



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Campus de São José do Rio Preto

AKISNELEN DE OLIVEIRA TORQUETTE

SEGMENTAÇÕES NÃO-CONVENCIONAIS DE PALAVRAS
EM ESCOLAS PÚBLICAS E PRIVADAS

São José do Rio Preto

2016

AKISNELEN DE OLIVEIRA TORQUETTE

SEGMENTAÇÕES NÃO-CONVENCIONAIS DE PALAVRAS
EM ESCOLAS PÚBLICAS E PRIVADAS

Tese apresentada ao Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista, Campus de São José do Rio Preto, para obtenção do título de Doutor em Estudos Linguísticos (Área de Concentração: Análise Linguística).

Orientador: Prof. Dr. Lourenço Chacon Jurado Filho

São José do Rio Preto

2016

Torquette, Akisnelen de Oliveira.

Segmentações não-convencionais de palavras em escolas públicas e privadas / Akisnelen de Oliveira Torquette. -- São José do Rio Preto, 2016
146 f. : il., gráfs., tabs.

Orientador: Lourenço Chacon Jurado Filho
Tese (doutorado) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas

1. Linguística. 2. Análise linguística (Linguística) 3. Língua portuguesa (Ensino fundamental) 4. Palavra (Linguística) 5. Segmentação ortográfica. 6. Escolas públicas. 7. Escolas privadas. I. Jurado Filho, Lourenço Chacon. II. Universidade Estadual Paulista "Júlio de Mesquita Filho". Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU – 41

Ficha catalográfica elaborada pela Biblioteca do IBILCE
UNESP - Campus de São José do Rio Preto

AKISNELEN DE OLIVEIRA TORQUETTE

SEGMENTAÇÕES NÃO-CONVENCIONAIS DE PALAVRAS
EM ESCOLAS PÚBLICAS E PRIVADAS

Tese apresentada ao Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista, Campus de São José do Rio Preto, para obtenção do título de Doutor em Estudos Linguísticos (Área de Concentração: Análise Linguística).

Orientador: Prof. Dr. Lourenço Chacon Jurado Filho

Comissão Examinadora

Titulares

Prof^o. Dr. Lourenço Chacon Jurado Filho (orientador – UNESP)
Prof^a. Dr^a Luciani Ester Tenani (UNESP)
Prof^a. Dr^a Fabiana Cristina Komesu (UNESP)
Prof^a. Dr^a Raquel Salek Fiad (UNICAMP)
Prof^a.Dr^a Ana Paula Nobre da Cunha (UFPeL)

Suplentes

Prof. Dr^a Elaine Cristina de Oliveira (UFBA)
Prof^a. Dr^a Cristiane Carneiro Capristano (UEM)
Prof^a. Dr^a Erotilde Goreti Pezatti (UNESP)

São José do Rio Preto

04 de março de 2016

Dedico este trabalho:

*A Deus,
fonte de vida, consolo, sabedoria, inspiração e força.*

*A minha Mãe,
motivo e sustentação da minha caminhada, meu porto seguro.*

AGRADECIMENTOS

Sobretudo ao meu orientador, Lourenço Chacon, que me abriu as portas para um horizonte muito maior, sempre preocupado com minha formação. Agradeço pelas orientações e por tudo que me ensinou sobre linguística, leitura, escrita e didática;

À Luciani Tenani, por todo apoio e incentivo. Agradeço por ter me ajudado a organizar meu banco de dados e por me explicar fonologia tão pacientemente. Agradeço pelas leituras dos textos iniciais, pelos debates e por todas as valiosas contribuições para esta pesquisa;

À Fabiana Komesu, por ter me incentivado desde 2006 a fazer pós-graduação, por todas as discussões teóricas, que me fizeram amadurecer muito academicamente. Agradeço pelas leituras e contribuições para esta pesquisa;

À Ana Paula Nobre da Cunha, pelas contribuições dadas no debate do SELIN e na defesa desta pesquisa;

À Raquel Salek Fiad, pela participação e contribuições na defesa desta tese;

À Cristiane Carneiro Capristano, por ter me ajudado a escolher e a construir a proposta de coleta de dados;

À professora Maria Cláudia Cabrini Grácio, pela ajuda na constituição da amostra estratificada que subsidiou a pesquisa que embasou a presente tese;

À professora Tânia Suely Antonelli Marcelino Brabo, pelos esclarecimentos e discussões acerca da complexidade da variável *gênero* na análise dos dados;

À Lilian Maria Silva, por ter me ajudado compartilhando todo material de leitura de que eu precisava, por tirar minhas dúvidas e me explicar conceitos. Agradeço pela companhia, pelas interlocuções, por me acalmar e por me incentivar a continuar durante todo esse processo;

À Aline Vassoler, pela amizade, por abrir as portas da sua casa para mim em Rio Preto. Agradeço por estar sempre comigo me dando força, apoio e incentivo;

À Thaís Chiaramonte, por ter me ajudado com a tabulação dos dados e por todos os momentos de interlocução da pesquisa. Agradeço também por sua amizade e de sua mãe, Maria José, que me receberam como parte da família, cuidando de mim quando estive com dengue e me fazendo companhia. Por causa de vocês meus domingos foram felizes em Marília;

À Cristyane Sampaio Villega, pela ajuda com a estatística, por me ensinar a fazer os testes e interpretá-los;

À Secretaria de Educação Municipal de Marília e à Diretoria Estadual de Educação por terem aceitado esta pesquisa e cedido os dados necessários. Agradeço também aos colégios, diretores, coordenadores, professores e alunos que aceitaram nossa pesquisa, que gentilmente nos receberam e se dispuseram a ajudar na aplicação, desenvolvimento e recolhimento dos dados;

À Ana Cândida, Suelen, Cristina e as meninas do grupo de pesquisa de Marília, pelas companhias e por dividirem comigo as angústias e os desafios. Agradeço pelos cafés e jantares animados.

À minha irmã Aline, minha tia Sandra e toda minha família, por me ajudarem e me incentivarem durante esse e outros desafios da minha vida;

Às minhas amigas Elizângela, Francine, Elienay e a todos os meus amigos de Marília, por terem me ajudado, me consolado, cuidado de mim. Agradeço pelos momentos de descontração, de companhia e por terem cuidado de mim no pós-cirúrgico e na dengue;

Às minhas amigas Adélli, Maísa, Tatiane – e a todas as outras que seria impossível nomear aqui – pela companhia e incentivo. Com vocês, tudo é mais fácil e bom;

Aos professores e ao Programa de Pós-graduação em Estudos Linguísticos, da Unesp/São José de Rio Preto. E também aos professores e aos funcionários da Unesp/Marília, principalmente ao Departamento de Fonoaudiologia, que me recebeu tão bem.

À Capes pelo apoio financeiro;

A cada um que esteve comigo durante esses quatro anos.

SUMÁRIO

INTRODUÇÃO	17
1 FUNDAMENTAÇÃO TEÓRICA	27
1.1 A noção de palavra	27
1.1.1 Noção de palavra do ponto de vista fonológico	34
1.1.2 Noção de palavra do ponto de vista morfossintático	40
1.1.3 Noção de palavra do ponto de vista (orto)gráfico	48
1.1.4 A dificuldade de definição de uma noção de palavra	51
1.2 Segmentações não-convencionais de palavras	54
1.2.1 Hipersegmentações	59
1.2.2 Hipossegmentações	61
1.2.3 Mesclas	62
2 MATERIAL E METODOLOGIA	64
2.1 A concepção e a criação do Banco de Dados <i>EscIn</i>	68
2.1.1 Caracterização do Banco de Dados <i>EscIn</i>	73
2.2 Constituição da amostra	77
2.3 Procedimentos de análise dos dados	83
2.3.1 Análise estatística	88
3 APRESENTAÇÃO DOS RESULTADOS	92
3.1 Tipos de escolas	93
3.2 Anos letivos	95
3.3 Sexo/gênero	106
4 DISCUSSÃO DOS RESULTADOS	114
4.1 Tipos de escolas	114

4.2 Anos letivos	119
4.3 Sexo/gênero	125
CONCLUSÕES	133
REFERÊNCIAS	139
ANEXO	146

LISTA DE FIGURAS

Figura 01 - Exemplo de segmentação de palavras para mais que a escrita convencional	17
Figura 02 - Exemplo de segmentação de palavras para menos que a escrita convencional ...	18
Figura 03 - Exemplo de segmentação de palavras para mais e menos que a escrita convencional	18
Figura 04 - Exemplo de segmentação de palavras para menos que a escrita convencional	39
Figura 05 - Exemplo de segmentação de palavras para menos que a escrita convencional ...	39
Figura 06 - Exemplo de segmentação de palavras para menos que a escrita convencional ...	39
Figura 07 - Mapa do Brasil com destaque em São Paulo e mapa de São Paulo com destaque em Marília	64
Figura 08 - Infográfico do Produto Interno Bruto de Marília, de São Paulo e do Brasil	65
Figura 09 - Infográfico do número de escolas de Marília, de São Paulo e do Brasil	66
Figura 10 - Mapa do Índice de Desenvolvimento Humano dos municípios de São Paulo	67
Figura 11 - Exemplo de texto com hipo e hiper segmentação	85
Figura 12 - Exemplo de hipossegmentação	86
Figura 13 - Exemplo de hipossegmentação	86
Figura 14 - Exemplo de texto não identificado.....	87

LISTA DE GRÁFICOS

Gráfico 01 - Distribuição do número de textos da amostra.....	81
Gráfico 02 - Percentual de ocorrências de segmentações não-convencionais em escolas públicas e privadas	95
Gráfico 03 - Percentual de ocorrências de segmentações não-convencionais nas escolas públicas e privadas e anos letivos (1º ano, 2º ano, 3º ano, 4º ano e 5º ano)	103
Gráfico 04 – A razão da diminuição do percentual de ocorrências de segmentações não-convencionais em escolas públicas e privadas e anos letivos (1º ano, 2º ano, 3º ano, 4º ano e 5º ano)	104
Gráfico 05 - Percentual de ocorrências de segmentações não-convencionais em escolas públicas e privadas e sexo/gênero (masculino e feminino)	112

LISTA DE QUADROS

Quadro 01 - Distribuição do Banco de Dados por zonas do município	75
Quadro 02 - Número total de textos do Banco de Dados <i>EscIn</i> provenientes de escolas públicas e de escolas privadas	76
Quadro 03 - Número total de textos do Banco de dados <i>EscIn</i> de escreventes femininos e escreventes masculinos	77
Quadro 04 - Organização do Banco de Dados e sorteio da amostra	78
Quadro 05 - Organização do Banco de Dados e sorteio da amostra	79
Quadro 06 - Número total de textos da amostra estratificada de escolas públicas e de escolas privadas	81
Quadro 07 - Número total de textos da amostra estratificada de escreventes femininos e escreventes masculinos nos anos letivos	82
Quadro 08 - Número total de textos da amostra estratificada de escreventes femininos e escreventes masculinos	82
Quadro 09 – Dados resultantes do 5º ano do Ensino Fundamental I	104
Quadro 10 – Dados qualitativos resultantes do 5º ano do Ensino Fundamental I	105

LISTA DE TABELAS

Tabela 01 - Relação de escolas públicas municipais de Marília - SP.....	70
Tabela 02 - Relação de escolas privadas de Marília – SP.....	71
Tabela 03 - Número e percentual de participantes de escola pública (2012)	73
Tabela 04 - Número e percentual de participantes de escola privada (2012)	74
Tabela 05 - Distribuição do número de textos final do <i>EscIn</i> pelos critérios de tipos de escolas, anos letivos e sexo/gênero dos escreventes	76
Tabela 06 - Distribuição do número de textos final da amostra estratificada pelos critérios de: tipos de escolas, anos letivos e sexo/gênero dos escreventes	80
Tabela 07 - Distribuição das médias, das medianas e dos desvios padrão de ocorrências de segmentações não-convencionais em escolas públicas e privadas.....	93
Tabela 08 - Distribuição do número de palavras nos textos por tipos de escolas: públicas x privadas	94
Tabela 09 - Distribuição do percentual das segmentações não-convencionais por tipos de escolas: públicas x privadas	94
Tabela 10 - Distribuição das médias, das medianas e dos desvios padrão de ocorrências de segmentações não-convencionais em tipos de escolas (públicas x privadas) e anos letivos (1º x 2º x 3º x 4º x 5º)	96
Tabela 11 - Distribuição do número de palavras dos textos no 1º ano por tipos de escolas: públicas x privadas	97
Tabela 12 - Distribuição do percentual das segmentações não-convencionais no 1º ano por tipos de escolas: públicas x privadas	97
Tabela 13 - Distribuição do número de palavras no 2º ano por tipos de escolas: públicas x privadas	97
Tabela 14 - Distribuição do percentual das segmentações não-convencionais no 2º ano por tipos de escolas: públicas x privadas	98
Tabela 15 - Distribuição do número de palavras no 3º ano por tipos de escolas: públicas x privadas	98
Tabela 16 - Distribuição do percentual das segmentações não-convencionais no 3º ano por tipos de escolas: públicas x privadas	98

Tabela 17 - Distribuição do número de palavras no 4º ano por tipos de escolas: públicas x privadas	99
Tabela 18 - Distribuição do percentual das segmentações não-convencionais no 4º ano por tipos de escolas: públicas x privadas	99
Tabela 19 - Distribuição do número de palavras das segmentações não-convencionais no 5º ano por tipos de escolas: públicas x privadas	99
Tabela 20 - Distribuição do percentual das segmentações não-convencionais no 5º ano por tipos de escolas: públicas x privadas	100
Tabela 21 - Distribuição do número de palavras nos textos por anos letivos nas escolas públicas (1º x 2º x 3º x 4º x 5º)	100
Tabela 22 - Distribuição do percentual das segmentações não-convencionais por anos letivos nas escolas públicas (1º x 2º x 3º x 4º x 5º)	101
Tabela 23 - Distribuição do número de palavras nos textos por anos letivos nas escolas privadas (1º x 2º x 3º x 4º x 5º)	102
Tabela 24 - Distribuição do percentual das segmentações não-convencionais por anos letivos nas escolas privadas (1º x 2º x 3º x 4º x 5º)	102
Tabela 25 - Distribuição das segmentações não-convencionais por tipos de escolas (públicas e privadas) e sexo/gênero (masculino x feminino)	106
Tabela 26 - Distribuição das médias, das medianas e dos desvios padrão de ocorrências de segmentações não-convencionais em escolas públicas por anos letivos (1º x 2º x 3º x 4º x 5º) e sexo/gênero (masculino x feminino)	107
Tabela 27 - Distribuição das médias, das medianas e dos desvios padrão de ocorrências de segmentações não-convencionais em escolas privadas por anos letivos (1º x 2º x 3º x 4º x 5º) e sexo/gênero(masculino x feminino)	107
Tabela 28 - Distribuição do percentual das segmentações não-convencionais por anos letivos (1º x 2º x 3º x 4º x 5º) de escreventes femininos nas escolas públicas	108
Tabela 29 - Distribuição do percentual das segmentações não-convencionais por anos letivos (1º x 2º x 3º x 4º x 5º) de escreventes masculinos nas escolas públicas	108
Tabela 30 - Distribuição do percentual das segmentações não-convencionais por anos letivos (1º x 2º x 3º x 4º x 5º) de escreventes femininos nas escolas privadas	109
Tabela 31 - Distribuição do percentual das segmentações não-convencionais por anos letivos (1º x 2º x 3º x 4º x 5º) de escreventes masculinos nas escolas privadas	109

Tabela 32 - Distribuição do número de palavras nos textos de escreventes femininos em escolas públicas e privadas	110
Tabela 33 - Distribuição do percentual de segmentações não-convencionais de escreventes femininos em escolas públicas e privadas	110
Tabela 34 - Distribuição do número de palavras nos textos de escreventes masculinos em escolas públicas e privadas	111
Tabela 35 - Distribuição do percentual das segmentações não-convencionais de escreventes masculinos em escolas públicas e privadas	111
Tabela 36 - Distribuição do número de palavras nos textos por sexo/gênero (masculino x feminino)	101
Tabela 37 - Distribuição do percentual das segmentações não-convencionais por sexo/gênero (masculino x feminino)	112

RESUMO

No presente trabalho, tivemos como objetivo comparar a distribuição das segmentações não-convencionais de palavras encontradas em textos do 1º ao 5º ano do Ensino Fundamental I, a fim de traçar um perfil dessas ocorrências em função das variáveis: (a) tipos de escolas (públicas e privadas); (b) anos letivos (1º ano, 2º ano, 3º ano, 4º ano e 5º ano); e (c) sexo/gênero (masculino e feminino). Para tanto, compusemos um banco de dados – denominado EscIn – composto por 4.230 textos resultantes da aplicação de uma proposta de atividade de escrita, do 1º ao 5º ano do Ensino Fundamental I, de cinco escolas privadas e de dez escolas públicas do Município de Marília (SP), durante o ano de 2012. A partir desse banco de dados, formamos uma amostra estratificada de 10% dos dados de cada variável de nosso objetivo, perfazendo um total de 427 textos do *corpus* a ser analisado. Fizemos análise estatística descritiva e inferencial dos dados de segmentação não-convencional, comparando as variáveis. Os resultados das ocorrências de segmentações não-convencionais mostram que: (1) ocorrem em maior percentual nas escolas públicas do que nas escolas privadas; (2) tendem a diminuir em ambos os tipos de escolas com a progressão dos anos letivos, mas, de forma diferente – nas escolas públicas a queda é contínua, com um pico do 2º para o 3º ano, já nas escolas privadas a queda é acentuadamente do 1º para o 2º ano e, a partir do 3º ano, o percentual zera; e (3) são mais numerosas entre os meninos do que entre as meninas nas escolas públicas, mas não nas escolas privadas. A forma de segmentar dos escreventes mostra-se, pois, como heterogênea, na medida em que é condicionada pelos tipos de escolas: públicas ou privadas. Os resultados das variáveis anos letivos e sexo/gênero indiciam que há outros fatores que influenciam no modo de segmentar dos escreventes, não sendo possível – ou pouco proveitoso – analisar essas variáveis de forma isolada.

Palavras-chave: língua portuguesa, palavra, aquisição da escrita, segmentações não-convencionais, escola pública/escola privada.

ABSTRACT

In this study, our objective was to compare the distribution of unconventional segmentations words found in texts of the 1st to 5th grade of elementary school in order to draw a profile of these occurrences depending on variables: (a) type of school (public and private); (b) school years (1st year, 2nd year, 3rd year, 4th year and 5th year); and (c) sex/gender (male and female). To this end, we wrote a database - called EscIn - composed of 4,230 texts resulting from the application of a proposal for a writing activity, from 1st to 5th year of elementary school, five private schools and ten public schools in the city of Marilia (SP) during 2012. From this database, we formed a stratified sample of 10% to each variable of our goal, making a *corpus* total of 427 texts to be analyzed. In descriptive and inferential statistical analysis data of unconventional segmentation, we compare the variables. The results of unconventional segmentations occurrences show that: (1) occur in higher percentage in public schools than in private schools; (2) tend to decrease in both types of schools with the progression of school years, but differently - in public schools fall is continuing, with a peak in the 2nd to the 3rd year, already in private schools fall is sharply from the 1st to the 2nd year and, from the 3rd years, the percentage resets; and (3) are more numerous among boys than among girls in public schools, but not in private schools. The way of to segment writers is shown since, as heterogeneous as far as it is conditioned by the school types: public and private. The results of the school years and sex/gender variables indicate that there are other factors that influence the way to segment the writers, it is not possible - or little profitable - to analyze these variables in isolation.

Keywords: Portuguese, word, writing acquisition, unconventional segmentations, public school / private school.

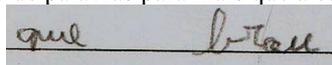
INTRODUÇÃO

Os escreventes, em aquisição de escrita, transitam por diferentes caminhos em busca da aprendizagem da **palavra** gráfica e de seus limites, pois, mesmo reconhecendo diferenças entre fala e escrita, saber onde ocorrem as fronteiras de palavras na escrita (delimitadas por espaços em branco ou hífen) gera grande dúvida. Assim, textos de escreventes em aquisição da escrita são, segundo Abaurre (1991), unidades de análise privilegiadas, pois fornecem, ao pesquisador, pistas de processos que subjazem a essa aquisição.

A dificuldade de conceber o que é uma palavra gera conflitos quanto aos critérios de definição de seu conceito, uma vez que, para a definição desse conceito, concorrem aspectos como: (i) o fonológico; (ii) o morfossintático; (iii) o semântico; e/ou (iv) o (orto)gráfico. Para Abaurre (1989), essa dificuldade de saber o que é uma palavra faz com que o escrevente tenha a necessidade de testar hipóteses, que podem ser sustentadas, por exemplo, em conhecimentos fonológicos e/ou semânticos. Na perspectiva dessa autora, a segmentação não-convencional de palavras na escrita infantil indicia hipóteses sobre o que poderia ser – ou não – uma palavra e sobre quais seriam os limites entre elas de acordo com a convenção ortográfica.

Essas dificuldades geram escritas como as apresentadas nas figuras 1, 2 e 3, a seguir:

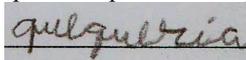
Figura 01-Exemplo de segmentação de palavras para mais que a escrita convencional



Fonte: banco de dados *EscIn*¹.

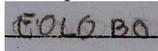
¹O banco de dados *EscIn* (Escrita Infantil) foi desenvolvido nesta pesquisa com textos de crianças do Ensino Fundamental I de escolas públicas e privadas do município de Marília-SP. No capítulo II, serão detalhadas as características do *EscIn*, bem como a forma com que foi constituído.

Figura 02 - Exemplo de segmentação de palavras para menos que a escrita convencional



Fonte: banco de dados *Escln*.

Figura 03 - Exemplo de segmentação de palavras para mais e menos que a escrita convencional



Fonte: banco de dados *Escln*.

As estruturas que apresentamos correspondem ao que se convencionou categorizar, na literatura, respectivamente, como: (a) na figura 1, um exemplo de *hipersegmentação* (segmentações de palavras *para mais*) “que brou” – para *quebrou*; (b) na figura 2, um exemplo de *hipossegmentação* (segmentações de palavras *para menos*) “quequeria” – para *que queria*; e (c) na figura 3, um exemplo de *mescla* (estruturas que combinam partes hiper e hipossegmentadas)² “eolo bo” – para *e o lobo*.

No início da aquisição de escrita, a instabilidade encontrada nos limites gráficos entre as palavras já é esperada, uma vez que decorre, por exemplo, de conflitos entre convenções ortográficas e constituintes prosódicos – tais como aqueles propostos por Nespor e Vogel (1986). Essas relações entre convenções ortográficas e constituintes prosódicos – investigadas nos trabalhos de Abaurre (1988, 1989, 1991, 1996, 1999, 2013), Cunha (2004, 2007, 2009, 2010, 2012), Chacon (2004, 2005, 2006), Capristano (2004), Tenani (2004, 2006, 2009, 2011) – indiciam a relação entre convenções ortográficas (como as de que unidades morfológicas se

² *Mesclas* são a ocorrência de hipo e hipersegmentações em uma mesma estrutura – classificação assumida pelos pesquisadores do grupo de pesquisa Estudos sobre a Linguagem – GPEL (CHACON 2004; TENANI, PARANHOS, 2011). *Híbridos* são os processos de segmentação não-convencional em que ocorrem hipo e hipersegmentação em uma mesma sequência – classificação assumida pelos pesquisadores do grupo de estudos sobre a aquisição da linguagem escrita – GEALE (CUNHA, 2010, p. 347).

separam umas das outras por meio de espaços em branco) e constituintes prosódicos (tais como eles podem ser apreendidos nos enunciados falados) (CHACON, 2013).

Devemos destacar, também, que no conflito das relações entre as convenções ortográficas e os constituintes prosódicos emergem representações dos escreventes das palavras escritas, que podem variar tanto de criança para criança, quanto para uma mesma criança – já que, em um mesmo texto, observam-se flutuações que se expressam em diferentes grafias de uma mesma palavra (TENANI, 2009; CAPRISTANO, 2003; CHACON, 2013). Essas flutuações, em uma perspectiva discursiva, – a qual assumimos neste trabalho – são entendidas como fruto do trânsito do escrevente por práticas orais e letradas.

As flutuações, em uma perspectiva construtivista – diferentemente da qual assumimos neste trabalho –, podem revelar parte do conhecimento construído desde o início do contato das crianças com a língua, bem como hipóteses que elas fazem para reelaborar seu sistema linguístico ao adquirirem a escrita (FERREIRO; PONTECORVO, 1996; CUNHA, 2004; MIRANDA, 2006).

A segmentação não-convencional de palavras, na aquisição de escrita, tem se mostrado como um importante lugar de pesquisa em diferentes áreas, como comprovam estudos em: (a) Fonoaudiologia (ZORZI, 1998; MENESES; LOZI; SOUZA; ASSENCIO-FERREIRA, 2004; ANDREAZZA-BALESTRIN; CIELO; LAZZAROTTO-VOLCÃO; LAUSCH, 2012); (b) Linguística (ABAURRE, 1991; Silva, 1994; CAPRISTANO, 2003; CHACON, 2005; TENANI, 2011); e (c) a interface Linguística-Educação (CUNHA, 2004; CUNHA; MIRANDA, 2008).

Nos estudos sobre a aquisição de escrita infantil desenvolvidos no Brasil – com sedes em diferentes universidades e regiões do país, – destacam-se trabalhos oriundos de três vertentes de pesquisa, a saber: (1) do projeto integrado *A relevância teórica dos dados singulares na aquisição de escrita*, coordenado pelas pesquisadoras Maria Bernadete Marques

Abaurre, Maria Laura Trindade Mayrink-Sabinson e Raquel Salek Fiad – do Instituto de Estudos da Linguagem da Universidade Estadual de Campinas (IEL/UNICAMP); (2) do Grupo de Pesquisa *Estudos sobre a linguagem* (GPEL/CNPq), coordenado pelo pesquisador Lourenço Chacon – da Universidade Estadual Paulista (UNESP); e do *Grupo de estudos sobre a aquisição da linguagem escrita* (GEALE/CNPq), coordenado pela pesquisadora Ana Ruth Moresco Miranda – da Universidade Federal de Pelotas (UFPel).

O aspecto comum a essas três vertentes de estudos é o de destacarem, na aquisição da escrita, a importância dos aspectos fonético-fonológicos da língua, além de, em maior ou menor grau e com diferentes concepções, abordarem as relações sujeito/linguagem.

Os pesquisadores que fazem parte das vertentes de estudos apresentadas têm chamado a atenção para (entre outros fenômenos da escrita infantil) a importância de estudos sobre segmentações não-convencionais de palavras. No entanto, mesmo produzindo um volume significativo de trabalhos, esses pesquisadores têm mostrado que a riqueza desse fenômeno requer ainda mais investigações para que haja melhor compreensão da aquisição da escrita e, conseqüentemente, dos seus aspectos linguísticos e do seu ensino.

O estudo das segmentações não-convencionais de palavras se mostra especialmente importante, também, porque, em um dos documentos oficiais mais importantes do Ministério da Educação, os Parâmetros Curriculares Nacionais (PCN) – norteador do ensino escolar no Brasil –, encontra-se a recomendação de que os alunos deveriam terminar o Ensino Fundamental I sendo capazes de “escrever textos com domínio da separação de palavras” (BRASIL, 2000, p.80). Dito de outro modo, de acordo com os PCN, os escreventes deveriam iniciar o Ensino Fundamental II sem flutuações em relação ao registro da fronteira gráfica de palavra.

Entretanto, não é o que demonstram Paranhos e Tenani (2011), Tenani (2011), Paranhos (2014) e Silva e Tenani (2014). Com efeito, nesses trabalhos, as autoras apontam

grande recorrência de dados de segmentação não-convencional de palavras mesmo em textos de escreventes do Ensino Fundamental II. Segmentar palavras de acordo com a ortografia convencional, portanto, não se mostra tarefa fácil para crianças e adolescentes.

Embora, como já destacamos, a dificuldade em segmentar as palavras não seja um assunto novo de investigação, diversas questões relativas à segmentação não-convencional de palavras merecem ser ainda investigadas.

Observando as diferenças de condições educacionais na realidade brasileira, sobretudo quando se as observam escolas públicas e privadas de ensino, nos impulsionou a ter como hipótese que essas diferenças se mostram também na maneira como escreventes do Ensino Fundamental I, oriundos desses dois tipos de escolas, segmentam palavras em seus textos escritos. Testá-la – acreditamos nós – possibilitaria ampliação do conhecimento de possíveis variantes sociais que influenciariam esse fenômeno, uma vez que

O fator sócio-econômico parece também influenciar os rumos tomados pela escrita infantil com relação aos processos de segmentação. 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).

Essa tendência parece consensual quando se observa a quantidade de ocorrências de segmentações não-convencionais de palavras; porém não se é mostrado quão diferente é a proporção dessas ocorrências.

Com efeito, em um estudo mais aprofundado, baseado em dois mil textos oriundos de uma escola pública e uma escola privada, Cunha (2010, p. 331) aponta que acontecem mais erros nos textos da escola pública do que nos textos da escola particular. Porém, por esse não ser seu objetivo principal de investigação, a autora não estima percentuais dessa diferença. Em outras palavras, a autora mostra que há diferença entre escolas públicas e privadas e que essa diferença pode ser provada de forma numérica, sem, no entanto, investigá-la de forma mais aprofundada.

Mesmo havendo consenso entre os trabalhos de Abaurre (1991) e de Cunha (2010), quanto ao maior número de segmentações não-convencionais de palavras em escolas públicas, há, ainda, pouca pesquisa sobre a relevância do fator socioeconômico na aprendizagem da segmentação na realidade brasileira, o que reforça, pois, o merecimento – para o qual Abaurre (1991) chama a atenção – de uma pesquisa mais aprofundada, que investigue a relevância de fatores sociais (como escolas de ensino público e privado) no fenômeno das segmentações não-convencionais de palavras.

O que se pode verificar nos estudos de Abaurre (1991) e de Cunha (2010) é que há uma diferença na quantidade de erros dos textos oriundos de escreventes de escolas públicas, que mostra maior quantidade de ocorrências de segmentações não-convencionais de palavras em relação aos textos oriundos de escreventes de escolas privadas, que mostram menor quantidade de ocorrências.

Entretanto instiga-nos pensar o quanto varia essa diferença em função desses dois tipos de escolas. Dito de outra forma, qual é a proporção da diferença entre o número de ocorrências de segmentações não-convencionais de palavras de textos de escreventes oriundos de escolas públicas e de textos de escreventes oriundos de escolas privadas? Investigar quantitativamente essa diferença, justamente, é a proposta da presente pesquisa.

A importância de nosso estudo se configura, então, na ampliação do conhecimento sobre possíveis diferenças na distribuição da segmentação não-convencional de palavras em escolas públicas e privadas, mostrando quanto é essa diferença e se ela é estatisticamente significativa, ou não. Temos como objetivo, a partir desses dados, pensar sobre as diferenças de segmentações não-convencionais de palavras de escolas públicas e privadas no que concerne à realidade que investigamos.

No interior desses dois tipos de escolas (públicas e privadas), pelo menos dois fatores mereceriam um olhar especial. O primeiro deles é o dos anos letivos, fator para o qual

chamam a atenção Capristano (2003), Tenani e Paranhos (2011) e Paranhos (2014). O destaque que essas autoras deram a esse fator se deve aos resultados de suas investigações mostrarem significativa queda de ocorrências de segmentação não-convencional de palavras: Capristano (2003), em textos de escreventes do Ensino Fundamental I da primeira à quarta série; Tenani e Paranhos (2011) e Paranhos (2014), em textos de escreventes do Ensino Fundamental II da quinta à oitava série³.

Também em outra situação de investigação – a da ortografia e das patologias relacionadas à escrita – Capellini (2010) detecta a mesma diminuição das ocorrências de segmentação não-convencional de palavras. Fato que reforça a necessidade de se olhar com mais atenção para o fator anos letivos. Uma recomendação da autora é que,

Os dados deste estudo apontam para o desenvolvimento de mais estudos na área, com a finalidade de estabelecer o perfil de aquisição e desenvolvimento ortográfico de escolares em fase inicial de alfabetização, pois somente dessa forma será possível identificar quais os erros comuns à apropriação do sistema de escrita e aqueles erros que são persistentes e sugestivos de transtornos de aprendizagem, como a dislexia e o distúrbio de aprendizagem (CAPELLINI, 2010, p. 159).

Merece, pois, ser melhor investigado o efeito dos anos letivos sobre a quantidade de ocorrências de segmentação não-convencional de palavras ao longo do Ensino Fundamental I, já que as pesquisas que relatam diferença na quantidade de ocorrências têm objetivos qualitativos de investigação diversificados e não se destinam a quantificar essa diferença. Mas como nossa hipótese é a de que a diferença de tipos de escolas (públicas e privadas) provoca diferenças na maneira como estudantes de escolas públicas e privadas segmentam palavras em sua escrita, devemos investigar possíveis diferenças quantitativas, também, no efeito dos anos letivos – uma vez que, fundamentalmente, os trabalhos sobre segmentação não-convencional de palavras desenvolvidos no Brasil seguem a tradição instaurada por Abaurre (1989, 1991) de buscar, nelas, seus vínculos com constituintes prosódicos, verificando diferenças

³ Os três trabalhos citados (CAPRISTANO, 2003; TENANI, PARANHOS, 2011; PARANHOS 2014) foram desenvolvidos com bancos de dados constituídos antes da mudança do Ensino Fundamental para nove anos e da mudança da nomenclatura de série para ano, instituídas pela lei nº 11.274/06.

qualitativas nas segmentações. Torna-se, pois, importante buscar dados quantitativos mais precisos decorrentes da variação de tipos de escolas (públicas e privadas), bem como de anos letivos (1º ano, 2º ano, 3º ano, 4º ano e 5º ano).

Um terceiro fator, para o qual chamam a atenção Meneses (2004) e Andreazza-Balestrin *et al* (2012), merece também um olhar especial em nossa investigação: o do gênero/sexo. Nesses dois estudos, a relevância desse fator para a explicação de desempenhos em tarefas – conforme a perspectiva de investigação, se mostram como relevantes para a aquisição da escrita – foi discordante. Essa falta de concordância e, conseqüentemente, de caracterização do fenômeno acerca da variável sexo/gênero nos provocou o interesse em verificar em que medida meninos segmentariam diferentemente seus textos do que meninas.

Em virtude da pouca existência de discussões e de pesquisas que investigam as segmentações não-convencionais de palavras em escolas públicas e privadas, da confirmação sobre a diminuição das ocorrências por anos letivos ao longo do Ensino Fundamental I e da necessidade de uma melhor caracterização desse fenômeno quanto ao gênero/sexo do escrevente, no presente trabalho, temos o objetivo geral analisar como as ocorrências de segmentações não-convencionais de palavras se distribuem em função da variação dos *tipos de escolas* (públicas e privadas). Em decorrência desse objetivo geral, temos como objetivos específicos analisar, dentro da variável *tipos de escolas*, a distribuição das segmentações não-convencionais dos *anos letivos* (1º ano, 2º ano, 3º ano, 4º ano e 5º ano) e do *sexo/gênero* dos escreventes (masculino e feminino).

Para tanto, no capítulo 1, *Fundamentação teórica*, discutiremos na primeira seção a noção de palavra de forma geral e em três abordagens, a saber: do ponto de vista fonológico, do ponto de vista morfossintático e do ponto de vista (orto)gráfico – atentando para a complexidade desse conceito nos estudos linguísticos; na segunda seção, discutiremos o

conceito de segmentações não-convencionais de palavras, apresentando suas características e classificações.

No capítulo 2, *Material e metodologia*, trataremos do percurso metodológico. Inicialmente justificaremos a escolha e importância da cidade de aplicação da pesquisa, bem como descrevemos a realização do banco de dados, mostrando as dificuldades e as soluções que encontramos para elas. A partir da constituição do *EscIn* (nome dado ao nosso banco de dados), relataremos os critérios que adotamos para a formação de uma amostra estratificada. Em seguida, abordaremos os procedimentos de análise pautado em análises estatísticas, descritiva e inferencial.

No capítulo 3, *Apresentação dos resultados*, mostramos os resultados estatísticos, descritivos e inferenciais, em função das nossas variáveis: (1) tipos de escolas (públicas e privadas); (2) anos letivos (1º ano, 2º ano, 3º ano, 4º ano, 5º ano); e (3) sexo/gênero (masculino e feminino).

No capítulo 4, *Discussão dos resultados*, trataremos das tendências apontadas pela análise estatística dos resultados. Primeiramente, discutiremos os resultados de que há mais ocorrências de segmentações não-convencionais de palavras em escolas públicas do que em escolas privadas. Em seguida, discutiremos os resultados da variável anos letivos, em função da qual detectamos uma diminuição das ocorrências de segmentações não-convencionais de palavras ao longo dos anos, mas de forma diferente em escolas públicas e privadas. Por fim, discutiremos, também, a variável sexo/gênero, levantando hipóteses sobre o porquê de os textos de escreventes masculinos apresentarem mais ocorrências de segmentações não-convencionais de palavras do que os textos de escreventes femininos em escolas públicas e tal fato não ocorrer em escolas privadas.

Na conclusão, faremos uma síntese das discussões consideradas como mais relevantes. Também apontaremos possíveis desdobramentos, que não pudemos abordar em função do nosso recorte, vinculados às reflexões que serão desenvolvidas no decorrer deste trabalho.

1 FUNDAMENTAÇÃO TEÓRICA

“TRADICIONALMENTE, tôda e qualquer descrição da língua portuguesa leva em conta a existência do <vocábulo>, e nisso se baseia.” (CAMARA JR, 1971, p. 34). Com base nessa afirmação de Camara Jr. (1971), apresentaremos (para o desenvolvimento deste trabalho) investigações teóricas que discutem as noções de palavra e de segmentações não-convencionais de palavras.

A noção de palavra é discutida por muitos campos dos saber e variadas áreas⁴. Os estudos realizados por essas áreas se dão nos diferentes planos da linguagem, seja o fonológico ou o morfossintático, bem como os aspectos (orto)gráficos da língua.

Refletir, então, sobre a noção de palavra é importante para o nossa pesquisa, na medida em que é um desafio – tanto para o pesquisador em Linguística quanto para o (pequeno) escrevente – identificar nos enunciados suas fronteiras de palavras, já que se mostram como instáveis.

1.1 Noções de palavra

A noção de palavra foi por muito tempo (e ainda tem sido) tratada como óbvia, como já sabida ou como uma verdade inquestionável. Porém, em se tratando de linguagem, sabemos não haver verdades inquestionáveis, pois, a depender do olhar, o fenômeno estudado pode ser definido de variadas maneiras.

⁴ Reconhecemos que há outras abordagens linguísticas que dedicam estudos à noção de palavra (como as da lexicografia e da terminologia), bem como estudos decorrentes de abordagens de outras áreas (como a pedagogia, a fonoaudiologia e a psicologia); porém, neste trabalho, nos restringiremos aos aspectos fonológico, morfossintático e (orto)gráfico da palavra.

Definir, então, o que é uma palavra não é tarefa fácil. Muitos linguistas têm enfrentado esse desafio investigando a noção de palavra nos níveis fonológico, morfossintático, bem como em aspectos (orto)gráficos da língua. No geral, é costumeiro haver noções ancoradas em critérios pouco esclarecedores, como as expressas em dicionários de grande circulação nacional, segundo os quais palavra é

1. Fonema ou grupo de fonemas com uma significação; termo, vocábulo: ‘Agora vão ... aparecer algumas palavras difíceis, mas quem não souber, dever ir ver no dicionário’ (José Saramago, *A maior Flor do Mundo*). 2. Sua representação gráfica. 3. Manifestação verbal ou escrita. 4. Faculdade de expressar ideias por meio de sons articulados; fala. 5. Modo de falar (FERREIRA, 2011, p. 651 – grifo do autor).

Mesmo em manuais gramaticais – desde os mais tradicionais até os que se apresentam como reflexivos⁵ – não há preocupação explícita com a noção de palavra. As explicações encontradas (cf. CUNHA & CINTRA, 1985; NICOLA, 2004; BECHARA, 2006; CEREJA & COCHAR, 2009) ocorrem em meio à apresentação de outras noções (como o de fonema) ou no contexto de discussões sobre morfologia e ortografia.

É comum encontrarmos em discursos pedagógicos ou em livros didáticos de alfabetização a noção de palavra como uma unidade da língua que possui sentido ou função gramatical.

No discurso pedagógico veiculado na reportagem “É para escrever tudo junto ou separado?”, publicada na Revista Nova Escola, edição 261, de Abril de 2013, são relatadas propostas de ensino para o trabalho com segmentações não-convencionais de palavras. A noção de palavra é caracterizada pelo critério gráfico, na medida em que se ressaltam os espaços em branco e a falta de correspondência entre escrita e fala.

Destaca-se, na reportagem, que os alunos reproduzem “no papel o que escutam e falam. Mas essa estratégia nem sempre funciona” (MEIRELLES, 2013, p. 1). A separação dos

⁵ Chamamos de tradicionais as gramáticas que, segundo Perini (1996) e Possenti (1996), apresentam suas definições sob ótica normativa ou prescritiva.

elementos, a qual ocorre por “um caráter sumamente abstrato” (FERREIRO, E; TEBEROSKY, A, 1985, p. 106 *apud* MEIRELLES, 2013, p. 1), é discutida, com base em Ferreiro e Teberosky (1985), sem menção na reportagem ao que não está claro e/ou concreto para a autora. As atividades propostas para trabalhar as segmentações não-convencionais são: (i) separar uma estrofe de música dada sem os espaços devidos; (ii) encontrar erros de separação de palavras em textos do gênero bilhete; (iii) escrever a partir de um ditado com focalização. Pelas atividades, confirmamos a concepção do conceito de palavra pautado no critério gráfico. No entanto, as atividades (i) e (iii) são incoerentes com a afirmação de que não há correspondência entre fala e escrita. A atividade (i) porque, pedir ao aluno que separe as palavras de uma estrofe de uma música, vai (no mínimo) deixá-lo confuso, pois, para ele, em muitos casos, as pausas que ocorrem na música são diferentes daquelas presentes na letra da canção. A atividade (iii) porque as focalizações que privilegiariam as fronteiras de palavras não serão as mesmas de uma leitura corrente (em que ocorrem fusões de/entre palavras), o que torna a atividade artificial.

No material didático apostilado *Uno* (utilizado como apoio didático-pedagógico em uma das escolas privadas que fizeram parte do banco de dados constitutivo desta pesquisa), editado pela Editora Bernardi – uma afiliada da Editora Moderna –, a gramática trabalhada é indutiva, já que não há exposição de regras explícitas no material do aluno. Na apresentação, os autores antecipam: “você vai descobrir como as letras se combinam para formar palavras e aprender como usar as palavras para se comunicar por escrito”(s/p).

No primeiro ano, a apostila inicia com um texto sobre nomes de pessoas e propõe, como atividade, que as crianças copiem os nomes que iniciam com a mesma letra de cada alternativa do exercício, destacando no trabalho a atenção aos grafemas. São propostos sete diferentes exercícios com nomes próprios e suas diferentes formas de escrita, tais como: quantas letras há nos nomes?; Qual é o maior nome?; E qual é o menor nome?. Apesar de

contextualizar os exercícios com a realidade do aluno, a concepção de língua presente no material é tradicional, pois a preocupação é com a estrutura da palavra, o tamanho e a quantidade de letras.

O último exercício da unidade pede para a criança escrever o nome de quatro colegas e depois escrever “uma palavra para cada um dos quatro colegas” (p. 15). No entanto, não há explicação no material sobre o que seria essa palavra, nem o que seria uma palavra. Pelo que vinha sendo trabalhado na unidade, podemos supor que a concepção subjacente de palavra para os autores e a qual supostamente seria identificada pelas crianças nas atividades trabalhadas é a de conjunto de grafemas.

No Ensino Fundamental I público do município de Marília-SP (região do centro-oeste paulista), cidade na qual desenvolvemos a pesquisa, o material didático utilizado para a alfabetização é desenvolvido pelas professoras com base em pressupostos apresentados em uma Proposta Curricular do município. O documento é um regulador para os conteúdos e objetivos a serem trabalhados nas escolas municipais de Ensino Fundamental I. Esse documento surge no final de 2008 com a necessidade de implantação, em 2009, da mudança do Ensino Fundamental para nove anos. O objetivo apresentado é o de que “A Proposta Curricular irá subsidiar o fazer pedagógico de cada professor e garantir uma unidade de ação em nível de Secretaria Municipal da Educação” (MARÍLIA, 2008, p. 02). Destaca-se, porém, que não se está traçando um modelo homogêneo de ensino, mas, sim, uma ferramenta norteadora de ações em busca de melhorias.

Para todas as disciplinas, há expectativas de aprendizagem, conteúdos e orientações didáticas. Referente à Língua Portuguesa, as expectativas e conteúdos estão, em maior número, voltados à oralidade. No entanto, quanto à escrita vemos indicações de “Compreender a natureza do sistema da escrita; e, no mínimo, produzir escrita silábica com ou sem valor sonoro convencional;”, também “Conhecer e nomear as representações das

letras do alfabeto de imprensa maiúsculo;” e, ao final do ano, “Escrever o próprio nome e utilizá-lo como referência para a escrita;” (MARÍLIA, 2008, p. 4). A orientação didática que permeia o ensino da escrita é de que as atividades dadas às crianças façam sentido para elas e que isso possibilite que os alunos percebam o caráter social da escrita.

É interessante observar que a atividade de a criança escrever o próprio nome aparece tanto no livro trabalhado em uma escola privada quanto na proposta da escola pública. Porém, a atividade é a primeira do ano na escola privado e uma das últimas para as escolas públicas.

Ainda sobre discursos pedagógicos, Chacon (2010) analisa a noção de palavra que professoras alfabetizadoras e seus alunos têm. As análises foram construídas com base nas respostas de alfabetizadoras sobre o que seria palavra e nos discursos dos alunos dessas professoras, fazendo a mesma pergunta a eles. A noção de palavra para os escreventes, então, remete aos modos de enunciação – ora generalizado para a escrita ora generalizado para a fala. Assim, o autor apresenta que “Fruto desse trânsito [entre o oral e o escrito], as crianças parecem detectar vínculos entre o que lhes chega como o fenômeno ‘palavra’ – em contexto escolar e fora dela – e a própria complexidade da linguagem” (CHACON, 2010, p. 180). As crianças reconhecem que há palavras na fala, fazendo menção em suas respostas a atos de fala; porém, sobrepõem-se os discursos de professoras e alunos quando se trata da palavra na escrita. A ênfase dada pelos escreventes à noção de palavra vinculada à escrita é a mesma detectada nos discursos das professoras, que se fundamentam no critério ortográfico para explicar a noção de palavra.

Vimos que, nos discursos pedagógicos veiculados por materiais didáticos e, também, no discurso de professores e de alunos, não são estabelecidos os mesmos critérios para a noção de palavra. Essas flutuações nos critérios de definição de palavra podem ser encontradas mesmo nos estudos linguísticos como os de filologia de Camara Jr. Com efeito,

em seu dicionário de fatos gramaticais, o autor define *palavras* como “Vocábulo provido de uma significação (v.) externa, concentrada em seu radical (v.). Ou, noutros termos, vocábulo provido de um semantema.” (CAMARA JR., 1956, p. 158). Porém, além de critérios semânticos, posteriormente, em outra definição, o autor se apoia em critérios gramaticais: “as palavras são, essencialmente – a) os nomes, na sua tríplice função de substantivos, adjetivos e advérbios, e – b) os verbos.” (*idem, ibidem*). Em seguida, há uma crítica ao uso irrestrito do termo *palavra* pelos gramáticos da época ao invés de *vocábulo*, quando meramente relacionado a questões fonológicas.

Para o termo *vocábulo*, ainda no mesmo texto, o autor reconhece que essa é uma noção complexa, uma vez que, em seu interior, identificam-se segmentos menores com significação. Em outras palavras, o *vocábulo* seria uma unidade da língua que poderia ser dividida em unidades menores de significação. A explicação para o que seja *vocábulo* é fundamentada em três planos da língua, na escrita e na fala. Com efeito, (1) no plano fônico, são esmiuçadas as características fonético-fonológicas dos *vocábulos* no português brasileiro referentes: ao número de sílabas; à acentuação e posição do acento nos *vocábulos*; aos processos e às regras que envolvem o sistema fonológico do português brasileiro. Já (2) no plano gramatical, o autor tece sua explicação distribuindo os *vocábulos* em categorias ou classes gramaticais. (3) No plano semântico, o autor faz uma separação entre *vocábulos* de significado e *vocábulos* gramaticais. (4) Na escrita, eleger como ponto principal de identificação do *vocábulo* o espaço em branco. E, (5), por fim, na fala, Camara Jr (1972) configura o *vocábulo* como uma série fônica que engloba o contínuo das sílabas:

As séries fônicas que englobam o contínuo de sílabas, entre duas pausas, se subdividem em VOCÁBULOS FONÉTICOS, assinalados por uma delimitação típica ou JUNTURA, que corresponde, conforme a língua, a variados fenômenos fonéticos (CAMARA JR., 1972, p. 84 – destaques do autor).

Vimos, então, que o autor – por ter ciência da complexidade da noção do que seja um *vocábulo* – busca explicações em variados planos da língua e nos modos de enunciação

escrito e falado, dando indícios de que, para tal definição, seria necessário considerar mais de um critério explicativo.

Essas explicações nos diferentes planos e modos de enunciação da língua foram uma tentativa de sanar o que, para o autor, era uma falha metodológica, pois, apesar de ser muito utilizada a distinção da unidade de palavra, não se costumava definir em que se consistia essa unidade (o que continua acontecendo):

O mau hábito das nossas gramáticas de partir explícita ou implicitamente da língua escrita [...] não as deixou tomar consciência dessa falha metodológica, porque na escrita se entende por <vocábulo> o conjunto de letras que fica entre dois espaços em branco (CAMARA JR., 1971, p. 34).

Essa maneira de definir palavra por um critério gráfico – “conjunto de letras que fica entre dois espaços em branco” – se mostra insuficiente e encontra uma série de problemas, como aponta o autor. Entretanto, é a definição que mais encontramos nos discursos de professores, em explicações de materiais didáticos e em manuais gramaticais.

De acordo com Camara Jr (1972), há necessidade de distinguir palavra morfológica de palavra fonológica, já que, em muitos casos, não há coincidência entre essas duas noções de palavra – há palavras morfológicas que correspondem a mais de uma palavra fonológica e vice-versa. Em virtude dessa falta de correspondência termo a termo das palavras fonológicas com as palavras morfológicas, são necessários critérios linguísticos para diferenciá-las, tais como: os fonológicos, os morfossintáticos e os (orto)gráficos. Com base nesses critérios, seguindo Camara Jr. (1972), é que trataremos, a seguir, a discussão do conceito de palavra: na subseção 1.1.1 do ponto de vista fonológico; na subseção 1.1.2 do ponto de vista morfossintático; na subseção 1.1.3 do ponto de vista (orto)gráfico; e, ainda, na subseção 1.1.4 apresentaremos as dificuldades encontradas para a definição de um conceito de palavra.

1.1.1 A noção de palavra do ponto de vista fonológico

A noção de palavra é uma questão complexa – não só do ponto de vista teórico, mas, também, do ponto de vista de discursos e práticas pedagógicas (como vimos) – pela dificuldade de saber qual(is) unidade(s) pode(m) ser considerada(s) palavras, bem como definir e justificar os critérios pertinentes para essa identificação (BASILIO, 2004).

Na tentativa de conceituação de palavra, “a noção do vocábulo assenta na identificação parcial que se faz dentro da diferença global das frases” (CAMARA JR., 1972, P. 87), identificação que ocorre por meio de critérios diferenciados na escrita e na fala. Ainda para esse autor, deve haver uma distinção entre: (i) as unidades do vocábulo fonológico – divisão espontânea na cadeia de emissão vocal; e (ii) o vocábulo significativo – segmento fônico individualizado em função de um significado específico que lhe é atribuído na língua (CAMARA JR, 1972). Embora relacionados, os vocábulos fonológicos e significativos são distintos e podem não coincidir.

Outra diferença entre o vocábulo significativo e o fonológico está em suas fronteiras – sejam elas limites gráficos na escrita ou pausas na fala. Na escrita (entre os vocábulos significativos), há espaços em branco ou hifens; já na fala (entre os vocábulos fonológicos), a pausa é marcada por uma interrupção na emissão da voz.

O vocábulo, encarado meramente do ponto de vista da fonação ou seja – o vocábulo fonético, é êsse corpo sensorial, e nada mais. Do ponto de vista da fala, é, ao contrário, significado, símbolo, representação, ou seja – vocábulo significativo (CAMARA JR, 1972, p. 85).

A falta da coincidência entre o elemento na fonação e o elemento na escrita fez com que alguns linguistas negassem que o vocábulo significativo (ou vocábulo propriamente dito) fosse uma entidade natural linguística e, por isso, o consideravam como qualquer coisa de convencional, imposto à nossa consciência pelas formalidades do ensino e da língua escrita (CAMARA JR, 1972, p. 86).

O grande argumento para sustentar o caráter supostamente convencional do vocábulo é a falta de coincidência entre êle, como elemento significativo, e o vocábulo meramente fonético (CAMARA JR, 1972, p. 87).

Na fala, as pausas são marcadas por grupos de forças, que são sequências de vocábulos sem pausa (CAMARA JR, 1971, p. 63). Desse modo, não há pausa dentro de um sintagma, mesmo havendo dois ou mais vocábulos fonológicos; como ocorre, por exemplo, no sintagma nominal “a bela doce menina” em que há três vocábulos fonológicos, mas não há pausa entre eles por fazerem parte de um único grupo de força (CAMARA JR, 1971, p. 34).

A discussão sobre a distinção do vocábulo fonológico, proposta por Camara Jr (1972), é retomada por Bisol (2000, 2004, 2005), para quem uma das complexidades de se definir palavra é por haverem formas lexicais átonas, chamadas de clíticos – por exemplo: *e*, *me*, *de*, *o*. A complexidade instaurada se deve ao fato de os clíticos serem considerados estruturas dependentes (não possuem acento) e, por isso, marcam um lugar problemático na noção de palavra.

Esse lugar problemático é marcado quando os vocábulos significativos não possuem acento prosódico, pois dependem de um vocábulo acentuado – anterior ou posterior – para formar um grupo de força.

É ainda a falta de coincidência entre o vocábulo fonético e o significativo que levou a gramática grega ao conceito de vocábulos CLÍTICOS, isto é – vocábulos significativos que na enunciação linguística são sistematicamente integrados num vocábulo fonético maior e subordinados ao acento que se dá a êste individualidade fonética (CAMARA JR, 1972, p. 86 – destaque do autor).

O estatuto do clítico tem sido bastante discutido do ponto de vista fonológico e do ponto de vista morfológico. Há muitas discussões sobre a natureza desses elementos, isto é: se os clíticos pertencem à categoria de palavra, à categoria de afixos, ou ainda se formam uma categoria independente.

Os clíticos não pertencem a nenhuma classe morfossintática específica, como podemos notar nos exemplos: *e* = conjunção; *me* = pronome; *de* = preposição; *o* = artigo. Por os clíticos não possuírem acento nem pertencerem a uma classe morfossintática específica,

são considerados formas dependentes – tanto do ponto de vista fonológico quanto do ponto de vista sintático. Do ponto de vista fonológico, os clíticos são dependentes, pois, em razão de não apresentarem acento, apoiam-se no acento da palavra de conteúdo precedente ou seguinte (BISOL, 2005); já do ponto de vista morfossintático são considerados dependentes – e não presos – por não poderem ocorrer sozinhos no enunciado (BLOOMFIELD, 1926).

O comportamento dos clíticos e seu estatuto prosódico são apontados por Câmara Jr. (1988) ao discutir a composição do vocábulo fonológico no português do Brasil. As partículas átonas, pronominais e não-pronominais, “não têm status de vocábulo fonológico” (CAMARA JR, 1988, p.63), uma vez que dependem de um item lexical com acentuação própria para junto com ele formar uma unidade fonológica. Associadas ao vocábulo que as antecedem, as partículas átonas valem como uma sílaba postônica desse vocábulo; quando se associam ao vocábulo que as seguem valem como uma sílaba pretônica.

Outra característica dos clíticos é quanto às vogais que se manifestam neles. Quando o clítico coloca-se na posição postônica, a vogal que lhe pertence enquadra-se no sistema vocálico átono final, ou seja, deve ser uma das três vogais /a, i, u/. No entanto, quando ele se coloca na posição pretônica, observa-se a neutralização entre as vogais médias e as vogais altas, em proveito das altas, comportando-se como se fossem vogais átonas finais (CAMARA JR, 1988).

Conforme Bisol (2005), deve-se considerar o clítico como uma palavra independente que se junta a uma palavra de conteúdo na formação de uma unidade prosódica: o grupo clítico – como em “osporquinhos” ou “mechamaram”, em que os clíticos *os* e *me* se juntaram à palavra de conteúdo seguinte: *porquinhos* e *chamaram*, respectivamente.

Outra evidência que pode comprovar a proposição de existência da unidade *grupo clítico* são as regras de sândi externo, de acordo com as quais há, nas fronteiras das palavras,

uma transformação das estruturas, causada, em geral, pela queda de vogais – conforme em “damiga” (da amiga) ou “seu” (se eu).

O apagamento de uma vogal consiste no que se chama de elisão – como quando há um *a* no final de palavra e a palavra seguinte se inicia por uma vogal diferente de *a*, como em “umestrada” (uma estrada) ou “delite” (da elite). Como se vê, a elisão é um processo que não se aplica no interior de um vocábulo, mas se aplica no interior do grupo clítico igualmente ao que ocorre entre palavras fonológicas.

Em uma abordagem teórica explicativa distinta da apresentada com base em Bisol (2000, 2004, 2005), Simioni (2008) apresenta, para o português brasileiro, uma proposta de representação do clítico na estrutura prosódica.

Bisol (2005) levanta evidências de o clítico constituir com a palavra de conteúdo com que se relaciona um constituinte prosódico pós-lexical (o grupo clítico), nos moldes de Nespor e Vogel (1986). Diferentemente, Simioni (2008), apoiada em uma concepção de hierarquia de restrições, propõe que o clítico é anexado a um domínio superior ao da palavra, ou seja, à frase fonológica, colocando-se, assim, contrária à necessidade de um constituinte prosódico que abrigue o clítico e seu hospedeiro.

Outra perspectiva ainda sobre os clíticos é a proposta por Bloomfield (1926), contrastando com a verificação de que eles são formas não acentuadas e, por isso, não livres, porém, não possuem comportamento de morfemas, portanto, não seriam formas presas também. Em decorrência dessa discussão, Câmara Jr (1972) propõe um acréscimo na classificação estabelecida para que pudesse melhor abarcar os vocábulos do português brasileiro:

As formas mínimas podem ser – LIVRES (vocábulos), DEPENDENTES (clíticos ou partículas), PRÊSAS (partes de vocábulos). A sua apreensão constitui a ANÁLISE MÓRFICA, cuja técnica, para cada língua, é delicada e complexa (CAMARA JR, 1972, p. 96 – destaque do autor).

O autor mantém a teoria proposta por Bloomfield (1926) acerca do vocábulo formal, mas amplia a classificação com *dependentes*. Os clíticos apresentam um comportamento totalmente diferenciado das formas livres e presas, por isso, são chamados de formas dependentes por Camara Jr (1972), relacionando a falta de acento e sua necessidade de dependência ao vocábulo acentuado anterior ou posterior.

Mesmo havendo abordagens distintas a respeito do domínio em que o clítico é prosodizado à estrutura prosódica, os clíticos no PB são interpretados seguindo uma mesma tendência, ou seja, parecem ter *status* autônomo em relação à palavra hospedeira.

Por não serem considerados (i) pertencentes a uma classe morfossintática específica; (ii) dependentes do ponto de vista fonológico, pois não têm acento; e (iii) dependentes do ponto de vista sintático, pois não podem ocorrerem sozinhos em um enunciado – os clíticos se mostram como uma das mais recorrentes partículas envolvidas em estruturas de segmentação não-convencional na escrita, não apenas na aquisição (ABAURRE, 1991; CUNHA, 2004; CHACON, 2004), mas perdurando no desenvolvimento do aprendizado da escrita (TENANI, 2009; SILVA, 2014; PARANHOS, 2014).

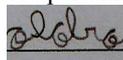
O escrevente não supõe que a escrita é mera transcrição da fala, nem redige em seu texto diretamente a variedade linguística que utiliza. Ele faz hipóteses sobre como ocorrem: (a) a distribuição de espaços em branco; e (b) as relações entre enunciados falados e escritos; ou seja, como a escrita é constituída (TENANI, 2008, p. 240).

Pode-se, pois, depreender do que expusemos até aqui que a noção de palavra é uma questão complexa, pois há dificuldade de se saber que unidades podem ser consideradas palavras, bem como definir e justificar os critérios pertinentes para essa identificação. Especialmente porque, embora relacionados, os vocábulos fonológicos e significativos são distintos e podem não coincidir, já que, na fala, seus limites são marcados por grupos de forças e, na escrita, por espaços em branco. Outra complexidade apresentada de se ter uma

noção de palavra é a de haver formas lexicais átonas, chamadas de clíticos, estruturas dependentes do ponto de vista fonológico por não apresentarem acento e, por isso, se apoiarem no acento da palavra de conteúdo precedente ou seguinte.

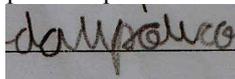
Essas não-coincidências no estatuto da palavra são frequentemente indiciadas nas segmentações não-convencionais de palavras encontradas nos dados da presente pesquisa, como o mostram as ocorrências abaixo. Na ocorrência exposta na Figura 4, a estrutura hipossegmentada corresponde à sequência entre um clítico e uma palavra posterior acentuada: *o lobo*; já naquela da Figura 5, a hipossegmentação corresponde a uma frase fonológica reestruturada, no interior da qual o verbo (*dá*) se junta o grupo clítico *um pouco* (na junção, a coda nasal não foi registrada ortograficamente); por fim, na Figura 6, na estrutura hipersegmentada, a sílaba pré-tônica se mostrou como um clítico, isolado de uma estrutura dissilábica que pode corresponder a uma palavra fonológica da língua.

Figura 04 - Exemplo de segmentação de palavras para menos que a escrita convencional



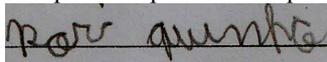
Fonte: AC_12_3C_23M, banco de dados *Escln*.

Figura 05 - Exemplo de segmentação de palavras para menos que a escrita convencional



Fonte: AC_12_3C_23M, banco de dados *Escln*.

Figura 06 - Exemplo de segmentação de palavras para menos que a escrita convencional



Fonte: AC_12_3C_23M, banco de dados *Escln*.

Discutidas, então, as dificuldades do ponto de vista fonológico em se definir um conceito de palavra, daremos continuidade a nossa investigação, na subseção seguinte, buscando o que apresentam sobre esse conceito os estudos morfossintáticos.

1.1.2 A noção de palavra do ponto de vista morfossintático

Os níveis morfológico e sintático da língua, bem como os estudos que se voltam para suas relações – os estudos morfossintáticos –, necessitam de uma noção do que seja palavra por ser uma unidade fundamental para investigações nesses níveis – em sua organização e/ou função. Os manuais gramaticais (que se dedicam aos estudos do funcionamento da língua, seja com o objetivo de prescrever regras ou de analisar como os falantes a usam) têm a mesma necessidade, pois prescrevem/descrevem/analisa regras da língua utilizando esse conceito. Logo, delimitar a unidade *palavra* se torna uma tarefa importante.

O nível morfológico da língua é, hierarquicamente, tratado como o segundo nível da língua – posterior ao nível fonológico. Sandalo (2001) discute que a morfologia é tratada por muitos gramáticos como o principal componente do estudo gramatical e que, então, o nível morfológico da linguagem teria papel fundamental em todos os outros níveis. O objeto de estudo da morfologia são as partes de uma palavra, “mas o que seria uma palavra?” – destaca a autora. Desse modo, nos estudos morfológicos, a dificuldade em se definir palavra se instaura no mesmo ponto que nos estudos fonológicos – os critérios para essa definição.

Os critérios analisados por Sandalo (2001) para se definir palavra se mostram insuficientes ou pouco produtivos e ela os desconstrói com argumentos de que: (1) no critério semântico é possível usar uma palavra ou uma expressão para designar o mesmo significado (construtor/quem constrói); (2) no critério fonológico (como discutido em 2.1.1) há acento

principal e secundário (O que é detergente? É o ato de prender pessoas)⁶; (3) no critério sintático, segundo o qual uma sequência de sons só será uma palavra se for resposta a uma pergunta e se puder ser usada em várias posições sintáticas, ou seja, os clíticos (como *o*, *na*, *de*) não seriam considerados palavras nessa concepção por não serem respostas (como *lá*, *ele*, *sim*) e terem lugares fixos sintaticamente. Assim a autora assume a definição de que “palavra é a unidade mínima que pode ocorrer livremente”, o que significa que nessa perspectiva os clíticos não são considerados como palavras, pois não podem ocupar todas as posições sintáticas de uma frase.

Na *Gramática normativa da Língua Portuguesa*, Lima (1969) tem como proposição para a morfologia tratar da classificação e da flexão das palavras, bem como de sua estrutura e formação (p.10). Para o autor,

A morfologia não examina as palavras uma por uma, tais como encontra-se no Dicionário. Grupa-as em certo número de classes, de acordo não só com a significação senão também com os caracteres de forma e flexões que elas tenham em comum. (p. 63)

Os critérios usados para a explicação do que trata a morfologia são semântico e gramatical, voltando-se para a formação e flexão das palavras. Vale ressaltar que não há – nem na apresentação nem ao longo de todo o texto – uma definição ou explicação do que seja uma palavra ou de como identificar seus limites; a preocupação é exclusivamente com a classificação e flexão das partes de uma palavra, mas sem referir-se ao que ela seja.

Já no nível sintático, a palavra aparece vinculada ao conceito de sentença, como se pode ver, por exemplo, em Berlinck, Augusto, Scher (2001, p. 207), uma vez que a organização da sentença supõe, para os autores, uma combinação de palavras organizadas sob a forma de regras – o que faz da palavra elemento importante para a definição de sentença.

⁶ O exemplo dado pela autora – da palavra *detergente*: sabão líquido; e da expressão *deter gente*: ato de prender pessoas – pode ser explicado porque a estrutura segmental e a sequência de grafemas são iguais, variando-se, então, a estrutura rítmica das duas expressões. Na palavra *detergente*, há um acento principal em “gen” e um acento secundário em “de” seguido por uma sílaba complexa “ter”, que não recebe acento. Já na sequência *deter gente*, os acentos se mostram como seguidos – em “ter” e em “gen”.

Diferentemente, porém, de Sandalo (2001), Berlinck, Augusto, Scher (2001) não estabelecem o que consideram como palavra, nem definem critérios nos quais pudessem se apoiar para conceituá-la.

Na sintaxe, prevalece o valor funcional da palavra como parte da oração e a sua relação de dependência (LIMA, 1969, p. 10). Entretanto – como no capítulo sobre morfologia de sua obra –, não há nenhuma definição ou explicação do que seria uma palavra, apenas se menciona como palavras se relacionam e quais funções podem assumir em uma oração.

Nos estudos morfossintáticos realizados por Bloomfield (1926), a palavra é a forma livre mínima: uma forma que pode ocorrer isoladamente, por si só, constituindo um enunciado, não podendo ser totalmente subdividida em formas livres. Por ser livre, a palavra se distingue dos morfemas presos (radicais, afixos ou clíticos); sendo mínima, como forma livre, distingue-se dos sintagmas oracionais, que podem conter mais de uma forma livre.

Nessa perspectiva de Bloomfield (1926), a palavra é tida como unidade da estrutura do enunciado: dado um enunciado, podemos segmentá-lo em palavras. Entendida dessa forma, a noção de palavra não problematiza a dificuldade prática de reconhecer os limites de uma palavra na fala e na escrita, bem como o problema teórico de se distinguirem palavras de suas diferentes manifestações de caráter flexional.

O critério de Bloomfield (1926) é suficiente para explicar palavras que apresentam correspondência entre fala e escrita; entretanto, uma palavra composta não é totalmente divisível em formas livres. A solução apontada pelo autor para resolver essa dificuldade é a de que a composição é definida como uma construção que contém dois ou mais radicais e não como uma construção lexical baseada em duas ou mais palavras, como encontramos nas gramáticas normativas. Assim, o autor considera que, enquanto formadores de palavras compostas, os radicais correspondentes seriam presos.

Como observa Basílio (2000, p. 11), não há aplicabilidade da definição para os compostos do português: em *sofá-cama*, essa forma mínima – e forma livre – pode se subdividir em duas formas livres: *sofá* e *cama*. A palavra composta *sofá-cama* tem como significado um móvel com duplo funcionamento (um sofá, que se abre e vira uma cama). Se analisarmos as duas palavras isoladamente, haveria dois outros significados distintos, mas que se somam na composição do significado de *sofá-cama*. Assim, as palavras *sofá* e *cama* não têm radicais e significados distintos de *sofá-cama* e, desse modo, “não podemos dizer que a forma, enquanto parte da composição, se caracteriza de modo diferente, como presa, ou que apresenta uma mudança dramática de significado” (BASÍLIO, 2000, p. 11).

Ainda segundo a autora, a palavra é não apenas uma unidade morfológica, mas, sobretudo, uma unidade lexical – o léxico pode ser definido como o conjunto de palavras de uma língua. Os compostos são tidos, então, como conjuntos de palavras que funcionam lexicalmente como uma palavra só. Por fim, para Basílio (2000), há necessidade de se distinguirem as dimensões morfológica e lexical da palavra, além da dimensão gráfica – a qual não poderia deixar de ser mencionada, pela importância que a gramática normativa dá à escrita.

A distinção entre léxico e morfologia se mostra como importante porque as palavras lexicais estão relacionadas ao significado, ao conteúdo referencial, como *amar* ou *gostar*, palavras diferentes porque apresentam significados distintos. Portanto, desse ponto de vista, palavras ortográficas e palavras lexicais se associam por sua relação com o significado – por essa razão, palavras com a mesma estrutura fonológica, como *seção*, *sessão* e *cessão* correspondem a palavras diferentes com significados distintos. Já as palavras morfológicas estão relacionadas à flexão, ao uso de diferentes morfemas. Assim *amamos* e *amei* correspondem, morfológica e ortograficamente, a palavras diferentes, porque têm flexões diferentes que geram diferenças de significados relacionadas aos morfemas, embora o

significado da raiz da palavra seja mantido. Também há questões de grafia que envolvem os morfemas, como em *beleza* e *duquesa*, morfemas com a mesma estrutura fonológica, mas com diferenças morfossemânticas.

A não-coincidência entre dimensões morfológica e ortográfica da palavra é apontada, por exemplo, por Paranhos (2014). Essa autora constata serem as preposições a classe mais envolvida na grafia dos clíticos tanto nas hipo quanto nas hipersegmentações em dados longitudinais – como em Tenani (2011), em dados transversais. A ausência de fronteira gráfica entre palavra e a presença de fronteira gráfica dentro de palavra podem, então, ser relacionadas a conflitos entre as dimensões ortográfica e morfossintática da palavra (PARANHOS, 2014, p. 121).

É interessante observar como alguns gramáticos brasileiros tratam a noção de palavra. Muitas vezes, o conceito de palavra não é colocado em questão; é tratado como natural, mesmo sendo a palavra, no modelo clássico de descrição gramatical, a unidade mínima de análise linguística. É o que observamos, por exemplo, em Lima (1969), em que a noção de palavra não é definida ou explicitada, somente referida nas explicações sobre morfologia e sintaxe.

No entanto, diferentemente desse autor, Pereira (1958), em *Gramática expositiva*, apresenta um conceito para palavra, destacando, nele, o que considera como suas principais características: (i) é tanto uma unidade da fala quanto da escrita; (ii) relaciona uma forma material (som ou letra) a uma ideia ou significação. Vê-se, assim, que essas características consideram o critério físico (da materialidade) e o semântico (da significação) na definição do que seria uma palavra.

A *Nova gramática do Português contemporâneo* apresenta uma explicação confusa ou pouco explícita para a noção de palavra.

1. Uma língua é constituída de um conjunto infinito de frases. Cada uma delas possui uma face sonora, ou seja, a cadeia falada, e uma face significativa, que corresponde ao seu conteúdo. Uma frase, por sua vez, pode ser dividida em unidades menores de som e significado – as PALAVRAS – e em unidades ainda menores, que apresentam apenas a face significante – os FONEMAS. As palavras são, pois, unidades da língua menores que a frase e maiores que o fonema (CUNHA; CINTRA, 1985, p.75).

Um problema dessa definição calcada no “tamanho” de uma palavra é a variabilidade de possibilidades: os monossílabos compostos por apenas uma sílaba; os artigos definidos compostos de um único fonema; e palavras polissílabas de longa extensão. Assim, essa definição se mostra pouco produtiva em decorrência da diferenciação proposta “menor que uma frase/ maior que um fonema”, já que, no português brasileiro, há a possibilidade de palavras com um único fonema e/ou de frases com uma única palavra.

Já em Bechara (2006), detecta-se a preocupação do autor com as não-coincidências na noção do que seria uma palavra:

Na língua escrita, as fronteiras são demarcadas pelos espaços em branco: *se aluga ou aluga-se*. Quando, além da parte fônica (significante), se considera a parte significativa (significado), também há diferenças que devem ser levadas em conta. A questão se põe quando se pergunta quantas palavras existem no conhecido verso de Olavo Bilac: *Cheguei. Chegaste. Vinhas fatigada*. Responderemos que há quatro ou três palavras conforme o conceito por que tomemos do termo *palavra*. Serão quatro se entendermos *cheguei, chegaste* como palavras de significação gramatical diferentes, já que expressam uma o pretérito perfeito do indicativo na 1ª pessoa do singular (*eu cheguei*) e a outra o pretérito perfeito do indicativo na 2ª pessoa do singular (*tu chegaste*). Tais significações se dizem ‘gramaticais’ porque decorrem de classificações estabelecidas na gramática da língua portuguesa. Assim, *cheguei e chegaste* são duas *palavras gramaticais*. Dizemos que só há no referido verso três palavras se entendermos *cheguei, chegaste* como palavras de uma só significação lexical, já que têm em comum o significado lexical “atingir o final de ida ou vinda”, um significado dado pela língua, mas que se situa fora da gramática, pois se trata de uma realidade do mundo em que vivemos. Assim, *cheguei e chegaste* são duas formas, duas ‘flexões’ de uma mesma *palavra léxica*, o verbo ‘*chegar*’ (BECHARA, 2006, 333-334 – grifos do autor).

A constatação da dificuldade em se definir palavra leva, portanto, Bechara (2006) a fazê-lo de mais de um ponto de vista. O autor faz uma discussão sobre palavra lexical e morfológica, mostrando que “*cheguei e chegaste*” são palavras morfológicas distintas, chamadas por ele de palavras gramaticais e que essas podem ser consideradas flexões de uma mesma palavra lexical. Do ponto de vista ortográfico, o autor trata sobre os espaços em

branco que marcam as fronteiras das palavras, porém, mesmo usando o exemplo “aluga-se” não menciona que a fronteira de palavra também pode ser marcada por hífen, como discutido por Tenani (2011).

Nicola (2004) também observa que para se ter uma noção de palavra é preciso analisar o conceito de mais de um ponto de vista. O autor destaca características ortográficas da palavra:

Inicialmente, gostaríamos de esclarecer que a grafia de uma palavra pode estar ligada à sua história (ou seja, à sua etimologia) ou à sua sonoridade (priorizando o aspecto fonológico) (NICOLA, 2004, p. 183).

No entanto, a preocupação do autor é com a dimensão ortográfica da palavra e com a questão do tipo de motivação para a ortografia. Ao definir o seu conceito, o autor prioriza aspectos das motivações para a ortografia: etimológicos, quando se refere ao fato de que a escrita de uma palavra depende de sua história, ou seja, de sua formação morfológica e/ou semântica, pois a etimologia remete também a significado; e fonético-fonológicos, quando se refere à sua base material na fala. Mesmo tendo como interesse tratar de questões ortográficas, Nicola (2004) não define o que seria uma palavra desse ponto de vista.

Situação semelhante se observa, também, em Cereja e Cochar (2009):

muitas palavras de nossa língua tem sua grafia definida não por causa de um conjunto de regras coerentes, mas por sua etimologia. [...] São poucas as regras ortográficas existentes, por isso convém conhece-las ou revisá-las. Contudo, em ortografia, é essencial a memória visual, o treino e a consulta ao dicionário sempre que houver necessidade (CEREJA; COCHAR, 2009, p.73).

São acrescentadas características morfológicas e/ou semânticas, quando, a exemplo de Nicola (2004), os autores justificam a escrita de uma palavra por aspectos morfossintáticos:

Você observou que uma palavra pode ser segmentada em letras e sílabas e que essas unidades não são portadoras de sentido. E também que outra forma de segmentação de palavras possibilita obter unidades portadoras de sentido (NICOLA, 2004, p. 96).

Para tratar da segmentação, os autores definem palavra, efetivamente, pela única vez, como “unidades portadoras de sentido”, usando um critério semântico para essa noção, além de um critério (algo ambíguo) de natureza fonológica, ao destacarem a sílaba.

Por fim, em uma perspectiva funcionalista, a gramática de usos de Neves (2000) destaca o papel de algumas palavras como juntores discursivos:

Algumas palavras da língua que pertencem à esfera semântica das relações e processos que atuam especificamente na junção dos elementos do discurso, isto é, ocorrem num determinado ponto do texto indicando o modo pelo qual se conectam as porções que se sucedem. Esses elementos podem ter seu estatuto determinado dentro da estrutura da oração ou dentro de subestruturas dela (caso das preposições, das conjunções subordinadoras e das conjunções coordenadoras), além de poder determinar-se fora da estrutura oracional, ou seja no âmbito textual (caso das conjunções coordenadoras) (NEVES, 2000, p. 601).

Como se trata de uma gramática de usos, é considerado o discurso como unidade de análise. Nessa circunstância, então, o aspecto semântico da palavra é abordado não preferencialmente quanto à significação, mas quanto ao estabelecimento de relações micro e macroestruturais. Também o aspecto sintático é visto no contexto do que Neves (2000) chama de discurso, uma vez que as funções gramaticais destacadas são aquelas que as palavras podem assumir em uma frase ou texto. Porém, a preocupação não está em se definir o que seja uma palavra.

Desse modo, vimos, nesta subseção, que, além de diversos, são bastante flutuantes os aspectos de que se servem linguistas e gramáticos para definirem o que entendem como palavra. É preciso, no entanto, reconhecer que, mesmo com essa flutuação, mostram-se presentes, nos enunciados escritos de crianças, palavras que correspondem às convencionais, mas sem que essa presença implique “supor que elas já compartilhem dos critérios morfossintáticos e semânticos utilizados pelos adultos na identificação das palavras.” (ABAURRE, 1991, p. 02).

Em vista das menções ao critério (orto)gráfico na noção de palavra presente em manuais gramaticais discutidos nesta subseção, apresentaremos, na subseção seguinte, estudos que tratam a palavra de um ponto de vista (orto)gráfico.

1.1.3 A noção de palavra do ponto de vista (orto)gráfico

Olhar para a palavra do ponto de vista da (sua) escrita significa, normalmente, vê-la segundo suas delimitações por meio de espaços em branco.

A palavra gráfica é

a apresentação do vocábulo na escrita (...) pelo critério formal. Deixa-se entre eles, obrigatoriamente, um espaço em branco, porque, mesmo quando sem pausa entre si num único grupo de força, cada um é considerado uma unidade mórfica de per si (CAMARA JR, 1997, 69).

O ponto principal do vocábulo na escrita, então, é o espaço em branco, que ocorre mesmo sem que haja uma pausa num grupo de força, dissociando-se, assim, a fala da escrita.

A noção de palavra (bem como de seus limites) não é normalmente discutida nos estudos da (orto)grafia. É, portanto, quase axiomático que determinada unidade forma uma palavra da língua e expressa algum sentido de acordo com seus usos possíveis no interior de um texto – além de apresentar função sintática específica em relação à construção de uma oração (DONADEL, 2013).

De modo geral, nesses estudos a palavra é definida como um conjunto de letras entre dois espaços em branco (DONADEL, 2011), um item lexical listado em dicionário, ao qual se pode atribuir significado, seguindo a convenção ortográfica vigente. Assim, a palavra ortográfica acaba sendo, segundo essa autora, um artefato cultural, um objeto que se constitui pelo acordo entre aqueles que a usam, sendo considerada como adequada ou inadequada, em algumas abordagens, ou como certa e errada, em outras (DONADEL, 2011, p. 204-205)

A noção de palavra, segundo Donadel (2011; 2013), do ponto de vista da ortografia é decorrente do percurso da escrita. Nessa perspectiva, portanto, a concepção de palavra é subjacente à palavra gráfica. A autora propõe interpretar a escrita como um processo “que se movimenta da fala (língua) para a norma, o que implica uma concepção de palavra que perpassa todos os níveis da gramática e que também se define por assentamentos da ordem da língua escrita” (DONADEL, 2013, p. 6).

A ideia de que a palavra pode ser entendida como fruto da escrita também pode ser localizada em Zorzi (1998). É o que se detecta em sua explicação de que a noção de palavra implica a capacidade de pensá-la “não só do ponto de sua estrutura acústica, mas também a partir de um referencial visual, considerando a forma gráfica que as palavras têm, ou seja, a convenção” (ZORZI, 1998, p. 87).

Sobre a noção de palavra, Ferreiro e Pontecorvo (1996) consideram a necessidade de se criar um conceito para ela no processo de alfabetização. Para as autoras, quando a criança começa a refletir sobre *o ortográfico dentro do alfabético*, ela precisa entender o que é uma palavra para que possa se inserir em práticas letradas escolares – em uma concepção de palavra como sendo apenas as que estão de acordo com a convenção ortográfica do português brasileiro. Diferentemente do que considera Tenani (2008), para quem as palavras grafadas não-convencionalmente podem ser indícios da “representação de dizer, de recortes da realidade que são simbolizados por meio da grafia” (TENANI, 2008, p. 243).

Na concepção explicitada por Ferreiro e Pontecorvo (1996), questões como a da segmentação de um enunciado na escrita vão se impor, já que a criança necessitará dominar minimamente os critérios morfossintáticos, que sustentam a convenção ortográfica, para ser bem sucedida em suas tarefas de escrita. As segmentações não-convencionais corresponderiam, então, segundo o que pensamos, a tentativas, parte de um processo, em que o escrevente está em trânsito para chegar à escrita das palavras de acordo com a convenção ortográfica do português brasileiro.

Embora mesmo antes do início do Ensino Fundamental “as crianças sejam capazes de produzir e compreender enunciados, seu conhecimento lexical é implícito e inconsciente” (BARRERA, 2000, p. 69). Para essa autora, no curso de sua escolarização, a apropriação da noção de palavra pela criança será gradual e dependente de uma reflexão metalinguística que se desenvolve com a escrita. Assim, a noção de palavra é entendida como subjacente ao

domínio da convenção ortográfica vigente, contrariamente ao modo como entendemos, em que assumimos, juntamente com Tenani (2008, p. 234), que “A palavra ortográfica também é objeto de estudo e a grafia de uma palavra não é dada puramente por uma motivação morfológica/ lexical, mas também historicamente constituída”.

Para os estudos da noção de palavra gráfica, é importante discutir o papel do hífen: um sinal gráfico pertencente apenas à escrita que une dois vocábulos formais – que podem ser um único vocábulo fonológico (*comê-lo*); ou dois vocábulos formais unidos, sendo dois vocábulos fonológicos (*tico-tico*), demonstrando o não-isomorfismo entre palavra fonológica e palavra morfológica discutido por Bisol (1996). Essa não-isomorfia, como os estudos em aquisição da escrita sempre procuram demonstrar, costuma causar bastante dúvida para as crianças.

Acerca da grafia de palavras com hífen, Tenani (2011) analisa dados de segmentação não-convencional de palavras de alunos do Ensino Fundamental II, verificando que os casos de hipossegmentações ocorrem na maioria dos dados encontrados. Segundo Tenani (2011, p. 99), quatro tipos de casos são os mais comuns: (i) a presença de hífen em estruturas em que convencionalmente não há (**estava-mos*); (ii) presença de hífen entre duas palavras (**certo-dia*); (iii) ausência de hífen substituída por hipossegmentação (**despistalos*); (iv) ausência de hífen substituída por espaço em branco (**mata los*).

A mesma autora, bem como Paranhos (2014), demonstram que esse fenômeno está atrelado ao processo de letramento dos escreventes, pois, ao se depararem com a escrita formal, eles tentam usar os recursos gráficos que nela detectam, mas sem que, necessariamente, se deem conta de como operam esses recursos na escrita convencional.

Diante do exposto, vimos que os estudos sobre ortografia definem preferencialmente a palavra como um conjunto de letras entre dois espaços em branco. Em outras palavras, o critério principal de identificação de uma palavra na escrita é o espaço em branco. Há, no

entanto, menção dos autores ao fato de que nem sempre esses limites são claros na aquisição da escrita – situação bem ilustradas pelas segmentações não-convencionais de palavras. Diferentemente, porém, de vê-las como erros cruciais, resultantes de um não-domínio, por parte do escrevente, dos critérios morfossintáticos que sustentam a convenção ortográfica, preferimos vê-las, tal qual Tenani (2008), como tentativas que fazem parte da aquisição e do desenvolvimento da escrita, processo para se chegar à escrita das palavras de acordo com a convenção ortográfica do português brasileiro.

1.1.4 A dificuldade de definição de uma noção de palavra

O problema da formulação de uma noção para palavra é discutido na Linguística moderna, pois os estudiosos tradicionais consideravam o vocábulo “qualquer coisa de convencional, imposto a nossa consciência pelas formalidades do ensino e da língua escrita. O elemento significativo seria, a rigor, exclusivamente a própria frase” (CAMARA JR, 1967, p. 86). Por não se ter um consenso acerca da noção de palavra, os estudos tomavam como unidade de análise a frase. Para esse autor, era uma falha os estudos descritivos de Língua Portuguesa considerarem a existência do vocábulo sem definir em que ele consiste (CAMARA JR, 1967, p. 34).

Porém, mesmo que os estudos linguísticos tentem definir uma noção de palavra, o que vimos é que, em diferentes vertentes, as tentativas de definição privilegiam ora características sonoras, ora características funcionais, ora, ainda, características gráficas. Não nos parece haver, ainda, um consenso teórico e/ou metodológico para tal noção.

Na discussão sobre a dificuldade de se conceber o que é uma palavra por escreventes, Tenani (2008) aponta possíveis motivações: (i) a circulação dos escreventes por práticas de oralidade e de letramento; (ii) a representação dos escreventes de sua realidade de linguagem – como destacam Abaurre e Silva (1993), a propósito de o escrevente recortar “não apenas

sua representação da realidade, mas também a linguagem, sistema simbólico através do qual tal representação adquire expressão e materialidade” (ABAURRE; SILVA, 1993, p. 101); e (iii) o processo de aprendizagem das convenções ortográficas.

No entanto, a literatura atual sobre as segmentações não-convencionais na aquisição da escrita tem mostrado que as ocorrências de espaços em branco a mais ou a menos, em desacordo com a convenção gráfica, não se manifestam aleatoriamente. Como observa Tenani (2008, p. 241) “(...) as segmentações não-convencionais nos dão pistas de que o branco pode ser, também, usado pelo escrevente para construir sentidos do texto”. Trata-se, portanto, de outra forma de olhar/analisar as segmentações não-convencionais, não mais como erros e falta de conhecimento das convenções ortográficas, mas como indício do conflito do escrevente com aspectos do falado e do escrito que ele detecta em práticas de oralidade e letramento que o constituem como escrevente.

Nessa mesma perspectiva, para Chacon (2013), flutuações encontradas em segmentações não-convencionais de palavras são marcas tanto da subjetividade do escrevente quanto da heterogeneidade constitutiva da escrita – conforme proposta por Corrêa (2004). Também para Capristano (2004), segmentações não-convencionais de palavras podem ser resultantes de momentos em que os escreventes pressupõe existir “uma relação unívoca entre aspectos prosódicos da fala e fatos de segmentação da escrita” (CAPRISTANO, 2004, p. 157), bem como da imagem que os escreventes supõem ser a escrita valorizada pela escola.

Como a criança reconhece e produz palavras de sua língua primeiramente em seu modo de enunciação falado e como esse reconhecimento pode se assentar em todos os níveis da gramática (fonológico, morfológico, sintático e semântico), ao redigir seu texto, a criança estabelece relações entre palavra escrita e palavra falada, baseando-se, intuitivamente, em critérios fonológicos, semânticos e sintáticos. Os limites entre essas múltiplas facetas da

palavra, como mencionado, muitas vezes, não coincidem; por isso as oscilações, flutuações entre grafias que se prestam a diferentes interpretações:

a palavra escrita não é uma noção de fácil aquisição e, possivelmente, os limites do que seja palavra para o falante não coincidem, em muitos casos, com os limites da palavra escrita (TENANI, 2008, p. 233).

Pode haver um conflito entre a palavra fonológica, a morfológica e a ortográfica. As segmentações não-convencionais de palavras que ocorrem, sobretudo, na escrita infantil podem, então, ser pistas do trânsito do escrevente por esse conflito que ocorre na língua com relação aos limites da palavra:

(...) a análise e a discussão de segmentações não-convencionais ora apresentadas dão bases para tratar também do conceito de palavra, um objeto de reflexão para o escrevente, cujos significados e limites não são dados necessariamente por uma convenção ortográfica, mas são construídos a medida que o escrevente segmenta a realidade e a representa por meio da linguagem (TENANI, 2008, p. 243).

Portanto, como nossa análise se voltará para dados de segmentação de palavras na escrita inicial de crianças, a noção de palavra escrita, neste trabalho, será a de uma construção inacabada, oriunda de hipóteses que emergem de um sujeito que se defronta com as possibilidades de sua língua e oscila entre elas. A flutuação na escrita infantil, quanto à palavra, pode ser, em última análise, um indício da busca pelo próprio estatuto de palavra (CHACON, 2006).

Por meio das reflexões apresentadas, acreditamos que investigações sobre as segmentações não-convencionais de palavras possibilitam tematizar a(s) relação(ões) entre o sistema fonológico, o sistema morfossintático e o sistema ortográfico do português. Vejamos, então, na seção seguinte, como os diferentes estudos abordam as segmentações não-convencionais de palavras.

1.2 A segmentações não-convencionais de palavras

O objetivo desta seção é caracterizar o registro escrito de palavras que não está de acordo com as convenções ortográficas – as segmentações não-convencionais de palavras – bem como explicitar (de uma perspectiva linguística) o modo como entendemos esse tipo de registro. Vale ressaltar que não pretendemos fazer uma exaustiva exposição dos trabalhos que tratam do assunto, mas, sim, apresentar as publicações com as quais conversamos neste trabalho e as recentes considerações feitas acerca das segmentações não-convencionais de palavras.

Vários são os pesquisadores brasileiros e vasta é sua produção científica acerca das segmentações não-convencionais de palavras: Abaurre (1988, 1989), Cunha (2010), Capristano (2004), Chacon (2010), Tenani (2011), Paranhos (2014), dentre outros. O que há de comum nesses trabalhos – e que compartilhamos – é o olhar para as segmentações não-convencionais de palavras como uma busca do escrevente pela escrita convencional, parte da aquisição da escrita.

Durante a aquisição e o desenvolvimento da escrita, os escreventes, em contato com práticas de letramento, desenvolvem hipóteses sobre o que seja uma palavra gráfica e sobre quais são seus limites. Mesmo reconhecendo diferenças entre fala e escrita, os escreventes têm dúvidas de onde ocorrem as fronteiras das palavras – dúvida marcada pelo uso não-convencional dos espaços em branco não-convencionais e do hífen (TENANI, 2011).

O fenômeno da segmentação das palavras é complexo, pois os critérios em que se assenta cada escrevente para segmentá-las decorrem de inúmeros fatores, nem sempre recuperáveis pelo pesquisador. No entanto, frequentemente, os escreventes propõem soluções que parecem baseadas em algum tipo de informação relativa às formas faladas das palavras, percepção que eles já mostram a respeito da organização rítmica e prosódica dos enunciados:

Reconhecer, portanto, que unidade de natureza segmental, como as sílabas (ou sequências ainda maiores), constituem-se em suportes que possibilitam a manifestação do ritmo e da entonação, parece ser condição necessária para que se explique de forma adequada a relação que as crianças estabelecem com a fala ao se proporem a segmentá-la, em diferentes situações (ABAURRE, 1989, p. 5-6).

O fato de o escrevente reconhecer aspectos fonético-fonológicos presentes nas palavras e segmentá-las com base nesses aspectos não significa que ele escreve com base no que fala ou como fala – conforme acreditam muitos linguistas e professores – como discute Abaurre (1988). Para a autora, dizer que o escrevente escreve como fala é uma suposição ingênua e equivocada:

A tarefa que aguarda o aprendiz de escrita é bem mais complexa do que “escrever a fala” e o que é importante registrar é que ele demonstra perceber logo no início tal complexidade [...] Em maior ou menor grau, a convencionalidade começa logo a ser incorporada pelas crianças, e é evidente que quanto maior for seu contato com essas atividades, no ambiente em que vive, mais atentas elas estarão para os aspectos convencionais da escrita (ABAURRE, 1988, p.137).

O escrevente, afetado pela língua, marca em sua escrita a complexidade existente na delimitação de palavras, transitando por critérios ora fonético-fonológicos, ora morfossintáticos ou ortográficos – a depender da atividade proposta e/ou de sua inserção em práticas de letramento (CAPRISTANO, 2007).

A variabilidade de critérios em que os escreventes se apoiam pode ser vista nos resultados obtidos por Cunha (2010, p. 350). A autora aponta, como motivação para ocorrências de espaços em branco além do esperado para a escrita convencional: (i) reconhecimento da sílaba inicial como uma palavra gramatical (na mora); e (ii) preferência por formações de palavras dissílabas e paroxítonas (resol veram). Esses resultados demonstram que os escreventes também segmentam com base em: critérios morfossintáticos (reconhecendo palavras gramaticais na língua); critérios ortográficos quando buscam segmentar como dissílabas (reconhecendo que no português grande parte das palavras são grafadas assim); e também critérios fonológicos ao segmentarem com preferência em palavras paroxítonas – mais recorrente – percebendo o acento secundário da palavra, como o exemplo

citado pela autora “resol veram” em que há convergência de dois critérios, o ortográfico por serem dissílabas as partes e o fonológico por haver mais de um acento na palavra.

Com base no reconhecimento da multiplicidade dos critérios para definição de palavras, tomamos como concepção que as segmentações não-convencionais de palavras são indícios da relação entre fatos do oral/falado e do letrado/escrito (CORRÊA, 2004), ou seja, do modo como o escrevente, em seu texto escrito, projeta características de enunciados falados.

Essa relação entre o oral/falado e o letrado/escrito fundamenta o que Corrêa (2004) propõe como modo heterogêneo de constituição da escrita, de acordo com o qual não há relação de prioridade, ou maior importância, de um membro sobre o outro dessa relação. A fala e a escrita, para esse autor, são tidas como modos de enunciação da língua, não homogêneos, já que se trata da enunciação em sociedades com tradição, ao mesmo tempo, oral e letrada.

Em outras palavras, esses chamados **erros** de escrita podem ser interpretados como pistas do trânsito do escrevente por práticas de linguagem orais/faladas e letradas/escritas. Os indícios desse movimento entre o oral/falado e o letrado/escrito na escrita ressaltam, pois, o caráter constitutivamente heterogêneo da produção escrita (CORRÊA, 2004). Como inicialmente assinalou Chacon (2005) e Capristano (2007) para dados de segmentações não-convencionais de palavras na escrita infantil.

Nessa perspectiva, os textos dos escreventes (de séries iniciais no caso desta pesquisa) permitem ao pesquisador refletir sobre a língua, bem como sobre os processos que subjazem à aquisição e ao desenvolvimento da escrita. Segmentações não-convencionais de palavras seriam, portanto, *marcas linguísticas* que apontam para o funcionamento linguístico da escrita infantil e para processos de subjetivação do escrevente, o que faz desses enunciados importante fonte de pesquisa e de reflexão sobre esse fenômeno (CAPRISTANO, 2007) –

ideia verificada também em Chacon (2005), Tenani (2009) e Paranhos (2014), na medida em que, para esses autores, segmentações não-convencionais de palavras indiciam pistas sobre a interação sujeito/língua, por meio da forma de segmentar as palavras realizada pelos escreventes.

Capristano (2003) compartilha da concepção de que as segmentações não-convencionais de palavras são decorrentes de os escreventes se apoiarem em variados critérios para segmentar palavras, bem como da heterogeneidade constitutiva da escrita proposta por Corrêa (2004). Por meio de uma análise discursiva, a autora identifica pelo menos dois fatores que parecem atravessar as segmentações não-convencionais: um primeiro, mais ligado a aspectos prosódicos e vinculados à representação da gênese da escrita; um segundo, mais ligado à representação que os escreventes teriam da escrita com base no código institucionalizado como discutido por Corrêa (2004) e retomado por Capristano (2003) em dados de escrita infantil (CAPRISTANO, 2003, p. 157). Sobre a representação da gênese da escrita, a autora considera, ainda, que há uma pressuposição dos escreventes de uma relação unívoca entre aspectos prosódicos da fala e fatos da segmentação da escrita. Já sobre a representação da escrita como código institucionalizado, o escrevente estaria adequando sua escrita ao que ele acredita ser o recomendado pelas práticas escolares (CAPRISTANO, 2003, p. 157-158).

Essas representações que podem ser inferidas em dados de segmentação não-convencionais de palavras não são, no entanto, fruto de uma ação consciente dos escreventes em busca de uma noção de palavra, pois “a noção de ‘palavra’ subjacente às convenções escritas não é (...) uma ‘realidade’” (CAPRISTANO; CHACON, 2013, p. 15) para os escreventes em aquisição de escrita. Dito de outro modo, para os autores, a noção de palavra não é uma *realidade*, mas uma busca por *unidades*, na qual os escreventes se deparam com o equívoco e com a dispersão (CAPRISTANO; CHACON, 2013, p. 16).

Chacon (2006) analisa dados de segmentações não-convencionais de palavras e os considera “como resultado do movimento transitivo-dialógico do escrevente por diversas práticas de linguagem, sejam elas orais ou letradas” (CHACON, 2006, p. 1253). Essa reflexão é pautada nas tendências de as segmentações não-convencionais serem fruto da percepção dos escreventes de fronteiras dos constituintes prosódicos e também da percepção dos escreventes sobre a organização ortográfica da escrita.

A diversidade de combinações entre sílaba, pé rítmico, palavra fonológica e grupo clítico (que, nos enunciado orais, pode se marcar por relações de proeminência) atuam na percepção do aprendiz da escrita, podendo influenciar o movimento de reelaboração da segmentação em função de pontos de fronteiras de constituintes prosódicos. Mas, de outro lado, as soluções encontradas pelos aprendizes para a distribuições de espaços em branco podem ser vistas como fruto do trabalho de reflexões desses aprendizes acerca da organização gráfico-visual do fluxo da escrita, influenciado pela sua inserção em práticas de letramento desenvolvidas especialmente (mas no somente) em contexto escolar (CHACON, 2006, p. 1253).

A relação entre as segmentações não-convencionais e aspectos prosódicos também é estudada por Tenani (2004) em dados de desenvolvimento de escrita com alunos de ensino Fundamental II. Para a mesma autora (em TENANI, 2009), as segmentações não-convencionais estão ancoradas em informações prosódicas e letradas, conclusão a que chega Paranhos (2014).

A segmentação de palavras na escrita caracteriza-se pela utilização de espaços em branco ou de hifens (TENANI, 2011), cuja distribuição se baseia em critérios morfossintáticos, gradativamente adquiridos pelos escreventes. Sobretudo na escrita inicial, detecta-se, nas segmentações não-convencionais de palavras, forte vinculação entre constituintes prosódicos da língua e características das convenções ortográficas, pois os escreventes realizam separações além ou aquém das previstas pela ortografia convencional – fatos caracterizados como: (1) hipersegmentações – quando, numa estrutura, há mais fronteiras de palavras do que o definido pela ortografia convencional, como em “por quinho”; (2) hipossegmentações – quando, numa estrutura, há menos fronteiras de palavras que o

previsto pela ortografia convencional, como em “osporquinhos”; e (3) mesclas – quando, em uma mesma estrutura, encontram-se partes em que há hiper e hipo segmentadas⁷, como em “erauma” (CHACON, 2006).

Passamos, a seguir, a tratar dessas três possibilidades de classificação dos dados.

1.2.1 Hipersegmentações

As hipersegmentações podem ocorrer porque os escreventes percebem que a maioria das palavras no português brasileiro é dissílábica e constituída de sílabas compostas com a estrutura CV (consoante-vogal), como, por exemplo, em “casa”, “mesa”, “bola” etc. (ABAURRE, 1991). Assim, quando uma palavra foge a essa característica, os escreventes tendem a segmentar o enunciado em palavras não-convencionais constituídas de duas sílabas, o que, para essa autora, revela pistas de sua reflexão sobre o seu sistema fonológico e sobre a escrita escolar. Assim, a constante presença, nos dados de hipersegmentação, de palavras trissílabas com acento na penúltima sílaba (cf. “por quinho”)⁸, permite considerar que a criança tem a representação de uma forma canônica de palavra, ou seja, que ela identifica como palavra a forma mínima dissílaba paroxítônica (ABAURRE, 1991, p. 208).

Em estudos sobre hipersegmentações, características prosódicas na escrita infantil também são destacadas por Chacon (2005). Com efeito, em um trabalho sobre hipersegmentações de trissílabos, esse autor observa que os escreventes não hipersegmentam aleatoriamente, já que, nos momentos de ruptura,

(...) as crianças parecem tentar plasmar em sua escrita contrastes rítmicos que certamente detectam, sobretudo, a partir do relevo prosódico do acento da oralidade e que, por sua recorrência, padronizam-se sob forma de constituintes prosódicos da própria língua. (CHACON, 2005, p. 82).

⁷ As estruturas categorizadas como mesclas por Chacon (2006) são categorizadas como híbridos em Cunha (2004).

⁸ Exemplos obtidos no banco de dados desta pesquisa *EscIn*.

Dessa forma, vê-se que a hipersegmentação indicia um reconhecimento das características prosódicas da língua; essas estruturas não são, pois, um mero erro.

As características prosódicas presentes nas considerações que investigam hipersegmentações são extraídas do modelo proposto por Nespor e Vogel (1986). As autoras propõem que todas as línguas apresentam características fonológicas comuns, como a organização de seu componente prosódico em constituintes dispostos em relação hierárquica. Nessa hierarquia, um constituinte é *exaustivamente* contido pelo constituinte superior, na seguinte ordem: sílaba (σ) < pé-métrico (Σ) < palavra fonológica (ω) < grupo clítico (C) < frase fonológica (φ) < frase entonacional (I) < enunciado fonológico (U).

Silva e Tenani (2014) verificaram que a maioria dos escreventes começa e termina o Ensino Fundamental II escrevendo com hipersegmentações. Os resultados obtidos pelas autoras mostram que informações fonológicas importantes ligadas aos constituintes prosódicos, como a presença de acento, constituem um lugar privilegiado para inserção de um limite gráfico não-convencional na fronteira onde há proeminência – resultados que corroboram os de Chacon (2005) sobre dados de Ensino Fundamental I.

Dentre esses constituintes, dois deles se mostram como mais diretamente envolvidos nas segmentações não-convencionais: a *palavra fonológica* e o *grupo clítico*. Esse maior envolvimento é detectado, por exemplo, em Paranhos (2014), em dados de Ensino Fundamental II, como também em Chacon (2005) e em Capristano (2007), em dados do Ensino Fundamental I. Outros constituintes se mostram, também, envolvidos nas segmentações não-convencionais, como a sílaba e o pé métrico, apontados também por Silva e Tenani (2014). Mas não apenas. Ao analisar ocorrências de hipersegmentação na escrita infantil, Capristano (2007) salienta o reconhecimento, pelo escrevente, da presença de padrões da escrita, na medida em que “os pontos de corte, nesses dados, parecem ser momentos em que a criança reconheceu unidades que na escrita aparecem com frequência separadas: *em*

bora, na quela ou *em com tado*”, ou seja, a hipersegmentação ocorreria pelo reconhecimento de palavras gramaticais. Ressalta-se, pois, além da ancoragem dos escreventes em características prosódicas da língua em sua escrita hipersegmentada, também sua ancoragem em fatos relacionados com a aquisição das convenções ortográficas.

1.2.2 Hipossegmentações

Também nas hipossegmentações se detecta ancoragem por parte dos escreventes em aspectos fonológicos e gráficos da língua. Cunha (2004), por exemplo – ao analisar dados de segmentações não-convencionais de palavras em textos de escreventes de uma escola pública e de uma escola privada – constata que a maioria dos casos de hipossegmentação envolve a ausência de um espaço em branco entre um clítico (correspondente a monossílabos não-acentuados como pronomes, preposições, artigos, conjunções) e a palavra de conteúdo seguinte. Ainda segundo a autora, por volta da quarta série do Ensino Fundamental⁹, os casos de hipossegmentações tendem a diminuir, pois o maior contato com a escrita possibilita aos escreventes aprenderem a diferença entre o letrado e o oral.

Por sua vez, ao investigarem a ocorrência de hipossegmentações em discurso direto e em outros contextos discursivos, Capristano e Ticianel (2014) observaram que, fora da instância enunciativa do discurso direto, seus dados confirmam os resultados de Cunha (2004). Porém, acrescentam que esses resultados podem mudar quando as hipossegmentações são analisadas no discurso direto. Isso porque os constituintes prosódicos encontrados nas estruturas hipossegmentadas no discurso direto são hierarquicamente maiores, chegando a haver casos de enunciados fonológicos – o que não ocorre fora do discurso direto nos dados

⁹ Esse e os demais estudos referidos foram realizados antes da mudança para o ensino de nove anos, instituído pela lei federal 11.274/06. Entenda como 1º ao 5º ano do Ensino Fundamental I equivalente a 1ª a 4ª séries do Ensino Fundamental I e 6º ao 9º ano do Ensino Fundamental II equivalente a 5ª a 8ª séries do Ensino Fundamental II.

das autoras. Assim, a reflexão apresentada é de que as características prosódicas das hipossegmentações dependem da instância enunciativa da qual elas fazem parte.

Mas desde as primeiras hipossegmentações que ocorrem em textos de escreventes que ingressam no Ensino Fundamental II (e mesmo no final do Ensino Fundamental I) já se pode observar “que os limites das estruturas das crianças sempre coincidiram ou com o início ou com o final de uma palavra escrita da língua – o que aponta para um entrecruzamento de critérios de natureza prosódica e critérios de natureza gráfica na composição dessas estruturas” (CHACON, 2011, p. 260).

1.2.3 Mesclas

Por fim, nas estruturas caracterizadas como mesclas, observam-se, ao mesmo tempo, híper e hipossegmentações, como nas estruturas “e uvou” ou “resol veuir”. Em ambas, há uma palavra gráfica hipersegmentada (respectivamente, “eu” e “resolveu”); porém, parte da hipersegmentação resultante se une à palavra seguinte, formando uma hipossegmentação na mesma estrutura. Também nas mesclas, como ocorre com as híper e com as hipossegmentações,

(...) o que parece ficar atestado (...) é justamente o modo heterogêneo de constituição da escrita, na medida em que é possível atribuir a essas porções de escrita não apenas a ação da (e da reflexão sobre a) organização prosódica da língua mas também a ação das (e da reflexão sobre as) convenções ortográficas do português brasileiro, resultantes do trânsito do escrevente por práticas de linguagem orais e letradas (CHACON, 2004, p. 230).

Outra observação a ser feita sobre as mesclas é a de sua distribuição. Como observado por Chacon (2004), em dados de crianças de primeira a quarta série do Ensino Fundamental I, e por Tenani e Paranhos (2011), em dados de quinta série, as mesclas ocorrem em número muito menor do que as híper e das hipossegmentações.

Tenani e Paranhos (2011) também observam que, na maioria dos textos do Ensino Fundamental II, além de praticamente já não ocorrerem mais mesclas, em apenas um dos dados desse tipo de segmentação não-convencional houve ruptura de sílaba – rupturas bastante comuns nesse tipo de segmentação em textos do Ensino Fundamental I.

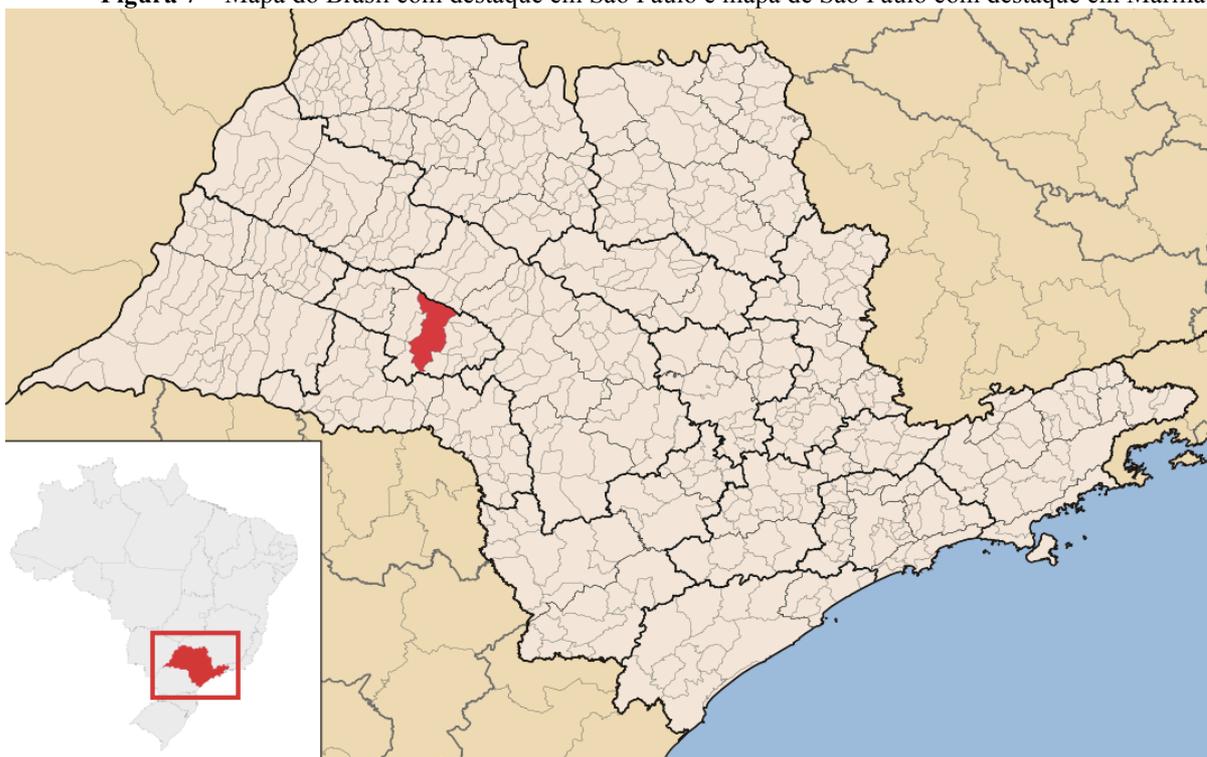
Na descrição dos três tipos de segmentações não-convencionais, segundo a visão que orienta o desenvolvimento da presente pesquisa, deve-se, pois, ressaltar que as estruturas que delas resultam mostram-se como fruto da ancoragem do escrevente tanto em informações que extrai de características da fala (a que, certamente, teve acesso por sua inserção em práticas de oralidade) quanto em informações que extrai de características ortográficas (a que possivelmente tenha tido acesso por sua inserção direta e/ou indiretamente em práticas de letramento).

2 MATERIAL E METODOLOGIA

Destinamos este capítulo a caracterizar nossa pesquisa, apresentando desde o início do processo com o planejamento da coleta dos dados, formando o banco de dados *EscIn*, bem como os critérios de seleção da amostra analisada e a forma de análise dos dados.

Nossa pesquisa foi realizada no município de Marília, situado no Centro-Oeste do estado de São Paulo.

Figura 7 – Mapa do Brasil com destaque em São Paulo e mapa de São Paulo com destaque em Marília



Fonte: wikipedia¹⁰

A escolha dessa localidade se deu em virtude de ser realizada lá grande parte das

¹⁰ Disponível em: https://upload.wikimedia.org/wikipedia/commons/b/ba/SaoPaulo_Municip_Marilia.svg. Acesso em 25/11/2015 às 20:00hrs

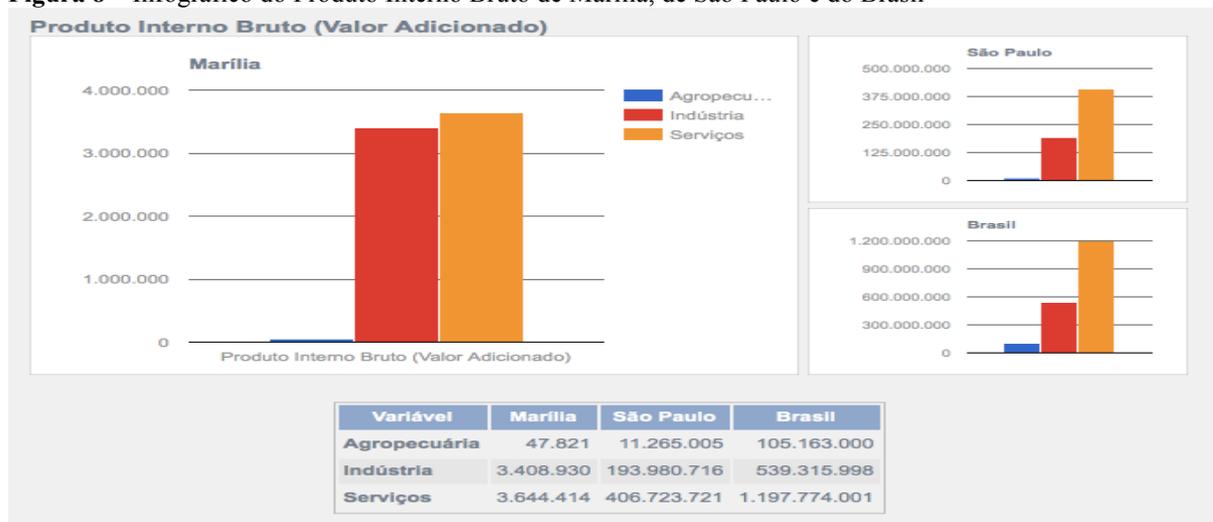
pesquisas e trabalhos do Prof^o. Dr. Lourenço Chacon, orientador desta pesquisa.

A importância de Marília-SP no cenário estadual e nacional se dá pelo perfil do município, que segundo dados do IBGE (Instituto Brasileiro de Geografia e Estatística), divulgado em dezembro de 2010¹¹, mostra-se da seguinte forma:

- População total: 216 745 habitantes.
- Urbana: 207.727 (95,84%)
- Rural: 8.974 (4,16%)
- Mulheres: 112.019 (51,68%)
- Homens: 104.726 (48,32%)
- Taxa de alfabetização: 194 604 pessoas - 95,35%
- População em creches ou escolas: 60 028 pessoas;
- Densidade demográfica (hab./km²): 185,21
- Índice de Desenvolvimento Humano (IDH-M): 0,821
- IDH-R Renda: 0,885
- IDH-L Longevidade: 0,822
- IDH-E Educação: 0,962

A economia de Marília-SP mostra grande força no ramo de serviços e também no ramo da indústria, principalmente a alimentícia.

Figura 8 – Infográfico do Produto Interno Bruto de Marília, de São Paulo e do Brasil

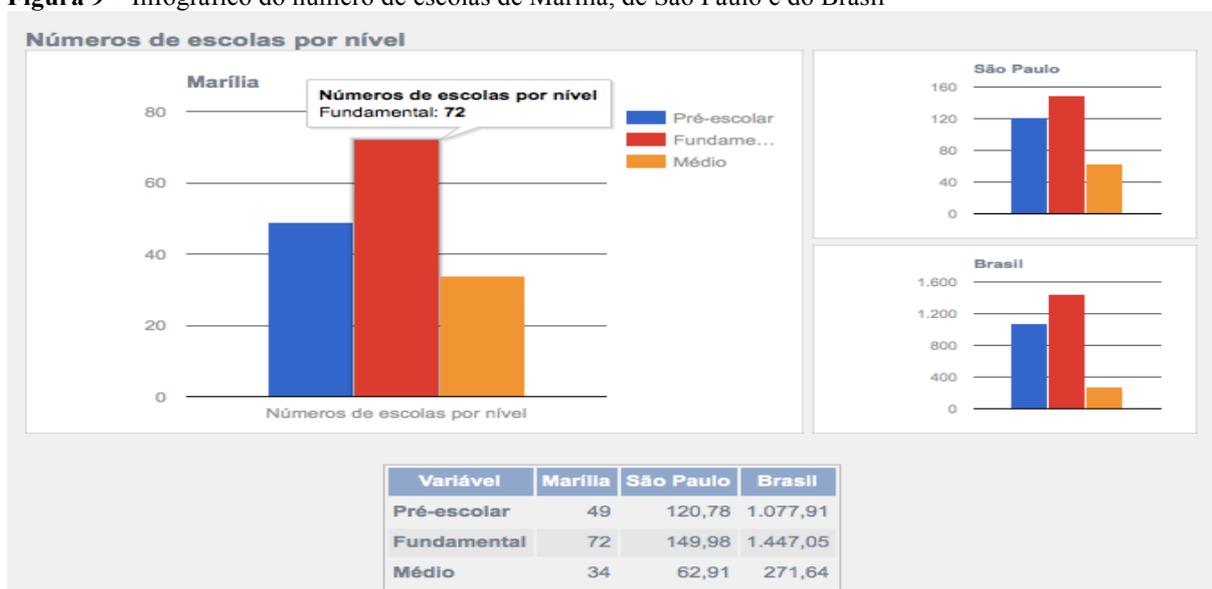


Fonte: IBGE

¹¹ Disponível em: <http://www.cidades.ibge.gov.br/sao-paulo/marilia/sintese-das-informacoes-> Acesso em 26/11/2015 às 10:00hrs.

Quanto ao número de escolas por nível de ensino, o município se assemelha ao padrão do Estado.

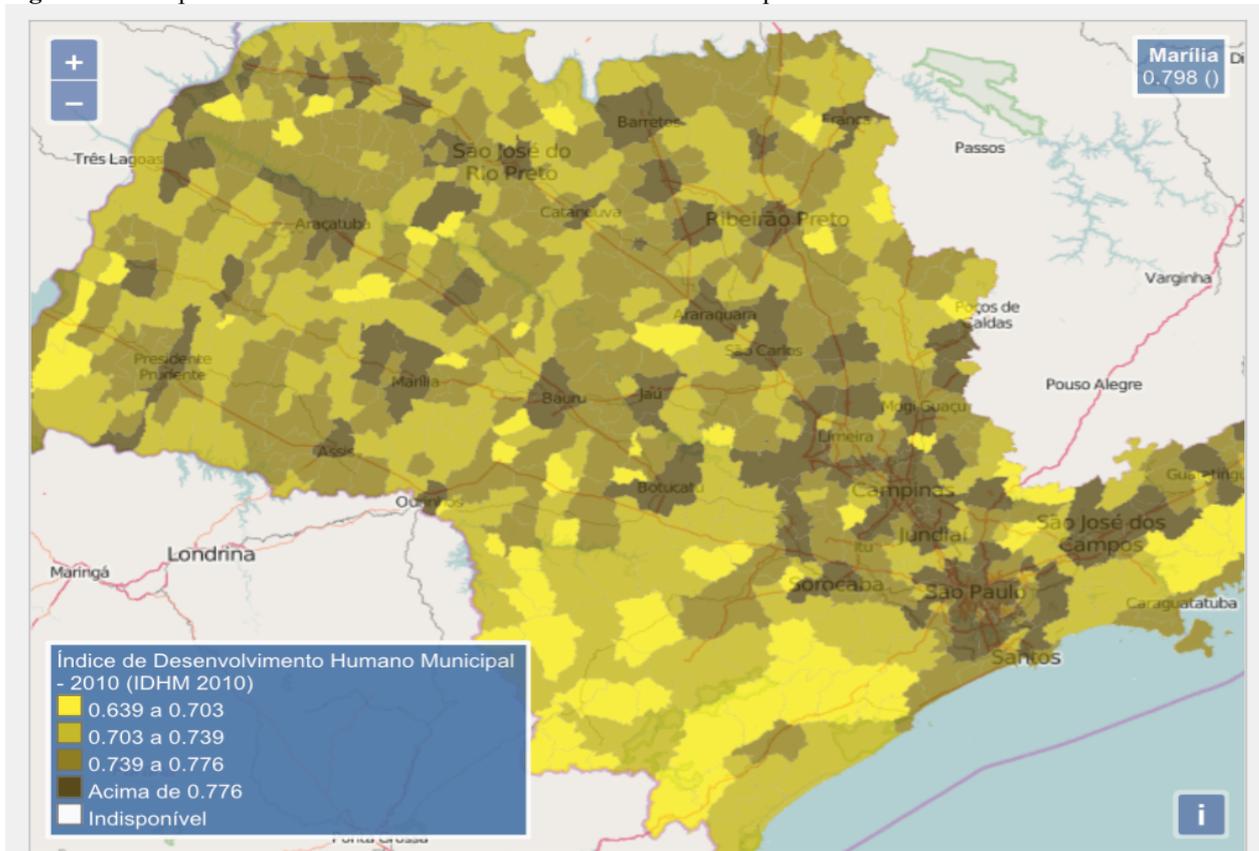
Figura 9 – Infográfico do número de escolas de Marília, de São Paulo e do Brasil



Fonte: IBGE

Quanto ao IDH-M (índice de desenvolvimento humano municipal), Marília destaca-se como parte dos melhores municípios do Estado.

Figura 10 – Mapa do Índice de Desenvolvimento Humano dos municípios de São Paulo



Fonte: IBGE ¹²

Os dados do IDEB (Índice de Desenvolvimento da Educação Básica) 2013 referentes a 4ª série/5º ano¹³ mostram também que Marília-SP possui resultados melhores que os nacionais e estaduais.

Brasil: 5.2; Pública: 4.9 / Privada: 6.7

São Paulo: 6.1; Pública: 5.7 / Privada: 7.3

Marília: 6.5; – não há dados discriminados

¹² Disponível em: <http://www.cidades.ibge.gov.br/cartograma/mapa.sao-paulo|marilia|sinthese-das-informacoes>. Acesso em 26/11/2015 às 8:00hrs.

¹³ Disponível em: <http://ideb.inep.gov.br/resultado/>. Acesso em 07/12/2015 às 10:00hrs.

Pelas razões econômicas e educacionais apresentadas, o município de Marília-SP é representativo de muitas cidades com perfis parecidos no estado de São Paulo ou em outros estados com características similares.

2.1 A concepção e a criação do Banco de Dados *EscIn*

Nossa pesquisa foi pensada a partir de leituras sobre segmentações não-convencionais de palavras e a constatação de haverem poucas informações sobre comparações de tipos de escolas, anos letivos e gênero/sexo dos escreventes. A partir da inquietação concernente a essas variáveis decidimos desenvolver um banco de dados com o maior número de participantes possível para verificar as variáveis e, posteriormente, dar subsídios a outras pesquisas.

Inicialmente, no primeiro semestre de 2012, planejamos o material de apoio para uma proposta textual, para ser entregue aos professores dos sujeitos da pesquisa. A primeira etapa da produção do material foi escolher uma história que contemplasse interesses dos escreventes (estudantes do primeiro ao quinto ano do Ensino Fundamental I), para que pudéssemos estabelecer parâmetros igualitários. Para realizar essa proposta textual, contamos com o auxílio da pesquisadora Prof^a. Dr^a. Cristiane Carneiro Capistrano¹⁴ que construiu um banco de dados longitudinal de textos infantis com quatorze propostas textuais diferentes nos anos de 2000 a 2004. Dentre elas, para a pesquisadora, uma das que mais despertou grande interesse das crianças foi o conto “A verdadeira história dos três porquinhos”¹⁵ de Jon Scieszka. Nossa hipótese para a escolha desse conto foi a de que ele envolveria todos os escreventes – desde os mais novos aos que são mais velhos e estão terminando o Ensino

¹⁴ A Prof^a Dr^a Cristiane Carneiro Capistrano atua na Universidade Estadual de Maringá, no Programa de Pós-graduação em Letras, e é integrante do Grupo de Pesquisa *Estudos sobre a Linguagem* (GPEL/CNPq).

¹⁵ SCIESZKA, J. **A verdadeira história dos Três Porquinhos**. São Paulo: Companhia das letras, 2005.

Fundamental I –, pois ele contesta um conto muito conhecido e querido pelas crianças “Os três porquinhos”.

Após definida essa história, a etapa seguinte foi discutirmos uma proposição da atividade de produção textual concernente com a concepção de escrita assumida pelo GPEL, conforme Corrêa (2004), de que o escrevente deixa em seu texto marcas da relação entre o oral/falado e o letrado/escrito. De acordo com essa concepção, portanto, os erros de escrita podem ser interpretados como pistas dessa relação.

Assim, a atividade proposta consistiu em o escrevente recontar, a seu modo, “A verdadeira história dos três porquinhos”. A proposta (em anexo) foi encaminhada com uma carta de apresentação aos professores contendo esclarecimentos sobre a pesquisa e orientações para a aplicação da atividade.

O material de apoio consistiu, então, de: (1) carta de apresentação e esclarecimento da pesquisa aos professores; (2) Termo de Consentimento Livre e Esclarecido; (3) cópia do conto original “Os três porquinhos”; (4) cópia do conto “A verdadeira história dos três porquinhos”; e (5) folha de redação.

Em seguida, encaminhamos à Secretaria Municipal de Educação de Marília-SP um requerimento para o desenvolvimento da pesquisa. A proposta foi aceita por essa secretaria, que emitiu uma autorização para que fôssemos às escolas e forneceu uma lista com telefones e endereços de todas as escolas do município distribuídas por zonas na cidade. O desenvolvimento desta pesquisa considerou, primeiramente, todas as escolas públicas municipais de Marília – SP, o que totalizaria 19 escolas com 7.633 alunos, discriminados por escola e ano escolar, na tabela a seguir:

Tabela 01 - Relação de escolas públicas municipais de Marília – SP

Nº	Escolas Públicas	1º ano	2º ano	3º ano	4º ano	5º ano	Total
1	Américo Capelozza	119	142	132	138	102	633
2	Antonio Egea	8	14	2	10	4	38
3	Antonio Moral	66	77	94	80	56	373
4	Antonio Ribeiro	139	142	124	139	84	628
5	Cecília Guelpa	99	90	104	109	69	471
6	Célio Corradi	57	63	76	87	55	338
7	Chico Xavier	80	87	70	91	41	369
8	Edméia Braz Sola	93	85	90	86	72	426
9	Geralda C. Vilardi	69	60	71	77	49	326
10	Gov. Mário Covas	89	85	88	99	67	428
11	Isaltino de Campos	72	87	84	72	45	360
12	Myrthes Negreiros	90	83	82	93	39	387
13	Nelson Gabaldi	59	57	56	72	43	287
14	Nicácia Garcia Gil	42	39	63	56	32	232
15	Nivando Mariano	87	70	88	83	90	418
16	Olímpio Cruz	156	153	143	169	90	711
17	Paulo Freire	123	120	112	106	75	536
18	Reny Cordeiro	81	62	81	66	28	318
19	Roberto Cimino	84	57	100	72	41	354

Fonte: Secretaria Municipal de Educação de Marília-SP 2012

A Tabela 02, a seguir, apresenta uma lista de todas as escolas privadas de Marília – SP cedida pela Diretoria Estadual de Educação. Há uma relação de 12 escolas com 2.743 alunos distribuídos por escolas, por ano escolar e por gênero.

Tabela 02 - Relação de escolas privadas de Marília – SP

Escolas Particulares		1º Ano	2º Ano	3º Ano	4º Ano	5º Ano	Total
1	Colégio Água Viva	43	57	39	36	36	211
2	Colégio Assembleiano de Marília ¹⁶	24	11	14	17	14	80
3	Colégio Criativo	58	71	80	70	66	345
4	Colégio Cristo Rei	71	104	85	75	82	417
5	Colégio Esquema Único	36	39	45	31	30	181
6	Colégio Sagrado Coração De Jesus	50	54	71	71	66	312
7	Escola Fundação Bradesco	59	61	69	77	79	345
8	Escola Orbe	56	37	34	34	24	185
9	Sesi - Centro Educacional	64	64	64	64	70	326
10	Instituto Educ. Sagrada Família	11	3	8	5	8	35
11	Colégio CNA De Marília	2	2	3	5	5	17
12	Colégio Bezerra de Menezes	67	76	56	52	38	289

Fonte: Diretoria Estadual de Educação de São Paulo 2012

Ainda no primeiro semestre de 2012, entramos em contato, via telefone, com todas essas escolas do Município, agendando reunião com a direção de cada uma delas para fins de esclarecimento e aprovação do trabalho. Durante esse processo, encontramos muitas dificuldades de acesso a alguns diretores. Nos casos em que os diretores ou vice-diretores não eram encontrados ou não podiam atender as ligações, protocolamos, junto à secretaria da escola, um ofício de requerimento anexado ao projeto da pesquisa. Quando os ofícios não eram respondidos, entrávamos em contatos via telefone novamente. Porém, ainda assim, alguns diretores não foram encontrados, o que resultou em exclusão da escola na participação da pesquisa.

¹⁶ Não houve participação, dessa lista apresentada na Tabela 02, do Colégio Assembleiano de Marília e do Instituto Educacional Sagrada Família por dificuldades de contato com essas instituições. Todos os demais colégios foram contatados, mas alguns não tiveram interesse em participar da pesquisa.

Outra grande dificuldade foi algumas escolas já fazerem parte de outros projetos vinculados a uma universidade ou a estágios de docência. Essas escolas consideraram a participação em mais uma pesquisa como excessiva, já que, a seu ver, essa participação poderia prejudicar o andamento das aulas e o cumprimento dos conteúdos pragmáticos e, por isso, também não fizeram parte da coleta.

A não-aceitação de algumas escolas ocasionou em uma redefinição dos objetivos inicialmente traçados, pois, primeiramente, pretendíamos aplicar a pesquisa em todas as escolas do município; entretanto, a pesquisa não foi aceita por todas elas. Das dezenove escolas municipais, dez aceitaram participar de nossa pesquisa e das dez escolas privadas, cinco aceitaram.

Pretendíamos, também, ter uma amostragem de todas as regiões da cidade; no entanto, não foi possível porque há uma única escola pública em sua Zona Leste, que não aceitou participar, o que acarretou em essa região ser representada apenas por escolas privadas, diferentemente das demais regiões (Norte, Sul, Oeste), por essa razão não pudemos verificar essa variável.

Após esse contato prévio, realizamos uma reunião com a direção e a coordenação de cada escola que se interessou em participar da pesquisa para esclarecermos os objetivos, procedimentos e possíveis dúvidas sobre o trabalho. Com a autorização da direção e da coordenação de cada escola, agendamos uma reunião com os professores para realizarmos os mesmos esclarecimentos. Essas reuniões foram marcadas em horários de estudo coletivo (HEC) para que pudéssemos obter aceitação dos professores para a realização das atividades e, também, saber a quantidade de professores e de alunos para elaborarmos o material de apoio. O prazo de entrega dos textos aplicados pelos professores foi livre, pois tentamos não atrapalhar o andamento das atividades das escolas. Algumas delas combinaram internamente

realizar a atividade em um mesmo dia; outras, na mesma semana; outras, ainda, deixaram livre a cada professor a escolha da data de aplicação da proposta.

No início do segundo semestre de 2012, entregamos cópia do material de apoio a cada professor e aguardamos o contato para o recolhimento dos textos produzidos pelos escreventes – que ocorreu nos meses de outubro, novembro e dezembro daquele mesmo ano.

Os textos recolhidos foram arquivados por turma, ano e escola¹⁷. Os critérios de sistematização das variáveis de organização das informações do banco são os seguintes: (1) tipos de escolas (públicas/privadas); (2) anos letivos (do primeiro ao quinto); (3) gênero/sexo (meninos/meninas); (4) zonas urbanas (norte, sul, leste e oeste), além de uma escola de zona rural¹⁸.

2.1.1 Caracterização do Banco de Dados *EscIn*

Na Tabela 03, a seguir, são apresentados os números e os percentuais de todos os alunos da rede pública e, mais especificamente, dos alunos e das escolas que participaram da pesquisa.

Tabela 03 - Número e percentual de participantes de escola pública (2012)

Número de alunos/ Ano	1º ano	2º ano	3º ano	4º ano	5º ano	Total
Número total de alunos do município	1613	1573	1660	1705	1082	7633
Número de alunos participantes	876	872	865	924	603	4140
Percentual de alunos participantes	54,30	55,43	52,10	54,19	55,73	54,23

Fonte: elaboração própria

¹⁷ As variáveis de turma e zonas urbanas foram excluídas da amostra analisada nesta pesquisa, como explicado nesta seção. Ainda assim, decidimos organizar o banco de dados com essas variáveis, para que, em pesquisas futuras, elas possam, de alguma maneira, ser consideradas de modo a contribuir para a análise dos dados.

¹⁸ Essa classificação como zona rural é realizada pela Secretaria Municipal de Educação, com base em que essa escola se localiza em um distrito do município e atende, principalmente, crianças que moram em propriedades rurais próximas ao local.

Vê-se na Tabela 03 que, em todos os anos escolares, houve participação de mais de 50% das escolas da cidade, o que confere ao *EscIn* maior credibilidade e, por consequência, aos trabalhos de análise realizados com base em seus dados, já que representam pelo menos metade da população escolar do município investigado.

A Tabela 04 apresenta o número total de alunos de escolas privadas de Marília, o número de alunos que aceitaram participar da pesquisa, bem como os percentuais da participação desses alunos em todos os anos escolares.

Tabela 04 - Número e percentual de participantes de escola privada (2012)

Número de alunos/ Ano	1º Ano	2º Ano	3º Ano	4º Ano	5º Ano	Total
Número total de alunos do município	541	579	568	537	518	2743
Número de alunos participantes	267	330	296	265	252	1410
Percentual de alunos participantes	49,35	56,99	52,11	49,34	48,64	51,40

Fonte: elaboração própria

Nas escolas privadas, embora os primeiros, os quartos e os quintos anos não tenham atingido o índice de 50%, estiveram próximos desse percentual. No entanto, o resultado final supera esse índice almejado, permitindo que as comparações entre os contextos públicos e privados sejam realizadas, já que em ambos os tipos de escolas há representatividade maior do que 50% dos alunos de todo o município.

O quadro 01 especifica o total de textos devolvidos pelas escolas que aplicaram a pesquisa – destaque-se que ocorreram problemas com ausências e não entrega de textos. Também é apresentada a distribuição das escolas participantes públicas (PU) e privadas (PV):

Quadro 01 - Distribuição do Banco de Dados por zonas do município

PU	Américo Capelozza	553	}	Total 3.020
PU	Cecília Guelpa	159		
PU	Mário Covas	351		
PU	Antônio Ribeiro	520		
PU	Myrthes Negreiro	300		
PU	Nicácia Gil	174		
PU	Nivando Mariano	221		
PU	Paulo Freire	439		
PU	Chico Xavier	269		
PU	Antônio Égea	34		
PV	Sagrado Coração	321	}	Total 1.210
PV	Água Viva	158		
PV	Esquema Único	115		
PV	Cristo Rei	395		
PV	Bezerra de Menezes	221		
Total geral de textos		4.230		

Leia-se: PU: escola pública; PV: escola privada.

Fonte: elaboração própria

Na organização interna que fizemos do banco, os textos foram separados por zona do município, escola e turma, seguindo a distribuição vinda da escola. Entretanto, como já explicamos, a variável zona não pode ser analisada por não haver participação de escolas públicas e privadas em todas as regiões da cidade. Problema semelhante houve com a variável turma, pois não havia um número comum de turmas nos anos em todas as escolas, por isso não consideramos essa uma variável possível de ser analisada quantitativamente por não ser possível fazer comparações. Porém, mantivemos a separação dessas variáveis no banco de dados para realizações de pesquisas e reflexões futuras.

Na tabela 05, é apresentada uma síntese do total final dos textos efetivamente pertencentes ao Banco de Dados *EscIn* distribuídos pelos critérios de tipos de escolas, anos letivos e gênero/sexo dos escreventes, apresentando discriminadamente o número de textos em cada variável e o percentual relativo ao total geral de textos do banco *EscIn*.

Tabela 05 - Distribuição do número de textos final do *EscIn* pelos critérios de tipos de escolas, anos letivos e sexo/gênero dos escreventes

Ano gênero/sexo		1°		2°		3°		4°		5°	
		F	M	F	M	F	M	F	M	F	M
Escolas Públicas	N°	317	346	279	312	335	327	174	193	380	357
	%	7,49	8,18	6,59	7,37	7,92	7,73	4,11	4,56	8,98	8,44
Escolas Privadas	N°	117	126	147	114	108	110	122	110	119	137
	%	2,76	2,98	3,47	2,69	2,55	2,60	2,88	2,60	2,81	3,24
Totais por gênero/sexo	N°	434	472	426	426	443	437	296	303	499	494
	%	10,26	11,16	10,07	10,07	10,47	10,33	7	7,16	11,80	11,68
Total por ano	N°	906		852		880		599		993	
	%	21,42		20,14		20,80		14,16		23,47	

Leia-se: F: feminino; M: masculino.

Fonte: elaboração própria

O Quadro 02 mostra o resultado final da coleta de textos do Banco *EscIn*, comparando o número de textos das escolas públicas com o número de textos das escolas privadas:

Quadro 02 - Número total de textos do Banco de Dados *EscIn* provenientes de escolas públicas e de escolas privadas

Banco de Dados <i>EscIn</i>		
Tipo de escola	Número de textos	Percentual
Pública:	3.020	78,39
Privada:	1.210	28,61
Total	4.230	100

Fonte: elaboração própria

A comparação entre os números de textos mostra que o acervo é composto de 2,49 vezes mais de textos de escolas públicas em relação ao número de textos de escolas privadas,

diferença que acompanha bem de perto aquela entre o total de alunos de escolas públicas e privadas de todo o município, a saber: 2,78.

Por fim, o quadro 03 apresenta o resultado final do número de textos de escreventes femininos e o número de escreventes masculinos no banco:

Quadro 03 - Número total de textos do Banco de dados *EscIn* de escreventes femininos e de escreventes masculinos

Banco de Dados		
Gênero/sexo	Número de textos	Percentual
Femininos:	2.098	49,60
Masculinos:	2.132	50,40
Total	4.230	100

Fonte: elaboração própria

Há no *EscIn* equilíbrio entre o número de textos de escreventes femininos e escreventes masculinos, já que a diferença (levemente favorável aos meninos) é de apenas 0,98. Não é possível comparar esses dados com a realidade total do município, como fizemos com o tipo de escolas, pois não há separação de gênero/sexo nos dados da Secretaria municipal de Educação.

2.2 A constituição da amostra

Diante da impossibilidade de analisar quantitativa e qualitativamente todos os 4.230 textos que compõe o *EscIn*, realizamos um recorte representativo da totalidade dos textos presente no banco. Esse recorte obedeceu a um critério de 10% do total de textos de cada variável, considerado estatisticamente representativo.

A ordem estabelecida de distribuição da quantidade de textos foi por: (a) zona do município (Norte, Sul, Leste, Oeste e Rural); (b) tipo de escolas (públicas ou privadas), obedecendo a uma ordem alfabética de seus nomes; (c) anos letivos (1º ano, 2º ano, 3º ano, 4º ano e 5º ano); (d) turmas (A, B, C, D etc.), conforme a quantidade que havia em cada escola;

(e) gênero/sexo dos escreventes (feminino e masculino). Lembramos novamente que as variáveis zona e turma não foram analisadas por não terem um padrão para comparação, mas foram mantidas para outras possíveis análises futuras.

Os textos que constituíram a amostra da pesquisa – obedecendo-se ao atendimento de 10% do total em cada variável, considerado estatisticamente representativo – foram escolhidos por sorteio (por meio de ferramenta do Excel) conforme a lista de presença em ordem alfabética dos escreventes em cada turma. O quadro 04 mostra a organização estabelecida para realizar o sorteio da amostra, obedecendo o padrão do banco:

Quadro 04 - Organização do Banco de Dados e sorteio da amostra

Zona	Escola	SÉRIE TURMA	FEMININO	Tamanho da amostra	Nº sorteio	MASCULINO	Tamanho da amostra	Nº sorteio	TOTAL
------	--------	----------------	----------	--------------------------	---------------	-----------	--------------------------	---------------	-------

Fonte: elaboração própria

Já no quadro 05, que se segue, mostra-se a distribuição dos textos em função das variáveis acima descritas.

Assim, como critério de seleção dos textos, consideramos o percentual de 10% somente nas variáveis: anos letivos e gênero/sexo. Por exemplo: na EMEF Américo Capelozza, foram coletados 119 textos no primeiro ano, 59 de escreventes femininos e 60 de escreventes masculinos. Desse total, foram selecionados por sorteio 12 textos, 6 de escreventes femininos e 6 de escreventes masculinos, assegurando o percentual desejado e arredondando os números fracionados para o número inteiro superior mais próximo.

O sorteio dos textos foi feito com a ferramenta do Excel. De acordo com essa ferramenta, é sorteado entre o 1º e o número máximo de textos. Por exemplo, no caso do primeiro ano, turma A, do colégio Américo Capelozza, há 12 escreventes femininos, dos

quais selecionamos 1 texto. Com base entre o primeiro e o número máximo (12) de textos, o programa sorteou o 4º texto. Então buscamos na lista de presença da turma esse texto de escrevente feminino. Em seguida, o digitalizamos e ele foi incluído como *corpus* da pesquisa.

Quadro 05 - Organização do Banco de Dados e sorteio da amostra

ZONA NORTE	AMÉRICO CAPELOZZA	SÉRIE TURMA	FEM	Tamanho da amostra	Nº sorteio	MASC	Tamanho da amostra	Nº sorteio	TOTAL
ZN	AC	1ºA	12	1	4	7	1	3	19
ZN	AC	1ºB	12	1	2	9	1	9	21
ZN	AC	1ºC	9	1	2	7	1	5	16
ZN	AC	1ºD	7	1	2	8	1	5	15
ZN	AC	1ºE	5	0	-	11	1	2	16
ZN	AC	1ºF	7	1	2	14	1	6	21
ZN	AC	1ºG	7	1	7	4	0	-	11
ZN	AC	1º	59	6	-	60	6	-	119
ZONA NORTE	CECÍLIA GUELPA	SÉRIE	FEM	Tamanho da amostra	Nº sorteio	MASC	Tamanho da amostra	Nº sorteio	TOTAL
ZN	CG	1º A	8	1	3	7	1	5	15
ZN	CG	1º C	5	0	-	15	2	4; 10	20
ZN	CG	1º	13	1	-	22	3	-	35
ZONA NORTE	SAGRADO CORAÇÃO	SÉRIE	FEM	Tamanho da amostra	Nº sorteio	MASC	Tamanho da amostra	Nº sorteio	TOTAL
ZN	SC	1º A	10	1	7	11	1	9	21
ZN	SC	1º B	8	1	2	17	2	9; 15	25
ZN	SC	1º	18	2	-	28	3	-	46

Fonte: elaboração própria

A tabela 06 sintetiza a distribuição do número de textos da amostra pelos critérios de tipos de escolas, de anos letivos e do gênero/sexo dos escreventes, apresentando discriminadamente o número de textos em cada variável e o percentual relativo ao total geral de textos da amostra:

Tabela 06 - Distribuição do número de textos final da amostra estratificada pelos critérios de tipos de escolas, anos letivos e sexo/gênero dos escreventes

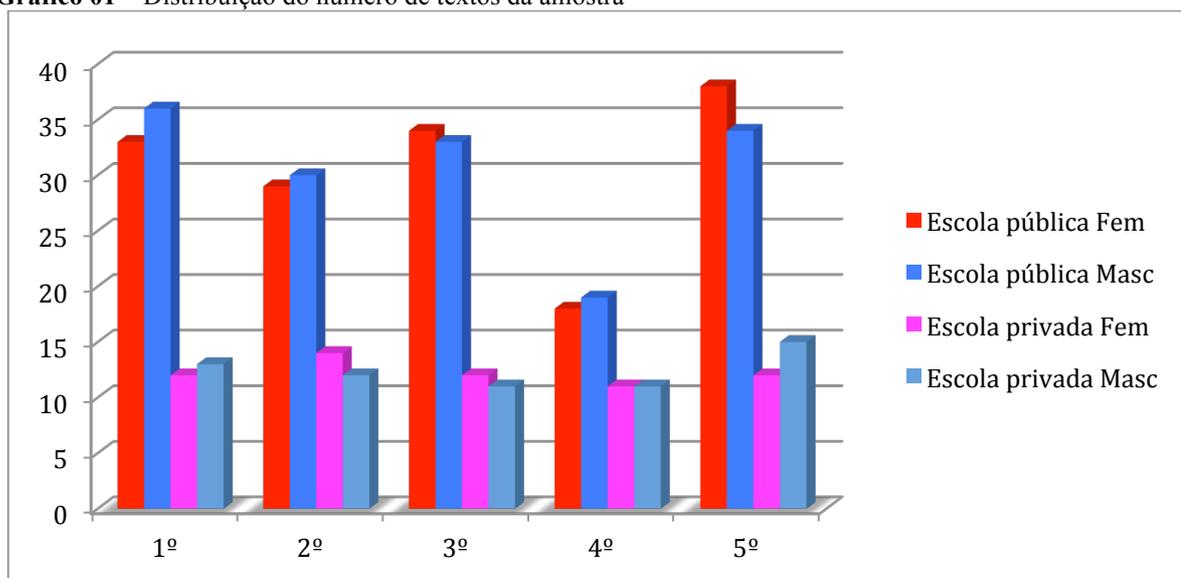
Ano	gênero/sexo	1°		2°		3°		4°		5°	
		F	M	F	M	F	M	F	M	F	M
Escolas Públicas	N°	33	36	29	30	34	33	18	19	38	34
	%	7,73	8,43	6,79	7,02	7,96	7,73	4,21	4,45	8,90	7,96
Escolas Privadas	N°	12	13	14	12	12	11	11	11	12	15
	%	2,81	3,04	3,28	2,81	2,81	2,58	2,58	2,58	2,81	2,51
Totais por gênero/sexo	N°	45	49	43	42	46	44	29	30	50	49
	%	10,54	11,47	10,07	9,84	10,77	10,30	6,79	7,02	11,71	11,47
Total por ano	N°	94		85		90		59		99	
	%	22,01		19,91		21,08		13,82		23,18	

Leia-se: F: feminino; M: masculino.

Fonte: elaboração própria

O gráfico 01 ilustra as informações da distribuição dos textos da amostra presente na tabela 06, mostrando o quão semelhante e regular é a quantidade de textos de escreventes femininos e de escreventes masculinos, tanto das escolas públicas quanto das escolas privadas.

Gráfico 01 – Distribuição do número de textos da amostra



O quadro 06 sintetiza a quantidade final do número de textos da amostra pelo critério de tipos de escolas.

Quadro 06 - Número total de textos da amostra estratificada de escolas públicas e de escolas privadas

Amostra Estratificada		
Tipo de escola	Número de textos	Percentual
Pública:	303	70,96
Privada:	123	28,80
Total	427	100

Fonte: elaboração própria

A diferença de proporção da comparação entre os números de textos de escolas públicas e o número de textos de escolas privadas segue uma consonância entre o município (2,78), o banco de dados *EscIn* (2,49) e a amostra (2,46).

O quadro 07 sintetiza a quantidade final do número de textos da amostra pelo critério de gênero/sexo nos anos letivos.

Quadro 07 - Número total de textos da amostra estratificada de escreventes femininos e escreventes masculinos nos anos letivos

Ano gênero/sexo		1°		2°		3°		4°		5°	
		F	M	F	M	F	M	F	M	F	M
Totais por gênero/sexo	N°	45	49	43	42	46	44	29	30	50	49
	%	10,54	11,47	10,07	9,84	10,77	10,30	6,79	7,02	11,71	11,47
Total por ano	N°	94		85		90		59		99	
	%	22,01		19,91		21,08		13,82		23,18	

Fonte: elaboração própria

O quadro 08 sintetiza a quantidade final do número de textos da amostra pelo critério de gênero/sexo.

Quadro 08 - Número total de textos da amostra estratificada de escreventes femininos e escreventes masculinos

Amostra Estratificada		
Gênero/sexo	Número de textos	Percentual
Total de textos de meninos:	213	49,88
Total de textos de meninas:	214	50,12
Total	427	100

Fonte: elaboração própria

A tendência de equilíbrio entre o número de textos de escreventes femininos e de escreventes masculinos apresentada no bando de dados *EscIn* tem a proporção de 0,98, ou seja, levemente favorável aos textos de escreventes masculinos. Essa tendência é mantida na amostra com uma proporção de 0,99. Não foi possível a comparação entre o município e o banco de dados *EscIn*, como fizemos com a quantidade de textos de escolas públicas e privadas, pois a Secretaria Municipal de Educação não tem disponível essa diferenciação, o que nos impediu de realizar essa comparação.

A identificação de cada texto foi realizada como a codificação dos bancos de dados do grupo GPEL, obedecendo a uma sequência de informações, separadas por um sinal de *underline* (), que seguem da mais geral para a mais específica. Primeiramente, apresentam-se duas iniciais do nome da escola na qual o texto foi produzido – por exemplo: AC, para a escola Américo Capelozza, ou AR, para a escola Antônio Ribeiro. Em seguida, os dois últimos dígitos do ano da coleta (12), referente ao ano de 2012, o que é comum a todos os textos. Esse critério foi utilizado em virtude de possíveis trabalhos que visem comparar textos desse banco de dados com textos de outros bancos produzidos em tempos diferentes. Seguem-se depois, juntos, o ano e turma, mesmo a turma não sendo um critério de distinção para as análises. Utilizamos essa informação na identificação do texto para facilitar a organização em consideração a como eles foram escritos, pois cada turma teve uma professora diferente que aplicou a proposta aos seus alunos, o que em uma análise qualitativa precisaria ser considerado. A última informação é referente ao número do aluno na lista de presença da turma. Esse número é formado por dois dígitos, sucedidos de M, para texto de escrevente masculino, e de F, para texto de escrevente feminino. Assim, os códigos de texto são, por exemplo, “AC_12_1A_04F” para um texto da escola Américo Capelozza, do 1º ano, turma A, do quarto aluno da lista de presença (considerando todos os alunos que entregaram textos dessa turma) de escrevente do gênero feminino.

2.3 Procedimentos de análise dos dados

Da rede municipal e da rede privada da cidade de Marília constavam, no ano de 2012, 11.773 alunos. Desses, participaram, escrevendo um texto da constituição do Bando de Dados *EscIn*, 4.230 alunos. Desse total, selecionamos, então, uma amostra de 10%, constituída de 427 textos para a formação do *corpus* a ser analisado neste trabalho. Em cada texto, contamos o número de palavras escritas e o número de ocorrências de segmentações não-convencionais

de palavras e calculamos a média comparativa entre esses dois índices. Realizamos médias ponderadas com esse cálculo em todas as variáveis propostas para obtermos os perfis de cada uma e compará-las. Como esclarecido, consideramos como estruturas de hipossegmentação, hiperssegmentação ou mesclas entre os dois tipos, respectivamente a junção, a separação ou a junção-separação em uma mesma estrutura ortográfica.

A leitura dos textos para a identificação de segmentações não-convencionais requer cautela, pois alguns tipos de letra e a presença de rasuras causam dúvidas quanto à presença ou não de um espaço em branco. Pelo fato de os textos serem manuscritos, é importante salientar que houve uma avaliação pelos pesquisadores deste trabalho quanto aos tipos de distribuição dos espaços em branco. Nos textos em que os alunos usam letra cursiva, é mais clara a junção ou a separação das palavras. Ainda assim, é particular a cada texto essa delimitação, pois alguns escreventes, normalmente os que escrevem com letras grandes, costumam apresentar grandes espaços em branco entre as palavras. Já outros, principalmente os que escrevem com letras pequenas, costumam deixar pouco espaço entre palavras, o que, muitas vezes, dificulta ou, mesmo, inviabiliza uma avaliação, pelo tamanho do espaço em branco deixado pelo escrevente. Optamos por analisar, em particular a cada tipo de letra, se o espaço ou a falta dele era uma separação ou uma junção, de fato, das palavras.

O texto da figura 11 exemplifica como identificamos segmentações não-convencionais em meio às palavras escritas convencionalmente:

Figura 11 - Exemplo de texto com hipo e hiper segmentação

TEXTO

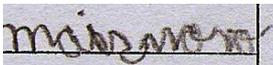
01 A VERDADEIRA HISTÓRIA DOS TRÊS PORQUINHOS
02 O LOBO ETAVA DO ETETI EGERIA FESE UMBELO
03 PARA SUA VOVOSINHA ELE PESAVA DE
04 ACUCA E FOI ATE A CASA DO VISINO
05 SEU VISINO ERA UM POQUINHO SEU NARIZ
06 COMESO SO ELE NO AGETO E PIRO
07 A CASA DESE POQUINHO ERA DE PAPA E
08 MOBO O LOBO ECOTO O POQUINHO
09 MOTINA DI SVA O LOBO A POVE COME
10 O PETOTANO TODINO ELE FOI PARA
11 A SEGUIDA CASA QU ERA MADERO MARI
12 COMES A COSA E DE UM GEDERPIRO
13 E A CASA UM GEDPIRO E A CASA CAIU
14 ELE COME O POQUINHO O LOBO LOBO
15 FOI PARA PARA TEST CASA QU ERA DE TIGOLO
16 ELE DE UM GERPIRO MAS A CASA NA
17 CASA NA CAIU O POQUINHO XIGO A VOVO
18 DELE E ELE FURSO ETETO AROBA A PATA
19 A TA QUE A POLASA XEGO TUDO

Fonte: banco de dados *EscIn* (AR_12_1°F_13F).

No texto da figura 11, encontramos tanto hipossegmentações – “equeria” (*e queria*), “umbelolo” (*um belo bolo*), “gederpiro” (*grande espirro*), “gedepiro” (*grande espirro*), “gerpiro” (*grande espirro*), “eteto” (*e tentou*) – quanto hipersegmentações “a pove” (*aproveitou*), “a ta” (*até*). Para entendermos a letra do escrevente e diferenciarmos as palavras escritas com segmentações convencionais das escritas com segmentações não-convencionais foi necessário analisar cada caso. Se observarmos com mais atenção, por exemplo, na linha 5, a escrita “e piro”, veremos que, embora essa escrita pudesse ser interpretada como uma hipersegmentação de *espirro*, observando-se os exemplos das linhas 11, 12 e 15, vemos que a criança escreve a palavra *espirro* como “piro”, fato que nos levou a interpretar “e piro” como *e espirrou*.

Outra ocorrência, mostrada na figura 12, foi, por nós, categorizada como uma hipossegmentação:

Figura 12 - Exemplo de hipossegmentação

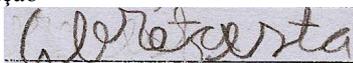


Fonte: banco de dados *EscIn* (SC_12_3ªA_14M).

Nessa ocorrência, o escrevente faz uma “ponte” entre as palavras, o que faz com que a estrutura seja reconhecida como apenas uma e, ainda, permite observar uma volta do sujeito sobre a (sua) escrita, bem como sobre o tamanho do limite entre as palavras.

Temos, na figura 13, mais um exemplo do que analisamos nos textos para considerar o que eram segmentações de palavras convencionais ou não-convencionais:

Figura 13 - Exemplo de hipossegmentação



Fonte: banco de dados *EscIn* (SC_12_2ªA_08F).

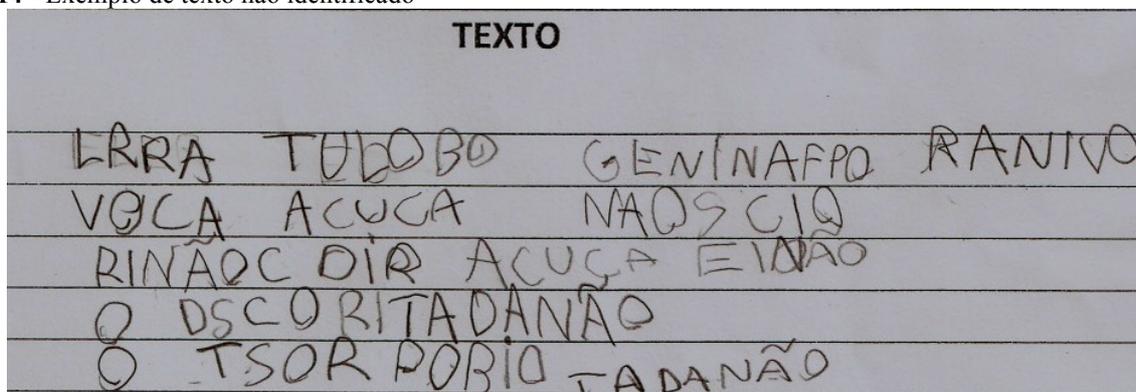
Na figura 13, há uma rasura que dificulta a leitura; porém, é possível ver que há um espaço, uma distância entre o limite das palavras “abre” e “esta” que faz com que essa estrutura não seja considerada uma hipossegmentação, como discute Capristano (2007).

Os casos de letra em caixa alta foram os que trouxeram maior dificuldade por, justamente, não haver uma ligação entre essas letras. Mais uma vez, porém, mantivemos o critério de analisar cada texto em particular para que não houvesse erros por generalizações.

Vale salientar que, além da interpretação dessas caligrafias constituir-se como uma das grandes dificuldades para a análise, algumas estruturas ou textos inteiros não puderam ser lidos. Nesses casos, marcamos, quando o texto todo não foi compreendido, como texto não-identificável. Esses textos fizeram parte da amostra, para que esse fato fosse constatado na análise. No entanto, não puderam fazer parte da análise quantitativa, por não fornecerem os

dados que nos propusemos analisar, como no exemplo do texto da figura 14, no qual não houve reconhecimento de nenhuma palavra da ortografia convencional de português.

Figura 14 - Exemplo de texto não identificado



Fonte: banco de dados *EscIn* (AV_12_1A_02F).

Alguns textos tinham apenas partes não-identificáveis, em que não era possível compreender o que o escrevente tentou redigir. Nesses casos, ignoramos essas partes e consideramos apenas as partes em que era possível reconhecer palavras do português. Para que não houvesse problemas na quantificação, nesses casos de não-identificação parcial, as estruturas não-identificáveis não foram contabilizadas nem como caso de segmentação não-convencional nem como número de palavras, para não haver prejuízo na média ponderada entre esses dois índices.

Assim, nossa análise foi primeiramente, quantitativa, o que – como se verá a seguir – permitiu diferenciar como ocorrem as segmentações não-convencionais de palavras nas escolas públicas e nas escolas privadas. Também por meio de uma análise quantitativa, realizamos uma caracterização de ambos os tipos de escolas, observando a variação de anos letivos (1º ano, 2º ano, 3º ano, 4º ano e 5º ano do Ensino Fundamental I) e a variação de gênero/sexo em textos de escreventes femininos e masculinos.

Essa caracterização numérica mostrou um perfil das variáveis que objetivamos investigar. Com base nos resultados encontrados, fizemos testes estatísticos, a fim de compararmos as variáveis descritas e apreendermos sua significância.

2.3.1 Análise Estatística

A análise estatística descritiva dos dados buscou responder ao objetivo de verificar como as ocorrências se distribuem em função: (a) tipos de escolas (públicas e privadas); (b) dos anos letivos (1º ano, 2º ano, 3º ano, 4º ano e 5º ano); e (c) do gênero/sexo dos escreventes (masculino e feminino).

Assim, quantificamos os números de ocorrências de segmentações não-convencionais de palavras nos textos de escolas públicas e nos textos de escolas privadas, em virtude de comparar e constatar semelhança ou diferença nos resultados obtidos.

Para a análise estatística descritiva, foi feita a leitura e a identificação dos textos, eles foram separados conforme apresentassem, ou não, segmentações não-convencionais. Primeiramente, foram identificadas as estruturas escritas como hipossegmentações, hipersegmentações e mesclas e quantificado o número de suas ocorrências em cada texto. Também foi quantificado o número de palavras de cada texto e, então, feita a razão entre o número de segmentações não-convencionais de palavras e o número de palavras do texto. Por exemplo: se em um texto de 32 palavras houve 4 ocorrências de segmentações não-convencionais de palavras, essa relação correspondeu a um percentual de 12,5% de estruturas não-convencionais nesse texto. Por sua vez, se em outro texto de 138 palavras houve 15 ocorrências de segmentações não-convencionais de palavras, esse percentual foi de 10,38%. Mostrou-se, para nós, importante essa forma de quantificar os dados de segmentação não-convencional para que não fosse criada a falsa impressão de que textos mais longos, mais

extensos apresentariam uma quantidade maior desses dados do que textos mais curtos, menos extensos.

Em seguida, para que houvesse uma média dos percentuais de cada ano, todos eles foram somados e divididos pelo número de textos produzidos no ano, o que gerou uma média de percentual de ocorrências de segmentações não-convencionais de palavras para cada ano.

Realizamos a comparação entre os tipos de escolas por meio de estatística descritiva utilizando a média estabelecida entre o número de ocorrências de segmentação não-convencionais e o número de palavras de cada texto para que não houvesse discrepância em função de o texto ser mais extenso e, por isso, conter mais ocorrências de segmentações não-convencionais.

Verificamos os dados da variação de anos letivos e de gênero/sexo a partir da variação de tipos de escolas, nesse sentido, buscamos o percentual de ocorrências de segmentações não-convencionais de palavras nos textos de 1º ano, 2º ano, 3º ano, 4º ano e 5º ano oriundos de escolas públicas e de escolas privadas, permitindo-nos comparar como é a distribuição das ocorrências ao longo de Ensino Fundamental I e se isso varia em função do tipo de escola.

Verificamos a variação de gênero/sexo quantificando o número de ocorrências de segmentações não-convencionais em textos de escreventes masculinos e em texto de escreventes femininos. Essa variação de gênero/sexo, também, foi subsequente da variação de tipos de escolas, o que possibilitou diferenciar a comparação entre masculino e feminino em escolas públicas e em escolas privadas, mostrando-nos se a diferença de gênero mantém-se ou muda em função do tipo de escola em que se apresenta.

A partir da estatística descritiva dos dados realizamos o tratamento estatístico inferencial com o uso do software Statística (versão 7.0), programa utilizado nos trabalhos do GPEL e que atendia às necessidades desta pesquisa. Realizamos, então, análise descritiva e análise inferencial dos dados.

A análise da primeira variável (descritiva e inferencial) buscou *verificar como as ocorrências de segmentações não-convencionais de palavras se distribuem em função dos tipos de escolas* e para isso foi utilizado o teste paramétrico *T-test* para amostras independentes, na comparação entre as ocorrências de segmentações não-convencionais de palavras em escolas públicas e em escolas privadas. A escolha do teste paramétrico se baseou na verificação da não-violação da curva do teste de normalidade, adotando-se o valor de $\alpha \leq 0,05$.

A análise da segunda variável (descritiva e inferencial) buscou *verificar como as ocorrências de segmentações não-convencionais de palavras se distribuem em função dos anos letivos dos escreventes*, foi utilizado, primeiramente, o teste não-paramétrico *T-test* para variáveis dependentes, comparando cada ano das escolas públicas 1º ano com seu respectivo ano das escolas privadas (1º ano das escolas públicas x 1º ano das escolas privadas; 2º ano das escolas públicas x 2º ano das escolas privadas; 3º ano das escolas públicas x 3º ano das escolas privadas; 4º ano das escolas públicas x 4º ano das escolas privadas; 5º ano das escolas públicas x 5º ano das escolas privadas). Outra análise que vislumbramos na variável ano letivo foi a comparação entre 1º ano, 2º ano, 3º ano, 4º ano e 5º ano das escolas públicas e das escolas privadas. Foi realizada a estatística com teste não-paramétrico *Friedman ANOVA and Kendall Coeff. of Concordance* para variáveis dependentes – devido à violação do teste de normalidade – na comparação entre as amostras: 1º ano, 2º ano, 3º ano, 4º ano e 5º ano.

A análise da terceira variável (descritiva e inferencial) buscou *verificar como as ocorrências de segmentações não-convencionais de palavras se distribuem em função do gênero/sexo dos escreventes em escolas públicas e em escolas privadas*. Na análise da variação de gênero/sexo, primeiro comparamos os dados de escreventes femininos de escolas públicas x escolas privadas e dados de escreventes masculinos de escolas públicas x escolas privadas. A comparação, primeiramente, foi realizada por meio da estatística com teste não-

paramétrico *Friedman ANOVA and Kendall Coeff. of Concordance* para variáveis dependentes – devido à violação do teste de normalidade – na comparação entre as amostras: 1º ano, 2º ano, 3º ano, 4º ano e 5º ano de escreventes femininos de escolas públicas e depois de escreventes masculinos de escolas públicas, o mesmo foi realizado para escreventes femininos e masculinos de escolas privadas. E, em seguida, para as análises, foi utilizado o teste paramétrico *T-test* para amostras dependentes, na comparação entre as ocorrências de segmentações não-convencionais de palavras encontradas em textos escritos por escreventes do gênero femininos e textos escritos por escreventes do gênero masculinos. A escolha do teste paramétrico se baseou na verificação da não-violação da curva do teste de normalidade, adotando-se o valor de $\alpha \leq 0,05$. Comparamos os dados de números de palavras e percentual de ocorrências de segmentações não-convencionais de palavras dos textos de escreventes masculinos com os dados dos textos de escreventes femininos.

3 APRESENTAÇÃO DOS RESULTADOS

Os resultados obtidos e apresentados neste capítulo são organizados em função de responder ao objetivo desta pesquisa, que consiste em analisar como as ocorrências de segmentações não-convencionais de palavras se distribuem em função da variação: do tipo de escolas (públicas e privadas), dos anos letivos (1º ano, 2º ano, 3º ano, 4º ano e 5º ano) e do gênero/sexo dos escreventes (masculino e feminino).

Conforme descrito na seção “Procedimentos de análise dos dados” (2.3), os dados que formam o *corpus* deste trabalho são referentes a segmentações não-convencionais de palavras encontradas em uma amostra de 427 textos. Para a realização da análise estatística descritiva, foram contabilizados o número de palavras de cada texto e o número de segmentações não-convencionais. A partir desses índices, foi realizado o cálculo da razão entre os dois, chegando-se, assim, à média percentual do número de ocorrências de segmentações não-convencionais de palavras de cada texto em relação ao número de palavras de cada texto. Conseqüentemente, são comparados, em cada texto que compôs a amostra, dois tipos de resultados: (1) o número de palavras de cada texto (Nº Pal); e (2) o percentual de segmentações não-convencionais de palavras resultante da razão entre o número de palavras e o número de ocorrências de segmentações não-convencionais (% SNC).

Posteriormente, cada variável dos tipos de escolas (públicas e privadas), dos anos letivos (1º ano, 2º ano, 3º ano, 4º ano e 5º ano) e de gênero/sexo dos escreventes (masculino e feminino) foi cruzada entre si e verificado o valor da média de Nº de Pal e do % de SNC.

Num primeiro momento, havíamos calculado as médias de cada valor; observamos, porém¹⁹, em virtude da grande variação entre os dados de cada texto, que a média nem sempre condizia com os dados do grupo, o que nos fez aplicar aos dados o cálculo de mediana e de desvio padrão²⁰ para que pudéssemos mostrar resultados mais fiéis e também a heterogeneidade do grupo.

3.1 Tipos de escolas

Quanto à variável tipos de escolas, investigamos a média, a mediana e o desvio padrão do número de palavras e do percentual das ocorrências, não separando, neste momento, os dados dos anos letivos e de escreventes masculinos e femininos.

Primeiramente, apresentamos, na Tabela 07, que segue, a estatística descritiva dos dados gerais – referentes à sua distribuição em escolas públicas e privadas:

Tabela 07 - Distribuição das médias, das medianas e dos desvios padrão de ocorrências de segmentações não-convencionais em escolas públicas e privadas

	Escolas Públicas		Escolas Privadas	
	Nº Pal	% SNC	Nº Pal	% SNC
Média	157,04	7,13	179,02	3,01
Mediana	157	1,78	180,05	0,49
D. Padrão	93,82	14,86	98	5,93

Leia-se: Nº Pal: número de palavras; Nº SNC: número de segmentações não-convencionais; % SNC: percentual de segmentações não-convencionais.

Fonte: elaboração própria

A proporção da diferença²¹ entre as escolas públicas e as privadas quanto ao número de palavras é de 1,15 vezes. E a proporção da diferença entre as escolas públicas e as escolas

¹⁹ Essa observação foi fruto da necessidade de resposta aos questionamentos da professora Dr^a Luciani Tenani em debate realizado no “Ciclo de palestras e debates da linha de pesquisa ‘Oralidade e Letramento’”, em 2013.

²⁰ Média é uma medida de posição, de tendência central, calculada a partir da soma dos dados dividida pelo número de dados. Mediana é uma medida de posição, de tendência central, calculada a partir da variável que ocupa a posição central de um conjunto de N dados ordenados. Desvio padrão é uma medida de dispersão, ele mostra o quanto de variância (ou dispersão) existe em relação à média, calculado a partir da raiz quadrada da variância.

²¹ Esse cálculo é realizado pela divisão da primeira variável pela segunda.

privadas quanto ao percentual de ocorrências de segmentações não-convencionais de palavras é de 3,63, que é ainda maior que a do número de ocorrências.

O resultado geral observado mostra que os escreventes em escolas públicas em relação às privadas escrevem menos; mesmo assim, sua escrita apresenta um maior percentual de segmentações não-convencionais de palavras.

Aplicamos aos dados a estatística inferencial, usando o T-test para a comparação dos dados de amostras independentes. A Tabela 08 apresenta a comparação do número de palavras de escreventes em escolas públicas com o de escreventes de escolas privadas. Como se pode observar, a diferença mostrou-se estatisticamente significativa.

Tabela 08 - Distribuição do número de palavras nos textos por tipos de escolas: públicas x privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	157,04	93,82	t-value = - 2,10 df= 409
Privadas	179,02	98	p= 0,04

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Por fim, a Tabela 09 apresenta a comparação do percentual de segmentações não-convencionais de palavras de escreventes das escolas públicas com o de escreventes das escolas privadas; mais uma vez, observamos que a comparação entre o percentual de segmentações não-convencionais de palavras de escolas públicas com escolas privadas é altamente significativa.

Tabela 09 - Distribuição do percentual das segmentações não-convencionais por tipos de escolas: públicas x privadas

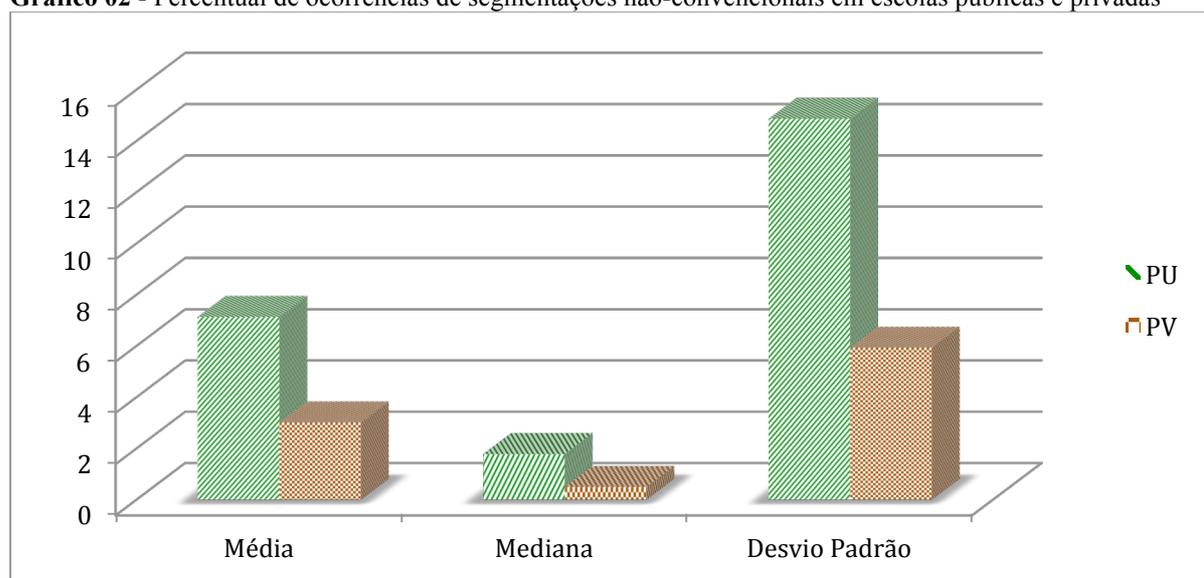
Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	7,13	14,86	t-value = 2,87 df= 409
Privadas	3,01	5,93	p= 0,00

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Em síntese, a comparação entre escolas públicas e privadas na estatística inferencial por meio de T-test se mostrou altamente significativa tanto em relação ao número de palavras, quanto ao percentual de ocorrências de segmentações não-convencionais de palavras – o que confirma os resultados observados na estatística descritiva.

E finalizamos nossa exposição de resultados sobre a distribuição geral das segmentações não-convencionais em função dos tipos de escolas com a apresentação do Gráfico 02, que sintetiza as informações relacionadas a essa distribuição:

Gráfico 02 - Percentual de ocorrências de segmentações não-convencionais em escolas públicas e privadas



3.2 Anos letivos

Quanto à variável anos letivos, comparamos média, mediana e desvio padrão do número de palavras e do percentual das ocorrências, não separando, neste momento, os dados de escreventes femininos e masculinos.

A Tabela 10 apresenta o cálculo da estatística descritiva dos dados relativos a essa comparação em escolas públicas e privadas:

Tabela 10 - Distribuição das médias, das medianas e dos desvios padrão de ocorrências de segmentações não-convencionais em escolas (públicas x privadas) e anos letivos (1º x 2º x 3º x 4º x 5º)

	Escolas Públicas						Escolas Privadas					
	Nº Pal ²²			% SNC			Nº Pal			% SNC		
	M	MD	DP	M	MD	DP	M	MD	DP	M	MD	DP
1º	48	28	54	20,4	12,8	24,7	90	41	78	8,4	7,3	7,3
2º	132	125	61	7,6	5,2	12,4	140	123	63	4,6	1,5	8,1
3º	182	169	74	3,3	1,7	5,4	205	205	64	0,9	0	1,9
4º	186	184	66	1,8	0,7	5,2	201	201	79	0,5	0	1,2
5º	233	226	78	1,3	0,5	2,3	259	251	96	0,3	0	0,4

Leia-se: Nº Pal: número de palavras; Nº SNC: número de segmentações não-convencionais; % SNC: percentual de segmentações não-convencionais; M: média; MD: mediana; DP: desvio padrão.

Fonte: elaboração própria.

Os dados (média, mediana e desvio padrão) dos anos letivos mostram que, em ambos os tipos de escolas, há um gradativo aumento no número de palavras ao longo dos cinco anos. Em ambos os tipos de escolas também ocorre uma gradativa diminuição do número de segmentações não-convencionais de palavras, bem como do percentual referente a elas, ao longo dos cinco anos. No entanto, os resultados reforçam, em todos os anos, o fato de que, nas escolas públicas, os escreventes escrevem menos e com mais ocorrências de segmentações não-convencionais de palavras. Porém, enquanto nas escolas públicas há uma diminuição contínua²³ – de 2,46 vezes do 1º ano para o 2º ano; de 3,06 vezes do 2º ano para o 3º ano; de 2,42 vezes do 3º ano para o 4º ano; e de 1,4 vezes do 4º ano para o 5º ano –, nas escolas privadas, além de uma queda acentuada do 1º ano para o 2º ano – de 4,86 vezes –, a diminuição do percentual de ocorrências de segmentações não-convencionais de palavras cai para zero a partir do 3º ano.

Aplicamos aos dados da variável de anos letivos a estatística inferencial, usando o T-test para a comparação dos dados de amostras independentes. Os resultados dessa análise são expostos, a seguir, para cada ano letivo – distribuídos em conjuntos de duas tabelas com

²² Os números de palavras foram aproximados para números inteiros em virtude de não haver espaço para mais de 3 casas numéricas na tabela.

²³ O cálculo foi realizado a partir dos dados de mediana do percentual de segmentações não-convencionais de palavras.

informações sobre o número de palavras e do percentual dessas segmentações. Iniciamos com as informações sobre o primeiro ano, comparando as escolas públicas e privadas:

Tabela 11 - Distribuição do número de palavras dos textos no 1º ano por tipos de escolas: públicas x privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	44,76	50,24	t-value = -2,95 df= 85
Privadas	89,63	78,16	p= 0,00

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Tabela 12 - Distribuição do percentual das segmentações não-convencionais no 1º ano por tipos de escolas: públicas x privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	20,38	24,66	t-value = 2,33 df= 85
Privadas	8,42	7,31	p= 0,02

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Como se pode depreender das informações mostradas na tabela 11, a diferença do número de palavras é acentuada, o que se confirma no valor de p como estatisticamente significativo. Na tabela 12, a distribuição do percentual das segmentações não-convencionais no primeiro ano das escolas públicas e privadas se mostrou estatisticamente significativa.

Passemos às informações sobre o segundo ano, comparando-se escolas públicas e privadas:

Tabela 13 - Distribuição do número de palavras no 2º ano por tipos de escolas: públicas x privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	132,07	60,91	t-value = -0,52 df= 81
Privadas	139,05	62,69	p= 0,61

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Tabela 14 - Distribuição do percentual das segmentações não-convencionais no 2º ano por tipos de escolas: públicas x privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	7,59	12,45	t-value = 1,09 df= 81
Privadas	4,59	8,03	p= 0,28

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Como se pode verificar pelos dados expostos nas tabelas 13 e 14, a diferença de número de palavras e de percentual desse tipo de segmentação não se mostrou significativa, quando se comparam os grupos do segundo ano das escolas públicas e privadas, o que mostra que eles escrevem de forma similar nesse ano letivo.

Passemos, a seguir, às informações sobre o terceiro ano, comparando-se escolas públicas e privadas:

Tabela 15 - Distribuição do número de palavras no 3º ano por tipos de escolas: públicas x privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	184,36	73,01	t-value = -1,16 df= 85
Privadas	205,30	64,14	p= 0,25

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Tabela 16 - Distribuição do percentual das segmentações não-convencionais no 3º ano por tipos de escolas: públicas x privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	3,36	5,41	t-value = 2,17 df= 85
Privadas	0,70	1,38	p= 0,03

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Os dados expressos na tabela 15 mostram, primeiramente, não ser significativa a diferença do número de palavras escritas entre os dois grupos do terceiro ano. No entanto, essa diferença se torna significativa quanto se trata do percentual de segmentações não-

convencionais de palavras no 3º quando comparados os grupos, nesse ano letivo, escolas públicas e privadas.

Passemos, agora, às informações sobre o quarto ano, comparando-se escolas públicas e privadas:

Tabela 17 - Distribuição do número de palavras no 4º ano por tipos de escolas: públicas x privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	185,81	65,79	t-value = -0,74 df= 53
Privadas	200,63	78,84	p= 0,46

F Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Tabela 18 - Distribuição do percentual das segmentações não-convencionais no 4º ano por tipos de escolas: públicas x privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	1,77	5,18	t-value = 1,02 df= 53
Privadas	0,54	1,18	p= 0,31

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Os dados mostrados nas tabelas 17 e 18 mostraram não ser estatisticamente significativa a comparação dos dois aspectos analisados nos dois grupos de quarto ano.

Por fim, passemos às informações sobre o quinto ano, comparando-se escolas públicas e privadas:

Tabela 19 - Distribuição do número de palavras das segmentações não-convencionais no 5º ano por tipos de escolas: públicas x privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	233,31	78,95	t-value = -1,34 df= 97
Privadas	258,70	96,33	p= 0,18

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Tabela 20 - Distribuição do percentual das segmentações não-convencionais no 5º ano por tipos de escolas: públicas x privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	1,36	2,34	t-value = 2,42 df= 97
Privadas	0,26	0,42	p= 0,02

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Como se pode depreender das informações dispostas nas tabelas 19 e 20, não é estatisticamente significativa a diferença do número de palavras entre os dois grupos do quinto ano, o que significa que ambos os tipos de escolares escrevem em quantidade similar. No entanto, essa diferença se torna significativa quando se trata do percentual de segmentações não-convencionais entre escolas públicas e privadas.

Aplicamos aos dados da variável de anos letivos a estatística inferencial, usando o teste *Friedman ANOVA and Kendall Coeff. Of Concordance* para variáveis dependentes, pois comparamos – dessa vez – os cinco anos entre si, mantendo o valor de $p \leq 0,05$. O teste foi aplicado para o número de palavras e para o percentual de segmentações não-convencionais. Exporemos, primeiramente, os dados de escolas públicas (nas tabelas 21 e 22) e, em seguida, para escolas privadas (nas tabelas 23 e 24):

Tabela 21 - Distribuição do número de palavras nos textos por anos letivos nas escolas públicas (1º x 2º x 3º x 4º x 5º)

	Média	Desvio Padrão	Friedman ANOVA and Kendall Coeff. Of Concordance	Sum of ranks
1º	43,28	51,25		43
2º	134,75	66,53	ANOVA Chi = 76,18	97
3º	193,36	76,57	df = 4	125
4º	185,80	65,79	p = 0,00	122
5º	245,97	82,76		153

Fonte: elaboração própria. Teste *Friedman ANOVA and Kendall Coeff. Of Concordance* para variáveis dependentes. ($p \leq 0,05$).

Como se pode observar, a diferença do número de palavras é acentuada ao longo dos cinco anos do Ensino Fundamental I, o que se confirma no valor de p como estatisticamente significativo. O aumento gradativo do número de palavras de ano para ano só é quebrado do 3º ano para o 4º ano – já que os textos de 3º apresentam mais palavras que os textos do 4º ano – mas sem que se comprometa a tendência crescente confirmada pela análise estatística.

Tabela 22 - Distribuição do percentual das segmentações não-convencionais por anos letivos nas escolas públicas (1º x 2º x 3º x 4º x 5º)

	Média	Desvio Padrão	Friedman ANOVA and Kendall Coeff. Of Concordance	Sum of ranks
1º	20,54	27,22		147
2º	7,01	15,17	ANOVA Chi = 37,82	122
3º	1,89	2,06	df = 4	108,5
4º	1,77	5,18	p = 0,00	87
5º	0,78	1,96		75,5

Fonte: elaboração própria. Teste *Friedman ANOVA and Kendall Coeff. Of Concordance* para variáveis dependentes. ($p \leq 0,05$).

Como se pode observar pelos dados da tabela 22, a acentuação de diferença do percentual de segmentações não-convencionais de palavras ao longo dos cinco anos é significativa, o que se confirma no valor de p como estatisticamente significante. Assim, em termos percentuais, há uma diminuição das segmentações não-convencionais de palavras, fato que reforça a tendência de redução desse tipo de segmentação quando se comparam os cinco anos do Ensino Fundamental I.

Expostos os dados de anos letivos relativos a escolas públicas, passemos à observação desse mesmo tipo de informação nas escolas privadas – conforme antecipamos, expostos nas tabelas 23 e 24:

Tabela 23 - Distribuição do número de palavras nos textos por anos letivos nas escolas privadas (1° x 2° x 3° x 4° x 5°)

	Média	Desvio Padrão	Friedman ANOVA and Kendall Coeff. Of Concordance	Sum of ranks
1°	75,05	60,79		26
2°	143,42	54,76	ANOVA Chi = 41,32	43,5
3°	213,05	55,44	df = 4	71
4°	200,63	78,84	p = 0,00	63,5
5°	270,58	90,63		81

Fonte: elaboração própria. Teste *Friedman ANOVA and Kendall Coeff. Of Concordance* para variáveis dependentes. ($p \leq 0,05$).

Como se pode observar pelos dados da tabela 23, a diferença do número de palavras vai se acentuando conforme se analisa esse número em função dos cinco anos do Ensino Fundamental I. Essa diferença se mostrou como estatisticamente significativa, o que se confirma pelo valor de p . O aumento gradativo do número de palavras de ano para ano, tal qual nas escolas públicas, só é quebrado do 3° ano para o 4° ano – já que os textos de 3° têm mais palavras do que os textos do 4° ano, mas sem que se comprometa a tendência crescente apontada pela análise estatística.

Tabela 24 - Distribuição do percentual das segmentações não-convencionais por anos letivos nas escolas privadas (1° x 2° x 3° x 4° x 5°)

	Média	Desvio Padrão	Friedman ANOVA and Kendall Coeff. Of Concordance	Sum of ranks
1°	8,84	7,46		82
2°	2,83	3,88	ANOVA Chi = 30,62	69
3°	0,73	1,40	df = 4	51
4°	0,54	1,18	p = 0,00	42
5°	0,25	0,47		41

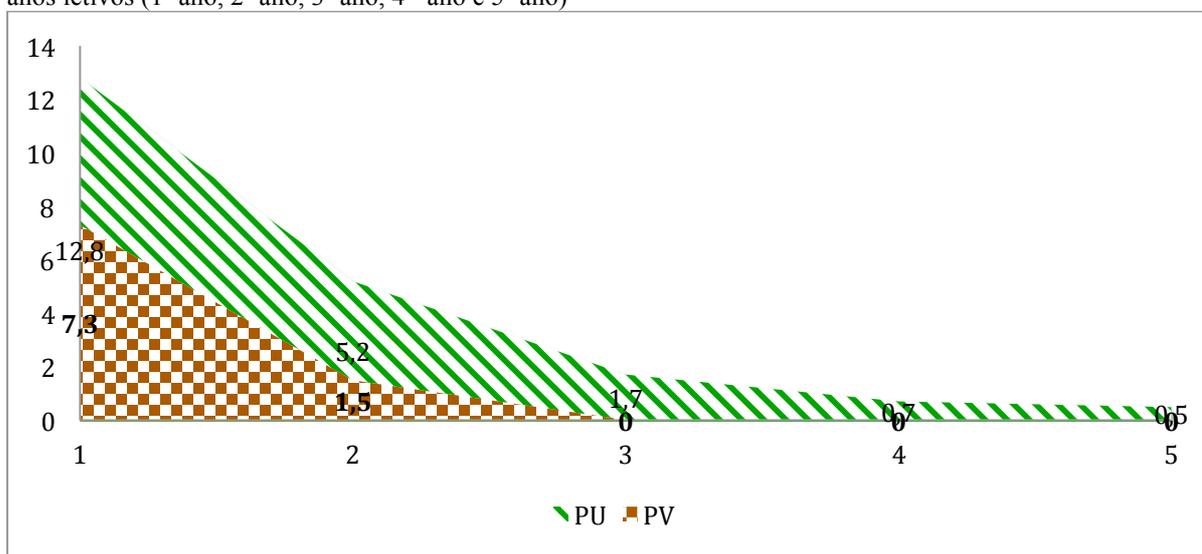
Fonte: elaboração própria. Teste *Friedman ANOVA and Kendall Coeff. Of Concordance* para variáveis dependentes. ($p \leq 0,05$).

Por fim, como se pode observar nos dados expostos na tabela 24, em termos percentuais ocorre diminuição de segmentações não-convencionais conforme se avançam os

cinco anos do Ensino Fundamental I nas escolas privadas. Esta diminuição se mostrou significativa, em termos estatísticos, como o atesta o valor de p .

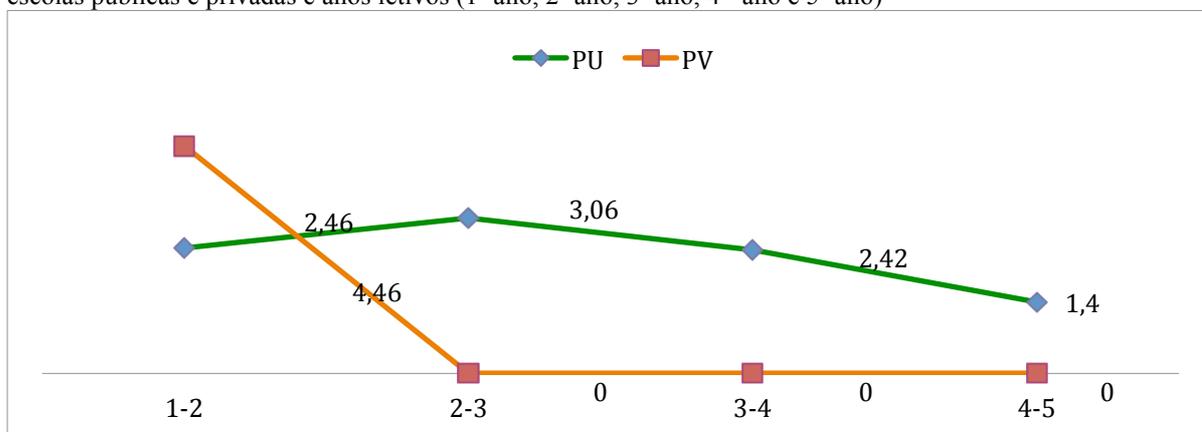
Assim, para sintetizar estes últimos passos de nossa descrição, apresentaremos o Gráfico 03 e o Gráfico 04, nos quais as semelhanças e as diferenças de tendências nos dados de anos letivos se mostram na comparação entre escolas públicas e privadas:

Gráfico 03 - Percentual de ocorrências de segmentações não-convencionais nas escolas públicas e privadas e anos letivos (1º ano, 2º ano, 3º ano, 4º ano e 5º ano)



Fato interessante a ser destacado no Gráfico 03 é o da diminuição mais acentuada entre o primeiro e o segundo ano nas escolas privadas, seguida de uma diminuição mais amena entre este e o terceiro ano, seguida, por fim de uma estabilização (tendendo a zero) nesse número deste ao quinto ano. Situação diferente vemos na diminuição nas escolas públicas: ela é mais contínua (embora se detecte uma queda mais acentuada do primeiro para o segundo ano) e chega ao quinto ano ainda com certo percentual de ocorrências não-convencionais. Essas informações se mostram realçadas quando se observa a razão da diminuição do percentual dessas ocorrências em ambos os tipos escolares analisados:

Gráfico 04 – A razão da diminuição do percentual de ocorrências de segmentações não-convencionais em escolas públicas e privadas e anos letivos (1º ano, 2º ano, 3º ano, 4º ano e 5º ano)



Como mostrado nos resultados, ao final do Ensino Fundamental I, nas escolas privadas, o percentual de segmentações não-convencionais de palavras chega a zero na mediana e 0,2 na média, pois ainda foram encontradas 13 ocorrências – cujo percentual não se mostra como significativo. Já nas escolas públicas, esse índice é de 0,5 na mediana e 1,3 na média. No Quadro 09, descreveremos esses dados:

Quadro 09 – Dados resultantes do 5º ano do Ensino Fundamental I

Dados do 5º ano do Ensino Fundamental I	Escolas públicas	Escolas privadas
Número de escreventes da amostra	72	27
Número de escreventes com ocorrências	34	8
Percentual de escreventes com ocorrências	47,22%	29,63%
Número de alunos com uma única ocorrência	12	5
Percentual de alunos com uma única ocorrência	16,67%	18,52%
Número total de segmentações não-convencionais de palavras	154	13

Fonte: elaboração própria

No Quadro 10, apresentamos o número de segmentações não-convencionais de palavras nas variáveis escolas públicas e escolas privadas classificadas em hipossegmentações, hipersegmentações e mesclas:

Quadro 10 – Dados qualitativos resultantes do 5º ano do Ensino Fundamental I

	ESCOLAS PÚBLICAS			ESCOLAS PRIVADAS		
	hipo	hiper	mescla	hipo	hiper	mescla
Número de segmentações não-convencionais	111	41	02	07	06	0

Fonte: elaboração própria

O número total de hipossegmentações foi de 118 ocorrências, das quais 91 corresponderam à estrutura de um grupo clítico – 77,12%. Analisando separadamente, o número de hipossegmentações em escolas públicas foi de 111 ocorrências, das quais 85 corresponderam à estrutura de um grupo clítico – 76,58%. Já o número de hipossegmentações em escolas privadas foi de 7 ocorrências, das quais 6 corresponderam à estrutura de um grupo clítico – 85,71%.

Já em relação ao número total de hipersegmentações, de 47 ocorrências, 40 apresentaram partes que poderiam corresponder a um clítico – 85,11%. Analisado separadamente por tipos de escolas, das 41 ocorrências de hipersegmentações nas escolas públicas 34 apresentaram partes que corresponderiam a um clítico – 82,93%. Já nas escolas privadas, nas seis ocorrências que permaneceram, todas apresentaram partes que poderiam corresponder a um clítico – 100%.

Como vimos no Quadro 10, as duas mesclas ocorreram em textos de escreventes de escolas públicas: “nun camais” (nunca mais) AR_12_5C_23M; “ea xou” (e achou) PF_12_5A_04M.

Passemos, finalmente, aos dados relativos ao último aspecto que levaremos em conta na distribuição das segmentações não-convencionais ao longo do Ensino Fundamental I: a variável gênero/sexo.

3.3 Sexo/gênero

Ao investigarmos os efeitos da variável sexo/gênero, observamos os valores de média, mediana e desvio padrão do número de palavras e do percentual de ocorrências de segmentações não-convencionais. Esses valores são expostos na tabela 25:

Tabela 25 - Distribuição das segmentações não-convencionais por escolas públicas e privadas e masculino x feminino

	Contexto Público				Contexto Privado			
	Masculino		Feminino		Masculino		Feminino	
	Nº Pal	% SNC	Nº Pal	% SNC	Nº Pal	% SNC	Nº Pal	% SNC
Média	221,21	2,03	244,13	0,75	230,47	0,3	294	0,21
Mediana	222,5	1,04	229,5	0,26	235	0	289,5	0
D. Padrão	75,61	2,61	81,29	1,91	89,94	0,38	95,87	0,49

Leia-se: Nº Pal: número de palavras; % SNC: percentual de segmentações não-convencionais; D. Padrão: desvio padrão.

Fonte: elaboração própria.

É possível verificar, de acordo com os números expostos nessa tabela, que os escreventes masculinos escrevem menos palavras do que os escreventes femininos, em ambos os tipos de escolas. No entanto, essa diferença varia em função dos tipos escolares em comparação: 1,03 vezes nas escolas públicas, menor do que a diferença de 1,23 verificada nas escolas privadas. Também é possível observar que a diferença se mostra, ainda, no que se refere ao número e ao percentual de segmentações não-convencionais de palavras.

Feita essa caracterização geral, vejamos uma síntese dos dados com as médias, medianas e desvios padrão de segmentações não-convencionais de palavras cruzando informações entre as variáveis sexo/gênero e anos letivos nas escolas públicas – na tabela 26 – e entre essas mesmas variáveis nas escolas privadas – na tabela 27:

Tabela 26 - Distribuição das médias, das medianas e dos desvios padrão de ocorrências de segmentações não-convencionais em escolas públicas por anos letivos (1º x 2º x 3º x 4º x 5º) e sexo/gênero (masculino x feminino)

Contexto Público

Ano	Masculinos						Femininos					
	Nº de Pal			% SNC			Nº de Pal			% SNC		
	M	MD	DP	M	MD	DP	M	MD	DP	M	MD	DP
1º	46,4	28	47,2	22,6	16,3	24,2	49,2	26,5	54,1	18	11,6	25,4
2º	125,2	118	53,5	7,5	5,9	6,1	139,2	128	67,9	7,7	3,6	16,8
3º	173,6	164	67,1	4,8	2,2	7,2	194,8	185	77,9	1,9	1,4	2,1
4º	182,7	190	67,7	1,2	0,9	1,3	188,9	178	65,6	2,4	0,5	7,3
5º	221,2	222,5	75,6	2	1	2,6	244,1	229,5	81,3	1,9	1,4	2,1

Leia-se: Nº Pal: número de palavras; % SNC: percentual de segmentações não-convencionais; M: média; MD: mediana; DP: desvio padrão.

Fonte: elaboração própria.

Tabela 27 - Distribuição das médias, das medianas e dos desvios padrão de ocorrências de segmentações não-convencionais em escolas privadas por anos letivos (1º x 2º x 3º x 4º x 5º) e sexo/gênero (masculino x feminino)

Contexto Privado

Ano	Masculinos						Femininos					
	Nº de Pal			% SNC			Nº de Pal			% SNC		
	M	MD	DP	M	MD	DP	M	MD	DP	M	MD	DP
1º	74,1	35	90	9,6	9,8	8,2	107,9	116	60,5	7	6,4	6,2
2º	123,8	120	68,4	8,8	1,3	11,9	149,3	155	59,3	2,1	1,8	2,7
3º	169	192	57,6	1,1	0,6	2	235	222	54,8	0,3	0,3	0,4
4º	186,3	178,5	98,7	0,8	0,2	1,6	216,6	208	49,8	0,2	0	0,4
5º	230,5	235	90	0,3	0	0,4	294	289,5	95,9	0,2	0	0,5

Leia-se: Nº Pal: número de palavras; % SNC: percentual de segmentações não-convencionais; M: média; MD: mediana; DP: desvio padrão.

Fonte: elaboração própria.

Observando-se os dados expostos nas tabelas 26 e 27, por meio da estatística inferencial, testamos a distribuição do percentual de ocorrências de segmentações não-convencionais de palavras cruzando a variável anos letivos com a variável gênero/sexo. Primeiramente, nas tabelas 28 e 29, expõem-se os dados das escolas públicas,

respectivamente, dos escreventes femininos e masculinos. Em seguida, nas tabelas 30 e 31, também respectivamente, são expostos aqueles de escreventes das escolas privadas:

Tabela 28 - Distribuição do percentual das segmentações não-convencionais por anos letivos (1° x 2° x 3° x 4° x 5°) de escreventes femininos nas escolas públicas

	Média	Desvio Padrão	Friedman ANOVA and Kendall Coeff. Of Concordance	Sum of ranks
1°	15,99	23,71		70
2°	3,92	3,85	ANOVA Chi = 10,94	58
3°	1,81	3,44	df = 4	53
4°	2,39	7,27	p = 0,03	46
5°	1,06	2,74		43

Fonte: elaboração própria. Teste *Friedman ANOVA and Kendall Coeff. Of Concordance* para variáveis dependentes. ($p \leq 0,05$).

A diminuição no percentual de ocorrências de segmentações não-convencionais de palavras se mostrou altamente significativa como mostra o valor de p . Há uma queda acentuada do primeiro para o segundo ano e uma queda para o terceiro e para o quinto. No quarto ano, há um aumento desse percentual, que pode ser observado também na estatística descritiva exposta na tabela 26.

Tabela 29 - Distribuição do percentual das segmentações não-convencionais por anos letivos (1° x 2° x 3° x 4° x 5°) de escreventes masculinos nas escolas públicas

	Média	Desvio Padrão	Friedman ANOVA and Kendall Coeff. Of Concordance	Sum of ranks
1°	23	25,03		79,5
2°	6,15	5,84	ANOVA Chi = 30,96	59,5
3°	6,45	9	df = 4	59
4°	1,16	1,29	p = 0,00	32
5°	2,45	2,88		40

Fonte: elaboração própria. Teste *Friedman ANOVA and Kendall Coeff. Of Concordance* para variáveis dependentes. ($p \leq 0,05$).

A diminuição no percentual de ocorrências de segmentações não-convencionais de palavras se mostrou altamente significativa, como o mostra o valor de p , assim como ocorreu

nos dados de escreventes femininos. Também houve, de forma semelhante aos dados de escreventes femininos, uma queda acentuada no percentual de segmentações não-convencionais de palavras do primeiro para o segundo ano. Porém, nos dados de escreventes masculinos há um leve aumento no terceiro ano e um aumento acentuado no quinto, diferentemente dos dados de escreventes femininos, grupo em que houve aumento apenas no quarto ano.

Passemos à descrição dos dados das escolas privadas.

Tabela 30 - Distribuição do percentual das segmentações não-convencionais por anos letivos (1° x 2° x 3° x 4° x 5°) de escreventes femininos nas escolas privadas

	Média	Desvio Padrão	Friedman ANOVA and Kendall Coeff. Of Concordance	Sum of ranks
1°	7,78	6,47		41,5
2°	1,30	3,91	ANOVA Chi = 24,46	35,5
3°	0,34	0,49	df = 4	22
4°	0,22	0,36	p = 0,00	19,5
5°	0,05	0,14		16,5

Fonte: elaboração própria. Teste *Friedman ANOVA and Kendall Coeff. Of Concordance* para variáveis dependentes. ($p \leq 0,05$).

A diminuição no percentual de ocorrências de segmentações não-convencionais de palavras se mostrou altamente significativa como mostra o valor de p . Há uma queda acentuada do primeiro para o segundo ano, seguida de uma queda contínua.

Tabela 31 - Distribuição do percentual das segmentações não-convencionais por anos letivos (1° x 2° x 3° x 4° x 5°) de escreventes masculinos nas escolas privadas

	Média	Desvio Padrão	Friedman ANOVA and Kendall Coeff. Of Concordance	Sum of ranks
1°	9,37	7,17		37
2°	8,76	11,89	ANOVA Chi = 9,23	30
3°	1,12	1,98	df = 4	23,5
4°	0,91	1,65	p = 0,05	24,5
5°	0,33	0,44		20

Fonte: elaboração própria. Teste *Friedman ANOVA and Kendall Coeff. Of Concordance* para variáveis dependentes. ($p \leq 0,05$).

A diminuição no percentual de ocorrências de segmentações não-convencionais de palavras se mostrou altamente significativa como mostra o valor de p . Há uma queda acentuada do segundo ano para o terceiro, diferentemente dos dados de escreventes femininos em que essa queda ocorre antes (do primeiro para o segundo ano). A diminuição a partir do terceiro ano é contínua.

Encerrada a comparação por anos letivos entre os escreventes de ambos os gêneros/sexo e em ambos os tipos de escolares, passemos à comparação global (que inclui todos os anos) entre eles. Os dados para essa comparação apresentam-se dispostos nas tabelas 32 e 33:

Tabela 32 - Distribuição do número de palavras nos textos de escreventes femininos em escolas públicas e privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	244,13	81,29	t-value = -1,77 df= 48
Privadas	294	95,87	p= 0,08

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Tabela 33 - Distribuição do percentual de segmentações não-convencionais de escreventes femininos em escolas públicas e privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	0,75	1,91	t-value = 0,97 df= 48
Privadas	0,21	0,49	p= 0,34

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Como se pode verificar pelos valores expostos nestas tabelas, nenhum dos aspectos comparados, entre os escreventes do gênero feminino, se mostrou como estatisticamente significativo.

Passemos, a seguir, ao mesmo tipo de comparação, desta vez feita entre escreventes do sexo/gênero masculino, cujos dados são mostrados nas tabelas 34 e 35:

Tabela 34 - Distribuição do número de palavras nos textos de escreventes masculinos em escolas públicas e privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	221,21	75,61	t-value = -0,37 df= 47
Privadas	230,47	89,94	p= 0,71

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Tabela 35 - Distribuição do percentual das segmentações não-convencionais de escreventes masculinos em escolas públicas e privadas

Tipos de escolas	Média	Desvio Padrão	T-test
Públicas	2,03	2,61	t-value = 2,55 df= 47
Privadas	0,3	0,38	p= 0,01

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Como se pode observar, no conjunto de todos os anos, não houve, para os escreventes do gênero masculino, diferença estatisticamente significativa no que se refere ao número de palavras de seus textos. No entanto, quando se trata do percentual de segmentações não-convencionais, a diferença entre os escreventes do gênero/sexo masculinos em escolas públicas e privadas se torna significativa em termos estatísticos.

Outra análise que realizamos – buscando nosso objetivo de traçar um perfil das ocorrências de segmentações não-convencionais de palavras – foi comparar exclusivamente os dados de escreventes masculinos com os de escreventes femininos, desconsiderando, neste momento, as variáveis de tipos de escolas e de anos letivos. Os resultados dessa comparação encontram-se nas 36 e 37, que se seguem:

Tabela 36 - Distribuição do número de palavras nos textos por sexo/gênero (masculino x feminino)

	Média	Desvio Padrão	T-test
Masculino	221,21	81,29	t-value = 1,23 df= 70
Feminino	244,13	75,61	p= 0,22

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Tabela 37 - Distribuição do percentual das segmentações não-convencionais por sexo/gênero (masculino x feminino)

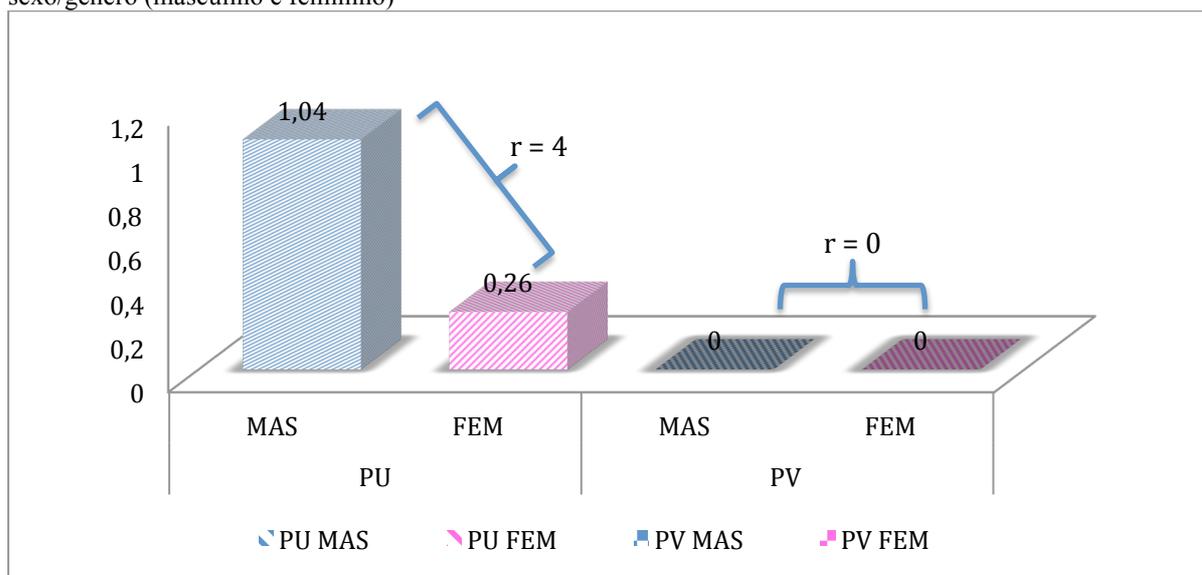
	Média	Desvio Padrão	T-test
Masculino	2,03	2,61	t-value = -2,39 df= 70
Feminino	0,75	1,91	p= 0,02

Fonte: elaboração própria. Teste T para amostras independentes ($p \leq 0,05$).

Nessa comparação, como se pode observar, embora a diferença do número de palavras de escreventes dos gêneros masculino e feminino não seja significativa, a diferença entre seu percentual de segmentações não-convencionais o é.

No entanto, quando acrescentamos a variável gênero/sexo à variável tipos de escolas, esses resultados apresentam grande modificação: enquanto que, nas escolas privadas, a diferença de gênero ($r = 0$)²⁴ não determina o percentual de ocorrências de segmentação não-convencional, nas escolas públicas essa diferença se mostra bastante acentuada ($r = 4$), como se pode acompanhar no Gráfico 05, a seguir:

Gráfico 05 - Percentual de ocorrências de segmentações não-convencionais em escolas públicas e privadas e sexo/gênero (masculino e feminino)



²⁴ r é a razão, resultado da divisão da primeira variável pela segunda.

Vimos, nos dados apresentados, indícios de outras informações relevantes e pertinentes aos estudos sobre ortografia, segmentações não-convencionais de palavras e prosódia. Porém, em decorrência do nosso recorte e do cumprimento de nossos prazos para o desenvolvimento desta pesquisa, voltamos nossa atenção apenas para as variáveis que propusemos analisar. Gostaríamos, entretanto, de destacar nosso interesse e incentivo de pesquisas futuras que investiguem outras variáveis com base em dados mais detalhados do que aqueles que sustentaram nossos resultados.

4 DISCUSSÃO DOS RESULTADOS

Descritos os resultados, passaremos, então, neste capítulo, à discussão das tendências para os quais eles apontaram. Essa discussão se dará em função das variáveis que foram comparadas no desenvolvimento da pesquisa.

4.1 Tipos de escolas

Os resultados referentes a escolas públicas e privadas confirmam nossa hipótese de que haveria diferença entre esses dois tipos de escolas – resultados que ratificam tanto a percepção de Abaurre (1991) quanto aqueles a que chegou Cunha (2004) para dados do interior gaúcho. Acreditamos, no entanto, que não apenas confirmamos o que já se supunha ou se tinha observado nesses dois trabalhos. Com efeito, ao trabalharmos com um grande conjunto de dados, coletados sob condições de certo modo controladas e submetê-los a crivo estatístico, mostramos que se trata de diferença bastante relevantes, já que o número de ocorrências de segmentações não-convencionais é 3,63 vezes maior nas escolas públicas do que nas escolas privadas – diferença que se mostrou altamente significativa em termos estatísticos. Vejamos o que dizem essas duas autoras sobre essa diferença de contexto.

Abaurre (1991, s/p) analisa um *corpus* constituído de centenas de textos oriundos de escolas públicas e particulares, de pré-escola e de primeira série, de regiões diversas do país e de diferentes classes econômicas. Essas variáveis, segundo a autora, influenciam consideravelmente o modo de segmentar dos escreventes. Quanto à variável *tipos de escolas*, é comum que se relacione a qualidade de ensino ou o bom desempenho dos alunos a questões sociais e econômicas tanto no Brasil como no resto do mundo, em Língua Portuguesa e em outras línguas. A educação é, então, fortemente relacionada com fatores sócio-econômicos, e

os resultados dessa relação levantam questões bastante problemáticas entre estratificação social e educação. Acerca dessa discussão, Abaurre (1991) observou que nas escolas privadas – em que estudam crianças de classes média e alta – os escreventes tendem a controlar melhor e mais cedo os critérios de segmentação.

A diferença entre as escolas públicas e as escolas privadas é percebida pela autora por ter um olhar sensível para as questões que envolvem a segmentação não-convencional de palavras, mas sem que tenha havido uma quantificação (numérica ou estatística) dessa diferença, o que levou a autora a reforçar a necessidade de pesquisas mais profundadas e detalhadas que investigassem essa diferença. Foi, particularmente, esse fato que nos chamou à atenção e nos fez buscar, nessa diferença, sua significância.

Como os tipos de escolas (públicas e privadas) está fortemente relacionado com questões sócio-econômicas e – conseqüentemente – com questões que dizem respeito à própria estratificação de como se distribuem (dentro e fora da escola) as práticas de letramento no país:

O fato de algumas crianças demonstrarem uma tendência para segmentar mais e mais cedo pode talvez ser explicado em termos de uma exposição quantitativa e qualitativamente diversa com relação à escrita, por parte dessas crianças. Muitas delas foram expostas mais cedo às atividades de leitura/escrita, o que fez com que provavelmente desenvolvessem um maior envolvimento com tais atividades e, conseqüentemente, uma maior curiosidade com relação ao funcionamento do próprio sistema alfabético da escrita do português (ABAURRE, 1991, s/p).

Conforme antecipamos, a diferença de desempenho na escrita relacionada aos tipos de escolas (públicas e privadas) também é investigada por Cunha (2004), em sua análise de 74 textos produzidos em duas escolas – uma pública e uma privada – de cinco escreventes de cada escola. Nos textos das escolas públicas pesquisadas pela autora, ocorreram 2067 segmentações não-convencionais; já nas escolas privadas foram 557 ocorrências. Como o objetivo da autora era analisar qualitativamente a ocorrência dessas segmentações, não há comparação dos percentuais obtidos nem tratamento estatístico dos dados, o que reforça a

importância dos dados obtidos nesta pesquisa, que mostraram essa diferença em significância estatística. A partir de seus dados, autora conclui que

as crianças da escola particular apresentam menos dúvidas do que as crianças da escola pública. No entanto, a proporcionalidade entre os tipos de erros é praticamente a mesma. Esse fato revela-nos que o processo de aquisição da escrita é semelhante, independentemente do tipo de escolar (p. 352).

Como ocorreu com nossos resultados, os de Cunha (2010) mostram haver uma quantidade de ocorrências menor nas escolas privadas. Essa diferença, porém, tanto no trabalho de Cunha quanto no nosso, não autoriza dizer diretamente que as escolas privadas são melhores e/ou que proporcionam melhor aprendizado, ou ainda que os estudantes nesse tipo escolar têm uma aquisição de escrita mais rápida. É importante lembrar que, em ambas as investigações, a análise volta-se apenas sobre um único aspecto das convenções ortográficas – a segmentação de palavras – que, por sua vez, é um dos aspectos da produção linguística escrita. Portanto, qualquer generalização sobre essa produção baseada apenas em ocorrências de segmentações não-convencionais de palavras seria precipitada e poderia, certamente, induzir a erros de avaliação da produção escrita nos dois tipos escolares analisados.

A diferença dos índices de desempenho entre escolas públicas e privadas também se mostrou significativa em um estudo de econometria aplicada. Nesse estudo, porém, essa diferença se mostra relativizada, já que é explicada em função de outros aspectos, que não apenas os de escolas públicas e privadas.

Analisando os colégios públicos federais e estaduais quanto à eficiência para diferentes níveis de entrada de aluno, mostrou-se que, para alunos com nível de entrada abaixo da média, há pouca diferença entre os sistemas de ensino. No entanto, para alunos com nível de entrada superior à média, o sistema público federal mostrou-se bastante superior ao sistema público estadual, aproximando-se bastante do sistema de ensino privado para os melhores alunos. Isto mostra que, para os melhores alunos, estudar em escolas privadas ou públicas federais não é o decisivo para seu aprendizado. Para as escolas públicas estaduais, o impacto sobre o aprendizado dos melhores alunos é inferior à média geral dos colégios públicos estaduais (GUIMARÃES; SAMPAIO, 2009, P. 65).

Como se vê, pode-se pensar numa estratificação dos tipos de escolas quando se acrescenta a esses tipos o âmbito da esfera pública envolvida na educação: federal ou

estadual. A explicação para que os colégios públicos federais estejam tendo resultados bons nos sistemas de avaliações se deve a vários fatores. Um deles diz respeito ao ingresso dos alunos, que ocorre por meio de concurso com grande concorrência, proporcionando o ingresso apenas dos que obtiveram melhores resultados no processo seletivo; outro fator é ter professores com melhor formação, pois, como há melhor remuneração no sistema federal do que no estadual, geralmente, os melhores e mais titulados professores é que ocupam os cargos do sistema federal. Esses fatores contribuem para que, muitas vezes, as escolas públicas federais obtenham desempenho igual ou superior do que muitas escolas privadas. Em outras palavras, os índices de avaliação não medem diretamente o desempenho dos alunos, mas as (diferentes) condições socioeconômicas que atravessam os tipos de escolas no qual sua formação é construída.

Assim a diferença numérica e estatística obtida como resultado nesta pesquisa não pode ser vista como parâmetro para julgar diretamente o ensino público como necessariamente pior do que o ensino privado. Basta pensar, como o faz Montt (2011), ao apresentar índices do Programa Internacional de Avaliação de Estudantes (PISA), que, se os alunos do ensino público tivessem as mesmas condições sociais que os do ensino privado, muito provavelmente apresentariam desempenho semelhante. As famílias com filhos no ensino privado tendem, talvez, a valorizar mais a educação e, por isso, tendem a incentivar os filhos a estudarem e não a trabalharem, como ocorre (muitas vezes) com famílias de baixa renda. O próprio fato de se pagar pela educação formal já significa valorizar a formação educacional – mesmo que, muitas vezes, em razão de um aspecto muito pontual: o de que seus filhos futuramente ingressem em uma bem valorizada socialmente instituição de ensino superior.

Vale ressaltar que, diferentemente do discurso de senso comum de que escolas privadas têm um melhor ensino porque têm alunos de classes sociais mais privilegiadas com

mais acessos a atividades que envolvem à escrita, acreditamos não ser apenas esse fator a provocar a diferença: um importante aspecto a ser considerado é também o da valorização das atividades de leitura e escrita, que nem sempre se mostra como um aspecto diretamente relacionado a uma ou outra classe social. A razão para acreditarmos que a valorização de atividades de escrita tenha importância a ser melhor considerada para uma avaliação mais precisa do (bom) desempenho de escrita é baseada no fato de que há grupos sociais economicamente privilegiados que tiveram ascensão social por meio de trabalhos que não requerem uso da escrita na norma-padrão, e, por isso, não valorizam a escrita culta ou a escolarização; já, por outro lado, famílias de classe baixa que acreditam que a única forma de ascensão social é a escolarização e o bom desempenho escolar, muitas vezes, valorizam a leitura e a escrita, incentivam seus filhos a estudarem e buscam bolsas de estudos em escolas privadas, acreditando ser este o caminho mais eficaz de oportunidade de vida melhor.

Outro fator que pode explicar a diferença entre as escolas públicas e as privadas “é resultado da concorrência e de níveis maiores de autonomia na definição do currículo e na alocação de recursos de que elas [escolas privadas] gozam” (MONTT, 2011, p. 11). Ao considerarmos essa autonomia do currículo e essa alocação dos recursos, não há como não relativizarmos os resultados dos desempenhos favoráveis às escolas privadas. A aplicação do currículo é diferenciada em escolas públicas e privadas. Há uma flexibilização muito maior nas escolas privadas pelo fato de as escolas não precisarem se submeter à burocracia que envolve o sistema público. Assim, nas escolas privadas, é possível haver mudanças físicas e educacionais de forma mais rápida e fácil.

Apesar dessa flexibilização maior nas escolas privadas,

É importante deixar claro (...) que o seu [da criança] comportamento diferenciado com relação à segmentação em textos espontâneos não pode ser atribuído à exposição a métodos específicos de alfabetização, já que esses métodos são similares nas escolas públicas e privadas (ABAURRE, 1991, s/p).

Os métodos de alfabetização seguem políticas públicas educacionais nacionais e igualitárias. Temos como principal orientação os PCN (1999) e em alguns estados há diretrizes curriculares que complementam as orientações, ausentes, porém, no estado de São Paulo.

Antes apenas as escolas privadas davam maior ênfase à ortografia para demonstrar resultados mais notados pelos pais – os quais, muitas vezes, confundem bom desempenho escrito com (apenas) bom desempenho ortográfico. Nos últimos anos, porém, também as escolas públicas têm adotado a mesma postura – sobretudo em decorrências de avaliações nacionais como PISA²⁵ (Programa Internacional de Avaliação de Estudantes), SAEB (Sistema de Avaliação da Educação Básica) e ANA (Avaliação Nacional de Alfabetização), que expõem socialmente seus resultados e classifica as escolas em *rankings*. Essa exposição tem, então, levado diretores e secretários de educação a buscarem estratégias não de melhoria do ensino-aprendizagem, mas de um melhor desempenho nessas provas – fato igualmente vivido no ensino médio, em que boas são as escolas classificadas pelo *ranking* do ENEM (Exame Nacional do Ensino Médio). Portanto, esses índices – ou posição no *ranking* de bom desempenho – precisam ser levados em conta se considerarmos o modo como são preferencialmente trabalhados os conteúdos relacionados à escrita nos dois tipos de escolas.

4.2 Anos letivos

A estatística descritiva e inferencial sobre os resultados de anos letivos mostra uma diminuição ao longo dos cinco anos do Ensino Fundamental I, como já havia observado Capristano (2003) e Capelini (2010). Porém, tal qual Cunha (2010), observamos que essa queda no número e no percentual de segmentações não-convencionais de palavras ocorre

²⁵ O significado da sigla é original do inglês: “Programme for International Student Assessment”.

tanto em escolas públicas quanto em escolas privadas, mostrando-se altamente significativa na estatística inferencial. Outra observação interessante que pudemos realizar é que a diminuição não ocorre da mesma forma nos dois tipos de escolas; nas escolas públicas há uma diminuição mais homogênea, com a maior queda do 2º para o 3º ano, já nas escolas privadas o percentual de ocorrências é zerado a partir do 3º ano e a queda acentuada ocorre do 1º para o 2º ano. Percebemos, então, que a queda acentuada é posterior nas escolas públicas.

Esboçaremos, neste momento, possíveis explicações para essa diferença de queda. Em princípio, para nós, a queda mais homogênea em escolas públicas confirma o que se prevê nos PCN, o fato de que a aquisição da escrita não se encerra ao final do quinto ano do Ensino Fundamental I.

Capristano (2003) – analisando segmentações não-convencionais de palavras de quinze propostas textuais de três sujeitos – observa que “nem sempre o número menor de segmentações não-convencionais está relacionado ao maior domínio das convenções ortográficas” (CAPRISTANO, 2003, p. 153). Assim, para a autora, a explicação para a diminuição das segmentações não-convencionais se deve:

(1) às diferentes histórias de inserção em práticas sociais orais/letradas; (2) aos diferentes modos pelos quais o sujeito circula por um imaginário em torno da escrita alfabética em sua dita modalidade padrão; (3) às diferentes formas de relação do sujeito com as propostas temáticas e com gêneros discursivos que elas mobilizaram; e, por fim, (4) às diferentes extensões (tamanhos e/ou volumes) das produções textuais consideradas para o nosso trabalho (CAPRISTANO, 2003, p. 156).

Consideramos essas conclusões da autora para nosso método de coleta (ver seção 2.1), a fim de verificarmos se a variável de anos letivos tem efeito sobre as ocorrências de segmentações não-convencionais. Em nossa coleta, todos os escreventes fizeram a mesma proposta temática com o mesmo gênero discursivo. Sobre as diferentes extensões, para minimizarmos esse efeito, calculamos a razão entre o número de palavras e o número de ocorrências de segmentações não-convencionais, comparando os percentuais e não o número de ocorrências.

Após o referido cuidado na coleta dos dados, a variável anos letivos se mostrou estatisticamente significativa, indiciando que as práticas de letramento escolar têm influência sobre a escrita dos alunos, seja em escolas públicas ou em escolas privadas.

Em um estudo sobre erros ortográficos com crianças de 1º a 4º ano, Capellini, Butarelli e Germano (2010) afirmam que “os resultados, também, apontam para o fato de os escolares, com o aumento da seriação, diminuírem a média de erros na escrita.” (CAPELLINI; BUTARELLI; GERMANO, 2010, p. 158), o que, para as autoras, corrobora dados já obtidos por Chacon (2004) e por Capristano (2004). As autoras fornecem a seguinte explicação para a queda ou ausência de segmentações não-convencionais:

A diminuição dos erros tipo hipossegmentação, ao longo da seriação, e a ausência do erro tipo hipersegmentação na escrita dos escolares deste estudo evidenciam a reflexão destes sobre as características da língua, apontando para um funcionamento normal da construção e apropriação da escrita (CAPELLINI, BUTARELLI, GERMANO, 2010, p. 160).

As autoras concluem que as práticas de letramento às quais os alunos têm acesso na escola influenciam no modo de escrever deles, aproximando, cada ano mais, à escrita convencional.

A variável anos letivos também chamou à atenção de Cunha (2010) ao analisar segmentações não-convencionais de palavras. A autora concluiu que

Analisamos a variável *série* por considerarmos a aquisição da escrita como um processo progressivo, o que se confirmou com a observação dos dados [...]. De uma maneira muito positiva, essa variável nos mostrou que a escola desempenha um papel importante no processo de aquisição da escrita, pois a cada série que passa o número de ocorrências de segmentações não-convencionais diminui tanto nos textos das crianças da escola pública quanto da particular e mais radicalmente na escola pública. (CUNHA, 2010, p. 352)

Essa evidência de que há uma diminuição das segmentações não-convencionais de palavras em ambos os contextos também foi obtida em nossos resultados. A diminuição mais “radicalmente” expressa nas escolas públicas, obtida nos resultados de Cunha (2010), foi em decorrência de a autora trabalhar com dados do número de ocorrências e não o percentual – como fizemos. Como os textos de escreventes de escola pública têm mais ocorrências de

segmentações não-convencionais de palavras a queda no processo de anos letivos também é maior, mesmo não chegando aos mesmos números da escola privada ao final do Ensino Fundamental I.

Silva e Spinillo (1998), em sua pesquisa, comparando escolas públicas e privadas, têm como resultado que os anos de escolaridade foram “o principal fator responsável pela distribuição e variação dos dados” (SILVA; SPINILLO, 1998, p. 9). Assim, a variável anos letivos se mostrou estatisticamente significativa, “isto indica claramente um efeito facilitador dos anos de escolaridade sobre a habilidade narrativa das crianças ao escreverem histórias.” (SILVA; SPINILLO, 1998, p. 10), tanto na escola pública quanto na privada. Para os escreventes de ambas as escolas, no primeiro ano escolar, escrever uma narrativa é uma tarefa difícil, porém, com o passar dos anos escolares há um aprendizado de como escrever essas narrativas.

No entanto, essa diferença dos anos de escolaridade na escola pública e na privada não ocorre da mesma maneira “embora haja um efeito positivo dos anos escolares em ambas as escolas, este efeito parece ser menos marcante na escola pública do que na escola particular” (SILVA; SPINILLO, 1998, p. 10). De acordo com as autoras, após três anos de escolaridade, os escreventes de escolas privadas têm textos com narrativas mais elaboradas.

A escola, embora amplie as habilidades narrativas iniciais identificadas no primeiro ano após a alfabetização, não desenvolve na mesma extensão as habilidades narrativas das crianças das duas classes sociais (SILVA; SPINILLO, 1998, p. 13).

Nossos resultados corroboram aqueles de Silva e Spinillo (1998), na medida em que também observamos um efeito positivo dos anos letivos nos textos dos escreventes. Enquanto na pesquisa desenvolvida pelas autoras os escreventes, após anos de escolaridade, foram capazes de escrever narrativas com enredos mais elaborados, em nossa pesquisa, os escreventes aumentaram o número de palavras escritas em seus textos e diminuíram o percentual de ocorrências de segmentações não-convencionais de palavras. Vale ressaltar que

não compartilhamos da mesma associação direta de Silva e Spinillo (1998) de tipos de escolas com classe social (escola pública/ classe baixa e escola privada/ classe alta). Isso porque acreditamos que na variável tipos de escolas (públicas e privadas) há influência de outra variável: região da cidade – a qual não pudemos investigar pela não participação de escolas de todas as regiões; não sabemos se é possível comparar escolas públicas centrais com periféricas.

Também, a nosso ver, não necessariamente se pode/deve interpretar como problemática a diferença entre o modo como se desenvolvem as duas curvas, já que a queda brusca do primeiro ao segundo ano nas escolas privadas pode decorrer de uma ênfase forte colocada nesse tipo de escola no caráter fônico e estrutural das palavras. Diferentemente, nas escolas públicas, essa queda mais regular e contínua pode indiciar uma ênfase não tão grande nas atividades de ortografia mas, muito provavelmente, nas tarefas de produção textual.

Passemos, então, a analisar o que resta de ocorrências nos dados do 5º e último ano do Ensino Fundamental I. Vimos que, dos escreventes da amostra, quase 50% deles, nas escolas públicas, ainda segmentaram de forma não convencional, enquanto que, nas escolas privadas, esse percentual reduz-se para 30% dos escreventes. Analogamente, ao analisarmos escreventes que apresentam apenas uma ocorrência, 64,71% dos escreventes de escolas públicas ainda têm mais de uma ocorrência em seus textos, enquanto nas escolas privadas esse percentual é de 37,5% – o que ratifica os dados já apresentados de que os textos dos escreventes das escolas públicas têm mais ocorrências de segmentações não-convencionais de palavras do que os textos de escreventes de escolas privadas. Mas não só isso. Enquanto os textos dos escreventes de escolas públicas tendem a ter mais de uma ocorrência de segmentação não-convencional de palavras, os de escolas privadas não. Diferenças que geram necessidades e modos de planejamento particulares para escolas públicas e para escolas

privadas no Ensino Fundamental II. Para tanto, é necessário saber a natureza dessas segmentações não-convencionais.

Tratando-se dessa natureza, também há diferença entre a forma de segmentar dos escreventes das escolas públicas e privadas. Nas escolas públicas, há mais do que o dobro de hipossegmentações em relação a hipersegmentações. Já nas escolas privadas o número é similar. Esses dados mostram que, para os escreventes de escolas públicas, o conflito entre juntar e separar palavras tende a levá-los a junções, ao passo que, para os escreventes de escolas privadas, o conflito se distribui igualmente entre as duas possibilidades. As ocorrências de mesclas só ocorreram em textos de escreventes de escolas públicas e como foram apenas dois casos não se mostram como um problema a que deva ser dada atenção, como é o caso das hipossegmentações.

No interior desses dois tipos de ocorrências que ainda restam no último ano do ensino fundamental, um aspecto estrutural dessas ocorrências é comum aos dois tipos de escolas: a característica de elas, sempre, envolverem o que é e o que se pode interpretar como um clítico fonológico. Esse aspecto em comum antecipa, pois, a persistência dessa tendência nos anos escolares subsequentes (do Ensino Fundamental II), observada por Silva e Tenani (2014).

As questões de segmentação que persistem ao final do 5º ano do Ensino Fundamental I parecem, portanto, se relacionar mais diretamente com questões fonológicas relacionadas à palavra ortográfica, sobretudo aquelas que envolvem a possibilidade de haver palavras desprovidas de acento – característica essencial daquelas que se comportam como clíticos – e, ainda, de essas palavras muitas vezes serem idênticas a sílabas (especialmente pré-tônicas) de outras palavras da língua. Temos como exemplo os casos de hipossegmentação: *seder* (se der); *porai* (por aí); *comelo* (comê-lo); *medar* (me dar); e os casos de hipersegmentação: *a soprou* (assoprou); *ba teu* (bateu); *coren do* (correndo).

Uma diferença, no entanto, é digna de destaque quando se comparam as ocorrências de hipossegmentações dos textos de 5º ano de escolas privadas e públicas. Com efeito, das 07 ocorrências de escreventes de escolas privadas, 04 delas envolveram o emprego do hífen (uso inadequado ou ausência); já, nas públicas, de 111 ocorrências, apenas 03 o envolveram. Portanto, o aspecto que mais gera dúvidas nos escreventes de escolas privadas é de natureza fundamentalmente ortográfica, já que, como destaca Tenani (2011), o uso do hífen é decorrente de práticas letradas e a instabilidade quanto a seu uso mostra que os aspectos do letramento são os que mais fortemente estão em questão para esses escreventes, mesmo no uso não-convencional desse sinal. Em síntese, o que resta no ano final do Ensino Fundamental I advém de questões fonológicas e ortográficas bastante específicas.

4.3 Sexo/gênero

Quanto à variável sexo/gênero, nossos resultados mostraram que, nas escolas privadas, a mediana é zero tanto para o gênero masculino quanto para o gênero feminino. Já nas escolas públicas, a razão do gênero masculino para o gênero feminino é de 4, ou seja, os textos de escreventes masculinos têm 4 vezes mais ocorrências de segmentações não-convencionais de palavras.

Essa diferença de sexo/gênero presente nas escolas públicas e não nas escolas privadas explica, a nosso ver, a falta de consenso entre pesquisas que investigam a escrita de meninos e meninas. Porém, nos instiga a pensar por que os textos de meninos em escolas públicas apresentam mais ocorrências de segmentações não-convencionais do que os textos das meninas – diferença, como vimos, não existente em escolas privadas.

Sabemos que a diferença entre homens e mulheres é remetida preferencialmente às suas características biológicas (físicas); por isso, vistas como pré-estabelecidas e imutáveis. No entanto, a essa diferença física, ao se analisarem diferenças de desempenhos que

envolvem a diferença de sexo/gênero, deveriam, a nosso ver, ser acrescentadas as diferenças sociais que envolvem o sexo/gênero, já que homens e mulheres, tradicionalmente, exercem papéis diferentes na sociedade, com deveres e direitos diferentes atribuídos a eles.

A questão de sexo/gênero vem sendo discutida há pouco tempo. Em meados do século passado, as mulheres começaram a reivindicar seus direitos e, com isso, houve mudanças na sociedade (VIANNA; UNBEHAUM, 2004). A partir do movimento feminista em 1960, as mulheres – engajadas em movimentos sociais – deram origem ao conceito de gênero, diferenciando-o de sexo. A desigualdade de sexo/gênero, então, se tornou um tema a ser discutido, pois era (e ainda é) necessário rever os direitos pessoais e sociais das mulheres (VIANNA; UNBEHAUM, 2004, p. 22).

A Constituição Federal, a Lei de Diretrizes e Bases e o Plano Nacional da Educação não fazem menção à questões de sexo/gênero e desigualdades nesse âmbito. O que há nesses documentos oficiais é uma orientação ao trabalho com as diferenças, entre elas, a de sexo.

Na esfera educacional,

as políticas que são implantadas nas escolas e as identidades subjetivas que muitas vezes sustentam (...) procuram reverter o modelo dominante de masculinidade/feminilidade, como um modo de dar significado às relações de poder estabelecidas e difundidas pelas políticas educacionais nas suas mais variadas esferas, níveis e modalidades de ensino (VIANNA; UNBEHAUM, 2004, p. 23).

Na Inglaterra, os dados de testes e provas têm mostrado um desempenho pior dos meninos em relação às meninas nas últimas três décadas. Esses resultados têm sido vistos de forma tão alarmante, que esse desempenho dos meninos vem se tornando uma das prioridades da questão de sexo/gênero (ORR, 2000, p. 23). Porém, tal como no Brasil, esses resultados são divulgados sem que se leve em consideração a heterogeneidade que, socialmente, constitui os gêneros. Assim, comparar meninos de classes sociais privilegiadas com meninos de classes populares se torna bastante questionável (POVEY, 2004).

Tanto na Inglaterra como no Brasil, as questões relativas ao ensino têm mudado. Os indivíduos oriundos de classes baixas tiveram mais acesso à alfabetização e as mulheres estão apresentando índices melhores no desempenho da maioria das disciplinas – salvo as disciplinas da área de exatas (POVEY, 2004, p. 160). Porém, o que em um olhar inocente parece um ganho para as mulheres, em um olhar mais atento vemos que as condições de vida e de trabalho continuam desprivilegiadas, já que elas ainda têm cargos inferiores e salários menores que os homens.

Mas por que, então, mesmo com desempenho escolar melhor, as mulheres não ocupam, ainda, os empregos melhores? Uma possível resposta está relacionada ao condicionamento das meninas a determinadas áreas e/ou profissões, supostamente mais condizentes ao perfil feminino.

O “mau” desempenho escolar dos meninos também tem sido discutido com relação à forma de aceitação do comportamento masculino tradicionalmente instituído pela sociedade em que o homem – ou, no caso, os meninos – são fortes, agressivos, gostam de esportes e não de estudar, não conseguem se concentrar, são bagunceiros e, por isso, não se adaptam bem à rígida disciplina da escola (FARIA, NOBRE, 1997). Ao homem, sempre foi imposto/esperado que sejam os provedores, os mantenedores da família, assim desde crianças são cobrados por trabalhar. Embora as meninas também estejam sujeitas a cobranças sociais, os propósitos (ainda) são bastante diferentes – as meninas precisam trabalhar em casa, para cuidar e manter a casa em ordem, já os meninos precisam trabalhar para ganhar dinheiro, ajudar no sustento da família (FARIA, NOBRE, 1997). Esse condicionamento do comportamento de meninos e de meninas surge desde antes mesmo de a criança nascer, quando são comprados os brinquedos: das meninas, bonecas e panelinhas; dos meninos carrinhos e ferramentas.

Essa questão de as meninas serem estimuladas nas brincadeiras a profissões de cuidados nos faz entender por que profissões como pedagogia, letras, enfermagem, secretariado são dominadas por mulheres.

Em reportagem do *site* G1.com²⁶, foi divulgada a história das trigêmeas campeãs das Olimpíadas de Matemática das Escolas Públicas (OBMEP). As meninas são oriundas de uma família de classe social baixa e rural, filhas mais novas, com uma irmã e um irmão mais velhos. Na família, com a chegada das trigêmeas, a mãe e o filho mais velho foram para a roça ajudar o pai no sustento da família, enquanto a filha mais velha cuidava das meninas. Mesmo assim, a filha mais velha conseguiu estudar, ganhou bolsa para fazer faculdade de enfermagem (sendo a primeira pessoa de toda a região a conseguir fazer graduação) e hoje faz doutorado. Os efeitos dessa formação escolar acabaram chegando às irmãs mais novas, com bons resultados. Entretanto, o que nos chama à atenção é que, enquanto as filhas meninas conseguiram mudar sua perspectiva de futuro com o estudo, não se falou mais na reportagem sobre o filho homem, o que nos faz supor que, por estar trabalhando na roça, ele não tenha conseguido seguir com os estudos e ter o mesmo destino das irmãs.

Esse caso é bastante comum em famílias de baixa renda e amplia as discussões sobre oportunidades e estímulos à educação diferenciada entre meninos e meninas nesse tipo de contexto social.

Carvalho (2004, p. 181) mostra que, em 1960, o número de anos de escolarização era maior entre os homens. Já em 1996, esse número aumentou entre homens e mulheres, mas as mulheres passaram a ter mais anos de escolarização do que os homens. No entanto, para a autora, essa mudança não significa um ganho para as mulheres, porque, mesmo com mais anos de escolarização, as mulheres ainda não ocupam os melhores cargos nem têm salários

²⁶ Reportagem “Ouro na matemática, trigêmeas do ES não tinham nem internet em casa” publicada no site G1.com, de Naiara Airpini. Disponível em: <http://g1.globo.com/espírito-santo/educacao/noticia/2015/07/ouro-na-matematica-trigemeas-do-es-nao-tinham-nem-internet-emcasa>. Acesso: 09/07/2015 às 10:00.

melhores que os homens. Há, portanto, fatores bastante complexos quando se trata diferença de gênero em qualquer análise de dados em que ela se cruza com a escolarização:

Devemos considerar que esses dados são todos muito gerais e que, para serem efetivamente entendidos, precisariam ser divididos por regiões do país, por área urbana e rural, por classe, raça/cor e etnia dos alunos e alunas, para que pudéssemos perceber quais os grupos mais atingidos pelas dificuldades escolares (CARVALHO, 2004, p. 183).

A autora também questiona testes aplicados em esfera nacional com categorias amplas como país, estado, município, escola, seriação, nos quais as diferenças intergêneros estão envolvidas:

Pensar a partir da categoria teórica de gênero nos dá a possibilidade de entender as relações entre os sexos também no âmbito da cultura, do simbólico, das representações, e isso é muito importante quando se pensa em educação, porque, quando trabalhamos nessa área, reconstruímos a cultura, os valores, os símbolos nas novas gerações, transmitindo ou recriando, reproduzindo ou transformando as hierarquias, as diferentes importâncias atribuídas socialmente àquilo que é associado ao masculino e ao feminino (CARVAHO, 1999, p. 09).

Como mostramos com nossos resultados, a categoria sexo/gênero, analisada isoladamente, não retrata a real diferença no aprendizado de meninos e meninas. A desigualdade de sexo/gênero é, portanto, uma construção política e social, ratificada por resultados generalistas e difundida por discursos conservadores quanto à diferenciação de papéis sociais relacionados ao sexo/gênero. A construção social dessa diferença é feita principalmente no interior da família, que, especialmente nos setores menos favorecidos socialmente, permite à menina estudar, mas ao menino trabalhar e ajudar no sustento da casa.

Nossos resultados indiciam não haver, pois, simplesmente uma diferença de gênero/sexo, mas uma diferença socialmente construída, a partir de discursos que geram expectativas e diferentes construções do gênero/sexo nos diferentes estratos sociais. A famílias de classes sociais mais privilegiadas sonham que seus filhos meninos sejam médicos, engenheiros; já famílias de classe mais desfavorecidas socialmente precisam que seus filhos, principalmente os meninos, ajudem na renda doméstica – levando assim os meninos a

trabalharem cedo, com a expectativa de que sejam, no máximo, bons profissionais técnicos em áreas com a mecânica e a construção civil. Com as meninas há também essa diferença: meninas de estratos mais favorecidos são criadas para serem profissionais com nível superior em carreiras valorizadas socialmente; já meninas de famílias de estratos menos favorecidos, além da ajuda nos cuidados com a casa, são criadas com a expectativa de que sejam secretárias, vendedoras ou, algumas vezes, professoras. A diferença de desempenho intergêneros em escolas públicas pode ser explicada, portanto, pelo fato de que as expectativas projetadas nas meninas desse contexto envolvem trabalhos que mais exigem fala e escrita na norma padrão, o que as incentivam a estudarem mais, enquanto os meninos, para exercerem melhor suas tarefas (menos verbais) fazem cursos profissionalizantes.

As expectativas do futuro profissional para o que esses meninos e meninas serão indiciam que a linguagem deixa de ser apenas produto de uma atividade psicológica do sujeito para ser uma atividade constitutiva do sujeito no mundo (DE LEMOS, 1989), inserção que influencia ou determina os resultados de suas práticas sociais e escolares. É a interpretação do adulto que coloca a criança no funcionamento da língua, porque para ela essa é a possibilidade constitutiva de enquadrar-se na fala do outro:

No momento em que é interpretada pelo outro, a criança produz rearranjos na sua relação com a língua escrita e são por essas reformulações que ela é posta no funcionamento da estrutura da língua (OLIVEIRA; RENZO, 2009, s/p).

A relação entre a fala da criança e do adulto é tida por De Lemos (2002) como um processo dialógico no qual a

aparente coesão e progressão dialógicas que emergiam como efeito tanto do enunciado com que o adulto interpretava o que a crianças fazia ou falava de modo ainda informe e fragmentado, quanto pela dependência que essa fala fragmentada mostrava ao se ancorar na fala do adulto, muito particularmente na fala da mãe (DE LEMOS, 2002, p. 46).

A criança é significada pelo outro porque é ele quem estrutura a sua entrada e posição na língua (DE LEMOS, 1992) – e, conforme pensamos, no mundo. Essa entrada se dá por meio de “processos constitutivos do diálogo e da aquisição da linguagem” (DE LEMOS,

2002, p. 46). O primeiro processo é o de especularidade, que se caracteriza pela “presença na fala da crianças de parte do enunciado da mãe que o antecede, assim como pela incorporação da fala da criança no enunciado da mãe” (DE LEMOS, 2002, p. 46); o segundo processo é o de complementaridade, que indicia a “relação da pergunta da mãe com a resposta da criança e, principalmente, pela relação formal entre partes mutuamente incorporadas que parecem completar-se, compondo uma unidade ou instanciando uma ‘sentença’” (DE LEMOS, 2002, p. 46-47); por fim, o terceiro processo é o de reciprocidade, que se caracteriza pela “retomada pela criança do papel da mãe, iniciando a interação, desencadeando com sua fala uma fala que refletiria e completaria a sua” (DE LEMOS, 2002, p. 47).

Nessa interação entre a fala da criança e a do adulto, e pensando-a no que se refere à escrita,

É do sistema do outro que vêm os significantes com os quais a criança comporá sua escrita. A singularidade desta se faz na medida em que esses significantes, presentes na escrita da criança por incorporação ou espelhamento, não são considerados uma mera repetição do mesmo (BOSCO, 2002, p. 72).

A relação que essa autora faz do processo de espelhamento proposto por De Lemos (1992) nos ajuda a pensar, de forma mais ampla, que a criança (menino ou menina) espelha a fala do adulto, bem como seu discurso, suas expectativas. Assim, reproduz não só a estrutura linguística, como também o papel social construído pelos discursos aos quais é exposta:

É, pois, o outro (adulto falante) quem primeiro interpreta e significa os sons produzidos pela criança, inserindo-os na rede da linguagem. Desse modo, acredito, constitui-se uma noção de subjetividade/alteridade na relação entre a criança e o adulto falante (KOMESU, 2002, p. 59).

É nesse processo de subjetividade/alteridade que a criança traz para sua escrita o outro (adulto) que a constitui e a interpreta como sujeito, bem como a expectativa do papel social e do papel da escrita projetado para ela (a criança). Esse adulto do qual falamos tanto pode ser os pais, a família, quanto os professores ou meio social. Assim

a interação da criança com o outro (...) que possibilitará à primeira realizar movimentos interpretativos em direção à constituição de sua escrita. Movimentos da ordem da língua estarão em jogo, efeitos de relações entre elementos de qualquer natureza e extensão, abertos a novas possibilidades de sentido (BOSCO, 2002, p. 75).

Então, de onde e de que lugar esse outro significa a criança?

A diferença apontada pelos resultados sugere uma possível antecipação/projeção dos lugares que meninos e meninas poderão/deverão vir a ocupar – seja em classes altas ou baixas. Nessa antecipação, no que se refere às escolas públicas, o que se pode indiciar pelos resultados é que os meninos são menos projetados/antecipados para virem a desenvolver tarefas/trabalhos que envolvem a escrita ou tarefas/trabalhos em que a escrita teria maior importância do que outras habilidades/conhecimentos. Já as meninas são projetadas para desenvolver tarefas em que a escrita pode ter um lugar de importância, mesmo em trabalhos com baixa remuneração como domésticas, secretárias ou vendedoras.

No que se refere às escolas privadas, a ausência de diferença na escola privada indicia que as projeções feitas para as crianças já supõem que suas futuras tarefas terão a escrita como um elemento de destaque – ou de bastante importância –, como as profissões com formação universitária privilegiadas socialmente.

CONCLUSÕES

Para finalização desta tese, gostaríamos de destacar o que consideramos como mais significativo da pesquisa que realizamos. Assim não serão retomadas as discussões teóricas e o percurso metodológico apresentados; enfatizaremos, aqui, as contribuições que – acreditamos – nossa pesquisa pode trazer aos estudos de segmentações não-convencionais de palavras, bem como aos estudos sobre as variáveis tipos de escolas (públicas e privadas), anos letivos e sexo/gênero (masculino e feminino).

A primeira parte de nosso objetivo foi analisar como as ocorrências de segmentações não-convencionais de palavras se distribuíam em função da variação dos *tipos de escolas* (públicas e privadas). Nossa hipótese foi de que havia diferença entre a forma de segmentar de escreventes de escolas públicas e escreventes de escolas privadas, mas não sabíamos como se constituiria essa diferença. Nossos resultados mostraram que a distribuição das segmentações não-convencionais de palavras em relação à variável *tipos de escolas* (públicas x privadas) apresentou médias e medianas com grande diferença – e estatisticamente significativa. Assim, respondendo ao nosso principal questionamento, obtivemos o índice de diferença de 1,15 vezes para o número de palavras e de 3 vezes para o percentual de segmentações não-convencionais de palavras. Depreendemos, então, que escreventes de escolas públicas e privadas não se distanciam muito com relação ao quanto escrevem, mas, sim, com relação a como segmentar palavras. Apesar de mais ocorrências de segmentações não-convencionais de palavras nos textos de escreventes de escolas públicas do que de escreventes de escolas privadas, isso não nos autoriza, nem é o que gostaríamos, pensar que os escreventes de escolas públicas têm mais problemas de ortografia ou escrevem pior. Nem poderíamos, pois, como procuramos mostrar, há questões de naturezas distintas imbricadas na (distribuição da) escrita e na do seu ensino.

Vimos que, embora os índices de segmentações não-convencionais de palavras tenham sido diferentes entre escreventes de escolas públicas e privadas, quando buscamos significado para essas diferenças, nossa análise mostrou que, em outros estudos, há mais questões mais complexas que precisariam/deveriam ter sido notadas, como, por exemplo, a heterogeneidade do sistema público de ensino em esfera municipal, estadual e federal. A partir dessas discussões, consideramos que os índices de desempenho de exames internacionais ou nacionais são necessários como um norte do que precisa ser pensado, mas não podem, nem devem, ser usados para taxar a aprendizagem dos escreventes, porque nesse processo questões históricas, políticas e sociais estão fortemente envolvidas e interligadas.

Outra questão que precisa ser considerada é o bom índice de desenvolvimento humano educacional (IDH-E: 0,962) de Marília – diferentemente de grande parte dos municípios brasileiros –, que alcança o patamar mais alto e ajuda a elevar o IDH do município, pois é mais elevado do que o IDH de renda (0,885) e o de longevidade (0,822). Esses dados mostram que, no município pesquisado, a renda e a saúde da população apresentam índices médios na escala de análise de IDH, juntamente com a maioria dos municípios do estado. A educação, porém, apresenta índice alto, alcançado por poucos municípios. Desse modo, nossos dados e resultados são provenientes de um sistema de ensino ótimo de um dos estados com melhor nível de ensino do país. Portanto, não é possível generalizá-los aos de outros contextos do Brasil, nem, tampouco, afirmar que as escolas públicas no Brasil têm resultados similares aos das escolas privadas. Os resultados que mostramos refletem o que ocorre em Marília-SP e/ou, provavelmente, em municípios com investimentos em educação como esse.

A segunda parte de nosso objetivo foi analisar como as ocorrências de segmentações não-convencionais de palavras se distribuiriam em função da variação dos *anos letivos* (1º ano, 2º ano, 3º ano, 4º ano e 5º ano). Embora nossa hipótese fosse a de que haveria uma diminuição ao longo dos cinco anos letivos do Ensino Fundamental I, queríamos saber se essa

diminuição seria diferente nas escolas públicas e privadas e, em caso afirmativo, qual seria essa diferença entre a forma de segmentar de escreventes de escolas públicas e escreventes de escolas privadas ao longo dos cinco anos letivos do Ensino Fundamental I. Os resultados que obtivemos, em relação à variável *anos letivos*, em ambos os tipos de escolas (públicas e privadas), mostraram que a diferença comparativa entre os anos se mostrou estatisticamente significativa nas escolas públicas e privadas. Porém, essa diminuição de ocorrências do primeiro para o quinto ano ocorreu de maneira diferentes nos tipos de escolas. Nas públicas observamos uma diminuição contínua, chegando a uma mediana de 0,5% de ocorrências no quinto ano; já nas privadas essa diminuição aconteceu de forma acentuada do primeiro para o segundo ano, chegando a 0 no terceiro ano. Entretanto, vale ressaltar que, no primeiro ano, os escreventes de escolas públicas iniciaram com um percentual de 12,8 e os de escolas privadas com 7,3. Desse modo, a trajetória do desempenho dos escreventes de escolas públicas apresentou ganhos que consideramos como superiores em relação aos de escreventes de escolas privadas, já que, ao final do percurso, ambos os tipos de escreventes mostraram desempenhos próximos.

Em outras palavras, embora a variável anos letivos tenha se mostrado como significativa para ambos os tipos escolares, é importante destacar a importância do acréscimo de um ano no Ensino Fundamental I para as escolas públicas, pois, como nesse tipo de escola a diminuição do percentual de segmentações não-convencionais foi contínua e gradativa, um ano a mais no novo sistema educacional pode ter favorecido a diminuição de diferença a que se chega no último ano, quando se comparam os dois tipos de escola.

Quanto às segmentações não-convencionais de palavras que sobraram no quinto ano, analisamos sua natureza, a fim de verificar quais seriam as principais dúvidas que seguiriam com os escreventes para o Ensino Fundamental II. Vimos que, nas escolas públicas houve maior recorrência de hipossegmentações, o que nos parece ser em virtude de dúvidas quanto a

questões fonológicas para a determinação das fronteiras de palavra. Diferentemente, nas escolas privadas, essa questão parece ter sido melhor resolvida, restando apenas aquelas que, mais fortemente, tocam em aspectos ortográficos como o do hífen.

A terceira, e última, parte de nosso objetivo foi analisar como as ocorrências de segmentações não-convencionais de palavras se distribuíam em função da variação do *sexo/gênero* dos escreventes (masculino e feminino). Com base na literatura sobre a escrita infantil, nossa hipótese foi de que os textos de escreventes masculinos teriam mais ocorrências de segmentações não-convencionais de palavras. Os resultados que obtivemos mostraram, porém, que tanto nas escolas públicas quanto nas privadas, ao longo dos cinco anos letivos, há uma diminuição do percentual de ocorrências de segmentações não-convencionais de palavras nos textos de escreventes femininos e masculinos. Vemos que os anos letivos melhoram a forma de segmentar tanto nas escolas públicas quanto nas escolas privadas, tanto para escreventes do sexo/gênero feminino quanto masculino.

Outra comparação que realizamos foi entre textos de escreventes femininos de escolas públicas e privadas – que não se mostrou estatisticamente significativa. No entanto, em relação aos textos de escreventes masculinos de escolas públicas e privadas, os resultados foram altamente significativos. Desse modo, os escreventes do sexo/gênero feminino segmentam de forma similar nas escolas públicas e nas escolas privadas, mas os escreventes do sexo/gênero masculino não.

Verificamos, ainda, que, nas escolas públicas, os escreventes masculinos apresentaram uma mediana quatro vezes maior de ocorrências do que aquela obtida para os escreventes femininos. Essa diferença foi verificada na estatística descritiva e se mostrou altamente significativa na estatística inferencial. No entanto, nas escolas privadas, a diferença entre escreventes masculinos e escreventes femininos, além de bastante pequena, não se mostrou como estatisticamente significativa. Esse resultado, a nosso ver, pode levantar mais elementos

para se explicar a falta de consonância entre as conclusões de pesquisas que comparam a escrita de meninos e meninas. Isso porque – como destacamos e gostaríamos de reforçar – ao analisar a escrita e/ou seu processo de aprendizagem, é necessário considerar as heterogeneidade das múltiplas tarefas envolvidas na relação que a constituição social dos gêneros tem com a distribuição social dos usos da escrita.

Em outras palavras, acreditamos ter mostrado que a forma de segmentar dos escreventes têm forte relação não apenas com o tipo de escola em que eles estudam, mas, também, com a complexidade da construção social dos gêneros e das tarefas sociais associadas a essa construção.

Em decorrência de nossos prazos e do volume de dados que analisamos, tivemos que fazer recortes em nossa análise. Há, porém, muitas outras questões que merecem mais discussões e trabalhos de pesquisa. Gostaríamos, então, de chamar a atenção para algumas delas – baseadas em observações que não constam da escrita desta tese, bem como em nossa percepção/intuição dos dados. Vamos a elas.

Quantificamos as hipo, hiper e mesclas em função das nossas variáveis (1 – tipos de escolas; 2 – anos letivos; 3 – gênero/sexo); porém, não foi possível tratar estatisticamente esses dados para termos uma noção mais real sobre a proporção da diferença.

Também ao analisarmos esses diferentes tipos de segmentação não-convencional, ocorrências, não fizemos uma análise mais minuciosa dos constituintes prosódicos envolvidos nesses diferentes tipos. No entanto, é de fundamental importância para um estudo mais qualitativo do tipo de dados para os quais nos voltamos que essa análise minuciosa seja feita, já que pode dar indícios mais seguros da diferença ou da semelhança de resultados em função das variáveis com que os analisamos. A título de exemplo, em observações não-sistematizadas, acreditamos ter detectado mais ocorrências de segmentações não-convencionais em sintagmas nominais do que nos sintagmas verbais – detecção que, se

correta, sugere que as relações entre questões fonológicas e gramaticais devem influenciar o modo de segmentar dos escreventes.

Ao final da leitura de todos os textos que compuseram nossa amostra, julgamos haver mais marcas de autoria e de interação nos textos dos anos iniciais do que nos textos dos anos finais. Se nosso julgamento puder ser confirmado, teremos indícios de que a variável anos letivos traz uma diminuição nas ocorrências de segmentações não-convencionais de palavras, mas, também, uma padronização de aspectos da escrita de narrativas.

Gostaríamos também de ter analisado a variável zona urbana e rural, por termos dados que possibilitariam essa análise. Porém, fazê-la levaria a outro grande conjunto de discussões cujo âmbito nos pareceu inalcançável para o tempo de desenvolvimento da pesquisa que resultou nesta tese.

Trabalhos futuros, no entanto, poderão se voltar para as questões que deixamos em aberto em nossa investigação.

REFERÊNCIAS

ABAURRE, M. B. M. The interplay between spontaneous writing and underlying linguistic representations. **European journal of psychology of education**, Portugal, v. III, n. 4. p. 415-430, 1988.

_____. Hipóteses iniciais de escrita: evidências da percepção, por pré-escolares, de unidades rítmico-entonacionais na fala, 12/1989, Científico Nacional, IV Encontro Nacional da ANPOLL, Vol. 13, PUC/SP, SP, BRASIL, 1989.

_____. A relevância dos critérios prosódicos e semânticos na elaboração de hipóteses sobre segmentação na escrita inicial. **Boletim da Abralín**, Curitiba, v. 11, p. 203-17, 1991.

_____. Os estudos lingüísticos e a aquisição da escrita. In: CASTRO, M. F. P. (Org.) **O método e o dado no estudo da linguagem**. Campinas: Editora da Unicamp. p. 111-78, 1996.

_____. Horizontes e limites de um programa de investigação em aquisição da escrita. In: LAMPRECHT, Regina (org.). **Aquisição da linguagem: questões e análises**. Porto Alegre: EDIPUCRS. p.167-186, 1999.

_____. Apresentação. **A construção fonológica da palavra**. 1ª ed. São Paulo: Contexto, 2013.

ABAURRE, M. B. M.; CAGLIARI, L. C. Textos espontâneos na primeira série: evidência da utilização, pela criança, de sua percepção fonética para representar e segmentar a escrita. **Cadernos Cedex**, v. 14. São Paulo: Cortez. p. 25-29, 1985.

ABAURRE, M. B. M.; FIAD, R. S.; MAYRINK-SABINSON, M. L. T. **Cenas de Aquisição da Escrita: O Trabalho do Sujeito Com O Texto**. 1ª. ed. Campinas, SP: Mercado de Letras. 204 p. 1997.

ABAURRE, M. B. M.; SILVA, A. O desenvolvimento de critérios de segmentação na escrita. **Temas em psicologia**. São Paulo, v. 1, p. 89-102, 1993.

ANDREAZZA-BALESTRIN, C.; CIELO, C. A.; LAZZAROTTO-VOLCÃO, C; LAUSCH, S. S. Habilidades em consciência fonológica: diferenças no desempenho de meninos e meninas. **Revista CEFAC**, v. 14, n. 4, p. 669-676, 2012.

BARRERA, S. D. **Linguagem oral e alfabetização**: um estudo sobre variação lingüística e consciência metalingüística em rianças da 1ª serie do Ensino Fundamental (tese de doutorado) São Paulo, 2000.

BASILIO, M. Em torno da palavra como unidade lexical: palavras e composições. **Veredas - Revista de Estudos Linguísticos**: Juiz de Fora, v. 4, n. 2, 2000. p. 9 a 18.

_____. **Formação e classes de palavras no português do Brasil**. São Paulo: Contexto, 2004.

BECHARA, E. **Moderna gramática portuguesa**. Rio de Janeiro: Lucerna, 2006.

BERLINCK, R. A. ; AUGUSTO, M. R. A. ; SCHER, A. P. . Sintaxe. In: Fernanda Mussalin; Anna Christina Bentes. (Org.). *Introdução à Linguística: domínios e fronteiras*. 1 ed. São Paulo: Cortez Editora, 2001, v. 1, p. 207-244.

BISOL, L. Sândi externo: o processo e a variação. In: KATO, M. (org.). **Gramática do português falado**. Convergências. Editora da UNICAMP, v. 5, p. 55-95, 1996.

_____. O clítico e seu status prosódico. **Revista de Estudos da Linguagem**, Belo Horizonte, v. 9, n.1, p. 5-20, 2000.

_____. Mattoso Câmara Jr. e a Palavra Prosódica. **DELTA**, São Paulo, v. 20, n. especial, p.59-70, 2004.

_____. O clítico e o seu hospedeiro. **Letras de Hoje**, Porto Alegre, v.40, n. 3, p.163- 184, 2005.

BLOOMFIELD, L. **A Set of Postulates for the Science of Language**. *Language* 2, 153-64., 1926. (In: *Readings in Linguistics*, 26-31).

BOSCO, Z. R. **No jogo dos significantes, a infância da letra**. Campinas, SP: Pontes, 2002.

BRASIL. Ministério da Educação. **Parâmetros Curriculares Nacionais: ensino médio**. Brasília: MEC, 2000.

CAGLIARI, Luiz Carlos. **Alfabetização & Linguística**. São Paulo: Editora Scipione, 2002.

CÂMARA JR, J. M. **Dicionário de fatos gramaticais**. Rio de Janeiro: Casa de Rui Barbosa, 1956.

_____. **Problemas de Linguística Descritiva**. 4a. ed. Petrópolis: Vozes, 1971.

_____. **Princípios de Linguística Geral**. Como Introdução aos Estudos Superiores da Língua Portuguesa. 4a. ed. rev., ampl.. Rio de Janeiro: Acadêmica, 1972.

_____. **Dicionário de Linguística e Gramática**. 11. ed. Petrópolis: Vozes, 1988.

_____. **Estrutura da Língua Portuguesa**. 26a. ed. Petrópolis: Vozes, 1997.

CAPPELLINI, S.; BUTARELLI, A.P.K.J.; GERMANO, G. D. Dificuldades de aprendizagem da escrita em escolares de 1ª a 4ª séries do ensino público. **Revista Educação em Questão**, Natal, v. 37, n. 23, p. 146-164, jan./abr, 2010.

CAPRISTANO, C. C. Aspectos de segmentação na escrita infantil. São José do Rio Preto, 2003. Dissertação. 213f. (Mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Unesp.

_____. A propósito da escrita infantil: uma reflexão sobre as segmentações não-convencionais. **Letras de Hoje**, Porto Alegre, v.39, n.3, p.245- 260, 2004.

_____. **Mudanças na trajetória da criança em direção à palavra escrita**. 2007. Tese (Doutorado em Linguística Aplicada) - Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Campinas, 2007.

CAPRISTANO, C CHACON, L. Relações metafóricas e metonímicas: notas sobre a “aquisição” da noção de palavra. In: TFOUNI, L. V.; TONETO, D. J. (Org.) **O (In)esperado de Jakobson**. Campinas: Mercado de Letras, 2014 (prelo).

CAPRISTANO, C. C. ; TICIANEL, G. F. . Discurso direto e hipossegmentações na escrita infantil. **Revista Filologia e Linguística Portuguesa** , v. 16, p. 233-259, 2014.

CEREJA , W. R. e MAGALHÃES, T. C. **Português: linguagens** – Ensino Médio. V. 1, 2 e 3. São Paulo-SP, Ed. Atual, 2009.

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

_____. Algumas palavras sobre a aquisição da pontuação. In: LAMPRECHT, R.R. (Org.) **Aquisição da linguagem**: questões e análises. Porto Alegre: Editora da PUCRS, 1999.

_____. Constituintes prosódicos e letramento em segmentações não-convencionais. **Revista Letras de Hoje**. Porto Alegre. v. 39, nº. 3, p. 223-232, 2004.

_____. Hipersegmentações na escrita infantil: entrelaçamentos de práticas de oralidade e de letramento. **Estudos Linguísticos**, Campinas, v. XXXIV, p. 77- 86, 2005.

_____. Prosódia e letramento em hipersegmentações: reflexões sobre a noção de palavra. In: CORRÊA, M. L. G.; BOCH, F. (org.). **Ensino de língua**: representação e letramento. Campinas (SP): Mercado de Letras. p. 155-167, 2006.

_____. Concepções de "palavra" em escolas de educação infantil. **Scripta** (PUCMG), v. 13, p. 165-178, 2010.

_____. Segmentações não-convencionais na escrita de pré-escolares: entrecruzamentos entre convenções ortográficas e constituintes prosódicos. In: LAMPRECHT, R. R. **Aquisição da linguagem**: estudos recentes no Brasil. Porto Alegre: EDIPUCRS, 2011.

_____. Flutuação na segmentação de palavras: relações entre constituintes prosódicos e convenções ortográficas na escrita infantil. **Revista Filol. linguíst. port.**, São Paulo, 15(2), p. 369-383, Jan./Jun. 2013.

CORRÊA, M. L. G. **O modo heterogêneo de constituição da escrita**. São Paulo: Martins Fontes, 2004.

CUNHA, A. N. A hipo e a hipersegmentação nos dados de aquisição da escrita: um estudo sobre a influência da prosódia. 2004. 130 f. Dissertação (Mestrado em Educação) - Universidade Federal de Pelotas. Orientador: Ana Ruth Moresco Miranda.

_____. As segmentações não-convencionais da escrita e sua relação com os constituintes prosódicos. **Cadernos de Educação** (UFPel), v. 35, p. 323-358, 2010.

CUNHA, A. P. N.; MIRANDA, A. R. A hipersegmentação nos dados de aquisição da escrita. **Anais do 4º SENALE – UCPEL**, Pelotas, 2006.

_____. A influência da hierarquia prosódica em hipossegmentações da escrita de crianças de séries iniciais. **Revista Virtual de Estudos da Linguagem**. v. 1, p.1-19, 2007.

_____. Problemas de segmentação da escrita em séries avançadas: em busca de soluções através de ensino reflexivo. In: XIV Encontro Nacional de Didática e Prática de Ensino, 2008, Porto Alegre. **Anais do XIV ENDIPE**. Porto Alegre: EDIPUCRS. v.1. p.1-19, 2008.

_____. As segmentações não-convencionais da escrita: a influência do traqueu silábico. In: VI Congresso Internacional da ABRALIN, 2009, João Pessoa. **Anais - VI Congresso Internacional da Abralín**. João Pessoa: ideia, 2009. v. 1. p. 315-322.

CUNHA, C. & CINTRA, L. **Nova gramática do português contemporâneo**. Rio de Janeiro: Nova Fronteira, 2001.

DONADEL, G. 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.

_____. As palavras dentro da palavra: segmentações não convencionais na escrita de estudantes do Ensino Médio e sua relação com o estatuto da palavra. 2013. 110 f. Tese (Doutorado em Teoria e Análise Linguística) – Universidade Federal do Rio Grande do Sul. Orientador: Luiz Carlos da Silva Schwindt.

FERREIRA, A. B. H. **Dicionário escolar da língua portuguesa**. Curitiba: Positivo, 2011.

FERREIRO, E. PONTECORVO, C. Os limites entre as palavras. A segmentação em palavras gráficas. In: FERREIRO, E. PONTECORVO, C. MOREIRA, N. HIDALGO, I. G. **Chapeuzinho Vermelho aprende a escrever**. São Paulo: Ática, p.38-66, 1996.

KOMESU, F. Diálogo e dialogismo no processo de aquisição de linguagem. São Paulo: **Alfa**, 2002.

LIMA, C. H. R. **Gramática normativa da língua portuguesa** — Curso médio. 14. ed. Rio de Janeiro: F. Briguet & Cia Editores, 1969.

MEIRELLES, E. Como trabalhar a segmentação convencional das palavras. *Revista Nova Escola*. Disponível em: <http://revistaescola.abril.com.br/fundamental-1/como-trabalhar-segmentacao-convencional-palavras-744682.shtml?page=1#>
Acesso em: 10/02/2015.

MENESES, M. S. D.; LOZI, G. P.; SOUZA, L. R. D.; ASSENCIO-FERREIRA, V. J. Consciência fonológica: diferença entre meninos e meninas. *Revista CEFAC*, v. 6, n. 3, p.242-246, 2004.

MIRANDA, A. R. Um estudo sobre a aquisição ortográfica das vogais do português. *Anais da ANPEDESul – UFSM, Santa Maria*, 2006.

_____. Aspectos da escrita espontânea e da sua relação com o conhecimento fonológico. In: LAMPRECHT, R. **Aquisição da Linguagem: estudos recentes no Brasil**. Porto Alegre: EDIPUCRS, 2008.

MONTT, G. Escolas privadas: quem sai ganhando? Disponível em: <http://www.oecd.org/pisa/pisaproducts/pisainfocus/48726101.pdf>
Acesso em: 27/07/2015.

NESPOR, M.; VOGEL, I. **Prosodic Phonology**. Dordrecht, Holland: Foris. 1986.

NEVES, M. H. M. **Gramática de usos do português**. São Paulo: Ed. UNESP, 2000.

NICOLA, J. **Gramática da palavra, da frase, do texto**. São Paulo: Scipicione, 2004.

ORR, P. Prudence and progress: national policy for equal opportunities (gender) in schools since 1975. In: MYERS, K. (ed.) **Whatever happened to equal opportunities in school? Gender equality initiatives in education**. Buckingham: Open University Press, 2000.

PARANHOS, F. C. Segmentações não-convencionais de palavra nos quatro últimos anos do Ensino Fundamental: um estudo longitudinal. São José do Rio Preto, 2014. Dissertação. 170f. (Mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Unesp.

PEREIRA, E. C. **Gramática expositiva**. São Paulo: Editora Nacional, 1958.

PERINI, M. A. **Gramática descritiva do português**. São Paulo: Ática, 1996.

POSSENTI, S. **Por que (não) ensinar gramática na escola**. Campinas – SP: Mercado das Letras, 1996.

POVEY, H. Dimensões de gênero e desigualdade na escola: o impacto das identidades.

SANDALO, M. F. Morfologia. In: MUSSALIN, F.; BENTES, A. C. (orgs.). **Introdução à Lingüística: Domínios e Fronteiras**. Vol. 1. São Paulo: Cortez Editora, 2001.

SELKIRK, E. On derived domains in sentence phonology. **Phonology Yearbook**, nº 3, 1986.

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

SILVA, A. **Alfabetização**: a escrita espontânea. São Paulo: Contexto, 1994.

SILVA, L. M. 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.

SILVA, L. M.; TENANI, L. Hipersegmentações de palavras no Ensino Fundamental. São Paulo: Cultura Acadêmica, 2014.

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

TENANI, L. **Domínios prosódicos no Português do Brasil**: implicações para a prosódia e para a aplicação de processos fonológicos. 2002. 317f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade de Campinas, Campinas, 2002.

_____. Segmentações não-convencionais e teorias fonológicas. **Revista Letras de Hoje**. Porto Alegre. V. 39, nº. 3, p. 233-244, setembro, 2004.

_____. Domínios prosódicos no Português Brasileiro: evidências rítmica, entoacional e segmental. **Estudos Linguísticos XXXV**, p.118-131, 2006.

_____. Notas sobre a relação entre constituintes prosódicos e ortografia. **Revista de Estudos da Linguagem**, v.16, p.231 - 245, 2008.

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

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

TENANI, L. E.; PARANHOS, F. C. Análise Prosódica de segmentações não-convencionais de palavras em textos do sexto ano do Ensino Fundamental. **Filologia e Linguística Portuguesa**, v. 13, p. 477-504, 2011.

VIGÁRIO, M. **The prosodic word in European Portuguese**. Berlin: Mouton de Gruyter 2003.

ZORZI, J.L. **Aprender a escrever**: a apropriação do sistema ortográfico. Porto Alegre: ArtMed, 1998.

ANEXO

PROPOSTA TEXTUAL "A verdadeira história dos três porquinhos"

Caro professor

Gostaria de agradecer por ter aceitado participar de minha pesquisa e dizer que, sem sua ajuda, nada seria possível.

Em virtude da necessidade de manter uma regularidade na aplicação da proposta de produção textual, faço, a seguir, alguns apontamentos:

- Primeiramente, perguntar aos alunos se eles conhecem a história dos três porquinhos;
- Em seguida, em caso de resposta afirmativa, pedir a uma criança para contar oralmente a história e, se necessário, complementar com fatos que, do seu ponto de vista, não foram mencionados. Em caso de resposta negativa ou se nenhuma criança quiser contar oralmente a história, você poderá ler a história dos três porquinhos anexada a essa carta (Texto 01);
- Terminada a contagem oral da história ou a leitura, discutir com as crianças que uma história pode ter versões diferentes e que, nem sempre, a versão que a gente conhece de um fato é a verdadeira; comentar que, às vezes, um irmão conta uma versão diferente do outro para a mãe de algo que aconteceu, ou algum outro exemplo que você se lembre de uma situação em que há duas versões de um mesmo acontecimento;
- Depois disso, ler a história "A verdadeira história dos três porquinhos", anexada a essa carta (Texto 02), para os alunos;
- Finalizada a leitura, pedir para que os alunos recontem, por escrito, a história da maneira com que eles acham que realmente aconteceu.

Observações:

- Os textos precisam ser identificados para a análise, porém, as identidades serão mantidas em sigilo;
- É imprescindível para a pesquisa que não sejam dados esclarecimentos sobre a ortografia das palavras. Se algum aluno perguntar sobre a maneira correta de registrar uma palavra, por exemplo, por gentileza, peça para o aluno escrever como ele acha ser o certo ou, ainda para ele usar a sua intuição.

Atenciosamente,

AKISNELEN DE OLIVEIRA TORQUETTE
Doutoranda em Estudos Linguísticos
IBILCE/UNESP

TERMO DE REPRODUÇÃO XEROGRÁFICA

Autorizo a reprodução xerográfica do presente Trabalho de Conclusão, na íntegra ou em partes, para fins de pesquisa.

São José do Rio Preto, ____/____/____

Assinatura do autor