Inicial e medial	Todos		<pat o ; boc a ; ch ave ; maç ã ; gele ia ; fac a rod a ; gal o ; nó ; len ha *; cal ha *; ex em pl o ; ar ar a*>
	Complexa		Oclusiva ou fricativa + líquida lateral	/pl, bl, tl*, dl*, kl, gl, fl, vl*/	<pl an ta; bibli ot eca; at l eta; cl or o; gl ob o; fl or >
			Oclusiva ou fricativa + líquida não-lateral	/pr, br, tr, dr, gr, fr, vr*/	<pr at o; br an co; tr at or; dr á cula; cr aq ue; gr oss o; liv ro >
Rima	Complexa	Medial	/S/, /N/, /r/ e /l/** (Bisol, 1999)		<cost a ; vent o ; vermel h o; salt o >

*Formas pouco produtivas na língua e presentes apenas em posição medial de palavra.

**Realizado como [w].

Figura 13 – Quadro de segmentos licenciados para o Português conforme o constituinte silábico.
Fonte: elaboração própria.

A respeito da frequência dessas estruturas no léxico do PB, a Figura 16 apresenta um levantamento realizado por Viaro e Guimarães-Filho (2007), que tomou como base 150.875 itens lexicais presentes no Dicionário Houaiss da Língua Portuguesa:

silaba	Total		Tônicas		Pretônicas		Pós-tônicas não-finais		Pós-tônicas finais	
CV	378340	60.6%	70493	46.7%	203185	59.3%	16697	90.7%	87965	78.0%
CVC	96019	15.4%	44675	29.6%	47172	13.8%	8	0.0%	4164	3.7%
V	52592	8.4%	7981	5.3%	37365	10.9%	439	2.4%	6807	6.0%
CCV	27767	4.4%	5127	3.4%	19674	5.7%	1251	6.8%	1715	1.5%
VC	26826	4.3%	5663	3.8%	21112	6.2%	1	0.0%	50	0.0%
CGV	12200	2.0%	298	0.2%	1361	0.4%	8	0.0%	10533	9.3%
CVG	11453	1.8%	5884	3.9%	5568	1.6%	0	0.0%	1	0.0%
CVGC	6633	1.1%	6477	4.3%	121	0.0%	0	0.0%	35	0.0%
CCVC	5171	0.8%	1920	1.3%	3153	0.9%	0	0.0%	98	0.1%
GV	1828	0.3%	503	0.3%	413	0.1%	9	0.0%	903	0.8%
VG	1770	0.3%	429	0.3%	1339	0.4%	0	0.0%	2	0.0%
CCVG	909	0.1%	289	0.2%	620	0.2%	0	0.0%	0	0.0%
CVCC	772	0.1%	116	0.1%	547	0.2%	0	0.0%	109	0.1%
CGVC	608	0.1%	214	0.1%	271	0.1%	0	0.0%	123	0.1%
CCVCC	422	0.1%	82	0.1%	329	0.1%	0	0.0%	11	0.0%
VCC	310	0.0%	20	0.0%	290	0.1%	0	0.0%	0	0.0%
GVC	308	0.0%	271	0.2%	28	0.0%	0	0.0%	9	0.0%
VGC	334	0.0%	238	0.2%	96	0.0%	0	0.0%	0	0.0%
outras	529	0,1%	195	0,1%	121	0,0%	0	0,0%	213	0,2%
total	624791	100%	150875	100%	342765	100%	18413	100%	112738	100%

Figura 14 – Frequência de estruturas silábicas do Português Brasileiro de acordo com a tonicidade. Fonte: VIARO; GUIMARÃES-FILHO, 2007, p. 31.

Da figura, percebe-se que o molde silábico mais frequente no léxico do PB é CV, tanto em contextoônico quanto em átono, totalizando um percentual médio de 60,6%. Os demais formatos silábicos, todos ramificados em *onset* ou rima (com exceção de V), distribuem-se nos 39,4% de contextos restantes. Nesse grupo, o molde CVC é destaque, com frequência superior à de V, e o *onset* ramificado em CCV, 4º colocado no *ranking* geral, carrega um percentual de apenas 4,4%.

Albano (2001) apresenta um levantamento especificamente relacionado à frequência dos grupos consonantais licenciados para formar o *onset* complexo no PB, tendo como *corpus* o Minidicionário Aurélio, e que pode ser verificado em (1):

$$(1) \text{tr} > \text{pr} > \text{br} > \text{gr} > \text{fl} > \text{kr} > \text{kl} > \text{pl} > \text{fr} > \text{dr}$$

A sequência apresentada mostra que os grupos formados pela líquida não-lateral são mais frequentes que aqueles formados pela líquida lateral, bem como que aqueles formados por oclusivas também são mais frequentes do que os compostos

por fricativas. Além disso, não parece haver preferência por combinações com base em traços de ponto e de sonoridade.

Tendo em vista a caracterização da sílaba feita até aqui, é possível sustentar aquilo que Nespor e Vogel [1986] (1994), conforme aponta Freitas (2017), atribuem à unidade como seu principal papel na gramática fonológica, qual seja: organizar a camada melódica (segmentos e traços), relacionando-a da forma mais básica à camada prosódica das línguas. Com isso, é possível melhor compreender os aspectos implicados na aquisição das estruturas silábicas, tema da próxima subseção.

2.3 A aquisição de estruturas silábicas complexas

Tendo clareza a respeito dos principais aspectos teóricos envolvidos no percurso de aquisição da linguagem, tanto falada quanto escrita, é possível compreender o que está em jogo no estudo específico sobre a aquisição de estruturas silábicas complexas. Nesta subseção, são apresentadas e revisadas as principais pesquisas que se dedicaram a estudar este tema nas áreas de aquisição fonológica e de aquisição da escrita, respectivamente.

2.3.1 Fonologia

Conforme apontam Matzenauer e Miranda (2012), o advento das fonologias não-lineares acarretou em uma mudança na compreensão sobre a aquisição silábica, passando esta a significar um processo de construção, pelas crianças, da estrutura da sílaba em seus constituintes menores. Anterior a isso, a aquisição da sílaba era majoritariamente analisada sob a ótica da Fonologia Natural (STAMPE, 1973), cujo pressuposto é o de que a aquisição fonológica significaria suprimir (ou manter) processos fonológicos inatos a fim de ajustar a gramática infantil às características específicas de uma língua. No caso da sílaba, como exemplificam Matzenauer e Miranda (2012, p. 108), “produções infantis como [‘pɛda] para ‘pedra’ e [‘fɔtʃi] para ‘forte’, por exemplo, seriam uma consequência da aplicação de processos de Redução de Encontro Consonantal e Apagamento de Consoante Final de Sílaba”. Assim, o trabalho da criança que está adquirindo uma língua que admite sílabas ramificadas, como é o caso do Português, seria suprimir esses processos.

A aquisição silábica no Português Brasileiro, tanto na perspectiva do desenvolvimento típico como do atípico, obteve grande contribuição das investigações realizadas no Centro de Estudos sobre Aquisição e Aprendizagem da Linguagem (CEAAL/PUC-RS), especialmente nas décadas de 1990 e 2000. As pesquisas realizadas no Centro tiveram como pressuposto comum a teoria gerativista em suas diferentes expressões adotadas para abordar a aquisição fonológica (Fonologia Natural, Fonologia Autossegmental, Teoria da Sílabas, Teoria da Otimidade) e forneceram subsídios importantes para o tratamento fonoaudiológico e para discussões caras à teoria linguística. A maior parte dos resultados obtidos está compilado em um volume organizado pela professora Regina Lamprecht e colaboradoras (2004), por meio do qual é possível ter acesso de forma sistematizada ao percurso de desenvolvimento fonológico pelo qual passam as crianças brasileiras²⁸.

Nesta subseção, são apresentados os principais resultados sintetizados na obra de Lamprecht et al. (2004) a respeito da aquisição fonológica da rima ramificada e do *onset* complexo no PB, respectivamente. Referências a outros estudos são feitas sobretudo quando há controvérsia na literatura envolvendo tópicos específicos.²⁹

2.3.1.1 Rima ramificada

A aquisição da rima ramificada, considerada a partir da proposta de Bisol (1999) para os diferentes segmentos que ocupam a coda no sistema do PB, bem como a posição na palavra em que ocorre (medial e final), tem em Mezzomo (2004)³⁰ uma descrição abrangente. A autora analisou dados de fala de crianças com idade entre

²⁸ Crianças brasileiras da região Sul do país.

²⁹ Na revisão sobre a aquisição fonológica das estruturas silábicas, optou-se por expor os resultados de estudos que descrevem aspectos mais gerais envolvidos no processo de aquisição, permitindo, assim, que seja traçado um perfil básico do processo de aquisição – o suficiente para os fins desta pesquisa. Em relação à aquisição do *onset* complexo, especificamente, é imprescindível referir a existência de outros importantes trabalhos no contexto brasileiro, com destaque para I. Miranda (2007), Vassoler (2016) e Barbieri (2019). Estes estudos têm como diferencial refinadas análises acústicas e articulatórias dos dados e se valem de abordagens teóricas distintas da teoria fonológica formal que embasa os estudos presentes em Lamprecht et al. (2004), como a Fonologia de Usos e a Fonologia Gestual.

³⁰ Durante esta subseção é utilizado o termo *coda* para se referir aos dados da sua pesquisa, em virtude de Mezzomo (2004) considerar a proposta de Bisol (1999) em sua análise e em razão da mesma autora utilizar o termo ao longo de todo seu texto.

1:2 e 3:8³¹ extraídos dos bancos de dados INIFONO e AQUIFONO³², os quais passaram por análise de oitiva e acústica. Mezzomo (2004) observou pelo menos dois tipos de diferenças importantes na aquisição deste constituinte silábico: a idade em que são adquiridos os diferentes segmentos que o ocupam e a posição na palavra em que a coda aparece.

A consoante nasal /N/ é a primeira a ser adquirida na coda considerando-se a média de idade observada para ambas posições na palavra. Na posição final, emerge ao 1:4 e estabiliza ao 1:7. As estratégias mais utilizadas são a omissão ([‘be] para ‘bem’) e a ditongação sem nasalização ([‘bow] para ‘bom’). Com relação aos contextos favorecedores à produção correta da nasal em posição final, são significativos a tonicidade da sílaba, a vogal /u/ como precedente e as palavras dissílabas.

Na posição medial, observa-se a mesma idade de surgimento, porém a estabilização é mais tardia, sendo verificada aos 2:2. As principais estratégias utilizadas pelas crianças enquanto não se observa a produção correta do segmento consistem na omissão da nasal ([‘petʃi] para ‘pente’) e a nasalização da vogal ([‘ãda] para ‘anda’). A tonicidade da sílaba, a vogal precedente /a/, as consoantes seguintes labiais e as palavras dissílabas constituem-se como contextos favorecedores à produção correta da nasal em posição medial.

É necessário lembrar o importante fato de que não há consenso a respeito da existência de uma consoante nasal em coda medial no sistema do Português (cf. MATEUS; ANDRADE, 2000; COSTA; FREITAS, 2001), conforme discutido na subseção 2.3. Assim, alguns pesquisadores propõem que as crianças adquirem, ao invés de uma consoante nasal, vogais nasais (cf. COSTA; FREITAS, 2001), motivados, também, pelo alto grau de omissões da suposta consoante assim como pela precocidade de sua aquisição. Esse aspecto será retomado na subseção que trata sobre a aquisição de sílabas complexas na escrita (2.4.2), onde a discussão

³¹ Convenção para designar a idade dos sujeitos da pesquisa. Neste caso, corresponde a 1 ano e 2 meses e 3 anos e 8 meses.

³² O Banco de Dados de Aquisição Fonológica (AQUIFONO) reúne entrevistas realizadas com 310 crianças das cidades de Pelotas (RS) e Porto Alegre (RS), com idade entre 2:0 e 7:1. Já o Banco de Dados de INIFONO reúne entrevistas realizadas com 100 crianças das mesmas cidades, com idade entre 1:0 e 2:0. A constituição de ambos os bancos foi uma iniciativa das professoras Carmen Matzenauer e Regina Lamprecht. Informações mais detalhadas podem ser conferidas em Lamprecht et al. (2004).

sobre a assimetria observada entre o sistema fonológico adulto e o infantil recebe contribuição singular à luz dos dados de escrita inicial.

A consoante lateral /l/³³ é a segunda consoante a ser adquirida em posição de coda. Em posição final de palavra, o /l/ surge antes mesmo da própria nasal, ao 1:2 e se estabiliza em seguida, ao 1:4. Enquanto não está estabilizado, sua produção pode ser omitida ([‘sɔ] para ‘sol’) ou apresentar um alongamento da vogal precedente ([‘go:] para ‘gol’). A probabilidade de produção correta aumenta quando o segmento estiver em sílaba tônica, quando a vogal precedente for /ɛ/ e quando as palavras forem monossílabas.

Já em posição medial, a lateral surge ao 1:6 e se estabiliza apenas aos 3:0, razão pela qual, através do cômputo da média, o /l/ termina sendo a segunda consoante em coda a ser adquirida pelas crianças. As estratégias utilizadas pelas crianças para lidar com essa estrutura consistem sobretudo na omissão da consoante ([‘kasa] para ‘calça’) e a coalescência ([‘fɔta] para ‘falta’), e os contextos favorecedores à sua produção correta são a posição tônica, a presença da vogal precedente /ɛ/ e a presença do traço [coronal] na consoante seguinte.

Em relação à aquisição da lateral em coda, é preciso fazer menção a uma controvérsia semelhante àquela existente em torno do estatuto da consoante nasal. Novamente, não há consenso a respeito de qual seja a forma subjacente, uma vez que o segmento é realizado na fala invariavelmente como [w], à exceção de poucos dialetos brasileiros (cf. CALLOU; MORAES; LEITE, 2013). Por essa razão, e pela precocidade observada na aquisição em se comparando às demais consoantes da classe das líquidas nas diferentes posições que podem ocupar na estrutura silábica, estudos como os de Hernandorena (1990) sustentam que as crianças adquirem um ditongo que ocupa o núcleo da sílaba, e não o /l/ que ocupa coda. O estudo de Bonilha (2005) sugere o mesmo apenas para o /l/ em posição final de palavra.

Em terceiro lugar no percurso de aquisição dos segmentos em coda está a fricativa /s/³⁴ que, no dialeto do Sul do Brasil, é realizada como [s] ou [z]. Uma

³³ Ver Azambuja (1998), que estuda especificamente a aquisição das líquidas laterais no PB, para maior detalhamento sobre questões concernentes à aquisição de tais segmentos em coda.

³⁴ Ver Savio (2001), que estuda especificamente a aquisição das fricativas /s/ e /z/ no PB, para maior detalhamento sobre questões concernentes à aquisição de tais segmentos em coda.

característica verificada no processo de aquisição desse segmento é a não-linearidade, isto é, são observadas diversos momentos de aparente “regressão” na produção da fricativa, formando a “curva em U”, típica do desenvolvimento infantil de modo geral (STRAUSS; STAVY, 1982).

Em posição final, a fricativa surge ao 1:6 e estabiliza aos 2:6. Ao longo desse intervalo, são observadas estratégias tais como a palatalização do [s] ([na‘liʃ] para ‘nariz’), a epêntese ([fe‘lizi] para ‘feliz’) e a omissão ([a‘χo] para ‘arroz’). A produção correta é favorecida pela presença de /a/, /e/ e /o/ como vogais precedentes, pela posição postônica da sílaba e por palavras dissílabas.

Em posição medial, o /S/ surge aos 2:0 e estabiliza aos 3:0, período no qual são especialmente observadas estratégias como a omissão ([‘kaka] para ‘casca’) e metátese ([si‘kada] para ‘escada’). A presença das vogais precedentes /a/, /e/ e /ε/, a posição tônica da sílaba, a presença de uma consoante seguinte com ponto [coronal] e palavras dissílabas favorecem a produção correta da fricativa em coda medial.

A última consoante a ser adquirida na coda é a líquida não-lateral /r/³⁵, que em posição final surge ao 1:11 e estabiliza somente aos 3:10. Nesse percurso, são observadas preferencialmente omissões ([kõⁿto] para ‘cantor’), semivocalizações ([ta’toj] para ‘trator’) e epênteses ([ku’εri] para ‘colher’), sendo a produção correta favorecida pela posição tônica da sílaba, pela precedência das vogais /a/ e /e/ e por palavras tri ou polissílabas.

Em posição medial, 2:2 é a idade de surgimento do /r/ cuja produção correta estabiliza-se também aos 3:10. As estratégias mais frequentes constatadas consistem na omissão ([‘kata] para ‘carta’), na metátese ([‘grafu] para ‘garfo’) e na semivocalização ([ugaw’zɪnu] para ‘lugarzinho’). Os contextos fonológicos que contribuem à produção conforme o alvo são as vogais precedentes /ε/ e /i/, a posição tônica da sílaba, uma consoante seguinte [dorsal] e palavras dissílabas.

³⁵ Ver Miranda (1996), que estuda especificamente a aquisição das líquidas não-laterais no PB, para maior detalhamento sobre questões concernentes à aquisição do segmento rótico em coda.

2.3.1.2 *Onset* complexo

A aquisição do *onset* complexo no curso de desenvolvimento típico do PB, considerando-se sua caracterização no sistema (ver Figura 13), é descrita por Ribas (2002), que analisa transversalmente dados de crianças brasileiras com idades entre 1;0 e 5;3, extraídos dos bancos de dados INIFONO e AQUIFONO, os quais passaram por análise de oitiva.

Um dos aspectos importantes, salientado por Ribas (2002), que difere o percurso de aquisição do *onset* complexo em relação ao da rima ramificada, é que não há condicionamento segmental, ou seja, não há diferença significativa no tempo de aquisição ao serem comparados os grupos consonantais com a líquida lateral e a não-lateral. Os grupos de *onset* formados por /l/ e /r/ são adquiridos ao mesmo tempo, apresentam curvas desenvolvimentais semelhantes e percursos marcados por vários períodos de aparente regressão. Tal fato é interessante, pois poder-se-ia ter como hipótese que a aquisição do grupo consonantal com /l/ seria anterior a do /r/, uma vez que tanto em coda quanto em *onset* simples a líquida lateral é estabilizada antes da não-lateral³⁶. No entanto, há diferenças entre os grupos no que se refere à produtividade de itens lexicais: o número de dados obtidos com o grupo obstruinte + /r/ é 8 vezes maior comparado ao grupo com /l/. Tal diferença corresponde ao que se verifica no léxico adulto, para o qual se observa proporção semelhante (cf. VIARO; GUIMARÃES-FILHO, 2007; ALBANO, 2001).

Outro aspecto que distingue os grupos refere-se aos contextos favorecedores à produção correta do *onset* complexo. Para o grupo da líquida lateral, a presença da vogal /a/ no núcleo da mesma sílaba e de /p/ na primeira posição do *onset* aumentam a probabilidade de produção conforme o alvo. Já para o grupo da líquida não-lateral, o sucesso na produção é favorecido pela presença de /b/ e /v/ na primeira posição do *onset*, pela posição medial da sílaba CCV, pelo contexto átono e pela presença das vogais /i u a/ no núcleo da mesma sílaba.

³⁶ Bonilha (2005), em estudo longitudinal de caso que analisa a aquisição do PB em relação ao segmento, à sílaba e ao acento à luz da Teoria da Otimidade de base conexionalista, encontra resultado diferente de Ribas (2002): no seu conjunto de dados, o encontro consonantal com a líquida lateral é adquirido antes do encontro com a líquida não-lateral. Bonilha (*op. cit*) explica este fato com base na escala de sonoridade proposta por Bonet e Mascaró (1997), que prevê sonoridade maior para líquida não-lateral, aproximando-a das vogais. Por essa perspectiva, há maior distanciamento em termos de soância entre vogal e lateral, favorecendo a produção do encontro com essa combinação segmental.

Uma discussão a respeito da aquisição do *onset* complexo é sobre existência ou não de estágios no seu percurso e sobre a caracterização desses estágios, caso existam. Para verificar este aspecto, Ribas (2002) observa as estratégias utilizadas pelas crianças quando produzem palavras cujo alvo contém estrutura silábica com *onset* ramificado. A autora propõe que a aquisição do *onset* complexo é caracterizada pela existência de 2 estágios: o primeiro marcado pela produção apenas de C¹V (isto é, com omissão da líquida) e o segundo pela produção correta da estrutura CCV. Não existem, portanto, estágios intermediários. Tal proposição é motivada especialmente pelo alto índice de redução do encontro consonantal verificado na amostra analisada, em comparação com outras estratégias observadas, cujos percentuais não passam de 3%.

Freitas (1997), em análise sobre a aquisição das estruturas silábicas no Português Europeu, propõe mais nuances no percurso: o primeiro estágio consistiria na seleção de itens lexicais cuja composição não apresentasse sílabas CCV. Em seguida, são produzidos itens lexicais compostos pelo molde CCV, a líquida é omitida da estrutura ou, ainda, todo o grupo consonantal é omitido. Antes ainda do estágio final, caracterizado pela produção correta, haveria um momento no qual se verifica a produtividade da epêntese vocálica entre a sequência de consoantes do *onset*.

Em texto recente, Freitas (2017) chama a atenção para o fato de se verificar essa assimetria entre o PB e o PE, especialmente no que se refere à produtividade da epêntese para além da omissão da líquida apenas. A motivação para isso, segundo a autora, pode estar ligada à maior distância entre a representação fonológica e as formas fonéticas disponíveis no *input* para as crianças portuguesas relativamente às vogais. É necessário lembrar que, no PE, especialmente em contexto átono, são verificadas altas taxas de supressão vocálica que geram sequências consonânticas, o que pode dificultar a identificação dos núcleos silábicos (por exemplo, [dʒʃprv'nir] para 'desprevenir'). Dessa forma, a epêntese surgiria como uma estratégia que evidencia um momento no qual o *onset* complexo é tratado pela criança, na verdade, como um caso de uma sequência consonantal decorrente de supressão vocálica.

Apesar de serem poucas as estratégias verificadas como de alta produtividade (omissão da líquida para PB e PE e epêntese apenas para o PE), Ribas (2002) destaca que as demais estratégias utilizadas pelas crianças podem indiciar

características importantes sobre o modo como vão sendo construídas as representações silábicas.

A metátese, a epêntese e a substituição da líquida, por exemplo, não são verificadas em fases mais precoces (como da faixa etária entre 1:0 e 1:11), período no qual predomina a omissão. Tais estratégias, que envolvem necessariamente a manipulação da líquida, vão emergir em períodos mais tardios, evidenciando o conhecimento que tem a criança do *onset* no formato ramificado, motivo pelo qual, se não se pode considerar o *onset* complexo ainda plenamente adquirido, é possível considerá-lo como já acessível em termos de representação subjacente.

Outra estratégia, destacada por Ribas (2002), que auxilia a corroborar o raciocínio precedente, e cujo relato já se verificava em Hernandorena (1988), é a não-palatalização de /t/ ou /d/ em contexto de omissão da líquida. Em uma sequência como [‘tiλu] para ‘trilho’ esperar-se-ia que /t/ fosse produzida na forma palatalizada, uma vez que o contexto para a regra está presente. Esse tipo de dado é interpretado por Hernandorena (1988) como um indício de que o constituinte silábico correspondente à ramificação do *onset* estaria presente na subjacência, atuando como bloqueador da regra de palatalização.

2.3.1.3 Síntese e aspectos para discussão

Nas precedentes subseções, foram expostas as principais características da aquisição fonológica de estruturas silábicas complexas. Na Figura 15, que segue, esse percurso pode ser visualizado de forma sintetizada:

Constituinte silábico	Segmentos	Posição na palavra	Faixa etária		Contextos favorecedores à produção correta	Estratégias
			Surgimento	Aquisição		
Coda	/N/	Final	1:4	1:7	Sílabas tônicas, vogal /u/ (precedente), palavras dissílabas	Omissão da nasal e ditongação sem nasalização
		Medial	1:4	2:2	Sílabas tônicas, vogal /a/ (precedente), consoantes labiais (seguinte), palavras dissílabas	Omissão da nasal e nasalização da vogal
	/l/	Final	1:2	1:4	Sílabas tônicas, vogal /e/ (precedente), palavras monossílabas	Omissão da líquida e alongamento da vogal precedente
		Medial	1:6	3:0	Sílabas tônicas, vogal /e/ (precedente), consoante [coronal] (seguinte)	Omissão da líquida e coalescência
	/S/	Final	1:6	2:6	Sílabas pós-tônicas, vogais /a e o/ (precedente), palavras dissílabas	Palatalização da fricativa, epêntese e omissão da fricativa
		Medial	2:0	3:0	Sílabas tônicas, vogais /a e e/ (precedente), consoante [coronal] (seguinte), palavras dissílabas	Omissão da fricativa e metátese
	/r/	Final	1:11	3:10	Sílabas tônicas, vogais /a e/ (precedente), palavras tri ou polysílabas	Omissão da líquida, semivocalização da líquida e epêntese
		Medial	2:2	3:10	Sílabas tônicas, vogais /e i/ (precedente), consoante [dorsal], palavras dissílabas	Omissão da líquida, metátese e semivocalização da líquida
Ataque	Oclusiva ou fricativa + líquida lateral	Inicial e Medial	1:8 – 2:2*	5:0	Vogal /a/ no núcleo adjacente e consoante /p/ na primeira posição do ataque	Omissão da líquida, substituição da líquida e metátese
	Oclusiva ou fricativa + líquida não-lateral				Sílabas átonas, posição medial, consoantes /b/ e /v/ na primeira posição do ataque, vogais /i u a/ no núcleo adjacente	

Figura 15 – Quadro-síntese do percurso de aquisição fonológica das estruturas silábicas complexas no Português Brasileiro.

Fonte: elaboração própria.

De forma análoga, a cronologia de aquisição dos segmentos conforme a posição que ocupam em cada constituinte tem sido sintetizada da seguinte maneira:

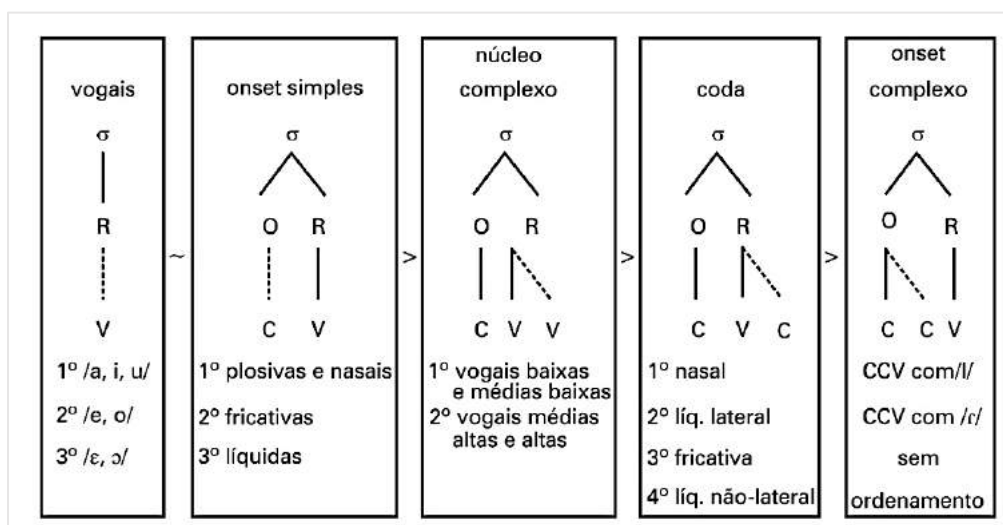


Figura 16 – Cronologia de aquisição dos segmentos do PB conforme posição ocupada no constituinte silábico.

Fonte: Ribas (2004), com base nos estudos de Rangel (2002), Lamprecht (1990), Hernandorena (1990), Miranda (1996), Azambuja (1998), Savio (2001), Oliveira (2002), Bonilha (2000), Mezzomo (1999) e Ribas (2002).

Dada a descrição realizada sobre a aquisição da rima ramificada e do *onset* complexo no PB, algumas particularidades observadas nos dados apresentados merecem mais algum comentário.

A primeira delas diz respeito às diferenças constatadas na aquisição dos segmentos em posição final e em posição medial de palavra, sobretudo no caso da rima ramificada, tanto no que se refere ao tempo de aquisição quanto ao tipo de estratégia utilizado quando o segmento ainda não está estabilizado.

Cabe notar que esse tipo de distinção pode ser observado em outros níveis da gramática. Por exemplo, Slobin (1980) relata um estudo sobre como crianças pequenas de diferentes línguas processam indicadores locativos. O autor comparou crianças que estavam adquirindo línguas que realizam esses indicadores por meio de preposições com crianças cujas línguas os realizavam por meio de posposições. O resultado mostrou que as crianças de línguas posposicionais adquiriram mais precocemente os indicadores locativos, em relação às preposicionais, o que levou Slobin (1980, p. 153) a considerar que, “por várias razões de atenção e memória, as crianças acham as terminações das palavras e enunciados mais salientes que os inícios e meios”. Assim, fatores de ordem psicolinguística parecem atuar para que posições finais de palavra sejam mais salientes que as mediais, justificando, no caso

da fonologia, a aquisição precoce dos segmentos em rima ramificada final em relação à medial.

Outra questão importante trata da diferença temporal existente entre a aquisição da rima ramificada e do *onset* complexo, bem como a diferença na aquisição dos mesmos segmentos em diferentes posições silábicas (como é o caso de /r/ e /l/ considerando-se apenas os constituintes neste trabalho descritos).

Tais distinções têm sido tratadas na literatura (cf. MATZENAUER; MIRANDA, 2012; FREITAS, 2017; MATZENAUER-HERNANDORENA, 2001) como uma evidência para a ideia de a aquisição segmental ser prosodicamente condicionada, isto é, depender da aquisição das estruturas silábicas (e de unidades mais altas da hierarquia prosódica). O fato de as crianças estabilizarem os segmentos em diferentes posições em momentos diferentes do desenvolvimento indicia que o que está em jogo não é propriamente uma dificuldade de realização fonética do segmento, mas sim a indisponibilidade do constituinte silábico.

Contribui com a sustentação desse raciocínio observar, também, a quase categórica influência do acento para a produção correta dos diferentes segmentos que ocupam a rima ramificada tanto em posição medial quanto final. Nota-se, neste caso, que a própria aquisição silábica, anterior ainda à segmental, parece sofrer um condicionamento da unidade imediatamente mais alta, o pé métrico, fato apreciado por Matzenauer-Hernandorena (2001). Mesmo no único caso no qual a sílaba em posição tônica não atua como contexto favorecedor à produção correta, mas sim a sílaba em posição postônica, a saber /S/ final, continua a se tratar de um caso cujo domínio é o pé métrico do acento primário da palavra. É semelhante ao que relata Miranda (1996) quando analisa a aquisição de /χ/: o segmento estabiliza primeiro em sílabas cujo domínio é o pé do acento primário (tanto em posição tônica como átona), e depois em sílabas fora do pé do acento.

Com isso, é possível recuperar e sustentar a ideia exposta na subseção 2.1.1 de que a aquisição de unidades fonológicas é desencadeada ou condicionada, em grande medida, por unidades superiores da hierarquia – o que não significa dizer que as unidades menores não afetarão o processamento em algum momento. Porém, a noção fundamental neste caso é a de proporcionalidade, isto é, de que o processamento seja *preponderantemente* de cima para baixo ou *top-down*, podendo, em alguns momentos do desenvolvimento, assumir a rota contrária *bottom-up*.

Por último, é importante ter em vista, em razão das discussões levantadas nesta dissertação e referidas nas subseções a seguir, as questões controversas acerca da composição segmental da rima ramificada. Dada a precocidade observada na aquisição das consoantes nasal e lateral em relação às consoantes rótica e fricativa, e considerando as diferentes propostas, para a fonologia adulta, no que se refere à representação da nasalidade medial no esqueleto silábico, as crianças adquiririam, de fato, as consoantes nasal e lateral em coda? As representações das fonologias infantil e adulta para esses segmentos diferem ou não? Em que medida?

2.3.2 Escrita

É importante notar que, por se tratar do estudo de uma unidade cuja natureza primária é fonológica, o tratamento adequado das sílabas complexas na escrita depende que se leve em consideração possíveis relações, não necessariamente simétricas, com a teoria fonológica bem como com os fatos observados na aquisição da fonologia. No entanto, apesar de o sistema de escrita alfabética operar a partir de um princípio que tem na fonologia da língua seu domínio de referência, é preciso considerar também suas especificidades. A principal delas refere-se à ortografia, que define por meio de convenção, nem sempre de forma transparente com a fonologia, as combinações de grafemas licenciadas para representação dos fonemas em um determinado sistema de escrita.

A aquisição da escrita, vista sob essa perspectiva, trata-se de um campo ainda pouco explorado e consolidado no Brasil, sobretudo se comparado ao da aquisição da fala. O pioneirismo desse tipo de investigação, segundo Miranda (2012, 2017), pode ser atribuído à linguista Maria Bernadete Abaurre, que enxergou nas escritas iniciais das crianças um material capaz de revelar aspectos sobre a representação linguística subjacente, especialmente no que diz respeito à informação fonológica. Inspirados nessa ideia que define um domínio linguístico formal para a investigação em aquisição da escrita, estão os estudos que passaram a ser desenvolvidos no GEALE, os quais procuram identificar não somente o conhecimento fonológico subjacente às escritas, mas também aqueles de outra ordem, como o ortográfico e o fonográfico, e também no GP^{EL}³⁷ que, à diferença do GEALE, parte de uma

³⁷ Grupo de Pesquisa Estudos sobre a Linguagem, coordenado pelo professor Lourenço Chacon (UNESP-Rio Preto/Marília).

concepção de linguagem como enunciação, agregando discussões acerca de efeitos das práticas de letramento, para além da análise da estrutura sonora da língua.

Nesta subseção, portanto, são apresentados e discutidos os principais estudos que se dedicaram a descrever e analisar a aquisição das sílabas complexas na escrita de crianças em fase inicial de escolarização sob a perspectiva das relações simétricas e assimétricas que podem se estabelecer entre teoria fonológica, aquisição fonológica, aquisição da escrita e aquisição da ortografia.

Assim como na subseção 2.4.1, a revisão começa pelos estudos que versam sobre a rima ramificada, seguindo, após, com aqueles que dedicados à análise do *onset* complexo, e culmina com uma síntese do exposto acompanhada de uma breve discussão acerca de questões relevantes envolvendo a aquisição das estruturas silábicas na escrita.

2.3.2.1 Rima ramificada

Os estudos de Abaurre (1991a, [1988] 2011), que deram origem às investigações cujo foco de interesse são as relações entre a fonologia e a escrita inicial, concentram-se na discussão sobre a grafia, pelas crianças, do que seriam codas nasais mediais, retomando a questão controversa sobre a assimetria atestada entre os dados de aquisição fonológica e as propostas para o sistema fonológico adulto (ver seções 2.3 e 2.4.1.1). O *insight* fundamental da pesquisadora é o de que a nasalidade medial seria interpretada pelas crianças como monofonêmica, o que levaria à consideração de vogais nasais no sistema, conforme proposta de Costa e Freitas (2001). No entanto, a ortografia representa a nasalidade fonológica com uma sequência de duas letras, o que levaria as crianças a reestruturarem sua representação subjacente, ajustando o estatuto da nasalidade ao formato bifonêmico VN, conforme a propostas de Câmara Jr. (1970) e Bisol (1999).

Tal ideia é recuperada por Miranda (2009a, 2009b, 2012, 2018, 2019), que procura sistematizar a discussão por meio da análise de um conjunto mais extenso de dados de escrita espontânea produzidos por crianças de 1ª a 4ª série, extraídos de textos do BATALE, além do acréscimo de dados das demais codas mediais do Português Brasileiro, a saber fricativa, rótica e líquida lateral, considerando-se a proposta de Bisol (1999). A autora aponta que, tanto em textos produzidos por alunos de escola pública quanto nos de escola particular, os erros ortográficos concentram-

se na grafia da líquida lateral e da nasal pós-vocálica, totalizando uma média de 40% e 15% das ocorrências respectivamente, enquanto a grafia das rimas com fricativa e rótica contabilizam apenas 5,5% e 4% de erros cada uma (MIRANDA, 2019).

Em termos de sua natureza, os erros são classificados pela autora de acordo com a proposta do GEALE, predominando aqueles de natureza fonológica em relação aos de natureza ortográfica. É importante notar que, para cada tipo de rima ramificada analisada por Miranda (2009a, 2009b, 2012, 2018, 2019), existem diferentes relações que se estabelecem entre fonologia e ortografia, o que dá motivo à divisão dos erros segundo diferentes naturezas (ver subseção 2.1.3).

Para a rima ramificada com lateral, Miranda (2019) considera que há relação biunívoca entre fonema e grafema se for considerada a forma fonológica /l/. Mas uma complexidade ortográfica é gerada em virtude da pronúncia semivocalizada, que iguala a nível fonético formas como ‘sal’ e ‘pau’. Assim, tanto erros como ‘gentiu’ para ‘gentil’ e ‘chapel’ para ‘chapéu’ seriam derivados da ambiguidade causada pela fala, que torna opaca esta relação específica entre fonema e grafemas e cuja aprendizagem se dá via memória gráfica.

No caso da nasal, além da complexidade fonológica, há uma complexidade ortográfica decorrente da observação que faz o sistema a respeito da assimilação de ponto a que está sujeita a consoante nasal: o grafema utilizado para representar a nasal é <m> antes de <p> e , em virtude da relação existente entre as duas consoantes adjacentes no que se refere ao traço [labial]; nos demais casos, é utilizada o grafema <n>. Por essa razão, não raro são registradas grafias como ‘tanto’ para ‘tanto’ e ‘bonba’ para ‘bomba’, as quais evidenciam a não-observância desta regra ortográfica definida via contexto fonológico.

A rima com fricativa, assim como a lateral, apresenta relações múltiplas entre fonema e grafemas. Neste caso, concorrem à representação de /s/ os grafemas <s> e <x>. No entanto, Miranda (2019) observa que o uso de <x> está condicionado à presença da vogal /e/ em contexto precedente, como na palavra ‘extenso’, o que torna esta convenção de caráter contextual e reduz o conjunto de palavras que seriam grafadas com ‘x’. Erros que decorrem da não-observância de tal relação são pouco frequentes, mas, quando ocorrem, aparecem especialmente em trocas de ‘x’ por ‘s’, como em ‘estra’ para ‘extra’.

A rima ramificada com rótica, por fim, é o caso mais simples em termos das relações fonografêmicas, não apresentando nenhum tipo de complexidade ortográfica. Por esse motivo, erros envolvendo sua representação são compreendidos potencialmente como de natureza fonológica ou fonográfica.

O tipo de erro fonológico mais frequente, em quaisquer dos tipos de rima ramificada, é a omissão do grafema que representa a consoante alvo ('grade' para 'grande'; 'feta' para 'festa'; 'pota' para 'porta'). No entanto, embora predomine a omissão da consoante, chama a atenção da autora a diversidade de estratégias utilizadas para grafar a nasalidade medial, que vão desde a inserção de um espaço em branco ('se tado' para 'sentado') até o uso do diacrítico til ('alevãto' para 'levantou'). Também se observa uma notável incidência de erros e de sua variedade sobre um item lexical específico da amostra analisada: a palavra 'grande'. Para Miranda (2009a), o arranjo estrutural composto de *onset* complexo mais nasalidade medial na sílaba /graN/ cria um contexto de supercomplexidade representacional para a criança, o que pode estar na base da alta taxa de erros concentrada neste item lexical. Tal interpretação é reforçada por Pachalski e Miranda (2018b) que, analisando o fenômeno da metátese na escrita, observam a influência da supercomplexidade presente em sílabas CCVN como um fator que pode motivar a ocorrência da metátese, como seria o caso em grafias como 'gerde' para 'grande' e 'palntando' para 'plantando'.

Outro tipo de erro peculiar observado por Miranda (2009a, 2018) no registro da nasal medial é a alteração na vogal adjacente, verificado em grafias como 'me~nga' para 'manga' e 'manta' para 'menta'. As vogais mais atingidas são /a/ e /e/, as quais, do ponto de vista fonológico, são bastante distantes em termos da sua caracterização interna. Entretanto, a autora se vale das considerações de Berti, Chacon e Pagliuso (2008), que apontam para a similaridade perceptual existente entre os dois segmentos quando afetados pela nasalidade. Assim, seria esse o motivo pelo qual se observam, na escrita, grafias que alternam entre o uso de 'a' e 'e' em contexto de nasalidade medial.

Os dados de aquisição da escrita analisados por Miranda (2009a, 2009b, 2012, 2018, 2019) revelam um comportamento contrário àquele observado na aquisição fonológica, uma vez que as rimas ramificadas com nasal e lateral são adquiridas precocemente em relação à fricativa e à rótica (cf. MEZZOMO, 2004). Nesse sentido, se a hipótese fosse a de que a AE segue o mesmo roteiro da AF, o esperado, na

escrita, seria que as crianças apresentassem mais erros na representação dos segmentos rótico e fricativo. Miranda (*op. cit.*), por entender que a escrita inicial reflete uma retomada do conhecimento fonológico internalizado, interpreta essa assimetria como um indicativo de que, na verdade, laterais e nasais não são consideradas como codas pelas crianças nesta etapa inicial do desenvolvimento da escrita – consequentemente, tampouco o são durante a aquisição fonológica. No entanto, considera que, com o decorrer da escolarização e do desenvolvimento ortográfico das crianças, a representação da nasalidade e do glide em núcleo silábico podem ser reprocessadas e, portanto, reestruturadas como consoantes nasal e lateral em coda.

Miranda (2009a, 2009b, 2012, 2018, 2019) sustenta a posição de que a *mudança representacional* é possível com base em dois pressupostos já explanados na subseção 2.1: (i) a aquisição da linguagem trata-se de um *processo de construção gradual* das estruturas da língua. Sendo a aquisição da escrita parte integrante desse processo, como uma espécie de segunda etapa, é plausível supor que haja não propriamente a criação de novas estruturas, mas sim algum tipo de remodelagem naquelas 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 o próprio conhecimento como objeto de atenção cognitiva, aspecto central no Modelo de Redescrição Representacional (MRR) proposto por Karmiloff-Smith (1994).

Além dos estudos de Miranda (*op. cit.*), vinculados ao GEALE, e de Abaurre (1991a, [1988] 2011), mais dois estudos merecem revisão, os quais tem como sujeitos crianças falantes de outras línguas e/ou variedades do Português.

O estudo de Vicente (2018) analisa dados espontâneos e controlados de escrita e de fala de crianças moçambicanas, as quais têm o Português moçambicano como L2 e o Changana como língua materna, uma língua bantu cujo molde silábico não inclui sílabas ramificadas. O objetivo do autor foi verificar o impacto da complexidade silábica no desenvolvimento fonológico e na aprendizagem da ortografia de crianças que frequentavam as 1ª, 2ª, 3ª e 5ª classes do Ensino Básico público de Moçambique. Em consonância com a literatura portuguesa (cf. COSTA; FREITAS, 2001) o autor não considera a consoante nasal como coda, portanto analisa apenas /S/, /r/ e /l/. Ele salienta que nenhuma das codas encontra-se estabilizada, na escrita, até a 5ª classe, sendo o apagamento das consoantes o tipo de erro mais frequente. No

desenvolvimento fonológico, Vicente (2018) aponta que a coda lateral não registra estabilidade em nenhum dos anos escolares, em contraposição à fricativa e à rótica sobre as quais a incidência de erros diminui consideravelmente. Resguardadas as diferenças entre as variedades do Português, bem como a especificidade dos sujeitos da pesquisa, que não tem o Português como L1, o resultado de Vicente (2018) vai ao encontro daqueles obtidos por Miranda (2019) para crianças brasileiras, que também apresentam maior dúvida na grafia da lateral. Vicente (2018), entretanto, não oferece uma interpretação sobre este resultado especificamente.

Ferreiro e Zamudio [2008] (2013), por fim, analisam dados de escrita controlada de crianças que tem o Espanhol como língua materna, a qual, em termos da estruturação silábica, é bastante semelhante ao Português. As autoras procuram compreender as dificuldades enfrentadas por crianças de 1ª série na escrita de sílabas CVC e CCV. No caso da rima ramificada, analisam apenas a grafia dos segmentos /r/ e /l/, a fim de manter conformidade com os segmentos licenciados para a segunda posição de *onset* complexo. São encontrados menos erros em sílabas do tipo CVC (43%), e os tipos de erro ortográfico mais frequentes verificados na grafia da rima são a omissão, seguida de epêntese, substituição (verificada apenas em coda lateral) e metátese, além de processos residuais.

2.3.2.2 *Onset* complexo

A aquisição do *onset* complexo na escrita, à diferença da rima ramificada, possui menor número de estudos dedicados à sua análise, provavelmente em virtude de não haver assimetrias importantes entre a aquisição fonológica e o sistema adulto.

Miranda (2019), analisando 2024 textos espontâneos do BATALE produzidos por crianças de 1ª a 4ª série, mostra que, em termos da distribuição geral, o *onset* complexo não chega a se apresentar como um grande foco de erros ortográficos, visto que em apenas 4% das palavras que continham essa estrutura foi verificado algum tipo de erro a ela relacionado. A autora destaca que a frequência dos grupos consonantais presentes na amostra analisada corresponde àquela verificada no léxico adulto (cf. ALBANO, 2001), uma vez que “os seis tipos mais frequentes na base do *Minidicionário Aurélio* - tr > pr > br > gr > fl > kr - são também os mais frequentes nos textos analisados” (MIRANDA, 2019, p. 14, grifos da autora). Apesar de não

apresentar a distribuição dos erros por grupo consonantal (consoantes acompanhadas de líquida lateral ou não-lateral), Miranda (2019) aponta quais as estratégias mais frequentes registradas no *corpus* como um todo, sendo dominante a omissão da líquida (média de 65,5% do total de erros), seguida da metátese (17,5%), substituição da líquida (11,5%) e epêntese (5,5%), resultado bastante semelhante àquele observado na aquisição fonológica (cf. RIBAS, 2002). Ao longo das séries escolares, os erros relacionados à grafia do *onset* complexo tornam-se cada vez menos frequentes, chegando a apresentar, já na 3ª série, apenas 1% de ocorrências.

Pachalski et al. (2014), em estudo com número de dados reduzido, também advindos do BATALE e produzidos por crianças de 1ª a 4ª série, apontam para resultado semelhante em termos da distribuição das estratégias para a grafia do *onset* complexo, com algumas nuances que merecem ser consideradas. Em quaisquer dos anos escolares, a omissão da líquida é a estratégia preferencial (média de 47% do total de erros), seguida da substituição da líquida (27,8%), metátese (10,9%), processos residuais (11%) e epêntese (1,7%).

Em estudo posterior no qual foram revisados os mesmos dados, Pachalski e Miranda (2017b) chamam a atenção para a oscilação observada na ocorrência da metátese entre os diferentes anos escolares: sua incidência é baixa nas 1ª e 2ª séries, com média de 6%. Na 3ª série o percentual de erros eleva-se a 20%, diminuindo para 10% na 4ª série. A maior incidência da metátese em faixas etárias mais avançadas é um fato, como notam as autoras (e como descrito na subseção 2.4.1.2), também observado na aquisição fonológica (cf. RIBAS, 2002) e pode ter explicação análoga nas duas modalidades de aquisição: se na fonologia a metátese pode evidenciar a incorporação do *onset* complexo ao inventário, uma vez que o uso dessa estratégia envolve necessariamente o deslocamento da líquida, na escrita a metátese evidenciaria o maior grau de explicitação das estruturas fonológicas que os sujeitos desenvolvem ao longo do processo de escolarização (PACHALSKI; MIRANDA, 2017b).

Santos (2013) analisa dados de fala e escrita (espontânea e controlada) de crianças portuguesas estudantes de 1º e 4º ano do Ensino Básico de Portugal, procurando investigar as diferenças de aquisição entre o *onset* ramificado e aquilo que chama de sequências consonânticas problemáticas (por exemplo, ‘**a**pto’ e ‘**p**sicologia’). De modo geral, os dados mostram que existem diferenças significativas entre a aquisição de *onsets* complexos e a de grupos consonânticos problemáticos,

entre a fala e a escrita e entre os dois níveis escolares testados. No tocante aos grupos consonantais do *onset* complexo, Santos (2013) verifica maior índice de erros nos grupos formados com a líquida lateral, tanto acompanhados de oclusivas quanto de fricativas.

Com relação à produtividade das estratégias utilizadas para grafar o *onset*, Santos (2013) registra a preferência pela epêntese, restando o segundo lugar para a omissão da líquida e o terceiro à metátese. Tal fato é interpretado por Freitas (2017) da mesma forma como a autora o interpreta para a aquisição fonológica do PE, período no qual também se verificam índices elevados de ocorrência da epêntese: a sua produtividade deve-se, possivelmente, à distância entre a representação fonológica das vogais e a sua realização fonética na variedade europeia do Português, caracterizada por altas taxas de supressão, o que gera sequências consonantais semelhantes ao *onset* complexo. Funcionaria, portanto, como estratégia que evidencia um momento no qual o *onset* complexo é tratado pela criança como um caso de uma sequência consonantal decorrente de supressão vocálica, o que acarreta na necessidade de demarcação dos núcleos silábicos. Além disso, ao comparar os dois anos escolares investigados, Santos (2013) registra um comportamento análogo ao verificado em Pachalski et al. (2015) no que se refere à metátese: trata-se de uma estratégia mais produtiva no 4º ano do que no 1º, contrariando a tendência geral dos erros ortográficos, qual seja, a de sua diminuição ao longo dos anos escolares. Tal evidência favorece, portanto, a hipótese interpretativa de Pachalski e Miranda (2017b) explanada anteriormente.

Cumprir observar novamente os resultados de Vicente (2018), que analisa dados espontâneos e controlados de escrita e de fala de crianças moçambicanas que tem o Português como L2. O autor mostra que as crianças apresentam mais erros na escrita da sequência *obstruinte + lateral* (CL), assim como o observado em Santos (2013). Em termos de estratégias, aquelas identificadas como mais produtivas foram a omissão da líquida, no caso dos grupos constituídos pela líquida não-lateral, e, nos casos dos grupos com a lateral, há um equilíbrio entre a omissão da segunda consoante e a epêntese vocálica. Vicente (2018) interpreta esses resultados com base em Veloso (2003, 2006), que defende haver uma representação distinta para as sequências de *onset* complexo na fonologia de crianças portuguesas ainda não alfabetizadas, isto é, para a fonologia infantil, sequências do tipo *obstruinte + vibrante* (CR) seriam consideradas tautossilábicas, mas sequências CL seriam

heterossilábicas. Assim, são esperados mais erros nesta sequência que tem representação distinta ao se comparar o sistema fonológico adulto, que serve de base para a composição do sistema ortográfico, com o sistema infantil. Também a epêntese é esperada como estratégia produtiva nesse sentido, pois ela gera como resultado a separação da sequência segmental e a consequente realocação dos segmentos na estrutura silábica (uma sílaba é transformada em duas) – uma sequência heterossilábica.

Por fim, é também necessário retomar o estudo de Ferreiro e Zamudio [2008] (2013), que analisam dados de escrita controlada de crianças mexicanas que tem o espanhol como língua materna. As autoras observam mais erros em sílabas CCV (74,5%), sendo a omissão da segunda consoante o tipo de erro mais frequente, seguido de metátese, substituição e epêntese, além de alguns processos residuais. A interpretação que Ferreiro e Zamudio [2008] (2013) dão aos resultados que encontram, tanto para a grafia de *onset* complexo quanto para a grafia de rima ramificada, é oportuna considerando o que se discute na próxima subseção, sendo, portanto, apresentada no texto que segue.

2.3.2.3 Síntese e aspectos para discussão

Nas precedentes subseções, foram expostas as principais características disponíveis na literatura a respeito da aquisição escrita de estruturas silábicas complexas. Algumas questões relevantes podem ser suscitadas a partir da observação dos resultados apresentados nos estudos revisados e de sua comparação com os resultados encontrados na aquisição fonológica, alimentando a discussão proposta nesta pesquisa. Com isso, mais alguns aspectos gerais a respeito da aquisição de sílabas complexas na escrita, sem sua pormenorização de acordo com cada constituinte, também merecem ser considerados, e são explorados, portanto, nesta subseção que encerra a fundamentação teórica do trabalho.

Antes, é importante esclarecer que a busca por estabelecer relações entre aquisição da fonologia e aquisição da escrita, neste trabalho, não pressupõe a ideia de que a segunda irá replicar processos ou estágios da primeira. Existe uma expectativa implícita na literatura da área (cf. SOARES, 2016; VICENTE, 2018; SANTOS, 2013) de que a aquisição da escrita poderia replicar padrões da aquisição fonológica. Tal expectativa parece vir, em parte, da ideia de *desenvolvimento*, que

integra ambos os processos de aquisição. Se um sujeito está em curso desenvolvimental, algumas etapas, estágios, níveis ou fases, a depender do modelo teórico, são esperadas como passíveis de verificação. No entanto, atuam em cada curso diversas variáveis que podem frustrar essas expectativas, obedecendo à natureza do conhecimento e da habilidade que estão em desenvolvimento. Entende-se, aqui, que é justamente este o caso entre aquisição da fonologia e da escrita, como em diferentes momentos já se esclareceu neste trabalho. Assim, ao se colocar os percursos de aquisição lado a lado, espera-se observar mais comportamentos distintos do que similares – e delinear essas (as) simetrias é necessário considerando a expectativa antes relatada de que a aquisição da escrita pode ser descrita e explicada pelos mesmos mecanismos da aquisição fonológica.

A observação de relações simétricas e assimétricas entre os dois processos de aquisição também se justifica por suscitar discussões acerca das diferenças existentes entre a fonologia infantil e a fonologia adulta, o que fica claro por meio dos estudos de Miranda (2009a, 2009b, 2012, 2018, 2019), revisados na subseção 2.4.2.1. Por sua vez, essas discussões também permitem que sejam fortalecidas ou enfraquecidas certas predições feitas pelos estudos de aquisição fonológica sobre o estatuto de determinadas unidades na fonologia infantil.

Com isso em vista, ao serem comparados os percursos de aquisição fonológica e escrita de sílabas ramificadas no PB, as simetrias são facilmente constatadas. A principal delas parece reservar-se ao tipo de estratégia utilizada para produzir as estruturas, com preferência pela omissão das consoantes que as constituem. Com o avanço dos anos escolares/faixas etárias, outras estratégias aparecem com maior incidência, sem, contudo, superar a preferência pelas omissões, como é o caso da metátese, das alterações vocálicas (especificamente na grafia da nasalidade medial) e das substituições.

As assimetrias entre os dois percursos também não são difíceis de mapear. A mais notável delas diz respeito à aquisição da rima ramificada: na escrita, testemunha-se o inverso daquilo que se observa na fonologia, uma vez que as consoantes nasais e líquidas concentram o maior índice de erros ortográficos, fato curioso que conta com ampla descrição e, sobretudo, explanação na literatura da área, conforme revisado na subseção 2.4.2.1.

Poder-se-ia perguntar, ainda, se, assim como existem diferenças cronológicas na aquisição fonológica do *onset* complexo em relação à rima ramificada, sendo

aquele de surgimento e aquisição mais tardios em relação à última (ver Figura 16), também existiriam tais diferenças na aquisição escrita, como suscita Soares (2016, p. 314):

[...] pouco se sabe sobre a ordem de aquisição dos demais padrões silábicos na aprendizagem da escrita pelas crianças [...]. No entanto, dados de estudos desenvolvidos sem esse objetivo precípua [...] e também práticas de alfabetização confirmam, de certa forma, ordem semelhante à identificada na aquisição da linguagem [...].

Nesse sentido, Vicente (2018), ao analisar os dados de escrita de crianças moçambicanas que têm o Português como L2, propõe uma escala de complexidade de padrões silábicos na aprendizagem da ortografia portuguesa: CV < CCV < CVC. Como se observa na proposta de Vicente (*op. cit.*), a rima ramificada é considerada mais complexa que o *onset*, pois nela estão concentrados mais erros ortográficos.

Para Ferreira e Zamudio [2008] (2013), que analisam dados de crianças que têm o Espanhol como língua materna, o padrão de complexidade na escrita seria CV < CVC < CCV. Porém, há de se considerar que as autoras incluem no escopo de análise da rima ramificada apenas os segmentos /r/ e /l/, o que pode influenciar os resultados no sentido de haver mais erros do que o verificado para a rima ramificada.

Pachalski e Miranda (2018a, 2019a) também apresentam algumas pistas a esse respeito, em estudos exploratórios no qual analisam dados de escrita espontânea e controlada extraídos do BATALE e produzidos por crianças brasileiras de 1º e 3º ano do Ensino Fundamental, chegando ao resultado de que não há diferença significativa entre erros envolvendo rima e *onset* ramificados, em ambos os anos escolares analisados. Esse resultado impediria, à princípio, sistematizar com precisão um ordenamento de aquisição de estruturas silábicas complexas na escrita, tal como existe para a aquisição fonológica. Para as autoras, é possível, talvez, que se estabeleça uma diferenciação temporal mais abrangente entre a aquisição escrita de sílabas ramificadas e não-ramificadas, mas não entre os constituintes específicos rima ramificada e *onset* complexo.

É necessário considerar, também, as particularidades dos dados de rima ramificada com nasal e lateral para estabelecer essa comparação. Se, como propõe Miranda (2019), as crianças não interpretam tais consoantes como licenciadas para ocupar a posição de coda no PB, na etapa inicial de desenvolvimento da escrita, não seria o caso de retirar esses dados do cômputo geral de erros envolvendo a coda e designar-lhes um estatuto mais específico? Dessa forma, não se veria,

provavelmente, uma queda considerável desse tipo de erro, tornando os erros em *onset* complexo mais numerosos em relação àqueles em rima ramificada?

Além disso, a questão sobre a cronologia de aquisição das estruturas silábicas exige que seja respeitada outra importante diferença relativamente à qualidade dos processos de aquisição fonológica e escrita, qual seja, o tipo de tarefa cognitiva imposta à criança quando da aquisição de cada sistema, aspecto detalhado nas subseções 2.1.1 e 2.1.3. Na elaboração da gramática fonológica, a criança precisa construir e incorporar representações fonológicas que se organizam hierárquica e contrastivamente. Nos termos de Karmiloff-Smith (1994), o formato de armazenamento desse conhecimento é inicialmente implícito, isto é, não está aberto ao sistema cognitivo para o acesso consciente do sujeito, motivo pelo qual o conhecimento não pode ser manipulado nem analisado. Trata-se, assim, de uma tarefa epilinguística. Na aquisição da escrita, o sujeito volta-se para a gramática fonológica que *já foi* construída, particularmente à camada fonêmica, e a ela relaciona representações gráficas correspondentes. O acesso consciente torna-se, neste caso, necessário à decomposição das unidades fonológicas até as unidades mínimas que são os fonemas. Trata-se, portanto, de uma tarefa metalinguística.

Esse tipo de diferença entre os processos de aquisição é destacado por Alves (2012), que estuda o efeito de propriedades segmentais sobre o desempenho de crianças portuguesas em tarefas de consciência segmental, de leitura e de escrita. A autora observa um efeito inverso das propriedades segmentais ao se comparar o desempenho das crianças nas tarefas testadas com o padrão do desenvolvimento fonológico do PE. Há melhor performance, por exemplo, em tarefas que envolvem fricativas do que aquelas que envolvem oclusivas, sendo que, na aquisição fonológica, as oclusivas são as primeiras a emergirem. Alves (2012, p. 415) interpreta esse tipo de assimetria, recorrente em seu estudo, considerando que “dependendo da tarefa a executar, as crianças parecem explorar de forma diferente o conhecimento de que dispõem”. Em outras palavras, as propriedades fonológicas – segmentais, no caso de Alves (2012) – têm efeito em ambos os processos de aquisição, mas esse efeito não se verifica e se interpreta da mesma maneira.

Tendo isso em vista, a tentativa de traçar uma cronologia de aquisição escrita das estruturas silábicas deveria ser relativizada, uma vez que, no caso da escrita, não se está construindo representações, ou efetivamente adquirindo-as, mas, antes, reprocessando representações já existentes e as tornando explícitas em termos de

sua constituição interna. A baixa incidência dos erros envolvendo o *onset* complexo em relação ao universo de contextos em que a estrutura aparece, por exemplo, conforme resultado apresentado por Miranda (2019) e exposto na subseção 2.4.2.2, pode significar mais um indício que corrobora tal raciocínio. O fato é que, na escrita, os erros são esporádicos, podendo um mesmo texto apresentar grafias corretas e incorretas para a mesma estrutura em palavras diferentes, enquanto, na fonologia, os erros (ou a não-produção) são mais sistemáticos enquanto a estrutura não está disponível em termos representacionais.

É nesse contexto que se fazem pertinentes as considerações de Abaurre (2001), em texto que aborda a temática da grafia das sílabas complexas na escrita inicial, sem diferenciar entre rima ramificada e *onset* complexo, explorando um pequeno conjunto de textos espontâneos. Para a autora, os erros envolvendo esse tipo de estrutura se devem, possivelmente, “a uma tentativa de análise da estrutura silábica conduzida pela própria criança”, que a coloca em dúvida sobre “o número de segmentos que devem ser representados, bem como a posição que devem ocupar na estrutura das sílabas” (ABAURRE, 2001, p. 69 e 75). Portanto, a principal questão que se coloca aos estudos que se dedicam a abordar essa temática seria, segundo a autora,

o que podem indicar esses dados sobre as representações fonológicas já construídas ou em construção pelas crianças nessa fase de desenvolvimento? Seria razoável afirmar que antes de entrarem em contato com a escrita alfabética os falantes de uma língua não analisam ainda as sílabas em segmentos, percebendo-as e produzindo-as holisticamente? (ABAURRE, 2001, p. 75).

Para Ferreira e Zamudio [2008] (2013), a motivação dos erros envolvendo sílabas CCV e CVC não estaria associada necessariamente ao nível da análise dos fonemas, mas, antes, à ênfase dada ao molde CV no contexto das práticas de alfabetização, especialmente nos chamados *métodos silábicos*³⁸. Para as autoras, isso explicaria tanto as omissões quanto as epênteses, bastante recorrentes nos dados que analisam:

[...] É como se se tratasse de preservar o modelo CV ou, melhor, de incorporar a palavra ouvida ao modelo CV. Esse modelo CV pode ser derivado das práticas escolares, mas ser assimilado pelas crianças como um modelo geral que poderia servir para resolver qualquer problema de escrita. [...] Talvez seja mais adequado pensar que impuseram um modelo gráfico já

³⁸ Para uma revisão e detalhada discussão a respeito dos métodos de alfabetização no contexto latino-americano, ver o capítulo introdutório de Ferreira e Teberosky [1984] (1999) e Soares (2016).

conhecido a um tipo de sílaba para a qual não tem modelo gráfico disponível (FERREIRO; ZAMUDIO, [2008] 2013, p. 238-239).

No modelo de Seymour (1997), revisado na subseção 2.1.3, haveria espaço para considerar a plausibilidade dessa ideia, uma vez que ele prevê a existência de *templates* gráficos que atuam como um tipo de conhecimento a que se recorre no processo de desenvolvimento ortográfico. No entanto, a hipótese de Ferreiro e Zamudio [2008] (2013), apesar de bastante interessante e inovadora, é difícil de sustentar, pois está ligada a uma variável cujo controle é igualmente difícil: as práticas escolares.

Enfim, além (ou junto) dessas questões, outras ainda podem ser suscitadas, com vistas a explorar de forma mais aprofundada o que está em jogo na grafia de sílabas complexas durante a aquisição da escrita. Tais questões são expostas e contextualizadas no capítulo que segue, o qual apresenta a metodologia da pesquisa.

3 Metodologia

Este capítulo tem a finalidade de descrever a metodologia da pesquisa, caracterizando-a quanto à sua natureza e esclarecendo sobre seus principais aspectos – os dados, os instrumentos de coleta, os sujeitos, os procedimentos de análise adotados, as variáveis e as perguntas de investigação. A organização do texto segue esse roteiro.

3.1 Caracterização da pesquisa

A pesquisa científica, segundo Marconi e Lakatos (2017), pode ser caracterizada de diversas maneiras, a depender do domínio de referência de que se esteja tratando. Assim, é muito comum classificar as pesquisas, por exemplo, segundo sua finalidade, incorrendo-se aos rótulos *básica* e *aplicada*; segundo seus objetivos, gerando a caracterização *exploratória*, *descritiva* e *explicativa*, ou, ainda, segundo seus procedimentos de análise e métodos, aspecto que dá origem às nomenclaturas *qualitativa* e *quantitativa*. Apesar de tais definições serem elucidativas no que se refere à organização do trabalho científico, convém lembrar, conforme salientam as autoras, que dificilmente uma pesquisa se encerra exclusivamente em uma ou outra natureza. Não raro ocorrem hibridismos, de forma que uma pesquisa básica pode alcançar, também, fins de aplicação; uma pesquisa descritiva pode vir a tornar-se explicativa, e, o que é mais comum, uma pesquisa pode adotar uma abordagem quanti-qualitativa.

Assim, pela perspectiva apresentada por Marconi e Lakatos (2017), é possível caracterizar a presente pesquisa como *básica*, uma vez que tem por finalidade preponderante o avanço do conhecimento teórico nas áreas de aquisição da linguagem (falada e escrita) e de teoria fonológica, sem a preocupação com a aplicabilidade imediata. Também é possível caracterizá-la como *descritiva*, visto que, partindo de estudos exploratórios já realizados, objetiva delinear as características de

um fenômeno observado na amostra de uma população (a grafia de sílabas complexas por crianças em fase de aquisição da escrita) e identificar relações entre variáveis que atuam sobre esse fenômeno, sem comprometer-se em, necessariamente, explicar a sua natureza. Por último, a esta pesquisa pode-se atribuir caráter quanti-qualitativo, posto que tanto se preocupa com a quantificação de dados e variáveis, quanto procura observar a qualidade do fenômeno estudado, objetivo que se serve, em boa medida, dos próprios dados quantitativos.

Por conta de tal caracterização, esta pesquisa, no que se refere ao levantamento de dados, utiliza da técnica de *documentação direta*, ou seja, coleta os dados no próprio local onde os fenômenos ocorrem (MARCONI; LAKATOS, 2017), aspecto que é detalhado na subseção que segue.

3.2 Os dados, os instrumentos de coleta e os sujeitos

Os dados analisados nesta pesquisa, isto é, as grafias de *onset* e de rima ramificados produzidas por crianças em fase inicial de escolarização, foram extraídos do Banco de Textos de Aquisição da Linguagem Escrita (BATALE).

O BATALE é um dos importantes resultados do projeto de pesquisa *Aquisição e Desenvolvimento da Escrita: ortografia e acentuação*, iniciado em 2001, o qual também originou as atividades do Grupo de Estudos sobre Aquisição da Linguagem Escrita (GEALE), grupo de pesquisa ao qual este trabalho está vinculado. Atualmente, o Banco possui 7423 textos espontâneos que estão organizados em 9 estratos. Essa distribuição e caracterização pode ser visualizada na Figura 17, conforme segue:

Banco de Textos de Aquisição da Linguagem Escrita (BATALE)							
2001-2019							
Estrato	Ano	Material coletado	Tipo de coleta	Série/Ano	Local da coleta	Escola	Número de textos
Estrato 1	2001-2004	textos espontâneos	transversal e longitudinal	1ª a 4ª série Ensino Fundamental	Pelotas/RS/Brasil	Pública e Particular	2024
Estrato 2	2008	textos espontâneos	transversal	1º a 4º ano Ensino Básico	Lisboa/Portugal	Pública	96
Estrato 3	2009	textos espontâneos	transversal	1ª a 4ª série Ensino Fundamental	Pelotas/RS/Brasil	Pública	507
Estrato 4	2009	textos espontâneos	transversal	1º a 3º ano Ensino Fundamental	Porto/Portugal	Pública	783
Estrato 5	2009	textos espontâneos e ditados de imagens	transversal	1º a 4º ano Ensino Fundamental	Pelotas/RS/Brasil	Pública	155
Estrato 6	2009	textos espontâneos	longitudinal	EJA	Pelotas/RS/Brasil	Pública	98
Estrato 7	2013-2015	textos espontâneos e ditados balanceados	transversal e longitudinal	1º a 5º ano Ensino Fundamental	Pelotas/RS/Brasil	Pública	1765
Estrato 8	2014-2015	textos espontâneos e ditados balanceados	transversal e longitudinal	1º a 3º ano Ensino Fundamental	Porto Alegre/RS/Brasil	Pública	1724
Estrato 9	2019	textos espontâneos	transversal	1º a 7º ano Ensino Básico	Maputo/Moçambique	Pública	271
Total de textos: 7423							

Figura 17 – Quadro com a configuração do Banco de Textos de Aquisição da Linguagem Escrita (BATALE).

Fonte: acervo do GEALE.

Por meio da Figura 17, nota-se que a maioria dos textos que compõem o BATALE foram produzidos por crianças brasileiras, residentes na cidade de Pelotas/RS e estudantes do 1º Ciclo do Ensino Fundamental da rede pública de ensino. Ademais, a maioria das coletas são de caráter transversal, com pequenas porções longitudinais. Também integram o BATALE textos produzidos por crianças portuguesas e moçambicanas, textos produzidos por adultos brasileiros alunos de programas de Educação de Jovens e Adultos (EJA), textos produzidos por crianças brasileiras residentes na cidade de Porto Alegre/RS, textos produzidos por crianças brasileiras estudantes de escola particular, e outros materiais, além de textos, em menor proporção, como ditados balanceados e de imagem.

Os textos do BATALE são obtidos por meio de coletas realizadas por integrantes do GEALE, preparados para este fim, e que obedecem a uma mesma metodologia, qual seja: ida às escolas parceiras para aplicação de oficinas de produção textual, as quais cumprem a seguinte rotina: *aquecimento/motivação*

(atividade de estimulação e de oferecimento de subsídios para a escrita do texto por meio de imagens ou diálogos), *produção textual* (atividade de escrita individual) e *socialização* (atividade de leitura das produções realizadas). A aplicação das oficinas ocorre sempre de forma coletiva, isto é, com todos os alunos que integram uma turma de determinado ano escolar.

Essa estruturação das oficinas não é fortuita: visa intencionalmente a coleta de textos *espontâneos*, os quais potencialmente fornecem os dados mais adequados para o tipo de pesquisa desenvolvida no GEALE, que busca compreender as hipóteses das crianças sobre a estrutura da língua e o sistema ortográfico.

O texto espontâneo, assim, é entendido como o instrumento mais apropriado, pois a interferência externa é minimizada, particularmente a do professor, do pesquisador ou mesmo da própria estrutura escolar. A preocupação da criança não incide, portanto, em atividade que reproduz os roteiros e rotinas escolares, mas sim orienta sua atenção à história a ser contada, ao fato a ser descrito ou ao argumento a ser construído, o que cria espaço para que expresse o modo como ela mesma concebe o objeto de aprendizagem, no caso, o sistema de escrita alfabética e também o sistema ortográfico. O texto espontâneo é, dessa maneira, de acordo com Abaurre (2011), um espaço de experimentação, de testagem de hipóteses e de solução de problemas. Nesse sentido, por exemplo, uma orientação importante que se procura assegurar durante as coletas de texto ocorre quando, não raro, as crianças perguntam como devem ser escritas determinadas palavras, do ponto de vista da norma ortográfica. O adulto que está presente no momento é orientado a não fornecer uma resposta explícita ou objetiva à pergunta; no máximo, devolverá outra pergunta que retorne a responsabilidade da decisão pela grafia das palavras à criança, tais como “escreve do jeito que achas que seja correto” ou “como *tu* achas que se escreve essa palavra?”.

Uma das oficinas de produção textual que serviu à coleta de parte dos textos analisados nesta pesquisa, por exemplo, chama-se “Animais Fantásticos”. Nela, como atividade de aquecimento/motivação, as pesquisadoras apresentaram imagens de seres mitológicos (centauro, minotauro, grifo etc.), e, por meio delas, interagiram com as crianças perguntando se conheciam os seres, se sabiam a quais histórias os seres estavam associados, se eles realmente existiram, por que eles foram criados, quais eram as suas características físicas e simbólicas. Por meio do diálogo, foram acrescentadas informações as quais, por vezes, as crianças não conheciam. Ainda como

parte do momento de aquecimento, foi apresentado um livro com folhas partidas em dois pedaços que, conforme vai sendo folheado, mostra animais a partir da combinação de duas metades, como se pode visualizar na Figura 18:



Figura 18 – Imagens do livro “Animais Fantásticos.
Fonte: acervo do GEALE; desenhos de Isabel Ramil.

Após apresentar o livro às crianças, relacionando-o com os seres mitológicos anteriormente exibidos, se explorou com elas as possibilidades geradas pela combinação das diferentes metades do livro, ao mesmo tempo em que se solicitava a descrição dos animais gerados: qual o nome, o que gosta de comer, qual sua habilidade principal, qual seu ‘poder’, onde vive, que par escolheria para se acasalar.

Do aquecimento surge, então, a proposta de produção de um texto expositivo: as pesquisadoras dizem às crianças que não se sabe se os seres mitológicos realmente existiram, mas o fato é que alguém os criou, ao menos em sua imaginação. Assim, é feito o convite às crianças para a criação de novos animais, a partir das combinações que o livro apresentado oferece. Elas foram orientadas a primeiro escolher a combinação que desejassem e depois escrever uma carta para alguém que eles imaginassem ser capaz de criar animais. As instruções para a elaboração da carta foram as seguintes: as crianças deveriam *descrever* as características do seu animal, ou seja, o que ele come, onde ele vive, o que ele faz, com quem ele se acasala

e outras coisas além que elas poderiam desejar. Porém, lembramos a eles que, como estavam pedindo um favor, deveriam *convencer* o “criador”, para quem estavam escrevendo, dos motivos pelos quais esse animal deveria ser criado e deveria existir. As crianças receberam as mesmas imagens utilizadas pelas mediadoras, porém em tamanho menor (Figura 19), de modo que pudessem manipulá-las e criar seu próprio livreto, auxiliando na produção do texto:



Figura 19 – Imagens do livro “Animais Fantásticos” em tamanho menor, manipuladas pelas crianças participantes da coleta.

Fonte: acervo do GEALE.

O encerramento da coleta se deu por meio da socialização dos textos produzidos com a turma, momento em que as crianças puderam realizar a leitura em voz alta daquilo que escreveram.

Em momento posterior ao da coleta, os textos são inicialmente armazenados em pastas catalográficas, próprias para a conservação adequada do material. Para sua melhor organização e catalogação, são criados códigos para cada texto dispostos em planilhas de Excel, os quais têm, por exemplo, a seguinte estruturação: 00004_080624_M_07_01_01_2015_TA_MFO_4A_A. Os campos correspondem, respectivamente, às informações *número do aluno*, *idade do aluno (ano, mês, dia)*, *sexo do aluno*, *número do estrato* a que pertence o texto do aluno, *número da coleta* a que pertence o texto do aluno, *número do texto* (para quando as crianças escrevem mais de um texto em uma mesma coleta), *ano* em que foi feita a coleta do texto, *tipo*

de texto produzido, *sigla da escola* à qual pertence o aluno, *série/ano escolar* do aluno e *turma* do aluno.

Depois de preparados os códigos, os textos são digitados em formato *Word*, respeitando-se a grafia utilizada pelas crianças e a troca de linhas, sem qualquer tipo de correção. Os textos são, também, digitalizados e salvos em formato PDF. Além disso, a identificação das crianças participantes das coletas é suprimida no processo de tratamento dos textos, de forma a preservar sua identidade bem como impedir qualquer forma de exposição, respeitando-se, assim, as normas éticas da pesquisa científica. Um exemplo de codificação, digitação e de digitalização de um texto resultante da coleta realizada a partir da oficina “Animais Fantásticos” pode ser visualizado na Figura 20:

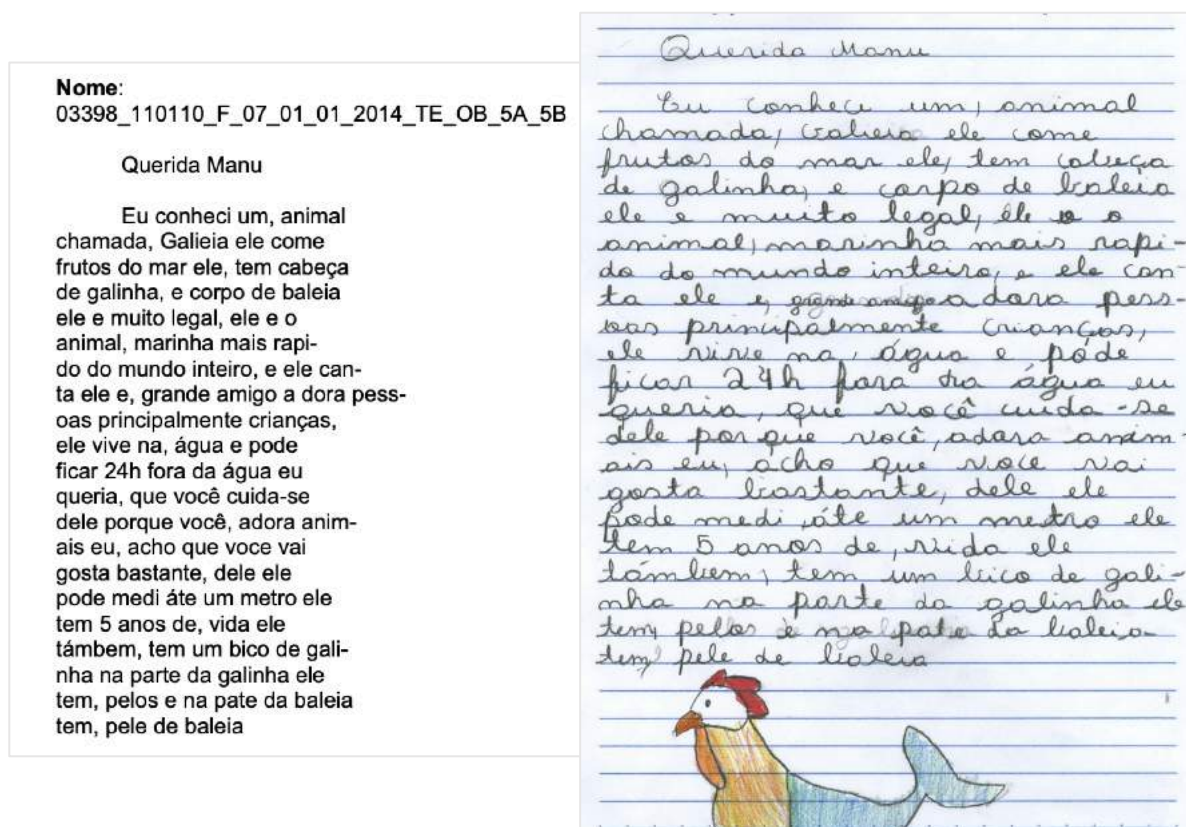


Figura 20 – Exemplo dos processos de tratamento de textos do BATALE: digitalização, codificação e digitação.
Fonte: BATALE (MIRANDA, 2001).

A Tabela 1 apresenta uma caracterização geral dos textos do Estrato 7:

Tabela 1 – Caracterização geral do Estrato 7.

Ano letivo	Escola	Anos escolares	Nº de textos coletados	Oficina de produção textual	Tipos de texto coletado	Ditado balanceado
2013	CMP	1º a 5º	582	“Os cães” (versão 1)	Narrativo, descritivo e argumentativo	Não
2014	CMP MFO OB	1º a 5º	215 98 188	“Animais Fantásticos” (versão 1)	Narrativo e expositivo	Sim
2015	CMP MFO OB	1º a 5º	104 26 54	“Os cães” (versão 2)	Expositivo	Sim

Fonte: elaboração própria.

No presente estudo, são analisados transversalmente dados extraídos textos espontâneos do Estrato 7, especificamente aqueles que foram coletados nos anos de 2014 e 2015, e produzidos por crianças estudantes de 1º, 2º, 3º, 4º e 5º anos de uma escola pública da rede municipal da cidade de Pelotas/RS (sigla CMP conforme Tabela 1).³⁹

A escolha por analisar os dados do Estrato 7 se deve ao fato de ser esta uma camada mais recente do BATALE que, por esse motivo, ainda não foi amplamente utilizada pelos estudos desenvolvidos no GEALE. A escolha pelos textos produzidos em 2014 e em 2015 se deve ao fato de se tratar de uma amostra que contemplou a coleta de ditados balanceados, os quais, apesar de não analisados nesta pesquisa, podem vir a ser utilizados em estudos futuros, de forma cotejada e complementar aos textos espontâneos. Além disso, mais três critérios foram considerados para a seleção do material analisado: os textos deveriam ser produzidos por crianças que não possuísem, à época da coleta, laudo médico que atestasse algum tipo de desenvolvimento neuroatípico; o texto deveria apresentar grafia alfabética⁴⁰, e os textos deveriam ser produzidos por crianças que cursavam pela primeira vez o respectivo ano escolar, respeitando o fluxo previsto para sua idade. Os textos que

³⁹ Em se tratando de uma pesquisa transversal, os sujeitos foram considerados de forma descontínua na computação dos dados. Por exemplo, uma criança que participou da coleta em 2014 cursando o 1º ano e que também participou em 2015, já no 2º ano, recebeu duas entradas diferentes nas planilhas, como se fossem dois sujeitos distintos.

⁴⁰ Conforme níveis de desenvolvimento da escrita propostos por Ferreiro e Teberosky [1984] (1999) – ver seção 2.3 deste trabalho.

respeitaram todos os critérios descritos, dos 319 textos coletados na escola CMP em 2014 e 2015, foram 280, sendo, portanto, este número efetivamente analisado.

Convém, ainda, um comentário sobre os anos escolares analisados. Os primeiros cinco anos do Ensino Fundamental de Nove Anos compreendem o 1º Ciclo desse período da Educação Básica e se consideram adequados à pesquisa cujo objeto é aquisição da escrita e da ortografia especialmente por representarem marcos importantes da aprendizagem formal desses objetos de conhecimento: o 1º ano é o período que marca o ingresso das crianças no Ciclo de Alfabetização assim como no Ensino Fundamental, simbolizando, assim, a sua introdução ao ensino formal do sistema de escrita; o 2º ano contempla o aprofundamento de aspectos que envolvem o sistema de escrita, motivo pelo qual se espera maior volume de grafias alfabéticas nos textos; o 3º ano compreende o término do Ciclo de Alfabetização, do que decorre a expectativa de que aspectos básicos envolvendo o processo de aquisição da escrita estejam consolidados e conhecimentos relativos ao funcionamento do sistema ortográfico já tenham sido introduzidos; o 4º ano compreende o aprofundamento de aspectos que envolvem o sistema ortográfico, e, por fim, o 5º ano representa a conclusão do 1º Ciclo do Ensino Fundamental, razão pela qual se espera que conhecimentos sobre o sistema ortográfico estejam consolidados⁴¹.

3.3 Os procedimentos de análise

Os procedimentos de análise adotados nesta pesquisa podem ser divididos em dois eixos que interagem entre si: o qualitativo e o quantitativo.

Do ponto de vista qualitativo, o primeiro procedimento adotado é a análise dos textos e a separação dos dados neles encontrados em duas grandes categorias: acertos e erros em sílabas complexas. Estes últimos são analisados a partir de categorias que auxiliam a descrevê-los segundo características que delimitam sua natureza. As categorias utilizadas nesta pesquisa têm amparo nos estudos desenvolvidos pelo GEALE desde 2001, através dos quais se chegou à proposta de categorização já apresentada na Figura 4 (subseção 2.1.3).

⁴¹ Conferir Lei de Diretrizes e Bases da Educação Nacional – LDB (BRASIL, 1996), Plano Nacional de Educação – PNE (BRASIL, 2015) e Direitos de Aprendizagem no Ciclo de Alfabetização (BRASIL, 2012).

Considerando o diagrama exposto na Figura 4 e o objeto de estudo desta pesquisa, isto é, a grafia de *onset* e rima ramificados, os erros (orto)gráficos encontrados nas grafias analisadas podem compreender duas naturezas principais e distintas: fonológica e ortográfica. Cada uma delas possui desdobramentos que possibilitam uma descrição mais precisa do dado analisado.

A primeira categoria considera a complexidade representacional de determinadas unidades melódicas e prosódicas da gramática fonológica, que pode provocar dúvida relativamente à sua representação gráfica, ainda que as relações fonografêmicas sejam biunívocas na maioria dos casos. É o caso de *onset* e rima ramificados, que têm representação fonológica complexa no nível da *sílaba*, mas que, à exceção de /N/ e /S/ em coda, têm representação ortográfica simples. Neste estudo, 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 escrita (ver subseções 2.4.2.1 e 2.4.2.2) e nos principais tipos de fenômenos verificáveis na fonologia infantil quando se trata da aquisição de sílabas complexas (ver subseções 2.4.1.1 e 2.4.1.2): omissão, metátese, epêntese e substituição⁴².

Os conhecimentos que a segunda categoria leva em conta dizem respeito às regras do sistema ortográfico, definidas contextual ou arbitrariamente – casos em que as relações entre fonemas e grafemas são múltiplas, o que, nas sílabas complexas, verifica-se apenas para as consoantes /N/ e /S/ em coda, como há pouco observado. A primeira tem sua representação gráfica regulada por uma regra contextual em posição de coda medial: antes de <p> e , usa-se <m>; para os demais contextos gráficos, usa-se <n>. A segunda consoante, a mais complexa do ponto de vista ortográfico para o PB, tem sua grafia regulada por uma regra arbitrária na posição de coda medial: é preciso lançar da estratégia mnemônica para optar corretamente entre <x> ou <s>. Neste caso, ainda há uma restrição contextual para o grafema <x> que se interpõe à arbitrariedade da regra: o <x> só é utilizado se houver antes dele a vogal coronal /e/, como em ‘**ext**ensão’. No entanto, ainda se trata de uma regra arbitrária em última instância, pois o grafema <s> também pode aparecer diante da vogal /e/, como em ‘**est**acionar’. Assim, considerando as estruturas em foco neste estudo, erros de

⁴² Conforme detalhado mais adiante, na subseção 4.4, a substituição possui características bastante heterogêneas na escrita, especialmente da rima ramificada. De modo geral, neste trabalho, a substituição remete a *alterações segmentais*, o que pode incluir tanto vogais quanto consoantes.

natureza ortográfica apenas são verificáveis na rima ramificada, e, mais especificamente, na grafia de /N/ e /S/.

As Figuras 21 e 22, que seguem, sistematizam as considerações feitas sobre as relações fonográfêmicas envolvendo as sílabas complexas e as categorias de erros:

Onset complexo*					Rima ramificada*⁴³			
C2	/r/	<r>	'prato'	biunívoca	/r/	<r>	'porta'	biunívoca
	/l/	<l>	'bloco'	biunívoca	/l/	<l>	'balde'	biunívoca ⁴⁴
C1	/p/	<p>	'empresa'; 'duplo'	biunívoca	/S/	<s>; <x>	'estudo'; 'extra'	múltipla arbitrária (apenas diante de /e/)
	/b/		'brinco'; 'blindar'	biunívoca	/N/	<n>; <m>	'conto'; 'compasso'	múltipla contextual
	/t/	<t>	'atraso'; 'atleta'	biunívoca				
	/k/	<c>	'crime'; 'esclarecer'	biunívoca*				
	/g/	<g>	'grama'; 'globo'	biunívoca*				
	/f/	<f>	'frente'; 'flor'	biunívoca				
	v	<v>	'palavra'	biunívoca				

*Os constituintes silábicos são analisados nos moldes VC, CVC, CCV e CCVC.

**A posição de *onset* complexo impõe aos fonemas /k/ e /g/ uma representação biunívoca, isto é, nesta posição /k/ só pode ser representado por <c> e /g/ apenas por <g>. Em outros contextos esses fonemas têm relações múltiplas, podendo ser também representados pelos grafemas <qu> e <gu>, respectivamente.

Figura 21 – Quadro com as relações fonográfêmicas das sílabas complexas no PB.

Fonte: elaboração própria, com base em Miranda (2020).

⁴³ Considerando que a “fonologia do sistema ortográfico” reflete a fonologia adulta, são considerados para a análise dos dados os quatro segmentos que, conforme Bisol (1999), ocupam a coda no PB. No capítulo seguinte, que analisa e discute os dados da pesquisa, a assimetria entre a fonologia do sistema ortográfico/adulta e a infantil são devidamente delineadas e distinguidas, sempre que necessário.

⁴⁴ Considerando, mais uma vez, que a “fonologia do sistema ortográfico” reflete a fonologia adulta, postula-se como fonema a lateral em coda (cf. BISOL, 1999). Por esse ponto de vista, é possível propor, portanto, que há uma relação biunívoca entre fonema e grafema neste caso. Contudo, há de se considerar, também, que a realização fonética da lateral como glide [w] cria uma ambiguidade na rima ramificada medial, pois formas homófonas são geradas entre o que seriam ditongos fonológicos, como em 'cauda', e a sequência vogal + lateral, como em 'calda'. Este fenômeno, ainda que não comprometa a natureza das relações fonográfêmicas (uma vez que estas se estabelecem entre *fonemas* e grafemas, e não entre *fonos* e grafemas), pode provocar dúvida ao aprendiz no momento de representar graficamente a lateral.

Categoria/natureza		Fonológica	Ortográfica
Descrição		Relacionada a algum tipo de complexidade representacional fonológica ou, ainda, à fala	Relacionada à complexidade ortográfica estabelecida pelas regras contextuais e arbitrárias do sistema ortográfico
Tipos de erros	OC	Omissão, metátese, epêntese e substituição	–
	RR		Não observância de regras: contextuais e/ou arbitrárias
Exemplos de erros	OC	‘bincar’ – ‘brincar’; ‘garde’ – ‘grande’; ‘espofoção’ – ‘explosão’; ‘refreção’ – ‘reflexo’	–
	RR	‘decobriram’ – ‘descobriram’; ‘enprotante’ – ‘importante’; ‘geniti’ – ‘gente’; ‘autura’ – ‘altura’	‘completo’ – ‘completo’; ‘estinação’ – ‘extinção’

Figura 22 – Quadro das categorias de erros (orto)gráficos utilizadas neste estudo (OC = *onset* complexo; RR = rima ramificada).

Fonte: adaptado de Miranda (2020).

Além das duas categorias enfatizadas, neste trabalho é considerada uma terceira categoria de natureza residual, denominada *outros*. Ela reúne, por um lado, erros que não correspondem, por excelência, aos conhecimentos fonológico e ortográfico tais como descritos acima, podendo ir ao encontro, em parte, daquilo que Miranda (2020) propõe como erros *fonográficos*⁴⁵; por outro, reúne erros que podem ser interpretados como de natureza fonológica, mas que não necessariamente se relacionam com o problema da ramificação em si ou, ainda, que não se enquadram nas subcategorias estabelecidas para os erros fonológicos, as quais correspondem a fenômenos típicos da fonologia sobretudo quando se trata das estruturas silábicas complexas (omissão, metátese, epêntese e substituição).

A Figura 23 exibe um quadro descritivo da categoria de erros residuais *outros*, esclarecendo sobre suas principais características e tipos de erros a ela relacionados:

⁴⁵ Lembrando que a categoria de erros *fonográficos* não é considerada neste estudo em virtude de ainda ser uma proposta embrionária que necessita de testes para melhor compreensão de sua natureza e de sua validade para erros do tipo que envolvem estruturas como as sílabas complexas.

Categoria/natureza		<i>Outros (residual)</i>
Descrição		Não relacionada à complexidade representacional, tanto fonológica quanto ortográfica, que envolve a ramificação silábica
Tipos de erros	OC	Alteração na C1; traçado
	RR	Hipersegmentação; omissão/inserção de constituinte ou sílaba
Exemplos de erros	OC	'trasinha' – 'pracinha'; 'ghede' – 'grande'
	RR	'com verçar' – 'conversar'; 'avntura' – 'aventura'

Figura 23 – Quadro descritivo da categoria de erros residuais *outros*.

Fonte: elaboração própria.

A classificação dos dados, tanto com base nas categorias mais abrangentes *acertos* e *erros* quanto nas categorias que definem os *tipos de erro*, prepara-os para a sua quantificação, etapa associada ao eixo quantitativo dos procedimentos de análise, que se subdivide em outras três etapas: o levantamento da frequência absoluta e relativa dos dados de acordo com as categorias estabelecidas, a estatística descritiva e a estatística inferencial.

A primeira etapa das três é realizada manualmente, por meio da verificação dos textos e da contagem dos dados segundo as categorias. A tabulação e a organização desses dados contam com o auxílio do *software Microsoft Excel* (versão 16.26). Um exemplo da utilização deste *software* para os fins da primeira etapa pode ser visualizado nas Figuras 24 e 25:

Página Inicial Inserir Desenhar Layout da Página Fórmulas Dados Revisão Exibir Conte-me										
Compartilhar Comentários										
C4										
	B	C	D	E	F	G	H	I	J	K
	número	escola	ano escolar	turma	ano letivo	erro lateral	acerto lateral	contextos laterais	frequência erros laterais	erro rotica
2	02354	CMP	5	B	2014	0	0	0	#DIV/0!	0
3	01188	CMP	4	B	2014	0	0	0	#DIV/0!	0
4	01188	CMP	5	B	2015	0	0	0	#DIV/0!	0
5	02336	CMP	3	D	2014	0	1	1	0	0
6	00655	CMP	2	B	2014	0	1	1	0	0
7	01770	CMP	4	B	2015	0	0	0	#DIV/0!	0
8	02389	CMP	3	A	2015	0	0	0	#DIV/0!	0
9	01781	CMP	3	C	2014	0	1	1	0	0
10	01190	CMP	5	B	2015	0	0	0	#DIV/0!	0
11	01190	CMP	4	B	2014	0	2	2	0	0
12	01248	CMP	5	B	2014	0	0	0	#DIV/0!	0
13	02337	CMP	3	D	2014	0	0	0	#DIV/0!	0
14	02399	CMP	5	B	2015	0	0	0	#DIV/0!	0
15	00622	CMP	2	B	2014	0	0	0	#DIV/0!	0
16	01191	CMP	5	B	2015	0	0	0	#DIV/0!	0
17	02390	CMP	3	A	2015	0	0	0	#DIV/0!	0
18	01192	CMP	4	B	2014	0	0	0	#DIV/0!	0
19	01192	CMP	5	B	2015	0	0	0	#DIV/0!	0
20	01786	CMP	3	C	2014	0	0	0	#DIV/0!	0

Figura 24 – Captura de tela com a imagem de parte de uma das planilhas de *Excel* utilizadas no processo de contagem, tabulação e organização dos dados da pesquisa (variáveis relativas a acerto e erro).

Fonte: elaboração própria.

Página Inicial Inserir Desenhar Layout da Página Fórmulas Dados Revisão Exibir Conte-me										
Compartilhar Comentários										
K26										
	B	C	D	E	F	G	H	I	J	K
	número	escola	ano escolar	turma	ano letivo	grafia	palavra alvo	fonol_RR_N (1)	fonol_RR_N (2)	ortog_RR_N
1	00655	CMP	2	B	2014	comprou	comprou	4		
2	01781	CMP	3	C	2014	brinquedos	brinquedos	1		
3	01248	CMP	5	B	2014	pese	pense	1		
4	02399	CMP	5	B	2015	viagem	viagem	1		
5	01193	CMP	4	B	2014	entrar	entrar			1
6	02349	CMP	5	B	2015	em loquizar	enloquecer			1
7	02349	CMP	5	B	2015	com verçar	conversar			1
8	02349	CMP	5	B	2015	grande	grande	4		
9	00600	CMP	3	A	2015	morendo	morendo			1
10	00601	CMP	3	A	2015	grado	grande	1		
11	01791	CMP	2	B	2015	gra de	grande	1		1
12	01791	CMP	2	B	2015	tabem	também	1		
13	01791	CMP	2	B	2015	tabem	também	1		
14	01791	CMP	2	B	2015	tabem	também	1		
15	00672	CMP	4	B	2015	garde	grande	1		
16	01251	CMP	5	B	2014	entredo nante	impressionante			1
17	01196	CMP	4	B	2014	en graçado	engraçado			1
18	01196	CMP	4	B	2014	tambem	também			1
19	01196	CMP	4	B	2014	tambem	também			1
20	01196	CMP	4	B	2014	tabem	também	1		
21	01196	CMP	5	B	2015	grande	grande			1
22	01196	CMP	5	B	2015	acomelhavel	aconselhável			1

Figura 25 – Captura de tela com a imagem de parte de uma das planilhas de *Excel* utilizadas no processo de contagem, tabulação e organização dos dados da pesquisa (variáveis relativas a tipo de erro).

Fonte: elaboração própria.

As últimas duas etapas foram realizadas por meio do *software SPSS Statistics* (versão 17.0) que, alimentado pelas informações tabuladas na primeira etapa, fornece os indicadores descritivos dos dados, como médias, medianas, desvio padrão e gráficos, além de realizar testes estatísticos de correlação e de diferença, os quais

permitem verificar a validade estatística das hipóteses da pesquisa, isto é, se os resultados podem ou não ser aplicados à população que a amostra analisada representa. Antes dos testes, os dados foram submetidos à análise exploratória, que indica o tipo de teste adequado à amostra: se ela for normal e homogênea, o teste será paramétrico; se for anormal e heterogênea, será não-paramétrico. O uso do software foi orientado pelo *Manual de Análise de Dados Quantitativos com Recurso ao IBM® SPSS: saber decidir, fazer, interpretar e redigir* (MARTINS, 2011), e um profissional da área de estatística também foi consultado a fim de sanar eventuais dúvidas relacionadas a esta parte da análise. As Figuras 26 e 27 mostram exemplos de como os dados e as variáveis desta pesquisa ficaram organizados no SPSS:

base spss textos_pos_ajustes.sav [DataSet1] - SPSS Statistics Data Editor

</

Figura 26 – Captura de tela com a imagem de parte de uma das planilhas de SPSS (aba *data view*) utilizadas no processo de levantamento estatístico descritivo e inferencial (variáveis relativas a acerto e erro).

Fonte: elaboração própria.

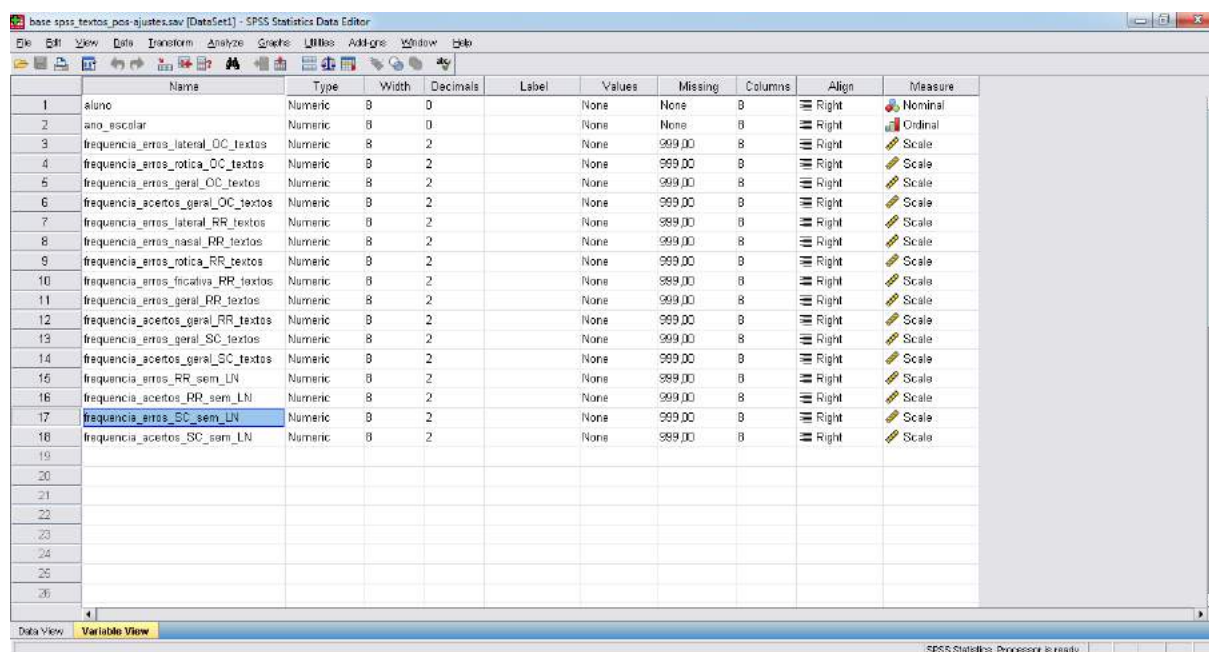


Figura 27 – Captura de tela com a imagem de parte de uma das planilhas de SPSS (aba *variable view*) utilizadas no processo de levantamento estatístico descritivo e inferencial (variáveis relativas a acerto e erro).

Fonte: elaboração própria.

Por fim, as Tabelas 2 e 3, que seguem, apresentam, respectivamente, uma relação dos testes estatísticos de diferença e de correlação utilizados em cada tipo de situação experimental deste estudo:

Tabela 2 – Testes de diferença utilizados na análise dos dados.

<i>Design</i>	Variável dependente	Distribuição das variáveis	Teste de diferença	Teste <i>Post Hoc</i> ⁴⁶	Valor de <i>p</i>	Valor de <i>p</i> marginalmente significativo
Intrassujeitos (2 momentos)	Intervalar	Anormal	<i>Wilcoxon</i>	—	< 0,05	< 0,1
Intrassujeitos (3 ou mais momentos)	Intervalar	Anormal	<i>Friedman</i>	<i>Conover</i>	< 0,05	< 0,1
Intersujeitos (2 grupos)	Intervalar	Anormal e heterogênea	<i>Mann-Whitney</i>	—	< 0,05	< 0,1
Intersujeitos (3 ou mais grupos)	Intervalar	Anormal e heterogênea	<i>Kruskal-Wallis</i>	<i>Mann-Whitney</i>	< 0,05	< 0,1

Fonte: elaboração própria.

⁴⁶ Com valores de *p* corrigidos pela correção de Bonferroni.

Tabela 3 – Testes de correlação utilizados na análise dos dados.

Variáveis	Distribuição das variáveis	Teste de correlação	Valor de p	Valor de p marginalmente significativo
Ordinal + Intervalar	Anormal	<i>Spearman</i>	< 0,05	< 0,1

Fonte: elaboração própria.

Conforme recomendações de Martins (2011) e do profissional de estatística consultado, o ponto de corte utilizado para inferência estatística foi de 0,05 (*valor de p*), mas também foram considerados os valores *marginalmente significativos*, isto é, aqueles inferiores a 0,10. Além disso, todas as variáveis da amostra apresentaram distribuição anormal e heterogênea, razão pela qual os testes utilizados foram não-paramétricos. Ainda, na consulta ao profissional de estatística, foi sugerida, em casos pontuais da análise, a realização de testes não disponíveis no SPSS (versão 17.0), sendo este o caso do teste *Post Hoc Conover*, que consta na Tabela 2. Em razão disso, nesses casos, foi utilizado o *software RStudio*. Todas essas informações são devidamente reportadas e explanadas no texto do capítulo 4.

3.4 As variáveis

Considerando os objetivos específicos da pesquisa e as perguntas da investigação que deles decorrem, as principais variáveis analisadas e subsumíveis aos testes estatísticos são apresentadas, nas subseções que seguem, conforme a abrangente classificação *dependente e independente*, acompanhadas de uma breve explicação que justifica seu controle na pesquisa.

3.4.1 Variáveis dependentes

Variáveis dependentes correspondem, em geral, ao objeto da pesquisa, e são aquelas sobre as quais se busca verificar o efeito ou influência que uma variável independente possa vir a exercer (MARCONI; LAKATOS, 2017). Neste estudo, a variável dependente é a grafia de sílabas complexas (*onset* e rima ramificados) na escrita de crianças em fase inicial de escolarização, podendo desdobrar-se em duas:

- a) *erro* na grafia de sílabas complexas: esta variável é organizada a partir do levantamento da frequência relativa de ocorrência dos erros ortográficos na grafia de *onset* e de rima ramificados encontrados nos textos analisados;
- b) *acerto* na grafia de sílabas complexas: esta variável é organizada a partir do levantamento da frequência relativa de ocorrência dos acertos na grafia em *onset* e rima ramificados encontrados nos textos analisados.

É importante notar que, na maioria das vezes, não interessa analisar erros e acertos de forma abrangente, isto é, desconsiderando se o erro ocorre em *onset* ou rima ramificados. Isso devido às diferenças que os estudos apontam relativamente ao comportamento de cada estrutura na aquisição fonológica (cf. RIBAS, 2002; MEZZOMO, 2004; Figuras 15 e 16), sobretudo ligadas a tempo de aquisição, estratégias utilizadas e condicionamento prosódico, e também na fonologia adulta (cf. CÂMARA JR., [1970] 1995; BISOL, 1999; MATEUS; ANDRADE, 2000; COSTA; FREITAS, 2001; Figura 13), especialmente no que se refere às controvérsias envolvendo o estatuto das codas.

Assim, são predominantes as perguntas de investigação que procuram observar o efeito de variáveis independentes sobre a grafia ou de *onset* complexo ou de rima ramificada, e, ainda, que buscam verificar se há diferenças significativas de erros e de acertos entre ambas estruturas. Por essa razão, as variáveis dependentes deste estudo podem ser também descritas da seguinte forma:

- c) erro na grafia de *onset* complexo;
- d) acerto na grafia de *onset* complexo;
- e) erro na grafia de rima ramificada;
- f) acerto na grafia de rima ramificada.

3.4.2 Variáveis independentes

Variáveis independentes correspondem a uma característica, condição ou causa para certo resultado, e são aquelas que podem influenciar ou não a variável dependente (MARCONI; LAKATOS, 2017). Nesta pesquisa, as variáveis independentes são as seguintes:

- a) *tipo de estrutura silábica*: no escopo mais amplo das sílabas complexas, é importante aferir se existem diferenças entre os tipos de estrutura silábica – *onset* ou rima ramificados – no que se refere à frequência de ocorrência dos erros (orto)gráficos. Assim, neste caso específico, as variáveis dependentes *erro em onset complexo* e *erro em rima ramificada* são comparadas entre si. A justificativa para essa testagem é amparada, primeiro, pelos mesmos motivos apresentados na subseção anterior para a divisão das variáveis dependentes conforme o tipo de estrutura silábica: os diferentes comportamentos que *onset* complexo e rima ramificada apresentam tanto na aquisição fonológica quanto na fonologia adulta. Com isso, se permite discutir algumas (as)simetrias que possam existir entre aquisição fonológica e aquisição da escrita⁴⁷, especialmente no que se refere à cronologia de aquisição das estruturas, aspecto revisado na subseção 2.4.2.3 (cf. SOARES, 2016; VICENTE, 2018; PACHALSKI; MIRANDA, 2018a, 2019a);
- b) *tipo de segmento*: esta variável remete à busca por aferir se os erros (orto)gráficos são mais ou menos frequentes a depender do tipo de segmento que está ligado ao constituinte ramificado (/S/, /N/, /r/ e /l/ para rima e /r/ e /l/ para o onset). O seu controle é fundamental considerando as controvérsias existentes acerca do estatuto fonológico dos segmentos em rima ramificada, particularmente no tocante às consoantes nasal e lateral. Conforme revisado na subseção 2.4.2.1, os estudos de aquisição da escrita, ao verificar o efeito desta variável, podem contribuir a elucidar essas questões controversas, conciliando assimetrias entre a fonologia infantil e a fonologia adulta (cf. MIRANDA, 2009a, 2009b, 2012, 2018, 2019), além de contribuir, também, com a compreensão sobre a motivação dos erros em sílabas complexas. Apesar de a sua importância estar mais ligada à rima ramificada, a variável *tipo de segmento* também é testada nos dados envolvendo *onset* complexo, a fim de que se certifique sobre a sua neutralidade no caso em análise;

⁴⁷ É importante lembrar que quando se compara aquisição fonológica com aquisição da escrita, **não** está implícita a ideia de que a última replicará processos da primeira. Na medida em que retoma representações fonológicas, é que pode se esperar alguma relação simétrica entre os diferentes cursos de aquisição, mas, sobretudo por demandarem tarefas cognitivas de natureza distinta, é que relações assimétricas também são esperadas.

- c) *tipo de erro (orto)gráfico*: esta variável remete às categorias de erros apresentadas na seção 3.3 (Figura 22), sobretudo aquelas mais abrangentes (fonológico e ortográfico). O objetivo principal é verificar se há diferença significativa entre a frequência de erros de natureza fonológica e erros de natureza ortográfica (o que é possível fazer apenas para a rima ramificada). Os estudos desenvolvidos pelo GEALE utilizando diferentes estratos do BATALE (cf. MIRANDA; MEDINA; SILVA, 2005; MIRANDA 2013, 2017, 2020; PACHALSKI; MIRANDA, 2019b) apontam esta variável como importante para a delimitação adequada da natureza dos erros e para a consequente compreensão de sua motivação na escrita. No conjunto geral dos dados, os mesmos estudos citados indicam uma recorrente predominância dos erros relacionados à fonologia, independente do fenômeno analisado. As subcategorias de erros dessa natureza – omissão, metátese, epêntese e substituição – também são consideradas para análise, por meio do que se deseja discutir qual tipo de erro é mais frequente e como esse resultado pode contribuir com a discussão sobre a motivação dos erros em sílabas complexas. Os estudos revisados nas subseções 2.4.2.1 e 2.4.2.2 apontam que a omissão é o tipo de erro mais frequente em ambas as estruturas complexas analisadas, mas não chegam a interpretar o resultado e discutir a sua implicação (cf. MIRANDA, 2009a, 2009b, 2012, 2018, 2019; PACHALSKI; MIRANDA, 2018a, 2019; PACHALSKI et al., 2014).
- d) *ano escolar*: esta variável compreende os anos escolares nos quais as crianças estavam matriculadas à época das coletas, ou agrupamentos que separam os anos iniciais de escolarização em dois momentos distintos de aprendizagem: o Ciclo de Alfabetização (1º, 2º e 3º anos) e os 4º e 5º anos. O seu controle permite inferir a respeito do efeito da escolarização sobre o processo de aquisição da escrita, ponto já salientado como importante por outras pesquisas que analisam dados de sujeitos falantes de outras variedades do Português (cf. SANTOS, 2013; VICENTE, 2018) e por estudos realizados no GEALE (cf. CUNHA, 2004; NEY, 2012; PACHALSKI; MIRANDA, 2018a; MIRANDA, 2020), estes últimos indicando que os erros especialmente de natureza fonológica tendem a diminuir à medida que avançam os anos escolares.

3.5 Perguntas e hipóteses de pesquisa

A decisão tanto sobre a organização da base de dados no *software* SPSS quanto sobre as variáveis que são controladas na pesquisa depende, em boa medida, das perguntas que orientam a investigação proposta, as quais decorrem, por sua vez, dos objetivos que o trabalho pretende atingir. Como forma de encerrar este capítulo que trata da metodologia do trabalho e de orientar a análise, a descrição e a discussão dos dados, esta subseção busca sistematizar as principais questões da pesquisa, acompanhadas de suas respectivas hipóteses – *de investigação, nula e alternativa* – relacionando-as com os objetivos geral e específicos, e com a pergunta e a hipótese gerais de investigação, conforme apresentados na introdução deste trabalho e nesta subseção retomados.

Considerando o problema de pesquisa geral deste trabalho – *o que motiva as crianças a omitir, inserir, substituir ou alterar a ordem dos segmentos de estruturas com onset e rima ramificados na escrita, se elas já adquiriram fonologicamente tais estruturas?* –, tem-se como objetivo geral da pesquisa *descrever e analisar a grafia de onset e de rima ramificados em textos espontâneos escritos por crianças em fase de aquisição da escrita, estabelecendo relações com a aquisição fonológica dessas estruturas e com a teoria fonológica*.

Desse objetivo abrangente decorrem os objetivos específicos, aqui retomados conforme sua apresentação no primeiro capítulo:

- a) descrever e analisar, quanti e qualitativamente, erros (orto)gráficos nos quais se verificam a alteração de estruturas de sílaba com *onset* e rima ramificados em textos espontâneos produzidos por alunos de 1º a 5º ano do Ensino Fundamental;
- b) caracterizar os erros (orto)gráficos partindo da proposta de categorização do GEALE, a fim de delimitar a sua natureza (fonológica e/ou ortográfica);
- c) verificar e comparar, quanti e qualitativamente, o efeito das variáveis *tipo de estrutura silábica, tipo de segmento, tipo de erro (orto)gráfico* e *ano escolar* sobre os erros (orto)gráficos;
- d) discutir sobre a natureza dos erros (orto)gráficos relacionados às sílabas complexas;

- e) discutir sobre simetrias e assimetrias existentes entre a aquisição da fala e da escrita das estruturas silábicas complexas;
- f) discutir sobre simetrias e assimetrias existentes entre a fonologia infantil e a fonologia adulta;
- g) propor uma formalização para as operações envolvidas na conversão das informações fonológicas *implícitas* em conhecimento *explícito*, especificamente no que se refere à representação subjacente da sílaba em crianças em fase de aquisição da escrita.

Assim como o objetivo geral se desdobra em objetivos específicos, a pergunta geral de pesquisa também se desdobra em diferentes questões de investigação. Objetivos específicos e questões de investigação se relacionam na medida em que contemplam diferentes aspectos envolvidos no fenômeno em foco, isto é, na medida em que contemplam as diferentes variáveis cuja interveniência no fenômeno analisado se deseja verificar, conforme expostas nas subseções imediatamente anteriores. Dessa forma, objetivos e questões orientam em detalhe a execução da metodologia do trabalho, sendo que as questões de investigação auxiliam a cumprir os objetivos traçados.

A Tabela 4 procura evidenciar como essa relação se estabelece neste trabalho, associando as principais questões de investigação⁴⁸ que decorrem do problema geral de pesquisa com suas respectivas hipóteses de investigação, nula e alternativa⁴⁹, e com os objetivos específicos:

⁴⁸ A Tabela 4 apresenta as *principais* questões e hipóteses de investigação. Aspectos menores e que exigem maior grau de detalhamento têm espaço na seção de análise do dados (capítulo 4). Além disso, cumpre assinalar que, ao longo da análise, surgem outras questões pertinentes à pesquisa, as quais são devidamente reportadas.

⁴⁹ Conforme Martins (2011), as hipóteses de *investigação* decorrem de revisão teórica e empírica, e ajudam a definir o(s) objetivo(s) e as variáveis da pesquisa. Já as hipóteses *nula* e *alternativa* estão relacionadas a procedimentos estatísticos inferenciais e às tomadas de decisão que deles decorrem. As hipóteses de investigação foram elaboradas com base no referencial teórico exposto no capítulo 2 e que é retomado no capítulo 4 à medida que cada questão é explorada e respondida.

Tabela 4 – Principais questões e hipóteses de pesquisa.

	Questão de investigação	Hipótese de investigação	Hipótese nula (H ₀)	Hipótese alternativa (H ₁)	Objetivos específicos relacionados
Bloco I	1 Existe diferença entre a frequência de erros e de acertos em sílabas complexas? ⁵⁰	Os acertos são mais numerosos em relação aos erros tanto em <i>onset</i> como em rima ramificados. ⁵¹	Não há diferenças entre erros e acertos em sílabas complexas.	Há diferenças entre erros e acertos em sílabas complexas.	(a), (e), (g)
	2 Existe diferença entre <i>onset</i> e rima ramificados no que se refere à frequência de erros?	Erros em <i>onset</i> e rima ramificados equiparam-se quanto à sua frequência. ⁵²	Há diferença entre <i>onset</i> complexo e coda no que se refere à frequência de erros ortográficos.	Não há diferença entre <i>onset</i> complexo e coda no que se refere à frequência de erros ortográficos.	(a), (c), (e), (g)
Bloco II	3 O tipo de segmento influencia na maior ou menor incidência de erros em sílabas complexas?	O tipo de segmento influencia na maior incidência de erros na rima ramificada, onde há mais erros envolvendo as consoantes nasal e lateral em relação às consoantes fricativa e rótica. ⁵³	Não há diferenças, para a rima, entre a frequência de erros considerando o tipo de segmento envolvido; para o <i>onset</i> , há diferenças.	Há diferenças, para a rima, entre a frequência de erros considerando o tipo de segmento envolvido; para o <i>onset</i> , não há diferenças.	(a), (c), (e), (f), (g)
Bloco III	4 O ano escolar influencia na frequência de ocorrência de erros ortográficos em sílabas complexas?	Os erros ortográficos diminuem à medida que avançam os anos escolares, tanto em <i>onset</i> quanto em rima ramificada. ⁵⁴	Não há diferença entre os anos escolares no que diz respeito à frequência de erros ortográficos.	Há diferença entre os anos escolares no que diz respeito à frequência de erros ortográficos.	(a), (c), (e), (g)

⁵⁰ É importante lembrar que as análises são feitas separadamente para *onset* e rima ramificados, conforme exposto na subseção 3.4.1 e como esclarece a redação das hipóteses de investigação na tabela.

⁵¹ Cf. Miranda (2020); Ney (2012); Cunha (2004); Adamoli (2012).

⁵² Cf. Pachalski; Miranda (2018a, 2019a).

⁵³ Cf. Miranda (2009a, 2009b, 2012, 2018, 2019).

⁵⁴ Cf. Cunha (2004); Ney (2012); Pachalski; Miranda (2018a); Miranda (2020).

Bloco IV	5	Existe diferença entre a frequência de ocorrência dos tipos de erro (orto)gráfico na rima ramificada?	Na rima ramificada, erros de natureza fonológica são mais frequentes em relação aos de natureza ortográfica. ⁵⁵	Não há diferenças entre a frequência dos diferentes tipos de erros ortográficos na rima ramificada.	Há diferenças entre a frequência dos diferentes tipos de erros ortográficos na rima ramificada.	(a), (b), (c), (e), (g)
	6	Existe diferença na frequência de erros de natureza fonológica e ortográfica se forem comparados dois grupos distintos no quesito ano escolar?	Na rima ramificada, erros de natureza fonológica são mais frequentes no Ciclo de Alfabetização, e erros de natureza ortográfica são mais frequentes no 4º e no 5º ano. ⁵⁶	Não há diferença entre a frequência das diferentes naturezas de erro se forem considerados o Ciclo de Alfabetização e 4º e 5º anos como grupos distintos.	Há diferença entre a frequência das diferentes naturezas de erro se forem considerados o Ciclo de Alfabetização e 4º e 5º anos como grupos distintos.	(a), (b), (c), (e), (g)
	7	Existe diferença entre a frequência de ocorrência dos tipos de erro de natureza fonológica em sílabas complexas?	A omissão de consoantes é mais numerosa em relação aos demais tipos de erro de natureza fonológica, tanto em <i>onset</i> como em rima ramificados. ⁵⁷	Não há diferença entre os tipos de erro de natureza fonológica no que se refere à sua frequência de ocorrência, tanto em <i>onset</i> quanto em rima ramificados.	Há diferença entre os tipos de erro de natureza fonológica no que se refere à sua frequência de ocorrência, tanto em <i>onset</i> quanto em rima ramificados.	(a), (b), (e), (g)

Fonte: elaboração própria.

Como é possível notar por meio da Tabela 2, alguns objetivos específicos estão associados a todas as questões de investigação, particularmente os objetivos (a) e (g). Isso ocorre em virtude de eles estarem relacionados a aspectos metodológicos mais abrangentes do trabalho, como descrever e analisar os erros ortográficos – o caso do objetivo (a) –, ou a aspectos teóricos mais abrangentes – o caso do objetivo (g). Dessa maneira, todas as questões propostas auxiliam, em alguma medida, no cumprimento de alguma faceta de tais objetivos, como o capítulo 4 torna claro. Além disso, convém reparar que as questões estão divididas em 4 blocos, organizados conforme o grau de abrangência das questões (de maior a menor) e por afinidade em termos das variáveis envolvidas. Essa organização orienta não somente as rodadas de análise estatística, mas a própria estrutura do capítulo seguinte.

⁵⁵ Cf. Miranda (2013, 2017, 2020); Pachalski; Miranda (2019).

⁵⁶ Cf. Miranda (2013, 2017, 2020); Seymour (1997).

⁵⁷ Miranda (2009a, 2009b, 2012, 2018, 2019); Pachalski e Miranda (2018a, 2019); Pachalski et al. (2014).

4 Descrição, análise e discussão dos resultados

Neste capítulo, são descritos e discutidos os resultados obtidos a partir dos procedimentos de análise de dados adotados, conforme expostos no capítulo 3. A descrição e discussão dos resultados está organizada em 4 subseções de acordo com as perguntas de investigação presentes na Tabela 4 (ver subseção 3.5). A última subseção do presente capítulo tem por objetivos sistematizar os resultados apresentados nas subseções precedentes e discutir as implicações desses mesmos resultados para a construção de uma resposta ao problema geral da pesquisa e para a checagem da hipótese geral da pesquisa.

4.1 As diferenças gerais entre acertos e erros (orto)gráficos envolvendo a grafia de sílabas complexas: questão 1

A primeira questão de investigação, presente na Tabela 4, se preocupa com a possibilidade de existência de uma diferença significativa entre a frequência de ocorrência de acertos e de erros (orto)gráficos envolvendo a grafia das sílabas complexas, tanto em *onset* como em rima ramificados. Tal questão é motivada pela busca de uma caracterização geral da amostra de dados analisada e, principalmente, pela busca de verificação da hipótese sustentada por estudos realizados no GEALE (cf. CUNHA, 2004; NEY, 2012; ADAMOLI, 2012; MIRANDA, 2020): no curso de aquisição da escrita ortográfica, os acertos são sempre mais numerosos que os erros.

A análise estatística realizada decorrente da questão 1 gerou os resultados descritivos e inferenciais apresentados nas Tabelas 5 e 6. A fim de verificar se é significativa a diferença observada entre as médias de acertos e erros, foi realizado o teste não-paramétrico *Wilcoxon*, uma vez que o *design* da amostra é intrassujeitos e as variáveis apresentam distribuição anormal. O resultado dos testes é expresso pelos valores de *Z* e de *p*, reportados nas tabelas:

Tabela 5 – Acertos e erros na grafia de *onset* complexo.

	Acertos (n = 145) ⁵⁸	Erros (n = 145)
Média (%)	94,60	5,40
Desvio padrão	19,72	19,72
Frequência absoluta	536/562	26/562
Z	-10,95***	

*** $p < 0,001$

Tabela 6 – Acertos e erros na grafia da rima ramificada.

	Acertos (n = 184)	Erros (n = 184)
Média (%)	89,11	10,90
Desvio padrão	23,80	23,80 ⁵⁹
Frequência absoluta	1428/1536	108/1536
Z	-10,66***	

*** $p < 0,001$

Por meio dos valores descritivos presentes nas Tabelas 5 e 6, é possível perceber, em ambas as estruturas silábicas complexas analisadas, um alto índice de acertos, com médias que giram em torno de 90%, em contraste com o pequeno índice de erros, com valores abaixo de 11%. Foram computados 562 registros de *onset* complexo, dos quais 536 foram grafados corretamente e 26 incorretamente, e 1536 registros de rima ramificada, dos quais 1428 foram grafados corretamente e 108 incorretamente. Os testes *Wilcoxon*, por sua vez, confirmaram a significância estatística das diferenças entre as médias de erros e acertos, permitindo reter as respectivas hipóteses alternativas (H_1) e rejeitar as respectivas hipóteses nulas (H_0) (ver Tabela 2), sendo $Z = -10,95$ e $p = 0,000$ para *onset* complexo, e $Z = -10,66$ e $p = 0,000$ para rima ramificada.

⁵⁸ O valor de n equivale ao número de sujeitos que ofereceram dados para computação. Isso significa dizer que, no caso do *onset* complexo, 145 sujeitos produziram pelo menos uma vez essa estrutura em algum dos textos que escreveram (certa ou errada); no caso da rima ramificada, 184 sujeitos produziram pelo menos uma vez essa estrutura em algum dos textos que escreveram (certa ou errada).

⁵⁹ Chamará a atenção do leitor ao longo do trabalho os elevados valores de desvio padrão registrados nas tabelas. Trata-se, de fato, de uma característica de toda a amostra analisada, e reflete um dado comum nos estudos de aquisição da escrita: a distribuição heterogênea das amostras, ou seja, uma distribuição pouco equilibrada das médias entre os sujeitos. No caso apresentado pela Tabela 6, por exemplo, a maioria dos 184 sujeitos que produziram rimas ramificadas apresentou uma média de 100% de acertos (ou 0% de erros), e alguns poucos sujeitos tiveram médias diferentes dessas, com valores variando entre 0% e 100%. Tais características significam que os dados estão concentrados em um polo da amostra (em 0% ou 100%) e alguns poucos esparsamente distribuídos, configurando-a como heterogênea. As implicações desse tipo de peculiaridade amostral não são amplamente discutidas neste trabalho, mas recebem um comentário no final do texto, na subseção 4.5.

Apesar de os erros chamarem à atenção pelo seu caráter heteróclito e por serem potencialmente reveladores da natureza do conhecimento que está sendo mobilizado pela criança durante a aquisição da escrita (MIRANDA, 2010), a grande marca desse período do desenvolvimento não são eles – é o que se pode inferir ao menos no que se refere à grafia das sílabas complexas, considerando os resultados obtidos com a amostra de dados analisada. Antes, ao contrário daquilo que o senso comum possivelmente consideraria, os acertos figuram em expressivo destaque numérico, apontando para o sucesso que as crianças têm, sob instrução escolar, na aprendizagem do funcionamento do sistema de escrita alfabética e das regras que regulam as relações entre fonemas e grafemas. Tal resultado também parece indicar que a estrutura silábica tem papel limitado no grau de complexidade que envolve a aquisição da escrita do PB, apesar de ser capaz, sim, de provocar dúvidas ao aprendiz – caso contrário não haveria qualquer índice de erros relacionado à grafia de sílabas ramificadas.

Além disso, a diferença significativa entre acertos e erros na grafia de ambas estruturas de sílaba analisadas aponta para uma assimetria entre a aquisição fonológica e a aquisição da escrita. Na aquisição da fonologia, quando um segmento ou uma estrutura não são considerados adquiridos ainda, eles podem sequer ser produzidos ou podem apresentar percentuais de produção conforme o alvo abaixo de 20%, especialmente em momentos iniciais do desenvolvimento. Este fato é recorrente em quaisquer dos estudos revisados e citados no capítulo 2 (cf. LAMPRECHT et al., 2004). Isso significa que os erros são mais frequentes e sistemáticos que os acertos durante a aquisição da fala, ao contrário do que aqui se observa para a aquisição da escrita, relativamente à grafia de sílabas complexas.

Possivelmente, esse fato se deva à principal característica que difere ambos os percursos de aquisição, recorrentemente salientada neste trabalho: o tipo de tarefa cognitiva demandada dos sujeitos. No período correspondente à aquisição fonológica, a demanda é por construção e efetiva aquisição de representações; na aquisição da escrita, as representações já construídas são retomadas e explicitadas ao sistema cognitivo, podendo, nesse processo, ser também reestruturadas. Na primeira, portanto, não há como produzir as estruturas conforme o alvo enquanto elas não estiverem devidamente elaboradas e posicionadas no sistema linguístico; na segunda, como é preciso “apenas” aprender a atualizar as mesmas representações

por meio de uma substância ou matéria alternativa à fala, uma considerável porção de trabalho já está realizada, fazendo com que os erros sejam bem menos frequentes.

4.2 O papel da estrutura silábica e do tipo de segmento na ocorrência de erros (orto)gráficos em sílabas complexas: questões 2 e 3

As questões de investigação seguintes à questão 1 exploram o caráter revelador dos erros (orto)gráficos acerca da natureza dos conhecimentos mobilizados pelas crianças que estão em processo de aquisição da escrita, particularmente no que se refere às estruturas silábicas complexas. As questões do Bloco II e que dão origem a esta subseção, particularmente, têm em comum a busca por verificar o efeito das variáveis *tipo de estrutura silábica* e *tipo de segmento* sobre a ocorrência de erros (orto)gráficos em sílabas complexas.

A questão 2 procura verificar a hipótese de que não há diferença significativa entre a frequência de ocorrência de erros (orto)gráficos em *onset* e rima ramificados – hipótese sustentada pelos estudos de Pachalski e Miranda (2018a, 2019a), revisados na subseção 2.4.2.3 deste trabalho. Como consta na referida subseção, trata-se de uma questão importante porque os resultados alimentam discussões acerca das simetrias e das assimetrias entre aquisição fonológica e escrita de sílabas complexas, especialmente no que se refere à cronologia de aquisição dessas estruturas.

A Tabela 7 apresenta os resultados obtidos a partir da análise estatística realizada, descritiva e inferencial, a fim de responder à questão 2. Como forma de verificar a diferença entre as médias das frequências de erros, foi utilizado o teste não-paramétrico *Wilcoxon*, uma vez que o *design* da amostra é intrassujeitos e as variáveis apresentam distribuição anormal. O resultado do teste é expresso pelos valores de *Z* e de *p*, reportados na última linha tabela:

Tabela 7 – Erros (orto)gráficos por tipo de estrutura silábica.

	Rima ramificada (n = 184)	Onset complexo (n = 145)
Média (%)	10,90	5,40
Desvio padrão	23,80	19,72
Frequência absoluta	108/1536	26/562
Z	-3,29**	

***p* < 0,01

A Tabela 7 revela que a média de erros na rima ramificada é maior que a média em *onset* complexo, com valores de 10,90% (108 erros de 1536 ocorrências no total) e 5,40% (26 erros de 562 ocorrências no total), respectivamente. Além disso, a tabela informa que a diferença constatada é significativa, sendo $Z = -3,29$ e $p = 0,001$, o que permite reter a hipótese nula (H_0) e rejeitar a hipótese alternativa (H_1). Assim, ao contrário do que se esperava, tendo por base o que apontam Pachalski e Miranda (2018a, 2019a), as frequências de erros observadas nas duas estruturas silábicas analisadas não se equiparam. É um resultado semelhante ao de Vicente (2018), que verifica mais erros na grafia da rima ramificada do que em *onset* complexo, ao analisar a escrita de crianças moçambicanas que tem o Português como L2 (ver subseções 2.4.2.1 e 2.4.2.2).

Com esse resultado, constata-se outro tipo de assimetria entre os processos de aquisição fonológica e escrita das sílabas complexas – não a mesma apontada por Pachalski e Miranda (*op. cit.*) – uma vez que, na escrita, a rima ramificada parece ser objeto de maior dúvida e, portanto, de mais complexidade representacional, enquanto na aquisição fonológica é o *onset* complexo a estrutura que mais demora a surgir e a estabilizar, evidenciando sua natureza representacional mais complexa em relação à rima ramificada (cf. LAMPRECHT et al., 2004; Figura 15 deste trabalho).

Essa assimetria direciona o olhar investigativo para a rima ramificada, também como uma forma de entender o que a diferença entre aquisição da fonologia e aquisição da escrita pode evidenciar sobre o fenômeno em foco. Assim, por que esta estrutura silábica concentra mais erros? Os estudos de Miranda (2009a, 2009b, 2012, 2018, 2019), revisados na subseção 2.4.2.1, dão pistas a esse respeito. Segundo a autora, para as crianças em fase de aquisição da escrita, existe um particular conflito entre a fonologia e a ortografia das consoantes lateral e nasal em coda, razão pela qual, também, a maioria dos erros ortográficos na rima ramificada concentra-se nesses segmentos. Isto é, a fonologia infantil, relativamente à estruturação silábica, não consideraria as consoantes lateral e nasal como ligadas à coda, mas sim como glide em núcleo ramificado e como vogal nasal em núcleo, respectivamente.

Essa discussão, por sua vez, conduz à questão de número 3, cuja hipótese, para a rima ramificada, está fundamentada nas considerações de Miranda (*op. cit.*). Na amostra de dados analisada neste trabalho, a natureza dos segmentos envolvidos também demonstra exercer influência na maior ou menor incidência de erros em

sílabas complexas, particularmente no que tange à rima ramificada? Os segmentos lateral e nasal de fato concentram mais erros (orto)gráficos em relação às consoantes fricativa e rótica?

As Tabelas 8 e 9 exibem, respectivamente, os resultados obtidos por meio das análises estatísticas descritiva e inferencial. Como forma de verificar as diferenças entre as médias das frequências de erros envolvendo cada segmento, foi realizado o teste não-paramétrico *Friedman* e, em seguida, o teste *post hoc Conover*⁶⁰ com correção de *Bonferroni*⁶¹ para checagem das diferenças entre cada par de segmentos, sendo o *design* da amostra intrassujeitos e as tendo as variáveis distribuição anormal:

Tabela 8 – Erros (orto)gráficos por tipo de segmento (rima ramificada).

	Lateral (n = 46)	Nasal (n = 160)	Fricativa (n = 136)	Rótica (n = 105)
Média (%)	18,15	13,02	3,30	2,10
Desvio padrão	34,68	27,00	15,40	11,57
Frequência absoluta	14/85	78/838	11/386	5/227
χ^2 (3)⁶²	10,98*			

* $p < 0,05$

Tabela 9 – Significância dos testes *Conover* para verificação das diferenças entre as médias de frequência de erros envolvendo cada segmento em rima ramificada.

	Lateral vs. Nasal	Lateral vs. Rótica	Lateral vs. Fricativa	Nasal vs. Rótica	Nasal vs. Fricativa	Rótica vs. Fricativa
Significância das diferenças	ns ⁶³	*	*	*	*	ns

* $p < 0,001$

⁶⁰ Teste rodado no software *RStudio* por não estar disponível na versão de SPSS utilizada, com orientação de profissional da estatística.

⁶¹ Segundo Martins (2011), a *Correção de Bonferroni* é um procedimento utilizado no contexto de testes não-paramétricos e consiste na divisão do “valor de ponto de corte convencional relativo à significância estatística, $p = 0,05$, pelo número de testes adicionais que temos de executar na sequência de um teste de diferenças significativo que comparou três ou mais grupos [ou momentos]” (MARTINS, 2011, p. 167, interpolações nossas). No caso em questão, é preciso dividir o valor padrão de p (0,05) por 6, valor este equivalente aos 6 testes *Conover* realizados. O valor de significância, assim, se torna mais exigente, baixando para 0,0083 o ponto de corte para que se possa considerar um resultado significativo.

⁶² Por ser um teste utilizado para amostras pareadas e multivariadas, o teste *Friedman* comparou apenas 36 casos válidos do conjunto total da amostra. Em outras palavras, o teste apenas considerou os alunos que apresentaram médias para os 4 segmentos, ou seja, os alunos que não possuem qualquer valor de *missing* nas variáveis em jogo. As médias apresentadas na Tabela 8 refletem as médias reais, considerando todos os casos válidos para cada variável (representados pelos valores de n na tabela).

⁶³ ns = não significativo ($p > 0,05$).

Por meio da Tabela 8, observa-se que a maior média de frequência de erros está na grafia da consoante lateral, com 18,15% (14 erros de 85 ocorrências totais de rima com lateral). Em seguida, com 13,02% de erros em média, está a grafia da nasal (78 erros de 838 ocorrências totais de rima com nasal); a grafia da fricativa figura em terceiro lugar, com 3,30% de erros (11 erros de 386 ocorrências totais de rima com fricativa), e a grafia da rótica, com 2,10% de erros, está na última posição (5 erros de 227 ocorrências totais de rima com rótica). Tais resultados, ainda que descritivos, vão ao encontro daquilo que Miranda (2009a, 2009b, 2012, 2018, 2019) também encontra nos dados que analisa.

O teste Friedman apontou diferença marginalmente significativa entre as médias de erros em cada segmento da rima ramificada, sendo $\chi^2(3) = 10,98$ e $p = 0,011$. Assim, a hipótese nula (H_0) pode ser parcialmente rejeitada e a hipótese alternativa (H_1), que corresponde em termos teóricos à hipótese de Miranda (*op. cit.*) e à hipótese de investigação deste trabalho, é parcialmente retida. Utiliza-se o termo *parcialmente* tendo em vista o valor de p marginalmente significativo.

A fim de averiguar entre quais tipos de segmentos estão localizadas as diferenças, são apresentados os dados da Tabela 9. Ela informa os resultados obtidos com o teste *Conover*, sendo significativas as diferenças entre as médias de erros na grafia da lateral e da rótica ($p = 0,000$), da lateral e da fricativa ($p = 0,000$), da nasal e da rótica ($p = 0,000$) e da nasal e da fricativa ($p = 0,000$).

Esses resultados permitem diferenciar dois grupos de segmentos no que se refere à grafia da rima ramificada: de um lado, as consoantes nasal e lateral, e de outro as consoantes fricativa e rótica. No primeiro grupo estão os segmentos com as maiores médias de erros e que, sendo testadas entre si, não apresentam diferença significativa. No segundo grupo, estão os segmentos com as menores médias de erros e que, também sendo testadas entre si, não apresentam diferença significativa. É na testagem de diferença entre cada grupo que os contrastes significativos aparecem – lateral e nasal não têm diferença entre si, mas ambas apresentam diferença se colocadas em oposição à fricativa e à rótica. Da mesma forma, fricativa e rótica não têm diferença entre si, mas apresentam diferença se contrastadas com lateral e nasal.

Dessa maneira, é possível sustentar a hipótese de Miranda (2009a, 2009b, 2012, 2018, 2019) de que, na aquisição da escrita, as crianças apresentam

comportamentos distintos em relação à grafia das consoantes lateral e nasal comparada à grafia das consoantes fricativa e rótica. E apresentam tal comportamento em virtude do conflito entre sua fonologia, que licencia como codas legítimas apenas as consoantes fricativa e rótica, e a ortografia, que elege grafemas consonantais em final de sílaba para representar aquilo que para a criança são glide em núcleo ramificado (no caso da lateral) e vogal nasal em núcleo (no caso da nasal).

Dado esse comportamento distinto observado nos dados das crianças relativamente às consoantes lateral e nasal, e tendo em vista a hipótese interpretativa, aqui sustentada, que Miranda (*op. cit.*) oferece para explicar esse comportamento, outra questão acaba por emergir, como uma forma de buscar reforçar ainda mais o resultado anterior: existiria diferença entre a frequência de erros na rima ramificada ao se considerar como coda apenas os segmentos fricativo e rótico, e a frequência de erros na rima ramificada ao se considerar como coda os segmentos lateral, nasal, fricativo e rótico? A hipótese, neste caso, é a de que sim, existe diferença.

A Tabela 10, a seguir, apresenta os resultados decorrentes das análises descritiva e inferencial realizadas a fim de responder essa questão. A significância da diferença entre as médias de frequência de erros em cada uma das modalidades de rima ramificada foi averiguada, novamente, por meio do teste não-paramétrico *Wilcoxon* uma vez que o *design* da amostra é intrassujeitos e as variáveis apresentam distribuição anormal. O resultado do teste é expresso pelos valores de *Z* e de *p*, reportados na última linha da tabela:

Tabela 10 – Erros (orto)gráficos na rima ramificada considerando os diferentes estatutos de coda.

	Coda com consoantes lateral e nasal (n = 184)	Coda sem consoantes lateral e nasal (n = 159)
Média (%)	10,90	2,55
Desvio padrão	23,80	10,74
Frequência absoluta	108/1536	16/613
Z	-4,49***	

*** $p < 0,001$

A Tabela 10 evidencia que, ao se considerar como coda apenas as consoantes fricativa e rótica, a média de erros na rima ramificada cai de 10,90% (108 erros do total de 1536 ocorrências de rima ramificada) para 2,55% (16 erros do total de 613

ocorrências de rima ramificada). Além disso, mostra que a diferença entre esses valores é significativa, sendo $Z = -4,49$ e $p = 0,000$, o que permite rejeitar a hipótese nula (H_0), reter a hipótese alternativa (H_1) e fazer valer a hipótese de investigação, qual seja, de que a frequência de erros em rima ramificada é significativamente menos numerosa quando se consideram como codas legítimas apenas os segmentos fricativo e rótico. Tal resultado é mais um indicativo que reforça a interpretação distinta que dão as crianças à estruturação silábica do Português Brasileiro, particularmente no que se refere ao estatuto da coda, em contraste com a fonologia adulta – considerando-se para esta as propostas de Câmara Jr. (1970) e de Bisol (1999).

Ainda como consequência do tratamento peculiar dado pelas crianças aos segmentos em coda, outra pergunta, não prevista na Tabela 4, torna-se relevante: se forem consideradas como coda apenas as consoantes fricativa e rótica, continuará existindo diferença entre a frequência de erros (orto)gráficos em *onset* complexo e rima ramificada? Mais uma vez, para a execução da análise, o teste *Wilcoxon* foi escolhido em razão das características de anormalidade e de *design* intrassujeitos das variáveis analisadas. Os resultados descritivos e inferenciais encontram-se na Tabela 11:

Tabela 11 – Erros (orto)gráficos em textos espontâneos por tipo de estrutura silábica (codas somente fricativa e rótica).

	Onset complexo (n = 145)	Rima ramificada (n = 159)
Média (%)	5,40	2,55
Desvio padrão	19,72	10,74
Frequência absoluta	26/562	16/613
Z	-1,60	

Para ler e interpretar os resultados expressos pela Tabela 11, é importante ter em vista as informações contidas na Tabela 7, que diz respeito às médias dos erros (orto)gráficos em textos espontâneos por tipo de estrutura silábica, considerando-se como portando estatuto de coda as consoantes lateral, nasal, fricativa e rótica. Neste caso, os erros envolvendo a rima ramificada contabilizam 10,90% de média. A Tabela 11, onde os valores para a rima consideram apenas os erros em codas fricativa e rótica, informa que a média de erros, neste caso, é de 2,55%. Esse resultado descritivo

é antagônico ao observado na Tabela 7, uma vez que, na Tabela 11, a maior média de frequência de erros está no *onset* complexo.

O teste *Wilcoxon*, no entanto, revelou não haver diferença significativa entre a frequência de erros envolvendo a grafia de *onset* complexo e a grafia da rima ramificada (considerando-se apenas as codas fricativa e rótica), sendo $Z = -1,60$ e $p = 0,110$, o que permite reter a hipótese alternativa (H_1) e rejeitar a hipótese nula (H_0) da pergunta original (questão 2).

Neste caso, a hipótese sustentada por Pachalski e Miranda (2018a, 2019a) é corroborada, ainda que as autoras, em sua análise, tenham computado erros nos quatro segmentos em coda, e não apenas em dois. Tal como especulado na subseção 2.4.2.3, esse resultado impediria sistematizar um ordenamento de aquisição de estruturas silábicas na escrita da mesma forma como existe na aquisição fonológica (ver Figura 16). O que se pode dizer é que, na escrita, tanto *onset* quanto rima ramificados, apresentam algum tipo de complexidade representacional, mais acentuado que as sílabas CV, mas não há uma diferença significativa entre ambas as estruturas silábicas no que concerne a essa complexidade, fato evidenciado pela equiparação das médias de erros (orto)gráficos envolvendo as duas estruturas. Esse resultado parece indicar, também, que a ramificação silábica não é tanto um problema para as crianças quanto é o conflito gerado pelas diferenças fonológicas do sistema infantil e do sistema ortográfico (equivalente ao adulto). A subseção 4.4 retoma tais considerações com uma discussão mais detalhada, a partir da questão de número 8.

Resta, ainda, dentro do objetivo desta subseção, averiguar os resultados atinentes ao *onset* complexo, especialmente no que se refere à questão de número 3, os quais foram deixados à parte em virtude do fluxo de discussão que demandava a rima ramificada. Neste caso, a questão 3 procura verificar se os dois grupos de *onset* complexo, determinados pelo tipo da segunda consoante – se lateral ou rótica –, diferem entre si no que concerne à frequência de erros (orto)gráficos.

A Tabela 10 apresenta os resultados obtidos por meio da análise realizada. Como forma de verificar a diferença entre as médias das frequências de erros, foi utilizado o teste não-paramétrico *Wilcoxon*, uma vez que o *design* da amostra é intrassujeitos e as variáveis apresentam distribuição anormal. O resultado do teste é expresso pelo valor de Z , reportado na última linha da tabela:

Tabela 12 – Erros (orto)gráficos em textos espontâneos por tipo de segmento (*onset* complexo).

	Rótica (n = 144)	Lateral (n = 23)
Média (%)	5,53	4,34
Desvio padrão	19,80	20,85
Frequência absoluta	24/528	2/34
Z	-0,53	

Os valores exibidos pela Tabela 12 para cada um dos grupos consonantais do *onset* complexo são bastante próximos, com uma diferença de apenas 1% entre cada um. Como se presumia com as hipóteses de investigação e alternativa, o teste *Wilcoxon* revelou não ser significativa essa diferença entre a frequência de erros envolvendo a grafia, em *onset* complexo, do grupo com a consoante lateral e do grupo com a consoante rótica, sendo $Z = -0,53$ e $p = 0,655$, o que permite rejeitar a hipótese nula (H_0) e reter a hipótese alternativa (H_1).

Esse resultado difere daquele obtido por Vicente (2018), em estudo que estabelece o mesmo tipo de comparação (ver subseção 2.4.2.2). Analisando dados de crianças moçambicanas, o autor encontra maior incidência de erros na grafia do grupo com a lateral, e atribui esse resultado a uma assimetria existente no Português Europeu entre a fonologia infantil e a fonologia adulta, em relação aos grupos consonantais em *onset* complexo (cf. VELOSO, 2003, 2006).

Os resultados exibidos pela Tabela 12 também revelam uma simetria entre os processos de aquisição fonológica e escrita do *onset* complexo. Conforme revisado na subseção 2.4.1.2, Ribas (2002) observa que, na aquisição fonológica, não há diferenças importantes no tempo de aquisição ao serem comparados os dois grupos consonantais, com líquida lateral ou não lateral. Os resultados apresentados na Tabela 12 permitem considerar algo semelhante para a escrita: não parece haver maior ou menor complexidade representacional – que geraria maior incidência de erros (orto)gráficos – em função do tipo de segmento que configura a segunda posição do *onset* complexo.

4.3 O papel da escolarização na ocorrência de erros (orto)gráficos em sílabas complexas: questão 4

A questão de número 4 busca verificar a hipótese de que há efeito da escolarização sobre a ocorrência de erros (orto)gráficos em sílabas complexas, ou seja, o avanço das crianças ao longo dos anos escolares – particularmente no 1º Ciclo do Ensino Fundamental – faria com que os erros (orto)gráficos diminuíssem. Essa hipótese encontra amparo em estudos do GEALE que versam sobre diferentes tipos de erro de natureza fonológica. Esses estudos evidenciam que os erros analisados diminuem à medida que avançam os anos escolares (cf. CUNHA, 2004; NEY, 2012; PACHALSKI; MIRANDA, 2018a; MIRANDA, 2019a).

Considerando a maneira como a questão está formulada e a natureza das variáveis envolvidas⁶⁴, dois tipos de testes estatísticos podem ser executados: um teste de correlação e um teste de diferença.

Com o primeiro tipo, “o objetivo é avaliar se duas (ou mais) variáveis têm alguma *relação* entre si” (MARTINS, 2011, p. 99, grifos da autora), relação que pode ser negativa ou positiva. Assim, no caso em análise, é possível investigar se o ano escolar tem relação negativa ou positiva com a frequência de ocorrência de erros (orto)gráficos. A hipótese levantada é a de que a relação seria negativa, isto é, à medida que aumenta o ano escolar diminuem os erros (orto)gráficos, tanto em *onset* quanto em rima ramificados.

Com o segundo tipo de teste estatístico, o objetivo é explorar se há *diferenças* entre grupos independentes no que se refere a um aspecto ou fenômeno específico, no caso de *design* intersujeitos (MARTINS, 2011). Portanto, neste caso, é possível verificar se há diferenças entre os anos escolares (que configuram os grupos independentes) no que diz respeito à frequência de ocorrência de erros (orto)gráficos em sílabas complexas. A hipótese correspondente é a de que há diferenças significativas entre os anos escolares em se tratando da frequência de ocorrência de erros (orto)gráficos, tanto em *onset* quanto em rima ramificados.

Assim sendo, os dois tipos de teste foram rodados para ambos os tipos de estrutura silábica analisados.

⁶⁴ A variável *ano escolar* é ordinal e a variável *frequência de ocorrência de erro (orto)gráfico* é intervalar.

Na Tabela 13, a seguir, encontram-se os resultados, relativos aos dados de *onset* complexo, decorrentes dos testes de correlação e de diferença utilizados, acompanhados da descrição das médias de frequência de erro (orto)gráfico por ano escolar. No caso do teste de correlação, por uma variável ser de natureza ordinal e outra de natureza intervalar, foi utilizado o teste *Spearman*, cujo valor pode ser verificado na penúltima coluna da tabela. No caso do teste de diferença, foi utilizado o teste *Kruskal-Wallis* em virtude das características de anormalidade e de heterogeneidade da amostra analisada. Seu valor pode ser verificado na última linha da tabela:

Tabela 13 – Erros (orto)gráficos na grafia de *onset* complexo conforme o ano escolar.

	1º ano (n = 8)	2º ano (n = 19)	3º ano (n = 43)	4º ano (n = 31)	5º ano (n = 44)
Média (%)	31,25	4,38	1,16	8,60	3,05
Desvio padrão	45,80	19,11	5,62	26,13	12,04
Frequência absoluta	4/11	5/48	2/136	6/142	9/225
r_s	-0,41				
$\chi^2(4)$	10,11*				

* $p < 0,05$

A Tabela 13 mostra que os erros (orto)gráficos em *onset* complexo têm uma média maior no 1º ano escolar, com 31,25% de frequência de ocorrência (4 erros de 11 ocorrências totais de *onset* complexo registradas neste ano escolar). Os valores de média diminuem consideravelmente até o 3º ano, etapa que concentra apenas 1,16% de erros (1 erro de 136 ocorrências totais de *onset* complexo registradas neste ano escolar). No entanto, no 4º ano, contrariando a tendência dos próprios dados até então, há um aumento percentual que chega a 8,60% (6 erros de 142 ocorrências totais de *onset* complexo registradas neste ano escolar), mas que volta a diminuir no 5º ano, etapa que conta com uma média de 3,05% (9 erros de 225 ocorrências totais de *onset* complexo registradas neste ano escolar).

O teste *Spearman* revelou não haver uma associação negativa significativa entre o ano escolar e os erros na grafia de *onset* complexo, sendo $r_s = -0,41$ e $p = 0,363$, o que permite reter a hipótese nula (H_0) e rejeitar a hipótese alternativa (H_1). Com isso, entende-se, com base na amostra estudada, que os erros (orto)gráficos em *onset* complexo não *necessariamente* diminuem à medida que avançam os anos escolares.

O teste *Kruskal-Wallis*, por sua vez, revelou haver uma diferença significativa entre os anos escolares e os erros na grafia de *onset* complexo, sendo $\chi^2(4) = 10,11$ e $p = 0,039$, valores que permitem rejeitar a hipótese nula (H_0) e reter a hipótese alternativa (H_1). Porém, resta saber entre quais anos escolares, especificamente, existem as diferenças. A Tabela 14, que segue, apresenta essa informação, exibindo as significâncias obtidas em cada um dos 10 testes não-paramétricos *Mann-Whitney* (realizados com *Correção de Bonferroni*), a fim de verificar as diferenças entre cada ano escolar:

Tabela 14 – Significância dos testes *Mann-Whitney* para verificação das diferenças entre as médias de frequência de erros em *onset* complexo em cada ano escolar.

	1º ano	1º ano	1º ano	1º ano	2º ano	2º ano	2º ano	3º ano	3º ano	4º ano
	vs. 2º ano	vs. 3º ano	vs. 4º ano	vs. 5º ano	vs. 3º ano	vs. 4º ano	vs. 5º ano	vs. 4º ano	vs. 5º ano	vs. 5º ano
Significância da diferença	<i>ns</i> ⁶⁵	*	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>

* $p < 0,005$

A Tabela 14 revela que existem diferenças significativas apenas entre o 1º e o 3º ano no que se refere à frequência de erros (orto)gráficos em *onset* complexo, sendo $U = 112,50$ e $p = 0,003$.

Apesar de a diferença estar localizada em apenas 1 par dos 10 testados, trata-se de um interessante resultado, uma vez que a diferença está justamente entre dois marcos do 1º Ciclo do Ensino Fundamental, conforme observados na subseção 3.2 deste trabalho. A diferença assinalada pelo teste estatístico aponta para a conformidade que têm as crianças relativamente à aprendizagem esperada delas, acerca do sistema de escrita, em cada um dos anos escolares que compõem o 1º Ciclo – pelo menos no que concerne ao *onset* complexo.

De fato, no 1º ano, são esperados mais erros (orto)gráficos, visto que é um momento de introdução ao sistema de escrita. É também esperada uma média significativamente menor de erros no 3º ano, já que este é o ano que encerra o Ciclo de Alfabetização e consolida, portanto, a aprendizagem de conhecimentos básicos relacionados ao sistema de escrita. E, embora os testes não tenham atestado uma

⁶⁵ ns = não significativo.

diferença significativa das médias em relação ao 5º ano, há de se notar a média deste ano escolar é consideravelmente mais baixa em relação à do 1º ano.

Nesse sentido, o resultado também indica que, ao longo do processo de aquisição da escrita – o qual acompanha o avanço do processo de escolarização –, há uma clara diferença de compreensão das crianças a respeito da grafia do *onset* complexo, sendo o 1º ano o período onde são evidenciadas mais dúvidas a respeito da representação gráfica dessa estrutura silábica, o que vai ao encontro dos resultados obtidos por Santos (2013) e Vicente (2018), que analisam dados de outras variedades do Português.

Outro aspecto interessante para discussão está relacionado ao aumento inesperado de erros no 4º ano, com uma média que só perde para a do 1º ano. Apesar de nenhuma diferença estatística ser apresentada pelos testes em relação ao índice do 4º ano, o aumento na média vai ao encontro de um aspecto característico dos cursos desenvolvimentais de modo geral: a *curva em U* (STRAUSS; STAVY, 1982), explorada na fundamentação teórica deste trabalho (ver subseção 2.1).

A Figura 28 apresenta um gráfico que ilustra como as médias exibidas na Tabela 13 realizam um movimento que sinaliza para a formação de um U:

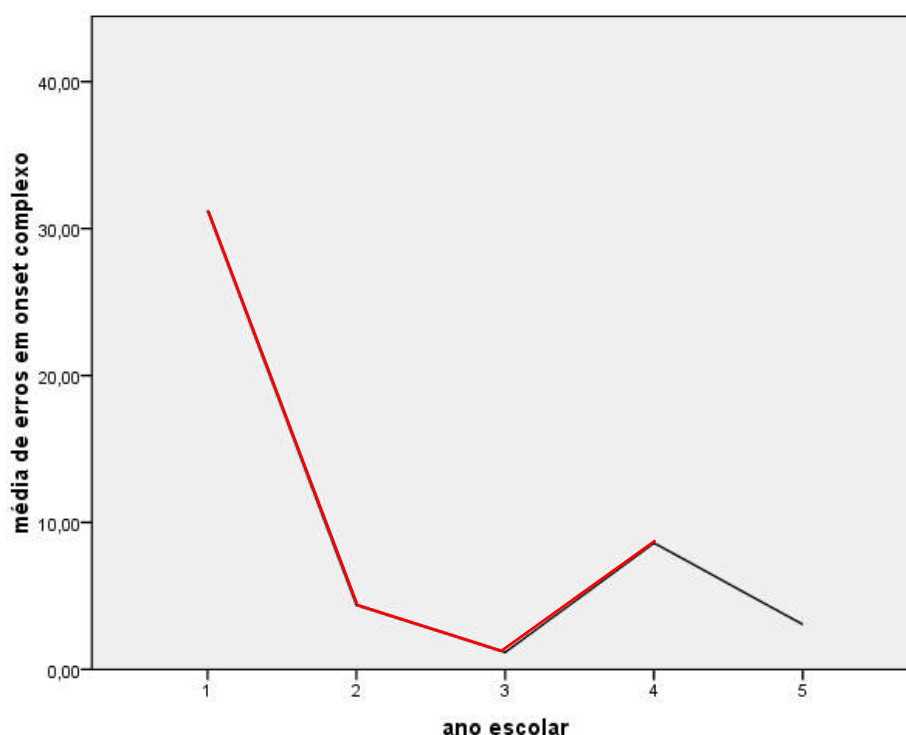


Figura 28 – Gráfico linear das médias de frequência de erros (orto)gráficos em *onset* complexo conforme cada ano escolar.

Fonte: elaboração própria (com auxílio do *software* SPSS Statistics).

O movimento que faz a linha do gráfico não chega a completar o traço da letra U, mas, ainda assim, o sugere. Para Karmiloff-Smith (1994), por exemplo, a existência de tais “erros tardios”, que dão origem à curva em U da aprendizagem, é fundamental na sustentação de sua teoria (ver subseção 2.1). Esse tipo de dado demonstra, para a autora, que o sistema cognitivo está lidando com representações de forma qualitativamente distinta de como antes lidava, ou seja, está agora manipulando e analisando as representações em suas partes componentes, o que termina por gerar novas condutas – “erradas”, do ponto de vista da forma-alvo esperada, mas precisas se forem levados em consideração os tipos de conhecimento e de processamento que estão sendo mobilizados. A curva em U também costuma revelar a entrada de uma nova informação no sistema cognitivo, informação que, ao ser processada, gera um rearranjo estrutural no sistema, o que pode desestabilizar momentaneamente um conhecimento já sedimentado. Essas são marcas do processo de explicitação do conhecimento, isto é, da progressiva abertura ao acesso consciente da informação. Como esse processo se dá no curso de aquisição da escrita, particularmente no tocante às sílabas complexas, é assunto discutido em detalhe na próxima subseção (4.4).

Até este momento, foram observadas as relações entre as variáveis ano escolar e frequência de erros (orto)gráficos em *onset* complexo. Resta ainda averiguar as relações entre ano escolar e erros (orto)gráficos na rima ramificada, considerando as 4 consoantes em coda (lateral, nasal, fricativa e rótica). Como se comportam os dados nesta estrutura silábica complexa?

A Tabela 15, que segue, exhibe os resultados, relativos aos dados de rima ramificada, decorrentes dos testes de correlação e de diferença utilizados, acompanhados da descrição das médias de frequência de erro (orto)gráfico conforme cada ano escolar. Pelos mesmos motivos relacionados aos dados de *onset* complexo, foram utilizados o teste *Spearman*, para verificação da correlação, e o teste *Kruskal-Wallis*, para verificação das diferenças. Os valores dos testes constam, respectivamente, na penúltima e na última coluna da tabela:

Tabela 15 – Erros (orto)gráficos na grafia da rima ramificada conforme o ano escolar.

	1º ano (n = 18)	2º ano (n = 34)	3º ano (n = 52)	4º ano (n = 35)	5º ano (n = 45)
Média (%)	16,66	19,31	10,27	8,06	5,11
Desvio padrão	38,34	31,83	22,61	17,95	8,78
Frequência absoluta	4/29	22/148	18/291	30/396	34/672
r_s	0,21				
$\chi^2(4)$	3,35				

A Tabela 15 mostra de forma descritiva que as médias de erros (orto)gráficos na rima ramificada diminuem conforme avançam os anos escolares: cada ano escolar apresenta uma média menor do que a média do ano precedente. Porém, essa relação não possui validade estatística, uma vez que o teste *Spearman* revelou não haver uma correlação negativa significativa entre erros na grafia da rima ramificada e o ano escolar dos sujeitos, sendo $r_s = 0,21$ e $p = 0,781$, o que permite reter a hipótese nula (H_0) e rejeitar a hipótese alternativa (H_1). Com isso, entende-se que os erros (orto)gráficos em rima ramificada não *necessariamente* diminuem à medida que avançam os anos escolares, assim como verificado para o *onset* complexo.

Da mesma forma, o teste *Kruskal-Wallis* revelou não haver diferenças significativas entre os anos escolares no que se refere à frequência de erros envolvendo a grafia da rima ramificada, sendo $\chi^2(4) = 3,35$ e $p = 0,501$, permitindo reter a hipótese nula (H_0) e rejeitar a hipótese alternativa (H_1). Tal resultado difere daquele encontrado para a grafia de *onset* complexo, estrutura para a qual são verificadas diferenças significativas entre os anos escolares no que se refere à frequência de erros (orto)gráficos.

O que se pode dizer, portanto, é que nesta amostra específica de dados, sem estender à população que ela representa, os erros (orto)gráficos tendem a diminuir conforme o ano escolar aumenta, o que cria graficamente uma linha descendente que liga as médias de erros, como ilustra a Figura 29:

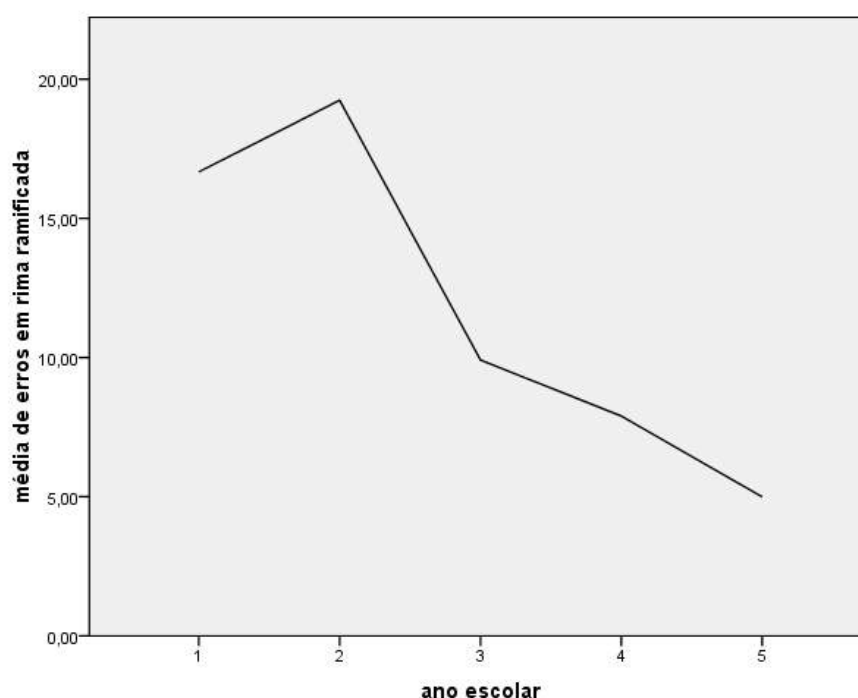


Figura 29 – Gráfico linear das médias de frequência de erros (orto)gráficos em rima ramificada conforme cada ano escolar.

Fonte: elaboração própria (com auxílio do *software* SPSS Statistics).

Este é um gráfico que causa uma impressão visual um pouco distinta daquela emitida pelo gráfico relativo às medias de *onset* complexo (Figura 28) em relação à variável ano escolar. O gráfico da Figura 29 cria um movimento descendente mais contínuo, com uma leve interrupção logo no início (no 2º ano); o gráfico da Figura 28 cria um movimento que sugere uma curva em U, com uma interrupção do movimento descendente mais tardiamente, em um momento já não esperado (no 4º ano).

Os resultados exibidos até então, ilustrados por meio dos gráficos apresentados nas Figuras 28 e 29, permitem concluir que há uma diferença entre *onset* e rima ramificados no que se refere ao papel que exerce o processo inicial de escolarização sobre a aquisição dessas estruturas silábicas na escrita. Para o *onset*, há diferença significativa entre alguns anos escolares na frequência de erros (orto)gráficos observados, mais especificamente entre o 1º e o 3º; para a rima, não se verificam essas diferenças, apesar de a média de frequência entrar em declínio contínuo a partir do 2º ano.

Por outro lado, há uma semelhança entre as duas estruturas silábicas a esse respeito: a frequência de erros em ambas não apresenta *necessária* correlação (negativa) com o passar da escolarização. Isso pode significar que, a qualquer

momento, curvas em U podem vir a surgir no curso de aquisição da escrita tanto da rima quanto do *onset*, como de fato ocorreu com este último, quando observou-se um aumento inesperado da média de erros no 4º ano. Mesmo no caso da rima ramificada, também observa-se um aumento na média de erros, o que, no entanto, não chega a se configurar idealmente como uma curva em U, pois o aumento se verifica logo no início do processo, do 1º para o 2º ano.

Assim, a hipótese de investigação para a questão 4, amparada por estudos do GEALE (cf. CUNHA, 2004; NEY, 2012; PACHALSKI; MIRANDA, 2018a; MIRANDA, 2019a) e por estudos que analisam dados de outras variedades do Português (SANTOS, 2013; VICENTE, 2018), se sustenta em parte no que diz respeito às sílabas complexas, uma vez que o papel da escolarização parece ter efeito significativo apenas sobre a grafia do *onset* complexo, e não sobre a rima ramificada. Essa discussão é retomada e ampliada na próxima subseção, quando a questão 7 é abordada.

4.4 A natureza dos erros (orto)gráficos em sílabas complexas: questões 5, 6 e 7

As subseções antecedentes abordaram o primeiro bloco de questões, as quais tinham em comum a discussão sobre aspectos ligados, principalmente, à estrutura fonológica das sílabas complexas: qual o comportamento de *onset* e rima ramificados em termos de frequência de erros e de acertos em sua grafia na aquisição da escrita, qual o papel dos segmentos na ocorrência dos erros nessas estruturas e, ainda, qual a relação dos erros em sílabas complexas com a escolarização, uma variável de caráter extralinguístico.

A presente subseção aborda o segundo bloco de questões da pesquisa, cuja preocupação reside sobre a natureza dos erros (orto)gráficos em sílabas complexas. Este aspecto é analisado e discutido considerando a proposta de categorização que vem sendo desenvolvida pelo GEALE desde 2001, apresentada na Figura 4 (subseção 2.1.3) e também, de forma mais específica, na Figura 22 (subseção 3.3) deste trabalho. É importante retomar, aqui, que, segundo essa categorização, os erros (orto)gráficos em sílabas complexas produzidos pelas crianças que estão em processo de aquisição da escrita podem advir de conhecimentos referentes a duas fontes principais e distintas: fonologia e ortografia. Além disso, importa saber que erros

residuais, isto é, aqueles nos quais a natureza é indefinida, são agrupados, neste trabalho, na categoria *outros* (ver Figura 23 na subseção 3.3).

A primeira questão do segundo bloco, de número 5, sugere a discussão sobre a natureza diversa dos erros, especialmente daqueles referentes à grafia da rima ramificada. Esta estrutura silábica comporta, em termos mais abrangentes, a existência de erros tanto de natureza fonológica quanto ortográfica, à diferença do onset complexo, que não possui complicadores ortográficos propriamente ditos, fazendo com que os erros envolvendo sua grafia sejam, por excelência, de ordem fonológica. Conforme explanado na subseção 3.3, é a representação das consoantes /N/ e /S/ que estão associados complicadores ortográficos na rima ramificada. Investigar o que propõe esta questão é importante para que se possa avaliar o peso que têm os diferentes tipos de conhecimento mobilizados na grafia desse tipo de estrutura complexa durante a aquisição da escrita.

Erros de natureza fonológica em rima ramificada são amplamente ilustrados na discussão envolvendo a questão 7, feita mais adiante nesta subseção. Por ora, é importante ter em vista que eles podem se configurar como omissão da consoante ligada ao constituinte ramificado ('atos' para 'altos'), alteração nos segmentos da rima – substituição ('camportou' para 'comportou') –, epêntese ('inisedio' para 'incêndio') e metátese ('enprotante' para 'importante'). Erros desse tipo são considerados de motivação fonológica pois correspondem a fenômenos fonológicos típicos materializados na fala, tanto adulta como infantil, os quais geralmente ocorrem em virtude de algum tipo de complexidade representacional, seja a nível de segmento ou sílaba (como é o caso da ramificação silábica).

Erros de natureza ortográfica, nos casos passíveis de análise, isto é, na grafia de /N/ e /S/, podem ser visualizados na Figura 30, que apresenta exemplos com excertos dos textos analisados nesta pesquisa:



Figura 30 – Exemplos de grafias com erros de natureza ortográfica em rima ramificada, feitas por alunos do 4º e do 5º ano: 'emgraçado' para 'engraçado', 'tanben' para 'também' e 'extinção' para 'extinção'. Fonte: BATALE – 7º estrato (MIRANDA, 2001).

Como é possível observar na Figura 30, erros de natureza ortográfica na rima ramificada podem envolver, no caso da grafia da nasal, o uso de <n> no lugar de <m> e vice-versa, a partir dos quais percebe-se que a regra contextual não está clara para as crianças, ou seja, não há o conhecimento de que, em posição final de sílaba dentro da palavra, existe um contexto bem definido que auxilia a eleger cada uma das duas opções gráficas existentes para a nasalidade (<m>, grafema que representa uma consoante labial, deve ser utilizado apenas antes dos grafemas correspondentes às, também, consoantes labiais /p/ ou /b/). É também possível considerar, no caso específico de ‘emgraçado’, que a criança esteja reconhecendo uma palavra gramatical na partícula ‘en-’ – a preposição ‘em’ – fenômeno comum durante o processo de aquisição da escrita e que geralmente é acompanhado da hipersegmentação, isto é, a inserção de um espaço em branco em lugar não previsto pela norma. Esse tipo de ocorrência é explorado com mais detalhe adiante nesta subseção.

Quanto à grafia da fricativa, é necessário considerar que a representação do fonema /s/ na escrita, independente da sua posição na sílaba, constitui o caso mais complicado dentro do sistema ortográfico do PB, uma vez que existem dez grafemas distintos elegíveis para sua representação. Miranda (2020), no entanto, observa que para cada grafema é possível estabelecer alguns contextos para sua utilização, o que ajuda a diminuir o conjunto de opções, reduzindo, assim, a sobrecarga na memória do usuário a cada grafia da fricativa. Ainda assim, restam casos de arbitrariedade. A concorrência entre <s> e <x> em final de sílaba dentro da palavra é bastante ilustrativa nesse sentido: há uma restrição contextual para o grafema <x> que se interpõe à arbitrariedade, uma vez que o <x> só é utilizado se houver antes dele a vogal coronal /e/, como em ‘**ext**ensão’. No entanto, ainda se trata de uma regra arbitrária em última instância, pois o grafema <s> também pode aparecer diante da vogal /e/, como em ‘**est**ender’. O exemplo dado na Figura 30, ‘**est**inção’ para ‘**ext**inção’, explica-se, dessa forma, por meio da incerteza gerada pela arbitrariedade em relação à escolha do grafema.

A Tabela 16 apresenta a distribuição das médias de erros ortográficos em rima ramificada considerando as diferentes naturezas de erro observadas neste estudo⁶⁶:

⁶⁶ Uma vez que nesta seção são discutidas as diferentes naturezas dos erros, o universo considerado para as análises é apenas relativo aos erros. Assim, na linha da frequência absoluta, deve-se ler que há 40 erros de natureza fonológica em rima ramificada de 78 erros no total, e que há 38 erros de

Tabela 16 – As diferentes naturezas de erro em rima ramificada.

	Fonológica (n = 47)	Ortográfica (n = 47)
Média (%)	51,24	48,76
Desvio padrão	47,71	47,71
Frequência absoluta	40/78	38/78
Z	-0,40	

Os dados da Tabela 16 exibem o resultado de médias muito próximas entre erros fonologicamente motivados em relação a erros ortograficamente motivados: 51,24% contra 48,76% de erros respectivamente. O teste *Wilcoxon*, expresso pelo valor de Z na última linha da tabela, confirma que a pequena diferença percentual observada não é estatisticamente significativa. O resultado descritivo ainda assim já chama à atenção, pois vai ao encontro da hipótese sustentada por estudos do GEALE – e que ampara a hipótese de investigação relacionada à questão 5 – é a de que os erros motivados pela fonologia predominam sobre aqueles motivados pela ortografia, considerando a média geral dos primeiros anos escolares (cf. MIRANDA, 2013, 2017, 2020).

Entretanto, outra ideia também defendida nesses mesmos estudos é a de que o conhecimento fonológico, enquanto insumo principal das crianças no período inicial da escrita alfabética, vai dando espaço ao conhecimento ortográfico à medida que os anos escolares vão avançando, ou seja, à medida que avança o próprio processo de aquisição da escrita. É assim que emerge a questão de número 6 desta pesquisa: há diferença na frequência de erros de naturezas fonológica e ortográfica se forem comparados os dois períodos distintos que caracterizam o 1º Ciclo do Ensino Fundamental, isto é, o Ciclo de Alfabetização e os 4º e 5º anos⁶⁷? Seriam as médias gerais, observadas na Tabela 16, fidedignas a essa nuance? Considerando tal

natureza ortográfica de 78 erros no total. É importante notar, ainda, que o total de 78 erros se refere a erros fonológicos e ortográficos (sem considerar a categoria residual *Outros*) somente em nasal e fricativa, pois são essas as consoantes que comportam, na rima ramificada, as duas diferentes naturezas de erro e que permitem a comparação realizada.

⁶⁷ A razão para a divisão dos sujeitos nos dois grandes grupos, ao invés de considerar cada ano escolar separadamente, deve-se ao fato de se considerar essa divisão mais proveitosa e econômica para a realização das rodadas estatísticas, uma vez que reúne mais casos válidos para a comparação das médias, além de reduzir o número de rodadas a serem realizadas (de cinco para duas). Além disso, o aspecto conceitual também torna essa divisão plausível e até mesmo desejável, uma vez que são esperadas aprendizagens distintas para ambos os grupos, em termos de aprofundamento do processo de aquisição da escrita – razão, inclusive, pela qual o primeiro grupo é considerado pela legislação como um ciclo contínuo de aprendizagem, a saber o Ciclo de Alfabetização.

indagação, as Tabelas 17 e 18 apresentam adiante as médias de erros de naturezas fonológica e ortográfica distribuídas conforme os dois períodos que compõem os anos iniciais do Ensino Fundamental, acompanhadas também do resultado dos testes não-paramétricos *Wilcoxon* e de seus valores de significância:

Tabela 17 – As diferentes naturezas de erro em rima ramificada no Ciclo de Alfabetização (1º a 3º anos).

	Fonológica (n = 25)	Ortográfica (n = 25)
Média (%)	70,00	30,00
Desvio padrão	45,64	45,64
Frequência absoluta	27/36	9/36
Z	-2,04*	

* $p < 0,05$

Tabela 18 – As diferentes naturezas de erro em rima ramificada no 4º e no 5º anos.

	Fonológica (n = 22)	Ortográfica (n = 22)
Média (%)	29,92	70,08
Desvio padrão	41,36	41,36
Frequência absoluta	13/42	29/42
Z	-1,87 ^{ms}	

ms = marginalmente significativo

Os resultados apresentados pelas Tabelas 17 e 18 indicam haver, de fato, importantes diferenças entre os dois períodos dos anos iniciais no que concerne ao tipo de conhecimento mobilizado pelas crianças durante a aquisição da escrita. Os valores de média observados para as duas naturezas de erro em rima ramificada são quase inversamente proporcionais ao se comparar o Ciclo de Alfabetização com os 4º e 5º anos. Os testes *Wilcoxon* assinalaram diferença significativa entre as médias de erros de natureza fonológica e de natureza ortográfica no Ciclo de Alfabetização, sendo $Z = -2,04$ e $p = 0,041$, e diferença marginalmente significativa nos 4º e 5º anos, sendo $Z = -1,87$ e $p = 0,061$ para os 4º e 5º anos. Tais resultados permitem rejeitar a hipótese nula (H_0) e reter a hipótese alternativa (H_1) correspondentes à questão 6.

Dessa maneira, percebe-se que as médias gerais obtidas para os anos iniciais, apresentadas pela Tabela 16, escondem uma notável característica do *processo* de aquisição da escrita. A hipótese defendida pelos estudos do GEALE, de que o

conhecimento fonológico seria mais acionado no período inicial do processo de aquisição da escrita, por ser um recurso de que a criança dispõe de antemão em virtude de sua já existente e bem estabelecida competência linguística, parece se sustentar com a comparação apresentada pelas Tabelas 17 e 18. O conhecimento ortográfico, por sua vez, apareceria com força mais tardiamente, tendo em vista se tratar de um conhecimento que depende preponderantemente da instrução formal, ou seja, a escola tem de apresentá-lo à criança e ajudá-la a construí-lo.

Se bem observado, essa ideia também encontra amparo no modelo de desenvolvimento ortográfico de Seymour (1997), apresentado e explanado na subseção 2.1.3 deste trabalho. Com o modelo, o autor sustenta que a principal contribuição do componente *Processo Alfabético*⁶⁸, o ponto de partida do desenvolvimento cognitivo da ortografia, é a formação do núcleo estrutural do sistema ortográfico. Isso significa dizer, em outras palavras, que sem o estabelecimento da forma mais básica de relação entre fonologia e ortografia – o princípio de funcionamento de um sistema de escrita alfabética – não há como elaborar uma estrutura ortográfica propriamente dita, com construção e organização, por exemplo, de regras que regem as relações múltiplas entre fonemas e grafemas. Nesse sentido, o conhecimento fonológico é condição para o acesso e o desenvolvimento do conhecimento ortográfico. O resultado exibido nas Tabelas 17 e 18 vai ao encontro desse raciocínio, ao evidenciar que o conhecimento ortográfico aparece de forma expressiva nos 4º e 5º anos do Ensino Fundamental, apenas depois de o conhecimento fonológico dominar a primeira etapa do processo de aquisição da escrita, representada pelo Ciclo de Alfabetização.

É interessante, ainda, traçar um comparativo entre esse resultado e aquele obtido como resposta à questão 4 desta pesquisa, que explora o papel da escolarização no processo de aquisição da escrita de sílabas complexas. Conforme exposto na subseção 4.3, os testes estatísticos realizados não apontaram nem para uma correlação negativa entre a escolarização e a frequência de erros (orto)gráficos em rima ramificada, nem para diferenças entre os anos escolares no que se refere à frequência dos erros nessa estrutura. Contudo, seguindo aquilo que há pouco foi

⁶⁸ O Processo Alfabético é um componente do processamento ortográfico que está fundado na compreensão do *princípio alfabético*, ou seja, na consciência sobre a realidade fônica da língua e de sua atualização em substância gráfica, o que possibilita o estabelecimento das relações iniciais entre fonemas e grafemas. Para sua formação, a consciência fonêmica é particularmente importante.

observado, a partir das Tabelas 17 e 18, ao serem escrutinadas as diferentes naturezas de erro passíveis de verificação na rima ramificada, a escolarização parece, sim, ter efeito significativo no processo de aquisição da escrita dessa estrutura silábica. Isso é possível constatar na medida em que se assume que o acesso ao conhecimento ortográfico, cuja proeminência se dá apenas nos anos finais do 1º Ciclo do Ensino Fundamental, é uma consequência do ensino formal. Disso decorre, portanto, que a escolarização tem influência sobre o tipo de conhecimento mais ou menos mobilizado pelas crianças ao longo da aquisição da escrita, ao menos no que concerne à rima ramificada.

Por fim, a última questão do bloco, de número 7, se ocupa especificamente dos erros de natureza fonológica, procurando verificar, dentro dessa grande categoria, quais são as estratégias preferenciais das crianças para grafar *onset* e rima ramificados. Conforme esclarecido na subseção 3.3, neste estudo, para ambas as estruturas silábicas analisadas, foram estabelecidos previamente 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 escrita (ver subseções 2.4.2.1 e 2.4.2.2) e nos principais tipos de fenômenos verificáveis na fonologia infantil quando se trata das sílabas complexas (ver subseções 2.4.1.1 e 2.4.1.2): omissão, metátese, epêntese e substituição.

A Tabela 19, que segue adiante, exhibe os resultados descritivos e inferenciais relacionados aos tipos de erros de natureza fonológica produzidos pelas crianças na grafia do *onset* complexo. Para verificar a significância da diferença entre a frequência de ocorrência de cada tipo de erro, foi utilizado o teste não-paramétrico *Friedman*, tendo em vista que as variáveis têm distribuição anormal e o *design* da amostra é intrassujeitos. Seu resultado está expresso na última linha da tabela:

Tabela 19 – Erros de natureza fonológica em *onset* complexo.

	Omissão (n = 8) ⁶⁹	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 absoluta	16/18	2/18	0/18	0/18
χ^2 (3)	22,00*			

* $p < 0,05$

A Tabela 19 revela que os erros de omissão da segunda consoante em *onset* complexo concentram a maior média, com 93,30% (16 ocorrências de 18 erros de natureza fonológica no total), seguida da metátese, que concentra 6,70% (2 ocorrências de 18 erros de natureza fonológica no total). Não foram constatados casos de epêntese e substituição⁷⁰. As Figuras 31 e 32 ilustram as estratégias registradas na amostra:



Figura 31 – Exemplos de grafias com omissão de C2 em *onset* complexo, por alunos do 2º e do 4º ano, respectivamente: ‘foresta’ para ‘floresta’ e ‘brinquedos’ para ‘brinquedos’.

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



Figura 32 – Exemplo de grafias com metátese em *onset* complexo, por alunos do 2º e do 4º ano, respectivamente: ‘furta’ para ‘fruta’ e ‘garde’ para ‘grande’.

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

O teste *Friedman* revelou haver diferença significativa entre as médias de frequência de ocorrência desses diferentes tipos de erros de natureza fonológica em

⁶⁹ O baixo valor de n é um fato que chama à atenção e reflete uma característica real da amostra: significa que, dos 145 sujeitos que produziram estruturas com *onset* complexo nos textos (ver Tabela 5), apenas 8 sujeitos efetivamente produziram erros (de natureza fonológica, no caso em análise na Tabela 19).

⁷⁰ Miranda (2019a), que encontra 11,50% casos de substituição ao analisar a grafia de *onset* complexo, apresenta como exemplo desse tipo de a forma ‘bruzéu’ para ‘blusão’. Pachalski et al. (2014) apresenta como exemplo de epêntese, que concentra 5% de média de ocorrência em sua amostra, o dado ‘outoros’ para ‘outros’.

onset complexo, sendo $\chi^2 (3) = 22,00$ e $p = 0,043$, o que permite rejeitar a hipótese nula (H_0) e reter a hipótese alternativa (H_1), a qual pressupõe haver diferença entre as médias de cada tipo de erro. Porém, entre quais estratégias a diferença é significativa? A fim de verificar essa questão, foi realizado o teste *post hoc* Conover⁷¹ com correção de Bonferroni. A significância do teste pode ser visualizada na Tabela 20:

Tabela 20 – Significância do teste Conover para verificação das diferenças entre as médias de frequência de cada tipo de erro de natureza fonológica em *onset* complexo.

	Omissão vs. Metátese	Omissão vs. Epêntese	Omissão vs. Substituição	Metátese vs. Epêntese	Metátese vs. Substituição	Epêntese vs. Substituição
Frequência de erros	*	*	*	ns	ns	ns

* $p < 0,0083$

ns = não significativo

Os resultados do teste Conover, exibidos na Tabela 20, revelam haver diferenças significativas entre as médias de erros do tipo omissão em relação às médias de erros dos tipos metátese, epêntese e substituição, sendo $p = 0,001$ (omissão vs. metátese), $p = 0,000$ (omissão vs. epêntese) e $p = 0,000$ (omissão vs. substituição). Tais resultados permitem rejeitar a hipótese nula (H_0) e reter a hipótese alternativa (H_1), a qual presume ser a omissão o tipo de erro significativamente mais frequente em relação aos demais tipos de erros de natureza fonológica observados em *onset* complexo.

No tocante aos tipos de erros de natureza fonológica produzidos pelas crianças na grafia da rima ramificada, observam-se, conforme apresenta a Tabela 21 adiante, tanto distinções quanto similaridades em relação aos resultados encontrados para o *onset* complexo:

⁷¹ Teste rodado no software RStudio por não estar disponível na versão de SPSS utilizada, com orientação de profissional da estatística.

Tabela 21 – 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 absoluta	34/58	2/58	0/58	22/58
$\chi^2 (3)$	43,82***			

*** $p < 0,001$

Os resultados descritivos exibidos na Tabela 21 revelam que a omissão na rima ramificada é o tipo de erro mais frequente, com percentual de 59,91% (34 ocorrências de 58 erros de natureza fonológica no total), tal como ocorre na grafia de *onset* complexo. Em segundo lugar está a substituição, com 37,00% de média (22 ocorrências de 49 erros de natureza fonológica no total), resultado que chama a atenção uma vez que, para o *onset*, não há sequer um caso de substituição registrado. Em seguida, com percentual bem mais baixo, vem a metátese, com 3,66% de ocorrências (2 ocorrências de 58 erros de natureza fonológica no total). Não foram registrados casos de epêntese⁷². As Figuras 33, 34 e 35 ilustram as estratégias encontradas na amostra:



Figura 33 – Exemplos de grafias com omissão em rima ramificada, por alunos do 2º e do 4º ano, respectivamente: ‘tabem’ para ‘também’ e ‘imãzinha’ para ‘irmãzinha’.
Fonte: BATALE – 7º estrato (MIRANDA, 2001).

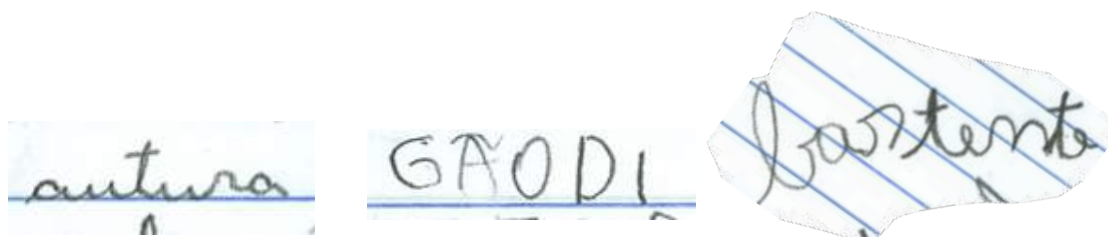


Figura 34 – Exemplos de grafias com substituição em rima ramificada, por alunos do 5º, do 1º ano e do 2º ano, respectivamente: ‘autura’ para ‘altura’, ‘gãodi’ para ‘grande’ e ‘bastente’ para ‘bastante’.
Fonte: BATALE – 7º estrato (MIRANDA, 2001).

⁷² Pachalski e Miranda (2018a) dão como exemplo de epêntese em rima ramificada o dado ‘inisedio’ para ‘incêndio’.

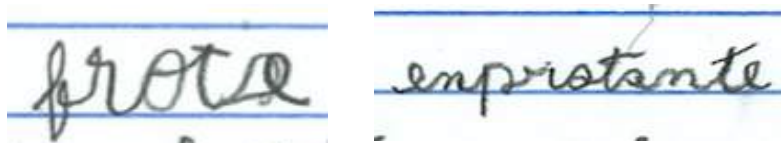


Figura 35 – Exemplo de grafia com metátese em rima ramificada, por alunos do 2º e do 4º ano, respectivamente: ‘frote’ para ‘forte’ e ‘enprotante’ para ‘importante’.
Fonte: BATALE – 7º estrato (MIRANDA, 2001).

Para verificar a significância da diferença entre as médias de frequência de ocorrência de cada tipo de erro na rima ramificada, foi utilizado o teste *Friedman*, tendo em vista que as variáveis têm distribuição anormal e o *design* da amostra é intrassujeitos. Seu resultado está expresso na última linha da Tabela 21, por meio do qual se entende que o teste revelou haver diferença significativa entre a frequência de ocorrência dos diferentes tipos de erros de natureza fonológica em rima ramificada, sendo $\chi^2(3) = 43,82$ e $p = 0,000$, permitindo que se rejeite a hipótese nula (H_0) e que se retenha a hipótese alternativa (H_1), que pressupõe a existência de diferença entre as médias dos diferentes tipos de erro.

A fim de averiguar precisamente entre quais tipos de erro a diferença existe, foi realizado o teste *post hoc* *Conover*⁷³ com correção de *Bonferroni*, cujas significâncias são expressas pela Tabela 22:

Tabela 22 – Significância do teste *Conover* para verificação das diferenças entre as médias de frequência de cada tipo de erro de natureza fonológica na rima ramificada.

	Omissão vs. Metátese	Omissão vs. Epêntese	Omissão vs. Substituição	Metátese vs. Epêntese	Metátese vs. Substituição	Epêntese vs. Substituição
Frequência de erros	*	*	*	ns	*	*

* $p < 0,0083$

ns = não significativo

Conforme apresenta a Tabela 22, o teste *post hoc* *Conover* revelou haver diferenças significativas entre todos os tipos de erros classificados, à exceção de metátese e epêntese, sendo $p = 0,000$ (omissão vs. metátese), $p = 0,000$ (omissão vs. epêntese), $p = 0,000$ (omissão vs. substituição), $p = 0,000$ (metátese vs.

⁷³ Teste rodado no software *RStudio* por não estar disponível na versão de SPSS utilizada, com orientação de profissional da estatística.

substituição), $p = 0,000$ (epêntese vs. substituição). Tais resultados permitem rejeitar a hipótese nula (H_0) e reter a hipótese alternativa (H_1). Assim, a omissão é significativamente mais frequente que os demais tipos de erros de natureza fonológica observados em rima ramificada. É interessante notar que, com esses resultados, a substituição também demonstra ter papel importante, pois sua média é significativamente maior que as médias de metátese e de epêntese, sendo apenas significativamente menor que a média da omissão.

Ao serem comparados os resultados obtidos para *onset* e rima ramificados, no tocante aos erros de natureza fonológica, observa-se que a omissão é o tipo de erro mais frequente em ambas as estruturas silábicas, aspecto que, devido às suas implicações, é discutido em detalhe mais adiante nesta subseção. As diferenças entre as estruturas silábicas estão localizadas nos demais tipos de erros: metátese e substituição, sendo este mais frequente na rima ramificada e aquele mais frequente em *onset* complexo.

Em pesquisa desenvolvida a fim de investigar exclusivamente a metátese na aquisição da escrita, Pachalski e Miranda (2018b) argumentam que o fenômeno está diretamente ligado à complexidade silábica, sobretudo em *onset*. As autoras mostram que a maioria dos casos de metáteses segmentais simples⁷⁴ ocorrem tendo sílabas constituídas de *onset* complexo como domínio do fenômeno, como em ‘barbo’ para ‘brabo’, ou como resultado do fenômeno, como em ‘codra’ para ‘corda’. Ferreiro e Zamudio [2008] (2013), em estudo revisado nas subseções de 2.4.2.1 a 2.4.2.3, verificam que, para a grafia de sílabas CCV no Espanhol, a metátese também é o tipo de erro mais frequente depois da omissão, diferente das sílabas CVC, para as quais se tem a epêntese como tipo de erro mais frequente após a omissão. Assim, os resultados obtidos neste estudo vão ao encontro daquilo que outros apontam sobre a metátese no contexto da grafia de sílabas complexas.

Em relação à substituição, primeiro é preciso observar as diferenças na qualidade que esse tipo de erro pode apresentar em cada constituinte silábico analisado. Os exemplos apresentados na nota 70, para o *onset* complexo, e na Figura 34, para a rima ramificada, denotam que, quando ocorre em *onset*, a substituição é referente apenas à troca de um grafema por outro na segunda posição desse

⁷⁴ Segundo Pachalski e Miranda (2018b), na aquisição da escrita, as metáteses podem ser do tipo segmental simples, como no exemplo de ‘barbo’ para ‘brabo’, ou do tipo segmental dupla, quando envolvem a permuta de dois segmentos em uma palavra, como em ‘tolenadas’ para ‘toneladas’.

constituente, particularmente a troca do grafema que representa uma líquida por um grafema que representa outra líquida (<l> por <r> ou <r> por <l>). Na rima, a substituição ganha mais nuances: pode ser a troca de um grafema por outro na última posição da rima – o caso em que <l> é substituído por <u>, em ‘autura’ para ‘altura’; pode ser a troca de grafemas que representam a nasalidade pós-vocálica medial por grafemas que representam um ditongo nasal, alterando a rima como um todo – o caso em que <an> é substituído por <ão>, em ‘gãodi’ para ‘grande’; pode ser, ainda, a troca de um grafema por outro na primeira posição da rima – o caso em que <a> é substituído por <e>, em ‘bastente’ para ‘bastante’. A diversidade de formas que a substituição é capaz de assumir na rima ramificada pode explicar, ao menos em parte, sua média de ocorrência distinta em relação à observada no *onset* complexo.

É interessante notar que a substituição é um tipo de erro que se manifesta, na amostra analisada, apenas no que seriam as codas lateral e nasal, como mostram os dados da Tabela 23:

Tabela 23 – Erros de substituição em rima ramificada por tipo de coda.

	Lateral (n = 17)	Nasal (n = 17)	Rótica (n = 17)	Fricativa (n = 17)
Média (%)	64,70	35,30	0,00	0,00
Desvio padrão	49,25	49,25	0,00	0,00
Frequência absoluta	13/22	9/22	0/22	0/22

A Tabela 23 revela que não há um dado sequer registrado nas codas rótica e fricativa que apresente a substituição como estratégia gráfica. Assim, as consoantes lateral e nasal concentram, juntas, 100% dos casos de substituição em rima ramificada. Esse resultado reforça o comportamento distinto das crianças na grafia dessas estruturas, aspecto já observado e discutido na subseção 4.2. As consoantes lateral e nasal não somente concentram mais erros relativamente às consoantes rótica e fricativa, em sentido geral, mas também diferem destas últimas na qualidade dos erros de natureza fonológica. Como consequência, a hipótese de Miranda (2019) também é fortalecida: em fase inicial de aquisição da escrita, as crianças interpretam de modo fonologicamente distinto, de um lado, o que seriam consoantes lateral e nasal, e de outro as consoantes rótica e fricativa, no contexto da rima ramificada. Ou seja, a fonologia infantil licencia como codas legítimas apenas as consoantes fricativa e rótica. Pode-se supor, assim, que aquilo que é /l/ e /N/ em coda para o sistema

adulto e para o sistema ortográfico, são, para as crianças, glide na segunda posição de núcleo ramificado e vogal nasal em núcleo simples, respectivamente.

É interessante, ainda, comparar a qualidade da estratégia de substituição entre os dois contextos fonológicos em que ela ocorre, isto é, entre as consoantes lateral e nasal. Em primeiro lugar, observa-se, com o auxílio da Tabela 23, que há mais erros desse tipo envolvendo a grafia da lateral. O *input* auditivo, que, neste caso, está em relação biunívoca com a organização da fonologia infantil, já que a produção é semivocalizada e a forma fonológica correspondente seria uma vogal alta que resulta em semivogal, possivelmente está na base do índice observado. Entretanto, apesar de haver menos casos de substituição para a grafia da nasal, observa-se mais heterogeneidade na forma como a substituição pode se manifestar neste ambiente: ou como um registro de ditongo nasal ou como uma alteração na qualidade da vogal envolvida na sequência VN.

Para Miranda (2009a), esse tipo de dado revela uma sensibilidade da criança a características fonéticas distintas que as vogais assumem nesse contexto: a sua produção nasalizada, o que por si só já as difere das vogais orais, também aproxima perceptualmente alguns segmentos vocálicos, sobretudo /a/ e /e/ (BERTI; CHACON; PAGLIUSO, 2008), o que explicaria, por exemplo, a grafia de <e> no lugar de <a> em ‘bastente’ para ‘bastante’. A autora ainda estabeleceu uma comparação entre a incidência de erros na grafia da nasalidade conforme as diferentes qualidades de vogais envolvidas, por meio do que constatou que as vogais /a/ e /e/ são aquelas que mais favorecem erros, concentrando, juntas, 58% de ocorrências do total.

Além disso, Miranda (2009a) chama a atenção para a heterogeneidade como uma característica marcante em geral dos erros envolvendo a grafia da nasalidade medial, não somente aqueles classificados como substituição. A autora expõe alguns exemplos, dentre eles erros envolvendo a palavra ‘grande’, que no *corpus* por ela analisado apresenta, ao menos, 10 formas gráficas distintas. Nesse item lexical específico, é preciso observar que, além da complexidade representacional ligada à nasalidade medial, há também uma associada ao *onset* complexo na mesma sílaba, o que pode intensificar as dúvidas das crianças. Na Figura 34, essa palavra aparece como exemplo de erro encontrado dentro da amostra analisada nesta pesquisa (‘gãodi’), e evidencia, de fato, dúvidas tanto em relação à representação do *onset* quanto de rima ramificados.

Miranda (2009a), ainda buscando ilustrar a diversidade de estratégias utilizadas pelas crianças, exhibe dados nos quais são verificados casos de hipersegmentação (a inserção do espaço em branco em lugar não previsto pela norma)⁷⁵ justamente no lugar onde é convencionado o uso do grafema <n>. Dados como esse também são encontrados na amostra desta pesquisa, como exhibe a Figura 36:

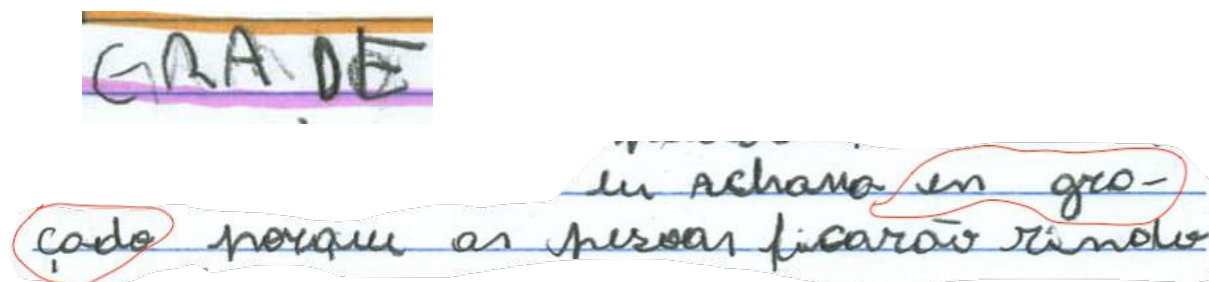


Figura 36 – Exemplo de grafia com hipersegmentação em contexto de nasalidade medial, por alunos do 2º e do 4º ano, respectivamente: ‘gra de’ para ‘grande’ e ‘en graçado’ para ‘engraçado’⁷⁶.
Fonte: BATALE – 7º estrato (MIRANDA, 2001).

Esse tipo de dado, ilustrado pela Figura 36, foi computado na categoria residual *outros*, uma vez que não se encaixa na definição de cada uma das duas naturezas principais de erro em sílabas complexas, ou seja, fonologia e ortografia. De todo modo, os exemplos da figura reforçam a ideia da complexidade especial envolvida na grafia da nasalidade pós-vocálica. Porém, as duas grafias apresentadas têm diferenças entre si que merecem alguns comentários.

Em ‘en graçado’ para ‘engraçado’, a hipersegmentação parece estar vinculada à sua motivação mais típica, gerando um resultado que Cunha (2004, 2010) aponta ser o mais frequente em textos de crianças brasileiras: uma palavra gramatical seguida de uma palavra fonológica. Palavras gramaticais, que costumam ser escritas com uma ou duas letras – o caso de preposições, artigos, pronomes e conjunções (em, a, os, e, que, para etc.) – são, no início do processo de aquisição da escrita, frequentemente unidas pelas crianças a palavras fonológicas, gerando o movimento

⁷⁵ Os erros de segmentação não-convencional, dentre os quais se inclui a hipersegmentação, são considerados um dos mais frequentes tipos de erro (orto)gráfico nos anos iniciais de escolarização (CUNHA, 2004, 2010; MIRANDA, 2013). No Brasil, Abaurre (1991b) é pioneira na abordagem fonológica das segmentações não-convencionais em dados de aquisição da escrita. Estudos mais recentes têm produzido análises das hipo e hipersegmentações em dados de alunos do Ensino Fundamental I (CUNHA, 2004, 2010; CHACON, 2004; CAPRISTANO, 2004) e do Ensino Fundamental II (TENANI, 2004).

⁷⁶ Sugestão de leitura: “[...] eu achava engraçado porque as pessoas ficaram rindo [...]”.

contrário ao dado ‘en graçado’: as hipossegmentações. Para Cunha (2004, 2010), isso ocorre em virtude de as crianças, no início do processo, não reconhecerem sequências de uma ou duas letras como palavras legítimas. Favorece esse comportamento o fato de essas sequências não serem fonologicamente acentuadas, o que na própria hierarquia prosódica⁷⁷ já é motivo de um processo de juntura, que termina por gerar um grupo clítico (C). Cunha (2004, 2010) mostra que, no decorrer do processo de escolarização, as crianças passam a conferir a essas estruturas o estatuto de palavra, razão pela qual os erros de hipossegmentação vão diminuindo e, também, os casos de hipersegmentação começam a aparecer. No caso de ‘en graçado’, assim, a criança demonstra reconhecer que uma preposição como ‘em’ tem estatuto de palavra gráfica, ainda que o faça em um contexto inadequado.

Já em ‘gra de’ para ‘grande’, de fato parece ser mais evidente que seria o conflito entre a fonologia infantil e a “fonologia do sistema ortográfico” o motivo principal da hipersegmentação, uma vez que, além dela, há também a omissão do grafema <n>. É como se estivesse sendo reservado um espaço gráfico para registrar um aspecto da língua cuja forma – fonológica e ortográfica – não está bem clara para a criança, embora ela perceba que existe. Reforça essa interpretação o fato de a hipersegmentação gerar uma ruptura no pé métrico, fenômeno raro como resultado desse tipo de processo, conforme as tendências apontadas por Cunha (2004, 2010). A autora encontra dados do tipo ‘gritan do’ para ‘gritando’ e ‘correm do’ para ‘correndo’, e sugere como análise a ideia de que as crianças reconhecem, de um lado, palavras de conteúdo reais da língua (‘gritam’ e ‘correm’) e, de outro, palavras gramaticais (‘do’). Cunha (2004, 2010) não descarta também a complexidade relacionada à grafia da nasalidade medial como uma variável que atue conjuntamente na motivação desses casos de hipersegmentação. Entretanto, é importante observar que no dado ‘gra de’ não há palavra de conteúdo à esquerda, e apenas o que poderia ser uma palavra gramatical reconhecida à direita, fato que parece corroborar ainda mais a ideia do conflito entre a fonologia infantil e a “fonologia do sistema ortográfico” como motivo principal da hipersegmentação registrada.

Para além de conferir robustez à argumentação sobre a complexidade representacional das estruturas em foco nesta pesquisa e, de forma específica, sobre a heterogeneidade constatada na grafia da nasalidade medial, é interessante analisar

⁷⁷ Ver nota 11 deste trabalho (subseção 2.1.1).

esse tipo de dado de forma mais minuciosa para também compreender melhor que qualidade de erro está, afinal, incluída na categoria *outros*, ao menos no que se refere à rima ramificada. Como é possível notar, são dados mais episódicos, motivados por outros conhecimentos que não o fonológico (como o reconhecimento de palavras gráficas), e que não necessariamente estão relacionados com o problema da representação das estruturas silábicas complexas em si, apesar de as envolverem em termos de ambiente onde o fenômeno acontece (o caso de ‘en graçado’, por exemplo).

Após essa discussão que começou com a estratégia de substituição e culminou com os dados heterogêneos envolvendo a grafia da nasalidade pós-vocálica, é necessário retomar e discutir aquele que se mostrou, de acordo com as Tabelas 19 e 21, o tipo de erro mais frequente, tanto em *onset* quanto em rima ramificados: a omissão. A principal questão, neste caso, é por que essa seria a estratégia preferencial das crianças? O que esse dado pode indicar acerca do que está em jogo nas representações fonológica e ortográfica das estruturas silábicas complexas durante o processo de aquisição da escrita?

O fato de a omissão ser o tipo de erro mais frequente é bastante relevante considerando a hipótese geral de pesquisa deste trabalho. Primeiro, a omissão assinala que, na tentativa de análise da estrutura silábica, a criança fica em dúvida menos sobre a posição e o tipo de segmento que deve ocupar a estrutura das sílabas, e mais sobre o número de segmentos que devem ser representados. Além disso, a omissão parece indicar, mais do que qualquer outro tipo de erro, que representar graficamente um fonema ligado a um constituinte silábico ramificado é um *problema*⁷⁸. Diferente dos outros fenômenos – metátese, epêntese e substituição –, na omissão o segmento aparenta passar despercebido durante o processamento da informação.

Sugerimos que um dos principais motivos para isso é que esse tipo de estrutura não está *plenamente explicitado*, ou, dito de outra maneira, porque há uma *explicitação incompleta*, como propomos que seja chamado esse fenômeno, o qual é amplamente explorado no decorrer do texto. Diante da parcial clareza e parcial

⁷⁸ O termo *problema*, neste contexto, remete ao seu conceito científico, de *questão a ser investigada*, à qual estão associadas hipóteses. As crianças têm, na grafia da ramificação silábica, um problema no sentido que esse aspecto lhes impõe uma questão a ser investigada. Para essa questão formulam hipóteses interpretativas: naturalmente provisórias, em processo de experimentação por meio da escrita e evidenciadas através dos erros (orto)gráficos. Essa perspectiva está diretamente associada à concepção piagetiana de *erro construtivo*, noção abordada na subseção 2.1.2 deste trabalho. Assim, *problema* não tem sentido negativo, como de “mau funcionamento de alguma coisa, que acarreta transtornos” (HOUAISS, 2004, p. 597).

acessibilidade à organização estrutural a que está submetida uma sequência segmental, a relação entre o constituinte e o segmento a ele associado em alguns momentos pode simplesmente ser ignorada, o que torna mais fácil omitir o segmento do que produzi-lo em posição ou qualidade alterada (o caso de metátese, epêntese e substituição).

É importante lembrar também que, com dados semelhantes, Treiman (1992), em estudo revisado na subseção 2.1.2.1, mostra que crianças pré-escolares têm dificuldade de analisar tanto a estrutura interna de *onset* quanto de uma rima quando esses constituintes são ramificados. Ademais, I. Y. Liberman et al. (1974) indicam que abstrair da linguagem falada os segmentos adequados à sua representação no sistema de escrita é uma tarefa cognitivamente mais desafiadora à medida que a unidade de segmentação do sistema de escrita é menos significativa do ponto de vista da estruturação do sistema linguístico. Assim, um sistema de escrita alfabético exige um grau de explicitação maior do que exigiria um silábico, por exemplo.

Outro aspecto, de ordem perceptual, pode contribuir como causa à proeminência da omissão no conjunto de erros de natureza fonológica em sílabas complexas. I. Y. Liberman et al. (1974), para sustentar a sua hipótese de pesquisa, chamam a atenção para a maneira como a fala é percebida e codificada pelo sistema linguístico humano, o que fazem baseados no estudo de A. M. Liberman, Cooper, Shankweiler e Studdert-Kennedy (1967). Segundo estes autores, a percepção do sinal acústico não corresponde biunivocamente à segmentação fonêmica, pois “os segmentos fonêmicos são codificados no nível acústico em unidades maiores aproximadamente de tamanho silábico” (I. Y. LIBERMAN et al, 1974, p. 203). Isso significa que a informação acústica a respeito das consoantes adjacentes a uma vogal fica contida na própria vogal e, assim, as pistas para identificar as fronteiras silábicas são veladas. Dessa maneira, explicitar e segmentar uma cadeia fonêmica com o número, o tipo e a posição correta das unidades que a compõe, como a presente em sílabas do tipo /mar/ e /tras/, é bastante desafiador na medida em que ela é percebida acusticamente como uma unidade só.

É necessário observar, ainda, outros fatores que adicionam nuances importantes a essa cara discussão. Um deles é o fato de que erros (orto)gráficos em sílabas complexas não são categóricos nos textos analisados, ou seja, é possível encontrar em um texto escrito pela mesma criança erros de omissão, por exemplo, e também acertos em sílabas complexas. Considere-se os exemplos da Figura 37:

Eu quero o pequeno ele é fofo e ele vai brincar comigo. Ele vai economizar ração ele vai vir comigo na mochila pro colégio porque ele é pequeno. Ele vai dormir comigo.

Eu não quero os outros porque eu tenho um médio e um grande não caberia no meu apartamento.

Taba
Um animal muito forte muito rápido come fruta e salva animais e vive na floresta e cava e é inteligente.

porque ele salva animais e vai ajudar floresta e vai salvar os ovos de passaros nunca vai para de ajudar todos.

Figura 37 – Exemplos de textos produzidos por alunos do 3º ano⁷⁹ e do 2º ano⁸⁰ do Ensino Fundamental, com erros e acertos em sílabas complexas.
Fonte: BATALE – 7º estrato (MIRANDA, 2001).

⁷⁹ Sugestão de leitura: “Eu quero o pequeno. Ele é fofo e ele vai brincar comigo. Ele vai economizar ração. Ele vai vir comigo na mochila pro colégio porque ele é pequeno. Ele vai dormir comigo. Eu não quero os outros porque eu tenho um médio e porque um grande não caberia no meu apartamento”.

⁸⁰ 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”.

No primeiro texto ilustrado pela Figura 37, observa-se um erro de omissão em *onset* complexo, em ‘bincar’ para ‘brincar’. Nas demais palavras onde se localizam estruturas silábicas complexas, tanto em *onset* como em rima, apenas acertos são verificados – o caso das palavras ‘dormir’, ‘outros’, ‘grande’ e ‘apartamento’.

No segundo texto da Figura 37, nota-se maior quantidade e também mais diversidade de erros em sílabas complexas. É o caso de ‘frote’ para ‘forte’, ‘furta’ para ‘fruta’, ‘foresta’ para ‘floresta’ e ‘nuca’ para ‘nunca’, nas quais são verificadas estratégias como omissão e metátese. Ainda assim, também são observados acertos em sílabas complexas, mais especificamente em rima ramificada, como nas palavras ‘salva’ e ‘inteligente’.

Essa observação vai ao encontro daquilo que se verifica em toda a amostra analisada, isto é, os erros (orto)gráficos em sílabas complexas apresentam índices significativamente inferiores ao número de acertos (ver Tabelas 5 e 6 na subseção 4.1). Da mesma forma, é possível observar, em um mesmo texto, como é o caso do segundo exemplo da Figura 37, a ocorrência de diferentes tipos de erros em sílabas complexas: além da omissão, ocorrem, em outras palavras do texto, erros como a metátese, os quais estão mais ligados às dúvidas que as crianças sobre a posição que os segmentos, uma vez identificados, devem ocupar na estrutura das sílabas.

Esse é um dos conflitos, aliás, que está na base da pergunta que move esta pesquisa: por que, em meio a uma abundância de acertos, ainda se observam erros? E, no mesmo sentido, por que em meio a erros que evidenciam maior sensibilidade à estrutura complexa (como a metátese), ainda são registradas ocorrências de omissão? Assim se chega à razão de uso do termo *incompleta* em *explicitação incompleta*. O problema envolvendo a representação gráfica das sílabas complexas não se refere a uma *total* inacessibilidade à organização estrutural a que está submetida uma sequência segmental, mas sim a um acesso (uma explicitação) restrito a essa organização. Do contrário, observar-se-ia um número muito maior de erros em relação ao que de fato se registra⁸¹, e erros de outras tipologias não coocorreriam com a omissão.

⁸¹ Esta é uma das importantes diferenças entre os percursos de aquisição da fonologia e da escrita. Como observado na subseção 2.4.2.3, enquanto a estrutura não está disponível fonologicamente, os erros ou a não-produção da estrutura são sistemáticos, e produções conforme o alvo são raras e esporádicas. Durante a aquisição da escrita, porém, o conhecimento fonológico é somente retomado e analisado, pois já foi construído, fazendo com que, desta vez, acertos sejam mais sistemáticos e erros mais esporádicos.

O MRR de Karmiloff-Smith (1994) parece ser útil para sustentar esse raciocínio. Primeiro, há de se lembrar que há um conhecimento implícito, o fonológico, principal insumo das crianças para o desenvolvimento da escrita alfabética. Conforme a terminologia da autora, este conhecimento está em um *nível I*, codificado em um formato que impossibilita sua manipulação, a sua análise e a interface entre representações mentais. Esse tipo de conhecimento permite às crianças, por ora, um bem-sucedido uso da linguagem atualizada em substância fônica.

Com o curso de aquisição da escrita alfabética, esse conhecimento precisa ser acessado ou explicitado, o que ocorre de forma progressiva. Por isso, mais uma vez, é possível falar em explicitação incompleta, isto é, existe um momento do percurso em que determinadas estruturas não estão totalmente abertas ao acesso consciente – nem totalmente inacessíveis, como no nível I. Uma estrutura que carece de explicitação completa, assim, é uma estrutura que estaria em um *nível E1* em termos de abertura ao acesso consciente, o nível seguinte ao *Implícito*, segundo o MRR. Uma das características que denotam a existência desse nível é a ocorrência de erros de execução em uma determinada tarefa que o sujeito costuma cumprir com *maestria procedimental*. Poderia se pensar, assim, que a coocorrência de erros e acertos em um texto produzido pelo mesmo sujeito, ou a coocorrência de erros de diferentes tipologias nesse mesmo texto, está incluída na caracterização feita por Karmiloff-Smith (1994): o sujeito costuma produzir conforme o alvo as sílabas complexas – tanto na fala quanto na escrita pois o conhecimento que serve como insumo para a produção é o mesmo – mas, por vezes e em um determinado 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, neste caso, ainda não está suficientemente explícito. É importante também ter em vista que o formato das representações no nível E1 são um pouco mais flexíveis em relação ao nível anterior e, por isso, *capazes de criar vínculos com outras representações*. Esta última observação se torna essencial considerando as ponderações que seguem no texto adiante.

O MRR dá pistas a respeito do porquê ocorrem explicitações incompletas ao propor, de maneira mais abrangente, que o desenvolvimento cognitivo ocorre por meio de um processo cíclico que transforma *progressivamente* representações implícitas *na mente* em representações explícitas *para a mente*. A abertura ao acesso consciente, assim, não ocorre de forma automática e em dois estágios apenas, mas

de forma progressiva e multinivelada. Mas o que sustenta que haja uma explicitação incompleta especificamente no tipo de dado que se está analisando, isto é, na aquisição da escrita de estruturas silábicas complexas? Aqui o modelo de Seymour (1997), revisado na subseção 2.1.3, pode ser retomado e relacionado ao de Karmiloff-Smith (1994). Como forma de guiar a retomada das ideias do autor, é importante também retomar a Figura 6 que mostra as unidades de representação silábica com as quais a criança lida no curso de desenvolvimento da consciência fonológica:

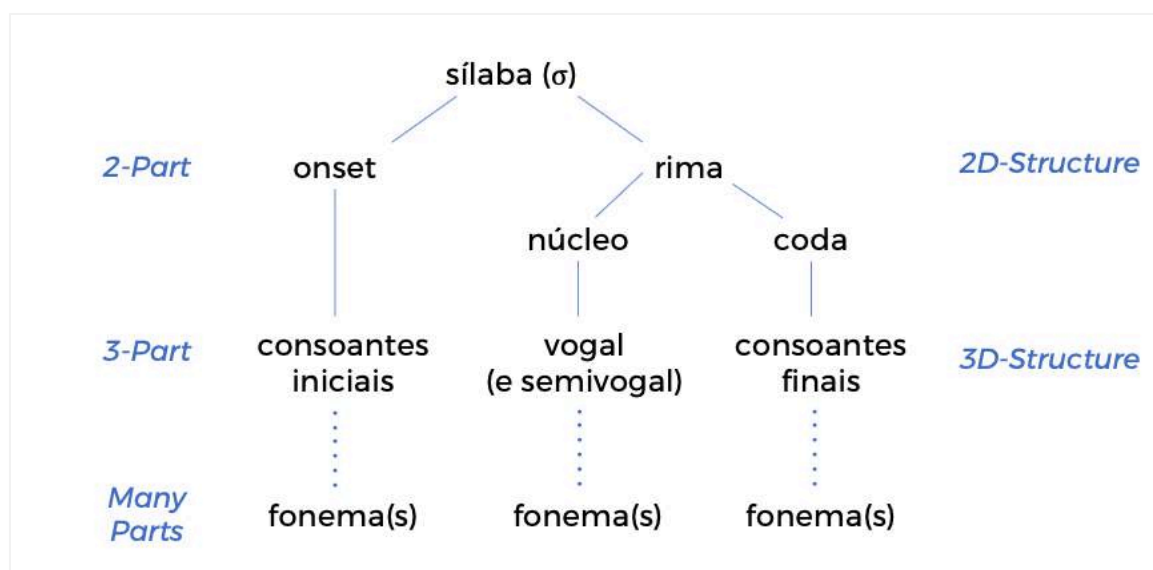


Figura 38 – Representação da hierarquia silábica adaptada à ideia de *explicitação incompleta*.
Fonte: adaptado de Seymour (1997).

Para Seymour (1997), durante a aquisição da escrita e o desenvolvimento ortográfico, há o seguinte movimento em termos do acesso consciente à fonologia: primeiro, uma tendência natural de processamento do tipo *top-down*, partindo de unidades maiores como a sílaba e passando pelas unidades intrassilábicas (camada *2-Part* conforme a Figura 38), o que é evidenciado pelos estudos de I. Y. Liberman et al. (1974) e Treiman (1992); em seguida, quando iniciado o processo de aquisição de um sistema de escrita alfabética, é como se esse movimento natural fosse interrompido momentaneamente, fazendo com que a sensibilidade à camada que corresponde à ramificação silábica (camada *3-Part*) seja precedida pela percepção da segmentação fonêmica da língua (camada *Many Parts*). Isso acaba implicando em uma percepção dos fonemas como unidades isoladas, que não necessariamente relacionam-se em termos de restrições fonotáticas. Visto que, nesse formato, não se tem acesso à informação de que estão *presos*, ou *ligados*, ou ainda *associados* a

nada, e, por isso, não estão submetidos a uma organização estrutural explícita, acabam sendo mais suscetíveis à sua omissão, à sua substituição ou a alterações na sua ordem. Essa dissociação é expressa, na Figura 38, pelas linhas pontilhadas que ligam a camada *3-Part* à camada *Many Parts* (o que configura uma adaptação da figura original).

Nesses termos, a pergunta feita anteriormente – *o que sustenta que haja uma explicitação incompleta especificamente na aquisição da escrita de estruturas silábicas complexas?* – seria assim respondida: primeiro, porque essa é a tendência natural seguindo o curso de desenvolvimento cognitivo mais abrangente conforme proposto por Karmiloff-Smith (1994), segundo o qual existem níveis cada vez mais sofisticados de acesso consciente que o sistema cognitivo vai atingindo, passando de um nível implícito, formato com o qual todo conhecimento inicia e por meio do qual atividades epilinguísticas são realizadas, até diferentes níveis de explicitação dos conhecimentos, momento em que atividades de caráter metalinguístico são passíveis de realização. Depois, o que explica de forma mais específica a existência da explicitação incompleta é a interposição de um processamento *bottom-up* em meio ao processo de aquisição da escrita alfabética, que seguia o curso de um processamento *top-down*. Este fato é ao mesmo tempo condição para que a explicitação ocorra e causa para que ela seja incompleta, uma vez que, conforme o modelo de Seymour (1997), se cria uma lacuna entre fonemas (ou grafemas a eles relacionados) e constituintes silábicos. A tarefa do sujeito, então, é de *relacionar as representações*, de criar vínculos entre elas, processo este que as tornará explícitas. É por essa razão que o MRR é preciso neste ponto: como dito anteriormente, o formato das representações do conhecimento que estão em nível E1 são mais flexíveis em relação ao nível I e, assim, começam a se tornar 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 *Many Parts* precisa associar-se à camada *3-Part*.

No entanto, ainda é necessário questionar se o raciocínio empregado até aqui para explicar não só a omissão enquanto tipo de erro distinto dos demais, mas a motivação dos erros em sílabas complexas como um todo, vale também para o caso das consoantes lateral e nasal na rima ramificada. Se as crianças interpretam de modo distinto essas estruturas do ponto de vista fonológico, a ponto de demonstrarem um comportamento gráfico também diferente em relação às demais consoantes da rima

ramificada, a ocorrência da estratégia de omissão seria contemplada pela argumentação anterior? A resposta que se propõe para essa questão é *parcialmente*, sendo necessário retomar e escrutinar algumas diferenças existentes entre lateral e nasal no que se refere à sua representação na fonologia infantil.

Em relação à lateral, defende-se, neste trabalho, que a interpretação dada pela criança inicialmente é de que há um glide em núcleo ramificado (ou mesmo em coda⁸²), formando um ditongo com a vogal antecedente. Em relação à nasalidade, por outro lado, não existe qualquer segmento ligado à coda tampouco à segunda posição de núcleo; apenas uma vogal nasal em núcleo simples. Assim, a omissão da consoante nasal na escrita, como no exemplo apresentado na Figura 37 ('nuca' para 'nunca'), refletiria a representação fonológica infantil. Todavia, se bem observado, a compreensão não seria exatamente a mesma para a lateral, uma vez que o conflito, neste caso, é mais sobre *a qualidade de segmento* que ocupa a posição pós-vocálica e não tanto sobre *se há algum segmento* em posição pós-vocálica. Possivelmente, essa seja uma das razões pelas quais a substituição seja mais frequente na grafia da lateral (63,64%), ao se comparar com a sua média de frequência na grafia da nasal (36,34%), como discutido a partir da Tabela 23 nesta subseção. Com o mesmo tipo de comparativo feito em relação à frequência da estratégia de omissão entre as diferentes qualidades de rima ramificada analisadas, verifica-se um resultado inverso daquele encontrado para a substituição, como assinala a Tabela 24:

Tabela 24 – Erros de omissão em rima ramificada conforme tipo de coda.

	Nasal (n = 26)	Rótica (n = 26)	Lateral (n = 26)	Fricativa (n = 26)
Média (%)	76,92	11,53	3,84	3,84
Desvio padrão	42,96	32,58	19,61	19,61
Frequência absoluta	28/33	3/33	1/33	1/33

Como é possível perceber, a grafia da nasalidade concentra mais da metade dos erros de omissão registrados na rima ramificada, com 76,92% de média (28 ocorrências de 33 erros classificados como omissão no total). Depois da nasal, a rótica apresenta o segundo percentual mais alto, com 11,53% (3 ocorrências de 33 erros classificados como omissão no total). Lateral e fricativa, por fim, dividem a fatia

⁸² Lembrando que, a depender do autor, os glides no PB podem ser considerados ligados ao núcleo ou à coda (ver subseção 2.3 sobre Teoria da Sílabas).

restante, com 3,84% cada (1 ocorrência para cada de 30 erros classificados como omissão no total). Esses resultados corroboram ainda mais, portanto, a hipótese proposta para as diferenças existentes entre a fonologia infantil e a fonologia adulta (correspondente à do sistema ortográfico), no que concerne ao estatuto de lateral e nasal em rima ramificada, e que é mais este conflito que está na base estratégia de omissão, especialmente quando considerada a grafia da nasalidade.

Neste caso, como propõe Miranda (2017), além de explicitado, o conhecimento fonológico infantil é reestruturado, processo para o qual o MRR de Karmiloff-Smith (1994) novamente oferece suporte⁸³. O modelo propõe a ideia de que seja possível haver redescrição representacional, ou seja, mudança no formato das representações, à medida que delas vai se demandando *atenção cognitiva* – portanto, acesso consciente. Conforme a criança tem acesso às novas formas de representar os segmentos em rima ramificada, via aprendizagem do sistema ortográfico, com um grafema consonantal correspondendo a um fonema consonantal, dela vai sendo demandada a *atenção* sobre as estruturas fonêmicas e silábicas envolvidas, já que há um conflito entre a sua estrutura representacional e a do sistema ortográfico. Essa demanda de atenção pode desencadear, assim, um processo de redescrição representacional, fazendo com que a representação da nasalidade no sistema fonológico infantil se redescreva (ou reestruture) em conformidade com o modelo que lhe é apresentado. As Figuras 39 e 40, que seguem adiante, apresentam uma proposta de formalização do conflito entre a fonologia infantil e a “fonologia do sistema ortográfico”, que resulta no processo de mudança representacional da vogal nasal em núcleo para consoante nasal em coda e do glide em núcleo ramificado para lateral em coda:

⁸³ A Figura 2 deste trabalho, na subseção 2.1.2, sintetiza os aspectos conceituais mais abrangentes envolvidos na mudança representacional.

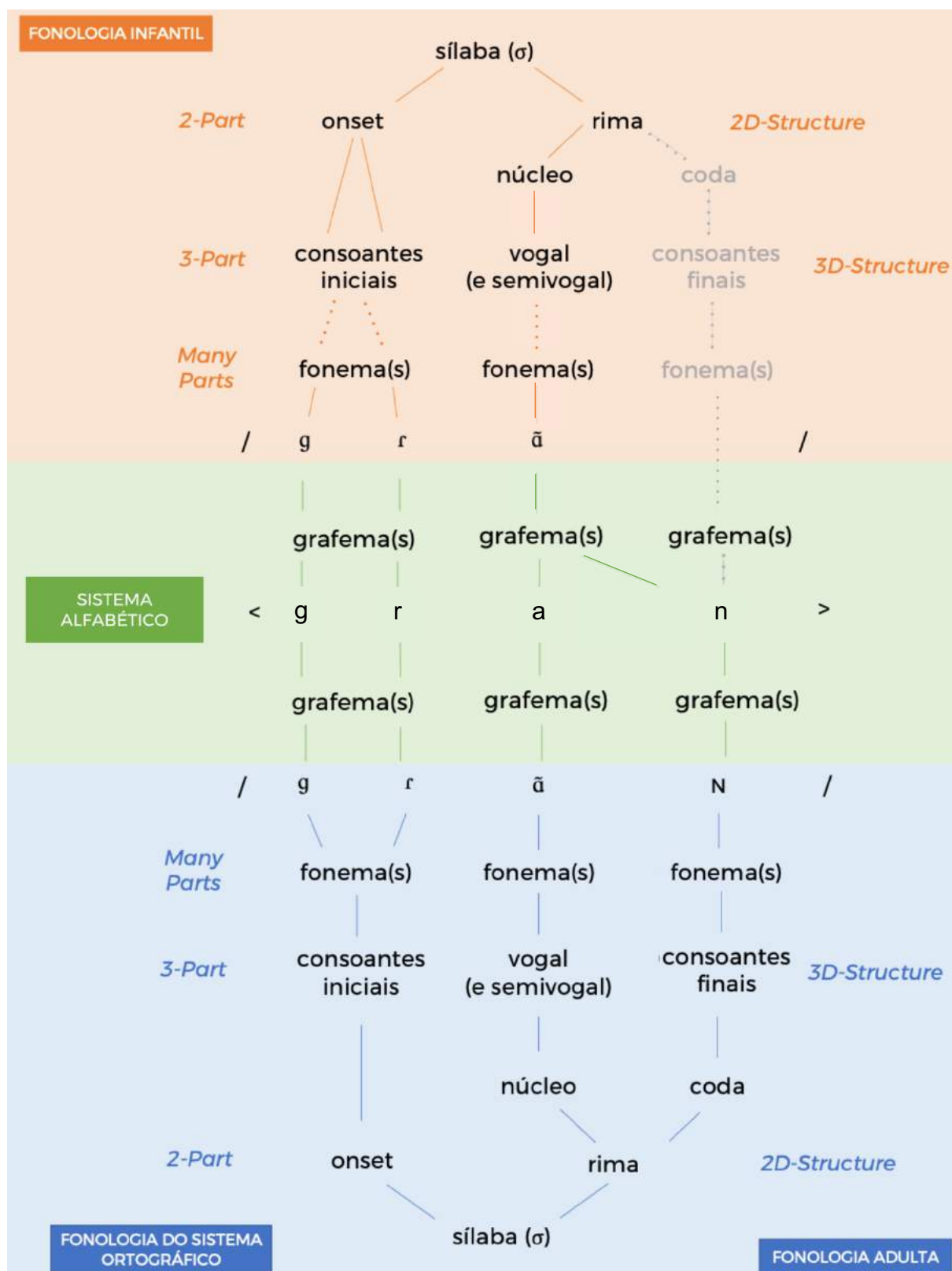


Figura 39 – Formalização do conflito entre a fonologia infantil e a fonologia do sistema ortográfico (correspondente à fonologia adulta alfabetizada), que resulta no processo de mudança representacional da nasalidade medial (exemplo com as sílabas /grã/ vs. /graN/ da palavra 'grande').
Fonte: elaboração própria.

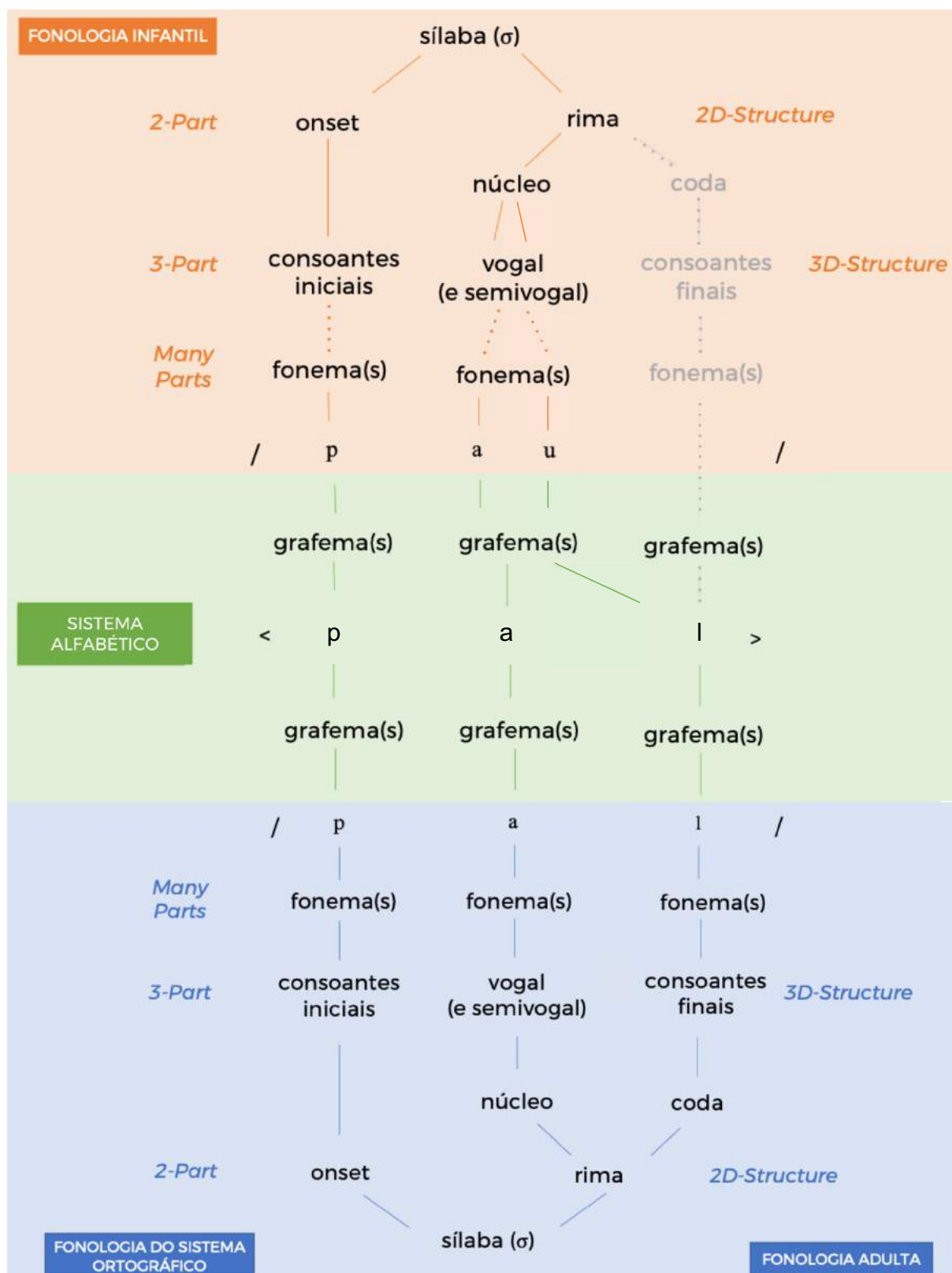


Figura 40 – Formalização do conflito entre a fonologia infantil e a fonologia do sistema ortográfico (correspondente à fonologia adulta alfabetizada), que resulta no processo de mudança representacional do glide em núcleo ramificado derivado de vogal alta (exemplo com as sílabas /pau/ vs. /pal/ da palavra 'palco').

Fonte: elaboração própria.

As Figuras 39 e 40 baseiam-se no modelo adaptado de Seymour (1997) para a representação silábica, com o uso de linhas pontilhadas para representar estruturas ou vínculos entre estruturas em processo de explicitação. Os elementos em cor cinza representam as estruturas que ainda não fazem parte da organização fonológica infantil, mas que, com a aquisição da ortografia, podem passar a incorporar seu inventário, por meio do processo de redescrição/mudança representacional, o qual inclui o processo de explicitação, razão pela qual também são utilizadas linhas pontilhadas. Tanto no caso da nasal quanto no da lateral, a estrutura não existente na fonologia infantil é a coda, em contraste com a fonologia do sistema ortográfico (considerando a proposta de Bisol (1999)).

4.5 Síntese dos resultados: respondendo ao problema geral da pesquisa e sistematizando as conclusões

Diante dos resultados apresentados e das discussões realizadas nas subseções precedentes, há a necessidade de refletir, de forma mais sistematizada, sobre a resposta ao problema geral desta pesquisa, qual seja, *o que motiva as crianças a omitir, inserir, substituir ou alterar a ordem dos segmentos de estruturas com onset e rima ramificados na escrita, se elas já adquiriram fonologicamente tais estruturas?* A hipótese teórica com a qual este trabalho operou se confirma após a descrição, a análise e a discussão dos dados, podendo configurar-se como resposta ao problema de pesquisa? Entendemos⁸⁴ que parcialmente sim, sendo que os resultados obtidos nos levam a considerar outros tipos de motivação para a ocorrência dos erros investigados não previstos pela hipótese original.

Na introdução deste trabalho, anunciou-se a hipótese de que os erros ortográficos em sílabas complexas ocorrem devido a processos cognitivos de conversão de uma informação *implícita* em conhecimento *explícito*, isto é, devido à passagem por níveis de acesso à consciência necessários à descoberta, pela criança, da existência e da forma como se constituem internamente as estruturas silábicas (KARMILOFF-SMITH, 1994; SEYMOUR, 1997). Na medida em que conhecimentos implícitos precisam ser explicitados, pode-se dizer, também, que há a passagem de uma atividade *epilinguística* a uma atividade *metalinguística* (GOMBERT, 1992).

⁸⁴ A primeira pessoa do plural é utilizada nesta subseção por tratar de argumentos defendidos pelas autoras deste trabalho (mestranda e orientadora).

Após verificar e discutir os dados, entendemos que a hipótese é parcialmente corroborada. Concluimos que os erros ortográficos em sílabas complexas produzidos por crianças em fase de aquisição da escrita podem apresentar três tipos de motivação não mutuamente excludentes, expostas a seguir por ordem de importância e de abrangência: (i) a explicitação incompleta da relação entre os constituintes silábicos *onset* complexo e rima ramificada e a camada fonêmica da língua, motivação mais estreitamente associada à hipótese original da pesquisa; (ii) nos casos da grafia de nasal e de lateral em rima ramificada apenas, o conflito entre a fonologia infantil e a fonologia do sistema ortográfico, sendo esta última equivalente à fonologia de adultos alfabetizados, as quais preveem estatutos distintos para as estruturas com nasal e lateral; (iii) nos casos da grafia de nasal e de fricativa em rima ramificada apenas, os conhecimentos relacionados a regras ortográficas contextuais e arbitrárias. As duas primeiras motivações são de natureza fonológica e são predominantes na primeira metade do processo inicial de escolarização – 1º, 2º e 3º anos –, ainda que se estendam, com frequência a cada ano menor, ao longo de todo o 1º Ciclo do Ensino Fundamental. A terceira motivação é de natureza ortográfica e é predominante na segunda metade do processo de escolarização, isto é, nos 4º e 5º anos.

Uma síntese visual do argumento pode ser visualizada na Figura 41, que segue:

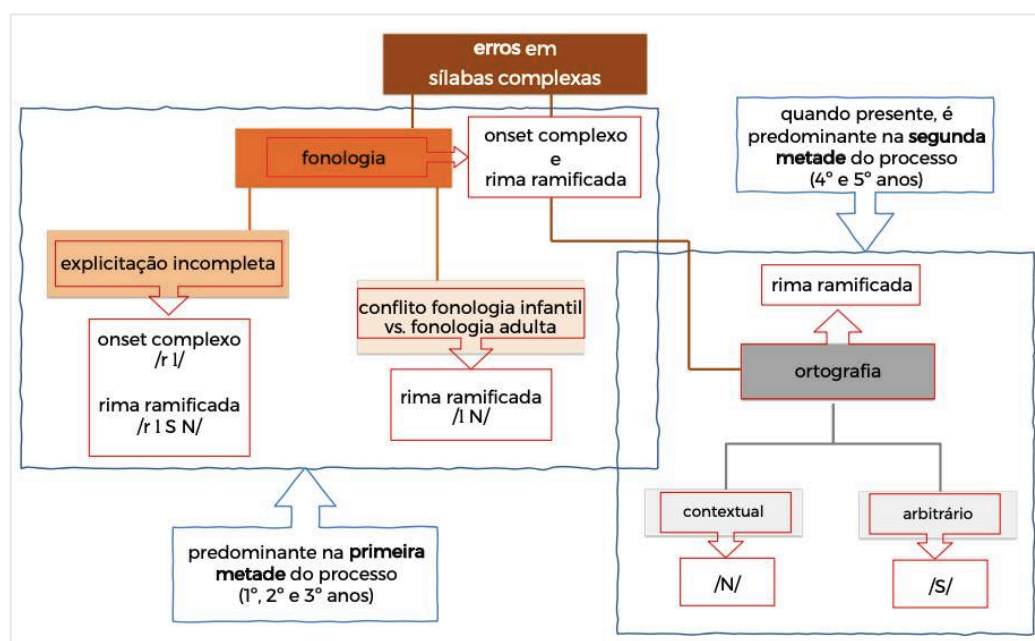


Figura 41 – Síntese visual do argumento da pesquisa.
Fonte: elaboração própria.

Em primeiro lugar, o que se defende com o argumento é que a motivação majoritária dos erros, expressa pelo item (i), não está ligada a uma análise *fonêmica* insuficiente, 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. Basicamente, três abordagens teóricas distintas, mas complementares, sustentam essa ideia: Karmiloff-Smith (1994), por oferecer um modelo abrangente de desenvolvimento cognitivo e que dá conta das gradações envolvidas na abertura ao acesso consciente da informação; Miranda (2017), por argumentar que o principal tipo de conhecimento a ser acessado no curso de aquisição da escrita é a gramática fonológica e que esta pode ser, nesse processo, reestruturada; Seymour (1997), por expor o tipo de processamento envolvido na aquisição da escrita alfabética e como ele envolve particularmente a organização interna da sílaba, com base em considerações de Treiman (1992) sobre a consciência silábica e intrassilábica.

A partir de Miranda (2017), cuja abordagem tem formalização expressa pela Figura 2 deste trabalho, podemos presumir que o conhecimento fonológico, construído e organizado de forma progressiva, tácita e implícita durante os primeiros anos de vida de uma criança (CHOMSKY, 1988; KIPARSKY; MENN, 1977; MACKEN, 1996), é retomado no processo de aquisição da escrita alfabética como objeto de atenção especial do sujeito, uma vez que o princípio que torna operacional a escrita alfabética é a relação entre unidades fonológicas (os fonemas, particularmente) e gráficas (LANDSMANN, 1995; GELB, 1952). Nesse sentido, considera-se a escrita alfabética, então, como uma substância secundária de atualização da língua, sendo a fala substância primária (SAUSSURE, 2012 [1912]; LYONS, 1968). Isso cria condições para sustentar que, através da aquisição da escrita, a gramática fonológica construída pela criança ao longo do seu desenvolvimento pré-escolar e realizada até então via fala, pode ser acessada de forma relativamente consciente, isto é, ser progressivamente explicitada, e, eventualmente, ser também manipulada e reestruturada (KARMILOFF-SMITH, 1994).

Seymour (1997), por sua vez, observa que, nesse processo de retomada do conhecimento fonológico, que corresponde à sua progressiva abertura ao acesso consciente, a tendência esperada é de haver um processamento do tipo *top-down*, uma vez que as crianças pré-escolares (não-alfabetizadas) já demonstram capacidade de segmentar a linguagem falada tanto em sílabas inteiras quanto nas

unidades intrassilábicas *onset* e rima (I. Y. LIBERMAN et al, 1974; TREIMAN, 1992). No entanto, Treiman (1992) aponta que as crianças não-alfabetizadas ainda não são capazes de segmentar a linguagem em unidades menores que *onset* e rima, o que inclui não somente a camada fonêmica, *mas também as ramificações desses constituintes silábicos*. Como estudos clássicos da área demonstram (MORAIS et al., 1979; READ et al., 1986), a consciência fonêmica é desenvolvida quando as crianças já estão com o processo de aquisição da escrita alfabética em curso, uma vez que ela demanda, em razão de seu princípio operacional, a sensibilidade à existência de fonemas.

Considerando esse roteiro, Seymour (1997) defende que, quando iniciado o processo de aquisição de um sistema de escrita alfabética, é como se o movimento natural *top-down* fosse interrompido, fazendo com que a sensibilidade à camada que corresponde à ramificação silábica (camada *3-Part* das Figuras 6 e 38) seja precedida pela percepção da segmentação fonêmica da língua (camada *Many Parts* das Figuras 6 e 38). Assim, um processamento do tipo *bottom-up* se interpõe ao *top-down*. Em citação já exposta na subseção 2.1.3, Seymour (1997, p. 327, tradução nossa) salienta que:

[...] Isso não significa que estruturas de ordem superior não sejam importantes no desenvolvimento da alfabetização, apenas que a sua importância emerge posteriormente ao invés de inicialmente. A razão para isso é que estruturas de ordem superior não se tornam relevantes até o estágio em que a estrutura ortográfica passa a se formar.

Este fato acaba implicando em uma percepção dos fonemas como unidades isoladas, que não necessariamente relacionam-se em termos de restrições fonotáticas. Visto que, nesse formato, não se tem acesso claro à informação de que estão *presos*, ou *ligados*, ou ainda *associados* a nada, e que, por isso, não estão submetidos a uma organização estrutural explícita – papel por excelência desempenhado pelas sílabas dentro da gramática fonológica –, acabam sendo mais suscetíveis à sua omissão, à sua substituição ou a alterações na sua ordem. Essa espécie de dissociação, que corresponde ao termo *explicitação incompleta*, é formalizada na Figura 32 deste trabalho, através linhas pontilhadas que ligam a camada *3-Part* à camada *Many Parts* (o que configura uma adaptação da imagem original, apresentada na Figura 6).

É nesse sentido específico, então, que se faz necessária a redescrição de informações implícitas – construídas de forma tácita, inconsciente e manifestadas em

atividades de caráter epilinguístico ao longo do desenvolvimento fonológico – em conhecimento explícito – manifestado por atividades de caráter metalinguístico realizadas em diferentes níveis de acesso consciente à informação antes implícita (KARMILOFF-SMITH, 1994; GOMBERT, 1992). Em outras palavras, é nesse sentido que se faz necessária a explicitação: especificamente, a explicitação da relação entre os constituintes silábicos *onset* complexo e rima ramificada e a camada fonêmica da língua. A tarefa do sujeito, então, é de *relacionar as representações*, de criar vínculos entre elas, processo este que as tornará explícitas.

Essa redescritção ou explicitação da informação fonológica ocorre de forma progressiva e multinivelada ao longo do processo de aquisição da escrita, conforme o Modelo de Redescritção Representacional (MRR) de Karmiloff-Smith (1994), e sob as condições de haver escolarização ou instrução formal, conforme demonstrado na subseção 4.3 deste trabalho. De acordo com o MRR, uma estrutura que carece de explicitação pode estar em diferentes níveis de abertura ao acesso consciente. Ela pode ser totalmente inacessível (*nível I*), parcialmente acessível (*níveis E1 e E2*) e totalmente acessível (*nível E3*). Por isso, mais uma vez, é possível falar em explicitação incompleta. Entendemos que o tipo de “lacuna” existente entre a camada fonêmica e a camada correspondente às ramificações silábicas corresponde a uma representação em nível E1 de explicitação: já *disponível* à criação de vínculos com outras representações, mas com esse processo ainda não efetivado. É um momento do percurso de desenvolvimento em que determinadas estruturas não estão totalmente abertas ao acesso consciente – nem totalmente inacessíveis, como no nível I. Isso explica, em boa medida, não só por que os erros (orto)gráficos em sílabas complexas podem acontecer, sobretudo os de natureza fonológica, mas também por que os acertos coexistem com os erros em um mesmo texto ou, ainda, por que coexistem diferentes tipos de erro de natureza fonológica em um mesmo texto, como abordado na subseção 4.4 deste capítulo.

A ideia de explicitação incompleta explica também a predominância da omissão do segmento ligado ao constituinte ramificado, tanto em *onset* quanto na rima, em relação aos outros tipos de erro de natureza fonológica, resultado observado nas Tabelas 19 e 21. Diante da parcial clareza e parcial acessibilidade à organização estrutural a que está submetida uma sequência segmental, identificar e consequentemente registrar todas as unidades dessa sequência é mais difícil. A

solução mais fácil e imediata, assim, é omitir o segmento do que produzi-lo em posição ou qualidade alterada (o caso de metátese, epêntese e substituição).

Além disso, consideramos que outro aspecto, de ordem perceptual, pode se associar à explicitação incompleta e contribuir como causa à proeminência da omissão no conjunto de erros de natureza fonológica em sílabas complexas. I. Y. Liberman et al. (1974), ao argumentar em favor de que a segmentação da linguagem em fonemas é mais custosa e tardia do que a segmentação em sílabas, chamam a atenção para a maneira como a fala é percebida e codificada pelo sistema linguístico humano, o que fazem baseados no estudo de A. M. Liberman, Cooper, Shankweiler e Studdert-Kennedy (1967). Segundo estes autores, “os segmentos fonêmicos são codificados no nível acústico em unidades maiores aproximadamente de tamanho silábico” (I. Y. LIBERMAN et al, 1974, p. 203). Isso significa que a informação acústica a respeito das consoantes adjacentes a uma vogal fica contida na própria vogal e, assim, as pistas para identificar as fronteiras silábicas são veladas. Dessa maneira, explicitar e segmentar uma cadeia fonêmica com o número, o tipo e a posição correta das unidades que a compõe, como a presente em sílabas do tipo /mar/ e /tras/, é bastante desafiador na medida em que ela é percebida acusticamente como uma unidade só.

Além de contribuir com a primeira parte do argumento principal deste trabalho, a omissão também revela aspectos importantes relacionados à segunda parte do argumento – item (ii) –, que diz respeito à motivação dos erros nos casos das grafias de nasal e de lateral em rima ramificada.

Defendemos, ao longo deste capítulo, que as crianças interpretam a composição da rima ramificada de forma distinta de como interpretam os adultos alfabetizados, ou de como interpreta o sistema ortográfico. Essa hipótese, levantada originalmente por Miranda (2009a, 2009b, 2012, 2018, 2019), leva em conta a precocidade de aquisição do que seriam as codas nasal e lateral ao se comparar com as codas rótica e fricativa, cuja aquisição é mais tardia (cf. MEZZOMO, 2004). Assim, abre-se a possibilidade para pensar que a fonologia infantil entende como segmentos licenciados para a coda apenas /r/ e /s/; em relação àquilo que para os adultos são um segmento lateral e um segmento nasal em coda (cf. BISOL, 1999), para as crianças são, respectivamente, um glide em núcleo ramificado (derivado de vogal

alta), formando um ditongo com a vogal antecedente, e uma vogal nasal em núcleo simples.

Dessa maneira, por meio dos resultados expostos na subseção 4.4, especialmente a partir da Tabela 24, que mostra a predominância da omissão como estratégia para a grafia da nasal pós-vocálica, entendemos que este tipo de erro reflete a representação fonológica infantil para a nasalidade, uma vez que, para as crianças, parece não existir qualquer segmento ligado à coda. Um pouco diferente é o caso da lateral, pois, para as crianças, existe um segmento após a vogal, ainda que não esteja ligado à coda e não corresponda, na subjacência, ao segmento /l/. Com isso, concluímos que o conflito, para a lateral, é mais sobre a *qualidade de segmento* que ocupa a posição pós-vocálica e não tanto sobre se *há algum segmento* em posição pós-vocálica, sendo este o caso da nasal. Por meio dessa observação, também pudemos interpretar por que a substituição é mais frequente como estratégia para a grafia da lateral, resultado expresso pela Tabela 23.

Outros resultados apresentados ao longo do capítulo 4 também colaboram com a construção e a sustentação do argumento relacionado à motivação distinta dos erros em nasal e lateral, e assim, também, permitem-nos atestar que a variável *tipo de segmento* tem importante papel para a ocorrência de erros ortográficos em rima ramificada, diferente daquilo que se observou para o *onset* complexo.

Com a Tabela 8, verificou-se que os índices de erros na rima ramificada concentram-se na grafia das consoantes lateral e nasal, comparadas à grafia de rótica e fricativa. Ao serem testadas as diferenças das médias de erro em cada tipo de segmento, aferiu-se que lateral e nasal não têm diferença entre si, mas ambas apresentam diferença se colocadas em oposição à fricativa e à rótica. Da mesma forma, fricativa e rótica não têm diferença entre si, mas apresentam diferença se contrastadas com lateral e nasal. Assim, fica demarcada a diferença de comportamento das crianças em suas escritas em relação a esses dois grupos de segmentos – de um lado lateral e nasal e de outro fricativa e rótica – o que, sob nossa interpretação, reflete o estatuto diferenciado que os dois grupos têm, na fonologia infantil, em relação à posição que ocupam na sílaba.

Na continuidade desse raciocínio, os resultados presentes na Tabela 10 mostram que, se forem computados apenas os erros envolvendo as consoantes fricativa e rótica, a frequência de erros em rima ramificada cai de 10,90% para 2,55%,

ou seja, mais de dois terços dos erros concentram-se na grafia de lateral e nasal. Ao atestar-se que a diferença entre esses percentuais é significativa, não somente se corrobora ainda mais a ideia de que há diferença entre a fonologia infantil e a fonologia adulta no que se refere aos segmentos licenciados para a coda, mas também que a ramificação silábica não é tanto um problema para as crianças quanto é o conflito gerado pelas diferenças fonológicas do sistema infantil e do sistema ortográfico (equivalente ao adulto).

Isso não implica desconsiderar a ideia de explicitação incompleta, que compõe a primeira parte do argumento deste trabalho, como uma motivação também presente na base dos erros envolvendo a grafia de lateral e nasal. Neste caso, o que ocorre é que, à medida que a relação entre os constituintes silábicos e a camada fonêmica vai sendo explicitada, o conflito entre a fonologia infantil e a fonologia do sistema ortográfico passa a ser objeto de atenção cognitiva da criança. Essa atenção pode desencadear, assim, um processo de redescrição representacional, nos moldes do MRR de Karmiloff-Smith (1994), fazendo com que a representação da nasalidade e do glide no sistema fonológico infantil se redescreva (ou reestruture) em conformidade com o modelo que lhe é apresentado, assim como defende Miranda (2009a, 2009b, 2012, 2017, 2018, 2019).

Essa questão conduz a um importante esclarecimento: ainda que lateral e nasal não sejam interpretadas como coda pelas crianças ao longo do desenvolvimento fonológico até o período inicial de aquisição da escrita, consideramos essas consoantes dentro do escopo de análise das sílabas complexas, porque essa é a *interpretação do sistema ortográfico*, e é a essa interpretação que as crianças se direcionam por meio do processo de aquisição da escrita ortográfica. O sistema ortográfico é o objeto em aquisição pelas crianças – e sua aquisição só se completa se elas passam por sucessivas reestruturações internas até ajustarem suas hipóteses sobre o objeto de acordo com o que ele de fato é e como ele de fato funciona, o que pode incluir, como argumentamos, a (re)organização da fonologia. A esse respeito, e em citação já feita outrora neste trabalho, Ferreiro e Teberosky [1984] (1999, p. 33, grifos das autoras) notam que:

[...] Na teoria de Piaget, o conhecimento objetivo aparece como uma aquisição, e não como um dado inicial. O caminho em direção a este conhecimento objetivo não é linear: não nos aproximamos dele passo a passo, juntando peças de conhecimento umas sobre as outras, mas sim através de reestruturações globais, algumas das quais são 'errôneas' (no que

se refere ao ponto final); porém, ‘construtivas’ (na medida em que permitem aceder a ele). Esta noção de *erros construtivos* é essencial.

Nessa perspectiva, como já dito, o conhecimento objetivo, em direção ao qual a criança caminha e do qual estamos tratando é o sistema ortográfico. Isso implica, entre outras coisas, que ela caminha em direção a uma fonologia diferente da sua em alguns aspectos, sendo os erros (orto)gráficos indícios das reestruturações globais por meio das quais a criança percorre esse caminho. Por meio da análise da grafia das sílabas complexas durante a aquisição da escrita, percebemos uma série de conhecimentos, os quais envolvem hipóteses que as crianças formulam sobre o objeto em aquisição, sendo reestruturados nesse processo: a estrutura fonológica, particularmente a sílaba, seus moldes e os segmentos licenciados para cada posição da hierarquia; o grau de acesso do sujeito à informação fonológica, ou o grau de explicitação dessa informação no sistema cognitivo do sujeito; a direcionalidade do processamento na explicitação da informação fonológica – de *top-down* para *bottom-up*; a estrutura ortográfica (SEYMOUR, 1997), que se reorganiza em um processo constante e cíclico ao longo da aquisição da escrita, especialmente por meio da estipulação das relações fonografêmicas, o que inclui informações tanto de ordem fonológica, como combinações de segmentos licenciados conforme a posição que ocupam na sílaba, quanto informações de ordem ortográfica propriamente dita, como as regras que estabelecem relações biunívocas, contextuais e arbitrárias – a relação da fonotática com a grafotática.

É à propósito da estrutura ortográfica, e particularmente das regras que regem as relações fonografêmicas, que chegamos ao terceiro tipo de motivação possível para os erros em sílabas complexas, conforme o item (iii) do argumento apresentado no início desta subseção. Este tipo de motivação tem um alcance restrito no contexto das sílabas complexas, pois só é possível para as consoantes nasal e fricativa em rima ramificada, uma vez que somente elas possuem complexidade ortográfica, além de fonológica: a grafia de /N/ é regulada por regra contextual e a grafia de /S/ é regulada por uma regra arbitrária. Considerando esse escopo de análise, observamos, especialmente por meio dos resultados exibidos nas Tabelas 17 e 18, que os conhecimentos de natureza ortográfica têm papel significativo na ocorrência dos erros em sílabas complexas, porém em contexto gráfico restrito e, sobretudo, mais tardiamente no processo de aquisição da escrita ou de escolarização. Isto é, os erros de natureza fonológica, em /N/ e /S/, são predominantes na primeira metade do 1º

Ciclo do EF (1º, 2º e 3º anos), já os erros de natureza ortográfica são predominantes na segunda metade (4º e 5º anos).

Com isso, são corroboradas ideias, já defendidas por estudos do GEALE (cf. MIRANDA, 2013, 2020; NEY, 2012; CUNHA, 2004), de que o conhecimento fonológico é o insumo principal das crianças no período inicial do processo de aquisição da escrita, e que esse tipo de conhecimento vai dando espaço ao conhecimento ortográfico à medida que os anos escolares vão avançando, ou seja, à medida que avança o próprio processo de aquisição da escrita. Nesse sentido é que se pode dizer que o conhecimento fonológico é condição para o surgimento do conhecimento ortográfico. O modelo de desenvolvimento ortográfico de Seymour (1997) é mais uma vez oportuno e auxilia a sustentar esse ponto: para o autor, a formação de uma estrutura ortográfica que envolva a organização de regras que regem as relações múltiplas entre fonemas e grafemas só é possível se antes o que ele chama de Processo Alfabético for mobilizado.

Esse resultado revelou também que a variável *tipo de erro (orto)gráfico* mostra-se bem articulada com a variável *ano escolar*, neste caso representada pelos dois grupos do 1º Ciclo do EF. Ainda que não se tenha atestado que, de modo geral, existem diferenças significativas entre os anos escolares no que se refere à frequência de erros (orto)gráficos em rima ramificada, é possível entender que a escolarização, se não influencia a frequência dos erros em geral, influencia o tipo de erro em rima ramificada. Um pouco diferente é o caso do *onset* complexo, que não comporta a variável *tipo de erro (orto)gráfico* no sentido mais abrangente, por só envolver complexidades de ordem fonológica, mas que apresentou diferenças significativas entre os anos escolares na frequência geral de erros, particularmente entre o 1º e o 3º anos. Apesar de não terem sido verificadas diferenças estatisticamente significativas, é importante notar que o resultado descritivo também apontou uma média de erros bastante baixa no 5º ano (3,05% de erros).

Notamos, primeiro, que a diferenças se localiza justamente entre dois marcos do 1º Ciclo do Ensino Fundamental, guardando certa semelhança com a rima ramificada, no que concerne às diferenças observadas entre os distintos períodos do 1º Ciclo. Além disso, entendemos o resultado como um indício de que, ao longo do processo de aquisição da escrita – o qual acompanha o avanço do processo de escolarização –, há uma clara diferença de compreensão das crianças a respeito da grafia do *onset* complexo, sendo o 1º ano o período onde são evidenciadas mais

dúvidas a respeito da representação gráfica dessa estrutura silábica, em consonância com os estudos de Santos (2013) e de Vicente (2018), que analisam dados de outras variedades do Português (ver subseção 2.4.2.2). Essa característica é esperada considerando que o tipo de erro em *onset* complexo está restrito à natureza fonológica: complexidades deste nível são um problema no início do processo de aquisição, especialmente em razão da explicitação incompleta, cuja causa está ligada à “descoberta” que fazem as crianças da camada fonêmica da língua, que por sua vez está alinhada à “descoberta” do princípio alfabético. Contudo, tendem a ser resolvidas à medida que o processo avança, uma vez que esse avanço implica em uma progressiva e cada vez mais completa explicitação das unidades fonológicas bem como de sua organização.

Se bem observado, mesmo em se tratando do *onset* complexo, é possível perceber uma associação entre a variável ano escolar e tipo de erro (orto)gráfico. Portanto, o que propomos, nesse sentido, é que a frequência de erros em uma estrutura tende a apresentar diferenças significativas ao longo do processo de escolarização se a sua natureza for fonológica, particularmente se a complexidade, neste nível, envolver apenas aspectos de ordem da explicitação incompleta; por outro lado, não é possível afirmar essa tendência se os erros em uma estrutura apresentarem a possibilidade do conhecimento ortográfico como motivação para sua ocorrência, além dos aspectos fonológicos – os quais, no caso da rima ramificada, envolvem não apenas a explicitação mas também o conflito entre a fonologia infantil e a fonologia do sistema ortográfico.

Outra característica observada relacionada à variável ano escolar, para ambas as estruturas silábicas, foi o resultado não significativo obtido nos testes de correlação, ou seja, a frequência de erros não apresenta associação negativa significativa com o avanço da escolarização. Interpretamos esse resultado como uma explicação tanto para a presença de uma curva em U nos dados de *onset* complexo no que se refere à frequência de erros (orto)gráficos conforme o ano escolar, quanto para a ausência dessa curva nos dados de rima ramificada. Na medida em que os erros não *necessariamente* diminuem à medida que avançam os anos escolares, curvas em U podem tanto surgir quanto não surgir no decorrer do processo. Se surgirem, são perfeitamente esperadas como parte do curso de desenvolvimento, conforme sustenta a literatura (cf. STRAUSS; STAVY, 1982; KARMILOFF-SMITH, 1994).

Ponderando sobre todos os resultados apresentados e as conclusões feitas a partir deles até então, podemos compreender, por fim, as diferenças e semelhanças gerais existentes entre *onset* complexo e rima ramificada no tocante à sua grafia por crianças em fase de aquisição da escrita – a variável *tipo de estrutura silábica*. Conforme a Tabela 7 exibiu, constatou-se uma diferença significativa entre erros envolvendo a grafia da rima ramificada, com média de 10,90%, em relação aos erros em *onset* complexo, com média de 5,40%, resultado que vai ao encontro daquele obtido por Vicente (2018) ao analisar dados de escrita de crianças moçambicanas que têm o Português como L2. Entendemos que essa diferença, nos dados desta pesquisa, se deve às duas diferentes motivações a mais que os erros em rima ramificada podem apresentar, expressas no argumento deste trabalho pelos itens (ii) e (iii): o conflito entre a fonologia infantil e a fonologia adulta alfabetizada (equivalente à do sistema ortográfico) em relação ao estatuto de lateral e nasal em posição pós-vocálica medial, e o conhecimento ortográfico, que envolve as regras que regem as relações múltiplas de /N/ e /S/. No caso do *onset* complexo, apenas a motivação expressa pelo item (i), relacionada à ideia de explicitação incompleta, pode ocorrer. Assim, no escopo das sílabas complexas e de sua aquisição escrita, o tipo de estrutura silábica tem importante papel sobre a ocorrência de erros (orto)gráficos, sendo que a rima ramificada apresenta maior complexidade que o *onset* complexo.

Essa conclusão conduz, finalmente, à discussão sobre as relações de simetria e de assimetria que podem ser estabelecidas entre a aquisição fonológica e a aquisição da escrita, um dos objetivos desta pesquisa. A ideia de que a rima ramificada apresenta maior complexidade em relação ao *onset* revela uma clara assimetria entre os dois processos de aquisição (cf. MEZZOMO, 2004; RIBAS, 2002; LAMPRECHT et al., 2004). As razões para isso, expostas no parágrafo anterior, auxiliam a elucidar a diferença de resultado, na medida em que observamos as *diferentes naturezas dos processos*.

No que se refere à motivação mais basilar para a ocorrência de erros (orto)gráficos em sílabas complexas, a explicitação incompleta, já está implicada a ideia de que a aquisição da escrita envolve, em grande parte, a retomada de conhecimentos implícitos na mente já construídos em um período anterior (MIRANDA, 2017), uma vez que só pode ser explicitado um conhecimento pré-existente. Diferente da aquisição fonológica, portanto, a aquisição da escrita trabalha com representações já construídas e, no máximo, pode reestruturar essas representações.

O mesmo tipo de observação pode ser feita no tocante ao segundo tipo de motivação para os erros em sílabas complexas, isto é, o conflito entre a fonologia infantil e a fonologia adulta. A aquisição da escrita mobiliza o conhecimento fonológico infantil já construído até então e, neste caso, o coloca em conflito com o conhecimento fonológico que baliza o novo sistema em aquisição, levando à reestruturação das representações específicas mobilizadas no conflito – no caso analisado, lateral *versus* glide (derivado de vogal alta) e nasal *versus* vogal nasal em rima ramificada.

Dessa maneira, é importante notar que, mesmo nas motivações que consideramos de natureza fonológica, a aquisição da escrita ainda assim apresenta diferenças com a aquisição da fonologia. Em outras palavras, a primeira não irá replicar processos da segunda, pois os processos têm identidade própria, ainda que envolvam a atualização do mesmo tipo de conhecimento. Além disso, mesmo que sejam encontradas estratégias semelhantes na produção falada e escrita das estruturas silábicas complexas, tanto em termos de qualidade quanto em termos de quantidade (omissão, metátese, epêntese e substituição), elas não podem ser interpretadas da mesma forma, porque, mais uma vez, embora a qualidade do *conhecimento* seja a mesma, a qualidade da *tarefa cognitiva* não é.

Como discutido na fundamentação teórica deste trabalho, na elaboração da gramática fonológica, a criança precisa construir e incorporar representações fonológicas que se organizam hierárquica e contrastivamente. Nos termos de Karmiloff-Smith (1994), o formato de armazenamento desse conhecimento é inicialmente implícito, isto é, não está aberto ao sistema cognitivo para o acesso consciente do sujeito, motivo pelo qual o conhecimento não pode ser manipulado nem analisado. Trata-se, assim, de uma tarefa epilinguística (GOMBERT, 1992). Na aquisição da escrita, o sujeito volta-se para a gramática fonológica que *já foi* construída, particularmente à camada fonêmica, e a ela relaciona representações gráficas correspondentes. O acesso consciente torna-se, neste caso, necessário à decomposição das unidades fonológicas até as unidades mínimas que são os fonemas. Trata-se, portanto, de uma tarefa metalinguística (GOMBERT, 1992).

Esse mesmo tipo de raciocínio é sustentado por estudos que investigam, por exemplo, o desenvolvimento da consciência segmental, como o de Alves (2012), revisado na subseção 2.4.2.3, que observa uma assimetria entre a emergência de propriedades segmentais no desenvolvimento fonológico e o papel que elas exercem

em tarefas de cunho metalinguístico – leitura, escrita e testes de consciência segmental. A autora conclui que

dependendo da tarefa a executar, as crianças parecem explorar de forma diferente o conhecimento de que dispõem. Os resultados deste estudo reflectem, portanto, a interferência de propriedades segmentais, tanto no desempenho da consciência segmental como em tarefas de leitura e de escrita, no entanto, não preditas pelos dados do desenvolvimento segmental (ALVES, 2012, p. 415).

A diferença entre os processos, relacionada à tarefa cognitiva que impõem ao sujeito, também explica por que são encontrados tantos acertos e poucos erros na escrita das sílabas complexas – os erros sempre com suas médias acompanhadas de um alto valor de desvio padrão –, e por que na fala ocorre o inverso: nesta, até que a estrutura seja efetivamente adquirida e esteja disponível em termos representacionais, a sua produção conforme o alvo é praticamente impossível; naquela, como as estruturas já estão adquiridas, é preciso resolver “apenas” o acesso a elas, o que viabiliza os acertos e deixa, esporadicamente, *vazar* erros (ABAURRE, 1999). É por essa razão, ademais, que os erros (orto)gráficos são considerados, neste tipo de estudo, como pista, indício, vestígio da fonologia (MIRANDA, 2010, 2017) – jamais produto direto.

Tendo tudo isso em vista, como exposto na subseção 2.4.2.3, deveria ser relativizada a tentativa de traçar uma cronologia de aquisição das estruturas silábicas na escrita, expectativa alimentada pelos estudos revisados de Ferreiro e Zamudio [2008] (2013), de Vicente (2018) e de Soares (2016), esta que diz:

[...] pouco se sabe sobre a ordem de aquisição dos demais padrões silábicos na aprendizagem da escrita pelas crianças [...]. No entanto, dados de estudos desenvolvidos sem esse objetivo precípua [...] e também práticas de alfabetização confirmam, de certa forma, ordem semelhante à identificada na aquisição da linguagem [...] (SOARES, 2016, p. 314).

A tentativa de elaboração de uma sequência de aquisição de padrões silábicos na escrita perde seu sentido na medida em que, conforme o argumento principal desta pesquisa, não é mais a ramificação silábica em si a causa da complexidade e dos consequentes erros (orto)gráficos, tampouco a sua não-disponibilidade no sistema linguístico.

O último tipo de motivação proposto para a ocorrência dos erros em sílabas complexas, que se restringe, na verdade, à grafia de /N/ e /S/ em rima ramificada, também contribui à discussão sobre as relações entre a aquisição fonológica e a aquisição da escrita. Diferente das outras duas motivações, que resguardam aspectos

mais sutis na revelação de assimetrias, esta motivação diz respeito a um conhecimento que não está ligado, de forma mais estreita, à fonologia e ao sistema linguístico como um todo. O conhecimento ortográfico concerne, em alguma medida, a convenções estabelecidas sem necessária ligação com o funcionamento orgânico do sistema linguístico, guardando também relações com a etimologia e com a quimera de critérios acumulada a partir dos diferentes acordos ortográficos realizados ao longo da história da língua portuguesa (LEMLE, 1987; MORAIS, 2002; FARACO, 2012; MIRANDA, 2020). Neste caso, a assimetria entre os processos fica mais evidente, isto é, em virtude das diferentes naturezas de conhecimento que podem ser mobilizadas em cada um.

Em resumo, na medida em que fala e escrita atualizam o mesmo tipo de conhecimento, é que pode se esperar alguma relação simétrica entre os diferentes cursos de aquisição – e, mesmo nesse caso, a interpretação dada aos resultados provavelmente será diferente. Porém, sobretudo por demandarem tarefas cognitivas de qualidade distinta, é que relações assimétricas são também e talvez mais ainda esperadas. Além disso, a aquisição da escrita inclui a aprendizagem de um sistema ortográfico, que mobiliza conhecimentos não necessariamente relacionados ao funcionamento do sistema linguístico.

5 Considerações finais

Esta pesquisa buscou descrever e analisar a grafia de *onset* e de rima ramificados em textos de crianças em fase de aquisição da escrita (1º a 5º ano do Ensino Fundamental), estabelecendo relações com a aquisição fonológica dessas estruturas e com a teoria fonológica, a fim de compreender a motivação dos erros (orto)gráficos que as crianças produzem nas estruturas silábicas complexas.

A hipótese de partida com a qual o trabalho operou é a de que erros em sílabas complexas ocorrem não porque existem estruturas que ainda precisam ser incorporadas ao inventário da criança, pois isso já ocorreu durante o período de aquisição da fala. Antes, o que majoritariamente está na base desses erros seriam processos cognitivos de conversão de uma informação *implícita* em conhecimento *explícito*, isto é, a passagem por níveis de acesso à consciência necessários à descoberta, pela criança, da existência e da forma como se constituem internamente as estruturas silábicas (KARMILOFF-SMITH, 1994; SEYMOUR, 1997). Na medida em que conhecimentos implícitos precisam ser explicitados, pode-se dizer, também, que há a passagem de uma atividade *epilinguística* a uma atividade *metalinguística* (GOMBERT, 1992).

Após a descrição e a análise dos dados, concluiu-se que os erros ortográficos em sílabas complexas produzidos por crianças em fase de aquisição da escrita podem apresentar três tipos de motivação não mutuamente excludentes, expostas a seguir por ordem de importância e de abrangência: (i) quanto às grafias de estruturas silábicas ramificadas em geral, a explicitação incompleta *da relação* entre os constituintes *onset* complexo e rima ramificada e a camada fonêmica da língua, motivação mais estreitamente associada à hipótese original da pesquisa; (ii) nos casos da grafia de nasal e de lateral em rima ramificada, o conflito entre a fonologia infantil e a fonologia do sistema ortográfico, sendo que esta equivale à fonologia de adultos alfabetizados, enquanto aquela sugere estatutos distintos para as estruturas

com nasal e lateral; (iii) nos casos da grafia de nasal e de fricativa em rima ramificada apenas, os conhecimentos relacionados a regras ortográficas contextuais e arbitrárias. As duas primeiras motivações são de natureza fonológica e são predominantes na primeira metade do processo inicial de escolarização – 1º, 2º e 3º anos –, ainda que se estendam, com frequência menor a cada ano, ao longo de todo o 1º Ciclo do Ensino Fundamental. A terceira motivação é de natureza ortográfica e é observada predominantemente na segunda metade do processo de escolarização, isto é, nos 4º e 5º anos.

Com tais resultados, este estudo, que se utilizou de métodos quantitativos, permite não só que seja corroborada a sua hipótese de partida, mas também contribui com a elucidação de outros tipos de conhecimentos envolvidos na ocorrência dos erros (orto)gráficos em sílabas complexas, não explicitamente previstos na hipótese original e expressos pelos itens (ii) e (iii) do parágrafo anterior.

Considerando especificamente as questões que este trabalho se propôs a responder, as quais se configuraram como desdobramentos da pergunta de pesquisa mais ampla e que permitiram que se chegasse às conclusões expostas anteriormente, destacam-se os resultados que seguem.

Com relação à questão 1 – *Existe diferença entre a frequência de erros e de acertos em sílabas complexas?* – verificou-se que os acertos são significativamente mais frequentes que os erros (orto)gráficos em ambas as estruturas silábicas analisadas – 94,60% contra 5,40% em *onset* complexo e 89,10% contra 10,90% em rima ramificada – corroborando a hipótese de investigação formulada.

No tocante à questão 2 – *Existe diferença entre onset e rima ramificados no que se refere à frequência de erros?* – verificou-se que os erros (orto)gráficos são significativamente mais frequentes em rima ramificada do que em *onset* complexo – 10,90% contra 5,40% – refutando a hipótese de investigação proposta, que previa não haver diferenças significativas entre as duas estruturas no que se refere à frequência de erros (orto)gráficos.

Em relação à questão 3 – *O tipo de segmento influencia na maior ou menor incidência de erros em sílabas complexas?* – verificou-se que os erros (orto)gráficos em rima ramificada são influenciados pelo tipo de segmento envolvido: a grafia das consoantes lateral e nasal concentra significativamente mais erros do que a grafia das consoantes fricativa e rótica. No *onset* complexo, não se verificou influência do tipo de segmento sobre a ocorrência de erros (orto)gráficos, ao serem comparados os dois

grupos consonantais possíveis, formados por lateral e rótica. Ambos os resultados corroboraram a hipótese de investigação proposta.

No que concerne à questão 4 – *O ano escolar influencia na frequência de ocorrência de erros ortográficos em sílabas complexas?* – verificou-se que a frequência de erros (orto)gráficos em *onset* complexo apresenta diferenças significativas entre o 1º ano e o 3º ano do Ensino Fundamental, evidenciando o papel significativo que tem a escolarização na ocorrência de erros nesta estrutura, previsto pela hipótese de investigação. O mesmo não foi verificado para a rima ramificada, estrutura que não apresentou diferenças significativas entre a média de erros (orto)gráficos em cada ano escolar, fazendo com que parte da hipótese de investigação seja refutada.

Na verdade, constatou-se que, na rima ramificada, o papel da escolarização está associado à natureza do conhecimento mobilizado pelas crianças na grafia dessa estrutura silábica, aspecto explorado a partir da questão 5 – *Existe diferença entre a frequência de ocorrência dos tipos de erro (orto)gráfico na rima ramificada?* – e da questão 6 – *Existe diferença na frequência de erros de natureza fonológica e de natureza ortográfica se forem comparados dois grupos distintos no quesito ano escolar?*. Ao serem comparadas as médias de erros de natureza fonológica e de erros de natureza ortográfica em todo o 1º Ciclo do Ensino Fundamental, não foi atestada diferença significativa, refutando a hipótese de investigação proposta para a questão 5. Contudo, ao serem comparadas as médias de erros de natureza fonológica e de natureza ortográfica considerando separadamente os dois períodos distintos que caracterizam o 1º Ciclo, foram constatadas diferenças significativas, corroborando a hipótese de investigação prevista para a questão 6. No primeiro período do 1º Ciclo do EF (1º a 3º anos), os erros de natureza fonológica são significativamente mais frequentes do que os erros de natureza ortográfica; no segundo período do 1º Ciclo (4º e 5º anos), os erros de natureza ortográfica são significativamente⁸⁵ mais frequentes que os erros de natureza fonológica.

Por fim, em relação à questão 7 – *Existe diferença entre a frequência de ocorrência dos tipos de erro de natureza fonológica em sílabas complexas?* – verificou-se que a omissão é o tipo de erro de natureza fonológica mais frequente em ambas as estruturas silábicas analisadas, corroborando a hipótese de investigação

⁸⁵ Diferença marginalmente significativa, conforme resultado expresso pela Tabela 18.

proposta. Na rima ramificada, a substituição também tem importância significativa, embora secundária em relação à omissão. Além disso, constatou-se que a ocorrência da omissão, na rima ramificada, está mais vinculada à grafia da consoante nasal e, conseqüentemente, ao conflito entre a fonologia infantil e a fonologia do sistema ortográfico sobre a representação da nasalidade medial (vogal nasal vs. vogal + consoante nasal). De forma análoga, a substituição, na rima ramificada, está mais associada à grafia da consoante lateral, também como decorrência do conflito entre a fonologia infantil e a fonologia do sistema ortográfico sobre a representação do segmento pós-vocálico (glide em núcleo ramificado vs. líquida lateral em coda). Essa localização das estratégias conforme o tipo de segmento levou à constatação de uma diferença na qualidade do conflito entre as fonologias infantil e adulta: para a lateral, o conflito corresponde a uma dúvida sobre a *qualidade de segmento* que ocupa a posição pós-vocálica; para a nasal, a uma dúvida sobre *se há algum segmento* em posição pós-vocálica.

A partir da descrição e da análise realizadas, não somente se chegou a uma resposta ao problema principal desta pesquisa – *o que motiva as crianças a omitir, inserir, substituir ou alterar a ordem dos segmentos de estruturas com onset e rima ramificados na escrita, se elas já adquiriram fonologicamente tais estruturas?* – mas também puderam ser discutidas outras questões relevantes, associadas ao objetivo central deste trabalho.

A primeira delas diz respeito às assimetrias existentes entre a fonologia infantil e a fonologia adulta alfabetizada, esta última equivalente à fonologia do sistema ortográfico. Conforme recorrentemente assinalado ao longo de todo o trabalho, e mesmo no presente capítulo, os resultados obtidos permitem considerar que as crianças interpretam a composição da rima ramificada de forma distinta de como interpretam os adultos alfabetizados, ou de como interpreta o sistema ortográfico – hipótese levantada originalmente por Miranda (2009a, 2009b, 2012, 2018, 2019) e que leva em conta a precocidade de aquisição do que seriam as codas nasal e lateral ao se comparar com as codas rótica e fricativa, cuja aquisição é mais tardia (MEZZOMO, 2004). Assim, abre-se a possibilidade para pensar que a fonologia infantil entende como segmentos licenciados para a coda apenas /r/ e /S/, diferente do que presume o estudo de Mezzomo (2004); em relação àquilo que para os adultos são um segmento lateral e um segmento nasal em coda (BISOL, 1999), para as

crianças são, respectivamente, um glide em núcleo ramificado (derivado de vogal alta), formando um ditongo com a vogal antecedente, e uma vogal nasal em núcleo simples. O processo de aquisição da escrita alfabética é o elemento que provoca o conflito entre as duas fonologias, bem como a consequente reestruturação da fonologia infantil em conformidade com a fonologia do sistema ortográfico, assim como defende Miranda (2009a, 2009b, 2012, 2017, 2018, 2019).

Outra questão discutida foi concernente às simetrias e às assimetrias entre a aquisição fonológica e a aquisição da escrita alfabética. De modo geral, todos os resultados obtidos, apontam, em alguma medida, para a existência de mais assimetrias do que simetrias entre os dois processos de aquisição, indo ao encontro daquilo que assinalam estudos como os de Alves (2012): na medida em que cada processo demanda um diferente tipo de tarefa cognitiva, as crianças exploram o mesmo conhecimento (no caso, o fonológico) de forma distinta. Na aquisição da fala, são construídas representações, armazenadas em formato implícito; na aquisição da escrita, essas representações já construídas são retomadas e explicitadas ao sistema cognitivo, de modo que podem vir a ser também manipuladas e reestruturadas (MIRANDA, 2017; KARMILOFF-SMITH, 1994).

Com isso em vista, não é recomendado esperar, na aquisição da escrita, pela replicação de processos fonológicos ou de padrões desenvolvimentais idênticos àqueles observados na aquisição da fala. Ainda assim, o cotejamento de resultados obtidos entre cada um dos processos de aquisição é interessante de ser feito na medida em que pode trazer à tona discussões como a levantada por Miranda (2009a, 2009b, 2012, 2018, 2019), a respeito das assimetrias entre a fonologia infantil e a fonologia adulta.

Foi discutida, também, a questão acerca da natureza dos erros (orto)gráficos. Nesse sentido, destaca-se a pertinência da proposta de categorização do GEALE (MIRANDA, 2017, 2020), que demonstrou não somente poder descritivo mas também explicativo. A delimitação dos erros com base nas categorias *Fonológico* e *Ortográfico*, considerando seus respectivos desdobramentos, orientou o processo de análise dos dados, ofereceu uma caracterização abrangente dos erros e contemplou as três diferentes motivações encontradas neste estudo para a ocorrência dos erros na grafia de sílabas complexas. A discussão sobre a natureza dos erros com base na categorização do GEALE também se integrou de maneira importante à discussão sobre as assimetrias entre a aquisição da fala e da escrita, pois aponta para a

mobilização do conhecimento ortográfico, exclusivo do processo de aquisição da escrita.

Finalmente, é digno de nota que, além das contribuições, lacunas são também deixadas por este estudo, o que abre espaço e possibilidade para a realização de novas pesquisas em torno do tema da grafia das sílabas complexas.

Nesse sentido, destaca-se o desenho metodológico desta pesquisa, que, por trabalhar com dados naturalísticos, perde em termos da robustez estatística, apesar de ganhar em aspectos como a discussão qualitativa do dado linguístico. Seria interessante, assim, desenvolver uma pesquisa experimental, incluindo ditados de palavras e de logatomas e oficinas de textos que estimulem a produção das estruturas-alvo da pesquisa. A obtenção de dados de percepção e de produção da fala também seria relevante para o cotejamento com os dados de escrita. De forma complementar ao desenho experimental, poderia se investir na realização de entrevistas clínicas, a fim de melhor explorar os dados produzidos pelos sujeitos nos testes aplicados, na perspectiva bem-sucedida do que já fizeram estudos desenvolvidos no GEALE (cf. FERREIRA, 2016; NEY, 2018).

Além disso, outras variáveis não controladas neste estudo podem ser relevantes para a descrição e a explanação do fenômeno em foco, e podem, assim, ser previstas na elaboração dos testes sugeridos anteriormente. O papel da frequência lexical e da frequência das estruturas silábicas no léxico, a tonicidade, a complexidade da palavra (relativa ao número de sílabas complexas que a palavra possui), a posição do constituinte silábico na palavra (medial e final), o método de alfabetização predominante a que as crianças foram submetidas no processo de aquisição da escrita, são algumas das variáveis cujo efeito ainda precisa ser avaliado e que podem suscitar novos estudos e discussões acerca da grafia das sílabas complexas na aquisição da escrita.

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]. Disponível em: <<http://letras.ufpel.edu.br/verbavolant/segundo/abaurre2.pdf>>.
- 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. Disponível em: <http://www.leffa.pro.br/tela4/Textos/Textos/Livros/Aquisicao_LM_e_LE.pdf>.
- ABAURRE, M. B. M. Horizontes e limites de um programa de investigação em aquisição da escrita. In: LAMPRECHT, R. R. (org.). **Aquisição da Linguagem: questões e análises**. Porto Alegre: EDIPUCRS, 1999. p.167-186.
- ABAURRE, M. B. M. Os estudos linguísticos e a aquisição da escrita. In: II Encontro Nacional sobre Aquisição da Linguagem, 1991a, Porto Alegre. **Anais [...]**. Porto Alegre: PUCRS, 1991a, p. 5-41.
- ADAMOLI, M. A. **Aquisição dos ditongos orais mediais na escrita infantil: uma discussão entre ortografia e fonologia**. 2006. Dissertação (Mestrado em Educação). Faculdade de Educação, Universidade Federal de Pelotas, Pelotas, 2006. Disponível em: <<https://wp.ufpel.edu.br/geale/files/2010/11/ADAMOLI-2006.pdf>>.
- ADAMOLI, M. A. **Um estudo sobre o estatuto fonológico dos ditongos variáveis [aj] e [ej] do PB a partir de dados orais e ortográficos produzidos por crianças de séries iniciais**. 2012. Tese (Doutorado em Educação). Faculdade de Educação, Universidade Federal de Pelotas, Pelotas, 2012. Disponível em: <<https://wp.ufpel.edu.br/geale/files/2010/11/TESE-MARCO-FINAL.pdf>>.
- ALBANO, E. C. **Os gestos e suas bordas: esboço de fonologia acústico-articulatória do português brasileiro**. Campinas: Mercado de Letras/ALB, 2001.
- ALEGRIA, J.; PIGNOT, E.; MORAIS, J. Phonetic analysis of speech and memory codes in beginning readers. **Memory and Cognition**, v. 10, p. 451-456, 1982. Disponível em: <https://www.researchgate.net/publication/16021998_Phonetic_analysis_of_speech_and_memory_codes_in_beginning_readers>.
- ÁVILA, M. M. **A escrita inicial de crianças brasileiras, moçambicanas e portuguesas: um estudo sobre a representação da nasalidade fonológica**. 2019. Dissertação (Mestrado em Educação). Faculdade de Educação, Universidade Federal de Pelotas, Pelotas, 2019. Disponível em: <<https://wp.ufpel.edu.br/geale/files/2019/09/Dissertac%CC%A7a%CC%83o-Mariana.pdf>>.

AZAMBUJA, E. J. **A aquisição das líquidas laterais do português**. 1998. Dissertação (Mestrado em Letras). Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1998.

BARBIERI, T. T. **Aquisição de encontros consonantais com *tap* no português brasileiro**: análises acústica e articulatória. 2019. Dissertação (Mestrado em Letras). Centro de Letras e Comunicação, Universidade Federal de Pelotas, Pelotas, 2019. Disponível em: <<http://guaiaa.ufpel.edu.br/handle/prefix/4485>>.

BERTI, L. C.; CHACON, L.; SANTOS, A. P. A escrita de /aN/ por pré-escolares: pistas acústico auditivas. **Cadernos de Educação**, Pelotas, v. 35, p. 195-219, 2010. Disponível em: <<https://periodicos.ufpel.edu.br/ojs2/index.php/caduc/article/view/1620>>.

BERWICK, R.; CHOMSKY, N. **Why only us**: Language and Evolution. Cambridge: MIT Press, 2016.

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.

BISSEX, G. L. **Gnys at Wrk**: a child learns to write and read. Cambridge: Harvard University Press, 1980.

BLEVINS, J. The syllable in Phonological Theory. In: GOLDSMITH, J. A. (ed.). **The Handbook of Phonological Theory**. Oxford: Blackwell Publishers, 1995.

BOERSMA, P. Cue constraints and their interactions in phonological perception and production. **Rutgers Optimality Archive**, n. 944, 2007. Disponível em: <https://www.researchgate.net/publication/254895984_Cue_constraints_and_their_interactions_in_phonological_perception_and_production>.

BOERSMA, P.; HAMANN, S. Loanword adaptation as first-language phonological Perception. In: CALABRESE, A.; WETZELS, L. W. (eds.). **Loanword phonology**. Amsterdam: John Benjamins, 2009. Disponível em: <https://www.researchgate.net/publication/241889148_Loanword_adaptation_as_first-language_phonological_perception>.

BONET, E.; MASCARÓ, J. On the representation of contrasting rhotics. In: MARTÍNEZ-GIL, F.; MORALES-FRONT, A. (eds.). **Issues in the phonology and morphology of the major Iberian languages**. Georgetown University Press: Washington D.C., 1997. p. 103-126.

BONILHA, G. F. G. **Aquisição dos ditongos orais decrescentes**: uma análise à luz da teoria da otimidade. 2000. Dissertação (Mestrado em Letras). Universidade Católica de Pelotas, Pelotas, 2000. Disponível em: <https://wp.ufpel.edu.br/ppgl/files/2018/11/Aquisicao_dos_ditongos-Giovana_Bonilha.pdf>.

BONILHA, G. F. G. **Aquisição fonológica do português brasileiro**: uma abordagem conexional 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>.

BRASIL. Lei nº. 9.394, de 20 de dezembro de 1996. Dispõe sobre as Diretrizes e Bases da Educação Nacional. **Diário Oficial da União**, Brasília, DF, 23 dez. 1996. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l9394.htm>.

BRASIL. Lei nº. 13.005, de 25 de junho de 2014. Aprova o Plano Nacional de Educação - PNE e dá outras providências. **Diário Oficial da União**, Brasília, DF, 2014. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l13005.htm>.

BRASIL. Ministério da Educação. Secretaria de Educação Básica. Diretoria de Currículos e Educação Integral. Coordenação Geral do Ensino Fundamental. **Elementos conceituais e metodológicos para definição dos direitos de aprendizagem e desenvolvimento do ciclo de alfabetização (1º, 2º e 3º anos) do Ensino Fundamental**. Brasília: MEC, SEB, 2012. Disponível em:
<http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=12827-texto-referencia-consulta-publica-2013-cne-pdf&category_slug=marco-2013-pdf&Itemid=30192>

BRYANT, P.; DEACON, H.; NUNES, T. Morphology and Spelling: what have morphemes to do with spelling?. In: JOSHI, M. R.; ARON, P. G (eds). **The Handbook of Orthography and Literacy**. New Jersey: Lawrence Erlbaum Associates, 2006. p. 601-616.

CAGLIARI, L. C. **Alfabetização e Linguística**. São Paulo: Scipione, 1987.

CALLOU, D.; MORAES, J. A.; LEITE, Y. Consoantes em coda silábica: /s, r, l/. In: CASTILHO, A. T. (coord.); ABAURRE, M. B. M. (org.). **A construção fonológica da palavra**: gramática do português culto falado no Brasil. Vol. 7. São Paulo: Contexto, 2013. p. 167-194.

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

CAPRISTANO, C. C. A propósito da escrita infantil: uma reflexão sobre as segmentações não-convencionais. **Letras de Hoje**, Porto Alegre, v. 39, n. 3, p. 245-260, 2004. Disponível em:
<<http://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/13918>>.

CARRAHER, T. N. Leitura e Escrita: Processos e Desenvolvimento. In: ALENCAR, E. S. **Novas Contribuições da Psicologia aos Processos de Ensino e Aprendizagem**. 4ª ed. São Paulo: Cortez. 2001. p. 15-48.

CARRAHER, T. N. **O método clínico**: usando os exames de Piaget. São Paulo: Cortez, 1989.

CHACON, L. Constituintes prosódicos e letramento em segmentações não-convencionais. **Letras de Hoje**, Porto Alegre, v. 39, n.3, p. 223-232, 2004. Disponível em:
<<http://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/13916>>.

CHOMSKY, C. Reading, writing and phonology. **Harvard Educational Review**, v. 40, p. 287-309, 1970.

CHOMSKY, C. Write first, read later. **Childhood Education**, v. 47, p. 296-229, 1971.

CHOMSKY, N. A review of B. F. Skinner's *Verbal Behavior*. In: LUST, B.; FOLEY, C. (eds.). **First language acquisition: the essential readings**. 1ª ed., vol. 4. Malden: Blackwell Publishing, 2004 [1959]. p. 25-55.

CHOMSKY, N. **Aspectos da teoria da sintaxe**. 2ª ed. Coimbra: Editor-Sucessor, 1978 [1965].

CHOMSKY, N. **Language and problems of knowledge: the Managua Lectures**. Cambridge: MIT Press, 1988.

CHOMSKY, N.; HALLE, M. **The sound pattern of English**. New York: Harper & Row, 1968.

CLEMENTS, G. N.; HUME, E. V. The internal organization of speech sounds. In: GOLDSMITH, J. (org.). **The handbook of phonological theory**. Oxford: Blackwell, 1995.

CLEMENTS, G. N.; KEYSER, S. J. CV Phonology: a generative theory of the syllable. In: KEYSER, S. J. (ed.). **Linguistic inquiry monograph**, n. 9. Cambridge: MIT Press, 1983.

COELHO, B. C. **Transposições ortográficas e estrutura da sílaba na escrita infantil**. 2016. Dissertação (Mestrado em Letras). Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista Júlio de Mesquita Filho, São José do Rio Preto, 2016. Disponível em:
<https://repositorio.unesp.br/bitstream/handle/11449/138142/coelho_bc_me_sjrp.pdf?sequence=3&isAllowed=y>.

COLLISCHONN, G. A sílaba em português. In: BISOL, L. (org.). **Introdução a estudos de fonologia do português brasileiro**. 5ª ed. Porto Alegre: EDIPUCRS, 2014 [1996]. p. 99-131.

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

COPLESTON, F. C. **A history of philosophy** – modern philosophy: the british philosophers from Hobbes to Hume. 1ª ed., vol. 5. New York: Doubleday, 1994.

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. Disponível em:
<http://www.leffa.pro.br/tela4/Textos/Textos/Livros/Aquisicao_LM_e_LE.pdf>.

CHRISTOPHE, A.; GUAISTI, T.; NESPOR, M. Reflections on phonological bootstrapping: its role for lexical and syntactic acquisition. **Language and Cognitive Processes**, v. 12, n. 5, p.585-612, 1997. Disponível em:

<https://www.researchgate.net/publication/239653441_Reflections_on_Phonological_Bootstrapping_Its_Role_for_Lexical_and_Syntactic_Acquisition>.

CRYSTAL, D. **Dicionário de linguística e fonética**. Rio de Janeiro: Jorge Zahar Editora, 1988.

CUNHA, A. P. N. **A hipo e a hipersegmentação nos dados de aquisição da escrita**: um estudo sobre a influência da prosódia. 2004. Dissertação (Mestrado em Educação). Faculdade de Educação, Universidade Federal de Pelotas, Pelotas, 2004. Disponível em: <<https://wp.ufpel.edu.br/geale/files/2010/11/anapaula.-disserta1.pdf>>.

CUNHA, A. P. N. **As segmentações não-convencionais da escrita inicial**: uma discussão sobre o ritmo linguístico do português brasileiro e europeu. 2010. Tese (Doutorado em Educação). Faculdade de Educação, Universidade Federal de Pelotas, Pelotas, 2010. Disponível em: <https://wp.ufpel.edu.br/geale/files/2010/11/TESE_CUNHA_20101.pdf>.

EHRI, L. C. Learning to read and learning to spell are one and the same, almost. In: PERFETTI, C. A.; RIEBEN, L.; FAYOL, M. (eds.). **Learning to spell**: research, theory and practice across languages. Mahwah: Lawrence Erlbaum, 1997. p. 237-269.

FARACO, C. A. **Linguagem escrita e alfabetização**. 1ª ed., 1ª reimp. São Paulo: Contexto, 2016 [2012].

FERREIRA, C. R. G. **Uma palavra é o nome de cada coisa**: um estudo sobre as percepções de crianças do ciclo de alfabetização acerca da palavra oral e gráfica. 2016. Tese (Doutorado em Educação). Faculdade de Educação, Universidade Federal de Pelotas, Pelotas, 2016. Disponível em: <<https://wp.ufpel.edu.br/geale/files/2016/07/Carmen-Ferreira.pdf>>.

FERREIRO, E. A desestabilização das escritas silábicas: alternâncias e desordem com pertinência. In: FERREIRO, E. **O ingresso na escrita e nas culturas do escrito**: seleção de textos de pesquisa. Trad. Rosana Malerba. São Paulo: Cortez, 2013 [2009]. p. 63-76.

FERREIRO, E. A representação da linguagem e o processo de alfabetização. **Cad. Pesq.**, São Paulo, n. 52, p. 7-17, 1985. Disponível em: <<http://publicacoes.fcc.org.br/ojs/index.php/cp/article/view/1357>>.

FERREIRO, E. Luria e o desenvolvimento da escrita na criança. **Cad. Pesq.**, São Paulo, n. 88, p.72-77, 1994. Disponível em: <<http://www.fcc.org.br/pesquisa/publicacoes/cp/arquivos/699.pdf>>.

FERREIRO, E. Nem tudo é ortográfico na aquisição da ortografia. In: FERREIRO, E. **O ingresso na escrita e nas culturas do escrito**: seleção de textos de pesquisa. Trad. Rosana Malerba. São Paulo: Cortez, [2001] 2013. p. 248-272.

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

FERREIRO, E; ZAMUDIO, C. A escrita das sílabas CVC e CCV no início da alfabetização escolar: a omissão de consoantes é uma prova da incapacidade para

analisar a sequência fônica?. In: FERREIRO, E. **O ingresso na escrita e nas culturas do escrito**: seleção de textos de pesquisa. Trad. Rosana Malerba. São Paulo: Cortez, 2013 [2008]. p. 219-243.

FRANÇA, A. I.; FERRARI, L.; MAIA, M. **A Linguística no século XXI**: convergências e divergências no estudo da linguagem. São Paulo: Contexto, 2016.

FREITAS, M. J. **A aquisição da estrutura silábica do português europeu**. 1997. Tese (Doutorado em Linguística). Universidade de Lisboa, 1997.

FREITAS, M. J. Aquisição da fonologia em língua materna: a sílaba. In: FREITAS, M. J.; SANTOS, A. L. (orgs.). **Aquisição de língua materna e não materna**: questões gerais e dados do português. Berlin: Language Science Press. 2017. Disponível em: <<https://langsci-press.org/catalog/book/160>>.

FRITH, U. Beneath the Surface of Development Dyslexia. In: PATTERSON, K.E.; MARSHALL, J.C.; COLTHEART, M. (orgs.) **Surface Dyslexia**: Neuropsychological and Cognitive Studies of Phonological Reading. London; Hillsdale: Lawrence Erlbaum, 1985. p. 301-330. Disponível em: <https://www.researchgate.net/publication/245583604_Beneath_the_surface_of_developmental_dyslexia>.

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>.

FROTA, S.; VIGÁRIO, M. Early intonation in european portuguese. In: Third conference on tone and intonation (tie 3). 2008, Lisboa. **Anais [...]**. Lisboa: Universidade de Lisboa, 2008. Disponível em: <https://www.researchgate.net/publication/266588834_Early_Intonation_in_European_Portuguese>.

GELB, I. J. **A study of writing**. Chicago: University of Chicago Press, 1952.

GENTRY, R. An analysis of developmental spelling in "GNYS AT WRK". **The Reading Teacher**, v. 36, p. 192-200, 1982.

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

GOUT, A.; CHRISTOPHE, A. O papel do bootstrapping prosódico na aquisição da sintaxe. In: CORREA, L. M. S. (org.). **Aquisição da linguagem e problemas do desenvolvimento linguístico**. Rio de Janeiro: EDIPUCRJ; São Paulo: Loyola, 2006. p. 103-128.

HAYES, B. **Metrical stress theory**: principles and case studies. Chicago: University Of Chicago Press, 1995.

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.

HERNANDORENA, C. L. B. M. **Uma proposta de análise de desvios fonológicos através de traços distintivos**. 1988. Dissertação (Mestrado em Letras). Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1988.

HOUAISS, A.; VILLAR, M. S. **Minidicionário Houaiss da língua portuguesa**. 2ª ed. Rio de Janeiro: Objetiva, 2004.

HYMAN, L. M. **A theory of phonological weight**. Dordrecht: Foris, 1985.

ILARI, R. O estruturalismo linguístico: alguns caminhos. In: MUSSALIM, F.; BENTES, A. C. (orgs.). **Introdução à Linguística: fundamentos epistemológicos**. Vol. 3. 5ª ed. São Paulo: Cortez, 2011 [2004], p. 53-92.

JUSCZYK, P. W; CUTLER, A; REDANZ, N. J. Infants' preference for the predominant stress patterns of english words. **Child development**, v. 64, n. 3, p. 675-687, 1993. Disponível em: <<https://www.jstor.org/stable/1131210?seq=1>>.

KAHN, D. **Syllable-based generalizations in English Phonology**. Tese (Doutorado, PhD). Cambridge, Mass: MIT, 1976.

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.

KOCH, I. V.; CUNHA-LIMA, M. L. Do cognitivismo ao sociocognitivismo. In: MUSSALIM, F.; BENTES, A. C. (orgs.). **Introdução à Linguística: fundamentos epistemológicos**. Vol. 3. 5ª ed. São Paulo: Cortez, 2011 [2004]. p. 251-300.

LAMPRECHT, R. R. **Perfil da aquisição normal da fonologia do português: descrição longitudinal de 12 crianças de 2:9 a 5:5**. 1990. Tese (Doutorado em Letras). Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1990.

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, A. M.; COOPER, F. S.; SHANKWEILER, D. P.; STUDDERT-KENNEDY, M. Perception of the speech code. **Psychol. Rev.**, 74, p. 431-461, 1967. Disponível

em:

<https://www.researchgate.net/publication/18994638_Perception_of_the_Speech_Code>.

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**, 18, p. 201- 212, 1974. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.602.5825&rep=rep1&type=pdf>>.

LURIA, A. R. The development of writing in the child. In: COLE, M. (ed.). **The selected writings of A. R. Luria**. New York: ME, Sharpe Inc., 1978 [1929]. p. 145-194.

LYONS, J. **As ideias de Chomsky**. São Paulo: Cultrix, 1970.

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

LYONS, J. **Lingua(gem) e Linguística**. São Paulo: LTC, 1987.

MACKEN, M. A. Developmental reorganization of phonology: a hierarchy of basic units of acquisition. **Lingua**, v. 49, p. 11-49, 1979.

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 Graphics, 1992. p. 249-269.

MALUF, M. R.; GOMBERT, J. E. Habilidades implícitas e controle cognitivo na aprendizagem da linguagem escrita. In: MALUF, M. R.; GUIMARÃES, S. R. K. (orgs.). **Desenvolvimento da linguagem oral e escrita**. Curitiba: Editora UFPR, 2008. p. 123-135.

MARCONI, M. A.; LAKATOS, E. M. **Fundamentos da metodologia científica**. 8ª ed. São Paulo: Atlas, 2017.

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

MARTINS, C. **Manual de análise de dados quantitativos com recurso ao IBM SPSS: saber decidir, fazer, interpretar e redigir**. 1ª ed. Braga: Psiquilíbrios Edições, 2011.

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

MATZENAUER, C. L. B. O modelo BiPhon e a fonologização de traços e segmentos na aquisição da linguagem. **Revista de Estudos Linguísticos da Universidade do Porto**, v. 10, p. 59-84, 2015. Disponível em: <<http://ojs.letras.up.pt/ojs/index.php/EL/article/viewFile/2670/2458>>.

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. Disponível em: <<http://www.periodicos.letras.ufmg.br/index.php/relin/article/view/2745>>.

MAWHIN, J. Henri Poincaré: a life in the service of science. **Notices of the AMS**, v. 52, n. 9, p. 1036-1044, 2005. Disponível em: <https://www.researchgate.net/publication/242022380_Henri_Poincare_A_Life_in_the_Service_of_Science>.

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.

MEHLER, J. *et al.* Infant recognition of mother's voice. **Perception**, v. 7, n. 5, p. 491-497, 1978. Disponível em: <<https://pdfs.semanticscholar.org/42ae/5ac303502ac3514036cad2e52740af15a662.pdf>>.

MIRANDA, A. R. M. **A aquisição do 'r'**: uma contribuição à discussão sobre seu status fonológico. 1996. Dissertação (Mestrado em Letras). Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1996. Disponível em: <<https://wp.ufpel.edu.br/geale/files/2019/02/MIRANDA-1996.pdf>>.

MIRANDA, A. R. M. A fonologia em dados de escrita inicial de crianças brasileiras. **Linguística (Madrid)**, v. 30, p. 45-80, 2014. Disponível em: <https://wp.ufpel.edu.br/geale/?page_id=1097>.

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. Disponível em: <https://wp.ufpel.edu.br/geale/?page_id=1093>.

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>>.

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. Disponível em: <https://wp.ufpel.edu.br/geale/?page_id=1428>.

MIRANDA, A. R. M. Aquisição da linguagem: escrita e fonologia. In: LAZZAROTTO-VOLCÃO, C.; FREITAS, M. J. (orgs.). **Estudos em Fonética e Fonologia**: coletânea em homenagem a Carmen Matzenauer. Curitiba: CRV, 2018. p. 335-364.

MIRANDA, A. R. M. **BATALE**: Banco de Textos de Aquisição da Linguagem Escrita. Pelotas: Faculdade de Educação, Universidade Federal de Pelotas, 2001. Disponível em: <<https://sistemavestigios.org>>.

MIRANDA, A. R. M. Informação fonológica na aquisição da escrita. In: RÉ, A.; KOMESU, F.; TENANI, L.; VIEIRA, A. J. (orgs.). **Estudos linguísticos**

contemporâneos: diferentes olhares. São Paulo: Cultura Acadêmica, 2013. p. 11-35. Disponível em: <https://wp.ufpel.edu.br/geale/?page_id=1093>.

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**. 1ª 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. Disponível em: <https://wp.ufpel.edu.br/geale/?page_id=1097>.

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

MIRANDA, A. R. M. Um estudo sobre o erro ortográfico. In: HEINING, O. L.; FRONZA, C. A. (orgs.). **Diálogos entre linguística e educação**. 1ª ed, v. 1. Blumenau: EDIFURB, 2010. p. 141-162. Disponível em: <https://wp.ufpel.edu.br/geale/?page_id=1093>.

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**. v. 48, n 3, p. 343-354, 2013. Disponível em: <https://wp.ufpel.edu.br/geale/?page_id=1097>.

MIRANDA, A. R. M.; MATZENAUER, C. L. B. Aquisição da fala e da escrita: relações com a fonologia. **Cadernos de Educação**, n. 35, p. 359-406, 2010. Disponível em: <https://wp.ufpel.edu.br/geale/?page_id=1176>.

MIRANDA, A. R. M.; SILVA, M. R.; MEDINA, S. Z. O sistema ortográfico do português brasileiro e sua aquisição. **Revista Linguagens e Cidadania**. v. 14, n. 1, 1-15, 2005. Disponível em: <https://wp.ufpel.edu.br/geale/?page_id=1097>.

MIRANDA, I. C. C. **Aquisição e variação estruturada de encontros consonantais tautossilábicos**. 2007. Tese (Doutorado em Linguística). Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2007.

MORAIS, A. G. **Ortografia:** ensinar e aprender. 4ª ed. São Paulo: Ática, 2002.

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.

MORGAN, J. L.; DEMUTH, K. (orgs.). **Signal to syntax:** bootstrapping from speech to grammar in early acquisition. 1ª ed. Mahwah: Lawrence Erlbaum Associates, 1996.

NAZZI, T.; BERTONCINI, J.; MEHLER, J. Language discrimination by newborns: towards an understanding of the role of rythm. **Journal of Experimental Psychology:** Human Perception and Performance, v. 24, n. 3, p. 756-766, 1998. Disponível em: <https://www.researchgate.net/publication/13657945_Language_Discrimination_by_Newborns_Toward_an_Understanding_of_the_Role_of_Rhythm>.

NESPOR, M.; VOGEL, I. **La prosodia**. Madrid: Visor Distribuciones, S.A., 1994 [1986].

NEY, L. A. G. **Acento gráfico e prosódico**: um estudo sobre as hipóteses de crianças dos anos iniciais. 2018. Tese (Doutorado em Educação). Faculdade de Educação, Universidade Federal de Pelotas, Pelotas, 2018. Disponível em: <https://wp.ufpel.edu.br/geale/files/2019/02/2018_Tese_Luanda_Alvariza_Gomes_Ney.pdf>.

NEY, L. A. G. **Acentuação gráfica na escrita de crianças de séries iniciais**. 2012. Dissertação (Mestrado em Educação). Faculdade de Educação, Universidade Federal de Pelotas, Pelotas, 2012. Disponível em: <https://wp.ufpel.edu.br/geale/files/2010/11/Disserta%C3%A7%C3%A3o_Luanda_Alvariza_Gomes_Ney.pdf>.

OLIVEIRA, C. C. **Aquisição das fricativas /f/, /v/, /š/ e /z/ do português brasileiro**. 2002. Dissertação (Mestrado em Letras). Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2002.

OTHERO, G. A.; KENEDY, E. (orgs.). **Chomsky**: a reinvenção da linguística. São Paulo: Contexto, 2019.

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. A metátese na aquisição da escrita: regularidades e possíveis motivações. In: XXV Congresso de Iniciação Científica da UFPel, 2016, Pelotas, **Anais eletrônicos** [...]. Pelotas: UFPel, 2016. Disponível em: <http://cti.ufpel.edu.br/siepe/arquivos/2016/LA_04230.pdf>.

PACHALSKI, L; MIRANDA, A. R. M. A metátese na aquisição da escrita: simetrias e assimetrias entre fonologia e ortografia. **Filologia e Linguística Portuguesa (Online)**, v. 20, n. 2, p. 233-256, 2018b. Disponível em: <<https://www.revistas.usp.br/flp/article/view/151909>>.

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>.

PACHALSKI, L; MIRANDA, A. R. M. Estruturas silábicas ramificadas na escrita de crianças do Ciclo de Alfabetização. In: V Jornada de Estudos da Linguagem, 2019, Pelotas, **Anais eletrônicos** [...]. Pelotas: UFPel, 2019. Disponível em: <<https://wp.ufpel.edu.br/jel/>>.

PACHALSKI, L.; RODRIGUES, J. C; VIEIRA, I. F.; MIRANDA, A. R. M. A produção de encontros consonantais tautossilábicos em dados de escrita inicial. In: XXIII Congresso de Iniciação Científica da UFPel, 2014, Pelotas, **Anais eletrônicos** [...].

Pelotas: UFPel, 2014. Disponível em:

<http://cti.ufpel.edu.br/siepe/arquivos/2014/LA_03222.pdf>.

PACHALSKI, L.; RODRIGUES, J. C.; VIEIRA, I. F.; MIRANDA, A. R. M. Simetrias e assimetrias entre a aquisição da fala e da escrita: o uso da metátese como estratégia para a grafia do onset complexo. In: XXIV Congresso de Iniciação Científica da UFPel, 2015, Pelotas, **Anais eletrônicos** [...]. Pelotas: UFPel, 2015. Disponível em: <http://cti.ufpel.edu.br/siepe/arquivos/2015/LA_04474.pdf>.

PIAGET, J. Psicogênese dos conhecimentos e seu significado epistemológico. PIATELLI-PALMARINI, M. (org.). **Teorias da Linguagem, Teorias da Aprendizagem**: o debate entre Jean Piaget e Noam Chomsky. São Paulo: Cultrix, EdUSP, 1983 [1979]. p. 39-49.

PIATELLI-PALMARINI, M. (org.). **Teorias da Linguagem, Teorias da Aprendizagem**: o debate entre Jean Piaget e Noam Chomsky. São Paulo: Cultrix, EdUSP, 1983 [1979].

PREMACK, D. Pedagogy and aesthetics as sources of culture. In: GAZZANIGA, M. (ed.). **Handbook of Cognitive Neuroscience**. New York: Plenum, 1984.

RANGEL, G. A. **Aquisição do sistema vocálico do português brasileiro**. 2002. Tese (Doutorado em Letras). Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2002.

READ, C. Lessons to be learned from the preschool orthographer. In: LENNEBERG, E. H.; LENNEBERG, E. (orgs.) **Foundations of Language Development: a Multidisciplinary Approach**. 1ª ed., vol. 2. New York: Academic Press, 1975.

READ, C. Preschool children's knowledge of english phonology. **Harvard Educational Review**, v. 41, n. 1, 1971. Disponível em: <<https://hepgjournals.org/doi/abs/10.17763/haer.41.1.91367v0h80051573?journalCode=haer>>.

READ, C.; ZHANG, Y.; NIE, H.; DING, B. The ability to manipulate speech sounds depends on knowing alphabetic writing. **Cognition**, v. 24, p. 31-44, 1986. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/001002778690003X>>.

REGO, L. L. B. O desenvolvimento cognitivo e prontidão para a alfabetização. In: CARRAHER, T. N. (org). **Aprender pensando**: contribuições da psicologia cognitiva para a educação. Petrópolis: Vozes, 1986. p. 32-49.

RIBAS, L. P. **Aquisição do onset complexo no Português Brasileiro**. 2002. Dissertação (Mestrado em Letras). Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2002.

SANTOS, R. N. **Aquisição de grupos consonânticos e seu impacto nos desempenhos escritos no 1º Ciclo do Ensino Básico**. 2013. Dissertação (Mestrado em Letras). Faculdade de Letras, Universidade de Lisboa, 2013. Disponível em: <<https://repositorio.ul.pt/handle/10451/12194>>.

SANTOS, R. S. **A aquisição do acento de palavra no português brasileiro**. 2001. Tese (Doutorado em Letras). Universidade Estadual de Campinas, Campinas, 2001.

SANTOS, R. S. Aquisição da fonologia em língua materna: acento e palavra prosódica. In: FREITAS, M. J.; SANTOS, A. L. (orgs.). **Aquisição de língua materna e não materna**: questões gerais e dados do português. 1ª ed. Berlin: Language Science Press, 2017. p. 95- 117. Disponível em: <<https://langsci-press.org/catalog/book/160>>.

SANTOS, R. S. **Uma interface fonologia-sintaxe**: o uso de sons preenchedores. 1995. Dissertação (Mestrado em Letras). Universidade Estadual de Campinas, Campinas, 1995.

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

SAVIO, C. B. **Aquisição das fricativas /s/ e /z/ do português brasileiro**. 2001. Dissertação (Mestrado em Letras). Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2001.

SCARPA, E. M. Interfaces entre componentes e representação na aquisição da prosódia. In: LAMPRECHT, R. R. (org.) **Aquisição da linguagem**: questões e análises. Porto Alegre: EDIPUCRS, 1999.

SCARPA, E. M. Learning external sandhi: evidence for a top-down hypothesis of prosodic acquisition. In: SORACE, A.; HEYCOCK, C.; SHILLCOCK, R. (eds.). **Proceedings of GALA' 1997 conference on language acquisition**: Knowledge representation and processing. Edinburgh: Edinburgh University Press, 1997. p. 272–277.

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. O desenvolvimento inicial da leitura em ortografias europeias. In: SNOWLING, M. J.; HULME, C. (orgs.) **A ciência da leitura**. Porto Alegre: Penso, 2013 [2005]. p. 306-333.

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. Disponível em: <https://www.researchgate.net/publication/10710502_Foundation_literacy_acquisition_in_European_orthographies_Electronic_version>.

SKINNER, B. F. **Verbal behavior**. Acton: Copley Publishing Group, 1957.

SLOBIN, D. I. **Psicolinguística**. São Paulo: EDUSP, 1980.

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

STRAUSS, S.; STAVY, R. U-shaped behavioral growth: implications for theories of development. In: HARTUP, W. W. **Child Development Research**. Chicago; London: The University of Chicago Press, 1982. p. 547-599.

TENANI, L. E. Segmentações não-convencionais e teorias fonológicas. **Letras de Hoje**, Porto Alegre, v. 39, n. 3, p. 233-244, 2004. Disponível em: <<http://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/13917/9231>>.

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>.

VASSOLER, A. M. O. **Coordenação gestual na produção de encontros consonantais em crianças com desenvolvimento típico e atípico**. 2016. Tese (Doutorado em Estudos Linguísticos). Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2016. Disponível em: <<https://repositorio.unesp.br/handle/11449/139429>>.

VELOSO, J. **Da influência do conhecimento ortográfico sobre o conhecimento fonológico**: estudo longitudinal de um grupo de crianças falantes nativas do português europeu. 2003. Tese (Doutorado em Linguística). Faculdade de Letras, Universidade do Porto, Porto, 2003. Disponível em: <<https://repositorio-aberto.up.pt/bitstream/10216/18030/2/4990TD01P000079024.pdf>>.

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. Disponível em: <https://www.researchgate.net/publication/259892028_A_lingua_na_escrita_e_a_escrita_da_lingua_Alguas_consideracoes_gerais_sobre_transparencia_e_opacidade_fonemicas_na_escrita_do_portugues_e_outras_questoes>.

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. Disponível em: <<https://doi.org/10.3989/loquens.2019.059>>.

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

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>>.