

Viviane Schier Martins Iachak

**Configuração prosódica das segmentações não-convencionais de
palavra no Ensino Fundamental I nos contextos público e privado.**

São José do Rio Preto -SP
2022

Viviane Schier Martins Iachak

**Configuração prosódica das segmentações não-convencionais de
palavra no Ensino Fundamental I nos contextos público e privado.**

Tese apresentada como parte dos requisitos
para obtenção do título de Doutora em
Estudos Linguísticos, junto ao Programa de
Pós-Graduação em Estudos Linguísticos, do
Instituto de Biociências, Letras e Ciências
Exatas da Universidade Estadual Paulista
“Júlio de Mesquita Filho” – UNESP –
Campus de São José do Rio Preto (SP).

Financiadora: CAPES

Orientador: Prof. Dr. Lourenço Chacon
Jurado Filho

São José do Rio Preto
2022

I11c

Iachak, Viviane Schier Martins

Configuração prosódica das segmentações não-convencionais de palavra no ensino fundamental I nos contextos público e privado. /

Viviane Schier Martins Iachak. -- São José do Rio Preto, 2022

131 p. : tabs.

Tese (doutorado) - Universidade Estadual Paulista (Unesp),
Instituto de Biociências Letras e Ciências Exatas, São José do Rio
Preto

Orientador: Lourenço Chacon

1. Escrita infantil. 2. Segmentação não convencional de palavras. 3.
Fonologia Prosódica. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Viviane Schier Martins Iachak

**Configuração prosódica das segmentações não-convencionais de
palavra no Ensino Fundamental I nos contextos público e privado.**

Tese apresentada como parte dos requisitos
para obtenção do título de Doutora em
Estudos Linguísticos, junto ao Programa de
Pós-Graduação em Estudos Linguísticos, do
Instituto de Biociências, Letras e Ciências
Exatas da Universidade Estadual Paulista
“Júlio de Mesquita Filho” – UNESP –
Câmpus de São José do Rio Preto (SP).
Financiadora: CAPES (Março/2017 a
Maio/2021)

Comissão Examinadora

Prof. Dr. Lourenço Chacon
UNESP – Marília
Orientador

Prof^a. Dr^a. Fabiana Komesu
UNESP – São José do Rio Preto/SP

Prof^a. Dr^a. Luciani Tenani
UNESP – São José do Rio Preto/SP

Prof^a. Dr^a. Elaine Cristina de Oliveira
UFBA - Universidade Federal da Bahia

Prof^a. Dr^a. Geovana Carina Neris Soncin
UNESP - Marília.

São José do Rio Preto
05 de maio de 2022.

À Lavínia por tudo o que tem me ensinado (e pelo que representa em minha vida)...

dedico este trabalho

AGRADECIMENTOS

Ao meu orientador professor Lourenço Chacon, a quem devo boa parte da minha trajetória na pós-graduação. Agradeço pelos compartilhamentos, não somente sobre a linguagem e sobre ciência, mas sobre a vida. Em especial, pelo apoio e pelo acolhimento que me foram imprescindíveis no período de desenvolvimento desta pesquisa.

À professora Luciani Tenani, por ter realizado a leitura do trabalho. Mais do que isso, agradeço por ter provocado inquietações importantes em mim, ao longo de minha trajetória na Unesp. Meu agradecimento, em especial, pela amizade.

À professora Fabiana Komesu por todos os esforços e orientações no âmbito da Coordenação do Programa. Ainda, agradeço pela leitura do texto e pelas valiosas contribuições no exame de qualificação.

À professora Jáima, pela leitura do trabalho. Sobretudo, pela amizade, pelas acolhidas, pelos compartilhamentos e pelos cafés quando estava em Marília.

À minha irmã Ana Cândida, que acompanhou de perto todo o processo de desenvolvimento da pesquisa e de produção do texto. Agradeço pelos diálogos, pelas risadas, pelos conselhos, pelas orações e por não ter me deixado sozinha em momento algum.

À minha mãe, por toda a ajuda com a Lavínia, apoio sem o qual não teria sido possível a revisão do texto.

À tia Isa, pelo apoio e pela ajuda.

À minha prima Jéssica, por não ter medido esforços para ajudar a mim e à Lavínia.

À minha amiga Cris (Formiga), a quem eu dedico boa parte do que aprendi ao longo desses anos em Rio Preto. Agradeço pelos compartilhamentos relacionados à escrita, à Fonologia Prosódica. Gratidão eterna pela acolhida em sua casa, pela amizade. Em especial pela amizade e por me inspirar.

Às colegas: Gabi, Larissa, Lídia, Lilian, Giovana, Nayra, Suelen, pelos diálogos, pela risada e pela amizade, o que tornou o processo mais leve e prazeroso. À Gabi, em especial, por te estado perto (mesmo longe!) com quem partilhei “os altos e baixos” da pós-graduação.

À Livia, pela amizade, pela escuta.

À Ana Carolina, a irmã “rio-pretense” que o doutorado me trouxe.

À Dona Tereza pelas orações e por todo o apoio.

A todos os familiares que torceram por mim.

Às professoras Elaine e Geovana pela leitura do trabalho e pelas contribuições.

A Deus, sem o qual a vida não faz sentido para mim.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, à qual agradeço.

RESUMO

Tomando como ponto de partida os resultados de pesquisas que apontam a existência de contrastes na escrita de alunos de Ensino Fundamental I de escolas públicas e de escolas privadas, quando comparados os dados de segmentação não-convencional de palavra – Abaurre (1991); Cunha (2004); Torquette (2016)- a presente pesquisa parte da hipótese de que a variabilidade, com relação às grafias dessa natureza, nos dados de escrita de alunos do Ensino Fundamental de contextos de ensino público e de ensino privado, pode se mostrar não somente em termos de quantidade (como apontam os estudos) mas também quanto à sua natureza, quando analisada a configuração prosódica das estruturas. Nesse sentido, a presente investigação tem como objetivo geral investigar a configuração prosódica de ocorrências de segmentação não-convencional de palavra em textos produzidos por crianças do 1º ao 5º ano de ensino fundamental I de escolas públicas e de escolas privadas. Mais especificamente, a partir da análise pretendemos, nesta investigação: a-) analisar prosodicamente as ocorrências de hipossegmentações e de hipersegmentações de palavras no dois contextos (público e privado) à luz da Fonologia Prosódica (Nespor e Vogel, 1986); (b) comparar os dados considerando as variáveis *tipo de escola* e *série*, a fim de verificar a existência de variabilidade na configuração prosódica das estruturas de hipossegmentação e de hipersegmentação. Este trabalho, de cunho qualitativo e de caráter transversal, tomou como *corpus* 423 textos produzidos por alunos de escolas públicas e de escolas privadas da cidade de Marília-SP. No caso das hipossegmentações, a percepção é pontual: os alunos de escola pública apresentam mais dificuldade na segmentação de palavras quando estão sendo mobilizados os domínios de frase fonológica e, sobretudo, o grupo clítico (quando os dados demonstram a existência de variabilidade estatística ainda mais acentuada do que no domínio da frase fonológica). No que diz respeito às hipersegmentações, a variabilidade hipotetizada inicialmente se mostra, também, de forma pontual: quando a estrutura segmentada é formada pela combinação de uma sílaba + pé (levando à formação preferencial de uma estrutura sílaba + pé troqueu). A conclusão mais geral a que chegamos, portanto, compreende ao fato de que, como hipotetizado, há diferenças quando comparados os dados de hipossegmentação e de hipersegmentação. Entretanto essa diferença se dá pontualmente e quando envolve estruturas e constituintes prosódicos específicos.

Palavras-chave: Segmentação de palavras. Prosódia. Ensino público. Ensino privado.

ABSTRACT

Taking as a starting point the results of research that point out the existence of contrasts in the writing of Elementary School I students from public and private schools, when comparing data on non-conventional word segmentation – Abaurre (1991); Cunha (2004); Torquette (2016)- this research starts from the hypothesis that the variability, in relation to spellings of this nature, in the writing data of elementary school students from public and private education contexts, can be shown not only in terms of quantity (as the studies point out) but also in terms of its nature, when analyzing the prosodic configuration of the structures. In this sense, the present investigation has the general objective to investigate the prosodic configuration of occurrences of non-conventional word segmentation in texts produced by children from the 1st to the 5th year of elementary school I in public and private schools. More specifically, based on the analysis, in this investigation we intend to: a-) prosodically analyze the occurrences of hyposegmentations and hypersegmentations of words in both contexts (public and private) in the light of Prosodic Phonology (Nespor and Vogel, 1986); (b) to compare the data considering the variables type of school and grade, in order to verify the existence of variability in the prosodic configuration of hyposegmentation and hypersegmentation structures. This qualitative and transversal work took as corpus 423 texts produced by students from public and private schools in the city of Marília-SP. In the case of hyposegmentations, the perception is punctual: public school students have more difficulty in word segmentation when phonological phrase domains and, above all, the clitic group are being mobilized (when the data demonstrate the existence of even greater statistical variability accented than in the phonological phrase domain). With regard to hypersegmentations, the variability initially hypothesized is also punctual: when the segmented structure is formed by the combination of a syllable + foot (leading to the preferential formation of a syllable + foot trochee structure). The most general conclusion we reached, therefore, comprises the fact that, as hypothesized, there are differences when comparing hyposegmentation and hypersegmentation data. However, this difference occurs occasionally and when it involves specific prosodic structures and constituents.

Keywords: Word segmentation. Prosody. Public education. Private education.

LISTA DE FIGURAS

Figura 1. O contínuo de gêneros textuais, conforme Marcuschi (2001).....	17
Figura 2. Ocorrências de segmentação não convencional na amostra.....	38
Figura 3. Hierarquia de constituintes.....	39
Figura 4. Exemplo de enunciado fonológico.....	40
Figura 5. Exemplo de frase entonacional	41
Figura 6. Exemplo de frase fonológica.....	42
Figura 7. Exemplo de grupo clítico.....	44
Figura 8. Exemplo de grupo clítico.....	44
Figura 9. Exemplo de grupo clítico.....	45
Figura 10. Exemplo de grupo clítico.....	45
Figura 11. Exemplo de palavra fonológica.....	46
Figura 12. Exemplo de pé métrico.....	48
Figura 13. Exemplo de pé métrico iambo.....	49
Figura 14. Exemplo de pé métrico troqueu.....	49
Figura 15. Exemplo de sílaba.....	50
Figura 16. Exemplo de sílaba.....	50
Figura 17. Exemplo de texto para contagem de palavras.....	62
Figura 18. Exemplo de texto não identificado.....	63
Figura 19. Translineação.....	63
Figura 20. Caso de hífen.....	64
Figura 21. Exemplo de mescla.....	64
Figura 22. Exemplo de hipossegmentação com ruptura prosódica.....	65
Figura 23. Exemplo de hipo em U.....	75
Figura 24. Exemplo de hipo em I.....	76
Figura 25. Exemplo de hipo em frase entonacional.....	76
Figura 26. Exemplo de hipo de frase fonológica.....	77
Figura 27. Exemplo de hipo em grupo clítico.....	77
Figura 28. Exemplo de hiper formada por pé+ pé.....	80
Figura 29. Exemplo de hiper formada por sílaba + sílaba	81
Figura 30. Exemplo de hiper formada por pé + sílaba.....	82
Figura 31. Exemplo de hiper formada por sílaba+ pé.....	82
Figura 32. Exemplo de hiper formada por pé + pé + sílaba.....	82
Figura 33. Exemplo de hiper formada por sílaba + sílaba+ sílaba.....	83

LISTA DE TABELAS

Tabela 1. O município de Marília – Dados do Censo 2010.....	53
Tabela 2. Índices do IDEB – Marília 2011-2019.....	57
Tabela 3. Distribuição do Banco de Dados do <i>Escin</i>	57
Tabela 4. Divisão da amostra.....	59
Tabela 5. Codificação dos textos.....	59
Tabela 7. O número de palavras escritas em cada contexto.....	59
Tabela 8. Comparação do número de palavras.....	68
Tabela 9. As segmentações não- convencionais nos contextos.....	71
Tabela 10. Distribuição das ocorrências de segmentações não-convencionais nos contextos, por séries, segundo o tipo.....	74
Tabela 11. Distribuição das Hipossegmentações nos contextos por série em cada domínio prosódico	80
Tabela 12. Distribuição das Hipersegmentações nos contextos por série em cada domínio prosódico.....	95

LISTA DE GRÁFICOS

Gráfico1. Perfil econômico do município de Marília.....	54
Gráfico 2. Escolas do município, por nível.....	55
Gráfico 3. Ideb do Município de Marília.....	57

SUMÁRIO

APRESENTAÇÃO.....	5
SEÇÃO 1: BASES TEÓRICAS.....	10
1.1 A propósito da escrita infantil enquanto objeto de estudo: breve retrospectiva.....	11
1.2 A propósito da relação entre fala e escrita.....	13
1.3 A propósito da noção de palavra no estudo das grafias de segmentação não-convencional: algumas considerações.....	20
1.4 Segmentação não-convencional de palavra: tendências para investigação.....	24
1.4.1 Tendência fonético-semântica.....	26
1.4.2 Tendência fonológico-letrada.....	29
1.4.3 Tendência fonológico-pedagógica.....	33
1.5 Sobre a fonologia prosódica.....	36
SEÇÃO 2: METODOLOGIA.....	52
2.1 Sobre o banco de dados e sua constituição.....	52
2.2 Sobre a amostra.....	58
2.3 Sobre os procedimentos metodológicos.....	60
2.4 Critérios para seleção dos dados.....	61
2.5 Procedimentos para análise estatística.....	65
SEÇÃO 3: DESCRIÇÃO DOS RESULTADOS.....	68
3.1 Sobre o números de palavras	69
3.2 Sobre a ocorrência de dados de segmentação não -convencional.....	71
3.3 Sobre as hipossegmentações e as hipersegmentações nos dois contextos: resultados da análise prosódica.....	73
3.3.1 Sobre as hipossegmentações.....	74
3.3.2 Sobre as hipersegmentações.....	80
SEÇÃO 4: DISCUSSÃO DOS RESULTADOS.....	87
4.1 Sobre a ocorrência de segmentações não-convencionais segundo o tipo.....	87
4.2 Hipossegmentações: configuração prosódica.....	94
4.3 Hipersegmentações: configuração prosódica.....	99
SEÇÃO 5: CONCLUSÕES.....	106
5.2 Respondendo ao objetivo geral.....	108

5.3 Respondendo aos objetivos específicos.....	109
5.4 Ultimas considerações.....	113
REFERÊNCIAS.....	116

APRESENTAÇÃO

Ao tomar o texto infantil como objeto de análise, estudos pioneiros como os de Abaurre et al. (1985) e de Cagliari (1986) mostraram, por meio da observação de registros escritos não-convencionais, que a relação entre fala e escrita é mais complexa do que a crença no fato de que a “criança escreve como fala”. Lançando o olhar para os chamados textos espontâneos, aqueles que resultam de produção não controlada experimentalmente, esses autores observaram que os erros de escrita ¹ não compreendem equívocos oriundos de uma interferência da fala na escrita, pois, na perspectiva desses autores, esses erros são, sobretudo, indícios da reflexão e da construção de hipóteses, por parte do escrevente, sobre a escrita e sobre sua relação com características da fala.

Arelados a essas discussões, pesquisadores brasileiros situados no âmbito dos estudos linguísticos vêm produzindo importantes reflexões sobre a relação entre fala e escrita, sobretudo no que tange às relações entre as características fonético-fonológicas da língua e aspectos variados da aquisição e do desenvolvimento da escrita.

Uma vez que neste trabalho analisamos as grafias não-convencionais de palavra que se caracterizam pela presença e/ou ausência de espaços em branco no interior da palavra ortográfica em lugares não previstos pela convenção - denominadas, na literatura, de segmentações não-convencionais de palavra – nos atentamos ao que tem sido produzido no âmbito da relação entre a prosódia e a segmentação de palavras² no rol dos trabalhos em Linguística no Brasil.

Nesse âmbito, podemos citar três diferentes tendências, que, embora sejam detalhadamente discutidas em seção posterior, merecem destaque neste momento: a) tendência fonético-semântica - que engloba os trabalhos pioneiros de Abaurre e de Cagliari – Abaurre et al. (1985), Cagliari (1986), Abaurre (1994, 1999 e 2001); Cagliari (1992, 1998, 2001, 2002 e 2004); b) tendência fonológico-letrada - a partir dos trabalhos de Capristano (2004), (2007) e (2010); Chacon (2004a), (2005) e (2006); Tenani (2004);

¹ Nos alinhamos à concepção da heterogeneidade da escrita (cf. Corrêa, 1997, 2004) para a qual os “erros” são compreendidos como grafias não previstas pela convenção ortográfica e não como supostos reflexos da interferência da fala na escrita, como propõe Koch (1997) ou como sintomas de transtornos de aprendizagem, tal qual analisado por Zorzi (1997), distanciando-nos, portanto, dessas concepções.

² As segmentações de palavras tem atraído o interesse de estudiosos de diversas áreas do saber, além da Linguística, tais como Fonoaudiologia (cf. Zorzi 1997) Educação (cf. Cox; Assis-Peterson, 2001).

c) tendência fonológico- pedagógica - proposta nos trabalhos de Cunha (2004, 2010); Cunha e Miranda (2007)³

Voltadas para as relações entre constituintes prosódicos – tais como postulados por Nespor & Vogel (1986) e comentados adiante – e as segmentações não-convencionais de palavras, as pesquisas realizadas sob essas três tendências têm se debruçado sobre a escrita produzida em diversas fases de escolarização

- (1) a escrita nas séries iniciais do Ensino Fundamental – EF I – como os trabalhos de Abaurre et al. (1985), Cagliari (1986), Abaurre (1994, 1999 e 2001); Cagliari (1992, 1998); Capristano (2003, 2007, 2010); Chacon (2004, 2005, 2006) Cunha (2004); Cunha, Miranda (2007);
- (2) a escrita nas séries finais do Ensino Fundamental - EF II– Fiel (2018), Tenani (2009a, 2010, 2011b); Tenani, Paranhos (2011);
- (3) a escrita na alfabetização de jovens e adultos (EJA) – como observada por Tenani (2004); Ferreira (2010).

Voltada para os dados de escrita infantil à luz de sua relação com a prosódia da língua, esta pesquisa dialoga mais fortemente com os estudos que consideram a escrita produzida nessa fase de escolarização, entretanto, esta investigação têm como referência investigações realizadas em outros contextos de escolarização, tais como Fiel (2018), Tenani (2009a, 2010, 2011b), dada a relevância desses trabalhos na discussão em torno da relação entre as segmentações não-convencionais (das hipersegmentações, mais especificamente) no Ensino Fundamental II e aspectos da prosódia da língua.

A presente investigação, de cunho qualitativo, se dá, portanto, a partir de uma análise transversal de hipossegmentações e de hipersegmentações em 423 textos de alunos do 1º ao 5º ano de escolas públicas e de escolas privadas da cidade de Marília-São Paulo. As amostras analisadas integram o *EscIn*, banco de dados construído pelos pesquisadores do Grupo de Pesquisa *Estudos sobre a linguagem* (GPEL/CNPq), liderado pelo Prof. Dr. Lourenço Chacon (UNESP/Marília e São José do Rio Preto).

O interesse por esses dados se deu mediante o fato de que, como já demonstrado por Abaurre (1991) a relação leitura / escrita na vida e na escola é desigual quando se

³ Reflexões desenvolvidas no âmbito do Grupo de Estudos sobre a Aquisição da Linguagem Escrita (GEALE – FaE/UFPeL)

trata de comparação de crianças oriundas de classes sociais desiguais, especificamente do ensino público e do ensino privado. Segundo a autora:

crianças de classe média e classe alta, de escolas particulares, talvez controlem melhor (e mais cedo) os critérios de segmentação na escrita espontânea do que as crianças de classe baixa, que frequentam escola pública. (...) isso aparece como uma tendência geral, o que sem dúvida merece uma pesquisa mais aprofundada (ABAURRE, 1991, p. 211).

Segundo a referida autora, as crianças provenientes das classes sociais mais favorecidas vivem já em um contexto onde as pessoas leem e escrevem (embora em graus diversos) e isso favorece o aprendizado da leitura e da escrita, uma vez que estão cercadas de estímulos que as acompanham durante os anos escolares.

De igual modo, em estudo posterior, Cunha (2004) e Torquette (2016), mesmo que sob propósitos distintos, chegaram a resultados semelhantes aos de Abaurre (1991): o número de segmentações não-convencionais de palavras em textos de alunos de escolas públicas é maior comparado aos alunos do contexto privado.

Ao desenvolver um estudo quantitativo-comparativo de dados de segmentação não-convencional em textos de crianças do 1º ao 5º ano Ensino Fundamental I do ensino público e do ensino privado da cidade de Marília/SP, Torquette (2016) demonstrou que o número de ocorrências de segmentações não convencionais é 3,63 vezes maior em textos de alunos do ensino público do que nos alunos do ensino privado. Tais respostas foram associadas, pela autora, aos seguintes fatores: (a) os escreventes de escolas privadas tendem a controlar melhor e mais cedo os critérios de segmentação; (b) há maior autonomia de aplicação do currículo no ensino privado do que no público; (c) os alunos do ensino privado teriam maior inserção em práticas de letramento.

Face a esses dados de Torquette (2016), levantamos à hipótese de que, se a variável *tipo de escola* (pública ou privada) apontou percentuais consideravelmente distintos quando comparadas a distribuição das ocorrências de segmentações não-convencionais nos textos de Ensino Fundamental I nos dois contextos, a análise da configuração prosódica das dessas segmentações, também poderia apontar diferenças quando consideradas as variáveis *tipo de escola* e *série*.

Em outras palavras, os resultados obtidos por Torquette (2016), nos despertaram o interesse para a realização de uma investigação qualitativa (e não quantitativa, como ela

já o fizera), e de caráter transversal⁴, das ocorrências de segmentação não-convencional de palavras- hiperssegmentações e de hipossegmentações- nos textos de alunos de 1º ao 5º ano de escolas públicas e de escolas privadas município de Marília- SP, suscitando os seguintes questionamentos:

- 1) À luz do modelo fonológico da Fonologia Prosódica, como se configuram prosodicamente as hipossegmentações e as hiperssegmentações produzidas por crianças do 1º ao 5º ano do ensino público? E do ensino privado?
- 2) Haveria alguma variabilidade na configuração prosódica das segmentações não-convencionais quando comparados os dados, considerando-se as variáveis *tipo de escola* e *série*?

Com base nessas questões, a presente pesquisa busca, de modo mais amplo, investigar as configuração prosódica de ocorrências de segmentações não-convencionais de palavras em textos produzidos por crianças do 1º ao 5º ano de ensino fundamental de escolas públicas e de escolas privadas. Mais especificamente, a partir dos questionamentos supracitados, pretendemos:

- a) analisar prosodicamente as ocorrências de hipossegmentações e de hiperssegmentações de palavras observadas nos dois contextos (público e privado);
- b) comparar os dados, considerando as variáveis *tipo de escola* e *série* a fim de verificar a existência de variabilidade na configuração prosódica das estruturas de hiperssegmentação e de hipossegmentação.

Ao nos debruçarmos sobre as relações entre as segmentações não-convencionais e aspectos da prosódia da língua, embasados em uma concepção de escrita que defende-a em sua constituição heterogênea, tal como proposto por Corrêa (1997, 2004), nos distanciamos de perspectivas que defendem a suposta interferência da fala sobre a escrita,

⁴ Dado o interesse de estabelecer um comparativo entre duas populações (público e privado) o estudo transversal fornece ao pesquisador a flexibilidade de considerar outras variáveis (tal como a *série*) em uma constante, com apenas uma variável, nesse caso *tipo de escola*, como foco.

bem como nos afastamos de concepções patologizantes (para as quais os desvios de ortografia são observados como sinalizações de possíveis transtornos de aprendizagem).

Calcada no estudo da relação entre constituintes prosódicos e segmentações não-convencionais de palavra à luz da Fonologia Prosódica, tal qual postulada por Nespor e Vogel (1986), esta investigação assume uma perspectiva fonológico-letrada - se alinha aos trabalhos de Capristano (2003, 2007), Chacon (2004, 2005, 2006), Tenani (2004, 2009a, 2010, 2011b- por entender que, nas segmentações não-convencionais podem ser identificadas informações linguísticas diversas, relativas não apenas ao seu funcionamento, mas também às práticas sociais de linguagem orais/faladas e letradas/escritas, embora, nesta pesquisa, por questões metodológicas lancemos olhar sob aspectos estritamente linguísticos (neste caso prosódicos) que atuam sobre a escrita.

Nesse rol, a presente pesquisa se justifica pela possibilidade de ampliação do conhecimento sobre configuração prosódica das hipossegmentações e das hipersegmentações na escrita de crianças em contextos de escolas públicas e de escolas privadas e por trazer reflexões que evidenciam o funcionamento linguístico da escrita infantil, para que entendamos melhor a relação entre o falado e o escrito nesses contextos.

O conteúdo do presente trabalho está apresentado em cinco seções. À primeira delas, intitulada “Revisão teórica”, apresentamos ao nosso leitor o arcabouço teórico que fundamenta esta pesquisa, demarcando nossa filiação e percorrendo sobre as aproximações e os distanciamentos de nosso posicionamento teórico em relação a outras perspectivas, cujo objeto de investigação o é, também, a escrita infantil, sob o viés fonológico.

À segunda seção dedicamos atenção à explanação dos procedimentos metodológicos adotados nesta pesquisa, ao passo que, na Seção 3, descrevemos os resultados obtidos de nossa comparação, apontando as tendências encontradas e apresentando dados de nossa amostra. A discussão dos resultados e as possíveis hipóteses explicativas para os resultados obtidos são apresentadas na Seção 4. Por fim, na Seção 5, desenvolvemos a conclusão a que chegamos a partir desta investigação.

SEÇÃO 1: BASES TEÓRICAS

Neste capítulo discorreremos sobre as bases teóricas que subsidiam a investigação de aspectos da segmentação de palavras na escrita infantil. Tal exercício se dá em função do interesse de apresentar o estado da arte e de demarcar as concepções teóricas assumidas nesta investigação.

Primeiramente, apresentamos um panorama sobre as principais perspectivas teóricas a propósito da escrita, delimitando, ao final da seção, o lugar a partir do qual fazemos esta investigação (seção 1.1).

À subseção 1.2 discorreremos sobre a relação entre fala e escrita nos estudos produzidos no âmbito dos estudos linguísticos.

À subseção 1.3 são apresentadas algumas considerações em torno da noção de palavra no estudo das grafias de segmentação não-convencional. As bases teóricas para o estudo das segmentações não-convencionais, por sua vez, são discutidas à subseção 1.4, elencando (de 1.4.1 a 1.4.3) as principais tendências de análise no estudo desses dados de escrita.

Ao final do capítulo, na seção 1.5, apresentamos os principais pressupostos que respaldam a análise das segmentações não-convencionais em constituintes prosódicos, tais como postulados por Nespor e Vogel (1986). Passemos às discussões.

1.1 A escrita infantil enquanto objeto de estudo: breve retrospectiva

Nesta subseção apresentamos uma breve retrospectiva das contribuições de autores que abordaram aspectos do período tradicionalmente recoberto pela expressão *aquisição da escrita*⁵, a qual tem sido empregada, de modo amplo, para fazer referência ao processo que leva o sujeito ao domínio do sistema simbólico da escrita. Embora não consensualmente empregada, tal expressão tem sido comumente utilizada para fazer referência a determinados momentos do processo de contato da criança com a escrita. Uma vez que os dados analisados nesta investigação recobrem a escrita de crianças em

⁵ Como assinala Abaurre (1999), essa não é a expressão mais apropriada, uma vez que ela supõe a existência de um sistema pronto a ser adquirido pela criança, além disso, como destaca Capristano (2003) o emprego desse termo, de modo amplo, parece não contemplar a existência do outro como instância representativa da língua, além de desconsiderar a própria complexidade do funcionamento da escrita.

fase inicial de escolarização (1º ao 5º ano do EF I), cabe uma breve discussão em torno da escrita infantil, enquanto objeto de estudo, em seus aspectos históricos.

Reconhecida pela abundância na produção de reflexões sobre o ensino da escrita, especialmente no que diz respeito à alfabetização (cujo interesse era o de promover o enfrentamento das práticas de alfabetização pautadas na tradição do ensino cartilhesco) a década de 80 foi marcada, no Brasil, pela publicação do trabalho *Psicogênese da Língua Escrita*, de autoria de Emilia Ferreiro e Ana Teberosky. Construída no âmbito das pesquisas em Psicologia, tal investigação caracteriza-se por explicitar o processo caracterizado como aquisição da escrita, relacionando-o à capacidade cognitiva de aprendizagem da criança, inspirando-se, de um lado, nos avanços dos estudos sobre a linguagem filiados à perspectiva gerativista e, de outro, no aporte construtivista de Jean Piaget.

Nessa construção, o interesse central das autoras era o discutir o processo cognitivo que leva à aprendizagem, deslocando o papel dos métodos de ensino para o papel da criança. Nesse processo, para as autoras, a criança “[...] constrói sistemas interpretativos, pensa, raciocina e inventa, buscando compreender esse objeto social complexo que é a escrita”. (FERREIRO; TEBEROSKY, 1985, p. 7).

Tendo a criança como centro de toda aprendizagem em detrimento do método e do conteúdo, as autoras compreendiam o sujeito tal qual ele o é na teoria piagetiana: intelectualmente ativo, que procura adquirir conhecimento, interessado na compreensão da realidade do mundo que o cerca e em buscar respostas às inquietações que dela decorrem. Ferreiro; Teberosky (1985), nesse sentido, põem à crítica o tratamento metodológico do ensino de leitura e de escrita que, à época, organizava-se partindo da premissa de que a criança nada sabia sobre a língua até ingressar na escola.

Dado o fato de considerarem que a criança traz diversas experiências com a escrita, as quais antecedem o seu ingresso no ensino formal, os estudos de Ferreiro; Teberosky (1985, 1999) assinalaram um significativo avanço para os estudos sobre a escrita infantil: uma vez consideradas, naquele momento, as diferentes hipóteses de escrita empregadas pelos alunos ao longo do processo denominado aquisição da escrita, os aprendizes passam a ser reconhecidos como sujeitos ativos da aprendizagem.

O interesse pelos temas amplos que cercam a alfabetização, sobretudo a aquisição da linguagem, não ficou restrito às pesquisas no interior da Psicologia ou da Educação. Ainda nos anos 80, orientados pelo anseio de formalizar considerações sobre como as crianças lidam com a escrita, estudiosos do domínio da Linguística, em subáreas

como a da Aquisição da Linguagem, foram impulsionados a investir em estudos sobre a chamada aquisição, de modo geral, o que levou a diversidade de abordagens teóricas, oriundas tanto da própria Linguística quanto da Psicologia. Conforme demonstra Tenani (2017 p. 583), ao tecer uma retrospectiva sobre os caminhos de investigação sobre escrita no Brasil, podem-se identificar, nas décadas finais do século passado:

pesquisas ancoradas em princípios da Teoria Gerativista, que define como questão central para os estudos linguísticos explicar a competência linguística (que compreende investigar relação entre mente e linguagem, cf. CHOMSKY, 2009), ao lado de pesquisas que mobilizam fundamentos da abordagem Sócio-interacionistas ou da Teoria Histórico-Cultural, baseadas nos estudos de Vygotsky (1988) sobre a relação entre pensamento e linguagem.

No rol das pesquisas em Linguística, a escrita produzida por crianças dos primeiros ciclos escolares também interessou a pesquisadores que investigavam aspectos fonéticos e fonológicos do português. Isso porque, como argumenta Tenani (2017), as produções escritas iniciais se mostraram como fonte de dados relevantes para as teorias linguísticas, de modo amplo, e para as teorias fonológicas, de modo específico.

Entre os interessados, no âmbito dos aspectos fonéticos e fonológicos, encontravam-se Maria Bernadete Marques Abaurre e Luiz Carlos Cagliari. Em seus trabalhos, Abaurre (1989) e Cagliari (1986) introduziram aos estudos uma visão sobre o processo de aquisição da escrita que, apesar de dialogar com o trabalho de Ferreiro; Teberoski (1985) em alguns pontos⁶, distanciava-se da perspectiva dessas autoras – em especial, da linearidade das etapas e da universalização do escrevente, uma vez que a psicologia genética preconizava um sujeito idealizado e universal. A produção escrita da criança era, na visão desses autores, um processo não-linear, não cumulativo, caracterizado por uma constante reelaboração, mediado pelo papel do outro, uma vez que, para esses autores, “a relação do sujeito com a linguagem é mediada, desde sempre, pela relação com o OUTRO” - interlocutor fisicamente presente ou representado e necessário ponto de referência para esse sujeito em constituição (ABAURRE et al., 1997, p. 41).

A partir da análise das grafias de segmentação não-convencional, dado linguístico que inicialmente interessou à Abaurre (1988), a autora postulou importantes reflexões a respeito da relação entre fala e escrita.

⁶ Aspecto comum a esses estudos é o de fazerem considerações sobre a reflexão sujeito/língua e de considerarem que os erros cometidos pelos aprendizes de escrita/leitura, na verdade, são hipóteses, indícios de um processo em curso, do que a autora chama de aquisição.

Uma vez que tal relação é um dos principais argumentos utilizados para explicar, dentre outras questões, a motivação para a ocorrência de grafias não convencionais (como as de segmentação em palavras, por exemplo) acreditamos que as noções de *fala* e *escrita* devem ser claramente definidas aqui, bem como a natureza da relação entre ambas. Portanto, na seção que segue, resenhamos alguns trabalhos produzidos no âmbito da Linguística que se debruçam sobre tais noções.

1.2 Da relação entre fala e escrita

Em estudo pioneiro sobre a escrita infantil, Abaurre (1988a) caracterizou-se por apresentar, em sua pesquisa, duras críticas à visão tradicional a qual defendia que as crianças tenderiam a escrever do modo como falavam, ou, em outras palavras, de que a escrita seria o reflexo da fala. Um dos principais pontos defendidos pela autora foi, portanto, o de que “ninguém se programa para simplesmente ‘escrever como fala’, em termos de uso das letras e das suas relações com o som, seja em termos de segmentação, seja em termos das próprias estruturas linguísticas.” (ABAURRE, 1988a, p.137).

Sob esse enfoque, a autora apontou, naquele momento, para a existência de duas perspectivas tradicionais a serem superadas na abordagem da relação fala/escrita. A primeira delas era aquela que defendia uma dependência radical da escrita à fala. A segunda, em contrapartida, defendia a autonomia (também extrema) da escrita em relação à fala.

Nesses paradigmas, aceitavam-se, respectivamente, que: (a) a escrita é reflexo da oralidade⁷; (b) a escrita é autônoma a ponto de não sofrer efeitos da oralidade. Ambas as proposições eram, segundo a autora, passíveis de questionamentos. Isso se deve, excepcionalmente, ao fato de que tratam do que a autora chama de oralidade e da escrita como fenômenos independentes.

Contrapondo-se a essas visões tradicionais, Abaurre (1988a) destacou, portanto, a partir da análise de um corpus composto por centenas de textos de crianças de escolas públicas e privadas, que a produção escrita infantil compreendia o momento em que a

⁷ De modo geral, na literatura que se volta para a escrita infantil, é comum que a escrita, enquanto fato linguístico, seja relacionada ao que os autores, genericamente e sem defini-la, chamam de oralidade (na maioria das vezes confundida e/ou entendida como a própria fala). Mais adiante, esclareceremos de que modo relacionamos fala, escrita, oralidade e letramento.

criança registraria suas hipóteses sobre a escrita, a partir das relações que estabelecia entre fala e escrita e que esta não poderia, portanto, ser observada como mera transcrição da fala.

Por outro lado, distanciado dessa concepção, está o ponto de vista defendido por Ingedore Koch. Em trabalho publicado em 1997 denominado “Interferências da oralidade na aquisição da escrita”, a autora defende que, antes de entrar para a escola, as crianças têm internalizado, como principal modelo, o texto oral, a partir do qual elas baseariam suas primeiras produções escritas. Isso explica o porquê de, na aquisição da escrita, segundo a autora, existirem muitas interferências da fala sobre o texto escrito. As características típicas da fala poderiam, ainda, perdurar na escrita das crianças por anos, durante todo o período de escolarização. Nesse viés, fala e escrita, enquanto duas modalidades distintas de uso da língua (homogêneas), possuem características próprias, embora, assim como Abaurre (1988), Koch (1997) acredite que a escrita não é uma mera transcrição da fala.

A partir de um levantamento de diferenças entre o texto oral e o texto escrito em atividades de produção escrita realizadas por alunos de 1ª e 2ª séries, Koch (1997) examina o que ela denomina “principais interferências” do oral na escrita infantil. Para tanto, ela divide-as em oito categorias, sendo elas: 1) os *mecanismos de referência* (marcados pelo uso de recursos anafóricos e catafóricos indevidamente; 2) as *repetições* (no texto falado, a repetição ocorre com extrema frequência, podendo mesmo ser considerada um dos mecanismos organizadores dessa modalidade textual. 3) a *organização textual com elementos continuadores típicos da fala* (uso frequente de expressões como *aí, daí, então, daí então*; 4) a *ausência de pontuação para ligar ideias*; 5) *discurso citado* (falta de marcação de discursos diretos); 6) a *segmentação gráfica em palavras*; 7) *grafias a partir da pronúncia da palavra*, as palavras são escritas da forma como se realizam foneticamente; 8) *correções comuns nos textos falado* (quando se observa o registro gráfico de hesitações da fala).

Estando o problema da segmentação gráfica figurando entre tais categorias elencadas por Koch (1997), a autora defende que esse tipo de interferência pode ser justificado da seguinte forma: 1) segmentações como *convocê, arainha, vamologo* resultam da apreensão, por parte das crianças, dos vocábulos fonológicos⁸ (M. Câmara,

⁸ O vocábulo fonológico, nos termos de Mattoso Câmara Jr. (1972) é delimitado pelo critério de presença de tonicidade e dentro de um grupo de forças em contraste. Para o autor, no registro do Português do Brasil há, a rigor, uma pauta acentual para cada vocábulo.

1972); 2) segmentações como *em bora, en con tado, na quela*, por sua vez, são explicadas a partir da afirmação de que “a criança, tentando efetuar a segmentação gráfica adequada, acaba, por vezes, caindo no extremo oposto, isto é, “picando” demais a palavra” (KOCH, 1997, P.37).

Lançando um olhar mais apurado para a análise realizada por Koch (1997), vemos que, especialmente no que diz respeito à segmentação gráfica (e aos exemplos de hipersegmentação), a explicação dada pela autora para esses fenômenos não se sustenta, uma vez que ela contradiz o princípio de interferência da fala na escrita sustentado em seus trabalhos. Em outras palavras, se em determinados momentos as crianças estão tentando realizar juma segmentação adequada com os padrões ortográficos (a autora utiliza a expressão “tentativas de efetuar a segmentação gráfica adequada” p. 37) então, nessas grafias essas crianças não estariam mobilizando apenas informações do texto falado, uma vez que, como sabemos, o uso de recursos gráficos para a divisão os enunciados em palavras é uma propriedade exclusiva da escrita, ligada tão somente ao aspecto gráfico-visual. Nesse sentido, essas grafias reafirmam a existência, na escrita, da atuação de fatores ligados à inserção da criança em práticas letradas⁹.

Sob uma outra perspectiva se alinha o trabalho de Marcuschi (1997). Ao tratar a relação entre fala e escrita esse afirma que “não é possível observar satisfatoriamente as semelhanças e diferenças entre oralidade e escrita sem considerar a distribuição de seus usos na vida diária” (p. 119). Nesse caminho, o que o autor propõe é uma revisão de concepção de língua e de texto (não centrados mais no código) que passam a ser vistos, nesse enfoque, como práticas sociais.

“mais urgente do que identificar primazias ou supremacias entre oralidade e escrita, e até mesmo mais importante do que observar oralidade e escrita como simples modos de uso da língua é a tarefa de esclarecer a natureza das práticas sociais que envolvem o uso da língua (oral e escrita) de um modo geral” (MARCUSCHI, 1997, p 120).

Para tratar da relação entre fala e escrita, Marcuschi (2001) assume, assim como Koch (1997), que essa se refere a duas diferentes modalidades da língua. Entretanto, para endossar um tratamento mais produtivo a essa relação, o autor propõe uma observação de caráter teórico, que compreende à distinção, sobretudo, entre duas dimensões de relações

⁹ Para tratar da problemática que envolve o conceito de letramento nos ancoramos em Corrêa (1997) para quem todo fato linguístico (fala/escrita) se vincula a uma prática social (oralidade/letramento). Tais noções são explicitadas na seção que segue.

no tratamento da língua falada e da língua escrita. De um lado há a necessidade, segundo o autor, de se distinguir duas práticas sociais (oralidade e letramento) e, de outro, há que se distinguir duas modalidades de uso da língua (fala e escrita).

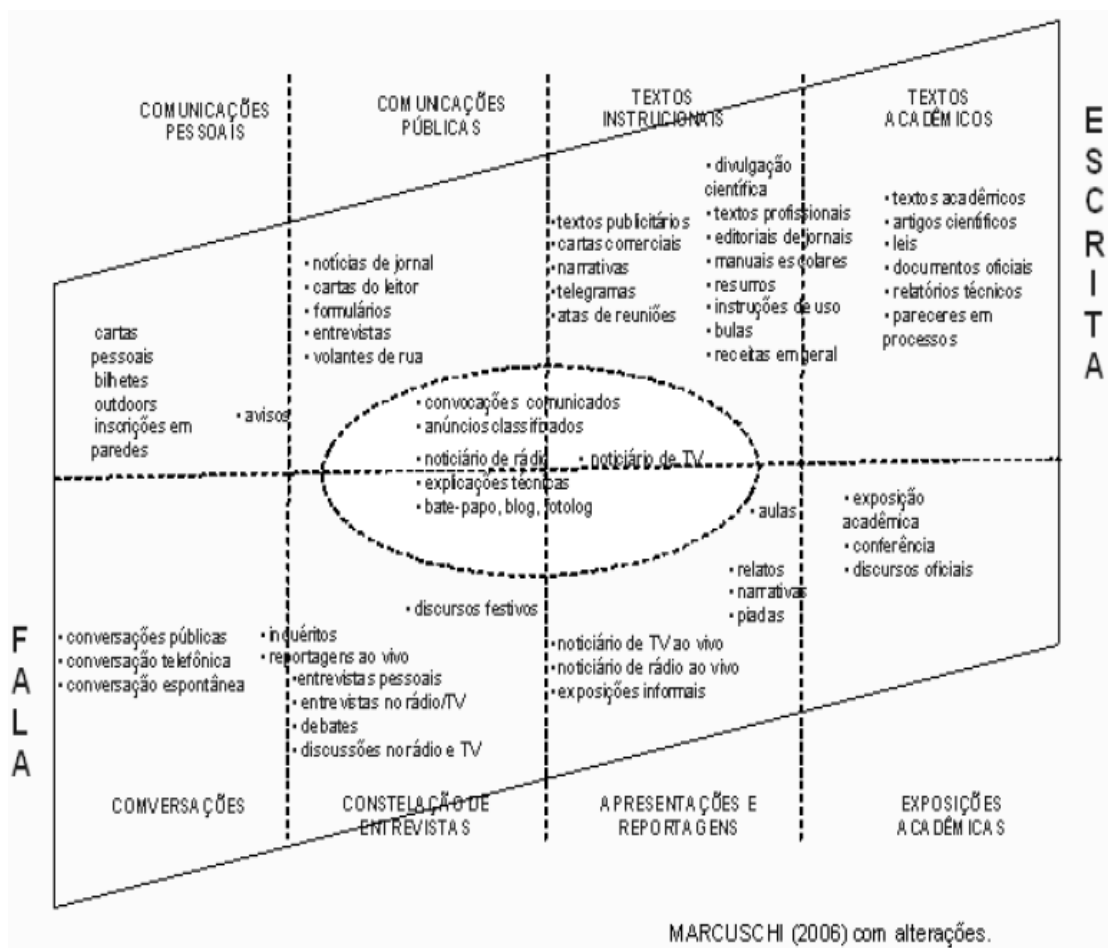
A oralidade, vista como uma prática social, se apresenta de variadas formas ou gêneros textuais que se dão nos mais variados contextos de uso (formais, informais) ao passo que o letramento, para Marcuschi (2001), é o uso da escrita na sociedade e pode ir desde a apropriação mínima da escrita (por exemplo um indivíduo que é analfabeto, mas sabe o valor do dinheiro, sabe distinguir mercadorias pela marca, mas não escreve cartas, nem lê jornal). De outro lado o autor distingue, ainda, fala e escrita:

“A fala seria uma forma de produção textual-discursiva oral, sem a necessidade de uma tecnologia além do aparato disponível pelo próprio ser humano. A escrita seria, além de uma tecnologia de representação abstrata da própria fala, um modo de produção textual-discursiva com suas próprias especificidades” (MARCUSCHI, 2001, p. 126).

Crítico da tendência tradicional que abordava as noções de fala e de escrita de forma dicotômica – Bernstein (1971) Labov (1972), Halliday (1985) –, Marcuschi (2001) afirma que tal modelo, ao separar forma e conteúdo, língua e uso, toma a língua unicamente como sistema de regras, o que faz com que, tal perspectiva assuma um caráter estritamente formal, que, embora forneça bons resultados em uma descrição empírica, ignora os fenômenos dialógicos e discursivos.

Nesse caminho, Marcuschi (2001) propõe que as relações entre fala e escrita não são óbvias nem lineares, uma vez que elas refletem um constante dinamismo, o qual se dá “dentro de um *continuum* tipológico das práticas de produção textual e não na relação dicotômica de dois polos opostos” (MARCUSCHI, 2001, p. 37). Marcuschi sustenta, por conseguinte, que “tanto a fala como a escrita apresentam um *continuum de variações*, ou seja, a fala varia e a escrita varia. Assim, a comparação deve tomar como critério básico de análise uma relação fundada no *continuum* dos gêneros textuais para evitar as dicotomias estritas” (MARCUSCHI, 2001, p. 42, grifos do original). O gráfico a seguir fornece uma noção dessa postura:

Figura 01- O contínuo de gêneros textuais, conforme Marcuschi (2001)



Nessa figura, vemos que fala e escrita se localizam em lugares opostos no contínuo (extremidades), ao passo que os gêneros textuais são distribuídos gradativamente no contínuo. Tais gêneros, falados ou escritos, podem ser homogêneos ou mistos (envolvem as duas modalidades -como os que aparecem no centro do círculo) e são distinguidos e correlacionados no contínuo tipológico de variações, a partir das quais Marcuschi (2001) ilustra as semelhanças e diferenças entre as modalidades¹⁰ falada e escrita.

¹⁰ Nos trabalhos de Koch (1997) e de Marcuschi (2001) encontramos o termo “modalidade” para fazer referência aos fatos da língua falados e escritos, os quais corresponderiam a duas modalidades distintas. Em Corrêa (2001), em cuja perspectiva o presente trabalho se assenta, o autor afirma a existência de *modos de enunciação* e não de modalidades, como veremos na seção posterior.

Segundo o autor,

tanto a fala quanto a escrita apresentam um continuum de variações, ou seja, a fala varia e a escrita varia. Assim, a comparação deve tomar como critério básico a análise de uma relação fundada no continuum dos gêneros textuais para evitar as dicotomias estritas (MARCUSCHI, 2001, p. 138).

Analizando a visão de Marcuschi (2001), percebemos que seu trabalho marca uma contribuição importante na tentativa de superar a dicotomia radical entre fala e escrita. Entretanto, segundo Corrêa (2001) tal contínuo parece recuperar, metodologicamente, essa dicotomização, isso porque a compartimentação de gêneros em um contínuo tipológico pressupõe a existência de polos (homogêneos) nos extremos do contínuo (em que fala e escrita são tratados como independentes, embora estejam relacionados). Portanto, para o autor, o contínuo tipológico proposto por Marcuschi (2001) estaria reafirmando as diferenças entre o que é oral/falado e o letrado/escrito, sem, todavia, postular a dicotomia.

Corrêa (2001), contrariamente às posições teóricas destacadas até o momento - Abaurre (1988), Koch (1997) Marcuschi (2001) – defende, por sua vez, que não existiriam, portanto, textos e/ou enunciados unicamente orais ou unicamente escritos, porque, para o autor há um *modo heterogêneo de constituição da escrita*, que assinala vínculos entre práticas orais e letradas.

A partir da noção de heterogeneidade da escrita, Corrêa (2001), defende que a presença do falado no escrito não pode ser explicada nem como uma questão de interferência e nem em relação à ocorrência de gêneros textuais mistos- como defendido por Koch (1997) e Marcuschi (2001) – mas se trata de uma evidência de que a escrita é heterogênea em sua constituição.

Argumentando favoravelmente a essa constituição heterogênea da escrita, o autor defende que tal constituição está ligada a uma particularização, para o domínio da escrita, do encontro das práticas orais/faladas e letradas/escritas (CORRÊA, 2004, p. 34). Nesse encontro, na perspectiva do autor, todo fato linguístico (falado ou escrito) está vinculado a uma prática social (oral/letrada), ou, em outras palavras, seria/ resultaria de uma prática social. Assim, a relação oralidade/fala/letramento/escrita fundaria, então, a própria heterogeneidade da escrita, o que significa dizer que a heterogeneidade é constitutiva (e não acessória) da escrita.

Essa postulação parte, por um lado, do interesse de superar as abordagens que tinham na relação entre o falado e o escrito um processo de interferência, por outro, o

autor intencionava afastar-se das “[...] avaliações estereotipadas sobre a escrita que tomam como modelo abstrato – literário ou não – de boa escrita” (CORRÊA, 2004, p. 86).

Para argumentar sobre a relação entre as práticas sociais orais/faladas e letradas/escritas nos modos de enunciação ¹¹falado e escrito, o autor propõe, portanto, a revisão das noções de letramento e de oralidade. Ao tratar do fundamento teórico que dá base à proposição da heterogeneidade da escrita, Corrêa (2004), afirma que, para discutir a problemática que envolve tais conceitos é necessário, primeiramente, desvincular letramento de alfabetização¹², uma vez que, para ele, a alfabetização não é condição primeira para o letramento. Para tanto, o autor propõe um conceito amplo de letramento, por meio do qual

“pretendo valorizar as habilidades atestadas por aqueles indivíduos que, mesmo não tendo acesso à alfabetização ou mesmo mantendo-se, na maior parte do tempo, alheios às práticas de leitura e escrita tal como foram consagradas, também fazem a história da língua e da sociedade por meio do modo oral de registro da memória cultural” (CORRÊA, 2004, p. 128).

A reconceptualização proposta por Correa (2004), dessa forma, diverge dos estudos cujo sentido de letramento é mais restrito, os quais têm o domínio do código escrito (alfabetização) como requisito para o letramento, ¹³ para ele, ao se caracterizar o indivíduo letrado, a ênfase têm recaído no tipo de prática social que envolve o uso da leitura e da escrita, e, nesse caso, mais especificamente, nas práticas canônicas de leitura e escrita. Nesse sentido para o autor, é a participação direta ou indireta do indivíduo nas práticas sociais, as quais, em última instância, envolvem codificação de sons em letras e decodificação dessas letras em sons (alfabetização), o que tem caracterizado as fronteiras do letramento.

A concepção de escrita como resultado do entrelaçamento de práticas orais/faladas e letradas/escritas, tal qual postulada por Corrêa, fundamenta a produção dessa investigação, uma vez que entendemos que os dados de escrita não são marcas de

¹¹ Fala e escrita, são vistos como *modos enunciativos* ligados intrinsecamente às práticas sociais de oralidade e letramento.

¹² Alfabetizar e letrar, na perspectiva de Soares (1998) são duas ações distintas, porém inseparáveis. Esse sentido é diferente do postulado por Corrêa (2004).

¹³ Uma das perspectivas que segue esse viés é o de Soares (1998) para quem “um indivíduo alfabetizado não é necessariamente um indivíduo letrado; alfabetizado é aquele indivíduo que saber ler e escrever, já o indivíduo letrado, indivíduo que vive em estado de letramento, é não só aquele que sabe ler e escrever, mas aquele que usa socialmente a leitura e a escrita.”

uma suposta intervenção e/ou interferência da fala na escrita e tampouco registram apenas a relação entre duas tecnologias diferentes. Dado o nosso recorte metodológico e o objetivo de nossa pesquisa - o de investigar a configuração prosódica das hipossegmentações e das hiperssegmentações de palavras nos contextos de ensino público e de ensino privado- lançamos o olhar, neste momento, para os fatos linguísticos, embora tenhamos consciência de que da escrita emergem informações diversas que estão para além da ordem da língua, exclusivamente.

1.3 A propósito da noção de palavra no estudo das grafias de segmentação não-convencional: algumas considerações.

Dado o nosso objeto – as segmentações não-convencionais de palavra –, não podemos nos furtar à amplitude envolvida na conceituação do que seja *palavra*, visto que o empreendimento de uma noção é uma tarefa complexa, sendo, pois, atravessada por critérios diversos selecionados para sua análise. Desse modo, a formalização do conceito de palavra é perpassada pelo nível de análise linguística em questão (prosódico, fonológico, morfológico, sintático, semântico) bem como pela perspectiva analítica assumida quando do empreendimento de uma definição.

Assim, haja vista que tanto a definição do termo *palavra*, quanto o reconhecimento dos seus limites se dão não apenas em função dos diferentes níveis de análise, mas também em função da abordagem teórica a que se filia cada estudo, se faz pertinente discorrer sobre a noção de palavra e explicitarmos nosso enquadramento teórico, uma vez que, delimitar o entendimento da noção de palavra faz-se necessário porque, conforme aponta DONADEL (2011), seus limites podem esbarrar em fronteiras completamente distintas:

Para a fonologia, palavra pode ser entendida como conjunto de fonemas organizados em sílabas e portador de relações de proeminência acentual. Para a sintaxe, é possível que palavra seja a unidade mínima de uma árvore sintática que pode expressar relações com outras palavras do sintagma. Para a morfologia, palavra pode ser entendida como um conjunto de morfemas associados, ou como unidade lexical abstrata que prevê variações a partir do acréscimo de outras informações possíveis, sejam elas lexicais (portadoras de sentido) ou gramaticais (que expressem funções sintáticas) (DONADEL, 2011, p. 216).

Uma vez que nosso objeto de análise são as segmentações não-convencionais de palavras e, na medida em que as estruturas que resultam desse tipo de segmentação estão fortemente ancoradas em aspectos prosódicos, nosso olhar para a descrição de

características do conceito de palavra é orientado por dois fatores: (a) o nível linguístico de análise (no nosso caso, o fonológico, uma vez que as palavras apresentam uma organização fônica que se mostra pertinente à interpretação das segmentações não-convencionais as palavra); (b) nossa filiação teórica (olhando para a segmentações não-convencionais e fatos linguísticos, tendo em vista a relação entre a dimensão prosódica da língua e as informações letradas que atravessam a noção de palavra dos escreventes). Isso se dá porque, como a literatura tem demonstrado, na grafia das palavras os escreventes, não levam em conta apenas os critérios convencionais.

No que tange à convenção de palavra escrita, sabe-se que a ortografia do Português Brasileiro, doravante PB, faz uso de noções morfossintáticas/semânticas, entretanto, para Bisol (2004) é necessário considerar a relevância de noções fonológicas no registro ortográfico dos limites de palavra, uma vez que, como mostrado por Abaurre (1991), as crianças conceituam a segmentação escrita de forma diferente dos critérios estabelecidos pelas convenções ortográficas.¹⁴

A discussão do estatuto de palavra, no trabalho pioneiro de Camara Jr. (1972) tem respaldo na distinção entre palavra fonológica (vocábulo fonológico) e palavra morfológica, visto, para o autor, não haver coincidência entre essas duas noções de palavras – os limites das palavras fonológicas são, segundo o autor, diferentes daqueles das palavras morfológicas: há palavras morfológicas que correspondem a mais de uma palavra fonológica e o contrário também. Para o autor, há diferentes critérios linguísticos em jogo na definição de palavra: fonológicos, morfossintáticos, (orto)gráficos.

Ao diferenciar essas noções, o autor sinaliza que palavra morfológica compreende palavras lexicais – como adjetivos, nomes, verbos, que correspondem a classes abertas – e palavras funcionais – como preposições e conjunções, que correspondem a classes fechadas –, ao passo que a palavra fonológica engloba palavras com acento (palavras fonológicas) e sem acento (clíticos).

Nesse mesmo sentido, Bisol (2004) defende a distinção entre palavras morfológica e prosódica.

A primeira compreende *palavras lexicais*, como nome, adjetivo e verbo, classes abertas, e *palavras funcionais* como preposição, conjunção e

¹⁴ Haja vista a existência de uma convenção, podemos dizer que a concepção de palavra é subjacente à de palavra gráfica, a qual seria resultado de um acordo entre aqueles que a usam, sendo considerada como adequada ou inadequada, em algumas abordagens, ou como certa e errada, em outras. Desse modo, escrever envolveria um movimento de transição da língua (na enunciação falada) a uma norma estabelecida convencionalmente.

determinativos, classes fechadas. A segunda distingue palavras com acento e sem acento, respectivamente palavras prosódicas e clítics (BISOL, 2004, p. 59).

A falta de correspondência entre os limites da palavra morfológica e prosódica pode ser vista, segundo a autora, na configuração das palavras compostas. Tome-se por exemplo, a palavra *guarda-chuva*. Vê-se nesse caso que, segundo a autora: a) morfollogicamente [guarda] e [chuva] funcionam juntas, constituindo uma única palavra – com um significado; b) fonologicamente, entretanto, [guarda] e [chuva] são duas palavras, unidades independentes, portadoras, cada uma, de um acento primário.

No que diz respeito à definição de palavra sob o ponto de vista prosódico, Bisol (2004) defende que uma das complexidades reside no fato de haverem formas lexicais átonas, chamadas de clítics¹⁵ – por exemplo: *e, me, de, o*, que, por serem considerados estruturas dependentes, não possuem acento, e, por isso, marcam um lugar problemático na noção de palavra.

Em se tratando da palavra no âmbito da Fonologia Prosódica, se faz pertinente explicitar que as noções usadas para discutir a formação desse constituinte não são as mesmas para todas as línguas. Nesse sentido, convém olhar para as regras de formação desse constituinte segundo o modelo de Nespor e Vogel (1986) e sua aplicação em alguns estudos sobre o domínio da palavra prosódica¹⁶ no PB – Bisol (2004); Toneli (2014). Convém comentar que nosso propósito, nesta seção não é o de explorar sistematicamente a definição e a caracterização desse domínio na variedade brasileira, uma vez que o empreendimento desta tarefa fora realizado em outros trabalhos - Toneli 2009, 2014 - mas o de fornecer os pressupostos acerca desse domínio prosódico tais quais os postulados por Nespor e Vogel (1986).

Tendo em vista que nesse modelo fonológico as línguas têm uma estrutura prosódica que é formada por constituintes que têm como referência a estrutura

¹⁵ Nas teorias fonológicas o lugar do clítico tem sido abordado sob diferentes pontos de vista. Para Simioni (2008) o clítico é prosodizado na frase fonológica, ou seja, ele se anexa a uma frase fonológica. Vigário (2007) por outro lado propõe uma nova proposta para resolver a problemática do grupo clítico, defendendo a existência de um constituinte autônomo superior à ω e inferior a ϕ . Conforme Bisol (2005), deve-se considerar o clítico como uma palavra dependente que se une a uma palavra de conteúdo na formação de uma unidade prosódica: o grupo clítico – como em “osporquinhos” ou “mechamaram”, em que os clítics *os* e *me* se juntaram à palavra de conteúdo seguinte: *porquinhos* e *chamaram*, respectivamente, isso porque, do ponto de vista fonológico [os] e [me] em razão de não apresentarem acento, apoiam-se no acento da palavra de conteúdo ou seguinte (BISOL, 2004).

¹⁶ O termo palavra prosódica aparece nos trabalhos de Bisol (2004), Toneli (2009), Vigário (2003) como referência ao conceito de palavra fonológica tal qual formulado modelo de Nespor e Vogel (1986) para demarcar uma filiação teórica que se distancia da perspectiva fonológica estruturalista.

morfossintática (porém não é isomórfica a ela), o domínio da palavra prosódica é construído com base em regras de mapeamento que fazem uso de noções não fonológicas.

Nesse sentido, a palavra prosódica, no modelo de Nespor e Vogel (1986) está relacionada a uma unidade linguística formada por um ou mais membros que estabelecem entre si uma relação de dominância. Dentro do domínio da palavra prosódica (ω), desse modo, podem se reajustar as sílabas e os pés métricos, respeitando-se os princípios universais e as restrições específicas da língua.

Nespor e Vogel (1986) defendem a existência de duas possibilidades de tamanho para a palavra prosódica: domínio de (ω) menor ou igual ao elemento terminal. Entretanto, no caso do PB, Bisol (2004) mostra que o nó sintático terminal pode incluir três opções: a-) **ser igual** a uma palavra morfológica, como ocorre em palavras lexicais simples (não compostas) e desprovidas de prefixo ou sufixo portador de acento primário, como em [prefixo]; b-) **ser maior**, como é o caso de palavras lexicais que sofreram ressilabação, como [mar] [aberto] - (ma)w (r aberto); c-) **ser menor** do que a palavra morfológica como em [(super) ω (mercado) ω].

Ao discutir a interação entre as unidades morfológica e fonológica, com ênfase nas funções da palavra prosódica no PB, Bisol (2004) apresenta evidências desse domínio no PB, afirmando primariamente que a condição mínima para se estabelecer uma palavra prosódica é que a unidade seja dotada de apenas um acento primário¹⁷

Diante o exposto, a seguir discorreremos sobre as bases teóricas que têm sustentado os estudos sobre as segmentações não-convencionais de palavra, apresentando tendências que têm suscitado reflexões por parte dos pesquisadores do Grupo de Pesquisa *Estudos sobre a linguagem* (GPEL/CNPq) e do grupo *Estudos sobre a Aquisição da Linguagem Escrita* (GEALE/UFPeL/CNPq).

1.4 Segmentação não-convencional de palavra: tendências de investigação

Para discorrer sobre a segmentação não-convencional, precisamos, primeiramente, retomar o que concebemos como *segmentação* na escrita. Para tanto, retomaremos discussões de Abaurre (1991), Silva (2004), Capristano (2007) e Tenani (2011).

¹⁷ À seção 1.5 a noção de palavra prosódica é retomada.

Ao elaborar hipóteses explicativas sobre a segmentação na escrita inicial, Abaurre (1991) destaca que, na definição dos espaços em branco entre as sequências de letras, a escrita alfabética do português faz uso de critérios morfológicos. No entanto, segundo a autora, por não dominarem tais critérios, as crianças conceituam a segmentação escrita de forma diferente daqueles estabelecidos pelas convenções ortográficas e, nesses momentos, elaboram hipóteses conflitantes para segmentar suas escritas. Por não atenderem às convenções ortográficas, a estruturas resultantes dessa escrita são denominadas segmentações não-convencionais.

Em um de seus trabalhos sobre aspectos de segmentação na escrita infantil, Capristano (2007) define-a como "[...] um recurso ligado ao aspecto gráfico visual do enunciado escrito que possibilita, de diferentes maneiras, a divisão do fluxo textual em porções menores – recursos como espaços em branco entre palavras, espaços em branco¹⁸ referentes a parágrafos, unidades de escrita delimitadas por sinais de pontuação, entre outros." (CAPRISTANO, 2007, p. 2).

A segmentação não-convencional de palavra gráfica pode ser definida, nos termos de Tenani (2011), a partir da ausência ou partir da presença do espaço em branco que delimita a palavra em local não previsto pelas convenções ortográficas.

Conforme Silva (1991) as segmentações não-convencionais dividem-se, basicamente, em dois grupos, *hipossegmentações*, ocorrências que apresentam falta de espaço entre fronteiras vocabulares, como em “derepente”, “nosofa”, “oque”; e *hipersegmentações*, ocorrências que apresentam inserção de espaço no interior da palavra, como em “toda via”, “a pareceu”. Chacon (2006) acrescenta a esses, um terceiro tipo de segmentação, denominado de *mesclas*, que caracteriza-se pela presença, em uma mesma estrutura, de hipo e hipersegmentação. A esse terceiro grupo de segmentações (para palavras grafadas tanto de forma hipossegmentada quanto hipersegmentada) Cunha de refere como *segmentação híbrida*, como em “escom deuce” (escondeu-se)

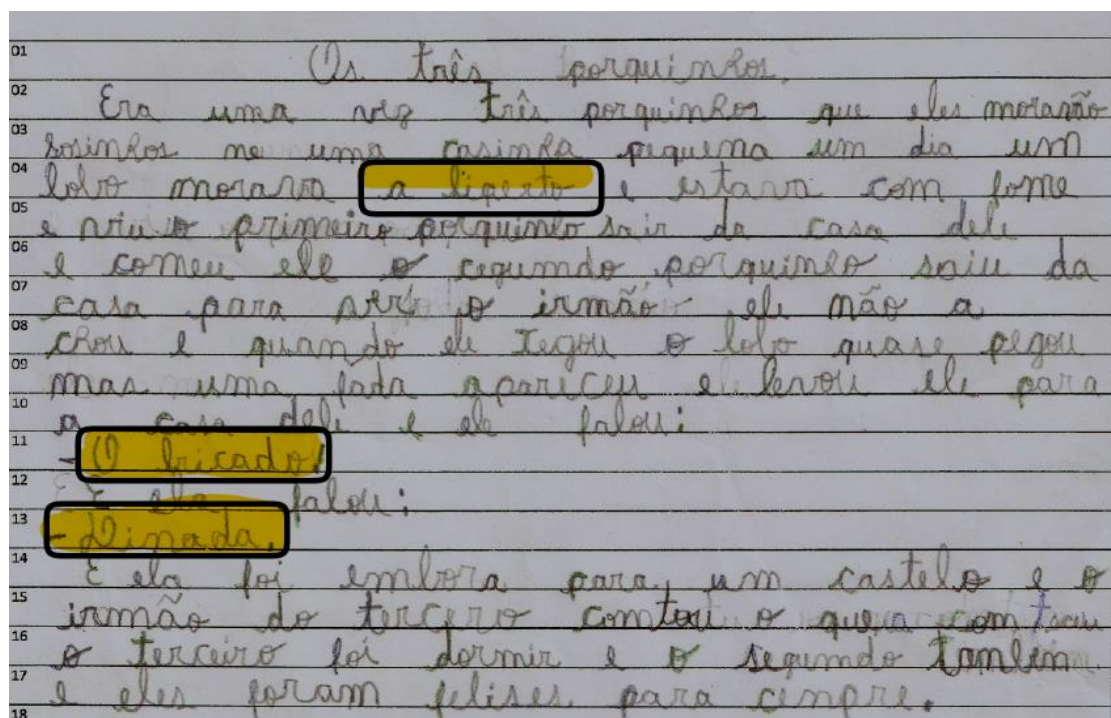
Com o fim de ilustrar os tipos acima descritos, uma vez que eles serão explorados em nossas análises, apresentamos, adiante exemplos desses tipos de estruturas que integram nossa amostra. Cabe comentar que, para a organização dos textos do banco, foi realizada uma codificação que levava em conta as variáveis que podem se prestar a análises (atuais ou futuras). Nesse sentido, o exemplo apresentado adiante, codificado como PB_PF_2F_141F corresponde ao seguinte: PB : texto do contexto de ensino

¹⁸ O branco constitui um signo e/ou sinal gráfico cujo funcionamento relaciona-se com a marcação dos limites e/ou fronteiras de palavras, frases, orações, sintagmas. (CAPRISTANO, 2003, p.23)

público; PF iniciais do nome da instituição escolar; 2F: turma 2º ano F; 141: número do texto dentre os 423 que compõem o banco; F: sexo feminino.

Dada essa explicação, apresentamos, na Figura 2, um texto que ilustra a ocorrência de segmentações não-convencionais de palavra em nossa amostra:

Figura 2. Ocorrências de segmentação não-convencional na amostra – PB_PF_2F_141F



Interpretação¹⁹

No texto acima, encontramos ocorrências dos três tipos de segmentação não-convencional que acabamos de apontar: (i) hipossegmentação, (ii) hipersegmentação e (iii) mescla. À linha 13, a estrutura “dinada” (de nada) corresponde a uma hipossegmentação, uma vez que o que a criança faz, nesse caso, é segmentar menos do que o esperado pela convenção ortográfica.

¹⁹ Era uma vez três porquinhos que eles moravam sozinhos em uma casinha pequena, um dia um lobo morava ali perto e estava com fome e viu o primeiro porquinho sair de casa dele e comeu ele/ o segundo porquinho saiu da casa para ver o irmão/ ele não achou e quando ele chegou o lobo quase cegou mas uma fada apareceu e levou ele para a casa dele e ele falou:

- Obrigado.

E ela falou:

- De nada.

E ela foi embora para um castelo e o irmão do terceiro contou o que aconteceu/ o terceiro foi dormir e o segundo também e eles foram felizes para sempre.

A hipersegmentação, por outro lado, aparece nas estruturas “o bricado” (obrigado), à linha 11. Nesse caso, há inserção de espaço em branco no interior da palavra ortográfica, do mesmo modo, destoante à convenção. À linha 04, a estrutura “a liperto” (ali perto) corresponde a uma mescla. Convém ressaltar que este último tipo de segmentação, denominado mescla, não integra as análises deste trabalho e o apresentamos apenas para exemplificar um outro tipo de dado de segmentação não-convencional que têm chamado a atenção de pesquisadores que investigam a escrita infantil.

Um olhar mais apurado sobre esses fatos será lançado em seção posterior, todavia, antecipamos a apresentação de nossos dados para ilustrar os tipos de segmentações. O que se pretende mostrar é que dados de escrita semelhantes a esses têm chamado a atenção de muitos pesquisadores das mais diversas áreas da Linguística – em especial, as da Fonética e da Fonologia.

Ao traçar um panorama²⁰ acerca dos estudos sobre as segmentações não-convencionais na escrita que se ancoram na Fonética e na Fonologia (tanto nos anos iniciais quanto nos finais do Ensino Fundamental), Silva (2014) assinala a existência de pelo menos três tendências de investigação, os quais resultam em estudos que descrevem e/ou analisam as segmentações sob enfoques: (1) fonético-semântico; (2) fonológico-letrado; e (3) fonológico-pedagógico. Em nossa revisão dos trabalhos desenvolvidos no país, que se voltam para as relações entre aspectos fonético-fonológicas da língua e segmentações não-convencionais, no campo dos estudos linguísticos, tomamos por base a proposta de Silva (2014). Passemos, então, a características de cada um desses três enfoques.

1.4.1 Enfoque fonético-semântico

A partir da análise de um *corpus* composto por centenas de textos de crianças de escolas públicas e privadas e ao abordar os critérios que levariam as crianças a segmentar, Abaurre (1988; 1989; 1991) apresenta critérios que, para a autora, seriam possíveis motivadores para a segmentação: (1) critérios ligados à percepção de grupos tonais, uma vez que os espaços em branco, por vezes, parecem querer registrar, na escrita, pausas da fala; (2) critérios ligados a princípios da organização interna do texto, já que as crianças tenderiam a segmentar mais nas passagens que introduziriam diálogos – isto é, no

²⁰ O trabalho de Silva (2014) não se reduz a esse panorama. Tal classificação é feita pela autora para contextualizar a análise que fará das hipersegmentações na escrita de estudantes do Ensino Fundamental II.

discurso direto; (3) critérios ligados à forma canônica da palavra em português, uma vez que, nesse sistema linguístico, as palavras podem receber acento fonológico na antepenúltima, penúltima ou última sílaba, sendo na antepenúltima o mais frequente, definindo assim o padrão paroxítono do português (; (4) critérios ligados a fatores socioeconômicos e possíveis graus de letramentos (ABAURRE, 1991, p. 11).

Em pesquisa realizada posteriormente, Abaurre (1996) acrescenta, a esses, um possível critério semântico, que, resultaria de a criança "[...] não querer abrir mão de algum significado conhecido a partes do enunciado." (ABAURRE, 1996, p. 8).

No que se refere ao primeiro critério, Abaurre (1989) destaca que a própria prosódia, cujos contornos são bastante salientes e perceptíveis para as crianças, é o que delimita, no âmbito dos enunciados linguísticos, as porções fônicas que tendem a corresponder a recortes possíveis da realidade. Essa observação é feita a partir de dados de hipossegmentação como “*tocuzimorino*” (estou quase morrendo) e *guenãovazia* (que não fazia), em que são observados limites de grupo tonal e de força.

No que se refere ao segundo critério, Abaurre (1991) destaca que, em textos narrativos, há uma tendência de as crianças segmentarem mais convencionalmente quando assumem o papel de narradoras em terceira pessoa do que quando empregam o discurso direto (momento em que demarcam a fala das personagens), situação em que tendem a segmentar menos do que a convenção prevê. Nesse último caso, o que se vê, segundo a autora, é um aumento do número de estruturas hipossegmentadas que podem ser justificadas devido à aproximação desse tipo de discurso com a característica contínua dos enunciados falados.

No que tange ao terceiro critério, as segmentações baseadas em formas canônicas do português foram explicadas pela autora a partir de questões rítmico-prosódicas que mostraram a influência do padrão rítmico preferencial do português brasileiro, a saber: os pés binários trocaicos. Nessa configuração, há alternância de um ritmo baseado em uma sílaba forte e uma fraca. Assim, segundo a autora, ao delimitarem as palavras, muitas vezes as crianças parecem combinar o padrão acentual preferencial da língua (paroxítono) com o número de sílabas por palavra. Esse tipo de critério foi exemplificado no trabalho a partir de estruturas como: “cofome” (com fome); “sero mano” (ser humano).

A respeito do quarto critério, o contexto em que se insere a criança é também, segundo Abaurre (1996), fator decisivo para a explicação da representação que a criança faria da escrita como contendo mais espaços em branco. Desse modo, um contato menor

com a escrita (fora do contexto escolar) faria com que a criança estivesse menos atenta a aspectos convencionais da segmentação da escrita do que uma criança que teria a oportunidade "[...] de contemplar esse objeto fora da escola, no contexto doméstico [...]" (ABAURRE, 1996, p. 140). É nesse sentido que a autora defende a existência de um critério de segmentação ligado à questão socioeconômica, uma vez que crianças com condições financeiras mais favoráveis tenderiam a ter contato com ler e manusear textos muito antes de irem para a escola, se comparadas às crianças de classes mais baixas.

Por fim, no que se refere ao quinto (e último) critério, Abaurre (1996) assume a existência de uma ação da escrita sobre as decisões infantis, não só nas propostas de segmentação, como ainda no uso das letras. Estruturas como "cala bolso" (calabouço) e "cata puta" (catapulta) funcionariam como indícios de que a criança pode estar operando com o reconhecimento de duas palavras distintas. Esse reconhecimento pode estar baseado não só na estrutura preferencial paroxítona, mas, segundo a autora, no interesse em tentar preservar o significado (re)conhecido das partes. O mesmo não ocorreria em "sero mano" (ser humano), em que a segmentação pode ser explicada apenas por meio da preferência por uma das formas canônicas das palavras no português, situação em que a criança segmentaria a sequência como se fosse constituída de duas palavras dissílabas em que o acento recai sobre a penúltima delas.

Silva (1991) buscou levantar, a partir de uma investigação que reuniu textos de crianças da 2ª série do antigo ensino fundamental, hipóteses sobre os aspectos da fala possivelmente envolvidos nos critérios de segmentação utilizados pelas crianças na elaboração de seus textos. Em seus resultados, o autor mostrou que a criança não se fundamenta apenas nos aspectos fonéticos da fala, mas ainda em características das convenções ortográficas e, também, em combinações entre aspectos da fala e das convenções, além de variações de outra natureza (como a oscilação na forma de grafar uma mesma palavra).

Em trabalho desenvolvido posteriormente, Abaurre; Silva (1993) buscaram discutir a relevância das segmentações não-convencionais como um meio de compreender como as crianças compreendem a relação linguagem/realidade por ela representada. A partir da hipótese de que crianças em aquisição da escrita estariam tentando identificar, no espaço de um mundo factual, aspectos que poderiam merecer o estatuto de entidades de existência autônoma (p. 1993), os autores destacam que as hiposegmentações seriam indícios de que as crianças não conseguiriam atribuir a alguns elementos da língua (como clíticos, conjunções, preposições, advérbios e formas

auxiliares do verbo) o estatuto autônomo; por isso, eles não mereceriam o recorte próprio da realidade.

As considerações de Abaurre (1988; 1989; 1991; 1996), de Abaurre; Silva (1993), Silva (1991, 1994) tiveram e têm grande expressividade no âmbito das reflexões sobre a segmentação não-convencional na escrita infantil, não apenas pelo seu pioneirismo, mas por suas abordagens distanciarem-se de hipóteses tradicionais que tomavam as segmentações como erro. Na perspectiva que assumem, ao olhar para as segmentações não-convencionais de palavra como resultados de hipóteses construídas pelas crianças em seu processo de aquisição e construção da escrita, esses autores mostraram que o que se caracterizava como erros são, na verdade, indícios do funcionamento linguístico da escrita infantil e de processos de subjetivação do escrevente. Por terem em sua base argumentos do aspecto *fonético* da prosódia e a relação desse aspecto com fatos da significação, tais estudos podem se enquadrar no enfoque a que Silva (2014) denomina de fonético-semântica. A seguir, apresentaremos características do segundo enfoque, que, à medida em que dialoga com aquele abordado na presente seção, volta-se, de modo particularizado, para o entrelaçamento de aspectos fonológicos das segmentações não-convencionais com aspectos oriundos da informação letrada.

1.4.2 Enfoque fonológico-letrado

Além de dialogar com estudos de pesquisadores destacados na descrição que acabamos de fazer, os estudos que poderíamos situar no segundo enfoque ancoram-se fortemente em postulações de Corrêa (2001)²¹.

Nesse sentido, Capristano (2003, 2007), Chacon (2004, 2005, 2006, 2008), Tenani (2004, 2009a, 2010, 2011b), dentre outros, têm se ancorado em reflexões de Corrêa (1997; 2001) ao relacionarem características fonológicas das segmentações não-convencionais a possíveis aspectos de práticas sociais de linguagem orais/faladas (sobretudo aqueles de natureza prosódica) e letradas/escritas (sobretudo aqueles prescritos nas convenções ortográficas). Em síntese, esses autores vinculam aspectos prosódicos das segmentações não-convencionais a aspectos das convenções ortográficas. Resultantes dessa escolha teórica, os trabalhos desses autores podem ser inseridos, nos termos de Silva (2014), no enfoque que a autora chama de fonológico-letrado.

²¹ A partir da concepção de heterogeneidade constitutiva da escrita.

A partir da análise de um conjunto de 45 textos de primeira série, Capristano (2003) detectou quatro tipos de funcionamento das segmentações não-convencionais, que resultariam: (1) de tentativas de escrita alfabética; (2) da oscilação entre diferentes trânsitos por constituintes prosódicos e informações sobre o código escrito institucionalizado; (3) da oscilação entre constituintes abaixo do domínio da palavra fonológica na hierarquia prosódica (sílabas e pé) e informações do código escrito institucionalizado; (4) de uma maior percepção de constituintes acima do domínio da palavra fonológica na hierarquia prosódica e, talvez em menor grau, de informações sobre o código escrito institucionalizado.

Dentre os aspectos mais significativos da proposta de Capristano (2003), cabe destaque à observação de que, como propôs Corrêa (1997), as ocorrências de segmentação não-convencional compreenderiam resultados do trânsito e/ou da co-ocorrência entre diferentes aspectos das práticas orais e letradas nas quais as crianças estavam inseridas. Dentre outros resultados, a autora sugere existirem pelo menos dois fatores que parecem atravessar as segmentações não-convencionais das crianças que analisou: (1) um mais ligado a aspectos prosódicos e vinculado ao que foi considerado, com base em Corrêa (1997), como representação pelos escreventes da escrita em sua suposta gênese; (2) outro mais ligado à imagem que as crianças teriam do que seria próprio da escrita e vinculado ao que Corrêa (2001) considerou ser uma representação baseada no código escrito institucionalizado.

Atendo-se a questões semelhantes e com vistas à relação entre fronteiras não-convencionais de palavras e fronteiras de constituintes prosódicos, Chacon (2004) analisa, nos mesmos 45 textos, marcas de segmentação que apresentam, segundo o autor, uma configuração que não atende nem aos princípios regulares como os de formação de constituintes prosódicos nem à convenção ortográfica. Nesse trabalho, o autor observa estruturas que resultam de oscilações entre constituintes abaixo do domínio da palavra fonológica na hierarquia prosódica e informações do código escrito institucionalizado. Chacon (2004) reforça a importância de se observarem dados como “acho rar” (a chorar) – ocorrência em que o espaço em branco foi colocado após o elemento que pode ter sido reconhecido como palavra – não somente como estruturas, mas como manifestações do modo de constituição heterogêneo da escrita, uma vez que, no caso desse dado em específico, a marca em branco pode ter sido “[...] fruto de uma ação retrospectiva do sujeito sobre sua escrita” (CHACON, 2004, p. 228).

Analisar estruturas dessa natureza sob esse enfoque seria um modo de

[...] mostrar não apenas a ação da (e a reflexão sobre) organização prosódica da língua, mas também a ação (e a reflexão sobre) as convenções ortográficas do português como resultantes do trânsito do escrevente por práticas de linguagem orais e letradas. (CHACON, 2004, p. 230).

Esse tipo de análise permitiria fugir à análise que vê essas estruturas como “[...] manifestações imperfeitas de um princípio mais geral da oralidade que atuasse, de algum modo, na escrita” (CHACON, 2004, p. 230).

O autor aponta, ainda, para a possibilidade de que, antes de tudo, haveria um funcionamento simultâneo (e não hierárquico) entre os constituintes prosódicos. Em seus dados, Chacon (2004) observa que, enquanto sequência, as estruturas não-convencionais evidenciam a ação de constituintes acima da palavra fonológica, ao passo que as porções demarcadas entre limites indiciam fronteiras de constituintes menores (abaixo da palavra fonológica: sílaba e pé). Esse funcionamento simultâneo é relacionado pelo autor a: (a) uma não-sincronização entre limites de constituintes e de prominências em constituintes; (b) uma ação mais forte do ritmo (mostrada nos dados pela ação do pé métrico) e menor de outros fatores prosódicos; (c) uma discrepância entre a idealização dos modelos e dados; (d) correspondências gráficas e sonoras relacionadas com palavras escritas da língua, pela ação do letramento sobre os sujeitos.

Segundo Chacon (2004), o que estaria motivando o interesse pelas relações entre convenções ortográficas e constituintes prosódicos é a demonstração de que

essas relações, em grande medida, indiciam a recuperação que os sujeitos escreventes fazem de suas vivências com enunciados falados (nos quais aspectos prosódicos contribuem enormemente para sua estruturação) e com enunciados escritos (nos quais, convencionalmente, unidades morfológicas como palavras têm realce, já que se separam umas das outras por meio de espaços em branco). (CHACON, 2004, p. 4)

Em trabalho posterior, Chacon (2006) retoma as pistas linguísticas que podem ser observadas nas segmentações não-convencionais e que marcam a presença das crianças em práticas de linguagem orais e letradas. As conclusões a que o autor chegou a respeito da aquisição da noção de palavra, nessa análise, retomam um fato já destacado em estudos anteriores: “padronização que a escrita codifica e com a qual o escrevente trabalha provém de práticas discursivas efetivas: em que se cruzam dois diferentes modos de enunciação verbal: o falado e o escrito” (CHACON, 2006, p. 165) e, portanto, a

aquisição da noção de palavra não apresenta apenas questões exclusivamente estruturais para a reflexão do escrevente.

Os fatos prosódicos e ortográficos que destacamos, na medida em que marcam no produto escrito a ação da língua sobre o sujeito e deste sobre a língua, remetem essa ação à inserção simultânea do sujeito em múltiplas práticas de oralidade e letramento (dentro de fora do contexto escolar), as quais tornam passível o acesso do sujeito escrevente não só à noção de palavra como também aos diferentes fatos de língua, tanto em seu modo de enunciação falado, quando em seu modo de enunciação escrito (CHACON, 2005, p. 165).

Na mesma direção, seguem os trabalhos de Tenani (2004, 2009a, 2010, 2011b). Interessada em dados de segmentação não-convencional produzidos por escreventes de séries do Ensino Fundamental II, Tenani (2004), ao tematizar as relações entre o sistema ortográfico e o sistema fonológico do português, a autora tem como fundamento as relações entre as práticas de oralidade e de letramento. Buscando investigar os limites de palavra, bem como as marcas que evidenciam as soluções buscadas pelos escreventes quando da delimitação das fronteiras, Tenani (2004) analisa as segmentações não-convencionais resultantes da oscilação entre diferentes trânsitos por constituintes prosódicos e informações sobre o código escrito institucionalizado.

Para Tenani (2004) as hipóteses dos escreventes acerca dos limites de palavra são ancoradas tanto em informações de natureza prosódica, cuja evidência relaciona-se à organização de elementos átonos (sílabas ou clíticos) em diferentes constituintes, quanto em informações de natureza letrada, a qual parece mostrar-se especialmente pela colocação do espaço em branco indicar possíveis palavras escritas.

Tenani (2011a) defende, em estudo posterior, que a existência de diferenças entre os dados de escrita do Ensino Fundamental I e do Ensino Fundamental-II, no que diz respeito à segmentação de palavras, tem seu fundamento em informações de natureza letrada. Um exemplo disso é que, quando comparados os dados desses contextos observa-se (por exemplo no dado [*ma telo*] do EF I) que nem sempre a sílaba pretônica coincide com algum clítico correspondente, por sua vez, com alguma classe gramatical, ao contrário do que ocorre nos dados do EF II, por exemplo “a migo” (amigo).

As discussões elencadas nesta seção se prestam a descrever a perspectiva fonológico-letrada. Os resultados dos trabalhos aqui citados serão retomados, todavia, em nossas análises, uma vez que nelas, percebemos tendências já destacadas nos estudos que se alinham a essa perspectiva.

Outra linha de análise é a que se constrói a partir das reflexões desenvolvidas do Grupo de Estudos sobre a Aquisição da Linguagem Escrita (GEALE – FaE/UFPeL), mais precisamente por meio das pesquisas de Cunha (2004, 2010), Cunha e Miranda (2009) Cunha e Ferreira (2011) os quais estabelecem, no estudo das segmentações não-convencionais, relações entre fatos fonológicos da prosódia e aspectos pedagógicos. Passemos à discussão dessa tendência.

1.4.3 *Enfoque fonológico-pedagógico*

Os trabalhos resenhados nesta seção, embora também analisem as segmentações não-convencionais com base na organização prosódica proposta por Nespor; Vogel (1986), diferem dos trabalhos apresentados à seção anterior, sobretudo por não comungarem a mesma concepção de escrita. Os trabalhos aqui expostos, diferentemente do que propõem os autores anteriores sobre o modo heterogêneo de constituição da escrita, orientam-se pelo viés do construtivismo piagetiano para a aquisição de conhecimento, tal qual defendido por Ferreiro; Teberosky (1999) e discutido em seção anterior (seção 1.1). Nessa tendência de estudo, destacam-se os trabalhos de Cunha (2004; 2010), Cunha; Miranda (2009) Cunha e Ferreira (2011) todos eles, como já mencionado, trabalhos resultantes de investigações desenvolvidas pelo Grupo de Estudos sobre a Aquisição da Linguagem Escrita (GEALE – FaE/UFPeL).

Em estudo semilongitudinal, Cunha (2004) compara textos produzidos por crianças de uma escola pública e de uma escola privada da cidade de Pelotas (RS). Analisando ocorrências de hiposegmentações, hipersegmentações e de estruturas que a autora chama de segmentações híbridas (misturas das duas anteriores) em textos de crianças com idades entre 6 e 11 anos que, à época, cursavam as quatro primeiras séries do ensino fundamental, a autora observou que tanto nas hiposegmentações quanto nas hipersegmentações o espaço em branco (seja por sua presença ou ausência) ocorre entre uma palavra gramatical e uma palavra fonológica (como em "tecomer" e "ter minar" – *te comer* e *terminar*). Segundo a autora, a hiposegmentação se justificaria pelo fato de que as crianças em início de alfabetização teriam dificuldades de reconhecer palavras formadas por uma ou duas letras como unidades autônomas da língua, sobretudo palavras desprovidas de acentuação fonológica. Ao contrário da hiposegmentação, na hipersegmentação, a criança, ao identificar a sílaba inicial de uma palavra como uma possível palavra gramatical, procura dar-lhe certa autonomia gráfica.

Outros resultados apontados por Cunha (2004) dizem respeito às variáveis contexto e série. Dentre as contribuições dessa autora, podem-se sintetizar os seguintes resultados: (a) as crianças da escola privada apresentam número consideravelmente menor de ocorrências de segmentações não-convencionais do que as crianças da escola pública, principalmente nas três primeiras séries; (b) há diferença entre os dados da escola pública e os da privada quanto ao tipo de segmentação não-convencional encontrada (em ambas as escolas existe maior incidência de dados de hipossegmentação do que de hipersegmentação, embora essa diferença seja muito mais aparente nos dados dos alunos da escola privada, havendo poucas ocorrências de hipersegmentações; (c) as crianças da escola privada, nos dois primeiros anos de estudo, apresentaram menos dúvidas quanto à segmentação de palavras do que as crianças da escola pública; d) a escrita seria um processo progressivo de aprendizagem, pois, quanto mais avançada a série, menor o número de ocorrências de hipo e hipersegmentações.

Por sua vez, com base em dados relativos à segmentação não-convencional de palavras extraídos de produções escritas de crianças em início de escolarização à luz da Fonologia Prosódica, Cunha; Miranda (2009) analisam hipossegmentações e hipersegmentações a partir de três variáveis linguísticas (*tipo de palavra, estrutura silábica, tonicidade*) e duas extralinguísticas: (*tipo de escola e série*). Em consonância com os dados obtidos por Cunha (2004), os resultados mostram que as hipossegmentações são preponderantemente influenciadas por constituintes mais altos da hierarquia prosódica (a palavra fonológica, a frase fonológica e a frase entonacional) ao passo que as hipersegmentações sofrem a influência dos constituintes mais baixos da hierarquia, tais como a sílaba e o pé métrico.

Tendo na aquisição da escrita um processo mais geral da aquisição da linguagem, as autoras defendem que: (a) as crianças atualizam conhecimentos já construídos sobre a gramática sonora de sua língua durante a aquisição da escrita; (b) o erro ortográfico²² é elemento chave para a descoberta de hipóteses das crianças sobre o sistema de escrita que estão a adquirir.

Em trabalho posterior, Cunha (2010) debruça sobre os dados de segmentação não-convencional produzidos por crianças brasileiras e portuguesas, de modo a descrever

²² Nos trabalhos do GEALE, a ideia de *erro* está respaldada na teoria piagetiana, para a qual o erro (seja ele de qualquer natureza) é parte integrante do processo de aprendizagem. Assim, o erro é relevante, na medida em que demonstra as hipóteses construídas pelos aprendizes na sua relação com o objeto a ser aprendido

os tipos de segmentação encontrados e a comparar os ritmos do português brasileiro e do português europeu.

Os resultados mostraram que, no caso das hipossegmentações, as grafias mais recorrentes envolviam a união de uma palavra fonológica e de uma palavra gramatical, como, por exemplo, “escondeuce” (escondeu-se). No que se refere às hipersegmentações, viu-se que a separação não-convencional, majoritariamente, se caracterizava por uma sequência de palavra gramatical e uma palavra fonológica, como, por exemplo, “da quela”²³. A análise das ocorrências evidenciou ainda que os domínios prosódicos reconhecidos foram: a) nas hipossegmentações: constituintes acima do nível da palavra fonológica na hierarquia; b) nas hipersegmentações: constituintes menores na hierarquia como o pé métrico e a sílaba.

No que tange à análise do ritmo, os resultados obtidos pela autora apontam evidências rítmicas entre o PB e PE (português europeu) tal qual a proposta de Abaurre e Galves (1998), para quem o ritmo do PB e do PE se distingue em relação à hierarquização de três restrições, a saber: integridade da palavra fonológica, binariedade do pé e pé trocaico.

Cunha (2010) assinala que a diferença entre o ritmo nas duas variedades do português está fundamentada na posição que o pé trocaico ocupa no interior da hierarquia - primeiro lugar no português europeu e último lugar no português brasileiro- pois em relação à integridade da palavra fonológica, em ambas as variedades, precede a posição da binariedade do pé. O ritmo em PE é baseado no pé troqueu, enquanto o ritmo brasileiro se constrói respeitando antes de mais nada as fronteiras de palavras fonológicas” (ABAURRE E GALVES, 1998, p.394 *apud* CUNHA, 2010, p. 74).

Ainda no rol dos trabalhos sob enfoque fonológico-pedagógico, encontramos a pesquisa de Ferreira (2011). A partir da análise de processos de segmentação não-convencional na escrita de sujeitos²⁴, no contexto da Educação de Jovens e Adultos (EJA), Ferreira (2011) constatou que, assim como os dados de produção escrita infantil, elencados em Cunha (2004), nos dados por ela analisados que: a) houve prevalência das grafias de hipossegmentação, em comparação aos dois outros tipos de grafias (hipersegmentação e híbridos); b) nas hipossegmentações havia ação de constituintes

²³ Essa análise corresponde à proposta por Cunha (2010). Convém destacar que “na” embora corresponda a uma palavra gramatical (preposição) prosodicamente ela está no domínio da sílaba. Portanto, a hipersegmentação “na quela” pode ser interpretada, também como uma construção composta por sílaba + palavra prosódica.

²⁴ Três sujeitos participaram da pesquisa, à época, eles tinham 44, 22 e 16 anos.

mais altos da hierarquia – frase fonológica, frase entonacional e enunciado fonológico – ao passo que, nas hipersegmentações os constituintes em ação eram os que estão abaixo da palavra – pé e sílaba); c) os elementos clíticos (palavra gramatical, nos termos de Cunha 2004) estão envolvidos tanto na hipossegmentações quanto nas hipersegmentações; d) processos fonológicos, tais como ditongação a degeminação ocorreram nos dados de hipossegmentação. e) a ocorrência de flutuações na forma de grafar uma mesma forma de palavra em um mesmo texto.

Os estudos destacados nesta seção, ainda que, sob propósitos distintos e, ainda, cuja preocupação mais estreita se dá com questões de ordem pedagógica, fornecem resultados importantes acerca da prosódia e da aquisição da escrita. Especialmente porque os resultados obtidos pelas autoras, em especial Cunha (2004), a respeito da influência do fator socioeconômico sobre os dados de escrita observados, traz fatos que, muito de perto, dialogam com o que buscamos em nossa pesquisa. Isso justifica, portanto, a inserção desses trabalhos no âmbito do arcabouço que dá respaldo a esta pesquisa.

Tendo em vista essas considerações, à seção que segue apresentamos, brevemente, os fundamentos que orientam a análise dos dados em constituintes prosódicos, conforme o modelo da Fonologia Prosódica proposto por Nespor e Vogel (1986).

1.5 Fonologia Prosódica e segmentações não convencionais: modo de análise

Conforme Scarpa (1999) as investigações sobre fatos prosódicos abarcam diferentes fenômenos, tais como: parâmetros de altura, intensidade, duração, pausa e velocidade de fala até o estudo dos sistemas de tom, entoação, acento e ritmo das línguas. Segundo a autora é possível identificar, portanto, dois polos de estudos: 1) fonéticos (se interessam pelo aspectos como altura, intensidade e quantidade, frequência, volume e duração, em termos individuais, por exemplo); 2) fonológicos: tem interesse pela organização e representação dos sistemas de ritmo, acento e entoação e a relação de interface desses com outros componentes gramaticais, como o morfológico, o sintático e o semântico.

Considerando que, ao escrever, a criança pode adotar critérios diversos para a segmentação, como propõem Abaurre (1991) e Silva (2004), e, partindo do pressuposto de que a instabilidade nos limites gráficos pode decorrer, dentre outros fatores, de

conflitos entre convenções ortográficas e constituintes prosódicos, selecionamos, para este trabalho, o modelo fonológico proposto por Nespor e Vogel (1986)²⁵.

Nesse modelo fonológico, no interior dos constituintes, são aplicadas regras fonológicas específicas, as quais são sempre orientadas por quatro princípios gerais:

Princípio 1. Um dado nó terminal da hierarquia prosódica, X^p , é composto de uma ou mais unidades da categoria imediatamente mais baixa, X^{p-1} ;

Princípio 2. Uma unidade de um dado nível da hierarquia é exaustivamente contida em uma unidade hierarquicamente superior da qual ela faz parte;

Princípio 3. As estruturas hierárquicas da fonologia prosódica são ramificações n-árias;

Princípio 4. A relação de proeminência relativa definida por nós irmãos é tal que a um só nó é atribuído o valor forte (s) e a todos os outros nós o valor fraco (w) (NESPOR & VOGEL, 1986, p.7)²⁶.

Nesse modelo, a estrutura interna de cada constituinte prosódico segue sempre a mesma regra de construção, denominada regra de *construção do constituinte prosódico*²⁷: “junte em ramificações n-árias X^p todos os X^{p-1} incluídos em uma cadeia delimitada pelo domínio de X^p .” (NESPOR; VOGEL, 1986, p.7)²⁸. A noção de constituinte, nesse sentido, está relacionada a uma unidade linguística formada de um ou mais membros que estabelecem entre si uma relação de dominância.

Partindo desse pressuposto, a representação mental da fala está organizada de maneira hierárquica, de modo que as sentenças de uma língua são segmentadas pelo componente fonológico em constituintes prosódicos. A hierarquia prosódica proposta por Nespor e Vogel (1986), é composta sete constituintes prosódicos, quais sejam: *enunciado fonológico* (U), *frase entonacional* (I), *frase fonológica* (φ), *grupo clítico*, (C) *palavra fonológica* (ω), *pé métrico* (Σ) e *sílaba* (σ).

²⁵ Alinhamo-nos à filiação partilhada por outros pesquisadores do GPEL para o estudo das segmentações não-convencionais, tais como Capristano (2003, 2004) Chacon (2004, 2005), dentre outros.

²⁶ “Principle 1. A given nonterminal unit of the prosodic hierarchy, X^p , is composed of one or more units of the immediately lower category, X^{p-1} . Principle 2. A unit of a given level of the hierarchy is exhaustively contained in the superordinate unit of which it is a part. Principle 3. The hierarchical structures of prosodic phonology are n-ary branching. Principle 4. The relative prominence relation defined for sister nodes is such that one node is assigned the value strong (s) and all the other nodes are assigned the value weak (w).” (NESPOR & VOGEL, 1986, p.7).

²⁷ Prosodic Constituent Construction (NESPOR & VOGEL, 1986, p. 7)

²⁸ “Join into na n-ary branching X^p all X^{p-1} included in a string delimited by the definition of the domain of X^p .” (NESPOR & VOGEL, 1986, p. 7).

Denominado *relation based*, nesse modelo fonológico as línguas têm uma estrutura prosódica que é formada por constituintes que têm como referência a estrutura morfossintática (porém não é isomórfica a ela).

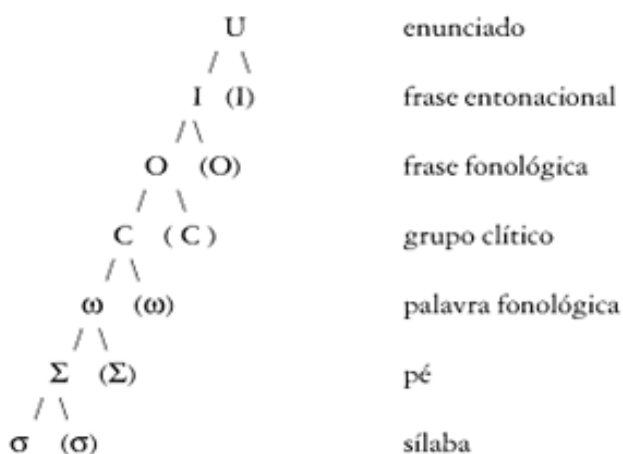
Para tratar da noção de relação à luz da Fonologia Prosódica, trazemos para o diálogo as explicações de Tenani (2017, p. 588) segundo a qual nesse as “relações sintáticas – ao lado de informações morfológicas e semânticas – são informações relevantes a partir das quais se configuram os constituintes prosódicos”. Nesse sentido, segundo a autora:

os constituintes prosódicos não são necessariamente isomórficos aos demais constituintes gramaticais. Por exemplo, um sintagma nominal (SN) constituído de substantivo e adjetivo constitui dois sintagmas fonológicos (ϕ) quando a *relação* for substantivo-adjetivo, como em: [livros bons]SN = [livros] ϕ [bons] ϕ . Porém, se a *relação* for adjetivo-substantivo, o sintagma nominal será mapeado como um sintagma fonológico, como em: [bons livros]SN = [bons livros] ϕ . Ou seja, um constituinte prosódico é criado a partir de informação de constituinte sintático, mas não há necessariamente isomorfia entre esses constituintes. (TENANI, 2017, P. 588).

Sob essa perspectiva, mesmo que em alguns casos os constituintes prosódicos venham a coincidir com outros constituintes linguísticos, essa coincidência não é obrigatória, pois a relação entre os constituintes prosódicos e os demais componentes da gramática não é do tipo isomórfica, tal como destacam Nepor e Vogel (1986).

Diante do exposto, apresentamos a seguir a hierarquia de constituintes prosódicos conforme, conforme Nespor e Vogel (1986):

Figura 3. Hierarquia de constituintes prosódicos



Para ilustrar esse tipo de análise, tomamos o exemplo extraído de Capristano (2007).

A casa de Ana fica em Marília

Tomando como modelo analítico a hierarquia proposta por Nespor e Vogel (1986) essa sentença poderia ser explicada, em termos de constituintes prosódicos, da seguinte maneira:

a) A **sílabas**: as sílabas /ka/ ou /za/ da palavra /kaza/ - cada uma possui um nó com valor forte, que em português é sempre uma vogal;

b) O **pé métrico**: a palavra /kaza/ é composta por duas sílabas que estabelecem entre si uma relação de dominância, em que uma se caracteriza por ser forte e outra fraca. Nesse caso /ka/ é dominante, caracterizando-se um pé troqueu – quando a sílaba mais forte está à esquerda. No caso da sílaba mais forte estar à direita, tem-se um pé iambo, como no caso da palavra “café”.

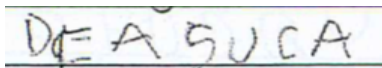
c) A **palavra prosódica**: não é isomórfica em relação à palavra morfológica, caracteriza-se pela presença do acento primário. No caso da sentença supracitada, segundo Nespor e Vogel (1986), /kaza/ é uma palavra prosódica. Fato que comprova a não isomorfia entre componentes fonológico e morfológico seria o caso da palavra “guarda-roupa”, em que se tem apenas uma palavra morfológica, mas duas palavras fonológicas, uma vez que há dois acentos na palavra composta.

d) O **grupo clítico**: é formado por uma única palavra de conteúdo acompanhada de clíticos (palavras átonas, como artigos, preposições, conjunções). No caso da sentença mencionada, tem-se quatro grupos clíticos: / A casa / da Ana/ fica / em Marília.

e) A **frase fonológica**: congrega um ou mais clíticos. Tem-se, na sentença, as seguintes frases fonológicas: [A casa da Ana] [fica em Marília].

g) A **frase entonacional**: é um domínio de contorno de entoação. Nesse caso tem-se uma (I): [A casa de Ana fica em Marília]

h) O **enunciado fonológico**: constituinte mais alto da hierarquia prosódica; indica o começo e o fim de um constituinte sintático. O (U) pode ser composto de uma ou mais frases entonacionais. No caso, é composto por apenas uma (I).

A partir desse modelo é possível a observação das possibilidades de organização prosódica das grafias de segmentação não-convencional e compreender as (possíveis) motivações para elas. Por exemplo o caso da grafia  (“de açúcar”) em que o clítico “DE” (que corresponde a uma preposição) é incorporado à palavra fonológica “ASUCA”, sendo interpretado pela criança como parte da palavra. Nesse caso dizemos que a segmentação não-convencional “DEASUCA” é constituída por grupo clítico (C).

Observada a disposição dos constituintes na hierarquia, cabe, neste momento a discussão sobre cada um deles, apresentando exemplos desses constituintes em nossa amostra.

- **Enunciado fonológico (U):**

Sendo este o mais alto constituinte da hierarquia, ele é formado por uma ou mais frases entonacionais (I) e é delimitado pelo começo e pelo fim de um constituinte sintático e pela proeminência relativa, que atribui valor forte ao nó mais à direita. Dominado pelo nó mais alto da cadeia sintática, o U aceita reestruturações e é marcado pela interação (não isomórfica) entre informações fonológicas, sintáticas e semânticas.

Para tanto, devem ser respeitadas algumas condições de ordem pragmática e fonológica:

Condições pragmáticas: (i) os dois enunciados devem ser produzidos pela mesma pessoa; e (ii) devem ser dirigidos ao(s) mesmo(s) interlocutor(res).

Condições fonológicas: (i) os dois enunciados devem ser relativamente curtos; e (ii) não deve haver pausas entre eles.

No exemplo abaixo temos uma hipossegmentação, extraída de nossa amostra, a partir da qual podemos observar uma estrutura formada a partir da junção de um enunciado fonológico (U)

Figura 4. Exemplo de Enunciado Fonológico



Texto: PV_BM_1B_70F

Interpretação: Eu acho que é do lobo porque ele falou a verdade

Análise prosódica: “[Eu acho] C] ϕ] [que é] C] ϕ] [do lobo] ϕ] C] I [[porque] ϕ [ele] C [falou] C] ϕ] [a verdade] C] ϕ] I]] U

A sentença acima ilustra uma estrutura de hipossegmentação que se dá pela junção de um enunciado fonológico. Na análise prosódica acima apresentada, observamos, hierarquicamente, que o enunciado fonológico [Eu acho que é do lobo porque ele falou a verdade] agrupa: duas frases entonacionais ([eu acho que é do lobo] I [porque ele falou a verdade] I); cada uma das frases entonacionais agrupa, por sua vez, três frases fonológicas ([eu acho] ϕ [que é] ϕ [do lobo] ϕ [porque] ϕ [ele falou] ϕ [a verdade] ϕ]; podem ser observados no exemplo, ainda, seis estruturas constituídas por clíticos, os quais se agrupam no interior das frases fonológicas.

- **Frase entonacional (I):**

A **I** é o constituinte imediatamente abaixo de U e é formada por uma ou mais frases fonológicas (ϕ). Segundo Nespor e Vogel (1986), uma frase entonacional se configura no nível de uma sentença raiz/estrutura oracional²⁹ e apresenta contorno entonacional reconhecível. Assim como U, em I também se verifica interação não isomórfica entre informações fonológicas, sintáticas e semânticas.

A possibilidade de reestruturação de I e a proeminência relativa desse domínio fazem com que ele tenha uma variabilidade estrutural característica.

A regra básica de formação de **I**, para Nespor e Vogel (1986, p. 188), é a de que ela é domínio de contorno de entoação e que os fins de frases entonacionais coincidem com posições em que pausas podem ser inseridas. No exemplo a seguir, percebemos que a hipossegmentação produzida pelo escrevente se dá pela ausência de espaço em branco entre as palavras “era”, “uma” e “vez” motivada pela percepção da frase entonacional.

²⁹Com base na teoria X-barras, Nespor; Vogel (1986) afirmam que as “[...] frases coordenadas são tipicamente sentenças raízes, enquanto frases subordinadas não são” (NESPOR; VOGEL, 1986, p.21). “[...] coordinated sentences are typically root sentences, while subordinated sentences are not.” (NESPOR; VOGEL, 1986, p.21).

Figura 5. Exemplo de Frase Entonacional

Texto: PV_CG_1A_26M

Interpretação: “Era uma vez”

Análise prosódica: [Era] C ϕ]C] [uma] C [vez] C] ϕ] I

- **Frase fonológica (ϕ):**

A ϕ é o domínio prosódico formado por um ou mais grupos clíticos (C). Nos termos de Nespor e Vogel (1986, p. 168) “O domínio de ϕ consiste de um grupo clítico que contém uma cabeça lexical (X) e todos os grupos clíticos em seu lado não recursivo”. O modelo de sintaxe adotado pelas autoras considera como cabeças lexicais (X) os verbos, os nomes e os adjetivos; já os nós irmãos das cabeças lexicais são considerados como complementos de X ou como especificadores de X³⁰.

Abaixo vemos um fragmento de um texto em que é possível observar duas hipossegmentações, ambas correspondem, cada uma, a uma ϕ :

Figura 6. Exemplo de Frase Fonológica

Texto: PB_CG_1C_27M

Tradução: “O lobo foi embora”

[[O lobo] ϕ foi embora] ϕ] I

³⁰ De acordo com Nespor; Vogel (1986) “Uma afirmação fundamental da teoria X-barra é que existe um paralelismo na estrutura interna das frases de uma língua. Assim, as categorias lexicais N, V, A [...] não precisam ser mencionadas nas regras de estrutura de frases, já que elas representam os valores que a variável X pode assumir.” (NESPOR; VOGEL, 1986, p. 20)

“O formato básico de uma regra de estrutura de frase é: $X_n \rightarrow \dots X_{n-1} \dots$, em que o valor máximo que n pode ter é uma questão empírica. Em princípio, também é possível que esse valor varie de língua para língua ou de categoria para categoria dentro de uma língua, ao longo de algum parâmetro. Para os fins deste livro, no entanto, vamos assumir que o valor máximo de n é 2. Cada frase terá, portanto, três níveis distintos: X, X1, e X2. X é a cabeça; os nós irmãos de X em X(1) são conhecidos como os complementos de X, e os nós irmãos de X(1) em X(2) são conhecidos como os especificadores de X.” (NESPOR; VOGEL, 1986, p. 20).

Nele verificamos uma frase entonacional composta por duas frases fonológicas [o lobo] ϕ [foi embora] ϕ . Nesse caso, podemos dizer que as segmentações não-convencionais “ulobo”, “voibora” compreendem a duas frases fonológicas, isso porque a presença e a ausência dos espaços em branco, marcam, justamente, as fronteiras do domínio de ϕ . No caso “ulobo” vemos, ainda, que a ϕ é uma estrutura formada por clítico + palavra de conteúdo, configurando uma estrutura que está, também, no domínio prosódico do grupo clítico (C), do qual falaremos adiante.

- **O grupo clítico (C):**

Conforme Nespor e Vogel (1986), o grupo clítico é composto por uma única palavra de conteúdo (não clítica) e todos os elementos clíticos³¹ (artigos, preposições, conjunções, pronomes pessoais adjacentes a ela). Sendo ele o domínio imediatamente abaixo da frase fonológica, é definido como o primeiro constituinte a apresentar interação entre informações fonológicas e morfológicas. Tal interação, entretanto, tal como já destacado nos demais domínios, não é isomórfica.

Ao assumirem a existência do constituinte grupo clítico (locução formada pela combinação clítico + palavra ou palavra + clítico) formando uma hierarquia de sete constituintes, as autoras assumem que o clítico, mesmo não portando acento, pode ser do mesmo tamanho que uma palavra prosódica. Na aceção dessas autoras, a ω se caracteriza por apresentar um acento primário. E, se, de acordo com a *Strict Layer Hypothesis* (SLH), todo XP consiste de XP-1, é preciso que o clítico figure como ω antes de fazer parte do grupo clítico.

Essa posição das autoras - de que os clíticos fonológicos são prosodizados em um constituinte específico, o grupo clítico (C)³² - entretanto, não é consensual³³, uma vez que, na hierarquia prosódica proposta por Selkirk (1984), por outro lado, o domínio grupo

³¹ Os clíticos são palavras funcionais que não pertencem a uma classe morfológica específica. Não são portadores de acento, o que os obriga a apoiarem-se no acento de palavras adjacentes e os impede de ocorrerem isoladamente.

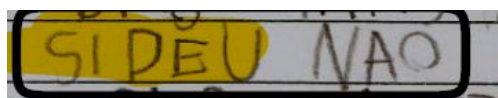
³² Esse posicionamento de que é necessário, na hierarquia prosódica, um domínio prosódico localizado acima da ω e abaixo da ϕ ω é defendido por Vigário (2007) em estudos sobre a prosódia do PE e por Bisol (2000, 2004), Toneli (2014, 2017) em estudos sobre os clíticos fonológicos no PB.

³³ Silva (2018) explica que a falta de consenso na definição do estatuto prosódico do clítico se dá pela sua própria natureza: sintaticamente eles podem representar nós terminais e encabeçar sintagmas (como algumas preposições) ao passo que prosodicamente eles só se integram à palavra se acompanhados de estruturas proeminentes.

clítico não existe. O clítico, segundo essa perspectiva, não pode ser considerado palavra fonológica, pois viola um princípio da Strict Layer Hypothesis (SLH): Princípio 2: uma unidade de um determinado nível da hierarquia está exhaustivamente contida na unidade hierarquicamente superior de que faz parte. Nesse sentido, Selkirk (1984) assume que os clíticos são sempre prosodizados junto à ω ou a ϕ .

Nos exemplos abaixo, ilustramos a ocorrência do grupo clítico nas segmentações não-convencionais de nossa amostra:

Figura 7. Exemplo de Grupo Clítico



Texto - PB_AR_1D_22M

Interpretação: “Se deu mal”

[[Se deu]C [mal] C] ϕ

A hipossegmentação ‘sideu’ se dá pela percepção de um C, formado pelo elemento clítico “si” e uma palavra de conteúdo “deu”. Um outro exemplo desse tipo de ocorrência é a construção “aminha” em que consideramos “a” um clítico (artigo definido) e “minha” uma palavra de conteúdo.

Figura 8. Exemplo de Grupo Clítico



Texto- PB_AC_1A_08M

Interpretação: “A minha”

[A minha] C

No que tange ao C formado por preposição + palavra, podemos citar, como exemplo a construção “nacara”.

Figura 9. Exemplo de Grupo Clítico

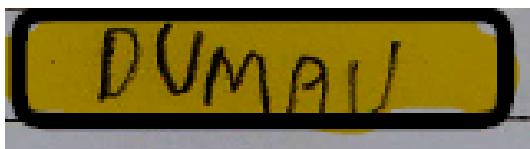
Texto - PB_AC_1F_11M

Interpretação: “Na cara”

[Na cara] C

Nessa situação, vemos que a hipossegmentação “nacara” é composta pelo clítico “na” (que resulta da contração entre preposição EM + artigo A) e a palavra de conteúdo “cara” formando um C (grupo clítico).

Ainda no que se refere à classe dos clíticos preposicionais o exemplo que segue ilustra a formação de um grupo clítico C cuja palavra de conteúdo é o monossílabo “mal” – grafado “mau”- à qual o clítico “du” (contração entre a preposição -de e do artigo -o –) se a um advérbio [mau] se une, levando à construção de uma estrutura hipossegmentada.

Figura 10. Exemplo de Grupo Clítico

Texto: PB_AC_1A_08M

Interpretação: “Do mal”

[Do mal] C]

No que se refere à proeminência relativa, Nespor e Vogel (1986) afirmam que, nesse domínio, o elemento considerado forte é a palavra de conteúdo, ou seja, o elemento considerado não clítico que compõe o grupo clítico. Nesse sentido, nessa sentença “mal” é o elemento forte do domínio, enquanto “do” é o elemento fraco.

As exposições até o momento realizadas acerca do domínio C, sobretudo as exemplificações, se dão em função da falta de consenso sobre a natureza e sobre o status desse constituinte³⁴, no âmbito da Fonologia Prosódica, daí o interesse de explicitarmos o ponto de vista do qual comungamos para a análise dos clíticos.

³⁴ Problematizadas em outros estudos, por exemplo, Vigário (2006), Bisol (2000 , 2004), Silva (2018).

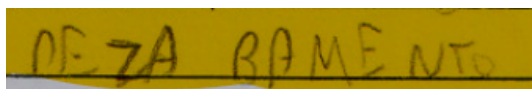
- **A palavra prosódica (ω):**

Imediatamente abaixo do grupo clítico, a ω é o mais baixo constituinte da hierarquia prosódica que apresenta interação com outros componentes da gramática. Nesse nível da hierarquia, as informações fonológicas interagem com o componente morfológico, embora tal interação, assim como ocorre nos demais constituintes, não necessariamente apresente isomorfia.

Para Nespor e Vogel (1986), não existem palavras prosódicas maiores que o elemento terminal de uma árvore sintática³⁵. Para ilustrar a incompatibilidade entre as divisões fonológicas e morfológicas, podemos citar, por exemplo, a palavra *salva-vidas*, que, do ponto de vista morfológico, é uma única palavra, pois corresponde ao elemento terminal de uma árvore sintática, além de ser um nome com significado; por outro lado, do ponto de vista fonológico, compreende duas palavras, visto que tanto *salva* quanto *vidas* são portadoras de acento primário.

Ainda para as autoras, retomamos: a ω apresenta duas características essenciais: domínio de restrições fonológicas e domínio de aplicação de regras. A *Strict Layer Hypothesis (SLH)* exige que uma unidade de um dado nível da escala prosódica esteja exaustivamente contida no nível superordenado do qual ela faz parte. Dessa forma, cada constituinte prosódico é a unidade composta de uma ou mais unidades da categoria imediatamente inferior.

Figura 11. Exemplo de palavra prosódica



Texto - PB_PF_5A_377M

Interpretação: “desabamento”

[desa] Σ [bamento] Σ] ω

³⁵ Ou seja, segundo as autoras, apenas as seguintes possibilidades existem: domínio de (ω) menor ou igual ao elemento terminal.

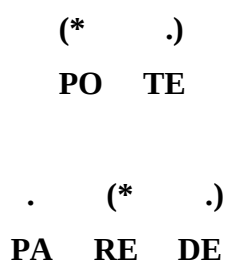
O exemplo explicitado acima mostra uma ocorrência de hipersegmentação que demonstra a separação de uma palavra prosódica (desabamento) em duas partes: “desa” e “bamento”. Uma vez que ω é a categoria que imediatamente domina o pé métrico (Σ), observamos que “desa bamento” é uma constituída por dois pé binários com proeminência à esquerda (pés troqueus) “desa” e “bamento” em que “de” e “men” correspondem às sílabas mais fortes.

Uma vez que o pé métrico³⁶, sendo o elemento determinante do acento, é elemento fundamental na ω , passamos à classificação desse domínio, a partir de exemplos de nossa amostra.

- **Pé métrico (Σ):**

O domínio Σ é o constituinte composto por uma ou mais sílabas, sendo uma delas acentuada, chamada de cabeça. É a partir da constituição dos pés que se dá a acentuação nas línguas, ou seja, a partir da noção de proeminência relativa, pois o pé se caracteriza pela relação entre duas sílabas em que uma delas será portadora de valor forte (acento) e a outra de valor fraco. O pé métrico é definido exclusivamente por informações fonológicas. Bisol (1996) afirma que, no caso do PB, a maioria das palavras tem formação de pés binários com proeminência à esquerda (denominados pés troqueus), como, por exemplo, em “pote”, “parede”, “catavento”.

Como vemos no esquema abaixo, há entre as sílabas uma relação de proeminência, em que o valor forte é atribuído à sílaba mais à esquerda (o asterisco representa uma sílaba acentuada ao passo que o ponto representa uma sílaba não acentuada):



³⁶ O pé métrico se define pelo nível imediatamente inferior, que é a sílaba, Ambos, sílaba e pé métrico, são elementos constitutivos palavra fonológica (BISOL, 2004).

(* .) (* .)
CA TA VEN TO

Além dos pés binários com cabeça à esquerda, há também os pés cujo acento recai sobre a sílaba mais à direita (denominados pés iambos), por exemplo [*café*], [*sofá*].

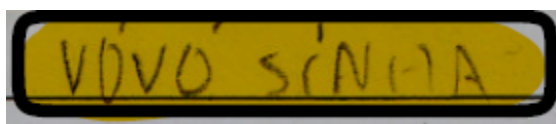
(. *)
CA FÉ

(. *)
SO FÁ

A formação de pé *espondeu* (formado por duas sílabas longas: uma pelo acento e outra em razão da coda silábica) é menos recorrente no PB, por exemplo [órgão]. Por fim, existem ainda, na composição das palavras do PB, os denominados pés ternários *dátilo*, como em “árvore” e *anapesto* como em “jacaré”. O primeiro integra a família do pé troqueu, uma vez que conserva proeminência à esquerda, ao passo que o segundo compõe o ritmo iambo (com a tonicidade à direita).

Para ilustrar o domínio do pé métrico selecionamos o caso de uma hiperssegmentação constituída por dois pés métricos “vovó” e “zinha”. Convém comentar que, uma vez considerados os princípios da hierarquia prosódica de Nespor e Vogel (1986), ambas as partes, “vóvó” e “zinha”, se constituem, também, como palavras prosódicas, visto que o constituinte ω domina imediatamente o constituinte inferior da hierarquia, que, no caso, é o Σ .

Figura 12. Exemplo de pé métrico



Texto: PB_MN_2A_123F

Interpretação: “Vovozinha”

[vovo] Σ [zinha] Σ ω

Nesse caso, vemos que a hiperssegmentação “vóvó sínha” é constituída de dois pés binários, nos quais o escrevente parece, por um lado, perceber um padrão rítmico do português: o pé troqueu. Essa interpretação é possível quando observamos que quando

escreve “sínha” a criança marca a sílaba forte por meio de um acento, demonstrando reconhecer a saliência prosódica. O mesmo ocorre na grafia “vóvó”. Vejamos que, embora a acentuação da primeira sílaba de “vovó” não seja convencional, a criança registra, em “vóvó” o mesmo padrão troqueu.

Nos exemplos que seguem, por outro lado, vemos dois casos distintos em que as grafias demonstram domínio do Σ :

Figura 13. Exemplo de pé iambo



Texto- PB_MN_2B_138F

Interpretação: “Chaminé”

Análise prosódica: [cha] [miné] Σ

Figura 14. Exemplo de pé troqueu



Texto- PB_MN_2B_112F

Interpretação: “Tijolo”

Análise prosódica: [ti] [jolo] Σ

Em “cha miné” há uma segmentação não convencional cujo espaço em branco separa uma sílaba e um pé métrico (iambo), em que a cabeça está à direita “miné”. No caso de “ti solo” em contrapartida, a grafia marca a separação de uma sílaba e um pé troqueu “solo”. Dado o nosso interesse em investigar a configuração prosódica das hipersegmentações e das hipossegmentações, a análise aqui empreendida considera a dimensão prosódica das segmentações não-convencionais, entretanto, como já exposto em capítulo anterior, essas grafias podem revelar - mais do aspectos do funcionamento linguístico - também aspectos relacionados às práticas sociais letradas pelas quais os escreventes transitaram. Nesse sentido, os dados dos exemplos 21 e 22 (“cha miné” e “ti solo”) podem ser observados a partir dessa outra perspectiva³⁷, segundo a qual “cha” e

³⁷ Capristano (2007), Chacon (2004), ancorados em Corrêa (2001) olham para dados segmentação na escrita infantil a partir dessa perspectiva.

“ti” correspondem, também, a palavras ortográficas, o que fornece evidências do reconhecimento, por parte do aluno, do código escrito institucionalizado.

- **Sílaba (σ):**

Na hierarquia prosódica, a sílaba é o menor constituinte. Apesar de ser consensual a relevância desse domínio na hierarquia prosódica, não há unanimidade quanto à sua estrutura interna³⁸. Para Nespor e Vogel (1986), embora a sílaba tenha seus constituintes internos, eles não têm papel na hierarquia.

Para o mapeamento da σ , a regra de proeminência relativa é determinante: assim como os demais constituintes prosódicos, a sílaba apresenta um cabeça (que se caracteriza por um nó com valor forte) e seus dominados (nós com valor fraco). No caso do PB, o nó com valor forte corresponde sempre às vogais, ao passo que os nós com valor fraco correspondem, preferencialmente, a consoantes e, eventualmente, a semivogais. No modelo de Nespor e Vogel (1986), as sílabas são agrupadas em pés.

O exemplos abaixo ilustram casos de hipersegmentação em que o domínio prosódico da σ pode ser reconhecido [$\sigma + \sigma$] :

Figura 15. Exemplo de sílaba



Texto - PB_AC_2B_95M

Interpretação: “Era”

[é] σ [ra] σ

Figura 16. Exemplo de sílaba



Texto - PB_AC_2B_95M

Interpretação: “Essa”

³⁸ Para uma discussão mais aprofundada em torno da estrutura da sílaba em PB, ver Collischonn (2001).

[e] σ [sa] σ

Tendo em vista as explanações feitas ao longo deste capítulo, a partir das quais foi possível explicitar como se dá a análise das segmentações não-convencionais nos termos de sua configuração prosódica, passemos à Seção 2, em que descreveremos a metodologia empreendida neste trabalho.

SEÇÃO 2. METODOLOGIA

O propósito desta seção é o de caracterizar as escolhas teórico-metodológicas referentes à amostra investigada e os critérios de análise tanto para a análise qualitativa, quanto para a estatística.

Na primeira seção, denominada “Formação do banco de dados” delineamos a constituição do *EscIn*, banco de dados do Grupo de Pesquisa *Estudos sobre a linguagem* (GPEL/CNPq), coordenado pelo professor doutor Lourenço Chacon/UNESP. À seção posterior, intitulada “Amostra analisada” abordaremos a seleção da amostra pesquisada, fornecendo as características do corpus investigado.

Na seção “Critérios de análise” apresentamos as decisões metodológicas assumidas para a análise dos dados de segmentação não-convencional levantados no corpus. Posteriormente na seção “Análise estatística”, apresentamos os procedimentos estatísticos empregados em nossa análise.

2.1 Sobre o banco de dados e sua constituição

O *corpus* desta investigação integra o banco de dados *Escrita Infantil (EscIn)*, que se caracteriza por um conjunto de textos coletados por pesquisadores do Grupo de Pesquisa *Estudos sobre a linguagem* (GPEL/CNPq), coordenado pelo pesquisador Dr. Lourenço Chacon (UNESP/Marília-SP/São José do Rio Preto-SP) na cidade de Marília – SP, no ano de 2012.

O *EscIn* é composto por 4.230 textos resultantes da aplicação de uma proposta de atividade de escrita produzida em situação escolar. A proposta consistia em o escrevente recontar, a seu modo, a “verdadeira história dos três porquinhos”, baseando-se no conto “A verdadeira história dos três porquinhos” de Jon Scieszka. Os textos que compõem o *EscIn* são de crianças do 1º ao 5º ano do Ensino Fundamental I, de cinco escolas privadas e dez escolas públicas do Município de Marília-São Paulo, perfazendo um total de 3.020 textos de alunos de ensino público e 1.210 textos de alunos da rede privada.

Para compreendermos um pouco mais sobre as condições que circundam a constituição dos dados do *EscIn*, se faz pertinente apresentarmos aspectos relacionados ao município de Marília (SP), onde o banco foi constituído.

Segundo dados obtidos pelo IBGE (Instituto Brasileiro de Geografia e Estatística), através do último censo demográfico (realizado em 2010)³⁹, Marília- SP mostra-se da seguinte maneira:

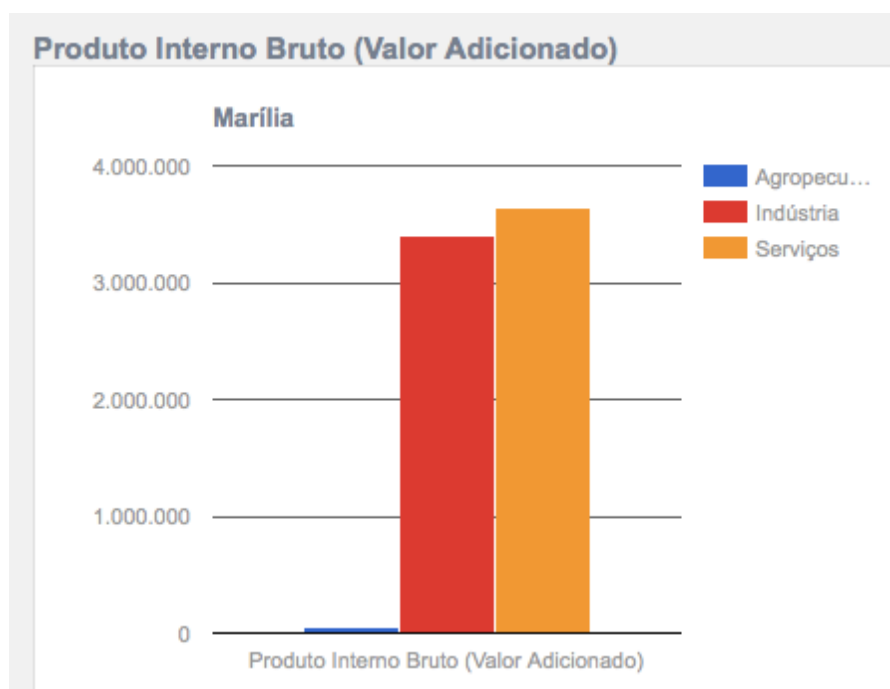
Tabela 1. Sobre o município de Marília - Dados do Censo Demográfico (2010)

População total	216.745
Urbana	207.727
Rural	8.974
Mulheres	112.019
Homens	104.726
Taxa de alfabetização	194.604
População em creches e escolas	60.028
Densidade demográfica (hab./km ²)	185,21
IDH- (M)	0,821
IDH (R)	0,885
IDH (longevidade)	0,820

No que diz respeito ao perfil econômico da cidade, tem-se que o PIB se concentra no ramo de serviços e indústrias, sobretudo alimentícias, conforme vemos no gráfico que segue:

³⁹ Disponível em: <http://www.cidades.ibge.gov.br/sao-paulo/marilia/sintese-das-informacoes-> Acesso em 16/05/2019 às 11:44hrs.

Gráfico 1. Perfil econômico do Município de Marília, segundo o IBGE

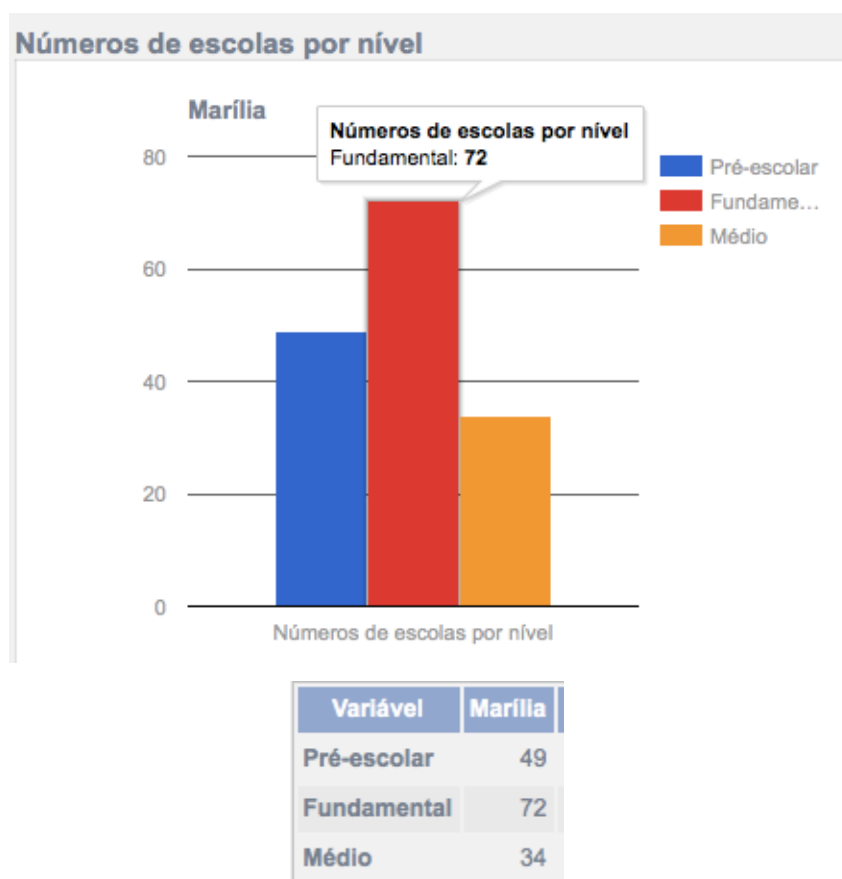


Fonte: IBGE

A municipalização da Educação no Estado de São Paulo foi iniciada em 1995, com base nos preceitos constitucionais definidos na Constituição Federal de 1988, a qual define no artigo 211, parágrafo 2º, a obrigatoriedade dos municípios atuarem com prioridade no Ensino Fundamental e na Educação Infantil. Em Marília, o processo de Municipalização do Ensino de 1ª à 4ª série do Primeiro Grau foi oficializado pela Lei Municipal nº 4.336, de 23 de outubro de 1997. Dessa forma, a Rede Municipal de Ensino Fundamental de Marília foi iniciada em 1998, contando com 3.230 alunos.

Em 2006, com a aprovação da Lei nº 11.274, que alterou a Lei de Diretrizes e Bases da Educação Nacional, o Ensino Fundamental passou a contar com nove anos de duração, com matrícula obrigatória a partir dos 6 anos de idade.

Em relação ao número de escolas, por nível, o município caracteriza-se, atualmente, da seguinte maneira da seguinte forma:

Gráfico 2- Escolas do município, por nível de educação

Fonte: IBGE

O Índice de Desenvolvimento da Educação Básica, no ano de 2011 (último resultado antes da coleta dos dados) mostrou que o IDEB do município de Marília era, à época da coleta, de 6,4.

Gráfico 3. IDEB do Município Marília

Fonte: QEdu.org.br. Dados do Ideb/ Inep (2011)

Os dados do Inep mostram, ainda, que, de 2011 a 2019, o IDEB do município teve um aumento significativo nas avaliações, tornando-se um dos municípios com maior índice de desenvolvimento educacional, comparado às médias do estado de São Paulo e do Brasil (para 4ª série/5º ano) que, em 2011, estavam em: 5,4 para o ensino público e 7,0 para o ensino privado; e 4,7 para o ensino público e 6,5 para o ensino privado, respectivamente.

Na tabela abaixo, apresentamos os números que mostram a progressão do IDEB do município de Marília entre os anos de 2011 (ano anterior à coleta dos dados) e 2019 (ano em que foi realizada a última avaliação).

Tabela 2. Índices do Ideb do Município de Marília - 2011-2019

2011	2013	2015	2017	2019
6.4	6.5	6.9	7.2	7.2

Fonte: INEP⁴⁰

Por meio dessas informações, é possível conhecer os perfis econômico e educacional do município de Marília e, dessa forma, conhecer as condições que permeiam os textos do *EscIn*.

A cidade de Marília contava, à época da coleta, com 19 escolas públicas municipais e 10 escolas privadas, com 7.633 e 2.743 alunos, respectivamente. Na formação do banco de dados, a participação das escolas se deu por parte de 10 escolas públicas e de 5 escolas privadas⁴¹, totalizando 3020 textos de alunos do ensino público e 1210 do sistema privado. Quando da constituição do *EscIn*, um dos interesses centrais era o de construir uma amostragem que englobasse os tipos de escola (públicas e privadas) de todas as regiões do município de Marília. Todavia, devido a recusa da única escola pública da Zona Leste da cidade, isso não foi possível, fazendo com que essa região do município fosse representada, no banco, apenas por escolas privadas.

Na tabela a seguir, são apresentados os totais de textos que integram o *EscIN* discriminados por escolas:

Tabela 3. Distribuição do Banco de Dados *EscIn*

Nome da escola	Rede	Número de textos
Américo Capellozza	Pública	553
Cecília Guelpa	Pública	159
Mário Covas	Pública	351
Antônio Ribeiro	Pública	520

⁴⁰ Disponível em: <http://ideb.inep.gov.br/resultado> Acesso em 23 de agosto de 2020, às 14:32hrs.

⁴¹ À época da coleta, os pesquisadores contataram todas as escolas do município, através de informações obtidas junto à Secretaria Municipal de Educação. Todavia, duas escolas públicas não se manifestaram e as demais se manifestaram contrariamente à participação, por estarem envolvidas em outros projetos/pesquisas.

Myrthes Negreiro	Pública	300
Nicácia Gil	Pública	174
Nivando Mariano	Pública	221
Paulo Freire	Pública	439
Chico Xavier	Pública	269
Antônio Egéa	Pública	34
		3020
Sagrado Coração	Privada	321
Água Viva	Privada	158
Esquema Único	Privada	115
Cristo Rei	Privada	395
Bezerra de Menezes	Privada	221
		1210

Fonte: elaboração própria

Esses números (3020 e 1210) correspondem a 78,39% e 28,61% dos dados que compõem o *EscIn*, divididos, respectivamente, entre escolas públicas e privadas.

Tendo como ponto de partida as considerações sobre a constituição do *EscIn*, na seção que segue, apresentamos as informações referentes à amostra selecionada para a realização desta investigação.

2.2 Sobre a amostra

A seleção dos textos do banco *EscIn* para a constituição da amostra se deu mediante um recorte⁴² representativo da totalidade dos textos, configurando, dessa forma,

⁴² Para o desenvolvimento desta pesquisa, nos valem da mesma amostra investigada por Torquette (2016) para estudo quantitativo. O recorte da totalidade dos textos para constituição da amostra é explicado mais detalhadamente em Torquette (2016).

uma amostragem estratificada. Esse recorte, realizado por meio de sorteio⁴³ através da ferramenta Excel, obedeceu a um critério de 10% do total de textos de cada variável, considerado estatisticamente representativo. Para a seleção dos textos a serem analisados, foram consideradas as variáveis (a) *tipo de escola* e (b) *série*.

A organização da amostra está representada na tabela a seguir:

Tabela 4. Divisão da amostra

	ESCOLA PÚBLICA	ESCOLA PRIVADA	TOTAL POR SÉRIE
1º ano	64	26	90
2º ano	59	26	85
3º ano	67	23	90
4º ano	37	22	59
5º ano	71	28	99
TOTAL			423

Fonte: elaboração própria

A amostra passou por organização e codificação, levando-se em conta as variáveis *tipo de escola*, *série* e *sexo*. Em nossa investigação não consideramos a variável *sexo*⁴⁴ nas análises, por entendermos que, embora os dados possam mostrar alguma diferença quando comparado o sexo dos escreventes, ela não se justificaria mediante o objetivo maior deste trabalho: analisar a configuração prosódica de segmentações não-convencionais de palavras, comparando-se os contextos de ensino público e de ensino privado. No entanto, para o desenvolvimento de estudo posterior, optamos por incluir a variável *sexo* na codificação dos textos. A tabela abaixo exemplifica como foi feita essa codificação:

Tabela 5. Codificação dos textos da amostra

Tipo de escola	Instituição	Série	Nº do texto	Sexo
PB	AC	1ª	01	F

⁴³ De acordo com essa ferramenta, é sorteado entre o 1º e o número máximo de textos. Por exemplo, com base no primeiro e o número máximo de textos, se o programa sortear o 4º texto, busca-se na lista de presença da turma esse escrevente, seleciona-se o texto e o inclui na amostra.

⁴⁴ Em seu trabalho, Torquette (2016) considerou a variável *sexo*. Uma vez que em nosso trabalho tal variável não é levada em conta, optamos por não apresentar aqui a divisão da amostra entre escreventes femininos e masculinos.

Legenda: **PB**: pública; **AC**: Américo Capelloza; **1A**: 1º ano A; **01**: número do texto, por ordem alfabética; **F**: feminino.

A partir dessa codificação, foi possível organizar, dentro da amostra, todos os textos segundo as variáveis que podem se prestar a análises (atuais ou futuras). Essa codificação (PB_AC_1A_01_F), por exemplo, é de um texto de escola pública, produzido por uma aluna do 1º ano A da Escola Municipal Américo Capelloza. O número do texto (**01**) respeitou a ordem alfabética, pós-sorteio desse texto para integrar a amostra.

2.3 Sobre os procedimentos metodológicos

Para realização desta pesquisa, optamos por uma análise transversal⁴⁵ dos textos relacionando as variáveis *tipo de escola* (pública ou privada) e *série* às possibilidades de organização prosódica⁴⁶ das hipersegmentações e das hipossegmentações.

No que concerne aos procedimentos adotados para a análise e interpretação dos dados, o delineamento dado à presente investigação envolveu, predominantemente, uma análise de natureza qualitativa, embora os dados tenham passado por quantificação para: (1) levantamento das ocorrências de segmentações não-convencionais de palavras nos textos; (2) classificação dos dados encontrados, segundo as variáveis extralinguísticas *tipo de escola* e *série*; (3) separação das segmentações não-convencionais segundo o tipo (hipossegmentação, hipersegmentação); (4) tratamento estatístico dos dados, posteriormente à análise prosódica, para comparação dos resultados.

A presente investigação se deu, portanto, mediante o cumprimento dos seguintes passos:

1. levantamento das ocorrências de segmentação não-convencional de palavras, considerando, respectivamente, o *tipo de escola* e a *série* e tabulação dos dados. Para realizarmos essa tarefa, observamos a quantidade de dados levando em conta: (i) o número de palavras convencionais e o número de segmentações não-convencionais em cada texto; e (ii) a relação entre a extensão dos textos (dada pelo número de palavras escritas) e o número de dados encontrados. Em cada texto, contabilizamos o número de

⁴⁵ Admitimos que um olhar longitudinal sobre essa mesma amostra pode fornecer informações relevantes a respeito do processo de escrita (especialmente das grafias de segmentação não-convencional) desses escreventes ao longo das séries ensino fundamental I. Todavia, a opção pelo delineamento metodológico transversal se deu em função de esse procedimento possibilitar a comparação dos dados de cada contexto (do público e do privado) por série escolar, permitindo um olhar mais focalizado sobre os dados.

⁴⁶ Conforme modelo fonológico proposto por Nespor e Vogel (1986).

- palavras escritas⁴⁷ e o número de ocorrências de segmentações não-convencionais, para posteriormente estabelecermos uma relação entre esses dois índices nos dois contextos (público e privado). Uma vez que os textos de escola pública estão em maior número na amostra, optamos por iniciar por eles. A fim de assegurar o olhar para o aspecto comparativo da pesquisa, adotamos como forma de trabalho a alternância entre dados de contexto público e de contexto privado. Dito de outra forma, para fins de organização e posterior comparação, o levantamento respeitou o seguinte critério: primeiramente foram levantados e quantificados os dados de 1ª ano de escola pública e, em seguida, os dados de 1ª ano de escola privada. Essa alternância foi adotada para os dados das demais séries;
2. classificação, dentro da variável *tipo de escola* (pública ou privada), das ocorrências de segmentação não-convencional segundo o seu tipo (hipersegmentação, hipossegmentação), de acordo com a mesma ordem de alternância entre público e privado, seguida de tabulação dos dados.
 3. análise das segmentações não-convencionais de palavras em constituintes prosódicos, observando-se, respectivamente, a natureza prosódica dos dados de: (1) hipossegmentações; (2) hipersegmentações; seguindo a alternância entre dados do ensino público e, depois, do privado.
 4. tabulação dos dados, após a análise prosódica, por *tipo de escola* e por *série* segundo as categorias prosódicas das hipossegmentações, das hipersegmentações.
 5. comparação dos dados por meio de tratamento estatístico;
 6. registro dos resultados.

Conhecidos os passos empreendidos para a realização deste estudo, à seção seguinte explicitamos os critérios que nortearam a seleção dos dados.

2.4 Critérios para seleção dos dados

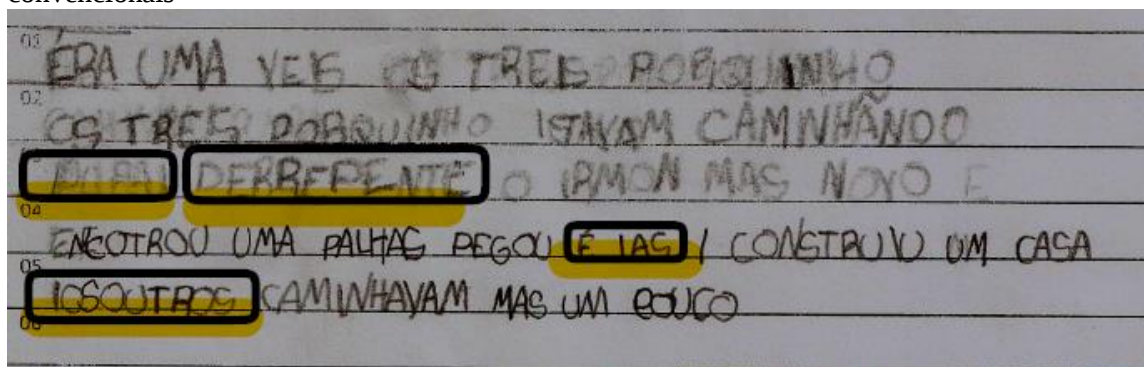
Para o levantamento do número de palavras escritas nos textos e posterior identificação das segmentações não-convencionais, baseamo-nos no critério gráfico, tomando como parâmetro as palavras convencionais (isto é, aquelas já estabelecidas pela convenção ortográfica). Entretanto, cabe salientar que alguns tipos ou tamanhos de letras,

⁴⁷ Para determinar essa contagem elegemos o critério gráfico. Nesse sentido, contabilizamos as palavras convencionais (estabelecidas pela convenção ortográfica) dos textos, para, posteriormente, observar a quantidade de segmentações não-convencionais de palavra.

ou mesmo rasuras⁴⁸, podem interferir quanto à identificação dos espaços em branco entre/nas palavras escritas. Nesse sentido, o trabalho de avaliação dos dados de escrita foi realizado em conjunto com outros pesquisadores⁴⁹ do GPEL, a fim de se evitarem possíveis equívocos de interpretação dos dados.

Na figura 3, ilustramos o um exemplo de como foi realizado o levantamento do número de palavras escritas e de ocorrências de segmentações não-convencionais:

Figura 17. Exemplo de levantamento do número de palavras escritas e de segmentações não-convencionais



TEXTO: PB_AC_1D_09M

Para o levantamento da quantidade de palavras escritas, levamos em conta que o texto acima é composto por 36 palavras convencionais⁵⁰. Há nele, entretanto, algumas palavras segmentadas de forma não-convencional: “puraí”, “derrepente”, “é las” e “iosoutros”. Nesses casos, consideramos a existência de 4 (quatro) ocorrências de segmentações não-convencionais, sendo três hipossegmentações (junção das palavras, em lugares não previstos pela convenção: “puraí”, “derrepente” e “iosoutros”) e uma hipersegmentação (separação de uma palavra convencional: “é las”).

A seleção dos dados de segmentação não-convencional de palavras nos textos levou em consideração, ainda, os seguintes aspectos: a possibilidade de interpretação da

⁴⁸ Os textos que apresentavam rasuras não foram excluídos do corpus. Nesse caso, optamos por considerar as partes de texto identificáveis e ignorar as partes que julgamos como não-identificáveis. As partes não-identificáveis foram desconsideradas tanto na contagem de palavras convencionais quanto no levantamento das segmentações não-convencionais.

⁴⁹ A avaliação dos dados aconteceu presencialmente entre os meses de julho e setembro de 2019, às sextas-feiras, no período da tarde, no Ibilce/Unesp – Campus de São José do Rio Preto (SP). Quando impossibilitados de ocorrerem na forma presencial, estes foram realizados por meio da ferramenta *Hangout*. O grupo de discussão era composto por mestrandos e doutorandos do PPGEL, integrantes do GPEL, cujas pesquisas têm como objeto a escrita infantil.

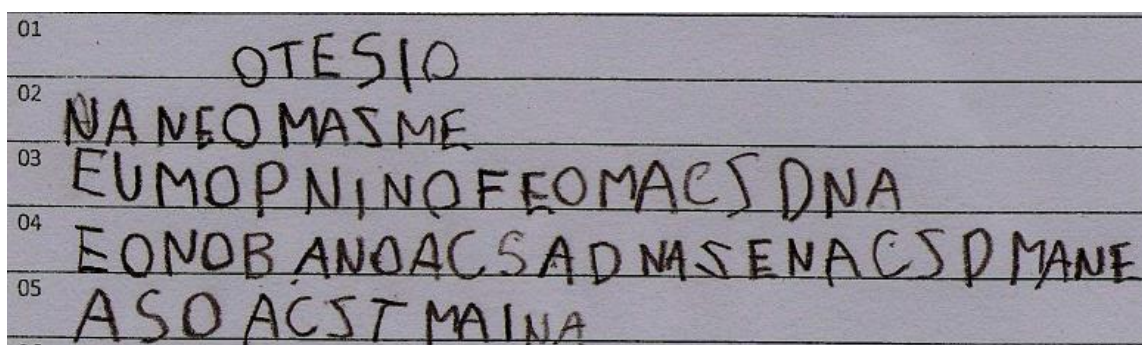
⁵⁰ Leia-se: “era uma vez os três porquinhos / os três porquinhos estavam caminhando / por aí de repente o irmão mais novo e / encontrou umas palhas e pegou elas e construiu uma casa / e os outros caminhavam mais um pouco”

caligrafia do escrevente; a translineação; o uso do hífen e o tipo de segmentação não-convencional envolvido na grafia. Além disso, dado o interesse dessa investigação se concentrar na configuração prosódica das estruturas, optamos por não incluir nas discussões, os dados segmentação não-convencional cujos limites não coincidem com nenhum dos constituintes prosódicos da hierarquia de Nespor e Vogel (1986)⁵¹, por entendermos que, de fato, as respostas e a explicação para essas ocorrências têm seu fundamento em aspectos da ordem do letramento.

Diante de todo o exposto, por questões metodológicas, excluimos do corpus:

- Grafias em que não é possível identificar com clareza (dada sua configuração ou a presença de muitas rasuras) a existência de dados passíveis de compreensão, como em:

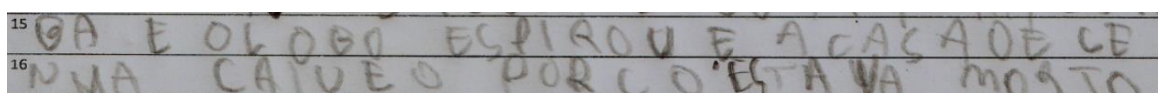
Figura 18. Exemplo de texto não identificado



Texto- PB_CX_1A_31M

- Os casos de translineação, como o que se vê na Figura 5 (“e o lobo espirou e acasade *le nha* caiueo porco estava morto”), em que a palavra “lenha” foi separada por espaço em branco:

Figura 19. Exemplo de caso de translineação



Texto - PB_PF_1A_85M

Interpretação: “e o lobo espirou e a casa de lenha caiu e o porco estava morto”

- As palavras em que se faz obrigatório o uso do hífen (palavras compostas e casos de ênclise, como “comê-los”):

Figura 20. Exemplo de caso de uso obrigatório de hífen

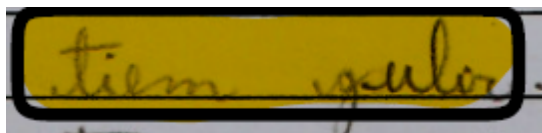


Texto- PB_ PF_3A_ 195M

Interpretação: “Comê-los”

- As segmentações não-convencionais denominadas *mesclas*⁵² (em que é possível identificar a ocorrência de separação e de junção de partes da palavra simultaneamente):

Figura 21. Exemplo de mescla



Texto - PB_ CX_4A_ 231M

Interpretação: “Te engolir”

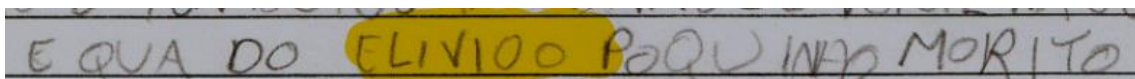
Uma vez elencados, à Seção 1.6 (*Fonologia Prosódica*), os constituintes prosódicos que se mostraram relevantes para a análise de segmentações não-convencionais de palavras feita neste trabalho, cabe explicitar aqui, outro aspecto de ordem metodológica. Uma vez que, dentre dados segmentações não-convencionais encontrados nos textos uma parte era composta de estruturas que não coincidiam com nenhum dos domínios prosódicos da língua - e portanto, não poderiam ser explicadas em função de categorias fonético-fonológicas⁵³ - optamos por excluí-las do corpus.

Um exemplo de dado dessa natureza é o que apresentamos adiante:

⁵² Essa natureza de dado foi explicitada à seção 1.

⁵³ Esses tipos de ocorrências, como já discutido em Chacon (2004), reforçam o princípio, defendido por Corrêa (2001), de que os dados de escrita – e aqui, especificamente, as segmentações não-convencionais – seriam marcas privilegiadas de observação da heterogeneidade da escrita.

Figura 22. Exemplo de hipossegmentação com ruptura constituinte prosódico



Texto:

Interpretação: “E quando ele viu o porquinho morto”

[E quando] ϕ [ele viu] ϕ [o porquinho morto] ϕ] I

Como vemos no exemplo acima, há dois casos de segmentação não-convencional de palavra, as quais correspondem a uma hipersegmentação e uma hipossegmentação, respectivamente: “qua do” e “elivioo”. A hipersegmentação “qua do” pode ser classificada, quanto à sua configuração prosódica, como uma estrutura em que a criança grafa duas sílabas (“qua” e “do”) como duas palavras distintas. Em contrapartida, no que diz respeito à hipossegmentação “elivioo” não é possível, na estrutura grafada, o reconhecimento de um domínio prosódico, pois, como pode ser observado, a segmentação não-convencional originada rompe com o domínio prosódico da frase fonológica. Como é possível observar as sentenças “ele viu” e “o porquinho morto” compreendem, prosodicamente, a duas frases fonológicas distintas: [ele viu] ϕ [o porquinho morto] ϕ . Quando a criança grafa a palavra “elivioo” ela rompe com a estrutura prosódica de ϕ , tomando, para a grafia da palavra, parte de uma e de outra ϕ .

Esse exemplo se presta à explicação de que, dentre as segmentações não-convencionais, há estruturas que não podem ser explicadas em função de categorias prosódicas, tal como mostrado por Chacon (2004)⁵⁴.

Por ocasião dessa explicação, “qua do” foi um dado considerado nas análises, ao passo que “elivioo” foi um dado excluído de nossa amostra. Explicitamos, a seguir, os procedimentos empreendidos no processamento estatístico dos dados.

2.5 Procedimentos da análise estatística

Para o tratamento estatístico dos dados foram utilizados o teste não paramétrico *Teste de Mann-Whitney* e o Intervalo de Confiança para Média. Na análise estatística foram utilizados os softwares: SPSS V20, Minitab 16 e Excel Office 2010.

O *Teste de Mann-Whitney* é usado quando há amostras independentes e em que se quer comparar sempre duas-a-duas as variáveis. O teste de Mann-Whitney teste foi

⁵⁴ Chacon (2004) defende, em seu trabalho, o ganho explicativo de se verem tais fatos de escrita sob a perspectiva de sua heterogeneidade de constituição.

desenvolvido primeiramente por F. Wilcoxon em 1945, para comparar tendências centrais de duas amostras independentes de tamanhos iguais. Em 1947, H.B. Mann e D.R. Whitney generalizaram a técnica para amostras de tamanhos diferentes. O teste de Mann-Whitney (Wilcoxon rank-sum test) é indicado para comparação de dois grupos não pareados para se verificar se pertencem ou não à mesma população. Nesses casos, verifica-se se há informações para se acreditar que valores de um grupo A são superiores a valores de um grupo B. O teste U pode ser considerado a versão não paramétrica do teste t, para amostras independentes. Ao contrário do teste t, que testa a igualdade das médias, o teste de Mann-Whitney (U) testa a igualdade das medianas. Assim, ele é composto pela seguinte hipótese:

$$\begin{cases} H_0 : \text{As duas amostras proveêm de uma única população} \\ H_1 : \text{As duas amostras são de populações diferentes} \end{cases}$$

O procedimento adotado compreende:

- a) ordenar todas as observações, independentemente de a qual amostra pertençam, e atribuir postos;
- b) obter

P_1 = soma dos postos das observações pertencentes ao grupo 1

P_2 = soma dos postos das observações pertencentes ao grupo 2

$$T = P_1 - \frac{n_1(n_1 + 1)}{2}$$

- c) calcular

$$T_0 = \frac{T - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

Onde:

n_1 = tamanho da amostra 1

n_2 = tamanho da amostra 2

d) obter, na tabela normal padrão, o valor $Z_{\alpha/2}$ tal que $P(Z > Z_{\alpha/2}) = \frac{\alpha}{2}$ e

$$P(Z < -Z_{\alpha/2}) = \frac{\alpha}{2}$$

e) se $T_0 > Z_{\alpha/2}$ ou $T_0 < -Z_{\alpha/2}$, rejeite H_0

O intervalo de confiança para a Média, por sua vez é uma técnica utilizada quando queremos ver o quanto ela pode variar numa determinada probabilidade de confiança. Essa técnica é descrita da seguinte maneira:

$$P\left(\bar{x} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

Onde:

$\bar{x}$ = média amostral;

$Z_{\alpha/2}$ = percentil da distribuição normal;

σ = variância amostral (estatística não viciada da variância populacional);

μ = média populacional;

α = nível de significância;

O resultado de cada comparação resulta em uma estatística chamada de *p-valor*. Esta estatística é que ajuda a concluir sobre o teste realizado. Caso esse valor seja maior do que o nível de significância adotado (erro ou α), concluímos que a H_0 (a hipótese nula) é a hipótese verdadeira; caso contrário, ficamos com H_1 , a hipótese alternativa.

SEÇÃO 3: DESCRIÇÃO DOS RESULTADOS

Nesta seção, serão apresentados os resultados obtidos nesta investigação. A organização desta seção responde ao nosso interesse de investigar as ocorrências de segmentações não-convencionais de palavras em textos produzidos por crianças do 1º ao 5º ano de ensino fundamental, de modo a estabelecer um comparativo entre os contextos de ensino público e privado. Para tanto, à subseção 3.1 apresentamos os resultados das quantificações relacionadas ao número de palavras escritas, ao passo que à seção 3.2 é apresentado o levantamento das segmentações não-convencionais, para mostrar como se distribuem em relação aos contextos e às séries. À seção 3.3 são mostrados os resultados do processamento estatístico dos dados após a análise das segmentações não-convencionais em constituintes prosódicos. Ao final, à seção 3.3, realizamos a discussão sobre as tendências encontradas.

3.1 Sobre o número de palavras escritas nos dois contextos

Dado o objetivo de compararmos a distribuição dos resultados em função do tipo de escola (pública ou privada), realizamos o levantamento prévio da quantidade de palavras escritas pelos alunos⁵⁵, por séries, nos dois contextos. Tal levantamento trouxe como resultados os seguintes números:

Tabela 6. Número de palavras escritas por séries nos contextos público e privado

ENSINO PÚBLICO			ENSINO PRIVADO		
	Nº Pal.		Nº Pal.		
1º ANO	3154		2087		
2º ANO	7631		3189		
3º ANO	6311		4696		
4º ANO	6821		3998		
5º ANO	16480		6910		

Leia-se: Nº Pal. = número de palavras.

⁵⁵ Relembremos que os critérios adotados para a realização desse levantamento encontram-se à seção 2.3.

A aplicação da testagem (U) Mann-Whitney sobre essas quantidades de palavras – comparados os contextos, por série – mostrou os seguintes resultados:

Tabela 7. Comparação do número de palavras de acordo com a seriação em escolas públicas e privada

			Média	Mediana	Desvio Padrão	Q1	Q3	N	IC	P-valor
Palavras	1º ano	Privada	90,7	49	72,9	38	139	23	29,8	0,268
		Pública	68,6	56	57,9	25	101	46	16,7	
	2º ano	Privada	133,0	115	66,1	90	167	24	26,5	0,609
		Pública	133,9	131	56,6	99	173	57	14,7	
	3º ano	Privada	154,5	155	37,5	141	168	2	51,9	0,478
		Pública	176,4	175	56,9	135	208	21	24,3	
	4º ano	Privada	181,7	190	83,3	109	214	22	34,8	0,654
		Pública	189,5	185	68,7	149	236	36	22,4	
	5º ano	Privada	246,8	238	89,4	178	295	28	33,1	0,555
		Pública	232,1	230	77,4	178	276	71	18,0	

Fonte: dados obtidos pelo pesquisador. Teste Mann-Whitney para amostras independentes (p -valor $\leq 0,05$). **Legenda:** Q1- 1º quartil; Q3- 3º quartil; N - IC- intervalo de confiança.

O Mann-Whitney testa a igualdade das medianas. A mediana é uma medida de posição que divide a amostra ao meio, resultando em que 50% dos indivíduos estejam acima do valor da mediana e 50% abaixo. Esta é uma estatística analisada em relação à média, pois quanto mais próximo seu valor for em relação à média, mais simétrica será a distribuição. Já uma distribuição assimétrica certamente mostra uma grande variabilidade.

A variabilidade é medida pelo desvio padrão. Quanto mais próximo (ou maior) esse valor for em relação à média, maior será a variabilidade, o que aponta para uma não homogeneidade dos dados. Os quartis são estatísticas descritivas de posição, ou seja, não são influenciadas por valores extremos (como a média e desvio padrão). O 1º quartil (Q1) mostra a distribuição até 25% da amostra e o 3º quartil (Q3) mostra a distribuição em até 75% da amostra.

O intervalo de confiança (IC) ora somado ora subtraído da média mostra a variação da média segundo uma probabilidade estatística. Também aqui esses limites não

têm relação com o cálculo de mais ou menos um desvio padrão em relação à média. Salienta-se que o IC é mais confiável pois indica uma probabilidade estatística associada em seu cálculo.

Nesse caminho, na análise comparativa da quantidade de palavras escritas nos dois contextos (público e privado), os resultados não apontaram diferença estatisticamente significativa entre as escolas nas cinco séries avaliadas. Nesse sentido é possível afirmar que, embora possa haver diferenças entre os valores (maiores no contexto público), eles não são representativos de variabilidade. Dessa forma, a diferença entre o número de palavras escritas pelos alunos de escolas públicas e o número de palavras escritas pelos alunos de escolas privadas não é significativa em termos estatísticos, porque, como mostra o teste, o **p-valor**⁵⁶ é $\leq 0,05$ em todas as séries, comparados os dois contextos.

Esse resultado, cabe destacar, vai de encontro àqueles obtidos por Torquette (2016), para quem a diferença no número de palavras de escreventes de escolas públicas e de escreventes de escolas privadas se mostrou altamente significativa na estatística inferencial⁵⁷, apontando que os escreventes de escolas públicas escreveriam menos, em relação aos escreventes de escolas privadas.

No entanto, em consonância com os dados obtidos pela autora, em nossa análise também é possível observar o aumento gradativo do número de palavras ao longo das séries. Essa tendência já foi observada, também, por Cunha (2004), ao estudar os processos de segmentações não-convencionais encontrados em textos de escrita espontânea, produzidos por alunos de 1ª a 4ª série, em que se observou que, com o avanço da seriação, o número de palavras escritas nos textos aumentava.

⁵⁶ Lembremos que o resultado de cada comparação apresenta uma estatística chamada de p-valor.

⁵⁷ A autora aplicou aos dados a estatística inferencial, usando o T-test para comparação dos dados de amostras dependentes.

3.2 Sobre a ocorrência de segmentações não-convencionais nos contextos público e privado

Tabela 8. Distribuição das ocorrências de segmentações não-convencionais nos contextos e nas séries

ENSINO PÚBLICO			ENSINO PRIVADO	
	Nº Pal.	SNC	Nº Pal.	SNC
1ºANO	3154	319	2087	102
2º ANO	7631	337	3189	81
3º ANO	6311	305	4696	29
4º ANO	6821	79	3998	22
5º ANO	16480	165	6910	18

Legenda: Nº Pal. = número de palavras; SNC = segmentações não-convencionais.

Os dados gerais mostram que as ocorrências de segmentação não-convencional aparecem sempre em maior número no contexto público, como se vê na tabela acima. Entretanto, faz-se necessário analisar a maneira como os dados estão proporcionalmente distribuídos, uma vez que não poderíamos fazer comparações quantitativas diretas. A seguir, apresentamos os resultados da aplicação do teste (U) Mann-Whitney aos números de segmentação não-convencional de palavra contabilizados nos dois contextos (público e privado):

Tabela 9. Comparação do número de segmentações não-convencionais em escolas públicas e privadas, por série.

			Média	Mediana	Desvio Padrão	Q1	Q3	N	IC	P-valor
SNC	1º ano	Privada	4,25	4	3,50	1,75	7,25	24	1,40	0,996
		Pública	5,50	3,5	6,28	0	8,75	58	1,61	
	2º ano	Privada	3,88	1	9,36	0	1,25	24	3,75	<0,001
		Pública	5,91	4	5,05	2	9	57	1,31	
	3º ano	Privada	1,50	1,5	2,12	0,75	2,25	2	2,94	0,579
		Pública	4,33	2	10,40	1	4	21	4,45	
		Privada	1,09	0,5	1,38	0	2	22	0,58	0,255

	4º ano	Pública	2,31	1	4,48	0	3	36	1,46	
	5º ano	Privada	0,64	0	0,99	0	1	28	0,37	0,019
		Pública	2,32	1	4,75	0	2	71	1,10	
%	1º ano	Privada	6,51%	6,0%	6,09%	1,8%	7,8%	23	2,49%	0,023
		Pública	11,63%	10,4%	9,42%	4,4%	16,8%	46	2,72%	
	2º ano	Privada	3,11%	0,7%	6,03%	0,0%	1,7%	24	2,41%	0,001
		Pública	4,89%	3,6%	4,06%	1,5%	8,0%	57	1,05%	
	3º ano	Privada	0,83%	0,8%	1,17%	0,4%	1,2%	2	1,62%	0,511
		Pública	2,44%	1,2%	4,83%	0,4%	2,2%	21	2,07%	
	4º ano	Privada	0,73%	0,3%	1,16%	0,0%	1,0%	22	0,48%	0,401
		Pública	1,87%	0,5%	5,18%	0,0%	1,2%	36	1,69%	
	5º ano	Privada	0,22%	0,0%	0,31%	0,0%	0,4%	28	0,12%	0,007
		Pública	1,21%	0,5%	2,58%	0,0%	1,1%	71	0,60%	

Fonte: dados obtidos pelo pesquisador no levantamento de dados da amostra

Na análise da quantidade de segmentações não-convencionais, identificamos diferença estatisticamente significativa no 2º ano com média de 3,88 nas escolas privadas contra 5,91 em Pública (p-valor <0,001) e medianas 1 e 4, respectivamente. Essa informação mostra que os alunos das escolas públicas mantêm a proporcionalidade de ocorrências de segmentação não-convencional mesmo com o avanço nas séries do 1º para o 2º ano. No contexto do ensino privado, essa proporcionalidade não se aplica, uma vez que os dados do 2º ano, nesse contexto, mostram redução significativa no número de segmentações não-convencionais em relação ao número de palavras escritas, comparado ao do 1º ano.

Essa diferença expressiva entre os contextos se mostra, também, nos dados de 5º ano. Nessa série, os dados de escolas privadas têm média de 0,64 contra 2,32 das públicas (p-valor = 0,019), mostrando que, ao final desse ciclo escolar (ensino fundamental I), os alunos de escola privada produzem segmentações não-convencionais em número inferior, se comparados aos alunos do 5º ano do contexto público.

Na análise do percentual de segmentações não-convencionais em relação ao número de palavras, encontramos diferença entre as escolas no 1º ano, 2º ano e 5º ano. Nas três séries notamos que o resultado foi sempre maior na escola pública, como entre os alunos do 1º ano, onde tivemos uma média de 11,63% em escola pública contra 6,51% em privada (p-valor = 0,023). Adiante à seção 4 explicitamos de modo mais aprofundado acerca dessa variabilidade.

3.3 Sobre as hipossegmentações e as hipersegmentações nos dois contextos: resultados da análise prosódica.

Concluída a apresentação dos dados obtidos no levantamento do número de segmentações não-convencionais, feita à subseção 3.2, passamos, neste momento, à descrição dos resultados da análise prosódica das hipossegmentações e das hipersegmentações respondendo aos objetivos específicos desta investigação de: a) analisar prosodicamente as ocorrências de hipossegmentações e de hipersegmentações produzidas nos dois contextos; b-) comparar os dois contextos, a fim de verificar a existência de variabilidade na configuração prosódica das estruturas analisadas.

Primeiramente, nesta seção, mostraremos como se distribuem os dados de segmentação não-convencional segundo o seu tipo. Adiante, apresentaremos os resultados das análises em constituintes prosódicos, conforme Nespor e Vogel (1986) e considerando os estudos feitos a partir desse arcabouço para o PB, elencando, em subseções distintas, os resultados obtidos para as hipossegmentações, e, em seguida, para as hipersegmentações.

Para termos uma dimensão de como ocorre a distribuição dos dados de segmentação não convencional na amostra segundo o seu tipo, apresentamos a tabela abaixo:

Tabela 10. Distribuição das ocorrências de segmentações não-convencionais nos contextos e nas séries por tipo.

	PÚBLICO		PRIVADO	
	HIPO	HIPER	HIPO	HIPER
1º ANO	218	62	75	12
2º ANO	247	70	66	10
3º ANO	192	109	10	19
4º ANO	51	27	13	9
5º ANO	114	47	6	12

Fonte: Dados obtidos pelo pesquisador. Legenda: HIPO = hipossegmentações; HIPER = hipersegmentações.

Os dados de hipossegmentação, de acordo com o que vemos na tabela acima, aparecem em maior número do que os demais tipos de dados. A hipótese de que, no início da aquisição da escrita, a criança parece partir da compreensão de unidades linguísticas maiores já foi levantada em estudos como os de Abaurre (1987) e de Cagliari (2002) e seria, portanto, a justificativa para prevalência das hipossegmentações nos dados de escrita infantil.

Com base no levantamento dos tipos de segmentação não-convencional e na observação de como eles se distribuem ao longo das séries, podemos ter uma dimensão dessas ocorrências na amostra. Às seções que seguem são expostos os resultados da análise prosódica das hipossegmentações (subseção 3.3.1), das hipersegmentações (subseção 3.3.2) bem como do processamento estatístico desses dados.

3.3.1 Sobre as hipossegmentações

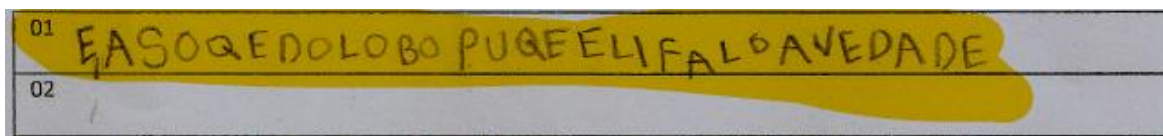
Das ocorrências de segmentações não-convencionais identificadas na amostra, 992 unidades correspondem a dados de hipossegmentação, sendo 822 delas no contexto do ensino público e 170 no ensino privado. Adiante, apresentamos exemplos de hipossegmentações analisadas prosodicamente, separando-os, hierarquicamente, de acordo com o domínios prosódicos identificados nessas estruturas⁵⁸

⁵⁸ A exemplificação proposta neste momento, se dá a título de retomada, uma vez que, à seção 1.5, quando da apresentação da hierarquia prosódica, já exemplificamos cada um dos domínios em estruturas de hipossegmentação nosso corpus.

Passemos aos exemplos, os quais apresentamos na seguinte ordem: enunciado fonológico (U); frase entonacional (I); frase fonológica (ϕ); e grupo clítico (C):

- **No domínio de U:**

Figura 23. Exemplo de hipossegmentação de enunciado fonológico



Texto: PB_AC_1B_07M

Interpretação: “eu acho que é do lobo porque ele falou a verdade”

[easoqedolobopuqelifaloavedade] U

[easoqedolobo] I [puqelifaloavedade] I] U

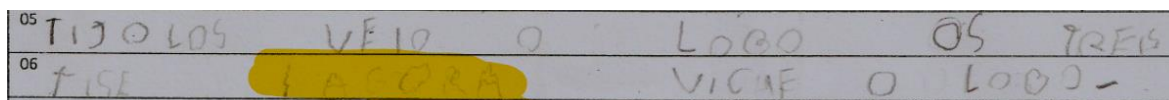
A figura acima mostra o fragmento de um texto produzido por um aluno do 1º ano de uma escola pública. Esse texto é composto por 11 palavras, mas se caracteriza como uma única estrutura hipossegmentada. Na classificação prosódica desse enunciado, observamos a existência de um enunciado fonológico (U) composto por duas frases entonacionais (I).

Hipossegmentações dessa natureza prosódica, no domínio do constituinte mais alto da hierarquia prosódica, como veremos adiante, se dão em um percentual muito pequeno em nossos dados e tendem a ocorrer especialmente na escrita de alunos de 1º ano.

No domínio de I:

O exemplo a seguir mostra uma ocorrência de hipossegmentação que marca os limites da frase entonacional (I). Para Nespor e Vogel (1986 p. 188)⁵⁹ “a frase entonacional é o âmbito de um contorno de entoação em que os finais das frases entonacionais coincidem com as posições em que se podem introduzir pausas em uma oração”. Observemos:

⁵⁹ (...) the intonation phrase is the domain of the intonation contour and [...] the ends of intonational phrases coincide with the positions in which pauses may be introduced in a sentence." (NESPOR; VOGEL, 1986, p. 188)

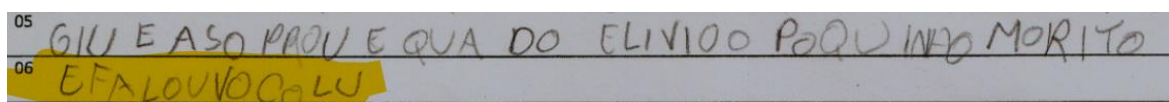
Figura 24. Exemplo de hipossegmentação de frase entonacional

Texto: PV_AV_1A_67M

Interpretação: (..) “veio o lobo/ os três disse/ e agora? / vigie o lobo!”

[IAGORA] I

A ocorrência “e agora?” é marcada por um contorno entonacional característico de interrogação. A frase entonacional pode ser composta por uma única frase fonológica, como é o caso do exemplo acima, ou por um conjunto de frases fonológicas, como vemos no exemplo abaixo, em que temos uma (I) composta por duas frases fonológicas: [e falou] ϕ [vou come-lo] ϕ .

Figura 25. Exemplo de hipossegmentação de frase entonacional

Texto: PV_PF_1B_56F

Interpretação: “e assoprou e quando ele viu o porquinho morto **e falou vou comê-lo**”

Análise prosódica: [E falou] I [vou comê-lo] I

- **No domínio de ϕ :**

O exemplo abaixo explicita uma ocorrência de hipossegmentação que se dá no domínio da frase fonológica ϕ . A frase fonológica é o constituinte imediatamente superior ao grupo clítico, ou seja, ela agrupa um ou mais clíticos. Dada a hierarquia, considera-se como grupo clítico tanto uma locução quanto apenas uma palavra fonológica. Na sentença apresentada a seguir, por exemplo, vemos uma ϕ formada por dois C: [a casa] C [do segundo] C. Nesse sentido, na figura abaixo, há uma ocorrência de hipossegmentação que se dá por meio da junção, por parte do aluno, de duas frases fonológicas: [a casa] ϕ [do segundo] ϕ .

Figura 26. Exemplo de hipossegmentação de frase fonológica



Texto: PB_MN_1A_43F

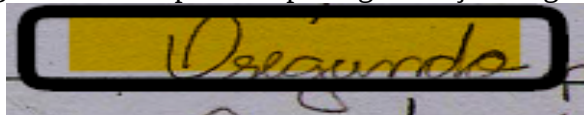
Interpretação: (...) quando espirrou / **a casa do segundo**

Análise prosódica: [[ACASA]C] ϕ [DOSEGUNTO] C] ϕ

- **No domínio de C:**

Para Nespor e Vogel (1986), os clíticos não se sustentam como palavra em um enunciado, pois dependem fonologicamente de uma outra palavra. Nesse caminho, Bisol (1996) define o grupo clítico como a unidade prosódica que contém um ou mais clíticos e uma só palavra de conteúdo. No exemplo que segue podemos ver uma hipossegmentação em que ocorre a junção do clítico “o” à palavra de conteúdo “segundo” formando um C.

Figura 27. Exemplo de hipossegmentação de grupo clítico



Texto: PV_BM_4A_294F

Interpretação: “O segundo”

[OSEGUNDO] C

Apresentados esses exemplos, adiante expomos a tabela em que constam os dados obtidos, em cada um dos domínios prosódicos exemplificados:

Tabela 11. Distribuição das Hipossegmentações nos contextos, por série e em cada domínio prosódico.

			Média	Mediana	Desvio Padrão	Q1	Q3	N	IC	P-valor
U	1º ano	Privada	0,000	0	0,000	0	0	24	- x -	0,520
		Pública	0,017	0	0,131	0	0	58	0,034	
	2º ano	Privada	0,042	0	0,204	0	0	24	0,082	0,123
		Pública	0,000	0	0,000	0	0	57	- x -	
	3º ano	Privada	0,000	0	0,000	0	0	2	- x -	1,000
		Pública	0,000	0	0,000	0	0	21	- x -	
	4º ano	Privada	0,000	0	0,000	0	0	22	- x -	1,000
		Pública	0,000	0	0,000	0	0	36	- x -	
	5º ano	Privada	0,000	0	0,000	0	0	28	- x -	1,000
		Pública	0,000	0	0,000	0	0	71	- x -	
I	1º ano	Privada	0,083	0	0,282	0	0	24	0,113	0,126
		Pública	0,328	0	0,758	0	0	58	0,195	
	2º ano	Privada	0,083	0	0,282	0	0	24	0,113	0,620
		Pública	0,123	0	0,683	0	0	57	0,177	
	3º ano	Privada	0,000	0	0,000	0	0	2	- x -	0,758
		Pública	0,143	0	0,655	0	0	21	0,280	
	4º ano	Privada	0,000	0	0,000	0	0	22	- x -	0,434
		Pública	0,028	0	0,167	0	0	36	0,054	
	5º ano	Privada	0,000	0	0,000	0	0	28	- x -	0,530
		Pública	0,028	0	0,237	0	0	71	0,055	
Φ	1º ano	Privada	0,417	0	0,776	0	1	24	0,310	0,536
		Pública	0,638	0	1,103	0	1	58	0,284	
	2º ano	Privada	0,667	0	2,297	0	0	24	0,919	0,215
		Pública	0,526	0	0,947	0	1	57	0,246	
	3º ano	Privada	0,000	0	0,000	0	0	2	- x -	0,655
		Pública	0,857	0	3,705	0	0	21	1,585	
	4º ano	Privada	0,000	0	0,000	0	0	22	- x -	0,019
		Pública	0,306	0	0,624	0	0	36	0,204	

	5º ano	Privada	0,000	0	0,000	0	0	28	- x -	0,114
		Pública	0,169	0	0,717	0	0	71	0,167	
C	1º ano	Privada	1,958	1,5	1,944	0	4	24	0,778	0,293
		Pública	1,517	1	1,828	0	3	58	0,470	
	2º ano	Privada	1,167	0	3,435	0	0,25	24	1,374	<0,001
		Pública	2,263	2	2,461	0	4	57	0,639	
	3º ano	Privada	0,000	0	0,000	0	0	2	- x -	0,070
		Pública	1,857	1	3,979	0	1	21	1,702	
	4º ano	Privada	0,409	0	0,796	0	0,75	22	0,333	0,892
		Pública	0,639	0	2,045	0	0,25	36	0,668	
		Pública	0,944	0	2,348	0	1	71	0,546	
	5º ano	Privada	0,107	0	0,315	0	0	28	0,117	0,002
		Pública	0,944	0	2,348	0	1	71	0,546	

Legenda: U = enunciado fonológico; I = frase entonacional; ϕ = frase fonológica; C = grupo clítico.

p-valores: considerados estatisticamente significativos perante o nível de significância adotado.

p-valores: que, por estarem próximos do limite de aceitação, são considerados como tendendo a ser significativos (até 5 pontos percentuais acima do valor do alfa adotado).

p-valores: considerados não significativos perante o nível de significância adotado.

- x -, quando não foi possível utilizar a estatística.

Fonte: dados obtidos pelo pesquisador.

Comparados os dados dos dois contextos (público e privado), não foram encontradas diferenças em relação às estruturas de hipossegmentação nos domínios prosódicos de enunciado fonológico (U) e frase entonacional (I) em nenhuma das séries. Em outras palavras, em se tratando desses dois domínios, não houve variabilidade considerável (em termos estatísticos) entre os dois contextos consideradas as variáveis *tipo de ensino* e *série escolar*, o que indica a existência de uma certa regularidade nos critérios de segmentação mobilizados pelos alunos.

Por outro lado, foram encontradas diferenças estatisticamente relevantes em dois domínios prosódicos, a saber, **frase fonológica (ϕ)** e **grupo clítico (C)** quando comparados os dois contextos⁶⁰. No domínio de ϕ , a variabilidade se mostra nos dados de 4º ano, ao passo que, no domínio de C, a variabilidade pode ser verificada nos dados de 3º, 4º e 5º anos.

⁶⁰ À seção 4 apresentamos uma discussão mais aprofundada dos resultados de acordo com a configuração prosódica.

3.3.2 Sobre as hipersegmentações

Das segmentações não-convencionais identificadas na amostra, os dados de hipersegmentação - caracterizados pela presença de espaços em brancos em lugares não previstos pela convenção ortográfica - correspondem a 423 ocorrências, sendo 311 no contexto do ensino público e 112 no ensino privado.

Antes de explicitarmos os resultados, destacamos que foram pertinentes para a análise das hipersegmentações os constituintes palavra fonológica (ω) pé métrico (Σ) e sílaba (σ). Assim como já observado por Cunha (2004), ao contrário das hipossegmentações, cuja estrutura se assenta nos constituintes mais altos da hierarquia prosódica, as hipersegmentações (alocação de espaços dentro dos limites da palavra) têm sua estrutura assentada pelos constituintes mais baixos da hierarquia, como a sílaba e o pé métrico.

Em seu trabalho sobre as palavras hipersegmentadas em textos produzidos nos anos finais do ensino fundamental, Silva (2014) destacou que, já que a principal característica das hipersegmentações é ocorrer no nível da palavra, as informações rítmicas – que se formam em seu interior por meio de proeminências prosódicas – tem forte influência na sua constituição. Nesse sentido, a autora observou, em suas análises, que os limites das marcas de hipersegmentação ocorrem em pontos fonologicamente condizentes com fronteiras de constituintes prosódicos, especialmente os de sílaba e pé métrico.

A seguir passamos à exemplificação das hipersegmentações, hierarquicamente, de acordo com os constituintes prosódicos identificados:

- **Estruturas formadas por pé (Σ) + pé (Σ)**

Figura 28. Exemplo de hipersegmentação formada por $\Sigma + \Sigma$



Texto: PB_ AC_4E_264M

Interpretação: “Colocaram”

[CÓLO] Σ [CARO] Σ

O exemplo exposto traz uma segmentação não-convencional que coloca em destaque a separação de uma palavra prosódica, em duas estruturas compostas por pés trocaicos⁶¹ [cólo] Σ [caro] Σ . Esse tipo de estrutura contabilizou 11 ocorrências e foi observada nos dados de escolas públicas e privadas e, nos dois contextos, a partir do 2º ano.

- **Estruturas formadas por sílaba (σ) + sílaba (σ)**

No exemplo exposto a seguir vemos um caso de um dissílabo segmentado não-convencionalmente, em razão da inserção de espaço em branco nos limites da sílaba, de tal modo que cada uma das sílabas da palavra forme duas partes [QUA] σ [DO] σ . Essa natureza de ocorrência predomina nos dados analisados, totalizando 160 unidades, dentre as hipersegmentações encontradas em todo o corpus.

Figura 29. Exemplo de hipersegmentação formada por $\sigma + \sigma$



Texto: PB_ MN_3A_212M

Interpretação “Quando”

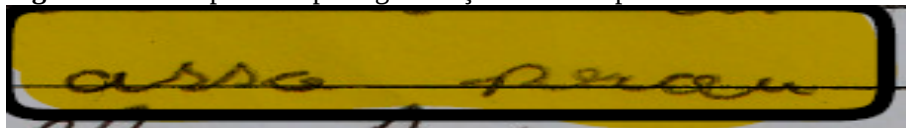
[QUA] σ [DO] σ

Além dessas estruturas, foram encontradas hipersegmentações que se dão por diferentes disposições dos constituintes sílaba e pé métrico, tal como vemos a seguir:

⁶¹ Ressalte-se: pé trocaico caracteriza-se pela formação de duas sílabas em que a proeminência (sílabas mais forte) está à esquerda, conforme mostra a representação, em que o ponto representa a sílaba átona e o asterisco, a sílaba tônica: (* .). Chama-se pé iâmbico o pé formado por duas sílabas com proeminência à direita: (.*)

- **Estruturas formadas por pé (Σ) + sílaba (σ):**

Figura 30. Exemplo de hipersegmentação formada por $\Sigma + \sigma$



Texto: PB_MN_4B_209F

Interpretação: “Assoprou”

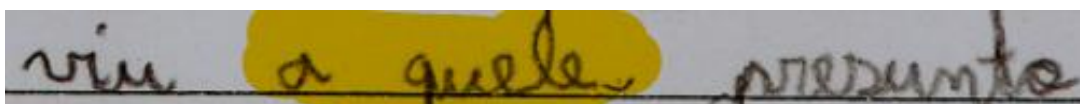
[ASSO] Σ [PROU] σ

Em “asso prou” a palavra foi segmentada de maneira a formar um pé métrico troqueu (Σ) e uma sílaba (σ). Esse tipo de estrutura não foi encontrada nos dados de segmentação não-convencional do ensino privado, entretanto somou, no ensino público, 13 ocorrências.

- **Estruturas formadas por sílaba (σ) + Σ (pé métrico):**

O exemplo a seguir ilustra uma hipersegmentação que pode ser analisada, em termos de constituição prosódica, como uma formação que marca o limite entre uma sílaba σ e um pé métrico do tipo troqueu (em que a proeminência está na sílaba à esquerda). Essa natureza de hipersegmentação corresponde a 154 ocorrências no corpus.

Figura 31. Exemplo de hipersegmentação formada por $\sigma + \Sigma$



Texto: PB_AR_5B_330M

Interpretação: “viu aquele presunto”

[A] σ [QUELE] Σ

- **Estruturas formadas por sílaba (σ) + sílaba (σ) + Σ (pé métrico):**

O exemplo exposto à figura acima corresponde a uma hipersegmentação em que a divisão da palavra em porções menores se dá em limites de sílaba e pé métrico. Trata-se da palavra “aconteceu”, a qual hipersegmentada em três partes distintas: [A] σ [COM]

σ [TECEU] Σ . Essa natureza de dado foi encontrada apenas no ensino público e apenas em dois momentos⁶².

Figura 32. Exemplo de hipersegmentação formada por $\sigma + \sigma + \Sigma$



Texto: PB_ PF_2E_141F

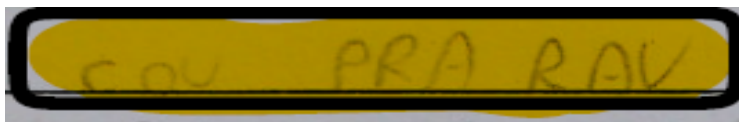
Interpretação: “Aconteceu”

[A] σ [COM] σ [TECEU] Σ

Há, ainda, na amostra, estruturas que se caracterizam pela segmentação de uma palavra trissílaba em três partes, correspondendo, cada uma delas a uma sílaba, como veremos no exemplo exposto a seguir.

- **Estruturas formadas por sílaba (σ) + sílaba (σ) + sílaba (σ)**

Figura 33. Exemplo de hipersegmentação formada por $\sigma + \sigma + \sigma$



Texto: PB_ PF_5A_377M

Interpretação: “Compraram”

[COU] σ [PRA] σ [RAU] σ

A hipersegmentação supracitada demonstra a inserção do espaço em branco em dois lugares no interior da palavra, levando a uma estrutura cuja separação se dá nos

⁶² A partir da análise de segmentações não convencionais em textos de alunos de Ensino Fundamental II, Tenani (2021) observou que ocorrências dessa natureza, por exemplo “*com ver sando*”, embora também não sejam tão recorrentes nos dados de EFII, têm em comum o fato de serem constituídas por partes que correspondem a monossílabos da língua “com”, “ver” e, nesse sentido, a hipersegmentação, poderia ser decorrente da confluência de informações prosódicas e semânticas. Isso ocorre, no exemplo “*a com teceu*” em que a inserção do espaço em branco marca fronteiras de estruturas que correspondem, separadamente, a três palavras ortográficas do PB.

limites da sílaba σ : [COU] σ [PRA] σ [RAU] σ . Ocorrências dessa natureza, isto é, que se caracterizam por apresentar dois espaços em branco não convencionais no interior da palavra, também são encontrados em raros momentos e, assim como o exemplo “*a com teceu*” vão ao encontro do que Tenani (2021) observou em dados de segmentação não-convencional de palavras na escrita de alunos do Ensino Fundamental II: as sílabas hipersegmentadas correspondem a monossílabos da língua, portanto, podem ser efeito de proeminências prosódicas e de relações morfossemânticas.

Encerrada essa descrição, apresentaremos, na tabela adiante os resultados obtidos do processamento estatístico das hipersegmentações, segundo os domínios prosódicos identificados. Os dados encontram-se organizados de modo a agrupar as ocorrências conforme os constituintes e a separá-las de acordo com o contexto e a série:

Tabela 12. Distribuição das Hipersegmentações nos contextos, por série, em cada domínio prosódico.

			Média	Me dia na	Desvio Padrão	Q1	Q3	N	IC	P- valor
$\Sigma + \Sigma$	2º ano	Privada	0,042	0	0,204	0	0	24	0,082	0,836
		Pública	0,053	0	0,225	0	0	57	0,058	
	3º ano	Privada	0,000	0	0,000	0	0	2	- x -	0,758
		Pública	0,048	0	0,218	0	0	21	0,093	
	4º ano	Privada	0,000	0	0,000	0	0	22	- x -	0,434
		Pública	0,028	0	0,167	0	0	36	0,054	
	5º ano	Privada	0,000	0	0,000	0	0	28	- x -	0,372
		Pública	0,028	0	0,167	0	0	71	0,039	
$\Sigma + \sigma$	1º ano	Privada	0,000	0	0,000	0	0	24	- x -	0,259
		Pública	0,069	0	0,317	0	0	58	0,082	
	2º ano	Privada	0,000	0	0,000	0	0	24	- x -	0,356
		Pública	0,053	0	0,294	0	0	57	0,076	
	3º ano	Privada	0,000	0	0,000	0	0	2	- x -	0,758
		Pública	0,048	0	0,218	0	0	21	0,093	
	4º ano	Privada	0,000	0	0,000	0	0	22	- x -	0,434
		Pública	0,028	0	0,167	0	0	36	0,054	

	5º ano	Privada	0,000	0	0,000	0	0	28	- x -	0,272
		Pública	0,056	0	0,287	0	0	71	0,067	
$\sigma + \Sigma$	1º ano	Privada	0,292	0	0,751	0	0	24	0,300	0,268
		Pública	0,448	0	0,820	0	1	58	0,211	
	2º ano	Privada	0,292	0	0,751	0	0	24	0,300	0,075
		Pública	0,579	0	0,925	0	1	57	0,240	
	3º ano	Privada	0,000	0	0,000	0	0	2	- x -	0,508
		Pública	0,238	0	0,539	0	0	21	0,231	
	4º ano	Privada	0,091	0	0,294	0	0	22	0,123	0,187
		Pública	0,278	0	0,566	0	0	36	0,185	
	5º ano	Privada	0,143	0	0,448	0	0	28	0,166	0,563
		Pública	0,183	0	0,457	0	0	71	0,106	
$\sigma + \sigma$	1º ano	Privada	0,333	0	0,816	0	0	24	0,327	0,762
		Pública	0,328	0	0,758	0	0	58	0,195	
	2º ano	Privada	0,167	0	0,381	0	0	24	0,152	0,483
		Pública	0,298	0	0,626	0	0	57	0,162	
	3º ano	Privada	0,500	0,5	0,707	0,25	0,75	2	0,980	0,695
		Pública	0,476	0	0,928	0	1	21	0,397	
	4º ano	Privada	0,318	0	0,780	0	0	22	0,326	0,435
		Pública	0,500	0	1,056	0	1	36	0,345	
	5º ano	Privada	0,286	0	0,713	0	0	28	0,264	0,948
		Pública	0,296	0	0,725	0	0	71	0,169	
$\sigma + \sigma + \Sigma$										
	2º ano	Privada	0,000	0	0,000	0	0	24	- x -	0,516
		Pública	0,018	0	0,132	0	0	57	0,034	
	3º ano	Privada	0,000	0	0,000	0	0	2	- x -	1,000
		Pública	0,000	0	0,000	0	0	21	- x -	
	4º ano	Privada	0,000	0	0,000	0	0	22	- x -	1,000
		Pública	0,000	0	0,000	0	0	36	- x -	
		Privada	0,000	0	0,000	0	0	28	- x -	1,000

	5º ano	Pública	0,000	0	0,000	0	0	71	- x -	
$\sigma + \sigma + \sigma$										
	3º ano	Privada	0,000	0	0,000	0	0	2	- x -	1,000
		Pública	0,000	0	0,000	0	0	21	- x -	
	4º ano	Privada	0,000	0	0,000	0	0	22	- x -	1,000
		Pública	0,000	0	0,000	0	0	36	- x -	
	5º ano	Privada	0,000	0	0,000	0	0	28	- x -	0,530
		Pública	0,014	0	0,119	0	0	71	0,028	

Legenda: Σ (pé métrico); σ (sílabas)

p-valores: considerados estatisticamente significativos perante o nível de significância adotado.

p-valores: que por estarem próximos do limite de aceitação, são considerados que tendem a ser significativos (até 5 pontos percentuais acima do valor do alfa adotado).

p-valores: considerados não significativos perante o nível de significância adotado.

- x -, quando não foi possível utilizar a estatística.

Fonte: dados obtidos pelo pesquisador.

Assim como já observado por Cunha (2004), os dados de hipersegmentação, independentemente da variável *tipo de escola* (pública ou privada), correspondem a um número muito menor nos dados de segmentação não-convencional, se comparados aos dados de hipossegmentações.

No que tange à comparação entre os dados de ensino público e de ensino privado, as hipersegmentações mostraram variabilidade apenas quando observada a seguinte configuração prosódica: $\sigma + \Sigma$. Essa variabilidade se dá no 2º ano e, cabe destacar, com um p-valor marginal de 0,075 que, por estar próximo do limite de aceitação é considerado como tendente a ser significativo. Nesse sentido, no que diz respeito às hipersegmentações que se formam a partir de outras combinações dos constituintes sílaba e/ou pé, os dados não mostraram a existência de variabilidade nas demais séries.

Feita a descrição dos dados, passamos, à seção adiante, à discussão desses resultados.

SEÇÃO 4: DISCUSSÃO DOS RESULTADOS

Nesta seção apresentamos a discussão dos resultados, apontando tendências que respondem aos objetivos específicos elencados nesta investigação, os quais retomamos: a- analisar prosodicamente as ocorrências de hipossegmentações e de hipersegmentações de palavras observadas nos contextos (público e privado); b-) comparar os dados, considerando as variáveis *tipo de escola* e *série* a fim de verificar a existência de variabilidade na configuração prosódica das estruturas de hipersegmentação e de hipossegmentação.

A discussão aqui posta segue a mesma organização adotada na descrição dos resultados. Primeiramente são discutidos os resultados do levantamento de ocorrências de segmentação não-convencional nos textos de alunos dos dois contextos, e, em seguida, são discutidos os resultados sobre a configuração prosódica das hipossegmentações e das hipersegmentações, separando-os conforme a sua natureza.

4.1 Sobre a ocorrência de segmentação não-convencional segundo o *tipo de escola* (pública ou privada)

No empreendimento desta pesquisa, para o fim de cumprimento do primeiro objetivo - qual seja o de analisar prosodicamente as ocorrências de hipossegmentações e de hipersegmentações de palavras observadas nos contextos (público e privado) - se fizeram necessários o levantamento dessas ocorrências em nossa amostra e a observação de como se distribuem de acordo com o *tipo de escola* e a *série*.

Quando do levantamento da quantidade de segmentações não-convencionais comparando-se os *tipos de escola*, identificamos diferença estatisticamente significativa, corroborando os resultados obtidos por Cunha (2004) e Torquette (2016) os quais já apontaram a predominância de dados de segmentação não-convencional na escrita de crianças de escolas públicas, em comparação à escrita de crianças da rede privada.

Cunha (2004) já havia demonstrado, a partir da análise de produções de alunos escolas públicas e privadas de Pelotas-RS que “crianças da escola pública apresentam uma quantidade maior de dúvidas quanto à segmentação da escrita, diferenciando-se das crianças da escola particular que apresentam uma menor incidência de casos.” (CUNHA, 2004, p. 65).

De igual modo, os resultados obtidos por Torquette (2016) mostram que as ocorrências de segmentações não-convencionais se dão em maior percentual nas escolas públicas se comparadas ao número de ocorrências de grafias segmentadas não-convencionalmente produzido no contexto de ensino privado.

Em nossos dados observamos a existência de maior instabilidade no domínio da segmentação de palavra no contexto das escolas públicas, especialmente nos anos iniciais e no ano final do ensino fundamental I – ou seja, para 1º ano, 2º ano e 5º ano. Os resultados mostram que as ocorrências de segmentações não-convencionais se dão em maior percentual nas escolas públicas e que tendem a diminuir em ambos os tipos de escolas com a progressão dos anos letivos.

Identificamos ainda, outra tendência observada por Torquette (2016), que corresponde ao fato de que a diminuição de ocorrências com a progressão dos anos letivos, em ambos os tipos de escolas, se dá de forma diferente. Em seu trabalho, a autora destacou que nas escolas públicas houve uma diminuição contínua de ocorrências com a progressão dos anos escolares; já nas escolas privadas essa queda acontece de forma acentuada do primeiro para o segundo ano, chegando a zero no terceiro ano. Em nossos dados, a mesma tendência pode ser vista: nas escolas públicas a queda no números de ocorrências de segmentação não-convencional é contínua do 1º ao 5º ano, já nas escolas privadas a queda é acentuadamente do 1º para o 2º ano.

A variabilidade constatada em nossa amostra nos dados de 2º ano, quanto ao levantamento do números de segmentações não-convencionais (média de 3,88 nas escolas privadas contra 5,91 em Pública (p-valor <0,001) e medianas 1 e 4, respectivamente) mostra que os alunos das escolas públicas mantêm uma constância no número de ocorrências de segmentação não-convencional no avanço do 1º para o 2º ano.

Nota-se que no 3º ano, entretanto, a variabilidade entre os dois contextos, visível em 1º e 2º anos já não ocorre mais nessa série. Isso significa que, embora o número de ocorrências possa ser maior, ainda, no contexto de ensino público, essa variabilidade é pequena quando comparados os contextos, e portanto, não é estatisticamente significativa na comparação das turmas de 3º ano de escolas públicas com as turmas de 3º ano de escolas privadas.

Para refletirmos sobre esse resultado, estabelecemos uma aproximação entre o que nos mostram esses dados e o que estabeleciam as diretrizes curriculares municipal e estadual vigentes à época da coleta dos textos a respeito da alfabetização dos alunos do

primeiros anos do Ensino Fundamental I. Segundo a Proposta Curricular Municipal de Marília –SP para o Ensino Fundamental (2008),

a Rede Municipal de Marília mantém como meta a alfabetização plena dos alunos até o final do 2º ano do Ensino Fundamental I, não restringindo as práticas alfabetizadoras apenas à decodificação, codificação e ortografização, mas garantindo que os alunos se apropriem das diferentes práticas de linguagem, participem da cultura letrada, construam conhecimentos e atuem com autonomia e protagonismo na vida social (p. 7).

Tomando por base essa afirmação, percebemos que, conforme as diretrizes municipais, a alfabetização plena dos alunos deveria ocorrer até o 2º ano do Ensino Fundamental I. Nesse sentido, uma vez que os documentos oficiais produzidos pelas instâncias governamentais orientam os currículos das redes pública e privada, pressupõe-se que a meta estabelecida nos documentos seja buscada por instituições escolares que integram os dois contextos. De igual modo, as Orientações Curriculares do Estado de São Paulo – EF Ciclo I (2008) estabeleciam que a apropriação do sistema de escrita alfabética deveria ocorrer nos dois primeiros anos do Ensino Fundamental I.

Sob essa perspectiva, mesmo que os conhecimentos produzidos nos diversos componentes curriculares, ao adentrarem às instituições, sejam recontextualizados de acordo com a lógica de quem dirige às instituições escolares⁶³ - portanto sejam transpostos diferentemente no contexto de ensino público e de ensino privado – seria possível dizer que a não existência de diferença estatística considerável no âmbito dos dados do 3º ano relaciona-se em alguma medida ao que é proposto nos documentos oficiais que norteiam o ensino-aprendizagem nessas séries.

Há necessidade de se considerar, entretanto, que o processo de aquisição da leitura e da escrita na escola não ocorre igualmente nos contextos de ensino público e de ensino privado, pois, como propõe a Lei de Diretrizes e Bases da Educação (2013) esse processo está fortemente relacionado aos usos sociais da escrita nos ambientes familiares de onde veem as crianças e pode demandar tempos e esforços diferenciados entre os alunos da mesma faixa etária. Nesse sentido, o esforço das instituições públicas na proposição do acesso à cultura escrita assume proporções maiores, em relação às escolas da rede privada, isso porque, como destacava o currículo paulista, ainda em 2008:

⁶³ Além do processo de didatização, os conteúdos curriculares passam a ter embutido, também, um sentido moral e político.

atualmente, as famílias que compõem a comunidade escolar da rede pública, em sua maioria, não tiveram acesso à cultura escrita. Isso não apenas torna mais complexa a tarefa da escola de ensinar seus filhos a ler e escrever, como também faz dela um dos poucos espaços sociais em que se pode intervir na busca da equidade para promover a igualdade de direitos de cidadania. E saber ler e escrever é um direito fundamental do cidadão (SÃO PAULO, 2008 p. 9).

Nesse sentido, é possível afirmar que à escola pública caberia a tarefa de propor situações de práticas sociais de uso da escrita às quais os alunos não tinham acesso antes do ingresso à vida escolar. Soares (1996) destaca, entretanto, que a escola não pode ser vista como uma redentora, isto é, como um espaço em que se possam superar as desigualdades sociais através do acesso de todos igualmente à cultura. A escola, segundo a autora deve ser comprometida com a luta contra as desigualdades e a luta para garantir às camadas populares a aquisição de conhecimentos e habilidades que as instrumentalizem para a participação no processo de transformação social. A condição socioeconômica não pode ser vista pela instituição pública como um pretexto ou um entrave no processo de ensino-aprendizagem, sob o risco de propagação de desigualdades.

Notadamente, embora possam existir contrastes entre a realidade de aplicação do currículo no que se refere ao tipo de escola (pública ou privada) os desafios em relação ao favorecimento do aluno à apropriação da escrita não são restritos às organizações escolares públicas, mas circundam a realidade da escola, enquanto instituição.

Ensinar a ler e escrever é um desafio que transcende amplamente a alfabetização em sentido estrito. O desafio que a escola enfrenta é o de incorporar todos os alunos à cultura do escrito, é o de conseguir que todos os alunos cheguem a ser membros plenos da comunidade de leitores e escritores (LERNER, 2002, p.17).

Cabe destacar, ainda, em relação às diretrizes municipais da cidade de Marília-SP, que o currículo passou, no ano de 2020, por uma atualização, resultado do interesse de alinhamento de sua proposta à nova Base Nacional Curricular Comum (2018) aprovada pelo Conselho Nacional de Educação (CNE) e homologada em 20 de dezembro de 2017. Dessa forma, em seu aporte, o documento passou a contemplar as competências gerais discriminadas pela Base Nacional Comum Curricular (BNCC), definindo, de modo

mais específico as competências e as habilidades essenciais para a formação integral do aluno.

No novo documento, denominado “Proposta Curricular: Ensino Fundamental – 1º ao 5º ano do Sistema Municipal de Ensino de Marília” diferentemente da proposta no currículo do ano de 2008, a Secretaria Municipal de Educação apresenta, ajustada à BNCC, propõe uma organização curricular dividida por Áreas do Conhecimento em que os componentes curriculares para cada ano são divididos em: Unidades Temáticas, Objetos de Conhecimento e Habilidades específicas da referida etapa de escolaridade. O que salta aos olhos, quando do exame do novo documento é que, diferentemente do documento anterior (vigente à época da coleta de dados) grande parte das habilidades previstas contemplam conhecimentos que relacionam aspectos fonético-fonológicos do português à escrita.

De um lado, o currículo do ano de 2008 trazia, dentre as doze expectativas de aprendizagem elencadas para o 1º ano, apenas 1 (uma) expectativa relacionada, mais especificamente à aquisição do código escrito: “compreender a natureza do sistema de escrita e, no mínimo produzir escrita silábica com ou sem valor sonoro convencional” (MARÍLIA, 2008, p. 4). Por outro lado, a Proposta Curricular: Ensino Fundamental – 1º ao 5º ano do Sistema Municipal de Ensino de Marília (2020) apresenta, de modo detalhado, habilidades que tocam diretamente em aspectos da relação entre fala e escrita, como as apresentadas abaixo, que se referem ao currículo do 1º ano do EF:

- a) Nomear as letras do alfabeto e recitá-lo na ordem das letras; b) Reconhecer o sistema de escrita alfabética como representação dos sons da fala; c) Segmentar oralmente palavras em sílabas. Identificar fonemas e sua representação por letras; d) Relacionar elementos sonoros (sílabas, fonemas, partes de palavras) com sua representação escrita. e) Comparar palavras, identificando semelhanças e diferenças entre sons de sílabas iniciais. f) Reconhecer a separação das palavras, na escrita, por espaços em branco (MARÍLIA, 2020, p. 90-94).

Dentre essas habilidades, uma, em especial, tange a um aspecto que o currículo anterior não contemplava de modo tão preciso: a segmentação de palavras. No caso das diretrizes atuais, há, nas orientações didáticas a respeito dessa habilidade o seguinte:

A segmentação de texto em palavras acontece pela prática da leitura e escrita, a fim de empregar corretamente a segmentação nos escritos produzidos. Assim, o trabalho envolvendo essa habilidade deve enfatizar processos que possibilitem a reflexão sobre os seguintes

aspectos: a) escrita e fala possuem critérios diferentes para segmentar as palavras, ou seja, a escrita da frase “*Vem cá comigo*” possui segmentações que não ocorrem quando a mesma é oralizada, tornando-se “*Vemcácomigo*” (MARÍLIA, 2020, p. 94).

Em relação às orientações didáticas propostas ao professor no currículo municipal de 2008, (2) duas se relacionam, de modo especial, com a escrita: 1) “O professor deve propor aos alunos que leiam e escrevam ainda que não o façam convencionalmente. Mas o fato de as escritas não convencionais serem aceitas não significa ausência de intervenção pedagógica para a construção da escrita convencional; 2) Organizar situações de aprendizagem que possibilitem a discussão e a reflexão e a sobre a escrita alfabética” (MARÍLIA, 2008, p. 6).

No que diz respeito às habilidades desenvolvidas pelo aluno, de acordo com o currículo paulista no decorrer do Ciclo I os alunos devem se tornar capazes, além de outras habilidades, de “compreender o funcionamento alfabético do sistema de escrita, ainda que escreva com erros ortográficos (ausência de marcas de nasalização, hipo e hipersegmentação, entre outros)” (SÃO PAULO, 2008, p. 9)⁶⁴.

Não pretendemos, neste trabalho, promover uma discussão aprofundada sobre as diferenças entre as duas propostas curriculares, tampouco, aqui, nos prendemos à avaliação da aplicação dos currículos, entretanto, nos parece coerente lançar o olhar para a proposta curricular municipal vigente à época da coleta de nossos dados, porque entendemos que, em alguma medida, as concepções de sujeito, de escrita, de aprendizagem atravessam as práticas de ensino. Sendo pois, os dados de escrita para os quais olhamos, resultados de práticas alicerçadas nesses documentos, não podemos nos furtar à responsabilidade de mirar nesses aspectos quando da observação da ocorrência de dados de segmentação não-convencional de palavra no âmbito nesse contexto.

Destacamos, nesse sentido, que os resultados obtidos quanto à ocorrência de segmentações não-convencionais podem resvalar em aspectos relacionados ao currículo e às distintas formas de transposição dessas propostas ao projeto político pedagógico das escolas da rede pública e privada. Essa hipótese encontra validade tanto nos momentos em que se observam diferenças estatísticas entre o ensino público e privado (no caso das

⁶⁴ Destaque-se, nessa citação as noções de hipo e hipersegmentações concebidas no rol da tradição, como “erros ortográficos”.

turmas de 1º, 2º e 5º anos), quanto no momento em que se observam aproximações (entre turmas de 3º e 4º anos dos de escolas dos dois contextos).

Consideramos importante retomar, ainda, um fator importante relacionado às relações entre o ensino público e privado no âmbito da município de Marília – SP: o índice do IDEB do ano de 2011 (ano anterior à coleta dos dados). Os índices educacionais para o referido município demonstravam que Marília continha uma média superior (6,4) se comparada à média paulista e que a média obtida pela rede de ensino privada era superior à das escolas públicas. Esses dados demonstrados na avaliação promovida pela SAEB, em alguma medida, podem se relacionar aos resultados de nossa pesquisa, quando olhamos para a ocorrência de segmentações não-convencional nos dois contextos. Na rede pública elas aparecem em maior número, sugerindo que os alunos dessa rede teriam menos domínio da convenção ortográfica comparando-se aos estudantes da rede privada. Entretanto, como discutido por Tenani (2021), uma importante problematização relacionada ao tema de desempenho dos alunos na escola diz respeito aos índices de mensuração de incidência dos erros ortográficos como parte das habilidades de escrita, avaliadas, por exemplo, pelo IDESP (através do Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo - SARESP) e pelo IDEB: a forma de avaliação dessa habilidade nesses testes, bem como a conceituação da relação entre fala e escrita que os subjaz, tem como traço desconsiderar relações entre fala e escrita, bem como práticas orais e letradas, relevantes para se interpretar a presença dos chamados erros ortográficos nos textos (TENANI, 2021, p. 21).

Nesse âmbito considerar as diferenças entre os contextos de ensino público e privado no que diz respeito à escrita tomando como ponto de partida apenas os índices do IDESP (no caso das escolas estaduais) ou do IDEB (no caso das escolas municipais) pode incorrer em interpretações atravessadas por formações ideológicas específicas e com interesses próprios⁶⁵, haja vista que os índices de desempenho se prestam, em grande medida, ao atendimento de preferências⁶⁶, sobretudo das instâncias governamentais, de

⁶⁵ Silva e Tfoufi (2009) no texto “A persistência do fracasso escolar: Desigualdade e ideologia” analisam os discursos sobre déficit e fracasso escolar, à luz da análise discursiva Pecheutiana e discutem os atravessamentos existentes nas diversas correntes ideológicas, as quais, perpassam discursos nas múltiplas instituições.

⁶⁶ A título de exemplificação cita-se a fala do secretário de Estado da Educação Paulo Renato Souza, em notícia veiculada ano de 2009, intitulada “57 escolas estaduais da região da Diretoria de Ensino de Marília, 80% das escolas avançaram no Idesp (Índice de Desenvolvimento da Educação do Estado de São Paulo) em 2009 no comparativo com 2008”: “O avanço no Idesp aponta que os programas educacionais implementados pelo governo estão corretos e já apresentam resultados”.

propagação de resultados (na maior parte das vezes positivos) das políticas educacionais bem-sucedidas implantadas em uma ou outra gestão, sem considerar, questões pertinentes à relação entre fala e escrita quando da avaliação.

Com base na argumentação até o momento proposta prosseguimos na discussão dos resultados. Adiante, discorreremos sobre os resultados da análise prosódica, primeiramente das hipossegmentações, e, posteriormente das hipersegmentações.

4.2. Hipossegmentações: configuração prosódica segundo o *tipo de escola e a série*

De acordo com o que vimos na Seção 3, os dados de hipossegmentação aparecem em maior número do que as hipersegmentações. A hipótese de que, no início da aquisição da escrita, a criança parece partir da compreensão de unidades linguísticas maiores já foi levantada em estudos como os de Abaurre (1987) e de Cagliari (2002) e seria, portanto, a justificativa para prevalência das hipossegmentações nos dados de escrita infantil.

A seguir apontamos os resultados da comparação das hipossegmentações nos dois contextos, conforme cada um dos domínios prosódicos analisados.

- **Hipossegmentações decorrentes da junção de U:**

As estruturas decorrentes da junção de U aparecem em um número muito pequeno em nossa amostra, sendo apenas 3 ocorrências, das quais duas foram produzidas no contexto de ensino público e uma no contexto de ensino privado, correspondendo a 0,24% e 0,52% respectivamente.

Em se tratando desse constituinte, convém destacar - além da inexistência de variabilidade, quando comparadas as cinco séries nos dois contextos- o fato de que, os dados encontrados estavam em textos, exclusivamente de 1º ano. Isso reforça a tese de Cunha (2004) de que as maiores instabilidades no que tange ao reconhecimento do que seja uma palavra se dão, especialmente, nas séries iniciais do EF I. Uma vez que o enunciado fonológico, como unidade prosódica, tem seus limites correspondentes aos limites de sentença, onde ocorrem as pausas, a segmentação não-convencional nos limites desse constituinte tendem a ser raros e, comumente só ocorram em dados de Ensino Fundamental I, como observado por Chacon (2004) e Tenani (2005).

- **Hipossegmentações decorrentes da junção de Frase entonacional (I)**

No que diz respeito às hipossegmentações em que é possível identificar a junção de frases entonacionais – e que são supostamente motivadas por uma linha entonacional decorrente de uma pausa – não foi observada variabilidade quando comparados os contextos público e privado em nenhuma das séries investigadas. Isso demonstra, portanto, a existência e uma certa regularidade nas hipossegmentação dessa natureza, as quais aparecem proporcionalmente nos dois contextos.

Conforme Cagliari (2002) as hipossegmentações decorrentes da junção de frase entonacional (por exemplo : “jalicotei” (já lhe contei) evidenciam a tendência da criança juntar todas as palavras usando a fala como referência primeira. Para o autor, seriam esses grupos tonais do falante, ou conjuntos de sons ditos em determinadas alturas os critérios utilizados pela criança na escolha da segmentação.

Um fato a considerar sobre as hipossegmentações dessa natureza, é que mesmo não havendo diferenças estatisticamente significativas no que tange às estruturas hipossegmentadas nos dois contextos, no âmbito do ensino privado, esse tipo de ocorrência só se dá até o 2º ano, ao passo que, no ensino público é possível identificar hipossegmentações decorrentes da junção de frase entonacional do 1º ao 5º ano.

Essa observação corrobora as discussões defendidas por Cunha (2004) de que as crianças da escola pública apresentam uma quantidade maior de dúvidas quanto à segmentação da escrita que permanece ao longo dos anos letivos, se comparadas às crianças do ensino privado.

Em contrapartida ao que observamos nos domínios de enunciado fonológico (U) e frase entonacional (I), a análise prosódica das hipossegmentações cuja configuração prosódica se dá nos domínios da frase fonológica (ϕ) e do grupo clítico (C), mostraram a existência de diferenças quando da comparação dos dados de ensino público e de ensino privado, como veremos adiante.

- **Hipossegmentações decorrentes da junção de frase fonológica (ϕ)**

Quando comparados os dados de hipossegmentação decorrentes da junção de frase fonológica, percebemos que em 1º, 2º, 3º e 5º anos não houve variabilidade, entretanto, nos dados do 4º ano se mostra uma diferença entre o público (0,306) e o privado (0,000). As hipossegmentações decorrentes da junção de frase fonológica, no 4º ano das escolas públicas corresponde a 23,5% e a 0% no ensino privado.

Se observarmos essas ocorrências de forma longitudinal, perceberemos que, no caso do ensino privado, elas têm um decréscimo significativo do 2º para o 3º ano e que, a partir do 4º ano, já não ocorrem mais nesse contexto. Já no ensino público, esse decréscimo parece ocorrer de forma gradual, diminuindo ano a ano, permanecendo, entretanto, até o 5º ano.

Esse resultado, em suma, permite dizer que:

- As crianças de 1º e 2º anos dos contextos público e privado não apresentam diferenças estatisticamente consideráveis quanto à hipossegmentação de frase fonológica, demonstrando a existência de uma regularidade na escrita dos escreventes nessas séries.
- No 3º ano, embora a diferença entre os dois contextos não seja estatisticamente significativa, uma vez que o p-valor não é indicativo de variabilidade, a ocorrência de segmentações não-convencionais dessa natureza no ensino privado é de 0,000, contra 0,857 do ensino público, indicando que, embora não haja variabilidade, no contexto de ensino público elas ainda ocorrem, ao contrário do ensino privado.
- No processo de escolarização, os alunos de escola privada deixam de apresentar hipossegmentações decorrentes da junção de frase fonológica antes do que as crianças do ensino público, uma vez que, a partir do 2º ano, os índices dessa natureza de segmentação não-convencional caem para 0,000 neste contexto, permanecendo, todavia, no contexto do ensino público.
- Considerando o decréscimo gradual da ocorrência dessa estrutura de hipossegmentação no ensino público, as análises dos textos de 5º ano não indicaram a existência de variabilidade entre os dois contextos. Isso significa que, tanto as crianças do ensino público quanto do ensino privado, nessa fase de escolarização, apresentam uma regularidade no que tange a segmentações não convencionais cuja estrutura corresponde à frase fonológica.

Como já exposto, Torquette (2016) observou - quando do levantamento quantitativo das ocorrências de segmentação não-convencional em textos do 1º ao 5º do EF I - que nas escolas públicas a diminuição de ocorrências ocorreu continuamente, ao passo nas escolas privadas essa diminuição aconteceu de forma acentuada do primeiro para o segundo ano, chegando a zero no terceiro ano. Essa mesma tendência é evidenciada

em nossos dados, visto que em se tratando especificamente de hipossegmentações de frase fonológica percebemos que, no caso do ensino privado, elas sofrem um decréscimo significativo do 2º para o 3º ano e que, a partir do 4º ano, já não ocorrem mais nesse contexto, ao passo que no ensino público elas acontecem até o 5º ano.

- **Hipossegmentações decorrentes da junção do Grupo clítico (C)**

Em relação às hipossegmentações decorrentes do grupo clítico, nossos resultados seguem a tendência já observada por Cunha (2004): as hipossegmentações decorrentes da junção de grupo clítico são a maioria, se comparadas às demais, nos dois contextos (público e privado). Em nossos dados, as hipossegmentações compostas por um clítico associado a uma palavra de conteúdo totalizavam 487 segmentações não-convencionais.

Em seu trabalho, Cunha (2004) inferiu que essa característica constatada na escrita encontra correspondência em parte do processo de aquisição do que ela define como linguagem oral. Diante disso, dois argumentos de ordem fonológica são elencados pela autora para explicar a incidência de hipossegmentações estruturadas por um grupo clítico: (1) na fala, as sílabas sozinhas sempre aparecem junto a emissões de outras duas ou mais sílabas – nesse sentido, os monossílabos tendem a se agregar a outras palavras; (2) o efeito de direcionalidade da língua, pois, no Português Brasileiro, verifica-se tendência em associar o clítico à palavra que está a sua direita, integrando a sílaba pretônica, de forma proclítica, dentro dos limites da palavra fonológica. De igual modo, Abaurre, Galves e Scarpa (1999) discutiram a correspondência entre os dados de fala e de escrita defendendo que, na aquisição da fala, as crianças que passam do balbúcio tardio às primeiras palavras têm a tendência de juntar o enunciado de uma sílaba às sílabas das palavras seguintes e isso ocorre também na aquisição da escrita, quando os monossílabos átonos tendem a ser juntados à palavra que vem adiante.

Essa tendência, destacada em Cunha (2004), a partir de Abaurre, Galves e Scarpa (1999), é detectada em nossa amostra. Embora tenhamos encontrado variabilidade em algumas séries (2º, 3º e 5º anos) no contexto do 1º ano não foi constatada variabilidade quanto ao domínio do grupo clítico quando da comparação dos contextos público e privado, demonstrando que as segmentações estruturadas por um grupo clítico seguem uma regularidade, em se tratando do início da alfabetização, que independe do contexto (público ou privado).

Em relação à variabilidade encontrada, são percebidos índices significativos na comparação dos dados de 2º, 3º e 5º anos dos dois contextos. Em nossos resultados identificamos diferença em no grupo clítico com maiores resultados na escola pública em três séries. Para os dados do 2º ano tivemos média de 1,167 contra 2,263, respectivamente, para privada e pública (p-valor = <0,001). No 5º ano, essa diferença se mostra em 0,107 para o ensino privado e 0,944 para o público (p-valor 0,002). Nos dados de 3º ano os números são de 0,000 para o privado e 1,857 para o ensino público (p-valor = 0,070) sugerindo uma diferença estatisticamente significativa, entretanto com um valor muito próximo do limite de aceitação.

Esses números, em suma, permitem concluir que:

- As hipossegmentações em cuja estrutura se detecta a junção de um grupo clítico nos dados de 2º ano são maiores no contexto de ensino público do que nos dados de ensino privado, correspondendo a 50,2% e 28,7% em termos percentuais.
- No 3º ano, a variabilidade entre os contextos público e privado se mantém, permanecendo a maior incidência das hipossegmentações estruturadas com junção de grupo clítico no ensino público. Entretanto, os números mostram que, na transição do 2º para o 3º ano, em ambos os contextos, há queda nas hipossegmentações dessa natureza.
- No 4º ano, não é verificada variabilidade entre os contextos, o que demonstra uma tendência para a estabilidade. No 5º ano, entretanto, novamente, os dados de hipossegmentação do ensino público, envolvendo o domínio do grupo clítico, passam a ser consideravelmente maiores do que no ensino privado.⁶⁷

Como já observado em outros trabalhos sobre segmentação de palavras nas séries iniciais⁶⁸, as hipossegmentações tiveram suas estruturas preponderantemente dominadas por constituintes mais altos da hierarquia. Prosseguindo em nossa discussão, adiante apresentamos os resultados obtidos na análise das hipersegmentações.

⁶⁷ As hipossegmentações encontradas nos dados de 5º ano mostram que as palavras segmentadas não-convencionalmente nessa faixa de escolaridade, não compreendem às mesmas as estruturas compostas por clítico + palavra de conteúdo tocam em aspectos ortográficos. Os casos de hipossegmentação encontrados, por exemplo, em grande medida compreendem a grafias nas quais a convenção permite (a depender do sentido) a inserção ou não do espaço em branco, se tratando, mais especificamente de palavras “novas” para os alunos, Por exemplo: “por que/ porque”; “se não/senão”. Isso foi observado, também, por Cunha (2010).

⁶⁸ Cunha (2004, 2010), Cunha e Miranda (2007; 2009).

4.3 Hipersegmentações: configuração prosódica segundo o *tipo de escola e a série*

Um fator que merece consideração nas discussões em relação à natureza das hipersegmentações levantadas nesta investigação, é o fato de que os resultados obtidos, se alinham ao que foi observado por Tenani (2021) a partir de dados de hipersegmentação do EF II: as palavras trissílabas se mostraram as formas mais segmentadas, como também apresentaram um maior conjunto de possibilidades de segmentação não-convencional, se comparadas às dissílabas e às polissílabas. Isso pode ser observado também em nossos dados: os trissílabos representam uma maioria em nossos dados, sendo 171 trissílabos hipersegmentados no ensino público (45,3%) e 27 trissílabos no ensino privado (43,5%), dentre as 377 hipersegmentações contabilizadas.

A seguir discutimos os resultados obtidos para as estruturas hipersegmentadas, levando em consideração as configurações prosódicas identificadas.

- **Hipersegmentações formadas por $\Sigma + \Sigma$**

Em seu trabalho, Chacon (2004) evidenciou que a hipersegmentação de palavras polissílabas aparece com baixa frequência nos dados de escrita infantil e tal observação foi feita por Silva (2014), também, para os dados de hipersegmentação na escrita de alunos do EF II. Em nosso *corpus*, as grafias dessa natureza correspondem a 5,03% das hipersegmentações.

Os resultados evidenciados nos trabalhos de Chacon (2004) e Silva (2014) se alinham à constatação feita em nossas análises, mais especificamente, no que diz respeito aos polissílabos segmentados em estruturas que combinam dois pés, como “colo caro”, que correspondem, em toda a nossa amostra a 11 ocorrências. Os números indicaram, em especial, que hipersegmentações dessa natureza corresponde a 2,53% no ensino público e 2,35% no ensino privado.

Nesse âmbito, foi possível observar que no que diz respeito à formação de estruturas hipersegmentadas a partir da combinação $\Sigma + \Sigma$ (por exemplo “nova mente”, “avi saro”, “constru ino”) não há indícios da existência de variabilidade quando comparados os dois contextos de ensino nas cinco séries.

Nesse sentido, é possível afirmar que hipersegmentações dessa natureza se dão proporcionalmente nos dois tipos de escola, uma vez que, como se vê no processamento dos dados, os números não indicam a existência de diferenças estatisticamente relevantes.

Convém comentar, ainda, outra regularidade observada: em ambos os contextos (público e privado) as hipersegmentações dessa natureza não são encontradas em dados de 1º ano, tendo sido observados, em ambos os contextos, a partir do 2º ano.

- **Hipersegmentações formadas por $\Sigma + \sigma$**

Na análise de hipersegmentações em textos de escrita infantil, Chacon (2004) observou que os trissílabos hipersegmentados constituíam fatos mais interessantes para a observação da prosódia se comparados às palavras dissílabas e polissílabas, isso porque, diferentemente dessas estruturas, os polissílabos são suscetíveis a diferentes possibilidades de separações não-convencionais que podem indicar princípios da oralidade, como o ritmo, já estruturado na língua, marca-se na oralidade, por exemplo, por meio de contrastes entre sílabas acentuadas e não-acentuadas” (CHACON, 2005, p. 81).

A característica rítmica se mostrou marcante nos dados de trissílabos analisados por Chacon, visto que das 136 ocorrências levantadas, 111 (81,62%) apresentaram a ruptura entre os limites de sílabas e de pés, a partir das quais: (a) 92 unidades (82,88%) delas promoveram a combinação de uma sílaba e um pé, como em “e Rita” (*irrita*); e (b) 19 unidades (17,12%) geraram a combinação contrária, ou seja, de um pé e uma sílaba, como em “colo que” (*coloque*).

Em relação às segmentações não-convencionais cuja estrutura hipersegmentada era composta pela sequência pé-sílaba (como “vonta de”, “corren do”) nossos dados vão ao encontro do que foi observado por Chacon (2004), pois, de 377 ocorrências de hipersegmentações encontradas no corpus, 38 unidades correspondem a estruturas constituídas de pé-sílaba (10,71%) ao passo que os trissílabos segmentados cuja combinação gera uma estrutura formada por sílaba-pé correspondem a 40,80% da amostra.

No que tange à comparação da configuração prosódica dessas estruturas de acordo com o tipo de escola (pública e privada) não foram encontradas diferenças estatisticamente relevantes em nenhuma das cinco séries. As segmentações constituídas pela combinação de uma sílaba e um pé corresponderam, na comparação dos tipos de escola (pública ou privada) aos seguintes valores: 10,61% das hipersegmentações produzidas nas escolas públicas e 9,58% das hipersegmentações produzidas nas escolas privadas.

- **Hipersegmentações formadas por $\sigma + \Sigma$**

Em relação às hipersegmentações dessa natureza, inicialmente, cabe comentar que, das 377 hipersegmentações contabilizadas (315 no ensino público e 62 no ensino privado) um total de 154 delas correspondiam a estruturas formadas por $\sigma + \Sigma$ (40,8 %).

Essa maioria, vai ao encontro do que foi observado por Chacon (2005) que, em análise de hipersegmentações em dados de escrita infantil constatou a preferência das crianças em separar palavras trissílabas de modo a formar a estrutura $\sigma + \Sigma$. Para o autor, nesses casos, as crianças parecem tentar plasmar em sua escrita contrastes rítmicos que certamente detectam, sobretudo, a partir do relevo prosódico do acento na fala e que, por sua recorrência, padronizam-se sob forma de constituintes prosódicos da própria língua.

Essa mesma constatação foi feita por Tenani (2021) para os dados de escrita de alunos do Ensino Fundamental II. Em seu estudo, a autora observou que a hipersegmentação dos trissílabos estaria sujeita à atuação do pé métrico e, segundo essa interpretação, grafias como “de rubo”, “na quela” seriam, prosodicamente, constituídas de sílaba átona seguida de pé métrico, resultando na configuração predominante entre os trissílabos. Nesse sentido, a autora observou que a localização de fronteira não-convencional de palavra entre trissílabos paroxítonos decorre, em parte, da configuração métrica de sílaba seguida de pé troqueu na medida em que o espaço em branco coincide com a fronteira prosódica entre sílaba e pé métrico. Além disso, a autora observou ainda, que outro fator parece incidir em algumas estruturas de hipersegmentações, que corresponde ao fato de que em casos como o “de pois” é possível a identificação de monossílabos átonos que corresponderiam a palavras gramaticais do PB. Nesse caminho, ao separar as palavras da mesma natureza de “de pois”, além da preservação de pés métricos, a criança pode estar interpretando-as como duas palavras lexicais integrantes do seu vocabulário.

No que tange à comparação entre os dados de ensino público e de ensino privado, as hipersegmentações não demonstraram diferenças nas seguintes séries: 1º, 3º, 4º e 5º ano. Entretanto, na comparação dos dados de 2º ano foi observada uma variabilidade estatisticamente relevante.

Observando a variável *série*, os resultados apontaram que hipersegmentações constituídas de $\sigma + \Sigma$, quando comparados os textos de 2º ano, ocorrem diferentemente, sendo em uma quantidade maior no ensino público: 41,5% das hipersegmentações produzidas no ensino público têm a configuração prosódica $\sigma + \Sigma$. No ensino privado esse valor corresponde a 27,05%.

Essa variabilidade se dá no 2º ano com um p-valor de 0,075 que, por estar próximo do limite de aceitação, é considerado tendente a ser significativo. Em nossos resultados, as estruturas predominantes foram as hipersegmentações formadas por sílaba $\sigma + \Sigma$ troqueu, como em [ES TAVA], [EM BORA], [ES PIRO], sendo as formações de sílaba $\sigma + \Sigma$ iambo, como [DE RUBOU], menos recorrente na amostra dos dois contextos. No ensino público as estruturas formadas por sílaba-pé troqueu corresponderam a 75,5%, e as estruturas formadas por sílaba-pé iambo corresponderam a 24,4%. No ensino privado, 69,5% e 30,5%, respectivamente.

O que se observa, nesse sentido é que embora os dados de 2º ano mostrem que as hipersegmentações que coincidem com estrutura sílaba + pé ocorram predominantemente nos textos de crianças de escola pública, a observação da configuração dessas segmentações mostra que as estruturas que preservam o troqueu silábico se dão nas mesmas proporções no ensino público e no privado (representando, nos dois contextos, a maioria das hipersegmentações formadas por sílaba + pé). Esse resultado segue uma tendência já observada em outros trabalhos⁶⁹: a maior parte das hipersegmentações ocorre, dentre outros aspectos, por influência do pé métrico portador do acento principal, em especial, o troqueu silábico.

- **Hipersegmentações formadas por $\sigma + \sigma$**

As estruturas de hipersegmentação que se dão pela separação de um dissílabo em duas partes, preservando o constituinte sílaba, contabilizaram 144 ocorrências. Das hipersegmentações dessa natureza, 120 unidades se mostraram em textos de alunos de escolas públicas e 24 unidades em textos de alunos de escolas privadas. O processamento estatístico dessas ocorrências não apontou para a existência de variabilidade significativa entre os dois contextos, quando comparados.

Convém comentar, ainda que na análise dessas estruturas foi possível verificar, assim como observado por Cunha (2004, 2010) que nos dados de hipersegmentação em geral pode-se constatar que, ao inserir um espaço dentro da palavra, a criança preserva as estruturas silábicas da língua.

Em se tratando da hipersegmentação de dissílabos, em textos de EF II Tenani (2021) observou que os dissílabos oxítonos (como por exemplo “então”) são mais

⁶⁹ Cita-se Cunha (2010) observa a influência do pé-métrico tanto em dados de segmentação não-convencional de em textos de crianças brasileiras quanto de crianças portuguesas.

hipersegmentados do que os dissílabos paroxítonos (como por exemplo “antes”)⁷⁰. Além disso, na avaliação de estruturas dessa natureza, os resultados obtidos por Tenani (2021) mostram que, no que diz respeito aos dissílabos oxítonos a sílaba átona dessas palavras corresponde a um clítico prosódico e a sílaba tônica, a palavras da língua, ao passo que, no que diz respeito aos dissílabos paroxítonos, o exame das hipersegmentação levou à constatação de esses têm motivação não na configuração métrica da palavra, mas na identificação de monossílabos como palavras ortográficas. Nessa interpretação, conforme a autora, uma, se não as duas sílabas, corresponde a palavras da língua: a sílaba átona do dissílabo tende a corresponder a um monossílabo que pertence ao conjunto de palavras gramaticais (preposição, artigo, conjunção), como “a sim” (“assim”); a sílaba tônica, a uma palavra lexical monossilábica, como “tem tei” (“tentei”).

Um fator a se considerar é o fato de que os resultados obtidos por Tenani (2021) para o EF II, no tocante à hipersegmentação dos dissílabos podem ser observados também em nossos dados: por meio da análise da configuração métrica de palavra percebe-se que característica principal da hipersegmentação é a segmentação não-convencional de sílaba quando essas podem corresponder a clíticos prosódicos monossilábicos.

Essa afirmação encontra fundamento a partir da predominância de dados dessa natureza em nossa amostra, como vemos abaixo:

“se não” – 52 ocorrências
“por que” – 14 ocorrências
“da qui” – 14 ocorrências
“em tão” – 6 ocorrências
“a sim” – 6 ocorrências
“por co” – 3 ocorrências
“sem pre” – 3 ocorrências
que brou – 3 ocorrências
“de poi” - 3 ocorrências
“com tou”- 3 ocorrências
“a chou” – 3 ocorrências
“o tro” – 3 ocorrências

As hipersegmentações “se não” e “por que”, encontradas nos dados de 4º e 5º em ambos os tipos de escola, anos sinalizam duas possibilidades gráficas na convenção

⁷⁰ Convém esclarecer que em PB há mais dissílabos paroxítonos do que dissílabos oxítonos e, segundo Tenani (2021) este fato leva à suposição de que a maior porcentagem de hipersegmentação de dissílabos oxítonos não seja necessariamente decorrente de um maior número de palavras desse tipo na língua.

ortográfica. Nesse sentido, a flutuação gráfica entre as formas “se não/senão” e “por que/porque” e sua predominância entre os dados de hipersegmentação pode ser explicada em função do fato de que tais palavras integram um conjunto específico de palavras do PB que geram dúvidas⁷¹ não apenas nos escreventes de EF I ou EF II, mas também em escreventes adultos e está relacionada ao domínio da convenção.

- **Hipersegmentações formadas por $\sigma + \sigma + \Sigma$**

Estruturas compostas por $\sigma + \sigma + \Sigma$, contabilizaram 2 unidades dentre as hipersegmentações levantadas, todas em 2º ano do ensino público, são elas “a com teceu” e “a cre dito”, não tendo sido possível, portanto, verificar a existência de variabilidade no processamento estatístico. A baixa frequência desse tipo de dado corrobora os achados da literatura em outras pesquisas sobre a escrita infantil.

- **Hipersegmentações formadas por $\sigma + \sigma + \sigma$**

Os dados dessa categoria, assim como as hipersegmentações citadas anteriormente se deram em um número muito pequeno dentro dos dados encontrados nesta pesquisa, contabilizando, neste caso, três ocorrências, todas no ensino público (0,95%). Essa baixa frequência de ocorrência de dados dessa natureza foi observada, também, por Cunha 2004, que identificou apenas três ocorrências em sua amostra.

Os dados obtidos através da análise prosódica das hipersegmentações evidenciaram que:

- embora as hipersegmentações se deem em maior número no contexto do ensino público, não há diferenças estatisticamente significativas em relação ao contexto privado. Ademais a distribuição de dados dessa natureza ocorre na mesma proporção em ambos os contextos;
- os escreventes tendem a apresentar regularidade, no sentido de demonstrarem a mesma percepção dos domínios prosódicos observados, independente do contexto.

⁷¹ Um estudo detalhado do complexo funcionamento de “senão” encontra-se em Tenani e Longhin-Thomazi (2015). No exame das ocorrências de “senão” nos textos do EF II, as autoras identificam quase absoluta preferência pela segmentação de “se não” (opção marcada) e a análise de seus usos mostra um padrão polissêmico muito peculiar, com uma multiplicidade de sentidos.

Tendo em vista os resultados aqui expostos, passamos, na seção que segue, às conclusões a que chegamos a partir do desenvolvimento desta investigação.

5. CONCLUSÕES

Como já apontado ao longo desta redação, a presente investigação buscou, de modo mais amplo, investigar a configuração prosódica de ocorrências de segmentações não-convencionais de palavras em textos produzidos por crianças do 1º ao 5º ano de ensino fundamental de escolas públicas e de escolas privadas, motivada por resultados de pesquisas⁷² que apontavam para a existência de contrastes de desempenho, quando comparados os rendimentos escolares de alunos de contexto público e de contexto privado⁷³; e, de modo especial por pesquisas que evidenciaram as diferenças quantitativas na segmentação de palavras quando consideradas as variáveis *tipo de escola* e *série*.

Para o empreendimento desta investigação, tendo em vista a pesquisa quantitativa realizada por Torquette (2016) no mesmo município, construímos a hipótese de que, em alguma medida, quando analisados qualitativamente os dados de segmentação não-convencional (mais especificamente os dados de hipossegmentação e de hipersegmentação) estes poderiam fornecer indícios de uma variabilidade não apenas no que diz respeito à distribuição das ocorrências segundo o *tipo de escola* e a *série* como já evidenciado por Torquette (2016), mas também no sentido de como se configuram prosodicamente essas ocorrências quando comparados os contextos (público e privado) por *séries*.

Nesse rol, nortearam nossa investigação os seguintes questionamentos:

- À luz do modelo fonológico da Fonologia Prosódica, como se configuram prosodicamente as hipossegmentações e as hipersegmentações produzidas por crianças do 1º ao 5º ano do ensino público? E do ensino privado?
- Haveria alguma variabilidade na configuração prosódica das segmentações não-convencionais quando comparados os dados, considerando-se as variáveis *tipo de escola* e *série*.

Para responder a eles, estabelecemos dois objetivos específicos:

⁷² Cunha (2004; Torquette (2006).

⁷³ Sampaio e Guimarães (2009).

- a-) analisar prosodicamente as ocorrências de hipossegmentações e de hipersegmentações de palavras observadas nos dois contextos (público e privado).
- b-) comparar os dados, considerando as variáveis *tipo de escola* e *série* a fim de verificar a existência de variabilidade na configuração prosódica nessas estruturas.

Antes de passarmos às conclusões as quais chegamos, cabe uma ressalva importante, no que concerne a aspectos de ordem metodológica que atravessaram o desenvolvimento desta investigação. Em um primeiro momento, quando do levantamento das ocorrências, observamos que, uma parte significativa das segmentações não-convencionais contabilizadas (288 ocorrências no ensino público e 37 no ensino privado) não eram passíveis de explicação a partir de aspectos fonético-fonológicos, por apresentarem estruturas que não coincidiam com domínios prosódicos da língua. Essas estruturas, por razão de nosso interesse, não foram analisadas nesta investigação. Entretanto, reafirmamos que um olhar para esses dados, sob uma perspectiva que os trata como indícios dos diferentes trânsitos dos escreventes por práticas sociais orais/letradas e faladas/escritas, tal como proposto por Corrêa (2001), pode vir a oferecer resultados interessantes sobre como essas relações atravessam os alunos de contextos público e de contexto privado.

Nesse aspecto, partilhamos da perspectiva teórica que defende a escrita como heterogênea em sua constituição, sobretudo porque essa postulação parte, por um lado, do interesse de superar as abordagens que tinham na relação entre o falado e o escrito um processo de interferência, ressaltando a heterogeneidade da escrita a partir das representações que o escrevente constrói sobre a escrita, sobre o interlocutor e sobre si mesmo. Por outro, porque a partir dessa concepção, Corrêa (2001) intencionava afastar-se das “[...] avaliações estereotipadas sobre a escrita que tomam como modelo abstrato – literário ou não – de boa escrita” (CORRÊA, 2004, p. 86).

Esse olhar, tal como proposto por Corrêa (2004) aponta para o fato de que as relações entre letramento/oralidade e fala/escrita são fenômenos de fala e escrita enquanto relações entre fatos linguísticos (relações fala vs. escrita) e enquanto relação entre práticas sociais (oralidade e letramento) (CORRÊA, 2004, p. 2). Nesse caminho, não haveria, como defende o autor, um material essencialmente escrito ou essencialmente falado, mas, heterogêneo, isto é, constitutivamente híbrido, em que fala e escrita se imbricam. Em

outras palavras, para Corrêa (2004), os textos escritos – que constituem nossa amostra – apresentam marcas linguísticas que funcionam como pistas sobre o imaginário do escrevente acerca da constituição heterogênea da escrita. No âmbito dessas proposições reiteramos que olhar que lançamos a nossos dados, nesta pesquisa, embora tenha se inclinado a questões de ordem prosódica, reconhecemos que nas segmentações não-convencionais produzidas nos dois contextos não estão implicadas questões exclusivamente de ordem linguística, mas também no que concerne às práticas de letramento do alunos.

Corroboramos, ainda, com o que já foi destacado em alguns estudos (como por exemplo Sampaio e Guimarães (2009)⁷⁴ com relação a alguns fatores serem determinantes no desempenho escolar, tais como o estabelecimento de ensino, a qualidade dos professores, a disponibilidade de laboratórios de qualidade, a disponibilidade de material de qualidade, a organização da estrutura escolar. Todos estes aspectos (mas não só eles) exercem influência sobre o desempenho do estudante. Com vistas a esse fato, apresentamos, adiante as conclusões alcançadas nesta pesquisa:

5.1 Respondendo ao objetivo geral:

- Observar a configuração prosódica de ocorrências de segmentações não-convencionais de palavras em textos produzidos por crianças do 1º ao 5º ano de ensino fundamental de escolas públicas e de escolas privadas.

O levantamento das ocorrências de segmentação não-convencional de palavra corroborou tendências já observadas em outros trabalhos:

- a) Nos textos de ensino fundamental I de escolas públicas o número de ocorrências de segmentação não-convencional de palavras é maior, se comparado ao número de ocorrências observadas no contexto do ensino-público, corroborando os achados de Abaurre (1987), Cagliari (2002), Cunha (2004), Torquette (2016).

⁷⁴ Trabalho cujo objetivo central era o de discutir as diferenças entre o ensino público e o ensino privado no Brasil.

- b) O número de hipossegmentações encontradas nos textos, independentemente do contexto de ensino, tende a ser superior ao que o número de hipersegmentações - (Chacon, 2004; Cunha (2004); Capristano (2007); Torquette (2016).
- c) Nos dados de segmentação não-convencional de Ensino Fundamental I é possível observar a atuação de todos os constituintes da hierarquia prosódica, como destacado por Capristano (2007), Chacon (2011) diferentemente do que ocorre nos dados de EF II e demonstrado em Tenani (2007). No caso dos dados do EF I, cabe ressaltar que as hipossegmentações se dão nos níveis mais altos da hierarquia (U, I, ϕ , C) ao passo que, nas hipersegmentações se dão nos constituintes mais baixos.

5.2 Respondendo aos objetivos específicos

Neste momento apresentamos as conclusões a que chegamos a partir da cumprimento dos objetivos específicos relacionados: a) análise prosódica das ocorrências de hipossegmentações e de hipersegmentações de palavras observadas nos dois contextos (público e privado) b-) comparação os dados, considerando as variáveis *tipo de escola* e *série* a fim de verificar a existência de variabilidade na configuração prosódica nessas estruturas.

Sobre as hipossegmentações decorrentes da junção de U:

- Não foi detectada variabilidade nas estruturas com essa configuração quando comparadas as cinco séries nos dois contextos, observamos que os dados encontrados só constavam em textos de 1º ano. Isso reforça a tese de Cunha (2004) de que as maiores instabilidades no que tange ao reconhecimento do que seja uma palavra se dão, especialmente, nas séries iniciais do EF I. Uma vez que o enunciado fonológico, como unidade prosódica, tem seus limites correspondentes aos limites de sentença, onde ocorrem as pausas, a segmentação não-convencional nos limites desse constituinte tendem a ser raros e, comumente só ocorram em dados de Ensino Fundamental I, como observado por Chacon (2004) e Tenani (2005).

Sobre as hipossegmentações decorrentes da junção I:

- Não há variabilidade nas hipossegmentações com essa natureza quando da comparação dos contextos nas cinco séries. Um fato a considerar é que mesmo não havendo diferenças estatisticamente significativas no que tange às estruturas hipossegmentadas nos dois contextos, no âmbito do ensino privado, esse tipo de ocorrência só se dá até o 2º ano, ao passo que, no ensino público é possível identificar hipossegmentações decorrentes da junção de frase entonacional do 1º ao 5º ano.

Sobre as hipossegmentações decorrentes da junção de ϕ :

- Foram encontradas evidências de variabilidade quando comparadas as séries dos dois contextos. Concluimos, a partir de nossos resultados que, no processo de escolarização, os alunos de escola privada deixam de apresentar hipossegmentações decorrentes da junção de frase fonológica antes do que as crianças do ensino público, uma vez que, a partir do 2º ano, os índices dessa natureza de segmentação não-convencional sofrem um decréscimo significativo nesse contexto (e caem para 0 no 4º e 5º ano do ensino privado) ao passo que no ensino público as hipossegmentações dessa natureza ocorrem continuamente do 1º ao 5º ano.

Sobre as hipossegmentações decorrentes da junção de C:

- Foram encontrados indícios de variabilidade quando comparados os contextos nas cinco séries. Em relação às hipossegmentações decorrentes do grupo clítico, nossos resultados mostraram que as hipossegmentações em cuja estrutura se detecta a junção de um grupo clítico são maiores no contexto de ensino público do que nos dados de ensino privado. As diferenças se dão em maiores proporções em 2º, 3º e 5º anos, mostrando que esse tipo de ocorrência predomina dentre as hipossegmentações feitas pelos alunos do ensino público. Cabe destaque o fato de que hipossegmentações dessa natureza, no ensino privado só ultrapassaram a marca de 50% dos dados no 1º ano. No ensino público, entretanto, além de a frequência desse tipo de ocorrência ser regular ao

longo dos cinco anos, no 5º ano, mais especificamente, as hipossegmentações com essa estrutura chegam a 58,7%. Concluimos então, que mesmo que apresentem uma redução relacionada à progressão escolar, a maior dificuldade dos alunos de escola pública permanece no domínio do grupo clítico. Isso não ocorre, entretanto, na rede privada. Seguindo à tendência já observada por Cunha (2004): as hipossegmentações decorrentes da junção de grupo clítico são a maioria, se comparadas às demais, nos dois contextos (público e privado). Em nossos dados, as hipossegmentações compostas por um clítico associado a uma palavra de conteúdo totalizavam 487 segmentações não-convencionais.

Sobre as hipersegmentações formadas por $\Sigma + \Sigma$

- Não há indícios da existência de variabilidade quando comparados os dois contextos de ensino nas cinco séries. Convém comentar, ainda, outra regularidade observada: em ambos os contextos (público e privado) as hipersegmentações dessa natureza não são encontradas em dados de 1º ano, tendo sido observados, em ambos os contextos, a partir do 2º ano.

Sobre as hipersegmentações formadas por $\Sigma + \sigma$

- A partir de nossas análises concluimos que não foram encontradas diferenças estatisticamente relevantes em nenhuma das cinco séries, quando comparados essas estruturas de acordo com o tipo de escola (pública e privada). As segmentações constituídas pela combinação de um pé e uma sílaba corresponderam a 10,61% das hipersegmentações produzidas nas escolas públicas e 9,58% das hipersegmentações produzidas nas escolas privadas.

Sobre as hipersegmentações formadas por $\sigma + \Sigma$

Na comparação dos dados foi observada uma variabilidade estatisticamente relevante quando considerada a variável *série*. Os dados apontaram que as hipersegmentações que levam à uma estrutura formada por sílaba-pé,

correspondem a 41,58 % das segmentações produzidas no ensino público, no ensino privado esse número corresponde a 23,05%. No ensino público as estruturas formadas por sílaba-pé troqueou corresponderam a 75,5%, ao passo que as estruturas formadas por sílaba-pé iambo corresponderam a 24,4%. No ensino privado, 69,5% e 30,5%, respectivamente. Concluimos, portanto, que, embora os dados de 2º ano evidenciem proporções maiores desse tipo de hipersegmentação no ensino público, as estruturas que preservam o troqueou silábico se dão nas mesmas proporções no ensino público e no privado (representando, nos dois contextos, a maioria das hipersegmentações formadas por sílaba + pé).

Sobre as hipersegmentações formadas por $\sigma + \sigma$

- Não foram encontradas evidências de variabilidade na hipersegmentações dessa natureza comparados os dois contextos nas cinco séries. Concluimos, portanto que não diferenças significativas entre as hipersegmentações formadas pela estrutura sílaba-sílaba. Essas ocorrências, entretanto, apareceram em um número expressivo, sendo, 120 unidades em textos de alunos de escolas públicas e 24 unidades em textos de alunos de escolas privadas.

Sobre as hipersegmentações formadas por $\sigma + \sigma + \Sigma$

- A análise dessas estruturas não demonstrou a existência de variabilidade. Estruturas compostas por $\sigma + \sigma + \Sigma$, contabilizaram 2 unidades dentre as hipersegmentações levantadas, ambas em 2º ano do ensino público, são elas “a com teceu” e “a cre dito.

Sobre as hipersegmentações formadas por $\sigma + \sigma + \sigma$

- Os dados dessa categoria, assim como as hipersegmentações citadas anteriormente se deram em um número muito pequeno dentro dos dados encontrados nesta pesquisa, contabilizando, neste caso, três ocorrências, todas

no ensino público (0,95%). Nesse sentido, não houve variabilidade estatística significativa quando analisados dados dessa natureza.

5.3 Últimas considerações

De posse desses resultados, apresentados individualmente, conforme a categoria, podemos, por fim, sintetizá-los, retomando a nossa hipótese de que se as variáveis *tipo de escola* (pública ou privada) e *séries* mostram diferenças quando comparada a distribuição das ocorrências de segmentações não-convencionais nos textos de Ensino Fundamental I no município de Marília- SP, tal como demonstrado por Torquette (2016), a análise da configuração prosódica das segmentações poderia também apontar a existência de variabilidade quando consideradas essas mesmas variáveis (*tipo de escola e séries*).

A conclusão mais geral a que chegamos, portanto, compreende o fato de que, como hipotetizado, há diferenças quando comparados os dados de hipossegmentação e de hipersegmentação. Entretanto essa diferença se dá pontualmente e quando envolve estruturas e constituintes prosódicos específicos.

No caso das hipossegmentações, a percepção mais geral é a de que os alunos de escola pública apresentam mais dificuldade na segmentação de palavras quando estão sendo mobilizados os domínios de frase fonológica mas, sobretudo, o domínio do grupo clítico. Os dados mostram que as dificuldades, no que diz respeito a esses domínios se mantém, mesmo com a progressão nas séries, porque ao final do 5º ano, essas ocorrências permanecem na escrita dos alunos do contexto público. Em contrapartida, no caso do ensino privado, essas hipossegmentações, não são visíveis nos textos dos alunos de 5º ano.

Como evidenciado por Abaurre (1991), Cunha (2004), Torquette (2016), em nossos dados os números de segmentação não-convencional passam por uma queda com o avanço nas séries nos dois contextos, todavia, no que diz respeito à natureza desses dados, vemos que, no caso do ensino público, em especial no que tange ao grupo clítico há uma regularidade que permanece ao longo dos anos. Isso vai ao encontro ao fato observado por Torquette (2016) de que o decréscimo nas ocorrências de segmentação não-convencional no ensino público ocorre de forma gradual, ao passo que no ensino privado esse decréscimo ocorre acentuadamente nos anos iniciais do EF. Esses achados corroboram com Abaurre (1991) quando observou que os alunos de escolas particulares,

talvez controlem melhor (e mais cedo) os critérios de segmentação na escrita espontânea do que as crianças de classe baixa, que frequentam escola pública. (...) isso aparece como uma tendência geral, o que sem dúvida merece uma pesquisa mais aprofundada (ABAURRE, 1991, p. 211).

No que diz respeito às hipersegmentações, a variabilidade hipotetizada inicialmente se mostra, também, de forma pontual (apenas no 2º ano) e quando a estrutura segmentada é formada pela combinação de uma sílaba + pé. Os dados evidenciam que as grafias dessa natureza predominam entre os dados de hipersegmentação produzidos pelos alunos o ensino público e, no caso do 2º ano com uma variabilidade estatisticamente significativa, demonstrando a preferência dos alunos de escolas públicas por esse tipo de estrutura. Entretanto, a partir de um olhar mais apurado percebemos que, embora os dados evidenciem proporções maiores desse tipo de hipersegmentação no ensino público, as estruturas que preservam o troque silábico se dão nas mesmas proporções no ensino público e no privado (representando, nos dois contextos, a maioria das hipersegmentações formadas por sílaba + pé). Ademais, a análise das hipersegmentações não revelou disparidades quando consideradas as variáveis *tipo de escola* e *séries*.

Por tudo o que foi exposto, esperamos, a partir deste trabalho contribuir com discussões nas mais diversas áreas (educacional, linguística) que porventura tenham como objeto a escrita infantil, demonstrando, sobretudo, que as grafias não-convencionais, tradicionalmente analisadas como “erros” são, na verdade, importantes ferramentas para entendermos como se dá a relação entre fala e escrita.

De igual modo acreditamos que este trabalho têm alcances significativos no sentido de auxiliar na desconstrução de uma visão estereotipada da escola pública, perpetuada no discurso do senso comum, de que haveria uma disparidade profunda entre esses contextos quando se trata da alfabetização. O que vemos, entretanto é que existência de contrastes, no que toca exclusivamente às segmentações não-convencionais de palavra, se revela mais expressivamente quando observada a ocorrência desses dados nos dois contextos de ensino, todavia, em se considerados em relação à prosódia da língua, vemos que a variabilidade se mostra menos expressiva.

Essas afirmações, contudo, têm validade quando considerada a situação de produção específica desses textos, isto é, um momento anterior à Pandemia instaurada pela COVID-19, no ano de 2020. Estudos seguem demonstrando ⁷⁵ que o isolamento

⁷⁵ Azevedo et al. (2020); apresentam simulações que indicam quedas no nível educacional, no aprendizado e nos futuros rendimentos de estudantes ao redor do mundo.

social (e o consequente fechamento das escolas) afetou de maneira mais intensa grupos socioeconômicos mais vulneráveis, levando, portanto, ao aumento das desigualdades, uma vez que, como destacado por Cavalcante, Komatsu e Filho (2020) com o fechamento das escolas, a maior parte do ensino precisou se materializar nas residências, com condições que não eram as mesmas para os alunos, com diferenças no acesso à internet, no ambiente de estudo, no acesso a livros e materiais escolares, etc.

Além das complexas questões pedagógicas relativas ao ensino remoto, a discussão envolve o tema da infraestrutura, das condições sociais e de saúde de toda a comunidade (GESTRADO, 2020). Dessa forma os efeitos no rendimento de estudantes ao redor do mundo é algo que pode levar anos para aparecer, todavia, neste ano, após a retomada das atividades de ensino presenciais, as primeiras consequências já começam a se mostrar e de fato, indiciam que “seja por problemas de infraestrutura ou de adaptação ao sistema de ensino à distância, os alunos pertencentes a grupos socioeconômicos mais vulneráveis tenham sentido mais fortemente os efeitos da Pandemia sobre sua vida escolar. Essa questão a nosso ver, não pode ser ignorada nas pesquisas de âmbito educacional, sobretudo àquelas que tocam nos aspectos de ordem socioeconômica relacionados à educação. Nesse quadro, entendemos que a investigação aqui empreendida pode estabelecer importantes diálogos com pesquisas sobre a escrita infantil desenvolvidas a partir deste momento, porque, como destacado, a Pandemia transformou profundamente a realidade educacional em diversos lugares do mundo.

REFERÊNCIAS

- ABAURRE, Maria Bernadete Marques. O que revelam os textos espontâneos sobre a representação que faz a criança do objeto escrito? In: KATO, Mary (Org.). **A concepção da escrita pela criança**. 2. ed. Campinas: Pontes Editores, 1988a, p. 135-142.
- ABAURRE, Maria Bernadete Marques. A relevância dos critérios prosódicos e semânticos na elaboração de hipóteses sobre segmentação na escrita inicial. **Boletim da ABRALIN**, Campinas, v. 11, 1991c, p. 203-17.
- ABAURRE, Maria Bernadete Marques. (Re)escrevendo: o que muda? In: ABAURRE, Maria Bernadete Marques; FIAD, Raquel Salek; MAYRINK-SABINSON, Maria Laura. **Cenas de aquisição da escrita: o sujeito e o trabalho com o texto**. Campinas: Mercado das Letras, 1997.
- ABAURRE, Maria Bernadete Marques. Posfácio: A aquisição da escrita do português – considerações sobre diferentes perspectivas de análise. In: ROJO, Roxane. **Alfabetização e letramento: perspectivas linguísticas**. São Paulo: Mercado das Letras, 1998, p. 205-232.
- ABAURRE, Maria Bernadete Marques; CAGLIARI, Luiz Carlos. **Textos espontâneos na primeira série: evidências da utilização, pela criança, de sua percepção fonética da fala para representar e segmentar a escrita**. Cadernos CEDES: Antropologia e Educação Interfaces do Ensino e da Pesquisa, São Paulo, v. 14, 1985, p. 25-29.
- ABAURRE, Maria Bernadete Marques; GALVES, Charlotte. Os clíticos no português brasileiro: uma abordagem sintáticofonológica. In: CASTILHO, Ataliba Teixeira de; BASÍLIO, Margarida. (Org.) **Gramática do português falado: estudos descritivos**. Vol. IV. Campinas: UNICAMP, 1996, p. 273-320.
- BISOL, Leda. Constituintes prosódicos. In: BISOL, Leda. **Introdução a estudos de fonologia do português brasileiro**. Porto Alegre: EDIPUCRS, 1996, p. 247-261.
- BISOL, Leda. **Mattoso Câmara Jr. e a palavra prosódica**. DELTA, São Paulo, v. 20, n. especial, 2004, p. 59-70.
- BISOL, Leda. **O clítico e seu status prosódico**. Revista de Estudos da Linguagem, Belo Horizonte, v. 9, n.1, 2000, p. 5-20.
- BISOL, Leda. **O clítico e o seu hospedeiro**. Letras de Hoje: Porto Alegre, v.40, n. 3, 2005, p. 163.
- CÂMARA JR., Joaquim Mattoso. **Estrutura da Língua Portuguesa**. Petrópolis, Rio de Janeiro: Vozes, 1970.

CAPRISTANO, Cristiane Carneiro. **Aspectos de segmentação na escrita infantil**. São José do Rio Preto. 166f. Dissertação (Mestrado em Estudos Linguísticos). Instituto de Biociências, Letras e Ciências Exatas, UNESP, 2003.

CAPRISTANO, Cristiane Carneiro. **A propósito da escrita infantil: uma reflexão sobre segmentações não-convencionais**. Letras de hoje: Porto Alegre. vol. 39, nº. 3, 2004, p. 245-260.

CAPRISTANO, Cristiane Carneiro. **Mudanças na trajetória da criança em direção à palavra escrita. Campinas**. 253f. Tese (Doutorado em Linguística Aplicada). Universidade Estadual de Campinas, 2007b.

CAPRISTANO, Cristiane Carneiro. **Por uma concepção heterogênea da escrita que se produz e se ensina na escola**. Cadernos de educação (UFPel). Pelotas, v. 35, 2010, p. 171-193.

CAVALCANTE, Vitor; KOMATSU, Bruno; FILHO, Naercio Menezes. **Desigualdades Educacionais durante a Pandemia**. Revista Policy Paper | Nº 51 Dezembro, 2020.

CHACON, Lourenço. **Ritmo da escrita: uma organização do heterogêneo da linguagem**. São Paulo: Martins Fontes, 1998.

CHACON, Lourenço. **Constituintes prosódicos e letramento em segmentações não-convencionais**. Letras de hoje: Porto Alegre. v. 39, nº. 3, 2004, p. 223-232.

CHACON, Lourenço. **Hipersegmentações na escrita infantil: entrelaçamentos de práticas de oralidade e letramento**. Estudos linguísticos. Campinas, 2005, v. 34, p.77-86.

CHACON, Lourenço. Prosódia e letramento em hipersegmentações: reflexões sobre a aquisição da noção de palavra. In: CORRÊA, Manoel Luiz Gonçalves. (Org.) **Ensino de língua: representação e letramento**. Campinas: Mercado de Letras, 2006, p.155-167.

COLLISCHONN, Gisela. **Acento secundário em português**. Letras de hoje: Porto Alegre, vol. 29, n. 4, 1994, p. 43-53.

CORRÊA, Manoel Luiz Gonçalves. **O modo heterogêneo de constituição da escrita. Campinas**. 422f. Tese (Doutorado em Linguística). Universidade Estadual de Campinas, 1997.

CORRÊA, Manoel Luiz Gonçalves. Letramento e heterogeneidade da escrita no ensino de português. In: SIGNORINI, Inês. (Org.) **Investigando a relação oral/escrito e as teorias do letramento**. Campinas: Mercado das Letras, 2001, p. 135-166.

CORRÊA, Manoel Luiz Gonçalves. **O modo heterogêneo de constituição da escrita**. São Paulo: Martins Fontes, 2004.

CUNHA, Ana Paula Nobre. **A hipo e a hipersegmentação nos dados de aquisição da escrita: um estudo sobre a influência da prosódia**. Pelotas. 132f. Dissertação (Mestrado em Educação) Universidade Federal de Pelotas, 2004.

CUNHA, Ana Paula Nobre. **As segmentações não-convencionais da escrita inicial: uma discussão sobre o ritmo linguístico do português brasileiro e europeu.** Pelotas. 188f. Tese (Doutorado em Educação). Universidade Federal de Pelotas, 2010.

CUNHA, Ana Paula Nobre; MIRANDA, Ana Ruth Moresco. **Indícios de reestruturação do conhecimento fonológico da criança em dados de reparo na escrita inicial.** Letras de Hoje: Porto Alegre. 2013, p. 343-354.

DELECRODE, C. R. (2002) **Uma abordagem linguística da construção da escrita por estudantes do ensino fundamental.** Marília, Relatório PIBIC.

DONADEL, Gabriela. **Breve reflexão acerca da escrita: a natureza das segmentações não convencionais.** Cadernos do Instituto de Letras. Porto Alegre, nº 42, junho de 2011. p. 201- 219.

FERREIRA, Carmen Regina Gonçalves. **Um estudo sobre a segmentação não-convencional na aquisição da escrita de alunos de EJA.** Pelotas. 163f. Dissertação (Mestrado em Educação). Universidade Federal de Pelotas, 2011.

FERREIRO, E.; TEBEROSKY, A. **Psicogênese da língua escrita.** Porto Alegre: Artmed, 1999.

GESTRADO; **Trabalho docente em tempos de pandemia.** Relatório técnico CNTE/CONTEE. Grupo de Estudos sobre política educacional e trabalho docente: UFMG: 2020.

KLEIMAN, Ângela. Modelos de letramento e as práticas de alfabetização na escola. In: (Org.) **Os significados do letramento: uma nova perspectiva sobre a prática social da escrita.** Campinas: Mercado de Letras, 1995, p. 15-64.

KOCH, Ingedore. Interferências da oralidade na aquisição da escrita. Campinas. **Trabalhos em Linguística Aplicada**, vol. 30, 1997, p. 31-38.

LEMOS, Claudia. Sobre a aquisição da escrita: algumas questões. In: ROJO, Roxane (Org.). **Alfabetização e letramento: perspectivas linguísticas.** São Paulo: Mercado de Letras, 1998, p. 13-32.

LERNER, Délia. É possível ler na escola? In: LERNER, D. **Ler e escrever na escola: o real, o possível e o necessário.** Porto Alegre: Artmed, 2002

MARCUSCHI, Luiz Antônio. Oralidade e letramento. In: _____. **Da fala para a escrita: atividades de retextualização.** São Paulo: Cortez, 2001, p. 15-43.

MARÍLIA. Secretaria Municipal da Educação. **Proposta Curricular Municipal para o 1º ano do Ensino Fundamental.** Marília, 2008.

MARÍLIA. **Proposta Curricular: Ensino Fundamental – 1º ao 5º ano do Sistema Municipal de Ensino de Marília.** Marília: Secretaria Municipal da Educação, 2020, Volume Único.

MIRANDA, Ana Ruth Moresco. Informação fonológica na aquisição da escrita. In: TENANI, Luciani *et. al.* (Org) **Estudos linguísticos contemporâneos: diferentes olhares**. São Paulo: Cultura Acadêmica, 2013, p. 11-36.

MIRANDA, Ana Ruth Moresco; MATZENAUER, Carmen. Aquisição da fala e da escrita: relações com a fonologia. **Cadernos de educação (UFPEL)**. Pelotas, v. 35, 2010, p. 359-405.

NESPOR, Marina; VOGEL, Irene. **Prosodic phonology**. Dordrecht: Foris Publications, 1986.

PAULA, Isis Fernanda Vicente de. **Movimentos na escrita inicial de crianças: um estudo longitudinal de hipersegmentações**. São José do Rio Preto. 154f. Dissertação (Mestrado em Estudos Linguísticos). Universidade Estadual Paulista, 2007.

PARANHOS, Fabiana Cristina. A segmentação não-convencional de palavras em textos de alunos de quinta série do Ensino Fundamental. **Relatório final de pesquisa de iniciação científica**. Inédito. FAPESP: Proc. 2009/14491-0. 2010.

PARANHOS, Fabiana Cristina; TENANI, Luciani. Análise prosódica de segmentações não-convencionais de palavras em textos do sexto ano do Ensino Fundamental. **Filologia e Linguística Portuguesa**, v.13, n.2, 2011, p. 477-504.

RAMOS, Gessica. SILVA, Laís. **Programa Paulista Ler e Escrever e o trabalho docente: entre orientações e padronizações**. Revista Práxis Educacional, Vitória da Conquista–Bahia –Brasil, v. 16, n. 40, p. 554-572, jul./set. 2020.

SÃO PAULO. Secretaria do Estado da Educação. **Orientações Curriculares do Estado de São Paulo: Língua Portuguesa e Matemática Ciclo I**. São Paulo: FDE, 2008. Disponível em: http://www.escoladeformacao.sp.gov.br/portais/Portals/36/arquivos/proposta_ciclo_I.pdf. Acesso em: 23/09/21.

SCARPA, Ester Miriam. Apresentação. In: SCARPA Ester Miriam. (Org.) **Estudo de prosódia**. Campinas: Editora da Unicamp, 1999, p 7-20.

SCHWINDT, Luiz Carlos. **O prefixo no português brasileiro: análise prosódica e lexical**. São Paulo: *DELTA*, v.17, n.2, 2001, p.175-207.

SELKIRK, Elisabeth. On derived domains in sentence phonology. **Phonology Yearbook**, nº3, 1986.

SILVA, Ademar da. **Alfabetização: a escrita espontânea**. São Paulo: Contexto, 1991.

SILVA, Lilian Maria da. **Um estudo longitudinal sobre as hipersegmentações de palavras escritas nos anos finais do Ensino Fundamental**. São José do Rio Preto, 2014. Dissertação. 155f. (Mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Unesp.

SIMIONI, Taíse. **O clítico e seu lugar na estrutura prosódica em português brasileiro.** *ALFA*, São Paulo, 52 (2), 2008, p.431-446.

TENANI, Luciani. **Domínios prosódicos no português do Brasil: implicações para a prosódia e para a aplicação de processos fonológicos.** Campinas. 331f. Tese (Doutorado em Linguística). Universidade Estadual de Campinas, 2002.

TENANI, Luciani. **Segmentações não-convencionais e teorias fonológicas.** *Letras de hoje: Porto Alegre*. V. 39, nº. 3, setembro, 2004, p. 233-244.

TENANI, Luciani. **Entre o grupo clítico e a palavra fonológica: os erros de segmentação não convencional de palavras.** *VI Congresso Internacional da ABRALIN*, 2009a.

TENANI, Luciani. A segmentação não-convencional de palavras: evidências de representações de relações entre enunciados falados e escritos. **Anais do II SIMELP: língua portuguesa: ultrapassar fronteiras, juntar culturas**, 2009b, p. 107-127.

TENANI, Luciani. A segmentação não-convencional de palavras e a aquisição da escrita. **Palestra apresentada no II Ciclo de Palestras sobre aquisição da linguagem e escrita.** Faculdade de Educação, Universidade Federal de Pelotas, Pelotas. 2010.

TENANI, Luciani. Letramento e segmentações não-convencionais de palavras. In: TFOUNI, Leda. **Letramento, escrita e leitura: questões contemporâneas.** Campinas: Mercado das Letras, 2011a.

TENANI, Luciani. A segmentação não-convencional de palavras em textos do ciclo II do Ensino Fundamental. *Revista da ABRALIN*, v.10, n.2, jul./dez. 2011b, p. 91-119.

TENANI, Luciani. **Sobre fronteiras: prosódia, escrita e palavras.** Araraquara: Letraria, 2021.

TORQUETTE, A.O. **Segmentações não-convencionais de palavra em escolas públicas e privadas.** São José do Rio Preto, 2016. Tese. 155f. (Doutorado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Unesp.

TFOUNI, Leda Verdiani. **Letramento e alfabetização.** São Paulo: Cortez, 1995.

VIGÁRIO, Marina. **The Prosodic Word in European Portuguese.** Tese de doutorado, Universidade de Lisboa, 2001.