

UNIVERSIDADE ESTADUAL PAULISTA  
“JÚLIO DE MESQUISA FILHO”  
Instituto de Biociências, Letras e Ciências Exatas

Lilian Maria da Silva

**Um estudo longitudinal sobre as hipersegmentações de palavras  
escritas nos anos finais do Ensino Fundamental**

São José do Rio Preto (SP)  
2014

Lilian Maria da Silva

**Um estudo longitudinal sobre as hipersegmentações de palavras escritas nos anos finais do Ensino Fundamental**

Dissertação 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 Mestre em Estudos Linguísticos (Área de Concentração: Análise Linguística).

Orientadora: Profa. Dra. Luciani Ester Tenani

São José do Rio Preto (SP)  
2014

Silva, Lilian Maria da.

Um estudo longitudinal sobre as hipersegmentações de palavras escritas nos anos finais do ensino fundamental / Lilian Maria da Silva. -- São José do Rio Preto, 2014

171 f. : il., gráfs., tabs.

Orientador: Luciani Ester Tenani

Dissertação (mestrado) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas

1. Linguística. 2. Palavra (Linguística) 3. Comunicação escrita. 4. Segmentação ortográfica. 5. Língua portuguesa – Versificação. 6. Língua portuguesa - Ortografia e silabação. 7. Letramento. I. Tenani, Luciani Ester. II. Universidade Estadual Paulista "Júlio de Mesquita Filho". Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU – 408.52

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

## **COMISSÃO JULGADORA**

### **Titulares**

Profa. Dra. Luciani Ester Tenani – Orientadora (UNESP)

Prof. Dr. Lourenço Chacon (UNESP)

Profa. Dra. Cristiane Carneiro Capristano (UEM)

### **Suplentes**

Profa. Dra. Ana Ruth Moresco Miranda (UFPel)

Profa. Dra. Fabiana Cristina Komesu (UNESP)

À minha *mãe* e ao *Bruno*, por tudo que representam na minha vida...

dedico esta dissertação!

Agradeço imensamente...

a *Deus*, por sempre me iluminar nos momentos de angústia

a minha *mãe*, por ser meu maior exemplo de vida e perseverança

ao *Bruno*, por sempre me compreender, ajudar e amar, apesar de qualquer adversidade

aos meus familiares, especialmente minhas primas-irmãs *Viviane e Raquel* e meus tios *Ana e Osvaldo*, pela torcida, orações e apoio constante

à professora *Luciani Tenani*, pela dedicada orientação desde a graduação e por ser um exemplo de profissional e mulher

à professora *Ana Ruth Moresco Miranda*, pelas sugestões valiosas no Selin e na Qualificação da dissertação

ao professor *Lourenço Chacon*, pelas importantes contribuições não só a este trabalho, mas por fazer parte da minha formação acadêmica

à professora *Cristiane Carneiro Capristano*, por aceitar contribuir com o trabalho, neste momento de defesa

a *Akisenelen, Bianca, Camila e Graziela*, amigas queridas que, dentro e fora do mundo acadêmico, estiveram sempre ao meu lado

a senhora *Maria Aparecida*, que me ensina constantemente a crescer como pessoa e, principalmente, pelo incentivo e compreensão

à FAPESP, pela bolsa concedida.

# SUMÁRIO

---

**Resumo** 10

**Abstract** 11

**Apresentação** 12

## **SEÇÃO 1: Sobre a noção de palavra** 19

---

*Apresentação da Seção* 19

---

1.1. Registros escritos da noção de palavra: grafias de segmentação não-convencional 20

1.1.1. A relevância dos dados de segmentação não-convencional de palavras 21

---

1.1.2. Tendências de análise 25

1.1.2.1. Tendência fonético-semântica 26

---

1.1.2.2. Tendência fonológico-letrada 31

---

1.1.2.3. Outras formas de análise 45

1.2. Abordagem linguística da noção de palavra 53

1.2.1. Analisando um primeiro “local problemático”: processos de formação de palavras 54

---

1.2.2. Analisando um segundo “local problemático”: clíticos fonológicos 58

---

*Resumo da Seção* 66

## **SEÇÃO 2: Sobre a concepção de escrita** 68

---

*Apresentação da Seção* 68

---

2.1. Formas de tratamento da relação fala/escrita 70

---

2.2. A proposta do modo heterogêneo de constituição da escrita 79

---

*Resumo da Seção* 86

## **SEÇÃO 3: Aspectos metodológicos** 88

---

*Apresentação da Seção* 88

---

3.1. Delimitação do material 88

---

3.2. Seleção do corpus e identificação dos dados 94

---

3.3. Procedimentos de análise dos dados 101

---

*Resumo da Seção* 107

## **SEÇÃO 4: Análise das hipersegmentações de palavras** 108

---

*Apresentação da Seção* 108

---

4.1. Aspectos “gerais” das hipersegmentações 109

---

4.2. Aspectos “particulares” das hipersegmentações 141

---

*Resumo da Seção* 157

## **SEÇÃO 5: Considerações finais** 160

**Referências bibliográficas** 165

**Anexos** 170

---

Anexo 1 170

*Anexo 2 170*

*Anexo 3 171*

*Anexo 4 171*

---

## Lista de gráfico

Gráfico 1: Relação entre hipersegmentação e extensão dos textos (pelo número de palavras), por escrevente 119

## Lista de tabelas

Tabela 1: Distribuição dos textos e das hipersegmentações nos anos escolares do EF-II 109

Tabela 2: Relação número de palavras escritas convencionais e número de hipersegmentações, por ano escolar 111

Tabela 3: Distribuição do número de textos e hipersegmentações produzidos, por escrevente 114

Tabela 4: Número de hipersegmentação em relação ao número de palavras por escrevente 115

Tabela 5: Número de palavras escritas e hipersegmentações por escrevente/ano escolar 118

Tabela 6: Classificação das hipersegmentações quanto ao recurso gráfico nos anos escolares 126

Tabela 7: Estruturas rítmicas de palavras mobilizadas nas hipersegmentações por ano escolar 130

Tabela 8: Hipersegmentações de espaço em branco e correspondências gráficas 134

Tabela 9: Correspondência gráfica e hipersegmentações: tipos de classes de palavras 135

Tabela 10: Estruturas rítmicas de palavras mobilizadas no 6º ano 170

Tabela 11: Estruturas rítmicas de palavras mobilizadas no 7º ano 170

Tabela 12: Estruturas rítmicas de palavras mobilizadas no 8º ano 171

Tabela 13: Estruturas rítmicas de palavras mobilizadas no 9º ano 171

## Lista de quadros

Quadro 1: Exemplo de proposta de produção textual do projeto de extensão 91

Quadro 2: Grafias de hipersegmentação produzidas por André 121

Quadro 3: Grafias de hipersegmentação produzidas por Viviane 121

Quadro 4: Grafias de hipersegmentação produzidas por Ana 121

Quadro 5: Grafias de hipersegmentação produzidas por Henrique 121

Quadro 6: Grafias de hipersegmentação produzidas por Camila 121

Quadro 7: Grafias de hipersegmentação produzidas por Fernando 121

Quadro 8: Grafias de hipersegmentação produzidas por Julia 122

Quadro 9: Grafias de hipersegmentação produzidas por Mateus 122

Quadro 10: Grafias de hipersegmentação produzidas por Maria 122

Quadro 11: Grafias de hipersegmentação produzidas por Lucas 122

Quadro 12: Grafias de hipersegmentação produzidas por Pedro 123

Quadro 13: Grafias de hipersegmentação produzidas por Bruno 123

Quadro 14: Grafias “particulares” (hipersegmentação e híbrido) produzidas de Bruno e Pedro 142

Quadro 15: Exemplos de hipersegmentações frequentes no EF-I e EF-II 144

## Lista de figuras

Figura 1: Exemplos de hipersegmentação ocasionadas pela presença do espaço em branco e hífen 13

Figura 2: Exemplos de hiposegmentação ocasionadas pela ausência do espaço em branco e hífen 13

Figura 3: Exemplo de dado híbrido 13

Figura 4: Dado “contra bando” (*contrabando*) 58

Figura 5: Dado “loca mente” (*loucamente*) 58

Figura 6: Dados “em tão” (*então*) e “na quele” (*naquele*) 66

Figura 7: Dados “a vó” (*avó*) e “falam do” (*falando*) 66

Figura 8: O contínuo de gêneros textuais, segundo Marcuschi (2001) 77

Figura 9: Primeiro problema de categorização do espaço em branco: caligrafia 96

Figura 10: Segundo problema de categorização do espaço em branco: translineação 97

Figura 11: Terceiro problema de categorização do espaço em branco: palavra estrangeira 98

Figura 12: Quarto problema de categorização do espaço em branco: nome próprio 98

Figura 13: Exemplo de hipersegmentação de palavra homônima 99

Figura 14: Problema de categorização do hífen: configuração de palavra composta 100

Figura 15: Exemplo de grafias do clítico “a” hipersegmentados 139

Figura 16: Exemplo de grafias do clítico “em” hipersegmentados 139

Figura 17: Dados “com ver sando”, “com versando” (*conversando*) e “com ver samos” (*conversamos*) 149

Figura 18: Dado “adisi o na” (*adicionar*) 151

Figura 19: Dados “come cheia” (*comecei a*) e “a olado” (*ao lado*) 153

## Resumo

---

Neste trabalho, analisamos as chamadas hipersegmentações, que dizem respeito a quando, no interior de uma palavra, identifica-se espaço gráfico não previsto pela ortografia (por exemplo: “da quela” e “mora-va”, que, de acordo com as convenções ortográficas, são grafadas, respectivamente, como *daquela* e *morava*). Por meio de um corpus longitudinal de textos escritos dos quatro últimos anos do Ensino Fundamental, desenvolvemos uma descrição geral dos dados e da trajetória de cada um dos escreventes investigados, com o objetivo de: (i) identificar possíveis regularidades linguísticas dessas grafias não-convencionais de palavras que caracterizariam o fim da etapa fundamental de formação escolar; e (ii) interpretar as ocorrências que não atenderam às regularidades mais gerais, mas que se mostraram relevantes para a compreensão de uma relação mais particular do sujeito com a (sua) escrita. Em função do objeto de investigação, fundamentamo-nos em discussões teóricas acerca das fronteiras não-convencionais de palavras (ABAURRE, 1991c; TENANI, 2010, CUNHA, 2004), da noção de palavra (SCHWINDT, 2001; BISOL, 2005; SIMIONI, 2008) e da constituição da escrita (CORRÊA, 2004). Os resultados do trabalho, em relação ao primeiro objetivo, indicaram que: (i) as ocorrências de hipersegmentação diminuíram ao longo dos anos escolares; (ii) a maioria dos escreventes concluiu o Ensino Fundamental produzindo hipersegmentações, embora o número delas tenha diminuído em relação ao tipo de palavra hipersegmentada e à extensão (dada em função do número de palavras) dos textos escritos; (iii) as palavras que continuaram a ser hipersegmentadas foram aquelas que, predominantemente, permitiram a relação entre sílabas pretônicas e possíveis classes gramaticais (como, por exemplo, artigos, preposições, pronomes) e entre sequências que poderiam ser interpretadas como sendo de verbo+pronome enclítico. Em relação a esse último resultado, constatou-se que informações linguísticas de natureza prosódica e da própria convenção ortográfica foram importantes para a ocorrência de um limite não-convencional em determinada posição da palavra. A respeito do segundo objetivo da pesquisa, chegamos à observação de que os escreventes que se diferenciam das tendências gerais: (ii) seguem mobilizando hipóteses características dos primeiros anos de escolarização; (ii) utilizam, em algumas circunstâncias, recursos gráficos que demarcam palavras como formas de destacar sentidos privilegiados do texto; e (iii) também se ancoram, de maneiras diferentes, em informações linguísticas que permeiam as regularidades observadas. Em suma, a relação entre possíveis tendências das grafias de hipersegmentação e de dados que, em alguma medida, não seguiram tais características, forneceu pistas de hipóteses gerais e particulares a respeito da noção de palavra escrita, hipóteses interpretadas, nesta pesquisa, como marcas de trânsito do escrevente por práticas de uso da linguagem, oral/falada e letrada/escrita, em que se inserem (CORRÊA, 2004).

**PALAVRAS-CHAVE:** hipersegmentação; palavra; prosódia; ortografia; oralidade; letramento.

## Abstract

---

In this paper, we analyze the calls hypersegmentation of words that relate to when, within a word, it identifies space of written not predicted by orthography (e.g. "da quela" and "morava", when at the limits convention are *daquela* and *morava*). Through a longitudinal study of written texts of the last four years of Elementary School, we develop a general description of the data and the trajectory of each of the written investigated, with the aim of: (i) to describe the distribution of the analyzed occurrences through the school years of Elementary School and written and to identify possible linguistic regularities of these non-conventional spellings of words that characterize the end of key stage of schooling, and (ii) to interpret the events that did not meet the more general regularities, but which were relevant in the understanding of a more particular subject with (his/her) writing. Depending on the object of investigation, we base ourselves in theoretical discussions about the boundaries unconventional words (ABAURRE, 1991c; TENANI, 2010, CUNHA, 2004), the notion of word (SCHWINDT, 2001; BISOL, 2005; SIMIONI, 2008) and the writing constitution (CORRÊA, 2004). The results of the paper related to the first objective, indicated that: (i) occurrences of hypersegmentation decreased through the school years; (ii) most written concluded Elementary School producing hypersegmentation, although hypersegmentation number has decreased compared to the type of word that was hipersegmented and extension (given by the number of words) of written texts; (iii) the words that remained to be hypersegmented were those ones that allowed the relation between unstressed syllables and possible grammatical classes and between sequences that could be interpreted as being verb + enclitic pronoun (according to the Portuguese grammar), and (iv) prosodic nature information and retrievable linguistic information about inserting the written in literacy/writings practices have been mobilizing an unconventional limit written in a certain position in the word. Regarding the second objective of this paper, we came to the observation that the written who differ from general trends: (i) they follow mobilizing hypotheses that are characteristic of the early years of schooling; (ii) they use, in some circumstances, graphical features that mark words as ways to highlight the privileged directions of the text, and (iii) they are based, in different ways, in linguistic information that permeated the observed regularities. In short, the relation between possible hyper trends and data, to some extent, did not follow such features, provided us with clues of general and specific hypotheses about the notion of the written word, interpreted hypotheses, in this research, as marks the written is living in by practices of language use, oral/literate and spoken/written, they take part in (CORRÊA, 2004).

**KEYWORDS:** Hypersegmentation; Word; Prosody; Orthography; Orality; Literacy.

# Apresentação

---

Nesta pesquisa, analisamos grafias não-convencionais de fronteiras de palavra caracterizadas pela presença de um recurso gráfico (espaço em branco ou hífen) no interior de palavras ortográficas. Na literatura, esse fenômeno é chamado de *hipersegmentação de palavras*. No bojo de uma reflexão mais ampla, as hipersegmentações compõem uma categoria maior de dados, a qual abriga outras formas não-convencionais de distribuição dos limites de palavra na escrita. Denominada segmentação não-convencional de palavras,<sup>1</sup> tal categoria engloba, ainda, as hipossegmentações – ausência da fronteira gráfica entre palavras – e os dados híbridos (CUNHA, 2010) – ocorrência da ausência seguida da presença do limite gráfico ou vice-versa em uma mesma sequência.<sup>2</sup> Nas Figuras 1, 2 e 3, apresentadas adiante, trazemos exemplos de cada um dos tipos de segmentação não-convencional mencionados.

---

<sup>1</sup> Neste trabalho, todas as referências feitas aos termos *segmentação* e/ou *segmentação não-convencional* correspondem, apenas, à organização do enunciado escrito em palavras, já que, na escrita convencional, a segmentação pode indicar, também, a divisão do fluxo textual em outras unidades (cf. CAPRISTANO, 2003, p.14).

<sup>2</sup> Cunha (2004, 2010) identificou dados híbridos em que primeiro ocorreu a hipossegmentação e depois a hipersegmentação; porém, em nosso cópula, as ocorrências híbridas apresentaram primeiro a hipersegmentação, depois a hipossegmentação. Desse modo, defendemos a possibilidade de haver dados híbridos que envolvam tanto a hipo seguida da hipersegmentação, quanto a hiper seguida da hipossegmentação. É importante lembrar que, a partir de um ponto de vista distinto, Chacon (2004) denominou por *mesclas* grafias que envolvem os dois tipos de segmentação não-convencional e, também, grafias que, dada a sua estrutura, não podem ser explicadas apenas em função da classificação em hipo ou hipersegmentação, uma vez que a ausência/presença do limite gráfico está presente “em partes nas quais deveria (ou se suporia) haver um *limite* ao mesmo tempo ortográfico e prosódico” (CHACON, 2004, p. 226, grifo no original). Segundo o autor, nesses dados haveria uma ação simultânea de constituintes prosódicos, no sentido de que em uma ocorrência como “fofim zida” (*foi visitar* – CHACON, 2004, p. 229) toda a sequência indicaria a ação de um constituinte frasal, como a frase fonológica, e as partes separadas entre limites, constituintes menores que a palavra fonológica, como o pé métrico (com a configuração de um iambo), no caso do exemplo mostrado. Não é um de nossos propósitos discutir a classificação de um terceiro tipo de segmentação não-convencional seja como *híbridos* ou *mesclas*, de modo que assumimos, de partida, a classificação proposta por Cunha (2010), visto que adotamos, nesta pesquisa, o mesmo critério da autora para definição de hipersegmentação de palavra.

**Figura 1:** Exemplos de hipersegmentação ocasionadas pela presença do espaço em branco e hífen<sup>3</sup>



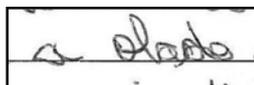
Fonte: Banco de Dados de Escrita do Ensino Fundamental II  
(textos: Z09\_6B\_26M\_07 e Z11\_8B\_15F\_04)

**Figura 2:** Exemplos de hipossegmentação ocasionadas pela ausência do espaço em branco e hífen



Fonte: Banco de Dados de Escrita do Ensino Fundamental II  
(textos: Z08\_5C\_39F\_03 e Z08\_8A\_07F\_01)

**Figura 3:** Exemplo de dado híbrido



Fonte: Banco de Dados de Escrita do Ensino Fundamental II  
(texto: Z10\_7B\_05M\_02)

É pertinente destacar que, enquanto objeto de investigação, as segmentações não-convencionais têm despertado o interesse de diferentes áreas do saber, como, por exemplo, a Fonoaudiologia (cf. ZORZI, 1997), a Educação (cf. COX E ASSIS-PETERSON, 2001), a Linguística (cf. CHACON, 2006; TENANI, 2011b) e, ainda, estudos na interface Linguística-

<sup>3</sup> A presença do hífen em “*arranja-se*” pode ser entendida como não-convencional, pois, conforme o enunciado do qual foi extraído o dado, a grafia deveria ser “*arranjasse*”, a saber: “e resolveu entrar para pedir a ajuda do nosso bom senhor, para que ele **arranja-se** um emprego”.

Educação (cf. CUNHA, 2010; CUNHA E MIRANDA, 2013).<sup>4</sup> Tendo, por sua vez, questões específicas de pesquisa, cada um dos campos do conhecimento abordam o fenômeno aqui investigado segundo propósitos de análise distintos. Concernente à Linguística – em particular, no viés norteador de reflexão adotado neste estudo –, as marcas de segmentação não-convencional têm auxiliado a argumentação de discussões relacionadas (i) à organização da fala em constituintes prosódicos; (ii) à relação constitutiva entre fala e escrita e oralidade e letramento; e (iii) às formulações que são construídas pelos escreventes sobre os limites (orto)gráficos de palavra. Antecipamos que uma reflexão detalhada desses três pontos está desenvolvida na primeira Seção deste trabalho.

Quanto ao tipo de material em que se identificam segmentações não-convencionais, são bastante frequentes pesquisas realizadas com base em produções de textos infantis (cf. ABAURRE, 1988a, 1991c; SILVA, A., 1991; CUNHA, 2004; CAPRISTANO, 2003; CHACON, 2005, 2006).<sup>5</sup> No entanto, pouco se conhece, em termos de pesquisa, sobre as formas de delimitação não-convencional das palavras em anos escolares mais avançados, como é o caso, por exemplo, dos anos finais do Ensino Fundamental (doravante, EF-II). Nossa pesquisa, nesse aspecto, pretende contribuir a área de pesquisas sobre o EF-II, à semelhança do que, recentemente, já vem sendo realizado por Tenani (2009a, 2010, 2011b) e Paranhos e Tenani (2011). A importância em se investigar dados do EF-II pode ser justificada não apenas em função dos poucos estudos realizados, mas, principalmente, por já ser esperado do escrevente, nesse período de escolarização, a produção de textos “*com domínio da separação em palavras*, estabilidade de palavras de ortografia regular e de irregulares mais frequentes na escrita e utilização de recursos do sistema de pontuação para dividir o texto em frases” (PCN, 1997, p. 80, destaque nosso). Por meio da análise de nossos dados, mostramos que escreventes que já concluíram, em termos formais, o período dos anos iniciais da

---

<sup>4</sup> Não é nosso intuito realizar um levantamento exaustivo de trabalhos desenvolvidos no interior dessas áreas.

<sup>5</sup> Ainda em contexto inicial de aquisição da escrita, existem trabalhos interessados na escrita de jovens e adultos em período de alfabetização (ver TENANI, 2008 e FERREIRA, 2011, por exemplo).

alfabetização em ambiente escolar continuam a ter dificuldades em reconhecer os limites de palavra escrita. Assim, o objetivo central deste estudo é o de investigar hipóteses linguísticas que permanecem e/ou passam a ser mobilizadas nos anos mais avançados do EF, a saber: os atuais 6º a 9º anos (antigas 5ª a 8ª séries).<sup>6</sup> Com o objetivo de buscar fatores linguísticos subjacentes às grafias de hipersegmentação, defendemos a hipótese de que a noção de palavra escrita é *complexa* e o domínio da convenção ortográfica exige do escrevente conhecimentos linguísticos de distintas naturezas (como: fonológico, morfossintático/semântico e, também, letrado).

Ao assumirmos a hipótese descrita, partilhamos com Capristano (2007b, p. 80) a visão de que a aquisição da escrita é um processo que “não poderia ser interpretado como um percurso de superação de etapas e/ou um percurso de desenvolvimento cognitivo, tampouco um percurso de exploração ou, ainda, tentativas de adequação às convenções escritas”, mas sim um processo que se modifica em função do “*Outro* como instância representativa da linguagem (e da escrita em particular), a *escrita* na complexidade de seu funcionamento (heterogeneamente constituída) e a criança<sup>7</sup> enquanto *sujeito* escrevente” (CAPRISTANO, 2007b, p. 160, destaques do original).

No que diz respeito aos objetivos específicos da pesquisa, com base em uma amostra longitudinal de produções escritas, almejamos desenvolver:

- (i) uma descrição das características gerais das hipersegmentações, primeiro para observarmos a distribuição dos dados pelos anos escolares do EF-II e escreventes investigados, e, segundo, para que possamos apontar possíveis regularidades linguísticas das grafias não-convencionais encontradas; e

---

<sup>6</sup> A Lei nº11.274, instituiu, a partir de 2009, o Ensino Fundamental para nove anos, implantando como obrigatória a matrícula das crianças de seis anos nesse nível de ensino. Dessa maneira, o EF passou a ser composto por um ano a mais em seu início e, em razão desse aumento, as antigas quinta/sexta/sétima/oitava séries passaram a ser nomeadas por sexto/sétimo/oitavo/nono anos, respectivamente.

<sup>7</sup> A palavra *criança*, nesse trecho, foi interpretada como uma expressão mais abrangente, que busca se referir aos *aprendizes de escrita*.

- (ii) uma análise de dados de hipersegmentação que, em relação à descrição geral dos dados, não tenham se assemelhado às tendências observadas, constituindo-se, em grande parte das vezes, em grafias únicas do córpus.

Em se tratando do modo de organização do trabalho, nas Seções 1 e 2 dedicamos atenção aos subsídios teóricos. Em virtude dos fundamentos assumidos, compreendemos as hipersegmentações enquanto um lugar importante de abordagem linguística e não como registros de “manifestações ‘imperfeitas’ de uma gramática ‘adulta’” (ABAURRE; FIAD e MAYRINK-SABINSON, 1997, p. 18).

Na Seção 1 enfocamos a *noção de palavra* entendendo-a sob duas vertentes. A primeira trata dos registros da palavra escrita com os dados de segmentação não-convencional. Dessa forma, damos início às discussões teóricas a partir do nosso objeto, em especial por julgarmos importante a descrição das formas de análise dedicadas às segmentações não-convencionais e por se fazer necessário explicitar a filiação teórica deste trabalho, dada a coerência com nossos propósitos investigativos e, também, com a concepção de escrita adotada. A segunda vertente reflete sobre a *noção de palavra fonológica*, visto termos observado, tanto em resultados de outros estudos sobre segmentações não-convencionais quanto em nossos próprios resultados, que as fronteiras de palavras escritas distantes da norma ortográfica guardam forte relação com fronteiras fonológicas. Nesse sentido, acreditamos poder mostrar que a noção de palavra escrita, registrada na forma de hipersegmentação, está fundamentada em outras noções que não apenas aquelas privilegiadas pela ortografia do português. Para além desse propósito, endossamos que, de uma abordagem fonológica, podemos deixar em destaque que a definição de palavra é, acima de tudo, um problema linguístico e não somente um problema gráfico.

Também, por considerarmos que as hipersegmentações revelam a complexidade da segmentação de enunciados escritos (e, desse modo, concebermos esses tipos de dados como

um lugar notável de observação do funcionamento da escrita), versamos, na Seção 2, sobre a *concepção de escrita*, na qual nos baseamos. Iniciamos a Seção situando a abordagem de escrita adotada em relação às concepções que tratam da relação entre fala e escrita e das quais ela se distancia, uma vez que discutir essa relação é fundamental para a proposta de a escrita ser constituída de modo heterogêneo. O nosso intuito, ao sinalizamos possíveis distinções, é o de reforçar as implicações, em nível de análise dos dados, de diferentes perspectivas teórico-metodológicas na compreensão de fatos não-convencionais de escrita. Ao fim, pois, apresentamos a concepção do modo heterogêneo de constituição da escrita, defendida por Corrêa (1997, 2004) e adotada nesta dissertação.

Entendemos que, ao assumirmos a perspectiva de escrita proposta por Corrêa (1997, 2004), fazemos a opção teórica de nos distanciarmos de concepções de análises, por exemplo, do tipo *normativa* (para a qual as segmentações não-convencionais são erros originados de uma suposta interferência da fala na escrita) e *patológica* (para a qual a ocorrência de segmentações não-convencionais pode ser sintoma de transtornos de aprendizagem), e afirmamos o nosso encontro com um enfoque de análise que, além de não se perfilar com os posicionamentos exemplificados, almeja compreender o modo como os escreventes deixam, no texto, marcas da sua relação com a linguagem em seu modo de enunciação escrita.<sup>8</sup>

Na Seção 3, descrevemos os *aspectos metodológicos* da pesquisa. Nesse ponto, abordamos o material de investigação, dando ênfase às particularidades do banco de dados de onde foram retiradas as produções escritas analisadas e aos critérios estabelecidos para a seleção do banco. Após a descrição do material, discorreremos sobre a composição do cópuz e sobre as decisões metodológicas que se mostraram, neste trabalho, como importantes para a identificação das grafias de hipersegmentação. Ao fim da Seção, tratamos das formas de análise dos resultados e, nessa parte, apresentamos os pressupostos do modelo de Fonologia

---

<sup>8</sup> Uma discussão aprofundada acerca de perspectivas de análise sobre os dados de segmentação não-convencional é encontrada em Tenani (2013).

Prosódica de Nespor e Vogel (1986), a partir do qual estruturamos a organização e interpretação dos dados.

Na Seção 4, realizamos a *análise dos dados*. Nesta parte do trabalho, buscamos descrever a distribuição dos dados e identificar, na estrutura das hipersegmentações, pistas de hipóteses linguísticas subjacentes às grafias não-convencionais. Também, apresentamos uma interpretação das hipersegmentações únicas, as quais mostram hipóteses que apontam para um processo mais particular de relacionamento do escrevente com a (sua) escrita.

Finalmente, na Seção 5, realizamos um *apanhado geral das discussões* desenvolvidas, dando destaque àquelas que consideramos serem as mais relevantes.

# SEÇÃO 1:

## Sobre a noção de palavra

---

### *Apresentação da Seção*

Ao tratarmos de grafias de hipersegmentação, estamos perpassando uma questão: quais limites definem uma palavra? No campo dos estudos linguísticos, essa não é uma questão nova, pois são bastante antigas as reflexões sobre a palavra, como objeto de pesquisa. Ainda que diferentes níveis de análise tenham se dedicado a estudá-la (por exemplo: morfologia, fonologia, sintaxe, semântica, lexicologia, lexicografia), não há definição consensual a seu respeito e uma das razões está relacionada às diferentes questões de interesse específicas de cada área. Na divergência dos aspectos considerados na formalização do conceito de palavra, uma das consequências é a de que os seus limites não são, em grande parte das vezes, coincidentes entre si, conforme apresentamos nesta Seção, quando consideramos a noção de palavra nos níveis morfológico e fonológico.

Como adiantamos, nesta Seção discutimos a noção de palavra. Iniciamos, pois, com estudos sobre dados de segmentação não-convencional (subseção: “*Registros escritos da noção de palavra: grafias de segmentação não-convencional*”), a fim de tratarmos de registros da noção de palavras em enunciados escritos. Os trabalhos resenhados, ainda que não tivessem a preocupação direta de discutir o conceito de palavra, apontaram, em suas análises, para fatos linguísticos subjacentes à conceituação dessa os quais teriam, de modo mais proeminente, mobilizado as grafias discordantes da convenção ortográfica. Para além de uma

reflexão mais ampla acerca da noção de palavra escrita, a consideração dos trabalhos selecionados permitiu a identificação de diferentes formas de análise quanto ao objeto que investigamos. Nesse sentido, revisitar pesquisas sobre os tipos de segmentação não-convencional foi fundamental, já que pudemos: (i) nos posicionar perante uma posição teórico-metodológica de análise mais coerente com nossos propósitos de pesquisa e com a concepção de escrita assumida (a ser apresentada na Seção 2); e (ii) demonstrar a relevância das segmentações não-convencionais aos estudos da linguagem.

Damos continuidade à Seção abordando a noção de palavra fonológica (subseção: *“Abordagem linguística da noção de palavra”*). A partir da discussão sobre os limites de palavra fonológica, demonstramos que a definição de palavra é, antes de tudo, um desafio linguístico, já que é difícil identificar, também nos enunciados falados, as fronteiras de palavras, por serem instáveis e/ou de difícil identificação. Acreditamos que, ao serem registradas palavras na escrita, marca-se essa instabilidade de fronteiras por meio de grafias como as hipersegmentações, por exemplo. Dito de outro modo, a reflexão proposta entre a relação de estudos fonológicos sobre o conceito de palavra e os tipos de dados analisados nesta pesquisa parte da premissa de que as hipersegmentações destacam a complexidade do estatuto de palavra, uma vez que sinalizam, em certa medida, a relevância de noções como palavra fonológica, palavra morfológica e clítico fonológico, por exemplo.

Finalizando essa apresentação inicial, a seguir passamos a expor as discussões sobre as grafias de fronteiras de palavra.

### **1.1. Registros escritos da noção de palavra: grafias de segmentação não-convencional**

Nesta pesquisa, o eixo central das discussões está respaldado na compreensão das hipersegmentações como frutos de “hipóteses conflitantes que os escreventes elaboram a partir de sua multifacetada relação com a linguagem falada e escrita” (TENANI, 2009b, p. 109). Privilegiamos, portanto, a investigação de aspectos linguísticos envolvidos nessas grafias, em detrimento de qualquer visão depreciativa que viria considerá-las como simples desvios da escrita padrão. Desse modo, na próxima parte do trabalho, apresentamos considerações que permitem observar a importância de dados de segmentação não-convencional para os estudos da linguagem.

### ***1.1.1. A relevância dos dados de segmentação não-convencional de palavras***

No Brasil, a partir de meados da década de 1980, muitos linguistas passaram a investigar um tipo de material de pesquisa restrito, até então, aos estudos pedagógicos: textos produzidos em início de aquisição escrita. Analisar esse tipo de material se mostrou importante, especialmente porque o modo como aprendizes analisariam a escrita e a representariam, em suas primeiras produções, expunha fatos que muito de perto interessavam à Linguística. Nessa reflexão, Abaurre e Cagliari (1985) defenderam a relevância da produção escrita realizada espontaneamente, pois, para os autores, nesse tipo de texto as crianças marcariam suas reais hipóteses sobre o novo objeto a que estavam refletindo, isto é, a escrita.<sup>9</sup>

---

<sup>9</sup> Merece explicação a ideia de texto espontâneo posta em Abaurre e Cagliari (1985). Essa ideia está fortemente relacionada a uma contraposição ao estilo de texto que era ensinado pelos métodos e materiais didáticos de alfabetização vigentes na época (as cartilhas): “As crianças da primeira série normalmente não produzem textos espontâneos. A relação que elas estabelecem com a escrita é, via de regra, extremamente artificial, porque se pressupõe que, uma vez que ainda não dominam a convenção ortográfica, elas não são ainda capazes de desenvolver, com a escrita, atividades significativas. Seus exercícios costumam ser, assim, absolutamente controlados pela professora, que reproduz a orientação pedagógica vigente. [...] Pode-se dizer, portanto, que a partir do seu primeiro contato com a escrita as crianças são submetidas de forma sistemática a uma série de exercícios que a distanciam progressivamente da noção de *texto*, para elas tão natural em termos de produção oral, quando ingressam na escola” (ABAURRE E CAGLIARI, 1985, p. 25, destaque no original).

Ao tomar o texto infantil como objeto de análise, uma das questões em investigação passou a ser a da segmentação em palavras. O interesse por esse aspecto gráfico voltou-se, predominantemente, para a descoberta de quais critérios linguísticos guiariam os aprendizes no reconhecimento dos limites de palavras escritas.

Interessada nos dados de segmentação não-convencional, Abaurre (1988a) sinaliza a importância da relação entre *fala e escrita*, para melhor compreensão das hipóteses infantis acerca dos limites de palavra. A autora aponta, em menção à relação destacada, a crença comum de que, nas primeiras produções escritas, os aprendizes tendem a escrever do modo como falam (visão que não parece ter mudado até os dias atuais!) e afirma ser essa uma consideração infundada, pois, como a própria autora observa, “ninguém se programa para simplesmente ‘escrever como fala’, seja em termos de uso das letras e das suas relações com o som, seja em termos de *segmentação*, seja em termos das próprias estruturas linguísticas” (ABAURRE, 1988a, p. 137, destaques nosso).

Com efeito, essas considerações de Abaurre (1988a) deixam entrever que a relação entre fala e escrita, a qual se dá à observação por meio dos registros escritos não-convencionais dos tipos mais variados, é bem mais complexa que a ideia de que “a criança escreve como fala” pressupõe. Na aceitação de tal complexidade, trabalhos mais recentes advogam sobre a relevância de dados não-convencionais de escrita, como os de segmentação (ver CHACON, 2005; TENANI, 2011b; CAPRISTANO, 2004), para uma proposta de entendimento da relação fala/escrita como aspectos de uma mesma constituição heterogênea (CORRÊA, 2004). O que, na verdade, esses trabalhos defendem é que a escrita se forma heterogeneamente e que, por isso, a presença de aspectos do falado, no produto escrito, não são marcas de interferência, mas pontos de emergência daquilo que, acima de tudo, a

constitui.<sup>10</sup> Assim, podemos apontar que as marcas de segmentação distantes da convenção permitem uma indagação mais sistemática acerca da relação dos sujeitos da linguagem com os modos enunciativos falado e escrito.

Questões que tocam à fonologia da língua são outro lugar de diálogo no estudo das segmentações não-convencionais. Segundo Abaurre (1991c), as grafias das fronteiras de palavras podem ser indícios importantes na “validação das unidades prosódicas propostas nos modelos fonológicos não-lineares como constitutivas de uma hierarquia que reconhece vários domínios prosódicos como significativos em termos das representações fonológicas subjacentes” (p. 204). A título de exemplo, nesse viés de pesquisa, tem sido desenvolvido grande parte dos estudos dos grupos de pesquisa *Estudos sobre a Linguagem* (GPEL/CNPq)<sup>11</sup> e *Grupo de Estudos sobre a Aquisição da Linguagem Escrita* (GEALE)<sup>12</sup>. Por ora, destaca-se que as diversas pesquisas realizadas por ambos os grupos têm mostrado pontos de intersecção entre dados não-convencionais de escrita e fatos fonológicos; no entanto, neste momento do trabalho não aprofundamos as conclusões de tais pesquisas, pois a essa tarefa nos dedicamos mais adiante, quando delineamos os diferentes tipos de hipóteses explicativas a respeito dos modos de análise dos dados de segmentação não-convencional.<sup>13</sup>

Outro ponto relevante para discussão linguística se refere aos limites de palavra escrita. Nessa reflexão, concordamos com Abaurre (1991c) quando essa propõe como problema trivial da escrita a necessidade de se conhecerem os critérios convencionais que definem palavra, para que assim essas possam ser separadas entre limites. A transformação simbólica operada pela escrita faz com que, muitas vezes, negligenciemos “qual é a relação

---

<sup>10</sup> A proposta de ver a relação fala/escrita como constitutivas do modo de enunciação escrito vai ser especificada na próxima Seção deste trabalho, em virtude de se tratar do eixo central de organização da noção de escrita que embasa esta pesquisa.

<sup>11</sup> Grupo de pesquisa coordenado pelos professores Lourenço Chacon (FFC/UNESP) e Luciani Tenani (IBILCE/UNESP).

<sup>12</sup> Grupo de pesquisa coordenado pela professora Ana Ruth Moresco Miranda (FaE/UFPeL).

<sup>13</sup> Alguns trabalhos que investigaram dados de segmentação não-convencional de palavras e sua relação com características fonológicas, dentro do GPEL: Capristano (2003, 2004), Chacon (2004, 2005, 2006), Tenani (2004, 2009a, 2010, 2011b); dentro do GEALE: Cunha (2004, 2010), Ferreira (2011), Cunha e Miranda (2013).

que aquele que não sabe ler tem com esses sinais [gráficos] que, para nós, apresentam-se como transparentes” (LEMOS, 1998, p. 17). É assim que muitas vezes estamos “de certa maneira, ‘surdos’ para as características mais contínuas dos enunciados orais” (ABAURRE, 1991c, p. 203) e acabamos por ouvir as palavras como unidades morfossintáticas já organizadas sintagmaticamente.

Ainda que seja inegável o fato de que, em uma sociedade predominantemente letrada como a nossa, os sujeitos (alfabetizados ou não) já possuam uma ideia, mesmo que apenas visual, de a escrita ser separada em unidades gráficas, não podemos afirmar que os aprendizes já “compartilhem dos critérios morfossintáticos e semânticos utilizados [...] na identificação das palavras” (ABAURRE, 1991c, p. 204), até mesmo porque não estão “habituaados à reflexão metalinguística que tal segmentação pressupõe” (ABAURRE E CAGLIARI, 1985, p. 27). Conforme discutimos na subseção seguinte, a palavra não é um objeto de investigação facilmente definível, já que suas funções e seus limites não são definidos *a priori*. Dessa maneira, as grafias não-convencionais das fronteiras de palavra são um indício dessa instabilidade pouco notada (e tantas vezes esquecidas) por parte dos usuários da escrita, especialmente quando esses tipos de dados permitem recuperar, justamente, as faltas de correspondências entre as diferentes formas de palavras, formas essas que, por sua vez, em seus respectivos níveis de análise, são também alvos de problema.

Com base no exposto, procuramos tematizar aspectos a partir dos quais acreditamos que a consideração dos dados de segmentação não-convencional podem ajudar a compreender. Por fim, sintetizamos os pontos relevantes: (i) questões referentes à relação fala/escrita; (ii) comprovação de pressupostos de teorias fonológicas e (iii) aspectos acerca da noção de palavra. Cabe destacar que esses pontos foram abordados separadamente por razões puramente expositivas, com o intuito de chamar a atenção para a importância linguística do

objeto analisado nesta pesquisa, visto que, em demais estudos sobre o assunto (sobretudo, os mais recentes), todos os aspectos estão considerados interligadamente.

Adiante discutimos formas de análises das grafias de segmentação não-convencional, com base em trabalhos que, à semelhança deste, direcionam-se pelos pontos relevantes explorados nesta subseção.

### ***1.1.2. Tendências de análise***

Entendendo as hipersegmentações como fenômeno de linguagem, buscamos, em pesquisas de orientação linguística, hipóteses explicativas que procuram interpretar tais ocorrências. É possível identificarmos, em todos os trabalhos, a relação entre a distribuição não-convencional dos limites de palavra e fatos linguísticos de natureza *prosódica*. A partir desse destaque geral, consideramos haver, nos estudos analisados, um desdobramento de tendências prototípicas de compreensão das motivações que guiam os escreventes na proposição de segmentações distantes das convenções ortográficas. Uma primeira tendência tem apoio em argumentos do aspecto *fonético* da prosódia e a relação desse aspecto com fatos de natureza *semântica* (chamaremos, daqui em diante, de orientação explicativa fonético-semântica), representativa, por exemplo, dos trabalhos de Abaurre e Cagliari (1985), Abaurre (1988, 1991c) e Silva, A. (1991). Uma segunda tendência está respaldada em fatos *fonológicos* da prosódia em conexão com informações *letradas* (que aqui denominaremos de tendência fonológico-letrada), e é possível de ser constatada, por exemplo, nas pesquisas de Capristano (2003, 2004), Chacon (2004, 2005, 2006), Tenani (2004, 2009a, 2010, 2011b) e Paula (2007). Abordamos, também, os trabalhos de Cunha (2004, 2010) e Ferreira (2011), os

quais relacionam fatos fonológicos da prosódia, buscando, em última instância, pensar a importância de seus resultados com questões pedagógicas.

Tratamos de cada uma das tendências nos próximos itens.

#### *1.1.2.1. Tendência fonético-semântica*

De modo pioneiro, Abaurre e Cagliari (1985) propõem que dados de segmentações em palavras divergentes da ortografia constituem-se em hipóteses das crianças acerca dos limites de palavra escrita. Para os autores, essas hipóteses não são categóricas e/ou estáveis, pois se, por um lado, em alguns momentos as grafias não-convencionais parecem ser guiadas por aspectos fonéticos da linguagem – especialmente quando, ao tomar como eixo norteador a pronúncia das palavras, produzem escritas baseadas em grupos tonais,<sup>14</sup> como as hipossegmentações de “umbripi” (*um príncipe*) e “abechou” (*a beijou*) –, por outro lado, em diferentes momentos são aspectos da própria escrita convencional que demonstram exercer peso nas decisões escritas tomadas pelas crianças para seus próprios textos, à medida que elas vão entrando em contato com outros textos escritos. Com base nessa consideração, os autores apontam para aspectos da própria ortografia, como forma de explicar as grafias em que as crianças segmentam “mais do que a ortografia exige [os casos de hipersegmentação], a partir da atribuição de conteúdos semânticos específicos a subparte das palavras” (ABAURRE E CAGLIARI, 1985, p. 28) – como destacam os autores, por exemplo, em relação ao dado “ador mesida” (*adormecida*).

Abaurre (1988a) reforça as argumentações do estudo anterior, sobre as hipóteses infantis não se mostrarem de forma decisiva para todos os problemas de escrita. Mais ainda,

---

<sup>14</sup> Grupo tonal é uma unidade fonética “maior do que o pé (por isso pode ser composto de um ou mais pés) e é delimitada por um padrão entoacional específico chamado tom” (SILVA, A., 1991, p. 40).

nesse trabalho a autora parece ampliar a rede de possibilidades que poderiam explicar as segmentações aquém/além do esperado pela convenção ortográfica:

Em que baseia suas propostas de segmentação dessa escrita, que se lhe apresenta como constituída de elementos discretos? Em critérios semânticos? Em critérios fonéticos? Na observação da própria escrita e em inferências que tal observação lhe permite fazer? Ou em todos esses critérios e ainda em outros, mais ou menos simultaneamente, manifestando um comportamento epilinguístico que lhe permite operar sobre esse objeto de forma a resolver localisticamente, segundo o critério que lhe parecer mais adequado a cada instante, os inúmeros problemas que a escrita lhe vai oferecendo? (ABAURRE, 1988a, p. 135-136).

Contudo, na análise apresentada, Abaurre continua a privilegiar os critérios *fonético* e *semântico*, ao explicar as motivações das ausências e presenças não-convencionais dos limites gráficos. Em relação ao primeiro critério, a autora detecta correspondência entre grafias de hipossegmentação e unidades fonéticas, como em “tuczimorino” (*estou quase morrendo*) e “guenãovazia” (*que não fazia*), em que são observados limites de grupo tonal e de força,<sup>15</sup> respectivamente. Quanto ao critério semântico, Abaurre assume existir uma ação da escrita sobre as decisões infantis, não só nas propostas de segmentação, bem como no uso das letras. Para a autora, a escrita opera no texto da criança “dinamicamente, para conduzir uma análise da própria fala e elaborar propostas de representação que com o tempo se constituirão em representações canônicas da língua” (ABAURRE, 1988a, p. 140).

Em publicação posterior, Abaurre (1991c) trabalha a ideia de representações canônicas da língua, quando discute que os aprendizes de escrita demonstram reconhecer, em alguma medida, formas preferenciais das estruturas das palavras do português, nos momentos em que produzem segmentações não-convencionais. De acordo com Abaurre, muitas vezes, ao delimitarem as palavras, as crianças, em suas hipóteses, pareceram combinar o padrão

---

<sup>15</sup> O grupo de força corresponde a um “suporte segmental de uma proeminência acentual possível em termos de enunciado (unidade fonológica – constituída de uma ou mais unidades morfológicas – onde ocorre uma sílaba acentuada)” (SILVA, A., 1991, p. 42).

acentual preferencial da língua (paroxítono) com o número, também preferencial, de sílabas por palavra. Seguindo essa hipótese, é frequente a ocorrência de novas palavras formadas em organização trissílaba e dissílaba, como, por exemplo, “docaro” (*do carro*) e “cata puta” (*catapulta*). Por meio dessas observações, a autora considera que “é possível dizer que elas [as crianças] começam muito cedo a elaborar, embora inconscientemente, algum conceito de palavra da língua” (ABAURRE, 1991c, p. 204).

Ainda, segundo Abaurre, as tentativas de registro das palavras segundo um critério de formas canônicas levantam questões rítmico-prosódicas interessantes, tendo em conta que as grafias demonstram a forte influência de pés binários trocáicos (configuração que alterna uma sílaba forte com uma fraca), unidades nas quais se fundamenta um padrão rítmico preferencial do português brasileiro (PB, posteriormente), como bem sinaliza a autora. Nota-se que, a partir dessa interpretação, Abaurre já assinala aspectos fonológicos subjacentes às hipóteses linguísticas sobre as fronteiras de palavra escrita.

Quanto à ocorrência de determinadas palavras hipersegmentadas, como “cata puta”, a autora destaca, também, a possível atuação de uma informação semântica, com base na qual a criança possa ter reconhecido duas possíveis palavras da língua, a saber: “cata” e “puta”. Segundo Abaurre, a divisão nessas duas unidades gráficas é corroborada em razão de *catapulta* ser uma palavra de uso bastante restrito na língua fato que, quando aliado à organização no padrão dissílabo-paroxítono, teria auxiliado a criança a reforçar a ideia de serem duas palavras de significado conhecido.

Observando a organização dos textos infantis, a autora sinaliza que determinados pontos da produção escrita, relacionados à disposição dos elementos da narrativa, estão mais propensos à ocorrência de segmentações não-convencionais. Abaurre aponta, por exemplo, que as crianças escrevem mais convencionalmente, no tocante à segmentação em palavras, quando, ao desenvolverem o enredo, se colocam como narradoras de terceira pessoa e

parecem segmentar menos que o previsto quando demarcam a fala das personagens por meio do discurso direto. Nesses momentos surgem maiores números de hipossegmentações, o que poderia ser explicado com a aproximação dos diálogos em relação à característica contínua dos enunciados falados e, por esse aspecto, seriam registrados em porções gráficas maiores. No entanto, Abaurre salienta: o levantamento dessa possível hipótese explicativa não significa afirmar que o restante das palavras escritas do texto foram segmentadas, pelas crianças, ao encontro da convenção.

Grande parte das considerações formuladas por Abaurre (1988a, 1991c) estão reunidas e sistematizadas em Silva, A. (1991). Esse autor desenvolve um estudo baseado na escrita de crianças da antiga primeira série (2º ano, atualmente) e, de modo específico, buscou observar a relação da fala com as ocorrências de hipo e de hipersegmentação. Nota-se, contudo, que os resultados de Silva, A. foram bem mais expressivos, demonstrando que a criança não se fundamenta apenas nos aspectos fonéticos da fala, mas em características da convenção ortográfica e, também, em combinatórias entre aspectos da fala e da escrita, além de variações de outra natureza (como a oscilação na forma de grafar uma mesma palavra).

Em relação à fala, Silva, A. argumenta sobre a atuação de grupos tonais e de força, especialmente nos casos em que o espaço em branco não empregado. O autor observa que as ocorrências por grupos tonais parecem preceder, na maioria dos casos analisados, as ocorrências fundamentadas em grupos de força. Sobre as presenças não-convencionais dos espaços em branco, Silva, A. identifica, nesse tipo de segmentação, uma busca da criança em registrar aspectos da escrita convencional, à exemplo de: (i) casos em que as crianças inferem, a partir de unidades como *o, a, um, da, de, em, com*, possíveis palavras escritas independentes, como ocorre em “da cela” (*daquela*); (ii) casos em que a interrupção da escritura, para acrescentar um determinado símbolo à palavra (como um acento gráfico, por exemplo), pode acarretar “muito facilmente, posições de corte para a criança nessa fase” (SILVA, A., 1991, p.

51), hipótese que pode explicar a escrita de “rão cou” para *roncou*; e (iii) casos em que a delimitação das fronteiras não-convencionais se dá em pontos de reconhecimentos de outras palavras, por exemplo, *felicidade* grafada como “feliz sidade”. O autor acrescenta, ainda, que o acento tônico das sílabas pode ser um indicativo para que ocorra a inserção de um espaço. Seguindo essa linha de raciocínio, Silva, A. destaca hipersegmentações como “com tente” (*contente*). Mas o levantamento dessa nova hipótese não abandona as outras já apontadas, no sentido de que as hipóteses explicativas podem *conviver*. No entanto, essa convivência não se localizaria em um mesmo dado, visto que é possível “prever em que circunstâncias ocorrem, ou qual é a mais preponderante” (SILVA, A 1991, p. 61). Não assumiremos, para este trabalho, essa visão proposta por Silva, A. (1991), pois, da perspectiva que assumimos, não há evidências para sustentar quando o aprendiz está dando mais ênfase a um critério em detrimento de outro. Entendemos que diferentes hipóteses atuam simultaneamente; porém, no momento não nos ateremos a essa reflexão, visto que teremos mais subsídios que a sustentam quando considerarmos os trabalhos de orientação analítica que definimos como fonológico-letrada.

Por todo o apresentado, reconhecemos a importância dos estudos de Abaurre e Cagliari (1985), Abaurre (1988a, 1991c) e Silva, A. (1991), não apenas por serem os primeiros estudos brasileiros a abordarem o tema das segmentações não-convencionais de um ponto de vista linguístico, mas, fundamentalmente, por, em suas formulações explicativas, estarem distantes de uma perspectiva de análise essencialmente normativa. Notamos que, ao se distanciarem dessa visão, aqueles pesquisadores, além de ressaltarem a importância desse tipo de dado para as investigações linguísticas, mostram que os textos de início de aquisição da escrita são, acima de tudo, um lugar de observação do trabalho do sujeito sobre a linguagem. As contribuições trazidas por esses trabalhos são vistas, ainda, em razão das suas discussões terem se constituído como ponto de partida para as reflexões mais recentes sobre o

tema, ampliando até mesmo o interesse de pesquisadores para esse fenômeno em produções escritas de escreventes com mais anos de escolarização e, também, de adultos em processo de alfabetização. Estamos nos referindo aos estudos que, no início desta subseção, denominamos como de orientação explicativa fonológico-letrada, os quais, a partir deste momento, passarão a ser explorados.

#### *1.1.2.2. Tendência fonológico-letrada*

Os trabalhos de Capristano (2003, 2004), Chacon (2004, 2005, 2006), Tenani (2004, 2009a, 2010, 2011b) e Paula (2007) são desenvolvidos sob a concepção do *modo heterogêneo de constituição da escrita*, conforme Corrêa (1997, 2004). Resultante dessa escolha teórica, os estudiosos têm identificado, nos tipos de segmentação não-convencional, informações linguísticas relativas às práticas sociais de linguagem orais/faladas e letradas/escritas. Para fins explicativos, os autores demonstram que, em diferentes graus, as hipo e hipersegmentações obedecem aos princípios reguladores “estabelecidos para a definição de constituintes prosódicos, tais como aqueles propostos por Nespor e Vogel (1986)”<sup>16</sup> (CHACON, 2004, p. 223) e, também, a estruturas possíveis dentro da convenção ortográfica. No entanto, Chacon (2005) faz uma ressalva: considerando que a análise das segmentações não-convencionais permitiria, ao investigador, a identificação de vários outros aspectos linguísticos-discursivos advindos de práticas sociais orais/letradas, privilegiar os aspectos prosódicos e ortográficos é apenas uma decisão metodológica que busca analisar mais atentamente a atuação da prosódia e da própria convenção ortográfica nas hipóteses que são construídas por escreventes em aquisição daquilo que seja o limite de palavra escrita.

---

<sup>16</sup> Todos os trabalhos apresentados nesta subseção, que se valeram de informações fonológicas para explicar estruturas de segmentação não-convencional, o fizeram com base no modelo de Fonologia Prosódica de Nespor e Vogel (1986).

Havíamos antecipado que um de nossos intuitos, ao revisitarmos as linhas de análise dispensadas ao objeto de nosso estudo, era o de podermos adotar aquela que mais dialogasse com os nossos objetivos e escolhas teóricas. Esta pesquisa, pois, se filia aos trabalhos desenvolvidos com base em análises fonológico-letradas. Desse modo, restringimos nossa investigação para a possível atuação da dimensão prosódica e gráfica na análise dos nossos dados. Daqui em diante, passamos a apresentar as contribuições dos trabalhos indicados no início deste item.

Iniciamos com o estudo de Capristano (2003). Na análise proposta, com base em textos de crianças da primeira série, Capristano reúne as marcas de segmentação não-convencional (hipo e hipersegmentação) em tipos de funcionamentos que apreenderiam, segundo a autora, propriedades dos dados capaz de “pensar o relacionamento entre aquilo que poderia ser considerado local (os *pequenos fatos* de escrita infantil) e aquilo que poderia ser considerado como global e/ou possível de ser generalizado” (CAPRISTANO, 2003, p. 97, grifos do original). Da reunião das marcas, traçam-se quatro funcionamentos distintos: 1) segmentações não-convencionais resultantes de tentativas de escrita alfabética; 2) segmentações não convencionais resultantes da oscilação entre diferentes trânsitos por constituintes prosódicos e informações sobre o código escrito institucionalizado; 3) segmentações não-convencionais resultantes da oscilação entre constituintes abaixo do domínio da palavra fonológica da hierarquia prosódica (sílabas e pé) e informações sobre o código escrito institucionalizado; e, por último, 4) segmentações não-convencionais resultantes de maior percepção de constituintes acima da palavra fonológica na hierarquia prosódica e, talvez, em menor grau, de informações sobre o código escrito institucionalizado. Destaca-se que os funcionamentos, propostos pela autora, estão fundamentados na relação das fronteiras não-convencionais de palavras com fronteiras de constituintes prosódicos, por um

lado, e com um dos eixos do imaginário sobre a escrita – *o código escrito institucionalizado* (cf. CORRÊA, 1997, 2004) – por outro.

Em publicações posteriores, Capristano (2004), Tenani (2004) e Chacon (2004) retomaram três dos funcionamentos discutidos em Capristano (2003), já que “os dados selecionados por Capristano (2003) permitem observar, de modo relativamente privilegiado, pistas das características prosódicas da língua” (TENANI, 2004, p. 235).

Capristano (2004) analisa o **quarto** funcionamento (segmentações não-convencionais resultantes de maior percepção de constituintes acima da palavra fonológica na hierarquia prosódica e, talvez, em menor grau, de informações sobre o código escrito institucionalizado), por se tratar do tipo correspondente ao maior número de dados. De acordo com Capristano, a localização dos espaços gráficos não-convencionais permite, de modo geral, chegar à conclusão de que as grafias são, sempre, resultado de uma ação simultânea e, em diferentes graus, de fatores de práticas sociais orais e letradas nas quais estão inseridas as crianças. De modo específico, a autora aponta dois fatores que parecem atravessar as estruturas hipo e hipersegmentadas, no funcionamento analisado, e que a permitiu alçar à sua conclusão mais geral. O primeiro fator funda-se em aspectos prosódicos, mais especificamente, em fronteiras de constituintes prosódicos, que, ao marcarem-se com certa saliência para as crianças, as levariam supor a possibilidade de transferência unívoca para o texto escrito. O segundo fator remete à possível ação de uma imagem que as crianças teriam do que são as estruturas próprias da escrita.

Tenani (2004) analisa o funcionamento **dois** (segmentações não convencionais resultantes da oscilação entre diferentes trânsitos por constituintes prosódicos e informações sobre o código escrito institucionalizado) relacionando os textos infantis com produções de adultos em alfabetização, no intuito de mostrar que os espaços em branco empregados de maneira não-convencional demarcam fronteiras e/ou proeminências de constituintes

prosódicos que evidenciam a “*possibilidade de sentidos* associados às diferentes segmentações da cadeia fônica” (TENANI, 2004, p. 236, grifos nosso). A esse respeito, Tenani argumenta que a possibilidade de sentido não é garantida na escrita unicamente pela delimitação das palavras, mas, também, por meio da disposição das sílabas proeminentes nos enunciados. A autora lembra que esse mecanismo de compreensão dos sentidos é o de que se valem, por exemplo, textos chistosos,<sup>17</sup> já que

entender a piada pede a descoberta de possíveis similaridades na cadeia fônica que estão “estrategicamente” articuladas com as dessemelhanças que se tornam visíveis por meio da consideração das possibilidades de localização das fronteiras de constituintes prosódicos e dos elementos proeminentes do enunciado (TENANI, 2004, p. 237).

Traçando um paralelo entre textos chistosos e aqueles produzidos durante a aquisição da escrita (por crianças e adultos), Tenani assinala que, na tarefa de delimitar palavras escritas – especialmente por escreventes sobre os quais se supõe não dominar os critérios convencionais que essa atividade exige –, é preciso lidar, dentre outros conhecimentos, com a descoberta de palavras “escondidas” dentro das possibilidades de significação de um mesmo contínuo fônico. À luz dessas considerações, a autora afirma não ser coincidência que apareçam, com grande frequência, ocorrências diferentes de segmentação para uma mesma palavra, no interior de um mesmo texto, ou em textos diferentes de um mesmo escrevente. Tenani enfatiza, ao fim, a não estabilidade e homogeneidade da língua, entendendo que os dados não-convencionais de escrita são relevantes por permitirem que se mostrem “‘as fendas e rachaduras’ das estruturas linguísticas” (TENANI, 2004, p. 242).

O trabalho de Chacon (2004) desenvolve-se a partir do **terceiro** funcionamento formulado por Capristano (2003) – ou seja, segmentações não-convencionais resultantes da oscilação entre constituintes abaixo do domínio da palavra fonológica da hierarquia prosódica

---

<sup>17</sup> Em trabalho anterior, Tenani (2001) dedicou-se, exclusivamente, à análise de mecanismos fonológicos na construção de textos chistosos.

(sílabas e pé) e informações sobre o código escrito institucionalizado. Chacon (2004, p. 225) seleciona, do conjunto de dados, onze marcas cujas características estruturais não se assemelham entre si nem em relação à convenção ortográfica. Os dados eleitos pelo autor apresentam, ainda, a particularidade de não parecer se explicar “com base em algoritmos de formação como aqueles que definem constituintes da hierarquia prosódica, tal como proposta por Nespor e Vogel (1986)”. No entanto, o autor não vê, por meio dessas características, a anulação de características mais regulares como a correspondência das grafias não-convencionais com palavras ou apenas partes de palavras do léxico da língua – que poderia ser explicado pela ação de práticas de letramento sobre as crianças, tanto aquelas ligadas diretamente à escolarização quanto aquelas que se desenvolvem fora desse contexto, mas que estão igualmente envolvendo os escreventes.

Chacon mostra que as segmentações não-convencionais analisadas não estão restritas a uma estrutura por meio da qual seria possível identificar, por exemplo, princípios regulares como os de formação de constituintes prosódicos. Considerar ocorrências como “*acho rar*” (*achorar*) e “*ficoa paxonado*” (*ficou apaixonado*), apenas enquanto estrutura, seria atribuir-lhes estatuto de “manifestações imperfeitas de um princípio mais geral da oralidade que atuasse, de algum modo, na escrita” (CHACON, 2004, p. 230). No entanto, essa interpretação não implica em negar por completo o funcionamento da prosódia no enunciado escrito, na medida em que com uma afirmação dessa natureza estaria subjacente a postura de a possível interferência da fala na escrita, o que, para o autor, significa assumir a existência de autonomia dos modos enunciativos da língua.

A proposta de Chacon é, então, chamar a atenção para a possibilidade de que haveria, antes de tudo, um funcionamento simultâneo (e não hierárquico) entre os constituintes prosódicos, pois a observação dos dados demonstra, de modo mais amplo, que toda a sequência não-convencional indicia a ação de constituintes maiores (como os localizados

acima da palavra fonológica), enquanto que, de modo mais específico, as partes demarcadas entre limites indiciam fronteiras de constituintes menores (como os formados abaixo da palavra fonológica). Para Chacon, a simultaneidade de funcionamento seria desencadeada: (i) por uma não-sincronização entre limites *de* constituintes e de prominências *em* constituintes; (ii) por uma ação mais latente do ritmo (evidenciado em seus dados pela ação do pé métrico) e menor de outros fatores prosódicos, pois como o próprio autor já havia defendido, em trabalho anterior (CHACON, 1998), toda a linguagem é organizada pelo ritmo; (iii) por uma dissensão entre os dados de linguagem (as segmentações não-convencionais, nesse caso) e os padrões dos modelos de análise; e/ou, ainda, (iv) por correspondências gráficas e sonoras relacionadas com palavras escritas da língua, pela ação do letramento sobre os sujeitos.

Em outros estudos, Chacon (2005, 2006) abordou mais detalhadamente o aspecto apresentado em (ii). Em se tratando do primeiro trabalho, Chacon (2005) analisa grafias de trissílabos hipersegmentados produzidos por crianças em início de aquisição da escrita, por enxergar na estrutura de palavras compostas por três sílabas fatos mais interessantes, quanto a questões prosódicas, do que em palavras polissílabas e dissílabas, pois: (i) dados a partir de palavras polissílabas apareceram com baixa frequência no cópulo analisado pelo autor; e (ii) as ocorrências de palavras dissílabas, embora em termos numéricos tenham ocorrido quantidade maior do que as palavras trissílabas, no que se refere à ação da prosódia mostram variação pouca significativa, dada a presença do limite gráfico ter ocorrido entre as sílabas das palavras, mostrando que apenas o constituinte prosódico sílaba foi relevante nessas ocorrências.

Chacon parte da tese de que qualquer ponto de um trissílabo pode oferecer contexto a um possível lugar de separação não-convencional. Porém, não foram em quaisquer lugares que se deram as rupturas; pelo contrário: o modo como os espaços em branco foram inseridos indicia que princípios recorrentes de estruturação da oralidade foram obedecidos, como, “é o

caso do *ritmo*, que, já estruturado na língua, marca-se na oralidade, por exemplo, por meio de contrastes entre sílabas acentuadas e não-acentuadas” (CHACON, 2005, p. 81, nosso grifo). Esses contrastes se marcam na língua, segundo o autor, em virtude de uma relação do tipo dominante/dominado que se estabelece e define o constituinte prosódico pé métrico.

A característica rítmica é marcante nos trissílabos analisados por Chacon, uma vez que das 136 ocorrências levantadas, 111 (81,62%) apresentaram a ruptura entre os limites de sílabas e de pés, a partir das quais: (i) 92 (82,88%) delas promoveram a combinação de uma sílaba e um pé, como em “e Rita” (*irrita*); e (ii) 19 (17,12%) geraram a combinação contrária, ou seja, de um pé e uma sílaba, como em “colo que” (*coloque*). Nas 25 ocorrências restantes do total de dados, as rupturas ocorreram nos limites silábicos, a exemplo de “sau da de” (*saudade*).

Sob o risco de deixar de contemplar a complexidade das hipersegmentações, Chacon passa a analisar, em um segundo momento do trabalho, as características dos dados que indiciam a circulação da criança por práticas letradas/escritas. Nesse ponto, o autor destaca que, ao romperem os trissílabos em qualquer um dos pontos descritos (ou seja, nos limites silábicos, nos limites de sílaba e pé ou nos limites de pé e sílaba), as crianças deixam pistas de que uma parte da palavra pode ter sido reconhecida como um monossílabo da língua, o qual, grande parte das vezes, é correspondente a categorias gramaticais; é o que exemplifica a sílaba à esquerda da grafia hipersegmentada da palavra *porquinho*: “por quinho”. Mas quando não é identificado um monossílabo, o reconhecimento pode ocorrer com o dissílabo restante, como, por exemplo, na possível palavra “cola” a partir da hipersegmentação de *escola*: “s cola”.

Como último destaque sobre as considerações do autor, apontamos o refinamento ortográfico que alguns dados demonstraram. Para ilustrar esse ponto, retomamos análise que Chacon faz da hipersegmentação “e Rita” (*irrita*). Para o pesquisador, esse dado marca

fortemente o trânsito do escrevente por práticas letradas, pois a criança elege grafar com a letra <e> um segmento que, em enunciados falados mais informais (esperado de uma criança), é pronunciado como [i]; mais ainda, é a partir de uma informação exclusivamente (orto)gráfica que advém o emprego de uma letra maiúscula (<R>) no início de uma palavra que a criança reconhece ser nome *Rita*.<sup>18</sup>

Posteriormente, Chacon (2006) reaborda as pistas linguísticas que podem ser observadas nas hipersegmentações e que marcam a presença das crianças em práticas de linguagem orais e letradas. Nesse trabalho, além das hipersegmentações em palavras trissílabas, o autor também analisou a ocorrência dessas marcas em palavras polissílabas e dissílabas. As conclusões a que o autor chegou nessa análise são muito semelhantes às observadas no estudo anterior – Chacon (2005) – por isso, não as retomaremos a exaustão. Simplesmente nos restringimos a destacar que, a respeito da aquisição da noção de palavra, o autor afirma que essa não apresenta apenas questões exclusivamente estruturais para a reflexão do escrevente; “pelo contrário, essa padronização que a escrita codifica e com a qual o escrevente trabalha provém de práticas discursivas efetivas: em que se cruzam dois diferentes modos de enunciação verbal: o falado e o escrito” (CHACON, 2006, p. 165). Ainda sobre esse aspecto, mais adiante o autor acrescenta:

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

---

<sup>18</sup> A escolha de letras, para registrar os limites das segmentações (tanto convencionais quanto não-convencionais), foi alvo da discussão de Tenani (2009b). A autora mostra, ao analisar textos de alunos dos dois primeiros anos do EF-II, que a maneira como o escrevente seleciona as letras do alfabeto é indício do modo como informações de práticas orais e letradas constituem hipóteses acerca dos limites de palavra. Com esse estudo, Tenani (2009b) conclui que o escrevente *flutua* tanto em relação às possibilidades de segmentação (segmentar convencional ou não) quanto no que se refere às possibilidades de preenchimento de uma dada posição da palavra por uma determinada letra. De acordo com a autora, essa flutuação indica motivações de natureza prosódica, “que dizem respeito à organização das sílabas átonas em diferentes constituintes e às realizações dos segmentos em diferentes contextos fonológicos” e de natureza letrada, “que dizem respeito à colocação de espaços em branco indicadores de palavras na escrita e às escolhas de grafemas” (TENANI, 2009b, p. 123).

língua, tanto em seu modo de enunciação falado, quando em seu modo de enunciação escrito (CHACON, 2005, p. 165).

O trabalho de Paula (2007) também analisa exclusivamente dados de hipersegmentação. Nessa pesquisa, a autora realiza um estudo longitudinal de textos de oito crianças dos três primeiros anos do EF. Com base nesse material, Paula busca demonstrar que as hipersegmentações da escrita infantil são marcas indissociáveis, ao mesmo tempo, de um processo mais geral do aprendizado institucional da escrita, perpassado pela participação da criança em práticas de oralidade e letramento, e de um processo mais característico de cada criança, mostrado em momentos nos quais se observa a manifestação de sua subjetividade.

A respeito do processo mais geral, Paula conclui que as crianças mobilizam estruturas dos enunciados falados, formalizadas em constituintes prosódicos, ao produzirem segmentações não-convencionais em seus textos. Em sua análise, a autora pôde observar que as hipersegmentações indicaram maior sensibilidade em relação aos limites de sílaba e de pé, os dois domínios prosódicos localizados abaixo do nível da palavra. Sobre esses constituintes, a presença mais frequente do espaço em branco ocorreu entre duas sílabas: “que ro” (*quero*) e a segunda mais comum se deu entre uma sílaba e um pé métrico: “es tava” (*estava*) – particularmente quando esse último apresentou a configuração de um troqueu. Paula, no entanto, não descarta, da constituição das hipersegmentações, aspectos representativos de enunciados escritos, “já que, na produção escrita, o sujeito parece recuperar elementos verbais (como os de natureza prosódica e os de natureza ortográfica) que circulam nessas práticas [letradas/escritas]” (PAULA, 2007, p. 91).

Em relação ao fato de os dados permitirem a reconstrução de momentos de subjetividade das crianças, Paula organiza em três movimentos as transformações pelas quais ela observa ter passado a escrita de cada escrevente. O primeiro movimento (que atingiu a escrita de cinco crianças investigadas) caracteriza-se por apresentar, inicialmente, uma forma de escrita comprometida quanto à atribuição de sentido às palavras, especialmente em função

de as letras terem sido empregadas sem qualquer marca de segmentação, e por apresentar, em um segundo momento, uma forma de escrita que, apesar de ser possível atribuir algum sentido às palavras, a caligrafia dificultou na identificação dos espaços em branco. Ao longo dos anos escolares, esse primeiro movimento desapareceu das produções escritas dos escreventes, visto que as distribuições dos espaços em branco e a atribuição dos sentidos passaram a se mostrar mais claramente.

A autora verifica, em relação ao segundo movimento (presente na escrita de duas crianças), a convivência entre uma escrita cujos blocos de letras comprometem a identificação dos sentidos e uma escrita em que os espaços em branco organizam-se irregularmente. A detecção desse movimento foi se tornando esporádica, sendo mais frequente nas produções escritas da primeira série.

Conforme observa a autora, apenas uma criança apresentou o terceiro movimento. Em termos das características, esse movimento delinea-se a partir de uma escrita que poderia ser aproximada das formas privilegiadas pela convenção ortográfica, tanto no que se refere às questões que envolvem a observação dos sentidos das palavras quanto em relação à clara identificação dos espaços em branco. Com o passar dos anos escolares a ocorrência do terceiro movimento foi diminuindo, bem como o número de grafias de hipersegmentação presentes nos textos da criança agrupada nele:

Isto nos leva a crer que estes movimentos resultariam, dentre outros fatores, da maior ou menor participação do escrevente em práticas sociais que envolvem a escrita, fato que colocaria essas crianças em diferentes planos de aquisição desse modo de enunciação da linguagem – visto que, no decorrer dos anos, as crianças que apresentaram o *primeiro movimento* se aproximaram das que apresentaram o *segundo movimento* e estas, por sua vez, aproximaram-se das que apresentaram o *terceiro movimento* (PAULA, 2007, p. 115).

Seguindo o mesmo eixo teórico-metodológico de análise dos trabalhos de Capristano (2003, 2004), Chacon (2004, 2005, 2006) e Paula (2007), Tenani (2009a, 2010, 2011a,b)

desenvolve reflexões acerca dos tipos de segmentação não-convencional que podem ser encontrados na produção escrita de alunos dos anos finais do EF. Em suas pesquisas, a autora busca mostrar que a dificuldade em se identificarem os limites de palavra escrita não está circunscrita aos escreventes em início de aquisição da escrita (adultos ou crianças), mas atinge até a escrita de escreventes com mais tempo de escolarização. Desse modo, os trabalhos de Tenani vêm procurando investigar, juntamente com outras pesquisadoras, frutos de orientação acadêmica da autora – cf. Paranhos (2010) e Silva, L. (2011) –, de quais soluções linguísticas escreventes já saídos do período de alfabetização se valem para propor delimitações às palavras. Partindo dessa questão maior, os trabalhos observam as características prosódicas e os aspectos da escrita que se evidenciam nas marcas de segmentação não-convencional e propõem refletir em que medida os dados presentes no EF-II convergem e/ou divergem em relação aos produzidos no início da aquisição da escrita. Quanto ao aspecto prosódico, a argumentação da autora tem se encaminhado na defesa dos constituintes prosódicos grupo clítico, palavra fonológica, pé métrico e sílaba para descrição e interpretação dos casos em que as segmentações em palavras fogem à convenção.

Em um de seus estudos, Tenani (2009a) assume a proposta de Bisol (2000, 2005) a respeito da pertinência do grupo clítico como um domínio prosódico relevante para a descrição de fatos fonológicos do PB. Ao analisar marcas de hipo e hipersegmentação, Tenani argumenta que as grafias encontradas nos anos que encerram o EF caracterizam-se por colocar em evidência o modo de organização dos elementos átonos da língua. Em outras palavras, a dúvida que parece ser frequente nas hipóteses dos escreventes é a de quando os elementos átonos são sílabas (pretônicas ou postônicas) de uma palavra fonológica e de quando são unidades (clíticos) que se comportam junto a uma palavra formando um domínio maior que, na interpretação de Tenani, é o grupo clítico.

No conjunto de textos analisados, a autora identifica que, nos casos das hipossegmentações, a ausência do espaço em branco é indício de que os escreventes tenham interpretado a cadeia fônica como uma só unidade lexical. Prosodicamente, esse resultado corresponde a um grupo clítico ter sido analisado como uma palavra fonológica. Nas ocorrências de hipersegmentação, a presença do espaço em branco é uma pista de que uma só sequência fônica foi analisada, pelo escrevente, como duas unidades lexicais independentes. Nesse caso, em termos prosódicos, uma palavra fonológica foi interpretada como um grupo clítico. Em virtude dessas análises, Tenani conclui que há flutuação entre os dois constituintes prosódicos (palavra fonológica e grupo clítico), explicada em função da imprecisão dos estatutos de sílabas átonas e de clíticos fonológicos.<sup>19</sup>

Ao fim de seu trabalho, Tenani reúne aquilo que considera as principais contribuições de seus resultados, aqui sintetizadas: (i) *indícios de hipóteses do que seja palavra*: para Tenani as hipóteses dos escreventes acerca dos limites de palavra são ancoradas tanto em informações de natureza prosódica, cuja evidência relaciona-se à organização de elementos átonos (sílabas ou clíticos) em diferentes constituintes, quanto em informações de natureza letrada, a qual parece mostrar-se especialmente pela colocação do espaço em branco indicar possíveis palavras escritas – (particularmente no caso das hipersegmentações); (ii) *pistas da reflexão dos escreventes sobre classes gramaticais*: a configuração dos dados evidencia, segundo a autora, que os escreventes ainda encontram dificuldades em identificar itens gramaticais como preposição, pronomes e conjunção. A dúvida consiste, exatamente, em identificar quando os elementos átonos são itens gramaticais e quando são sílabas de uma palavra; e (iii) *evidências da organização prosódica da língua*: desse aspecto, Tenani enfatiza a flutuação dos escreventes em organizarem prosodicamente as sílabas átonas: ora como parte de uma palavra fonológica ora como parte de um grupo clítico. Como já demonstramos, nas

---

<sup>19</sup> Em Tenani e Paranhos (2011), as autoras reafirmam essa linha de interpretação e a pertinência dos constituintes prosódicos “palavra fonológica” e “grupo clítico” para a descrição dos dados de segmentação não-convencional do EF-II.

hipossegmentações essa flutuação se mostra em relação ao clítico ser interpretado como sílaba da palavra. Inversamente, nas hipersegmentações a flutuação revela-se pela sílaba da palavra ser registrada como um clítico. Quanto a esse último tipo de segmentação, Tenani alerta sobre o fato de as hipersegmentações evidenciarem não só a leitura de uma sílaba pretônica como um clítico, correlacionado, por sua vez, a uma possível apreensão desses elementos com palavras independentes na escrita, mas, também, por sua estrutura evidenciar a saliência de um componente tônico e/ou, ainda, a relação conjunta desse aspecto tônico com a observação dos aspectos da escrita.

Fundamentada no aspecto tônico que pareceria figurar nas hipersegmentações, Tenani (2010) propôs uma segunda linha interpretativa para esse tipo de dado – sem excluir a hipótese formulada em Tenani (2009a). Nesse trabalho, a autora se aproxima das interpretações de Chacon (2005) e de Paula (2007), dedicadas aos dados infantis, para os quais as hipersegmentações mobilizariam informações prosódicas relacionadas aos constituintes sílaba e pé métrico. Particularmente, Tenani desenvolve um estudo transversal das hipersegmentações pelos quatro últimos anos do EF.

Os resultados da autora demonstram que as hipersegmentações foram motivadas, preferencialmente, pelos constituintes prosódicos sílaba e pé métrico e pelas diferentes combinações entre eles: sílaba + sílaba; sílaba + pé; pé + pé. Quanto à configuração prosódica dos dados, foi interessante observar que a ruptura na palavra ocorreu no ponto adjacente à sílaba de maior proeminência sonora da palavra, ou seja, no limite do acento primário (destaca-se que das cinco estruturas descritas, quatro apresentaram essa motivação). Tenani aponta, ainda, que dentre todas as combinações, a sequência de sílaba + pé troqueou, foi a única que se manteve em todas as séries, revelando a forte ancoragem dos escreventes – ainda em níveis finais de escolarização – em estruturas preferenciais da língua, algo já observado por Abaurre (1991c), Chacon (2005) e Paula (2007) para dados de escrita inicial. Cabe

pontuar que, aliados à questão rítmica, os dados analisados por Tenani parecem apontar, sempre, para aspectos de práticas letradas/escritas, em especial, quando as marcas linguísticas observadas nas grafias não-convencionais indiciam a tentativa de autonomia gráfica de partes de palavras: como, por exemplo, as formas verbais de “dizer” e “dar”, em “disse deram” (*decidiram*) e os possíveis itens gramaticais em “em quanto” (*enquanto*).

Tenani (2011a) destaca, em outra publicação, que as diferenças entre EF-I e EF-II, no que tange à segmentação em palavras, fundam-se em informações de natureza *letrada*. Segundo a autora, parece ficar mais evidente essa afirmação quando são comparados dados de segmentação não-convencional (de ambas as etapas escolares) que tiveram o mesmo contexto estrutural de palavra para a origem da grafia não-convencional. Por exemplo: “ma telo” (*martelo*, dado infantil) e “a migo” (*amigo*, dado do EF-II), em que se verifica que, no período do EF-I, nem sempre a sílaba pretônica, grafada à esquerda do espaço em branco, coincide com algum clítico correspondente, por sua vez, com alguma classe gramatical, característica essa que foi sempre observada nos dados de hipersegmentação do EF-II, conforme reforça Tenani. Desse modo, a autora argumenta que a correspondência que as sílabas segmentadas apresentam com itens gramaticais é um aspecto importante

não somente por dar visibilidade aos tipos de categorias gramaticais com as quais os alunos ainda têm dificuldades, mas por permitirem observar uma diferença qualitativa quando comparados estes dados com aqueles produzidos por crianças na fase inicial da aquisição da escrita infantil (TENANI, 2011a, p. 237).

As segmentações não-convencionais produzidas pelos escreventes do EF-II, ao serem analisadas sistematicamente nos diversos estudos de Tenani, despertam a atenção dessa autora para um fato novo nas pesquisas sobre o tema: a recorrência de dados que envolvem o emprego não-convencional do *hífen*. Tenani (2011b) observa que os contextos de colocação do hífen relacionam-se a: (i) no caso das hipossegmentações (a maioria dos dados

encontrados), ocorrências em estruturas *verbo+clítico*, como em “despistalos” (*despistá-los*); (ii) no caso das hipersegmentações, dois resultados estão em convivência: o primeiro ligado a estruturas de *palavras compostas*, como, por exemplo, “mau-humor”, e o segundo relacionado, também, com estruturas *verbo+clítico*, como em “estava-mos” (*estávamos*). Em todas as ocorrências, a autora atribui o uso do hífen ao processo de letramento dos escreventes, na medida em que interpreta que as estruturas em que esse recurso gráfico aparece estão relacionadas àquelas estruturas fortemente trabalhadas (e, mais ainda, valorizadas) nas atividades escolares de escrita, sobretudo, no período dos anos finais do EF.

#### 1.1.2.3. Outras formas de análise

Deixamos para o fim, a exposição dos trabalhos de Cunha (2004, 2010) e Ferreira (2011), os quais, apesar de embasarem suas análises nas discussões sobre a organização da fala em constituintes prosódicos, não partilham, com os trabalhos resenhados no item anterior, a concepção de escrita fundamentada na tese do modo de enunciação heterogêneo de constituição da escrita. Para além dessa diferença teórica, os trabalhos acima referidos, embora de cunho linguístico, buscam, ao fim de suas análises, estabelecer relação de seus resultados com práticas pedagógicas de ensino e aprendizagem de língua materna, a propósito do que encontramos em Cunha (2010):

Temos a convicção da necessidade de que os estudos em linguística devam ultrapassar os domínios da pesquisa acadêmica e, cada vez mais, tenham a devida repercussão em sala de aula. O professor de língua materna precisa conhecer o seu objeto de estudo (a língua) em todos os seus aspectos, para que lhe seja favorecida a possibilidade de trabalhar com seu aluno de forma mais adequada (p. 15).

Inseridos nas reflexões desenvolvidas do Grupo de Estudos sobre a Aquisição da Linguagem Escrita (GEALE – FaE/UFPeL), as pesquisas de Cunha e Ferreira têm como eixo norteador três ideias centrais:

[i] o processo de aquisição da escrita é parte integrante de um processo mais geral de aquisição da linguagem (ABAURRE, 1991); [ii] as crianças atualizam conhecimentos já construídos sobre a gramática sonora de sua língua durante a aquisição da escrita (MIRANDA, 2009a); [iii] o erro ortográfico<sup>20</sup> é elemento chave para a descoberta de hipóteses das crianças sobre o sistema de escrita que estão a adquirir (MIRANDA, 2009b) – (MIRANDA 2012, p. 11).

A partir dessas ideias centrais, Cunha (2004) desenvolve um estudo semilongitudinal e compara textos de um conjunto de crianças de uma escola pública e de uma escola particular da cidade de Pelotas (RS). Dividindo os dados de segmentação não-convencional em variáveis linguísticas (tipo de palavra,<sup>21</sup> tonicidade<sup>22</sup> e tipo de sílaba<sup>23</sup>) e extralinguísticas (série e tipo de escola), Cunha observa, em relação à variável tipo de palavra, uma tendência comum entre as hipo e hipersegmentações: em ambas, a ausência/presença do espaço em branco ocorre entre uma palavra gramatical e uma palavra fonológica, como nos respectivos exemplos de “olobo” (*o lobo*) e “em bora” (*embora*). Para a autora, a motivação para a hipossegmentação, nesse caso, diz respeito ao fato de as crianças, em início de aquisição da

<sup>20</sup> Nos trabalhos do GEALE, a ideia de *erro* está respaldada na teoria piagetiana, para a qual o erro (seja ele de qualquer natureza) é parte integrante do processo de aprendizagem. Assim, o erro é relevante, na medida em que demonstra as hipóteses construídas pelos aprendizes na sua relação com o objeto a ser aprendido. Ressalta-se que essa noção de erro é distinta da qual se vincula, por exemplo, o trabalho de Zorzi (1997), de quem trataremos na Seção 2. Para esse autor, o erro, na aquisição da escrita, pode ser indício de um problema ligado a alguma patologia de aprendizagem, especialmente quando as crianças, “apesar de estarem tendo oportunidades de interagir com a escrita parece que, ao contrário de seus pares, não conseguem sair dos níveis mais elementares ou superficiais de conhecimento” (ZORZI, 1997, p. 107).

<sup>21</sup> Cunha observa com essa variável as possíveis combinações: (i) palavra gramatical + palavra fonológica; (ii) palavra fonológica + palavra gramatical; (iii) palavra gramatical + palavra gramatical; e (iv) palavra fonológica + palavra fonológica. Por palavra gramatical a autora entende como “aquela que engloba segmentos que possuem apenas significado gramatical, como os clíticos” (CUNHA, 2004, p. 62). E, por palavra fonológica, Cunha (2004, p. 62-63) considera-a como aquela que “abarca a palavra lexical (que possui significado) e vai mais além, compreendendo todas as palavras que possuem um acento primário e que, embora não tenham significado conhecido na língua, são candidatas potenciais para tal”.

<sup>22</sup> Em relação à tonicidade, a autora considera tanto a sua atuação de proeminência dentro da palavra, como também, essa relação no interior de um sintagma.

<sup>23</sup> A autora busca observar com essa variável, dois aspectos: (i) ocorrências de possíveis processos de ressilabificação (a partir da aplicação de processos de sândi); e (ii) preservação ou não da estrutura da sílaba.

escrita, terem dificuldades de reconhecer palavras formadas por uma ou duas letras como unidades autônomas da língua, especialmente quando se trata de palavras desprovidas de acentuação fonológica. Já nos dados de hipersegmentação, a tendência descrita mostra o contrário: ou seja, de que a criança, ao identificar a sílaba inicial de uma palavra como uma possível palavra gramatical, procura dar-lhe certa autonomia gráfica. Sobre as demais variáveis linguísticas, Cunha chega ao resultado de que: (i) nas hipossegmentações, as crianças guiam-se por linhas entonacionais e, em contextos favoráveis, registram graficamente processos fonológicos como ditongação e degeminação (por exemplo: “siolharão” (*se olharam*) e “sisquecer” (*se esquecer*), respectivamente); e (ii) nas hipersegmentações, a tonicidade mostra que as crianças buscam preservar o pé responsável pelo acento primário da palavra e, de modo geral, sempre registram, nas grafias não-convencionais, a estrutura silábica do português.

Analisando os dados em função das variáveis extralinguísticas, Cunha conclui que as crianças da escola pública produziram mais ocorrências de segmentação não-convencional do que as crianças da escola particular. Esse resultado sugere, segundo a autora, “que o nível de escolaridade dos pais, e um provável contato maior com a escrita, antes do ingresso na escola, favorecem a ocorrência de menos segmentações não-convencionais” (CUNHA, 2004, p. 120). Além do mais, os resultados quantitativos demonstram que, com o passar dos anos escolares, os números de ocorrências diminuem nos dois tipos de escola. De acordo com Cunha, seus resultados são relevantes para a prática pedagógica que se desenvolve no período da alfabetização, pois auxiliam os professores na compreensão das hipóteses que formulam as crianças sobre a escrita, as quais são fundamentadas em conhecimentos linguísticos, e não são simples problemas de ortografia. Para a autora, a criança é um sujeito ativo no processo de aprendizagem e, por assim ser, formula hipóteses sobre o objeto aprendido, que devem servir de base para as propostas de ensino dos professores.

Em outro trabalho, Cunha (2010) compara os ritmos linguísticos do PB e PE (português europeu) a partir dos dados de segmentação não-convencional produzidos por crianças brasileiras e portuguesas, por acreditar que esses tipos de dados “podem revelar aspectos do conhecimento fonológico, oferecendo, portanto, argumentos inovadores para discussões linguísticas estabelecidas, até então, somente no campo da aquisição da linguagem oral” (CUNHA, 2010, p. 17). A fim de atingir o propósito principal de sua pesquisa, Cunha busca:

(i) descrever os tipos de segmentação não-convencional de palavras realizadas pelas crianças portuguesas. A realização dessa tarefa mostra que os dados de hipossegmentação foram predominantes em relação aos outros tipos de segmentação não-convencional (hipersegmentação e híbridos). Em termos estruturais, os dados de hipossegmentação mais recorrentes envolvem a união de uma palavra fonológica e de uma palavra gramatical, como, por exemplo, “escondeuce” (*escondeu-se*). Já nas hipersegmentações, a separação não-convencional se dá formando uma sequência de palavra gramatical e palavra fonológica como, por exemplo, “da quela” (*daquela*). Sobre os tipos de constituintes prosódicos, os resultados apontam para aqueles presentes acima do nível da palavra fonológica, no caso das hipossegmentações, e para os constituintes menores como o pé métrico e a sílaba, no caso das hipersegmentações;

(ii) cotejar os resultados encontrados em Cunha (2004), quanto à descrição das segmentações não-convencionais nos textos de crianças brasileiras, com as grafias identificadas nas produções lusitanas. Nesse ponto da análise, a autora chega a uma distinção entre os dados do PB e PE, no que se refere às ocorrências que envolvem a combinação entre palavras fonológica e gramatical. No PB, as hipossegmentações decorrem, em sua maioria, da união entre uma palavra gramatical e uma palavra fonológica, enquanto que, no PE, a junção foi entre palavra fonológica e palavra gramatical, revelando uma mudança importante, que é a

da direcionalidade em que o clítico se associa à palavra. Em relação às hipersegmentações, tanto PB quanto PE se assemelham, visto que as separações marcam uma combinação de palavras gramatical e fonológica;

(iii) encontrar nas grafias de segmentação não-convencional, do PB e PE, evidências de processos fonológicos. Em ambas as variedades, Cunha identifica que as ocorrências de hipossegmentação e híbrido proporcionam contextos favoráveis à aplicação de processos fonológicos como: a) vozeamento da fricativa. No PB: quando as junções foram entre uma palavra gramatical e uma fonológica; no PE: apenas um dado registrou o processo, em limites de duas palavras fonológicas; b) *tapping*. Em ambas, o contexto favorável foi hipossegmentações em que, à esquerda, ocorreu a palavra *por*; c) processos de sândi vocálico. A degeminação e a ditongação, no PB, foram aplicadas em hipossegmentações que uniram um pronome a um verbo (ou seja, uma palavra gramatical a uma fonológica); já a elisão ocorreu em poucos casos: no PE ela não foi identificada e a degeminação se aplicou mesmo quando a sílaba alvo do processo era a portadora do acento primário. Já a ditongação ocorreu em poucos casos, talvez, em razão de a realização fonética da vogal /i/, em contexto final de palavra, ser sempre /e/ no PE; d) haplologia. Esse processo foi inexistente no PE e bastante raro no PB, ficando restrito aos casos de reestruturação silábica; e) epêntese e monotongação. Ambos os processos foram incomuns nas fronteiras de palavras e, quando ocorreram, favoreceram a formação da estrutura silábica CV, tanto no PB quanto no PE;

(iv) verificar a ação do acento. Cunha observa que as três categorias de dados (hipossegmentação/hipersegmentação/híbridos) sofreram ação do acento. A esse respeito, as hipossegmentações se mostram mais suscetíveis ao acento secundário, enquanto que os dados de hipersegmentação e híbridos pareceram encontrar motivação no pé portador do acento primário, especialmente, da configuração do troqueu.

Por meio dos resultados das quatro etapas de descrição e análise dos dados, Cunha encontra subsídios para efetivar a análise do ritmo (objetivo central do trabalho). Para tanto, a pesquisadora adota a proposta de Abaurre e Galves (1998) sobre o estudo das diferenças rítmicas entre o PB e o PE. Para essas autoras, o ritmo do PB e do PE se distingue em relação à hierarquização de três restrições, a saber: integridade da palavra fonológica, binariedade do pé e pé trocaico. Quanto ao PE, o ritmo caracteriza-se pela hierarquia: 1) pé trocaico, 2) integridade da palavra fonológica e 3) binariedade do pé. Sobre o ritmo do PB, a hierarquia apresenta-se como: 1) integridade da palavra fonológica, 2) binariedade do pé; 3) pé trocaico – cf. Abaurre e Galves (1998, p. 393-394 *apud* CUNHA, 2010, p. 72-73). Junto com essas estudiosas, Cunha entende que a diferença entre o ritmo nas duas variedades do português está fundamentada na posição que o pé trocaico ocupa no interior da hierarquia (primeiro lugar no PE, último lugar no PB), pois em relação à integridade da palavra fonológica, essa, tanto no PB quando no PE, precede a posição da binariedade do pé. Desse modo, entende-se que “o ritmo em PE é baseado no troqueu, enquanto o ritmo brasileiro se constrói respeitando antes de mais nada as fronteiras de palavras fonológicas” (ABAURRE E GALVES, 1998, p.394 *apud* CUNHA, 2010, p. 74).

Ao comparar os resultados de Abaurre e Galves com os dados de segmentação não-convencional identificados nas duas variedades, Cunha acredita ter encontrado maior evidência do ritmo a partir de ocorrências de escrita que demonstram a direcionalidade com que a palavra gramatical se une à palavra adjacente. Por exemplo, a palavra *te* nas hipossegmentações “para tever melhor” (no PB) e “parate ver melhor” (no PE). Conforme já apresentamos em (ii), as hipossegmentações, no PB, seguiram a tendência de união entre uma palavra gramatical e uma palavra fonológica e, no PE, mostraram preferência pela união entre uma palavra fonológica e uma palavra gramatical, à semelhança do que ilustram os exemplos anteriores. No entanto, o que diferencia, a hipossegmentação “parate ver melhor”, no caso do

PE, é o fato de o monossílabo *te* ter se unido a uma preposição e não a um verbo, sendo a união a essa categoria a maior parte das ocorrências levantadas nessa variedade, a qual reflete a sintaxe da fala.

Então, de acordo com Cunha, o que explica a união não-convencional de “para tener melhor” e “parate ver melhor” é o *ritmo* (as barras inclinadas indicam limites de grupos rítmicos): a) para // tener // melhor (PB); e b) parate // ver me // lhor (PE). Com base na divisão em grupos rítmicos indicada por Abaurre e Galves, Cunha considera que, em (b), o monossílabo *te*, em função da sua atonicidade, insere-se no grupo anterior formando com ele um pé trocaico<sup>24</sup>, o que justifica, ainda, a presença da sílaba *me* (também átona) junto com o grupo rítmico a que pertence o *ver*. Mas, segundo Cunha (2010, p. 145), não se espera a grafia de *me* hipossegmentada ao *ver*, uma vez que a escrita não é “uma transcrição da fala; portanto, não pretendemos que as crianças escrevam obedecendo literalmente ao ritmo da língua”. Sobre o grupo representado em (a), a autora interpreta que a sua formação, além de ser esperada pelo tipo de preferência do PB em unir o pronome ao verbo pela próclise, também respeita a integridade da palavra fonológica em relação à formação de um pé troqueu. Assim, para Cunha (2010, p. 174), “a hierarquia proposta pelas autoras [ABAURRE E GALVES, 1998], para a análise do ritmo no PB e no PE, é igualmente adequada à análise rítmica das segmentações não-convencionais”.

Em sua pesquisa, Ferreira (2011) trata de processos de segmentação não-convencional no contexto da Educação de Jovens e Adultos (posteriormente, EJA). A autora analisa a escrita de três sujeitos que, à época da coleta dos dados, tinham 44, 22 e 16 anos de idade e cursavam a segunda etapa do EJA.<sup>25</sup> As grafias de hipossegmentação, hipersegmentação e híbridos são analisadas tendo por base as variáveis linguísticas, definida por Cunha (2004),

<sup>24</sup> Cunha (2010, p. 144) lembra que Abaurre e Galves “consideram trocaicos os pés com cabeça à esquerda, tanto os que possuem duas sílabas como os que possuem três e, indo mais além, incluem nessa categoria os pés degenerados”.

<sup>25</sup> A segunda etapa do EJA corresponde ao terceiro ano do EF (antiga segunda série).

para a análise dos mesmos tipos de dados em produções escritas de crianças. Essa decisão metodológica é tomada a fim de observar, em um primeiro momento, a relação entre os dados produzidos por adultos e crianças em alfabetização e, em um segundo momento, de verificar a adequação das categorias formuladas por Cunha na análise de textos de adultos em alfabetização.

Sobre o primeiro objetivo, Ferreira observou sete pontos em comum entre os dados de segmentação não-convencional por ela analisados e os dados de segmentação infantis descritos por Cunha (2004): 1) em termos quantitativos houve prevalência das grafias de hipossegmentação em comparação às de hipersegmentação e de híbridas; 2) ação de constituintes prosódicos mais altos da hierarquia prosódica (como a frase fonológica, a frase entoacional e o enunciado fonológico), especialmente nas hipossegmentações, e a ação de constituintes mais baixos (como o pé e a sílaba), em particular nas hipersegmentações; 3) a presença de elementos clíticos (palavra gramatical, na interpretação de Cunha (2004)) envolvidos tanto na hipo quanto na hipersegmentação; 4) ocorrências de processos fonológicos como ditongação e degeminação nos dados de hipossegmentação; 5) possível reconhecimento às formas de palavra escrita que levam a grafias não-convencionais do que se suporia ser palavras gramatical e lexical da língua; 6) não rompimento de constituintes basilares; e 7) ocorrência de flutuações na forma de grafar uma mesma forma de palavra em um mesmo texto.

Quanto ao propósito de avaliar a precisão da categoria “tipos de palavras” (CUNHA, 2004) na análise dos dados de segmentação não-convencional de EJA, Ferreira chega à conclusão de que essa categoria é plenamente pertinente para a compreensão dos processos de segmentação que ocorrem na escrita de adultos em alfabetização. Quantitativamente, os dados de Ferreira mostram, referente às hipossegmentações, que 71% do total foi de união não-convencional entre uma palavra gramatical e uma palavra fonológica; 18% de junção entre

duas palavras fonológicas; 7% de união entre uma palavra fonológica e uma palavra gramatical; e 4% de junção entre duas palavras gramaticais – cf. Ferreira (2011, p. 141). Também em função dos resultados quantitativos, no caso das hipersegmentações, 48% do total foi da presença do limite gráfico entre uma palavra gramatical e uma fonológica; 41% de separação entre duas palavras fonológicas; e 8% de separação entre duas possíveis palavras gramaticais. Por fim, quanto às ocorrências híbridas, a autora observa que, tanto na escrita inicial de adultos quanto na de crianças, esses tipos de dados são pouco frequentes.

Dando continuidade às reflexões sobre a noção de palavra, na próxima subseção a evocamos à luz dos estudos linguísticos.

## **1.2. Abordagem linguística da noção de palavra**

Como se sabe, a ortografia do PB faz uso de noções morfossintáticas/semânticas para definir a convenção de palavra escrita. Porém, por todo o discutido anteriormente, observamos que não são apenas os critérios convencionais que os escreventes levam em conta ao terem diante de si a tarefa de registrar (orto)graficamente os limites de palavras. Nesse sentido, defendemos, para este estudo, a relevância de noções *fonológicas* (na medida em que as palavras apresentam uma organização fônica que, pela observação de nossos dados, se mostra relevante para interpretar as grafias não-convencionais) e de noções *letradas* (que rondam os imaginários dos escreventes sobre o que é palavra de acordo com a convenção, a qual se constituiu sócio-historicamente por meio de práticas letradas/escritas). Assim, a seguir, apresentamos dois pontos de discussão que se destacam como “locais problemáticos” na definição de palavra; esses locais se constituem, em especial, no âmbito da fonologia em interface com informações morfossintáticas e, nos dados de hipersegmentação, pareceram, em

alguma medida, estarem registrados. Esclarecemos que outros aspectos como o morfossemântico e o letrado estão perpassados na análise dos dados.

### ***1.2.1. Analisando um primeiro “local problemático”: processos de formação de palavras***

Bisol (2004) propõe uma distinção entre palavras morfológica e fonológica. Segundo a autora:

A primeira [morfológica] compreende *palavras lexicais*, como nome, adjetivo e verbo, classes abertas, e *palavras funcionais* como preposição, conjunção e determinativos, classes fechadas. A segunda [fonológica] distingue palavras com acento e sem acento, respectivamente palavras fonológicas e clítics (BISOL, 2004, p. 59, grifos nosso).

A partir das definições da autora, abordamos a diferença entre as duas noções de palavras, observando a configuração de palavras compostas. Tomemos, pois, como exemplo a sentença: “O trabalho apresentado, na primeira *mesa-redonda*, foi um sucesso”. *Mesa-redonda*, nesse exemplo, ilustra a diferença entre as noções de palavras morfológica e fonológica, pois: (i) morfológicamente as duas bases, [mesa]β e [redonda]β, funcionam juntas na sentença, como uma só unidade lexical, portando um só significado; (ii) fonologicamente, no entanto, a palavra em análise configura-se como duas unidades independentes, [mesa]ω e [redonda]ω, já que cada uma é portadora de um acento primário, presente nas sílabas destacadas em *mesa* e *redonda*. Nesse aspecto, identifica-se uma diferença básica entre elas: “a primeira [palavra morfológica] está relacionada ao *significado*; a segunda [palavra fonológica], ao *ritmo*” (BISOL, 2004, 61, grifos nosso). Contudo, há casos em que os limites da palavra fonológica e morfológica coincidem, como ocorre em: “comprei uma *mesa*

*redonda* para minha casa”. Ao contrário daquela primeira sentença, nessa segunda, ambos os constituintes linguísticos são isomórficos, visto que o modo como *mesa* e *redonda* estão empregadas sintaticamente, leva as palavras a terem, em termos morfológicos, dois significados distintos, ou seja, são duas unidades lexicais independentes. Também, nos critérios fonológicos, as duas palavras são portadoras de dois acentos independentes. Existem, portanto, *duas* palavras morfológicas e *duas* fonológicas, nesse caso.

É um dos preceitos que regem a organização dos constituintes prosódicos<sup>26</sup> o uso de informações não-específicas, isto é, não-fonológicas, na construção de seus domínios. A palavra fonológica, em especial, é o primeiro constituinte prosódico a estabelecer interação com outro componente que não o fonológico, no caso, o morfológico (NESPOR E VOGEL, 1986, p. 109). Mas, como exemplificamos brevemente, essa interação não é sempre de natureza isomórfica; assim, os limites da palavra fonológica não têm a obrigação de coincidirem, sempre, com os limites do constituinte morfológico correspondente, muito embora, na maioria das vezes, o faça.

O funcionamento de determinadas palavras derivadas é outro lugar importante para se observar a interação flutuante entre morfologia e fonologia, no tocante à noção de palavra. De modo geral, palavras formadas por derivação apresentam um único acento primário e se constituem em palavras, fonológica e morfológica, cujos limites são idênticos. Por exemplo, na palavra *belo*, a sílaba tônica *be* é pronunciada com a vogal média-baixa, mas, com a inserção do prefixo *eza*, formando a palavra *beleza*, a sílaba *-be* sofre um levantamento da vogal (de vogal média-baixa para média-alta), sinalizando a mudança de acento dali para a sílaba seguinte (*beleza*).<sup>27</sup> É fundamental observar, que o deslocamento do acento presente na

<sup>26</sup> Essa afirmação está respaldada no modelo de Fonologia Prosódica de Nespor e Vogel (1986).

<sup>27</sup> Segundo Câmara Jr. (1970), em contexto tônico o PB faz a distinção vocálica de sete segmentos, a saber: /a/, /ɛ/, /ɔ/, /e/, /o/, /i/ e /u/. Em posição pretônica, esse número é reduzido para cinco (/a/, /e/, /o/, /i/ e /u/), ocorrendo, pois, a neutralização da oposição entre as vogais médias-baixas (/ɛ/ e /ɔ/) e vogais médias-altas (/e/ e /o/) no referido contexto (estamos considerando essa neutralização com base na variação que ocorre no estado de São Paulo).

sílaba *be*, em *belo*, para a sílaba *le*, na forma derivada *beleza*, não altera em nada o número de acentos primários (apenas um) de uma palavra para outra.

Em contrapartida, há palavras que preservam, em posição pretônica, a vogal média-baixa, como é o caso de *belamente*. Considerando que a regra geral de distribuição do acento no PB prediz que o acento primário só pode localizar-se em uma das três últimas sílabas de uma palavra (padrão proparoxítono, paroxítono, oxítono), no caso da palavra *belamente*, dada a sua configuração polissilábica, apenas três sílabas poderiam ser candidatas a receber o acento primário: *la*, *men* e *te* e, dentre essas possibilidades, o acento recai sobre a primeira sílaba do sufixo *-mente* (belamente). Seguindo essa regra, a palavra *belamente* passa a apresentar duas sílabas pretônicas e, quando produzida oralmente, recebe, na primeira delas, um *acento secundário* (COLLISCHONN, 1994), responsável pela alternância entre sílabas acentuadas e não-acentuadas, que garante o ritmo binário característico do PB. Nesse aspecto, a manutenção da vogal média-baixa, na base da palavra, é evidência da presença de mais de um acento (cf. QUADROS E SCHWINDT, 2008). Então, no que diz respeito ao estatuto fonológico, a palavra *belamente* é uma *palavra composta*, de modo tal que ela não apresenta coincidência com sua contraparte do componente morfológico, visto que, para a morfologia, ela se trata de uma só unidade formada por *derivação*, conforme destaca Boop da Silva (2010) ao analisar palavras com essa mesma configuração. Segundo a autora, palavras compostas, no âmbito da morfologia, são aquelas formadas por *duas bases*, como a propósito apresentamos anteriormente com o exemplo de *mesa-redonda*. Essa é uma informação importante, pois

a delimitação do que é um composto do ponto de vista morfológico tem sentido em função da diferenciação desse em relação aos compostos puramente fonológicos. Ou seja, compostos que apesar de apresentarem dois elementos providos de acento, são constituídos de apenas um elemento de força lexical. É o caso de palavras formadas por sufixos *-mente* (certamente) ou *-(z)inho* (potezinho), por exemplo, que portam acentos, mas não constituem palavras ou radicais autônomos na língua, sendo palavras derivadas e não compostas do ponto de vista fonológico (BOOP DA SILVA, 2010, p. 96).

O mesmo problema se estende aos prefixos acentuados, conforme analisou Schwindt (2001). Em seu estudo, Schwindt procura, dentre outros objetivos, categorizar os prefixos do PB quanto ao seu comportamento fonológico. Para tanto, submete-os à comparação por meio de três critérios, a saber: 1) presença/não-presença de acento; 2) oposição forma livre/forma presa; e 3) tipos de processos fonológicos a que estiveram propensos. Ao fim da comparação, Schwindt chega a duas classes de prefixos. A primeira é a dos *prefixos composicionais* (PCs)<sup>28</sup> – a qual é formada por prefixos que (i) têm configuração semelhante à de palavras fonológicas independentes, pois apresentam um acento primário (*pós-operatório*) e, no caso dos dissílabos, formam isoladamente um pé métrico (*autodidata*);<sup>29</sup> (ii) em determinados contextos, conseguem se instanciar como formas livres (*João reencontrou sua ex – ex-mulher*); (iii) sofrem processos fonológicos característicos de limites de palavras, como o sândi externo (*par[a]estatal ~ pa[es]tatal*) e a neutralização da átona final (*ant[e]projeto ~ ant[i]projeto*), indicando que, entre o prefixo e a base, há uma fronteira prosódica.

A segunda classe é a dos *prefixos legítimos* (PLs),<sup>30</sup> os quais, diferentemente dos prefixos do primeiro grupo, (i) demonstram comportamento de sílaba átona interior à palavra (*inesquecível*); (ii) não se sustentam sozinhos na sentença, igualando-se às formas presas (*João desfez as malas, antes que Maria \*des as dela*) e, por fim, (iii) são atingidos apenas por processos fonológicos que ocorrem no interior de palavras, como a neutralização da pretônica (*n[ε]ologismo ~ n[e]ologismo*), a harmonização vocálica (*r[e]fiz ~ r[i]fiz*) e a assimilação da nasal (*aN+alfabeto → a[n]alfabeto*).

A fim de ilustrar a importância de tais considerações para o presente estudo, antecipamos dados de hipersegmentação do cópuz que apresentam ancoragem nessa

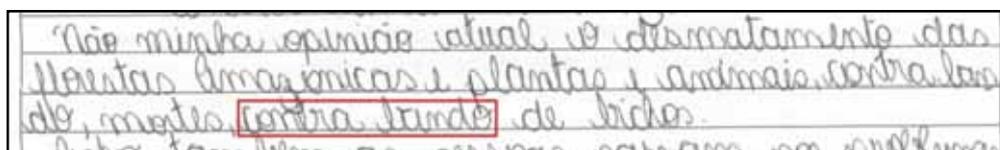
<sup>28</sup> A denominação *composicional* é dada já que, em virtude de a palavra fonológica não poder possuir mais que um acento, os prefixos acentuados ganham contornos semelhantes aos dos elementos de um composto.

<sup>29</sup> Todos os exemplos citados foram extraídos do próprio autor.

<sup>30</sup> *Legítimos*, pois, pela inacentuação, caracterizam-se como sílabas átonas que são afixadas à esquerda de uma base.

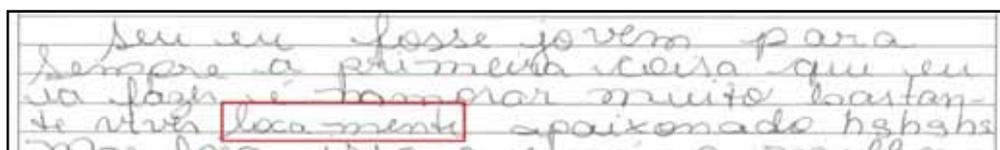
característica morfofonológica referente à formação de palavras. Vejamos, na Figura 4, um exemplo de hipersegmentação a partir de uma palavra formada por prefixo dissílabo, e na Figura 5, um exemplo de hipersegmentação motivada, como demonstraremos na seção de análise, do sufixo dissílabo acentuado. A análise que sustenta essa interpretação é feita na quarta Seção deste trabalho.

**Figura 4:** Dado “contra bando” (*contrabando*)



Fonte: Banco de Dados de Escrita do Ensino Fundamental  
(texto: Z11\_8D\_31F\_05)

**Figura 5:** Dado “loca mente” (*loucamente*)



Fonte: Banco de Dados de Escrita do Ensino Fundamental  
(texto: Z10\_7A\_31M\_01)

### 1.2.2. Analisando um segundo “local problemático”: clíticos fonológicos

Os clíticos fonológicos, de acordo com Bisol (2005, p. 164), apresentam as propriedades universais de atonicidade, comportamento de formas dependentes e o não pertencimento a uma classe morfológica específica. A primeira propriedade os diferencia de palavras fonológicas, as quais são sempre portadoras de acento primário; a segunda

propriedade decorre da primeira, pois, uma vez que os clíticos não apresentam acento, precisam se integrar/adjungir/anexar<sup>31</sup> a um hospedeiro que seja tônico para receberem status prosódico. Finalmente, a última propriedade refere-se ao fato de que todo clítico corresponde a diferentes classes morfológicas de palavra funcional (por exemplo: artigos – *o, a* –; preposições – *de, em* –; pronomes – *se, me* – conjunções – *e, ou*); contudo, nem sempre ocorre o inverso, vista a existência de palavras funcionais acentuadas, como *sobre*, por exemplo.<sup>32</sup>

Determinar o funcionamento prosódico dos clíticos, apesar de suas propriedades universais, não é simples, pois, a depender da língua em discussão, são notadas instabilidades em seu comportamento. Segundo Nespor e Vogel (1986, p. 145), os clíticos podem apresentar: (i) comportamento independente; (ii) comportamento de sílaba; e, ainda, (iii) comportamento específico. Essa natureza híbrida tem gerado dissensões teóricas quanto ao seu nível de inserção na estrutura fonológica. Assim, abarcando os citados tipos de comportamento, os clíticos podem ser integrados nos domínios da frase fonológica, da palavra fonológica e do grupo clítico<sup>33</sup>.

De acordo Nespor e Vogel (1986),<sup>34</sup> as diferentes sequências de clítico(s) + hospedeiro formalizam-se prosodicamente no nível do *grupo clítico*, uma vez que a análise de línguas como o grego e o turco, por exemplo, mostra a ocorrência de regras fonológicas na exata extensão das combinações de clítico(s) mais palavra. Os trabalhos de Bisol (2000, 2005) seguem esse mesmo posicionamento teórico e defendem, para o PB, a prosodização do clítico com a palavra adjacente no domínio do grupo clítico.

---

<sup>31</sup> Essas três possibilidades de união do clítico ao seu hospedeiro remete a formas de integração prosódica distintas, como mostraremos mais adiante.

<sup>32</sup> Há palavras funcionais dissílabas que, em funcionamento sintático, perdem informações segmentais, tornando-se clíticas (“para > pra > pa”, por exemplo). Nesses casos, a atonicidade não é estável, tal como ocorre com os clíticos monossílabos.

<sup>33</sup> Não é consensual a existência do grupo clítico como constituinte prosódico. No modelo prosódico formulado por Selkirk (1986), por exemplo, o nível que imediatamente domina a palavra fonológica é a frase fonológica, e não o grupo clítico. Porém, não entraremos nessa discussão, dado que, neste momento, ela extrapola os limites desta pesquisa. Relegamos essa questão a estudos futuros.

<sup>34</sup> Maiores detalhes sobre o modelo formulado por essas autoras estão desenvolvidos na Seção 3, deste estudo.

Antes de apresentarmos os argumentos defendidos por Bisol (2000, 2005), faremos uma observação: embora essa estudiosa, por um lado, assuma, junto com Nespor e Vogel (1986), a existência de um domínio prosódico “que se compõe de uma palavra fonológica e de um ou mais clíticos” (BISOL, 2000, p. 24), não avalia, por outro, como pertinente, o clítico ser considerado do modo como o é na análise de Nespor e Vogel: como uma palavra fonológica.<sup>35</sup> Bisol (2005) respalda seu posicionamento contrário nas três propriedades da palavra fonológica: (i) ser portadora de relações de proeminência; (ii) ser domínio de aplicação de regras lexicais;<sup>36</sup> e (iii) ser domínio de restrições fonotáticas. Em primeiro lugar, os clíticos não são portadores de proeminência relativa, nem internamente (em função da atonicidade inerente), nem externamente (já que presente em um constituinte maior, a proeminência recai sobre a palavra à qual o clítico está relacionado). Em segundo lugar, sequências de clítico+hospedeiro se mostram insensíveis para determinados processos que se aplicam no domínio da palavra fonológica (como a harmonia vocálica, conforme mostrou Brisolara (2008), para o PB). Em último lugar, no PB há a restrição fonotática de palavras não serem iniciadas por consoantes palatais; porém, na língua, há o clítico *lhe*. O não atendimento a essas propriedades indica que os clíticos, no PB, não podem receber o estatuto de palavra fonológica. Nesse sentido, Bisol (2000, 2005) defende que os clíticos são sílabas átonas que só adquirem *status* prosódico no nível pós-lexical, “formando um constituinte que, como a palavra fonológica, identifica-se pela presença de um só acento” (BISOL, 2005, p. 171).

Passando efetivamente à proposta de Bisol (2000, 2005), segundo a autora, há evidências de que, no PB, os clíticos são adjungidos “a uma palavra fonológica pronta, sem integrá-la, emergindo daí o primeiro constituinte pós-lexical” (BISOL, 2000, p. 19). Com

---

<sup>35</sup> Essa medida das autoras é para atender a *Strict Layer Hypothesis*, a qual exige que “uma unidade de um dado nível da hierarquia está exhaustivamente contida na unidade superordenada, da qual é uma parte” (Original: “A unit of a given level of the hierarchy is exhaustively contained in the superordinate unit of which it is a part” – NESPOR E VOGEL, 1986, p.7). Todas as traduções realizadas são de minha responsabilidade.

<sup>36</sup> Segundo Bisol (2005), a palavra fonológica não deve ser confundida com a palavra lexical, a qual está suscetível, também, a regras fonológicas pós-lexicais.

efeito, a defesa de seu posicionamento se faz em torno de três argumentos principais. O primeiro diz respeito aos clíticos serem insensíveis à restrição das três janelas, a qual impede que o acento de palavra recaia da quarta sílaba em diante (contando da direita para a esquerda). Por exemplo: *apresentávamos / apresentávamos-lhe* – “no entanto essas [últimas] formas são possíveis, apesar da restrição mencionada, indicando que sua configuração implica alguma sintaxe de frase, posterior à atribuição do acento” (BISOL, 2000, p. 24). O segundo argumento trata da característica de mobilidade posicional dos clíticos (particularmente os pronominais), uma característica que é típica de constituintes frasais: *Te dei um presente / Dei-te um presente*. Por fim, o terceiro é o de o clítico junto a seu hospedeiro estar propenso apenas à aplicação de regras pós-lexicais. A primeira regra, descrita pela autora, é a da neutralização da átona final: *s[e] conta ~ s[i] conta / conta-s[e] ~ conta-s[i]*. Segundo Bisol (2000, p.21), prefixos e sílabas pretônicas (que estão integradas a uma palavra) não sofrem a aplicação dessa regra: *preconceber ~ \*priconceber / lotação ~ \*lutação*. A segunda regra pós-lexical que atinge os clíticos é a nasalização, sobretudo, “a nasalização vocálica que não precisa de informação morfológica e que não tem exceções” (BISOL, 2005, p. 170). Os clíticos e as terminações átonas de palavras têm a nasal realizada como consoante ou glide: *homem (palavra): ˈomẽj~ˈomẽj~ˈomij~ˈome/ˈomi / em (clítico): ẽj~ẽj~ij*. A sonorização da fricativa, outra regra elencada por Bisol, ocorre dentro de palavras: *pa[z]mo* e fora desse domínio: *casa[z] bonitas*; afetando, também, os clíticos: *o[z] meninos* (BISOL, 2005, p. 170). Mais uma regra importante diz respeito à palatalização de /t/ e /d/ diante de /i/: (*[ˈtʃimi]*) – em palavras; (*[ˈtʃi vi]*) – em clíticos.

Mas a autora acredita serem as regras de sândi vocálico a principal evidência do grupo clítico. Ao observar o comportamento dos clíticos no processo de elisão, Bisol (2000, p. 26-27) observa, em relação à elisão da vogal /a/, que esse processo não se aplica no interior de uma palavra fonológica (*maometano \*mometano*), porém, se apresenta no interior de uma

frase fonológica (*casa escura* > *ca[zis]cura*) e de um grupo clítico (*venho pela estrada* > *venho pe[lis]trada*). Com base na elisão da vogal /e/, Bisol (2005) conclui que essa regra só se aplica na presença de um clítico, com exceção para quando diante do clítico está um elemento terminal designado (DTE), “elemento de proeminência relativa que projeta na grade uma posição forte, pronta para carregar o acento preponderante” (BISOL, 2005, p. 174), como ocorre em: *de amor \*damor*. A elisão de /e/ também não ocorre em frases fonológicas (*cidade antiga \*cidadantiga*), mas está restrita à ocorrência entre dois clíticos (*de+um: dum dia*) e entre um clítico e uma palavra funcional (*em outro dia > noutro dia*).

Assim, referente ao processo de elisão, a autora defende que o fato de esse

não se aplicar dentro de uma palavra é um indício de que o *clítico + hospedeiro* não se constitui uma palavra fonológica do tipo lexical, como *borboleta, parede, brinquedo*, embora o referido conjunto possua apenas um acento. Mas vale notar que, se for preciso fazer referência ao menor domínio de aplicação dessa regra, impõe-se a presença do grupo *clítico + hospedeiro*, pois, a partir daí é que a regra se estende para outros domínios. Isso é uma evidência de que o grupo clítico ocupa um espaço na literatura do português brasileiro como uma entidade prosódica, independentemente de possuir um *locus* específico na escala prosódica (BISOL, 2005, p. 173).

Simioni (2008), diferentemente de Bisol (2000, 2005), adota o posicionamento de que o clítico é anexado diretamente a um nó de frase fonológica, uma vez que, segundo a pesquisadora, não há processos, no PB, que confirmem a existência do grupo clítico. Na construção de sua argumentação, Simioni (2008), com base em Selkirk (2004 *apud* SIMIONI, 2008, p. 440), apresenta e discute quatro tipos de possibilidades de inserção prosódica dos clíticos, as quais são: **a**) (cl (lex)ω)φ; **b**) ((cl (lex)ω)ω)φ; **c**) ((cl lex)ω)φ; e **d**) ((cl)ω (lex)ω)φ. A defesa da autora é pela possibilidade de integração representada em (a), que é a do clítico anexado a uma frase fonológica; por consequência, as demais possibilidades são refutadas para o PB, já que, em (b), o clítico é tratado como afixal e estaria ligado à palavra hospedeira por adjunção. No entanto, a partir de processos fonológicos que atingem determinados

prefixos legítimos do PB (cf. SCHWINDT, 2001), a autora considera que esses apresentam, por exemplo, processos de ressilabação que não ocorrem na combinação de clítico+palavra. Para ilustração, prefixos terminados em segmento nasal seguidos de palavras iniciadas por vogal, passam a ter a sua consoante nasal como o ataque da próxima sílaba da palavra:  $(an+(aeróbico)\omega)\omega > (a[na]eróbico)\omega)\omega$ ; o mesmo fato não se aplica aos clíticos:  $(sem(esperança)\omega)\phi * (se([ne]sperança)\omega)\phi$  (cf. SIMIONI, 2008, p. 438). Diante dessas comparações, a autora defende que “essa diferença é explicada pelo fato de que prefixos se adjungem a uma palavra prosódica enquanto clíticos se anexam a uma frase fonológica” (SIMIONI, 2008, p. 438). A configuração em (c) representa a prosodização do clítico como integrado à palavra fonológica e, nesse sentido, o clítico se comporta, quando proclítico, como uma sílaba pretônica e, quando enclítico, como uma sílaba postônica. Para refutar também essa possibilidade de prosodização, Simioni mostra, em relação aos proclíticos, o fato de que esses sofrem processos característicos de sílabas átonas finais de palavra, como é o caso da neutralização da átona final, processo também discutido anteriormente por Bisol (2000, 2005), para reforçar a ideia de que o clítico possui certa independência em relação ao hospedeiro. Esse processo de neutralização, no entanto, não é esperado em sílabas pretônicas: *elefante*  $*[i]l[i]fante$ , com exceção para os casos de harmonia vocálica (*menino* ~  $m[i]nino$  – SIMIONI, 2008, p. 435). A respeito dos enclíticos, a autora observa que, quando uma palavra terminada em vogal média (/e/ e /o/) é seguida de um enclítico, aquela tem a sua vogal final alçada:  $(ped[e])\omega -se \sim (ped[i])\omega-se$ . Para Simioni, esses processos que acontecem em contexto final indiciam a existência de uma fronteira entre o clítico e o hospedeiro e que, portanto, o clítico não pode ser considerado uma sílaba integrada à palavra fonológica. Finalmente, a representação em (d) mostra um clítico com funcionamento de palavra fonológica. De acordo com Simioni, os clíticos do PB não atendem às propriedades que

caracterizam uma palavra fonológica e, dessa maneira, ela descarta a configuração (d) como a representação da prosodização do clítico na língua em análise.

Para formalizar a sua proposta, fundamentada na representação (a), Simioni se vale de um arcabouço teórico baseado em restrições, sobretudo, àquelas de Dominância Prosódica (Layeredness; Headedness; Exaustividade; Não-recursividade)<sup>37</sup> e de Alinhamento de Fronteiras.<sup>38</sup> Segundo Simioni, a configuração a qual o clítico se anexa diretamente a uma frase fonológica é a estrutura que, além de se mostrar a mais coerente para o PB a partir da observação de processos fonológicos, é, também, a possibilidade que viola menos número de restrições, ou seja, não atende apenas a *Exaustividade* (em virtude de uma frase fonológica dominar diretamente uma sílaba), e não há violação de Alinhamento. Trata-se, pois, da representação ótima do clítico na estrutura prosódica, de acordo com Simioni.

Por fim, a autora conclui que o grupo clítico não tem relevância para explicar fatos fonológicos do PB, uma vez que fatos como a elisão da vogal /e/ (defendida por Bisol (2005) como uma regra exclusiva do grupo clítico), podem ser explicados com base em restrições de alinhamento e, desse modo, não precisa de um constituinte prosódico específico para isso, contrariando a análise de Bisol (2000, 2005) para essa mesma língua.

Com base em diferentes abordagens teóricas, Bisol (2000, 2005) e Simioni (2008) fazem propostas distintas de representação da estrutura prosódica do clítico em PB. Bisol, por um lado, se alinha com a defesa de que existem fatos do PB que só se explicam na sequência exata de um clítico mais o seu hospedeiro; por isso, a necessidade em se considerar um

<sup>37</sup> *Layeredness*: Nenhum  $C^i$  domina um  $C^j$ ,  $j > i$ , por exemplo, “nenhuma sílaba domina um pé”. *Headedness*: Qualquer  $C^i$  deve dominar um  $C^{i-1}$  (exceto se  $C^i = \sigma$ ), por exemplo, “uma palavra prosódica deve dominar um pé”. *Exaustividade*: Nenhum  $C^i$  domina imediatamente um constituinte  $C^j$ ,  $j < i-1$ , por exemplo, “nenhuma palavra prosódica domina imediatamente uma sílaba”. *Não-recursividade*: Nenhum  $C^i$  domina  $C^j$ ,  $j = i$ , por exemplo, “nenhum pé domina um pé” (SELKIRK, 2004, p. 466-467 *apud* SIMIONI, 2008, p. 439).

<sup>38</sup> Restrições de alinhamento da palavra morfológica: *Align (Lex, E;  $\sigma$ , E)* – A fronteira esquerda de uma palavra lexical deve coincidir com a fronteira esquerda de uma palavra prosódica; *Align (Lex, D;  $\sigma$ , D)* – A fronteira direita de uma palavra lexical deve coincidir com a fronteira direita de uma palavra prosódica. Restrições de alinhamento da palavra prosódica: *Align ( $\sigma$ , E; Lex, E)* – A fronteira esquerda de uma palavra prosódica deve coincidir com a fronteira esquerda de uma palavra lexical; *Align ( $\sigma$ , D; Lex, D)* – A fronteira direita de uma palavra prosódica deve coincidir com a fronteira direita de uma palavra lexical (SELKIRK, 2004, p. 468-469 *apud* SIMIONI, 2008, p. 441).

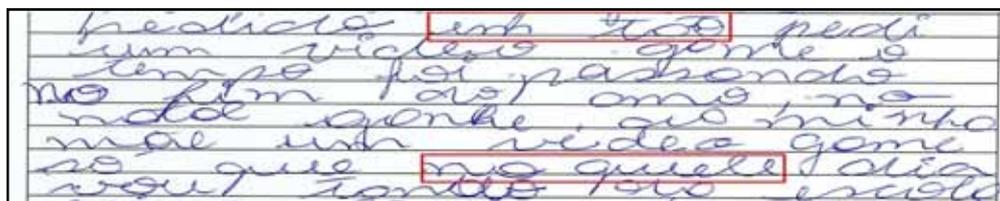
domínio como o grupo clítico. Por outro lado, Simioni apoia-se na ideia de que, por meio de restrições, o funcionamento dos clíticos pode ser englobado em um domínio prosódico superior ao da palavra (a frase fonológica), sem ter a necessidade de mais um constituinte prosódico exclusivo que abrigue o clítico e seu hospedeiro.

Seja o grupo clítico seja a frase fonológica o domínio em que o clítico é anexado à estrutura prosódica, entendemos que a formalização do comportamento prosódico dos clíticos, nas duas propostas apresentadas, parecem se encaminhar em uma mesma direção: a de que, no PB, os clíticos tendem a ser mais independentes com relação à palavra hospedeira, já que estão sempre ganhando *status* prosódico nos domínios que já estão em interação com o componente sintático.<sup>39</sup> Por hipótese, acreditamos que as fronteiras de palavras que se marcam para além do esperado na escrita poderiam ser um indício importante dessa autonomia maior que os elementos clíticos apresentam no PB. A esse tema retomaremos quando da análise dos dados. Por ora, somente expomos dois registros escritos que estariam, segundo interpretação que faremos mais à frente, em diálogo com essa discussão:

---

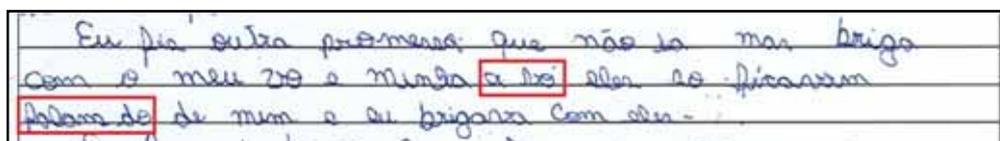
<sup>39</sup> Nota-se que essas posturas notadas para o PB são diferentes de outras propostas teóricas correntes na literatura sobre integração prosódica de clíticos, a exemplo da que foi desenvolvida por Vigário (2001) sobre os clíticos do PE. Afirmando haver assimetria na inserção prosódica de proclíticos e enclíticos, Vigário argumenta que os *proclíticos* unem-se ao hospedeiro por *adjunção* (nesse caso, seriam duas unidades), já os *enclíticos* são *incorporados* junto à palavra fonológica precedente (nesse outro caso, clítico e hospedeiro seriam uma só unidade). Para chegar a tal avaliação, a autora observou que regras fonológicas pós-lexicais que se aplicam na borda direita final das palavras, também se aplicam na sequência de palavra + enclítico, indiciando que o clítico, nessa posição, tem comportamento idêntico ao de segmento interno de palavra. Já a presença de um proclítico foi contexto favorável para a ocorrência de regras pós-lexicais na margem esquerda inicial da palavra, sinalizando, segundo a autora, de que entre o proclítico e o hospedeiro há a presença de uma fronteira prosódica: nesse sentido, o clítico não pode ser considerado parte da palavra. Enfim, de acordo com Vigário, enclíticos e seus respectivos hospedeiros são dominados por uma só palavra fonológica pós-lexicalmente, e os proclíticos ligam-se, por adjunção, às palavras prosódicas que o seguem, também, no nível pós-lexical.

**Figura 6:** Dados “em tão” (*então*) e “na quele” (*naquele*)



Fonte: Banco de Dados de Escrita do Ensino Fundamental  
(texto: Z08\_5C\_12M\_03)

**Figura 7:** Dados “a vó” (*avó*) e “falam do” (*falando*)



Fonte: Banco de Dados de Escrita do Ensino Fundamental  
(texto: Z09\_6B\_04M\_01)

### **Resumo da Seção**

No decorrer desta Seção, procuramos delinear as reflexões que compreendem a primeira parte dos subsídios teóricos de nossa pesquisa. Inicialmente lançamos mão, para tratar da noção de palavra, de trabalhos sobre as segmentações não-convencionais. Os trabalhos considerados analisaram dados produzidos por escreventes em diferentes momentos do processo de escolarização formal, mostrando que a delimitação das palavras escritas não é um desafio característico apenas da escrita inicial de crianças e adultos, mas atinge, também, escreventes com mais anos de escolarização. Essas constatações auxiliam a reforçar nossa hipótese de ser, a palavra escrita, uma noção complexa, a qual põe em jogo a decisão entre diferentes informações linguísticas.

Quanto ao estatuto dos registros das fronteiras de palavra, é importante destacarmos que, em todos os estudos, notamos um posicionamento comum que os consideraram como ocorrências representativas do funcionamento da linguagem e da relação dessa com os seus sujeitos. Assim, os dados de hipo e hipersegmentação deixam de ter relevância restrita a interesses pedagógicos de ensino e aprendizagem (enquanto um problema de alfabetização), e passam “a ser vistos como preciosa fonte de indícios sobre a natureza do trabalho realizado pelas crianças [e não só por elas, a nosso ver] com a linguagem” (ABAURRE, 1998, p. 208). Ainda dentro desse panorama, conseguimos identificar formas diferentes de análise dos dados (mas que convergiram entre si em alguns pontos), e optamos por adotar, para esta pesquisa, a que definimos por tendência fonológico-letrada, por se referir a um modo de análise que, à semelhança de nossos propósitos, identifica na estrutura dos dados aspectos da constituição heterogênea da escrita (CORRÊA, 2004).

Em seguida, abordamos discussões envolvendo a noção de palavra em função dos aspectos linguísticos presentes nos enunciados falados. A necessidade de perfilarmos tais reflexões sobre a noção de palavra fonológica com as hipersegmentações justificou-se pela hipótese de que as dificuldades em se definir os limites de palavra, quando considerados os enunciados falados, são refletidas, em alguma medida, nos enunciados escritos, quando são propostas as fronteiras não-convencionais de palavra. A esse respeito, mostramos, por exemplo, o estatuto morfossintático e fonológico de palavras derivadas por afixos acentuados, em dados como “contra bando” (para *contrabando*) e de elementos clíticos, em grafias como “na quele” (para *naquele*).

Na próxima Seção, damos continuidade à discussão dos pressupostos teóricos, abordando a concepção de escrita proposta por Corrêa (1997, 2004).

## SEÇÃO 2: Sobre a concepção de escrita

---

### *Apresentação da Seção*

Nesta Seção, tratamos da última parte dos fundamentos teóricos, explicitando a forma como as hipersegmentações são observadas em relação ao seu funcionamento no interior do texto escrito. Adiantamos que, por compreendermos esses registros não-convencionais das fronteiras de palavras como pistas do encontro indivisível entre enunciados falados e escritos e práticas sociais orais e letradas (CORRÊA, 1997, 2004), defendemos a importância de as noções de *fala* e de *escrita* serem claramente definidas, bem como a natureza da relação entre ambas. A justificativa para a ênfase nessas noções está baseada no fato de que, especialmente, certa concepção da relação fala e escrita é tomada como um dos principais argumentos utilizados para explicar, dentre outras questões, a motivação para a ocorrência de grafias não-convencionais, como as de segmentação em palavras, por exemplo.

Nessa linha argumentativa estão desenvolvidos os estudos de Zorzi (1997), Koch (1997) e Corrêa (1997, 2004). Ressaltamos, entretanto, a opção desses pesquisadores por diferentes posicionamentos teóricos sobre a relação fala/escrita. A fim de abordarmos essas diferenças, na primeira subseção (“*Formas de tratamento das relações fala/escrita*”) discutimos, a partir dos trabalhos de Zorzi (1997) e Koch (1997), a ideia de uma suposta interferência da fala na escrita, a qual está subjacente às análises desenvolvidas por esses autores acerca do que, para eles, se classificariam como diferentes erros ortográficos (incluindo, nessa classificação, as grafias de segmentação não-convencional) Adiantamos que

Zorzi (1997) e Koch (1997), ao conceberem a relação entre fala e escrita por meio da perspectiva da interferência, assumem, por consequência, fala e escrita como modalidades linguísticas homogêneas, no sentido de que as características que constituiriam cada uma delas seriam, por princípio, puras e opostas a da outra modalidade.

Neste trabalho, diferentemente, defendemos, para a análise das hipersegmentações, outra visão teórica sobre a relação fala/escrita, a qual se ancora na tese do modo heterogêneo de constituição da escrita, formulada por Corrêa (1997, 2004). Quando discutimos e comparamos a concepção de Corrêa com a proposta defendida por Zorzi, podemos nos situar contrários em relação a uma vertente de estudos fonoaudiológicos que buscam, na classificação e quantificação dos chamados erros de ortografia, demonstrações do que seria parte de um processo “normal” de aprendizagem da escrita, e do que nesses já seriam evidências de patologias da aprendizagem. Já, a partir da contraposição do trabalho de Corrêa com a abordagem em que se insere Koch, lançamos luz sobre posicionamentos distintos no interior da Linguística, concernente à relação fala/escrita. Nesse aspecto, no trabalho dessa autora, localizamos um tipo análise que, ao perpassar, por exemplo, o nosso objeto de estudo, se distancia, fortemente, daquela proposta desenvolvida na mesma área do conhecimento e com a qual nos filiamos.

Somam-se, aos nossos esclarecimentos, o fato de não nos aproximarmos das análises desenvolvidas por Koch e Zorzi, pois, embora nesses trabalhos seja possível identificar discussões que tangenciam a relação de dados de escrita com características da fala, esses autores não têm como preocupação a compreensão a investigação da relação fala/escrita, como também não adotam, em termos teóricos, nenhuma filiação com discussões aprofundadas a esse respeito, as quais tenham como ponto de partida, a definição dos conceitos de *fala* e de *escrita*. Assim, por negligenciarem aspectos dessa natureza,

compreendemos que as análises de Zorzi e Koch deixam de contemplar a complexidade a qual vemos constituir as grafias de hipersegmentação.

São, pois, pelas razões explicitadas, que, nesta dissertação, nos fundamentamos na proposta teórica desenvolvida por Corrêa (1997, 2004), a qual está sustentada sob a ideia de não oposição entre fatos linguísticos e práticas sociais (a ser apresentada na subseção “*A proposta do modo heterogêneo de constituição da escrita*”). De acordo com Corrêa, o texto escrito registra a circulação do escrevente pelas práticas de linguagem, registro que é observável por meio de marcas linguísticas (nesta pesquisa, as hipersegmentações), as quais permitem recuperar possíveis vínculos com aspectos linguísticos característicos de cada uma delas.

Nesse sentido, trazemos para discussão as pesquisas de Zorzi e Koch não só por essas ilustrarem tipos de análise muito frequentes acerca de objetos de investigação como o deste estudo (especialmente em diferentes correntes da Linguística, onde se situa este trabalho, bem como em um tipo de abordagem fonoaudiológica comumente adotada para o tratamento pedagógico da escrita), classificados apenas como erros ou desvios da chamada escrita padrão, mas, também, e principalmente, por marcarmos o ganho explicativo em adotar uma abordagem teórica dedicada ao tratamento exclusivo das questões que tocam a relação fala/escrita, enquanto um campo específico de investigação. A partir dessa perspectiva, então, privilegamos a observação de processos/hipóteses que mobilizam as grafias não-convencionais e que deixam marcas da “trajetória desses escreventes em sua inserção na escrita convencional” (CAPRISTANO, 2010, p. 175).

## **2.1. Formas de tratamento da relação fala/escrita**

A fim de compreender como crianças dos primeiros quatro anos da escolarização se apropriariam do sistema ortográfico, Zorzi (1997) propõe um estudo em que analisa diferentes tipos de erros ortográficos bastante característicos da etapa investigada. O autor opta por analisar a escrita de alunos de escolas particulares, por compreender que nesse tipo de instituição são proporcionadas condições privilegiadas de ensino e aprendizagem e, também, pelo fato de as crianças investigadas fazerem parte de famílias com boa situação econômica, as quais, “em geral, atribuem valores positivos ao aprendizado escolar e que, comumente, têm familiaridade com o tipo de linguagem que a escola exige e valoriza” (ZORZI, 1997, p. 21). Desse modo, “a escolha destas escolas teve a pretensão de evitar que problemas relativos ao desenvolvimento de uma nova linguagem pudessem ser atribuídos exclusivamente a fatores sociais, econômicos ou pedagógicos” (ZORZI, 1997, p. 21).

Para o levantamento dos erros, Zorzi desenvolve duas atividades distintas de coleta de dados. A primeira é a de aplicação de *ditados* de palavras, frases e textos (esses últimos selecionados de livros infantis). Já a segunda consiste na elaboração de *textos escritos* com base em temas previamente indicados. Segundo o autor, algumas crianças escrevem narrativas muito curtas e, nesse aspecto, o resultado dos ditados permite a construção de um *cópus* mais extenso, com a possibilidade de ampliação da análise dos conhecimentos ortográficos das crianças.

Os erros identificados são organizados e classificados pelo autor em dez categorias (cf. ZORZI, 1997, p. 25-34):<sup>40</sup> 1) *alterações ou erros devido à possibilidade de representações múltiplas* (casos em que a ortografia dá possibilidade de diferentes representações gráficas para um mesmo som, por exemplo: fonema /s/ representado pelos grafemas <s>, <ss>, <c>, <ç>); 2) *alterações ortográficas decorrentes de apoio na oralidade* (casos em que há desencontro entre a pronúncia da palavra e a forma convencional de grafá-la, por exemplo:

---

<sup>40</sup> Exemplos retirados do próprio autor.

palpite – “paupite”); 3) *omissões de letras* (casos em que a palavra é registrada sem alguma letra, por exemplo: comprou – “compou”); 4) *alterações caracterizadas por junção ou separação não convencional das palavras* (casos de hipossegmentação e hipersegmentação, por exemplo: se perder – “siperder” e quatrocentos – “quatro sentos”); 5) *alterações decorrentes de confusão entre terminações am e ão* (casos em que a ortografia prevê *am*, mas a grafia foi *ão* e vice-versa, por exemplo: falar**am** – “falar**ão**”); 6) *generalização de regras* (casos em que a criança utiliza uma regra ortográfica a qual já conhece em outros contextos semelhantes, por exemplo: cinema – “cenema”); 7) *alterações caracterizadas por substituições envolvendo a grafia de fonemas surdos e sonoros* (casos de trocas como em: pagando – “paganto”); 8) *acréscimo de letras* (casos em que a palavra grafada pela criança apresenta uma letra a mais do que a ortografia exige, por exemplo: machucar – “manchugar”); 9) *letras parecidas* (casos em que a troca de letras pode ser explicada em função da semelhança que se pode observar no traçado de algumas delas, por exemplo: tinha – “timha”); e 10) *inversão de letras* (casos que a letra está em posição invertida dentro da palavra, por exemplo: fraquinho – “farquinho”).

A partir dessas dez categorias, Zorzi chega à conclusão de que, embora, em um primeiro momento, os erros classificados como “escrita com apoio na oralidade”, “*dificuldades em separar palavras*” (as segmentações não-convencionais!) e “confusão entre as terminações *am* e *ão*” tenham sido analisados separadamente, esses erros convergem em relação a sua motivação, pois, “nestes três casos, há uma *forte influência de padrões de linguagem oral* determinando a ocorrências das alterações” (ZORZI, 1997, p. 87, grifos nosso). Para o autor, esse resultado pode ser bastante compreensível, já que “não é de imediato que a criança compreenderá a variação possível entre formas de falar e formas de escrever, diferenciando as características de cada uma destas duas línguas” (ZORZI, 1997, p. 10). Assim, na medida em que fala e escrita constituem duas diferentes línguas, a criança, por

dominar apenas as características da língua falada, tenta, nos seus primeiros contatos com a escrita, transpô-las para seus textos. Nesse sentido, “a direção da escrita [vai] sendo fortemente influenciada por padrões acústico-articulatórios, [ou seja] por uma referência fonética e não ortográfica” (ZORZI, 1997, p. 88).<sup>41</sup>

De acordo com Zorzi (1997, p. 10), as relações entre oralidade e escrita

podem ser discutidas de um ponto de vista *da oralidade para a escrita ou da escrita para a oralidade*. Tais relações têm se evidenciado quando a oralidade é entendida como uma primeira língua, ou como descreve Kato (1986), a “fala 1” que, inicialmente, serve de apoio para o primeiro nível de escrita, a “escrita 1”. Portanto, *em suas fases iniciais, a escrita sofre grande influência da oralidade*. Porém, na medida em que *a escrita vai se tornando mais independente da oralidade e adquirindo as características formais que a definem como modelo de língua padrão*, “escrita nível 2”, pode produzir transformações na própria oralidade que, assim, atingiria outro nível, “fala 2”. Como que num processo inverso, *a escrita torna-se “apoio” para a oralidade* (destaques nosso).

Dos destaques acima, observamos que a relação oralidade/escrita, na qual se fundamenta Zorzi, está baseada em uma concepção de interferência mútua, que se constrói de uma para outra (isto é, da fala para a escrita, da escrita para a fala). Por aprendermos a falar primeiro do que aprendemos a escrever, nos primeiros contatos com a escrita, ocorreria uma interferência da fala na escrita, conforme o autor conclui em sua análise. Com o passar do tempo, a escrita começaria a ganhar independência em relação à fala, à medida que as características convencionais passariam a ser aprendidas e sistematizadas pelos aprendizes. A escrita, ao fim, se consolidaria como a língua padrão, influenciando, dessa vez, a fala. Compreendemos que, para o autor, prevalece a ideia de separação entre fala e escrita, bem como o fato de o registro escrito (mas não qualquer um!) ser privilegiado em relação aos

---

<sup>41</sup> Vale lembrar a postura de Abaurre (1988a), já discutida neste trabalho, para quem a afirmação de que a criança se baseia apenas na fala para solucionar grande parte dos problemas que a escrita lhe impõe é ingênua, pois, segundo a autora: “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 de início tal complexidade. Embora as primeiras produções espontâneas sejam, em um primeiro momento, bastante idiossincráticas, é possível identificar desde cedo a incorporação de aspectos convencionais, o que só pode ser explicado pelo forte apelo social das atividades de leitura e escrita” (p. 136-137).

registros de fala, bem como a confusão de que o registro escrito está sempre marcado pela formalidade da variedade culta da escrita..

Não é difícil perceber que à escrita inicial – menos calcada nos preceitos convencionais da escrita padrão – está reservado o lugar de manifestação de características da fala, uma vez que, ao alcançar os critérios que regem a convenção ortográfica, os textos escritos estariam “blindados” a qualquer interferência que a fala poderia causar neles. A partir dessas constatações, a escolarização desempenharia um papel essencial na transição de interferência *da fala na escrita* para interferência *da escrita na fala*, pois, como salienta Capristano (2010, p. 178), com o processo escolar “os enunciados escritos pelos aprendizes passariam a ganhar características da escrita convencionalmente aceita, escrita que seria, supostamente, homogênea, sem ‘intervenção’ e/ou ‘interferência’ de fatos característicos de enunciados falados”.

Koch (1997) também assume a concepção de interferência da fala na escrita. Segundo a estudiosa, ao entrar para a escola, a criança já tem construído o modelo de texto falado, por meio do qual elas vão basear suas primeiras produções escritas e demais textos por vários outros anos. De posse do que atribui como as diferenças do texto falado em relação ao texto escrito (cf. KOCH, 1997, p. 33), a autora elenca os principais pontos de interferência encontrados nos textos infantis, dentre os quais, identificam-se: 1) *mecanismos de referência*, que correspondem à utilização indevida de recursos anafóricos e catafóricos; 2) *repetições*, casos comuns no texto falado, “podendo mesmo ser considerada um dos mecanismos organizadores dessa modalidade oral” (KOCH, 1997, p. 35); 3) *organização textual com elementos continuadores típicos da fala*, diz respeito ao uso frequente de expressões como *aí, daí, então, daí então*; 4) *ausência de pontuação para ligar ideias*, casos de textos escritos sem qualquer sinal de pontuação; 5) *discurso citado*, falta de marcação de discursos diretos; 6) *segmentação gráfica em palavras*; 7) *grafias a partir da pronúncia da palavra*, as palavras

passam a ser escritas a partir do modo em que se realizam foneticamente; 8) *correções comuns nos textos falados*, casos que se observa o registro gráfico de hesitações da fala.

No caso das segmentações não-convencionais, a autora explica que hipossegmentações, como “vamologo” (*vamos logo*), são resultados da apreensão das crianças de vocábulos fonológicos (ver CÂMARA JR., 1970). Porém, nas hipersegmentações, como “na que-la” (*naquela*), as crianças buscam efetuar uma segmentação convencional, por isso é que elas estariam, segundo Koch (1997, p.37), “caindo no extremo oposto, isto é, ‘picando’ demais a palavra”. Percebe-se que a explicação da autora não sustenta a sua própria interpretação calcada no argumento da interferência do texto falado no texto escrito, em razão de ser o único modelo de texto que a criança conhece ao ingressar na escola, visto que, como a própria autora reconhece, em determinados momentos as crianças estão tentando realizar uma segmentação adequada com os padrões ortográficos. Considerando que o uso de recursos gráficos a fim de dividir os enunciados em palavras é uma exigência exclusiva da escrita, é bastante compreensível que, quando os aprendizes arriscam segmentar as palavras convencionalmente, eles *não* estão lidando *apenas* com informações que organizariam o texto falado. As crianças, conforme mostram outros pesquisadores do tema, estão inseridas, também, em práticas letradas, por isso mobilizam conjuntamente informações dessas práticas para comporem os textos que escrevem. Frente ao que foi apresentado dos trabalhos de Zorzi (1997) e Koch (1997), fazemos uma relação entre a concepção de escrita em que se baseiam esses autores e o conceito cunhado por Street (1984) como *modelo autônomo de letramento*. Tfouni (1995) localiza o que esse estudioso definiu como modelo de letramento autônomo como uma retomada moderna da *teoria da grande divisa*, a qual tem por princípio básico a tese de que

em uma sociedade letrada, haveria separação radical entre usos orais e usos escritos da língua, caracterizando dois tipos específicos de discurso: o discurso oral e o escrito. No primeiro, teríamos contextualização, informalidade, causalidade,

envolvimento interpessoal e um tipo de raciocínio emocional e ambíguo; no segundo, teríamos perda do contexto imediato, estilo formal, evitação ou inexistência de envolvimento interpessoal, além de um tipo de raciocínio abstrato, descontextualizado e lógico (TFOUNI, 1995, p. 47-48).

Assim, no modelo autônomo, a escrita se concebe como uma *tecnologia* pronta em si mesma, cujo “processo de interpretação estaria determinado pelo funcionamento lógico interno ao texto escrito” (KLEIMAN, 1995, p. 22), sem qualquer relação com o evento real de produção de linguagem. Por meio dessa ideia, esse modelo desenvolve-se em torno do raciocínio de que basta a aquisição do sistema convencional de escrita (ou seja, o código) para participar das práticas sociais em que circulam a escrita. Além do mais, pela apropriação da escrita tem-se a abstração do pensamento e o desenvolvimento cognitivo, os quais, por consequência, trazem o progresso e a mobilidade social dos indivíduos.<sup>42</sup> Nessa abordagem, a escrita é um *valor* social necessário a todos. Para Street (1984), estudos que se desenvolvem no escopo do ideal de letramento autônomo desconsideram a não neutralidade da escrita e, portanto, baseiam suas análises dos fenômenos em argumentos infundados.

De outro ponto de vista, Marcuschi (2001) fez uma crítica à perspectiva de separação radical da fala e da escrita, sobre a qual ele atribui um caráter rigidamente formal capaz, apenas, de alçar “bons resultados na descrição estritamente empírica, [enquanto que] manifesta enorme insensibilidade para os fenômenos dialógicos e discursivos” (MARCUSCHI, 2001, p. 28), isto é, não se sustenta na análise de fatos linguísticos concretos. Buscando avançar em relação a essa abordagem, Marcuschi defende que fala e escrita são modalidades linguísticas que se realizam em práticas de linguagem (orais e letradas) por meio

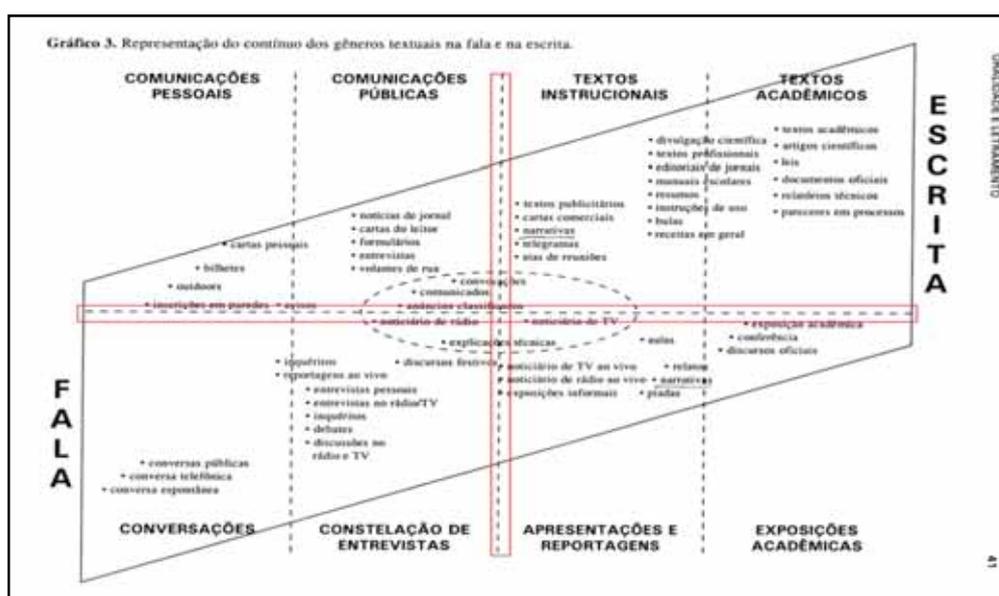
---

<sup>42</sup> A tese da relação entre aquisição da escrita e desenvolvimento cognitivo ganhou força, sobretudo, em investigações de orientação etnográfica, as quais buscavam comparar as estratégias de resolução de problemas por parte de grupos utilitários da escrita e grupos ágrafos (cf., a esse respeito, GOODY (1988/[1977])). Argumenta-se, ao realizar tal comparação, que os grupos que conhecem a escrita têm, mais habilidade em desenvolver as tarefas propostas. Street (1984), questionando esse argumento, defende ser impossível comensurar, com precisão, as performances lógicas de grupos de indivíduos, pois, segundo ele, as estratégias de resolução de problemas estão intimamente relacionadas com questões socioculturais e não com condições cognitivas.

de gêneros textuais. Por compreender a existência de práticas sociais na realização dos usos da linguagem, o autor chama a atenção para o fato de que a relação entre modalidades linguísticas e práticas sociais não são iguais e precisam ser distinguidas. Assim, na postura de Marcuschi, estudos que venham a se dedicar à reflexão sobre as formas de realização da linguagem, devem considerar, de um lado, oralidade e letramento (prática social) e, de outro, fala e escrita (modalidade da língua). Mais adiante mostramos o questionamento de tal proposição, em que modalidades linguísticas e práticas sociais são aproximadas, de modo a se constituírem como indivisíveis.

Endossando um tratamento mais produtivo à relação entre as modalidades linguísticas, Marcuschi (2001, p. 37) sugere que *“as diferenças entre fala e escrita se dão dentro do continuum tipológico das práticas sociais de produção textual e não na relação dicotômica de dois pólos opostos”* (destaques do autor). A posição do autor é elucidada no gráfico seguinte:

**Figura 8:** O contínuo de gêneros textuais, segundo Marcuschi (2001)



Fonte: MARCUSCHI, 2001, p. 41, modificado

O esquema acima explicita o contínuo de gêneros textuais, o qual se alicerça em *duas diferenciações*, representadas pelas linhas pontilhadas centrais (marcadas em vermelho) que o dividem. A linha horizontal propõe a distinção entre *meio de produção* (sonoro – gráfico) e o traço vertical apresenta a separação pela *concepção discursiva* (oral – escrita). Com base nesse princípio organizador, fala e escrita são localizadas nos extremos do contínuo e os gêneros textuais são distribuídos gradativamente, o que permite a observação de gêneros, tanto falados quanto escritos, mais homogêneos, bem como a identificação de gêneros mistos dessas modalidades (centrados em um círculo no meio do contínuo). Diante dessa disposição, que ilustra as diferenças e semelhanças da relação entre fala e escrita, Marcuschi reafirma sua ideia e sustenta que “tanto a fala como a escrita apresentam um *continuum de variações*, ou seja, a fala varia e a escrita varia. Assim, a comparação deve tomar como critério básico de análise uma relação fundada no *continuum* dos gêneros textuais para evitar as dicotomias estritas” (MARCUSCHI, 2001, p. 42, grifos do original).

Apesar de reconhecermos o ganho explicativo da perspectiva de Marcuschi, quando comparada à tese da separação radical entre fatos de fala e fatos de escrita, não deixamos de concordar com Corrêa (1997; 2001; 2004), que a proposta do contínuo de gêneros textuais recupera a separação entre fala e escrita, ao pressupor certa homogeneidade nos extremos do contínuo. Desse modo, nesta pesquisa, almejamos defender que o tipo de relação entre fala e escrita e, também, entre oralidade e letramento (as quais se mostram por meio do registro escrito) é de ordem distinta das apresentadas até então. Assim, na próxima subseção, apresentamos a proposta de Corrêa (1997, 2004), a qual nos auxilia a entender as hipersegmentações como dado linguístico que possibilita tematizar sobre a relação falado/escrito.

## 2.2. A proposta do modo heterogêneo de constituição da escrita

Dentre várias possibilidades teóricas de se abordar a escrita, e sua relação com a fala, adotamos um espaço singular de compreensão, o qual busca, antes de tudo, “questionar a delimitação do campo da escrita apenas pela constatação óbvia de um material específico – o gráfico – que lhe serve como base semiótica” (CORRÊA, 2004, p. 2). Em particular, a proposta do modo heterogêneo de constituição da escrita, formulada por Corrêa (1997, 2004),<sup>43</sup> firma-se como contraposição ao ideal de oposição radical entre fala e escrita e, também, ao de dicotomização enquanto recurso metodológico.<sup>44</sup> Para Corrêa, a presença de marcas da enunciação falada na enunciação escrita não se explica nem como uma questão de interferência direta, nem como pela ocorrência de gêneros textuais mistos; mais que isso: a presença do falado no escrito é evidência de que a escrita (e a linguagem de modo geral) não é caracterizada pela homogeneidade, mas sim por uma heterogeneidade que lhe é constitutiva.

O cerne dessa teorização, portanto, é o de que a heterogeneidade é própria **da** escrita (e não presente **na** escrita), ou seja, a “heterogeneidade como constitutiva da escrita e não como uma característica pontual e acessória desta. Noutros termos: *a heterogeneidade é [...] interior à escrita e não exterior a ela*”. (CORRÊA, 2001, p. 144, destaques nosso). De tal modo, toda e qualquer produção de escrita é, por excelência, heterogênea. Ao assumir essa perspectiva, Corrêa (1997, p. 86) se distancia daquelas “avaliações que tomam como parâmetro um modelo abstrato – literário ou não – de boa escrita” e acrescenta: “a consideração desse modo heterogêneo pode ser útil como uma contraposição ao preconceito

---

<sup>43</sup> A proposta do modo heterogêneo de constituição da escrita foi resultado da tese de doutoramento do autor (ver CORRÊA, 1997). Em trabalho posterior (ver CORRÊA, 2004) essas ideias foram retomadas e organizadas em forma de livro.

<sup>44</sup> A proposta teórica do contínuo de gêneros textuais foi apresentada por Corrêa (1997, 2004) como perspectiva de dicotomização metodológica.

comum com que se tomam as produções escritas consideradas como menos integradas a um padrão tido como legítimo”.

É por termos buscado, no tratamento das hipersegmentações, conforme afirmado anteriormente, distanciamento de algumas interpretações que as classificam, tão somente, como erros ortográficos<sup>45</sup> (por exemplo, KOCH, 1997) ou como indícios de algum problema de aprendizagem (por exemplo, ZORZI, 1997), que assumimos a premissa da heterogeneidade da escrita. A partir desse lugar teórico, identificamos a oportunidade de levantarmos outras hipóteses explicativas, especialmente às relacionadas à compreensão do modo como os escreventes deixam transparecer a sua relação com a linguagem em sua forma de enunciação escrita.

De acordo com Corrêa, fala e escrita, são *modos enunciativos* ligados intrinsecamente às práticas sociais de oralidade e letramento. Isso significa dizer que os fatos da língua (falado/escrito) só se realizam por serem modos de enunciação concretizados por meio de práticas de linguagem (oral/letrada). Dessa forma, o autor defende haver uma indissociabilidade entre ambos (modos enunciativos e práticas sociais) e, por essa razão, apresenta uma notação conjunta – *oral/falado e letrado/escrito* – ao tratar dos usos da linguagem.<sup>46</sup>

Desses pares vem a conceituação do modo heterogêneo de constituição escrita: “encontro entre as práticas sociais do oral/falado e do letrado/escrito, considerada a dialogia

---

<sup>45</sup> Como viemos conceituando ao longo de todo o trabalho, as hipersegmentações são vistas aqui como *fatos não-convencionais de escrita* e não como erros, seja-o na acepção de construção de hipóteses (conforme CUNHA, 2010) seja-o como desvios de escrita (conforme ZORZI, 1997). Nossa opção é decorrente de entendermos as hipersegmentações como *possibilidades gráficas no interior da língua*; nesse aspecto, elas se distanciam apenas em relação ao que está convencionalizado para a escrita.

<sup>46</sup> Veja-se que, nesse ponto, Corrêa (1997, 2004) se diferencia de Marcuschi (2001), para o qual fala e escrita são modalidades da língua pensadas paralelamente às práticas sociais de oralidade e letramento. Em estudo mais recente, Corrêa (2009) fez, também, uma crítica à ideia que subjaz o termo modalidade linguística. Segundo Corrêa, *modalidade* está relacionada a um juízo de código abstrato, o que significa em outros termos, de uma projeção do que seriam os modelos ideais de fala e de escrita. Circunscrever a escrita ao patamar de modalidade é, para o autor, negligenciar que, “a exemplo da modalidade oral, a modalidade escrita só se realiza como modo de enunciação e que, como tal, se reinstala, no mundo social, por meio de encontros, sempre inéditos, entre os participantes do discurso, para os quais importa a construção de sentidos e não, simplesmente, a obediência às restrições ligadas a um modelo de correção” (CORRÊA, 2009, p. 2).

com o já falado/escrito e ouvido/lido” (CORRÊA, 2004, p. 9). Em virtude dessa definição, o autor assume, também, enquanto articulador das práticas sociais oral/falada e letrada/escrita, o princípio bakhtiniano de dialogicidade da linguagem, tendo, pois, como ponto de partida, a concepção de que toda produção de linguagem é sempre resultado de um já dito/ouvido e escrito/lido.

Para a compreensão da não-distintividade entre as práticas sociais orais/faladas e letradas/escritas, Corrêa (2001) formula o conceito de *letramento amplo*. Nessa proposição, a ambição do autor é, ainda, a de

valorizar as habilidades atestadas por aqueles indivíduos que, mesmo não tendo acesso à alfabetização ou mesmo mantendo-se, na maior parte do tempo, alheios às práticas de leitura e escrita tal como foram consagradas, também fazem a história da língua e da sociedade por meio do modo oral de registro da memória cultural (CORRÊA, 2001, p. 141).

Nessa passagem, o autor esclarece que a ampliação do conceito de letramento tem como consequência o distanciamento de abordagens mais restritas do termo, segundo as quais a base gráfica é um pré-requisito para se pensar a forma de participação dos sujeitos em práticas letradas/escritas. Desse modo, uma implicação direta do conceito de letramento amplo é a ligação com práticas sociais orais/faladas. Para realizar essa aproximação, Corrêa constata que características como mobilidade no espaço e permanência no tempo – atribuídas como exclusivas dos registros gráficos – podem ser percebidas, também, nas formas de registros de sociedades de oralidade primária. Conhecer um fato, como defende o autor, independe do contato empírico com uma materialidade física, figurada em um tipo de escrita como a alfabética, por exemplo, mas pode se dar por formas distintas de relato oral. Com efeito, “em circunstâncias de línguas ágrafas, é, pois, a *palavra-testemunho* do relato oral que apresenta e faz o ouvinte compreender e vivenciar um fato à distância” (CORRÊA, 2001, p.

138, grifos nosso). Assim, as práticas do relato se configuram pelo entrecruzamento da linguagem e da memória e, nesse sentido, “pode-se postular na aparente fugacidade das práticas orais, a permanência no tempo e a mobilidade no espaço, a exemplo do que se costuma atribuir, muitas vezes com exclusividade, às práticas letradas” (CORRÊA, 2001, p. 139).

No desenvolvimento deste trabalho, participamos de discussões realizadas no âmbito do curso de extensão “Introdução aos estudos de letramentos”, ministrado entre os meses de abril a setembro de 2013, no IBILCE/UNESP (São José do Rio Preto/SP), o que nos motivou a entender e definir, para este estudo, o conceito de letramento. Nesse curso, foram realizadas discussões acerca do modo como as pesquisas linguísticas têm abordado o conceito. Segundo Tfouni (1995), a definição do que é letramento está diretamente relacionada ao tipo de posicionamento teórico que se assume e, assim, “pode-se dizer que, no estado atual, já existe uma polissemia relacionada à mesma, o que torna a sua conceituação complicada” (p. 31). Nesta pesquisa, adotamos o conceito de letramento amplo definido por Corrêa (2004), pois compreendemos que essa concepção demonstra a importância em se pensar o funcionamento linguístico dos modos enunciativos no interior de práticas sociais de linguagem, as quais se constituem muito além do vínculo direto com práticas de leitura e escrita institucionalizadas, no sentido em que considera apenas o tipo de escrita alfabética e as convenções ortográficas que a regem. Mas é importante considerar que o material e o dado desta pesquisa, bem como a perspectiva teórica adotada, nos leva a identificar aspectos da inserção do escrevente em práticas letradas/escritas e acaba por encaminhar-se na direção de o imaginário sobre um tipo de escrita esperado da escola, como lugar institucionalizado o qual valoriza, sobremaneira, textos escritos em pleno acordo com as convenções ortográficas. Nossos apontamentos, pois, sobre o tipo de informações letradas/escritas que parecem circundar o aparecimento de

hipersegmentações, não exclui que compreendamos o letramento como uma prática social que não se reduz a um só tipo de informação e sobre um só tipo de escrita.

Nota-se, ainda, outra questão fundamental em relação à caracterização da teorização do modo heterogêneo de constituição da escrita: a de que a heterogeneidade só se mostra por intermédio da relação do sujeito com a linguagem, pois “observar o encontro entre o falado e o escrito, portanto, não é tomar essas práticas como dados autonomamente observáveis, mas apreendê-las pelas marcas que o sujeito, assim constituído, imprime em seu texto” (CORRÊA, 2004, p. 294). Assim, por estar inserido em inúmeras práticas sociais de linguagem (tanto orais/faladas quanto letradas/escritas), o sujeito escrevente, quando enuncia, circula dialogicamente por elas, deixando *pistas* linguísticas do que julga ser a constituição da (sua) escrita, de si como escrevente e de seu(s) interlocutor(es).

Em uma análise de textos de vestibulandos, Corrêa pôde observar que os escreventes, em suas produções escritas, transitavam sempre pelos mesmos *imaginários*, os quais, segundo o autor, seriam socialmente partilhados. A identificação desse fato levou Corrêa a postular, o que ele denominou por *eixos de circulação dialógica do escrevente*, por meio dos quais seria possível ao pesquisador (ainda que conjecturalmente) reconstruir o processo subjacente ao texto escrito. Cabe ressaltar que, apesar do tipo específico de material a partir do qual Corrêa se valeu para propor esses lugares metodológicos, o autor lembra que, esse “processo de constituição da escrita podem ser retomados em qualquer época, na escrita de qualquer pessoa, em qualquer texto” (CORRÊA, 2004, p. 89).

Tratando dos eixos de circulação dialógica, esses são:

(i) *o da representação da (suposta) gênese da escrita*. Com base em Corrêa, esse eixo diz respeito à atribuição de um lugar específico para o oral/falado no letrado/escrito e compreende as marcas gráficas que, durante o processo de produção escrita, foram concebidas como possibilidades de registro biunívoco dos aspectos que constituem o enunciado falado no

interior de uma materialidade gráfica. No entanto, a respeito desse eixo, são necessárias duas ressalvas de ordem teórica. A primeira está relacionada à negação do termo *gênese* como sinônimo de *um ato fundador*. Segundo Corrêa, o uso de *gênese* é calcado na ideia de um “teatro de recomeços” (VERÓN, 1980, p. 118 *apud* CORRÊA, 2004, p. 86), visto que sua constituição pode ser apreendida em qualquer momento do processo de produção textual, não se restringindo, pois, “em algum ponto facilmente localizável do processo de alfabetização” (CORRÊA, 2004, p.86), mas estendendo-se, inclusive, à escrita adulta. A segunda recusa teórica é fundamentada na ideia de escrita enquanto representação da oralidade. Ao refutar essa postura, em favor da gênese da escrita como parte do imaginário que o escrevente faz do processo de constituição da (sua) escrita, o objetivo do autor é não se limitar “a localizar suas marcas em pontos determinados de uma sequência de estágios cronologicamente estabelecidos, nem tampouco as identifica[r] por meio de uma visão normativa que as tom[a] como *desvios*” (CORRÊA, 2004, p. 89, grifo nosso). Portanto, a busca por marcas da gênese da escrita não tem ligação no trabalho de Corrêa (nem tem em nossa pesquisa) com a identificação de possíveis pontos de interferência do falado no escrito, uma vez que essa procura daria margem à tese de modalidades puras de realização da linguagem e não da heterogeneidade constitutiva da escrita por meio da indivisão dos modos de enunciação com as práticas sociais. Em outros termos, os fatos relativos à gênese da escrita são indícios linguísticos da inserção do escrevente em práticas de linguagem orais/faladas;

(ii) *o da representação do código escrito institucionalizado*. Antes de qualquer consideração, o termo *código* merece esclarecimento:

Ao compor a expressão “código escrito institucionalizado”, a palavra “código” não remete, nem ao processo de codificação da língua pela escrita, nem à tecnologia da escrita, identificada, em geral, com a escrita alfabética; nem tampouco supõe, como trabalho de interpretação semiótica, a simples “decodificação” de um produto acabado. Com ela, pretendo significar o processo de fixação metalinguística da escrita pelas instituições, sujeito, portanto, aos movimentos da história e da sociedade. Como, nesse sentido, a institucionalização do código tem uma natureza dinâmica, excluo de consideração qualquer menção a um produto acadêmico

fechado, evitando, inclusive, restringir a sua institucionalização apenas à escola. Desse modo, a representação que o escrevente faz do código escrito institucionalizado deve ser entendida como a representação que ele faz do institucionalizado para a (sua) escrita, ficando aberta, portanto, a consideração de representações particulares, localmente atuantes. (CORRÊA, 2004, p. 10-11).

Para Corrêa (2004, p. 165), as marcas produzidas sob a imagem desse eixo reproduzem “a dinâmica social de institucionalização de valores para as diversas formas linguísticas”. No material de análise de que se vale o autor – e acreditamos, também, podermos estender ao material deste trabalho –, tal dinâmica vem marcada por uma relação tríade a qual vincula a língua com a escrita e essa última com a escola. Nesse sentido, o escrevente guia a sua escrita pelo o que confia ser “– a partir não só do que aprendeu na escola, mas, em grande parte, do que assimilou fora dela – *a visão escolarizada do código institucionalmente reconhecido*” (CORRÊA, 2004, p. 166, grifos nosso). Por suas características, esse eixo contrapõe-se à imagem que se tem da gênese da escrita, pois a “busca de um modelo leva o escrevente a exceder-se numa caracterização do texto baseada em características que ele supõe como próprias (e até exclusivas) da escrita” (CORRÊA, 2004, p. 166) e, nesse sentido, a forma como o escrevente se representa na escrita por meio da circulação dialógica nesse eixo, realça: (i) o seu lugar de reprodutor de uma prática de linguagem já edificada; (ii) a sua busca por atender as expectativas da instituição escolar em relação à escrita culta; (iii) a atribuição para si próprio de um lugar único enquanto escrevente, no interior de práticas letradas/escritas institucionalmente reconhecidas e em função da projeção de um interlocutor;

(iii) *o da dialogia com o já falado/escrito e ouvido/lido*. Com esse último eixo, o autor intenciona mostrar que todo e qualquer texto é sempre produto de um já-dito. Mais do que pistas intertextuais que poderiam marcar (mais diretamente) a relação do escrevente com o já falado/ouvido e o escrito/lido, Corrêa observa a relação da produção textual com o

interdiscurso e, retomando as palavras de Fiorin (1994, p. 35 *apud* CORRÊA, 2004, p. 231), lembra: “a intertextualidade não é um fenômeno necessário para a constituição de um texto. A interdiscursividade, ao contrário, é inerente à constituição do discurso”. No entanto, a dificuldade em recuperar relações intertextuais (pois, nem sempre mostradas) constitui-se em uma pista importante, já que sugere, por um lado, a dificuldade do escrevente em demonstrar linguisticamente a (sua) relação com outros discursos ou, por outro, que a apropriação de um determinado discurso se deu de um modo em que não é mais possível ser recuperado explicitamente. Assim, mesmo quando o escrevente, ao tentar alçar um padrão de escrita, que acredita ser o esperado por determinadas instituições – as quais ele projeta como interlocutoras – sua enunciação não deixa de apresentar marcas que indiciam sua circulação por outras práticas de linguagem e que o constituem como escrevente.

Na próxima parte desta Seção, trazemos um breve apanhado das discussões desenvolvidas, visando sintetizar as reflexões teóricas sobre a concepção de escrita adotada nesta dissertação.

### ***Resumo da Seção***

Iniciamos as discussões procurando tecer observações sobre aquilo que consideramos serem diferentes formas teóricas de entendimento da relação fala/escrita e com as quais não dialogamos. Examinamos, com base nesse propósito, trabalhos que permitem entrever os tipos de entendimento e a consequência desses tipos no tocante à análise de marcas linguísticas, como as de hipersegmentação analisadas por esta pesquisa. A nosso ver, os trabalhos consultados tendem a considerar, desde uma abordagem mais homogeneizante até uma que reconhece certa heterogeneidade, fala e escrita como fatos diferentes da linguagem.

Ao fim da Seção, lançamos luz sobre a concepção de escrita seguida para compreensão das produções textuais de nosso cópuz e das quais extraímos as marcas de hipersegmentação. Assumimos dos trabalhos de Corrêa (1997, 2004) a noção do modo heterogêneo de constituição da escrita e, a partir dela, entendemos a escrita enquanto um modo de enunciação que registra o encontro entre o oral/falado e o letrado/escrito. E é por esse modo de constituição que o autor afirma ser “sempre o produto do trânsito entre práticas sociais orais/faladas e letradas/escritas que nos chega como material de análise do modo de enunciação oral e do modo de enunciação escrito, ambos – como se sabe – manifestação de uma única e mesma língua” (CORRÊA, 2001, p. 142). Desse modo, defendemos que os usos não-convencionais das fronteiras de palavras são *registros* da heterogeneidade da escrita, o que, em outras palavras, equivale a dizer que essas grafias são fruto da imagem que os escreventes têm do que seja a constituição da (sua) escrita, a partir de suas inserções em práticas orais/faladas e letradas/escritas de uso da linguagem.

A Seção seguinte é dedicada aos aspectos metodológicos deste trabalho.

## SEÇÃO 3:

### **Aspectos metodológicos**

---

#### *Apresentação da Seção*

O propósito desta seção é o de caracterizar as escolhas teórico-metodológicas referentes ao material e ao corpus de investigação, aos critérios de identificação dos dados e à forma de análise das grafias de hipersegmentação.

Na primeira subseção (“*Delimitação do material*”) apresentamos o material de pesquisa, delineando as características do banco de dados do qual foram selecionadas produções escritas do EF-II, bem como os critérios estabelecidos para o primeiro recorte dos textos. Na sequência (subseção “*Seleção do corpus e identificação dos dados*”), tratamos dos critérios de definição do corpus e das decisões metodológicas adotadas para o levantamento das ocorrências de hipersegmentação. Após esses esclarecimentos, passamos a abordar os procedimentos de análise dos resultados e os fundamentos teórico-metodológicos que sustentam os procedimentos adotados (subseção “*Procedimentos de análise*”).

#### **3.1. Delimitação do material**

Nesta pesquisa, o material de investigação é composto por produções escritas realizadas em ambiente escolar, por alunos dos quatro últimos anos do EF, ou seja, dos 6º, 7º, 8º e 9º anos. Os textos pertencem ao Banco de Dados de Escrita do Ensino Fundamental II,<sup>47</sup> o qual é resultado do projeto de extensão universitária “Oficinas de Leitura, Interpretação e Produção Textual”. O projeto contou com a coordenação das professoras doutoras do IBILCE-UNESP Luciani Tenani e Sanderléia Longhin-Thomazi e desenvolveu-se no âmbito da escola pública “Zulmira da Silva Salles”, localizada em área urbana do município de São José do Rio Preto (SP). Participaram, também, da elaboração e execução das atividades do projeto junto à escola, bem como da informatização do banco de dados, estudantes de Licenciatura em Letras e do Programa de Pós-Graduação em Estudos Linguísticos (ambos do IBILCE-UNESP). Segundo as coordenadoras do projeto (cf. LONGHIN-THOMAZI E TENANI, 2013), a parceria universidade-escola buscou aprimorar atividades de leitura e de escrita oferecidas aos alunos do EF-II e, também, propiciar experiência e reflexão sobre a prática pedagógica de língua materna aos graduandos e pós-graduandos envolvidos.

O projeto foi desenvolvido em quatro anos (2008, 2009, 2010 e 2011)<sup>48</sup> e, nesse período, realizou atividades de dois caracteres distintos, porém interligados. Uma das atividades consistiu na realização de oficinas de produção textual, aplicadas por alunos da UNESP e/ou pelos próprios professores da escola. As oficinas sempre aconteciam durante as aulas de língua portuguesa e tinham a duração de cinquenta minutos, ou seja, nesse tempo os alunos deviam, junto com o aplicador responsável, ler e discutir proposta de produção textual e, em seguida, redigirem um texto sem possibilidade de reescrita. A análise dos textos, coletados nas oficinas, demonstrou dificuldades dos estudantes do EF-II em relação ao uso da

---

<sup>47</sup> O banco de dados é vinculado ao grupo de pesquisa *Estudos sobre a Linguagem* (GPEL/CNPq).

<sup>48</sup> Em 2008, o projeto recebeu financiamento da FAPERP (Fundação para o Desenvolvimento da Pesquisa de São José do Rio Preto). Nos demais anos (2009 a 2011) recebeu auspícios da Pró-Reitoria de Extensão da UNESP (PROEx). Os auxílios financeiros foram utilizados para custear as despesas com materiais oferecidos nas atividades do projeto e com bolsas para alunos de graduação. O projeto foi, ainda, premiado com o primeiro lugar, na área de Educação, no 6º Congresso de Extensão Universitária UNESP, ocorrido em 2011.

escrita. Por essa razão, outra atividade do projeto foi a de oferecimento de minicursos aos alunos, em horários alternados aos das aulas e de participação facultativa. Os minicursos visavam à sistematização dos conteúdos de português trabalhados em sala de aula, como gêneros textuais, coesão e coerência textual, pontuação, ortografia, dentre outros temas. De acordo com Longhin-Thomazi e Tenani (2013, p. 4), todas as atividades propostas pelo projeto de extensão tiveram como suporte teórico-metodológico a concepção de escrita defendida nos estudos de Corrêa (1997, 2001, 2004, 2007) e por esse motivo

as iniciativas do projeto foram guiadas pela valorização da multiplicidade de letramentos associada às práticas sociais vivenciadas pelos alunos, pela consideração dos gêneros discursivos, e pelo reconhecimento da importância de atentar ao processo de escrita do aluno, e não apenas ao produto final, visando ao alcance da escrita “institucionalizada”.

As propostas de produção textual eram formuladas em torno de um texto verbal e/ou não-verbal. A seguir, trazemos um exemplo de proposta:<sup>49</sup>

---

<sup>49</sup> A proposta textual apresentada é a primeira proposta aplicada em 2008, no 6º ano.

**Quadro 1:** Exemplo de proposta de produção textual do projeto de extensão

- Observe a tirinha e discuta com os seus colegas e professor(a) como o tema amoroso é tratado.



- A partir da discussão, escreva um texto em que dê continuidade à história, contando o que aconteceu com cada uma das personagens após a cena do último quadrinho. Para escrever seu texto, assuma a visão de uma das personagens.
- Seu texto deve conter de 20 a 25 linhas e deve ser escrito à tinta. Seu texto não deve ultrapassar os limites designados para a escritura.
- Dê um título a seu texto.

Fonte: Banco de Dados de Escrita do Ensino Fundamental II

Em 2008, primeiro ano do projeto, foram aplicadas propostas de redação em todas as turmas dos quatro anos do EF-II (ou seja, do 6º ao 9º anos). A partir de 2009, bem como nos demais anos, as oficinas de produção textual foram dirigidas apenas àquelas turmas de alunos que, em 2008, cursavam o 6º ano. Em virtude dessa estratégia de condução das atividades de coleta dos textos, o banco de dados se constituiu de modo a possibilitar a execução de pesquisas tanto por amostras transversais quanto por amostras longitudinais, como a nossa.

No geral, o banco de dados é composto por 5.468 textos. Essa quantidade está dividida em 2.752 textos destinados a estudos transversais e 3.614 textos reservados para investigações longitudinais.<sup>50</sup> Quanto à amostra longitudinal do banco, 124 estudantes foram os que

<sup>50</sup> As produções de textos do banco são identificadas por um código, cujas informações são: escola/ano letivo, série/turma, sujeito/sexo, proposta de texto. Na reunião dessas informações, os textos são organizados por uma sigla como “Z08\_5A\_01F\_01” – leia-se: Zulmira, 2008, quinta A, estudante 1, sexo feminino, proposta 1. Cabe lembrar que o início do projeto de extensão se deu quando o EF ainda era composto por oito anos de duração, o que explica as nomenclaturas antigas, 5ª, 6ª, 7ª, 8ª séries, no código de identificação dos textos. No entanto,

participaram integralmente do projeto de extensão e, desse total, 119 alunos redigiram textos em todos os anos, somando um valor final de 2.495 produções de textos (sendo: 611 textos no 6º ano; 726 no 7º ano; 477 no 8º ano; e 682 no 9º ano). Sobre a quantidade de textos redigidos por cada aluno em cada um dos anos e no geral das coletas, a maioria dos estudantes (52,1%) produziu entre 6 e 7 textos por ano e entre 21 e 23 textos ao todo. A partir desse resultado, as coordenadoras do projeto de extensão concluíram que “a participação dos alunos ao longo dos anos letivos não foi homogênea, embora tenha sido constante para esse grupo” (LONGHINTHOMAZI E TENANI, 2013, p. 9).

Tendo em vista nossa decisão por um tipo estudo de longitudinal, *a priori* já possuíamos um recorte dos textos para a composição do material, já que dentre todas as produções disponíveis do banco, 2.495 disseram respeito à produção dos estudantes que sempre realizaram textos nas oficinas do projeto de extensão. Considerando, desse modo, essa primeira eleição do número de sujeitos e textos (ou seja, 119/2.495), estabelecemos mais *dois* critérios de seleção para a definição do material de pesquisa, uma vez que a grande extensão da amostra longitudinal do banco de dados inviabilizaria a execução de uma pesquisa de mestrado, em função do tempo para a sua concretização (24 meses) e a dificuldade na organização e interpretação dos dados a partir de um cunho qualitativo. Conseqüentemente, os dois critérios adicionais de seleção tiveram por horizonte um recorte que contemplasse, simultaneamente, um conjunto relevante tanto em termos quantitativos, que possibilitasse a identificação de regularidades dos dados de hipersegmentação, quanto em termos qualitativos, que permitisse a análise da trajetória de cada um dos escreventes no decorrer do EF-II, em relação às grafias não-convencionais das fronteiras de palavras.

Explicitados os motivos que nos levaram ao segundo recorte do banco, apresentamos, a seguir, os dois critérios que embasam o material desta pesquisa: 1) produção, no sexto ano,

---

embora os escreventes do corpus tenham cursado oito e não nove anos no EF, usaremos a nomenclatura atual dos anos letivos, por motivo de este estudo estar sendo executado com a implementação dos nove anos do EF já em vigor.

de cinco e seis textos por um mesmo sujeito; e 2) produção, nos outros anos (7º, 8º e 9º), de pelo menos três textos por cada um dos sujeitos identificados pelo primeiro critério.

O primeiro critério foi colocado em virtude dos resultados obtidos em nossa pesquisa de iniciação científica<sup>51</sup> (SILVA, L., 2011). Nesse estudo, dados de segmentação não-convencional de palavras foram analisados em textos de 5º ano do EF, a partir dos quais pudemos concluir que os escreventes encerraram a etapa do EF-I sem saberem com exatidão os limites gráficos das palavras. Somaram-se às nossas conclusões, as observações de Paranhos e Tenani (2011) acerca da grande recorrência de dados de hipo e hipersegmentação em textos escritos de estudantes de 6º ano (início do EF-II). Por essas duas constatações, adotamos o mesmo critério de seleção feito por Paranhos e Tenani (2011), ao elegermos para análise os textos daqueles alunos que tinham produzido, no sexto ano, entre 5 e 6 textos (85% e 100% das coletas realizadas),<sup>52</sup> quantidade relativamente alta em relação ao número total de propostas de produção de textos que foram aplicadas naquele ano. Dessa maneira, o ponto de partida de seleção do material estabeleceu um diálogo com o que já foi descrito e analisado por Paranhos e Tenani (2011), sobre a alta ocorrência de hipersegmentação, objeto de investigação deste trabalho.

Definido o critério e identificados os alunos que os atenderam, procuramos os textos produzidos pelos alunos nos demais anos escolares do EF-II. Após o segundo levantamento, lemos todos os textos e pudemos constatar que, para os 7º, 8º e 9º anos, o primeiro critério não era pertinente para os objetivos de nosso estudo, uma vez que dados de hipersegmentação eram praticamente inexistentes na produção desses alunos. Esse cenário nos influenciou na adoção de mais um critério de seleção. Assim, com o estabelecimento do segundo critério – alunos que tenham produzido pelo menos 50% das propostas aplicadas –, buscamos ampliar o número de textos a serem analisados e, nessa direção, os possíveis números de ocorrências de

---

<sup>51</sup> FAPESP/Proc. 2010/17720-8.

<sup>52</sup> Nos anos de 2008 e 2010 (6º e 8º anos) foram aplicadas, em cada um dos anos letivos, seis propostas textuais. Já nos anos de 2009 e 2011 (7º e 9º anos), o projeto aplicou sete propostas de textos.

hipersegmentação. Formalizados os critérios, identificamos um total de **565 textos** para constituir o material de pesquisa.

Na próxima subseção, passamos a definir a composição do *cópus* e o levantamento dos dados.

### **3.2. Seleção do *cópus* e identificação dos dados**

A partir dos 565 textos que compõem o material, foram eleitas, para a formação do *cópus*, as produções escritas de alunos que apresentaram: 1) grafia de hipersegmentação no 6º ano; e 2) grafia de hipersegmentação em outro(s) ano(s) escolar. O primeiro requisito foi determinado para que pudéssemos verificar se os alunos que produziam hipersegmentação no início do EF-II, ao longo dos anos subsequentes continuavam a produzi-las ou atingiam a convenção de palavra escrita. Já o segundo requisito foi estabelecido com vistas a observar como se caracterizariam as hipersegmentações durante o restante do EF-II, tanto em relação à sua estrutura, quanto no que diz respeito às pistas de uma relação mais particular do escrevente com a escrita. Ao fim do levantamento, chegamos ao *cópus* deste trabalho, a ser composto por 12 escreventes (5 do sexo feminino e 7 do sexo masculino) que, ao longo dos quatro últimos anos do EF, produziram 266 produções textuais.

Na análise dos dados, fazemos referência aos escreventes do *cópus* por meio de nomes fictícios elegidos aleatoriamente. Com essa decisão, buscamos preservar a identidade dos alunos e, também, garantir, aos escreventes produtores dos textos, a imagem de sujeitos constituídos sócio-historicamente.<sup>53</sup> Os pseudônimos dos escreventes serão os seguintes: 1- *André*; 2- *Viviane*; 3- *Ana*; 4- *Henrique*; 5- *Camila*; 6- *Fernando*; 7- *Julia*; 8- *Mateus*; 9- *Maria*;

---

<sup>53</sup> Esclarecemos que essa forma de conceber os sujeitos da pesquisa não se aproxima nem de uma noção de indivíduo nem de uma noção de sujeito tal como é trabalhada, por exemplo, na Análise do Discurso.

10-Lucas; 11-Pedro; 12-Bruno (os números de 1 a 12 estão organizados em função do número de grafias de hipersegmentação produzidas).

Uma vez que o nosso objeto de estudo se caracteriza pela presença de um limite gráfico em posição não prevista pela convenção ortográfica, uma questão metodológica importante diz respeito à interpretação de quando a presença de um recurso gráfico pode ser considerada um caso de hipersegmentação. Tal questão é pertinente, sobretudo, em razão de nosso *cópus* ser constituído por textos manuscritos, os quais apresentam, muitas vezes, irregularidades na caligrafia dos escreventes.

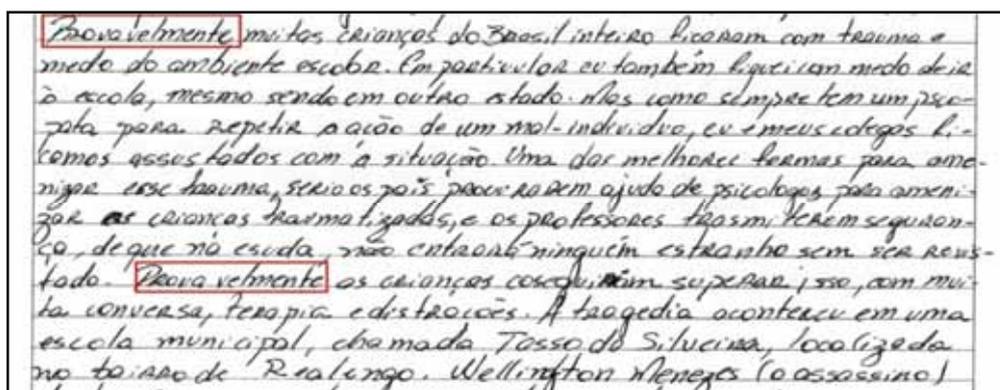
Para a identificação das hipersegmentações presentes em nosso *cópus*, tomamos como eixo norteador a proposta de Tenani (FAPESP/Proc. 2009/14848-6), publicada em Tenani (2011b). A autora sugere uma tipologia no levantamento de dados de segmentação não-convencional baseada tanto em um *critério gráfico* (espaço em branco: “na quela” ou hífen: “mora-va”) quanto em um *critério morfossemântico* (homonímias: “de mais”, quando a grafia é “demais”). Segundo Tenani (2011b), na literatura especializada, os dados de segmentação não-convencional têm sido identificados apenas em relação à ausência/presença do espaço em branco. Nesse sentido, o intuito de Tenani, ao estabelecer outros critérios no arrolamento dos dados, é o de explicitar novas questões linguísticas subjacentes no processo de delimitação gráfica das palavras.

De posse dessa metodologia de levantamento de dados, tratamos, em um primeiro momento, da categorização do espaço em branco e localizamos quatro problemas para a identificação das hipersegmentações: dificuldades em interpretar a caligrafia do escrevente, casos de translineação, ocorrências de palavra estrangeira e de nome próprio.

A respeito do primeiro problema, tomamos a decisão de excluir do *cópus* grafias cuja caligrafia nos gerou dúvida se a palavra em análise se tratava de uma hipersegmentação ou se consistia, apenas, de uma característica da forma como o escrevente, por um lado, registra as

letras (sem ligaduras) na folha de papel e, por outro, organiza o tamanho dos espaços em branco entre as palavras. A Figura 9, exposta a frente, ilustra esse problema. À primeira vista, as duas ocorrências da palavra *provavelmente* foram entendidas como um caso de hipersegmentação, em razão de parecer haver um limite gráfico maior entre “prova” e “velmente”. Entretanto, através de uma análise mais detalhada da distribuição dos espaçamentos entre as palavras no decorrer de todo o texto, consideramos que ambas as grafias não poderiam ser classificadas como hipersegmentações, pois outras palavras escritas do texto pareceram apresentar, também, a possível presença de um espaço; as palavras a que nos referimos são: “assustados” (linha 5), “procurarem” (linha 6), “terapia” (linha 10), “distrações” (linha 10) e “chamada” (linha 11). Assim, concluímos que a presença do espaço em branco, nesses casos, corresponde, de modo mais proeminente, a uma característica caligráfica do escrevente.

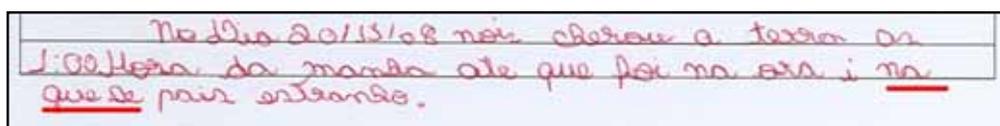
**Figura 9:** Primeiro problema de categorização do espaço em branco: caligrafia



Fonte: Banco de Dados de Escrita do Ensino Fundamental II  
(texto: Z11\_8B\_04F\_03)

O segundo problema encontrado na categorização do espaço em branco foi a translineação. Na escrita, a translineação ocorre com a necessidade de dividir uma palavra em duas linhas, de modo que uma parte da palavra fica no final de uma linha e o restante dela na parte inicial da linha seguinte. A regra para marcar a translineação é a inserção de um hífen do lado direito da palavra, indicando a sua divisão. A seguir, na Figura 10, trazemos um exemplo de translineação identificada no *cópus*. Nesse exemplo, ao segmentar em duas linhas a palavra *naquele*, o escrevente não o fez com a sinalização de um hífen, sendo que na linha em que foi grafada a sílaba “na” houve espaço suficiente para o registro de toda palavra, levando em conta o tamanho da letra do escrevente. Também, a parte da palavra localizada na linha inferior apresenta uma fronteira entre as sílabas “que” e “le”. No entanto, apesar das especificidades descritas da ocorrência de *naquele*, as quais poderiam caracterizá-la como uma hipersegmentação, não a classificamos dessa maneira, pois optamos por descartar todas as grafias de palavras que estivessem no contexto de translineação.

**Figura 10:** Segundo problema de categorização do espaço em branco: translineação

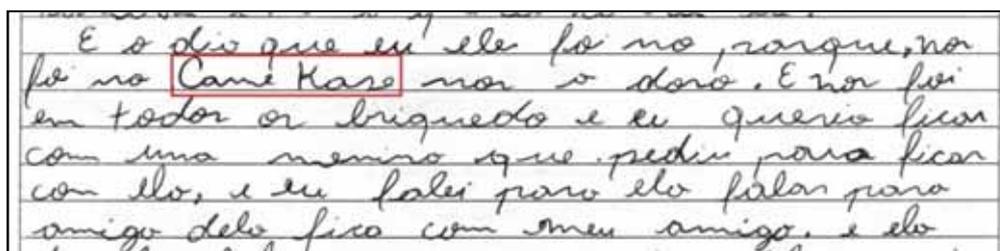


Fonte: Banco de Dados de Escrita do Ensino Fundamental II  
(texto: Z08\_5C\_09M\_05)

Nas Figuras 11 e 12, são expostos exemplos de grafias de palavra estrangeira e nome próprio (terceiro e quarto problema para a categorização do espaço em branco) cujas características gráficas apontariam para uma hipersegmentação. De partida, as ocorrências que se encaixaram nessas categorias (nome próprio e palavra estrangeira) foram excluídas do

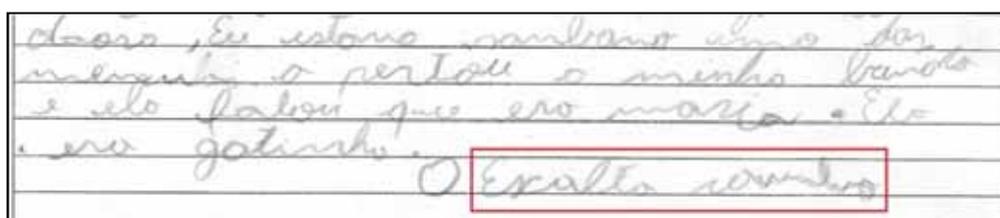
cópus, já que palavras estrangeiras seguem uma ortografia distinta da do português e nomes próprios não são, necessariamente, obrigados a atenderem a convenção ortográfica. Extrapolam, pois, nesses aspectos, os limites da análise almejada por este trabalho, quanto a grafias que se distanciam das normas ortográficas do PB. Na hipersegmentação “Came Kase” (ver Figura 11), o escrevente fez referência a um brinquedo comumente encontrado em parques de diversão; no entanto, a grafia esperada para ele é *Kamikase*, uma palavra de origem japonesa. Já, ao se referir a um grupo musical brasileiro (ver Figura 12), o escrevente registrou como “Exalta samba” (por serem duas palavras lexicais) um nome próprio cuja grafia é *Exaltasamba*.

**Figura 11:** Terceiro problema de categorização do espaço em branco: palavra estrangeira



Fonte: Banco de Dados de Escrita do Ensino Fundamental  
(texto: Z11\_8E\_09M\_01)

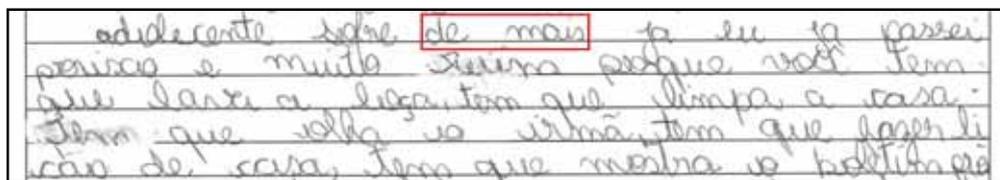
**Figura 12:** Quarto problema de categorização do espaço em branco: nome próprio



Fonte: Banco de Dados de Escrita do Ensino Fundamental  
(texto: Z10\_7C\_23M\_04)

No caso das palavras homônimas, utilizamos a estratégia de considerar todo o sentido do enunciado em que a grafia ocorreu, para verificarmos qual foi o significado privilegiado pelo escrevente em relação àquela palavra. A esse respeito, apresentamos, na Figura 13, a ocorrência “de mais” (cf. Figura 13). Na sentença: “adolescente sofre de mais” (linha 1), a palavra hipersegmentada tem a função sintática de adjunto adverbial, o qual está se referindo ao sofrimento excessivo dos adolescentes, conforme a opinião da escrevente. Portanto a grafia “de mais”, nesse caso, é um dado de hipersegmentação, já que a tentativa da escrevente foi a de grafar o advérbio de intensidade *demais* e não a locução *de mais* como em: “preciso de mais ajuda”.

**Figura 13.** Exemplo de hipersegmentação de palavra homônima

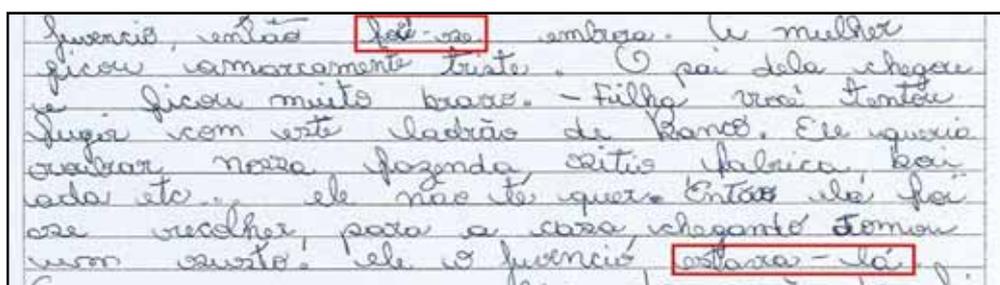


Fonte: Banco de Dados de Escrita do Ensino Fundamental II  
(texto: Z09\_6A\_27F\_05)

Em outro momento da seleção dos dados, tratamos da categorização do hífen. Quanto ao uso desse recurso gráfico em nosso corpúsculo, dois foram entendidos como não-convencionais, sendo um uso excluído da análise, em função da presença do hífen nos casos levantados não configurar como uma hipersegmentação, já que o critério principal que a define diz respeito ao emprego, para mais que o esperado, do limite gráfico no interior de palavras ortográficas. Embora possam ser considerados usos não-convencionais do hífen, conforme mostram as grafias destacadas na Figura 14, esses registros em nada alteram a

segmentação das palavras, dada que a presença do recurso está localizado de modo a unir/separar duas unidades morfossintáticas. Na verdade, essas ocorrências indicam um tipo de relação entre palavras que é semelhante à observada em palavras compostas. Destacamos, ainda, sobre o emprego do hífen em nosso cópua, que as grafias ilustradas na Figura 14, se distinguem, por exemplo, do dado “arranja-se”, mostrado na Figura 1, na Seção de apresentação desta pesquisa, pois no caso dessa última grafia não-convencional, o hífen hipersegmenta a palavra escrita *arranjasse* (ver nota três neste trabalho).

**Figura 14.** Problema de categorização do hífen: configuração de palavra composta



Fonte: Banco de Dados de Escrita do Ensino Fundamental II  
(texto: Z08\_5A\_23F\_01)

Ao descrevermos os problemas e as decisões metodológicas tomadas no levantamento das hipersegmentações, procuramos mostrar as dificuldades encontradas na identificação dos empregos não-convencionais dos recursos indicadores de palavra, especialmente em materiais manuscritos. Nas pesquisas de Paula (2007) e de Capristano (2003), por exemplo, há uma descrição de critérios que elegem dados de segmentação não-convencional em textos infantis. Por este trabalho abordar textos escritos do EF-II, compreendemos a necessidade de estabelecermos critérios distintos daqueles adotados pelos estudos de dados do EF-I. A análise das produções manuscritas com as quais trabalhamos revelou-nos características (quanto à forma de redação das letras e distribuição dos recursos gráficos) muito distintas das

que foram observadas nos textos de escrita inicial. Portanto, julgamos importante, para este estudo, um procedimento atento e particularizado de identificação dos dados, com vistas a garantir um levantamento das hipersegmentações com base em critérios explícitos e justificados, proporcionando, assim, a configuração de um *cópus* de hipersegmentação que caracteriza a escrita de escreventes com mais anos de escolarização.

### **3.3. Procedimentos de análise dos dados**

As hipersegmentações identificadas nos textos do EF-II serão submetidas às análises quantitativa e qualitativa. A adoção dessas formas de análise se deu com o intuito de descrevermos as características gerais dos dados, examinando a distribuição das ocorrências em cada ano letivo e para cada sujeito escrevente. A fim de verificarmos a distribuição das grafias de hipersegmentação, observamos a quantidade de dados considerando: (i) a comparação entre o número de palavras convencionais e de hipersegmentação dessas palavras; e (ii) a relação entre a extensão dos textos (dada pelo número de palavras escritas) e o número de dados encontrados. Por meio dessa organização dos dados, conseguimos identificar palavras que continuam a ser hipersegmentadas nos anos finais do EF e, a partir desse resultado mais geral, é possível examinar: (i) se os tipos de palavras convencionais e de hipersegmentações diminuem com o passar dos anos, como um possível efeito das práticas letradas/escritas; (ii) se essa possível diminuição se aplica, também, quando analisada a trajetória de cada um dos escreventes; (iii) se as palavras hipersegmentadas são sempre as mesmas em todo o EF-II e para todos os escreventes.

Para a caracterização dos dados, partimos da identificação das possíveis estruturas prosódicas que se mostraram mais recorrentes nas palavras convencionais e nas

hipersegmentações, organizando-as em tipos em função da propriedade estrutural da palavra. Assim, em termos estruturais, partindo da consideração de que a principal característica desse tipo de dado é ocorrer no *nível da palavra* (diferindo, nesse aspecto, das hipossegmentações, que sempre ocorrem entre duas unidades), consideramos a forma como se organizam as sílabas no interior das palavras convencionais e no interior das grafias não-convencionais. Por exemplo, a palavra cujo limite convencional é “embora” tem suas sílabas organizadas metricamente da seguinte maneira: (• \* •); em particular, essa palavra sempre foi hipersegmentada no cópuz como “em bora”, a partir da qual houve a reorganização das sílabas da palavra como: (•)σ (\* •)Σ, ou seja, como o registro de duas unidades prosódicas, que são a *sílaba* e o *pé métrico*, respectivamente.

Ao assumirmos, com Corrêa (2004), a noção de trânsito do escrevente pelo imaginário de representação do oral/falado na (sua) escrita, faz-se necessário explicitar que optamos por observar apenas o aspecto referente à *prosódia*<sup>54</sup> dos enunciados falados. Restringimos nossa investigação para uma dentre as várias regularidades, relacionadas às dimensões da linguagem, apontadas por Corrêa (2004, p. 92 e 114) durante a observação da manifestação da gênese da escrita nos textos dos vestibulandos. Nossa opção pela prosódia se alinha, também, com a escolha (já feita na primeira Seção deste estudo) pelo viés de análise proposto nos trabalhos de Tenani (2010, 2011b), Chacon (2004, 2005) e Capristano (2003, 2004), por exemplo, sobre a relação entre fronteiras não-convencionais de palavras e padrões rítmico-entoacionais que se formalizam em constituintes prosódicos da língua. Acreditamos, ainda,

---

<sup>54</sup> Na apresentação do livro “Estudos de prosódia”, Scarpa (1999) descreve que as investigações sobre fatos prosódicos abrangem diferentes fenômenos, que vão desde parâmetros de altura, intensidade, duração, pausa e velocidade de fala até o estudo dos sistemas de tom, entoação, acento e ritmo das línguas. Abarcando esses fenômenos, a autora identifica dois ramos de estudos: o primeiro (mais fonético) responde pelo interesse *acústico, mensurável, instrumental* da altura, intensidade e quantidade, correlatos perceptuais de frequência, volume e duração, em termos individuais, e, também, pelos correlatos acústicos dos sistemas de acento, entoação e ritmo, no âmbito das línguas. O segundo tipo de estudo (mais fonológico) tem interesse pela organização e representação dos sistemas de ritmo, acento e entoação e a relação de interface desses com os outros componentes gramaticais. Nesta pesquisa, privilegiou-se o aspecto fonológico da prosódia na relação com o texto escrito.

em acordo com os referidos pesquisadores, que fatos prosódicos não são exclusivos dos enunciados falados e, por isso, podem ser também recuperados na escrita.

Para subsidiar a descrição dos dados, sobretudo no que se refere à identificação de aspectos do oral/falado no texto escrito, fizemos uso das formulações do modelo de Fonologia Prosódica de Nespore e Vogel (1986).<sup>55</sup>

Em sua proposta, Nespore e Vogel (1986, p. 6) buscam distanciamento dos pressupostos da fonologia gerativa linear e sugerem um enfoque “que organiza uma dada sequência da língua em uma série de constituintes fonológicos arranjados hierarquicamente que, por sua vez, formam contextos para a aplicação de regras fonológicas”.<sup>56</sup> A definição da fronteira de cada constituinte prosódico é feita com base em evidências segmentais de aplicação de processos fonológicos que atingem as bordas esquerda e/ou direita dos domínios. Após vasta constatação de diferentes processos fonológicos em diferentes línguas, como o turco, o latim e o grego, por exemplo, Nespore e Vogel chegam à identificação de sete domínios prosódicos: sílaba ( $\sigma$ ), pé métrico ( $\Sigma$ ), palavra fonológica ( $\omega$ ), grupo clítico (C), frase fonológica ( $\phi$ ), frase entoacional (I) e enunciado fonológico (U) – organizados, nessa apresentação, em ordem crescente.

Conforme afirmamos anteriormente, os constituintes prosódicos baseiam-se em diferentes informações linguísticas e, desse modo, levam em consideração informações fonológicas e, também, informações de outros componentes da gramática, como o morfológico, o sintático e o semântico. Embora, em alguns casos, os constituintes prosódicos

---

<sup>55</sup> A eleição pela proposta de Nespore e Vogel (1986), frente aos demais modelos de Fonologia Prosódica, justifica-se (além da filiação à linha de trabalhos com os quais dialogamos no estudo das segmentações não-convencionais), pelo trabalho de Tenani (2002) no qual a pesquisadora aponta evidências dos domínios prosódicos formulados por aquelas duas autoras para a explicação de fenômenos do PB. Nesse sentido, é importante esclarecer que os objetivos deste trabalho não são o de discutir, a partir das hipersegmentações de palavras, a pertinência dos algoritmos de formação dos constituintes prosódicos para a explicação de dados de escrita. Nosso interesse é, antes de tudo, observar, a partir do arcabouço teórico da Fonologia Prosódica, como marcas ortograficamente discordantes podem ser evidências de estruturas linguísticas, ou seja, como grafias de palavras refletem possíveis fatos da língua.

<sup>56</sup> “that organizes a given strings of language into a series hierarchically arranged phonological constituents that in turn form the contexts within phonological rules apply”. Todas as traduções realizadas, nesta dissertação, foram de minha responsabilidade.

venham a coincidir com outros constituintes linguísticos, essa coincidência não é sempre obrigatória, pois a relação entre os constituintes prosódicos e os demais constituintes da gramática não é do tipo isomórfica. Assim, os constituintes da hierarquia prosódica são construídos com base no mapeamento de regras de vários constituintes gramaticais e se integram à estrutura prosódica pelo atendimento de quatro princípios universais:

*Princípio 1:* uma dada unidade não-terminal da hierarquia prosódica,  $X^p$ , é composta de uma ou mais unidades da categoria imediatamente inferior  $X^{p-1}$ ;

*Princípio 2:* uma unidade de um dado nível da hierarquia está exaustivamente contida na unidade superior da qual é uma parte;

*Princípio 3:* as estruturas hierárquicas da fonologia prosódica são estruturas de ramificação n-ária;

*Princípio 4:* a relação de proeminência relativa definida entre nós irmãos é tal que a um só nó é atribuído o valor forte (s) e a todos os outros nós é atribuído o valor fraco (w) (NESPOR E VOGEL, 1986, p. 7).<sup>57</sup>

Dentre os sete constituintes prosódicos, quatro se mostraram relevantes, nesta pesquisa, para a análise das hipersegmentações de palavras. São eles: a sílaba, o pé métrico, a palavra fonológica e o grupo clítico (doravante,  $\sigma$ ,  $\Sigma$ ,  $\omega$  e C, respectivamente).

A partir de Nespor e Vogel (1986), a  $\sigma$  e o  $\Sigma$  são os menores constituintes prosódicos e se definem, exclusivamente, por informações fonológicas. No seu interior, a  $\sigma$  agrupa segmentos vocálicos e consonantais e, no que se refere às relações de proeminência, esse constituinte possui como elemento de valor forte, em se tratando do PB, uma vogal e como elementos dominados (nós de valor fraco), as consoantes e/ou glides. A  $\sigma$  é dominada pela  $\omega$ , após ser organizada no interior de  $\Sigma$ s. Essa última categoria prosódica se caracteriza pela relação instituída entre duas ou mais  $\sigma$ s, de modo que, à semelhança do constituinte anterior, apenas uma das  $\sigma$ s será a portadora da proeminência (o acento, nesse caso). Analisando a

---

<sup>57</sup> *Principle 1:* a given nonterminal unit of the prosodic hierarchy,  $X_p$ , is composed of one or more units of the immediately lower category,  $X_{p-1}$ ; *Principle 2:* a unit of a given level of the hierarchy is exhaustively contained in the superordinate unit of which it is a part; *Principle 3:* the hierarchical structures of prosodic phonology are n-ary branching; *Principle 4:* the relative prominence relation defined for sister nodes is such that one node is assigned the value strong (s) and all the other nodes are assigned the value weak (w).

organização das palavras no PB, Bisol (1996) defende que a maioria delas é constituída por  $\Sigma$ s binários de proeminência à esquerda, um *troqueu*, como é a organização rítmica de palavras como *casa* e *mesa*, por exemplo. Existem, também, palavras cuja configuração é a de  $\Sigma$ s binários de cabeça à direita, formando um *iambo*, a exemplo de palavras como *sofá* e *café*. Em palavras como *dócil* e *órgão* há a formação de um pé *espondeu*, formado por duas sílabas longas: uma pelo acento e outra em razão da coda silábica; palavras com essa configuração são encontradas em baixa ocorrência no PB. Por fim, existem palavras constituídas por  $\Sigma$  ternário, por exemplo, um pé *dátilo*, como ocorre nas palavras com padrão de acento proparoxítono, como *lâmpada* e *árvore*; e um pé *anapesto*, como ocorre nas palavras oxítonas com duas sílabas pretônicas, em *jacaré*, *coração*.

Seguindo os princípios da hierarquia prosódica, os  $\Sigma$ s de uma dada sequência são dominados pela  $\omega$ . Nesse domínio, o  $\Sigma$  mais proeminente será aquele em que, sobre uma de suas sílabas, recair o acento primário. Como já discutimos, a  $\omega$  é caracterizada por ser portadora de um só acento, fato que a leva, muitas vezes, a não coincidir com os limites da palavra morfológica. A esse respeito, Nespor e Vogel salientam que a  $\omega$  é o primeiro constituinte prosódico a manter relação com noções de outros componentes gramaticais, mais especificamente, nesse caso, o da morfologia. Já, por fim, o C é o constituinte que, além de fazer uso de informações morfológicas, é o primeiro a contar com informações sintáticas na definição de seu domínio. Formado por um ou mais clíticos e uma  $\omega$  (que serve de hospedeira fonológica), o C tem como nó mais proeminente a  $\omega$ , já que internamente um clítico é desprovido de proeminência e, assim, não carrega a proeminência do constituinte.

Na análise desenvolvida, dispensamos maior destaque aos tipos de informações linguísticas que podem ser detectadas a partir de uma organização prosódica dos dados. Por meio de nossa decisão, não descartamos, por consequência, os aspectos letrados/escritos que parecem cercar as grafias de hipersegmentação, de modo que, em acordo com a perspectiva

teórica da concepção de escrita, não é possível afirmar que são apenas aspectos das práticas orais/faladas que estão subjacentes às fronteiras não-convencionais de palavras, mas há sempre um trânsito entre essas práticas e as práticas letradas/escritas que constituem todo e qualquer material escrito. Dessa maneira, a nossa detecção pelas marcas letradas segue duas direções: a primeira que trata do aspecto visual, ou seja, da atenção sobre as formas de delimitação das palavras, seja por espaço em branco ou hífen; a segunda direção identifica os aspectos morfossemânticos, privilegiados pela escrita no estabelecimento da convenção ortográfica, que demonstram ter peso nas formas de hipersegmentação encontradas.

Além da busca por regularidades linguísticas dos dados, analisamos, também, de ocorrências, muitas vezes, únicas da escrita dos escreventes. Nesse aspecto, tratamos, também, daqueles dados de hipersegmentação que, estruturalmente, não seguiram todas as regularidades observadas na descrição geral dos dados, mas que são relevantes para a compreensão de uma relação mais particular do escrevente com a linguagem em seu modo de enunciação escrito. Nossa estratégia metodológica, em relação a essas grafias, é a de recuperar as produções textuais em que houve as ocorrências de hipersegmentação, pois a consideração do texto pode ser relevante para a discussão e explicação dos dados, na medida em que alguns dos limites não-convencionais podem dar pistas de terem sido propostos como um modo de articular os sentidos do texto e de garantir uma prosódia ao dizer, como foi argumentado por Tenani (2008). Em suma, os dois métodos de análise foram adotados de modo complementar, estando perpassados em toda a análise dos dados, de modo que, neste trabalho, pudéssemos contemplar uma investigação tanto do geral quanto do particular das grafias de hipersegmentação.

### ***Resumo da Seção***

Discorremos, inicialmente, nesta seção, a respeito da caracterização do Banco de Dados de Escrita do Ensino Fundamental, um banco de dados inédito de textos do EF-II, o qual está disponível para pesquisas na UNESP, campus de São José do Rio Preto. Na mesma subseção, apresentamos, também, os critérios para a seleção do material, que levaram em conta um recorte já preestabelecido em relação à amostra longitudinal do banco de dados. Em seguida, tratamos da definição do *córpus* e dos aspectos gráficos e morfossemânticos relevantes para a atenta seleção dos dados de hipersegmentação. Atinente a essa última tarefa, durante a leitura das produções escritas do *córpus*, excluimos de análise aquelas grafias de espaço em branco que foram casos de translineação, palavra estrangeira, nome próprio e cuja caligrafia do escrevente pôs em dúvida se havia a presença de um espaço ou se apenas se tratava de uma característica relacionada à forma de grafia das letras. Expusemos, ainda, o critério para a seleção de palavras homônimas. Quanto ao emprego do hífen, tiramos de consideração o uso desse recurso entre duas palavras, por motivo de não vermos, nesses casos, a ocorrência da não-convencionalidade dos limites morfosintáticos da palavra.

Na terceira subseção, dedicamos atenção às formas de análise dos resultados e aos aspectos metodológicos que embasam a execução de tais formas. Esboçamos, então, os pressupostos do modelo de Fonologia Prosódica de Nespor e Vogel (1986), modelo que trata da organização prosódica das línguas, por meio de sete domínios, dentre os quais foram trazidas maiores informações acerca dos algoritmos de formação daqueles domínios que se fizeram pertinentes para análise e interpretação das hipersegmentações analisadas neste trabalho.

## SEÇÃO 4:

# **Análise das hipersegmentações de palavras**

---

### *Apresentação da Seção*

Nesta parte, expomos a análise dos dados de hipersegmentação. Por meio dessa análise buscamos discutir: (i) aspectos referentes à distribuição das ocorrências, os quais indicam hipóteses mais gerais dos escreventes em relação à noção de palavra escrita; e, também, (ii) aspectos mais particulares das grafias não-convencionais, tendo em conta hipóteses individuais de alguns dos escreventes analisados. Para tanto, as reflexões estão organizadas em dois momentos.

Iniciamos (subseção “*Aspectos ‘gerais’ das hipersegmentações*”) com a apresentação dos dados, levando em conta a distribuição desses em relação aos anos escolares do EF-II, aos escreventes, aos recursos gráficos (espaço em branco e hífen) e às estruturas prosódicas de palavras escritas convencionais e hipersegmentadas. A partir da descrição e análise do objeto com base nessas quatro variáveis, encontramos evidências que nos permitiram identificar, quanto à delimitação da palavra escrita, o que ainda é característico do fim da etapa fundamental de escolarização, tanto do que pode ser considerada tendência do corpus quanto do que está mais relacionado à trajetória particular de cada escrevente com a (sua) escrita. Também, sobre a análise da estrutura prosódica, o nosso objetivo é o de identificar a estrutura da palavra escrita convencional e da hipersegmentação mais frequente para, com base nesse levantamento, podermos discutir os tipos de informações linguísticas subjacentes às grafias não-convencionais, que permitem traçar regularidades acerca da escrita investigada.

O segundo momento da análise (subseção “*Aspectos ‘particulares’ das hipersegmentações*”) focaliza os dados que, durante a discussão das regularidades, indicaram outras motivações linguísticas e configurações estruturais, as quais atribuímos a um processo mais individual de aquisição da escrita, uma vez que essas grafias “particulares” podem ser localizadas quase que, predominantemente, nas produções de um único escrevente.

#### 4.1. Aspectos gerais das hipersegmentações

Dos 266 textos produzidos no decorrer dos quatro anos do EF-II foi possível extrairmos 113 grafias de hipersegmentação e mais duas ocorrências de híbridos que, por envolverem, simultaneamente, dois tipos de segmentação não-convencional, não foram classificados junto aos dados que só se caracterizaram pela presença da fronteira gráfica interna à palavra.<sup>58</sup>

A produção do número de textos e dados, em cada um dos anos escolares, é demonstrada na Tabela 1 (em que seguem os números absolutos e seus números percentuais correspondentes):

**Tabela 1:** Distribuição dos textos e das hipersegmentações nos anos escolares do EF-II

Ano escolar	Total de textos		Total de hipersegmentação	
	Nº	%	Nº	%
6º ano	67	25,1%	40	35,3%
7º ano	74	27,9%	24	21,3%
8º ano	61	23,0%	27	24,0%
9º ano	64	24,0%	22	19,4%
<b>Total</b>	<b>266</b>	<b>100%</b>	<b>113</b>	<b>100%</b>

Fonte: Elaboração própria

<sup>58</sup> Os dados de híbrido são: “a olado” (ao lado) e “come cheia” (comecei a) – ambos produzidos no 8º ano pelo escrevente Bruno. Adiantamos que esses dados estão analisados na próxima parte desta Seção.

Pela Tabela 1 é possível observarmos a distribuição das grafias de hipersegmentação, dando destaque para a diminuição no número de dados que ocorre do sexto ano – início do EF-II (40 dados, 35,3%) – para o novo ano – fim do EF-II (22 dados, 19,4%). No entanto, ao comparar, também, esse resultado com o total de dados encontrados nos outros dois anos escolares (sétimo e oitavo anos), notamos que a queda apontada parece não se dar de forma igual, ou seja, regularmente por todo o EF-II, pois, do sétimo para o oitavo ano, há um aumento na quantidade de hipersegmentações (de 24 dados (21,3%) para 27 dados (24%), respectivamente). Em decorrência desse resultado, entendemos a necessidade de investigar mais atentamente o oitavo ano em relação aos demais anos do EF-II, já que, também, outros estudos (ver TENANI, 2010, 2011b), sobre a análise do mesmo tipo de objeto e etapa escolar em análise chegaram a resultados parecidos com o nosso. Desse modo, quais fatores mobilizam essa diferença que, aparentemente, caracteriza os anos escolares do EF-II, no *cópus* investigado?

De modo a avaliarmos a confiabilidade desse primeiro resultado geral, sobre a distribuição das hipersegmentações, para, com base nele, podermos fazer generalizações e, ainda, comparações com outras pesquisas que analisaram o EF-II, realizamos o cotejamento do número total de dados e o *tipo de palavra escrita que deu origem à hipersegmentação* em cada ano. Em outros termos, a comparação se deu entre a palavra, sobre a qual foi proposto o limite não-convencional – por exemplo, a palavra escrita *atrás* –, e o número de hipersegmentação dessa mesma palavra – por exemplo, as ocorrências “a traz”, “a tras”, “a trais”.<sup>59</sup>

Ao considerarmos essa nova variável, nossa hipótese é a de que a elevação do número de dados, no oitavo ano, pode estar relacionada a grafias não-convencionais de um mesmo

---

<sup>59</sup> Exemplos retirados do sétimo e oitavo anos.

tipo de palavra. A partir dessa hipótese (caso venhamos a confirmá-la), acreditamos poder demonstrar que os escreventes analisados nesta pesquisa, embora sigam produzindo muitas ocorrências de hipersegmentação as fazem nas mesmas palavras. Portanto, as dificuldades enfrentadas quanto aos limites de palavra escrita que permanecem nessa etapa escolar tendem a se localizar em certas palavras e não, potencialmente, em qualquer palavra. Para ilustrar essa relação, entre o número de dados e palavra escrita convencional,<sup>60</sup> apresentamos a Tabela 2, a seguir:

**Tabela 2:** Relação número de palavras escritas convencionais e número de hipersegmentações, por ano escolar

Ano escolar	Total de palavras Convencionais		Total de Hipersegmentação	
	Nº	%	Nº	%
6º ano	31	34,1%	40	35,3%
7º ano	21	23,1%	24	21,3%
8º ano	22	24,1%	27	24,0%
9º ano	17	18,7%	22	19,4%
<b>Total</b>	<b>91</b>	<b>100%</b>	<b>113</b>	<b>100%</b>

Fonte: Elaboração própria

A Tabela 2 permite observarmos a relação entre a quantidade de palavras escritas convencionais e grafias de hipersegmentação e, ainda, afirmar que o número de palavras que geraram dúvidas, quanto ao limite gráfico, tende a diminuir com o passar do EF-II. Tanto no que se refere à distribuição do tipo de palavra escrita quanto das hipersegmentações, fica evidente a diferença quantitativa entre o sexto ano e os demais anos do EF-II. Apesar de termos apontado a elevação no número de grafias não-convencionais no oitavo ano, essa

<sup>60</sup> Os números referentes às palavras escritas convencionais foram obtidos por meio da contagem, em cada ano escolar, dos tipos de palavras em que os escreventes realizaram a hipersegmentação. Desse modo, quando, em um mesmo ano, foi identificada mais de uma hipersegmentação de um só tipo de palavra convencional, essa foi contabilizada apenas uma vez.

variação é muito menor quando comparada ao número de dados do sexto ano com quaisquer dos outros três. Esse estatuto do sexto ano, nos leva a supor que, em alguma medida, novas informações linguísticas introduzidas na nova etapa do EF podem ser um “gatilho” para que os escreventes formulem novas hipóteses sobre a escrita. Como observamos na Tabela 2, talvez, com o passar dos anos escolares, essas hipóteses sejam sistematizadas e se aproximem da convenção. Informamos que uma investigação exclusiva sobre dados de segmentação não-convencional presentes em textos do sexto ano já foi realizada por Paranhos (2010) e sintetizada em Paranhos e Tenani (2011), de modo que, nesta pesquisa, nos debruçamos a encontrar indícios que nos auxiliem na caracterização das hipersegmentações dos anos do EF-II como um todo.

Voltando à diferença expressa na Tabela 2 sobre a quantidade do tipo de palavra e do número de hipersegmentação, correspondente em cada ano, há palavras constantes, as quais possibilitam mais números de dados. No cópuz, ao todo, foram nove palavras que, em um mesmo ano e/ou entre diferentes anos, originaram mais de uma grafia de hipersegmentação. Essas palavras são: *demais*, *dele*, *aquele*, *queria*, *embora*, *comigo*, *depois*, *enquanto* e *atrás*. Merece comentário, particularizado, a esse respeito, a ocorrência da palavra *demais*, que, em todos os anos, foi responsável por 12 grafias não-convencionais (equivalente à 10,61% do total de dados). Particularmente, a palavra *demais* ocasionou dois dados, no sexto ano; outros dois dados, no sétimo ano; cinco ocorrências, no oitavo ano; e, por fim, três dados no nono ano. Quando excluídas as grafias não-convencionais dessa palavra, em cada ano, a quantidade final de dados cai, respectivamente para: **38**, **22**, **22**, e **19 dados**. Desse modo, as diferenças quanto ao número de dados ao longo do ano escolar, antes observadas, desaparecem quando se excluem da análise aqueles dados que sempre se repetem. Esse resultado nos leva a endossar a necessidade de, na análise desse tipo de dado, ter um olhar atento que leva em

conta não só a quantidade das ocorrências não-convencionais, mas fundamentalmente os tipos dessas ocorrências na relação com a palavra convencional que a origina.

Assim, verificamos que as grafias de hipersegmentação, nos anos que encerram o EF (pelo menos no cópuz analisado), diminuem com o passar dos anos escolares; mas não só elas: o número de palavras cujos limites se mostram como um desafio aos escreventes também decai. Nesse aspecto, podemos inferir que a imagem dos escreventes de que, na escrita, as palavras são marcadas por limites (dados ora pelo branco ora pelo hífen) se altera, aproximando-se mais da convenção ortográfica. Acreditamos que esse resultado talvez se deva em função das atividades letradas/escritas das quais os escreventes participam, no decorrer de todo o EF-II, nas aulas de Língua Portuguesa, privilegiadamente. Entretanto, não negamos, com essa afirmação, a inserção dos escreventes em outras práticas sociais de linguagem relevantes para a representação que construíram para a (sua) escrita. Pelo contrário, uma vez que concordamos com o fato de todo texto ser, sempre, resultado de diferentes práticas orais/faladas e letradas/escritas (CORRÊA, 2004). Nesse sentido, identificamos evidências de a escrita de nossos escreventes “mostrar-se, com mais clareza, como produto da inserção do aprendiz em práticas de letramento (dentro e fora do contexto escolar) que privilegiam um conhecimento metalinguístico sobre a língua em seu modo de enunciação escrito” (CHACON, 2005, p. 83).

Dado o caráter longitudinal deste estudo, na sequência, trazemos a Tabela 3, na qual constam informações sobre a distribuição dos textos e dados produzidos por cada um dos escreventes investigados:

**Tabela 3:** Distribuição do número de textos e hipersegmentações produzidos, por escrevente

Escreventes	Total de textos		Total de hipersegmentação	
	Nº	%	Nº	%
André	25	9,4%	2	1,8%
Viviane	23	8,7%	2	1,8%
Ana	20	7,6%	4	3,6%
Henrique	22	8,2%	4	3,6%
Camila	25	9,4%	4	3,6%
Fernando	20	7,6%	5	4,4%
Julia	25	9,4%	7	6,1%
Mateus	25	9,4%	7	6,1%
Maria	21	7,9%	9	8,0%
Lucas	24	9,0%	10	8,8%
Pedro	17	6,3%	22	19,4%
Bruno	19	7,1%	37	32,8%
<b>Total</b>	<b>266</b>	<b>100%</b>	<b>113</b>	<b>100%</b>

Fonte: Elaboração própria

Tratando do exposto na Tabela 3, abordamos as diferenças quanto à quantidade de hipersegmentação entre os escreventes. Notamos, a esse respeito, que o total de dados na relação intersujeitos apresenta grande variação entre si, diferenças que vão desde a quantidade máxima de 37 ocorrências (identificadas nas produções textuais do escrevente Bruno) até a quantidade mínima de dois dados (extraídos dos textos dos estudantes André e Viviane), uma variação que, nesse caso, é de 31%, em termos percentuais. Relacionado a esse fato, vale ressaltar, a propósito da relação entre o número de textos produzidos e o respectivo número de ocorrências, que os dois escreventes que mais produzem marcas de hipersegmentação de palavras (cf. Bruno: 37 ocorrências; e Pedro: 22 ocorrências) são os que redigem a menor quantidade de textos em todos os anos escolares (19 e 17 textos, respectivamente), enquanto que os dois alunos que apresentam menos dados (cf. André: duas ocorrências; e Viviane: duas ocorrências) compreendem os escreventes que produzem a maioria das propostas de textos aplicadas ao longo dos quatro anos do EF-II (respectivamente, 25 e 23 textos).

Duas hipóteses podem ser levantadas com base nesses fatos (de modo que uma não exclui a outra). A primeira é a de essas diferenças estarem ancoradas nos modos de participação dos estudantes em práticas sociais de linguagem, as quais os levam a elaborar representações sobre a escrita (sobretudo, aquela privilegiada pela instituição escolar) e que guiam os escreventes no uso das convenções ortográficas em seus textos (particularmente, do registro de fronteiras de palavras, em nosso trabalho). A segunda hipótese tem relação com a extensão dos textos produzidos. Nesse aspecto, o fato de um escrevente ter feito mais grafias hipersegmentadas se deva, em alguma medida, em função de ele ter escrito textos com mais palavras, o que amplia a possibilidade de aparecimento das grafias não-convencionais; como, na mesma direção, o motivo que tenha levado outro escrevente a grafar poucas ocorrências de hipersegmentação esteja ligado ao pequeno tamanho de seus textos e, nesse caso, restringindo as possibilidades de produção das hipersegmentações.

A fim de verificar a validade dessa última hipótese, observamos a quantidade total de palavras redigidas nos quatro anos escolares em proporção ao número, também total, de dados de hipersegmentação. O resultado a que chegamos está organizado na Tabela 4.

**Tabela 4:** Número de hipersegmentação em relação ao número de palavras por escrevente

Escreventes	Total de			% da hipersegmentação em relação ao número de palavras
	Textos	Palavras	Hipersegmentação	
André	25	2.122	2	0,09%
Viviane	23	3.196	2	0,06%
Ana	20	3.457	4	0,11%
Henrique	22	2.946	4	0,13%
Camila	25	3.261	4	0,12%
Fernando	20	1.759	5	0,28%
Julia	25	3.619	7	0,19%
Mateus	25	3.469	7	0,20%
Maria	21	3.774	9	0,23%
Lucas	24	3.231	10	0,30%
Pedro	17	1.588	22	1,38%
Bruno	19	1.847	37	2,00%
<b>Total</b>	<b>266</b>	<b>34.269</b>	<b>113</b>	<b>0,32%</b>

Fonte: Elaboração própria

A Tabela 4 permite observarmos a relação entre os escreventes que produzem mais/menos grafias de hipersegmentação, agora, a partir da extensão dos textos em função do número de palavras escritas. Em termos gerais, as marcas de hipersegmentação representam 0,32% de tudo o que foi escrito pelos 12 escreventes, o que indica que as fronteiras de palavras marcadas além do esperado pela ortografia, embora persistentes nos anos que encerram o EF, configuram-se em pequena quantidade nessa etapa escolar.

Referente à trajetória escrita de cada escrevente, o levantamento do número de palavras dos textos ressalta que as diferenças intersujeitos estão relacionadas, diretamente, ao quanto que cada aluno escreveu, pois há:

- i. escreventes que grafam quantidade idêntica de hipersegmentações e próximas de textos (ver André e Viviane: dois dados em 25 e 23 textos; Ana, Henrique e Camila: quatro dados em 20, 22 e 25 textos; Julia e Mateus: sete dados em 25 textos cada um), mas que apresentam percentuais distintos de representatividade dos dados em relação a toda sua produção ao longo do EF-II (ver André e Viviane: 0,09% e 0,06%; Ana, Henrique e Camila: 0,11%, 0,13% e 0,12%; Julia e Mateus: 0,19% e 0,20%);
- ii. escreventes que produzem mais dados em função do maior número de palavras redigidas (ver Julia, Mateus e Maria: sete, sete, e nove dados em 3.619, 3.469 e 3.774 palavras respectivamente) em contraste com alunos que produzem poucas ocorrências em virtude da menor quantidade de palavras escritas (ver André e Viviane: dois dados cada um em 2.122, e 3.196 palavras);
- iii. escreventes que se destacam pela alta representatividade das hipersegmentações na comparação com o tanto que escrevem em todo o EF-II (ver Fernando: 0,28%; Lucas: 0,30%; Pedro: 1,38% e Bruno: 2%).

Corroboramos para que afirmemos serem esses escreventes os que mais têm dúvidas sobre os limites de palavra, a consideração do número de palavras escritas em proporção ao total de textos produzidos, na medida em que o total de palavras foi baixo em comparação aos outros alunos que realizam uma quantidade de textos idêntica ou com pouca variação. A título de exemplificação, Ana (que redige, ao todo, 20 textos) grafa 1.698 palavras a mais que Fernando (que faz 20 textos) e 1.610 palavras a mais que Bruno (que elabora 19 textos). Nota-se, pois, que, embora Fernando, Lucas, Pedro e Bruno tenham participado com grande frequência das atividades de elaboração dos textos, suas produções, por terem extensão curta, deixam mais aparente o conflito de quando as palavras devem ter autonomia gráfica. Mais especificamente, ao observarmos a distribuição das hipersegmentações na escrita de Pedro e de Bruno, podemos presumir que, para esses dois escreventes, definir fronteira para as palavras na escrita é um desafio ainda maior do que para os demais escreventes. Até o fim desta análise, esperamos apresentar evidências que auxiliem na compreensão do(s) tipo(s) de fato(s) linguístico(s) que é (são) mobilizado(s) nas hipóteses desses dois escreventes e que fazem com que eles venham a se particularizar em relação ao restante dos escreventes.

Na sequência, vejamos a distribuição dos totais de palavras escritas e de dados produzidos, por cada escrevente e ano escolar:

**Tabela 5:** Número de palavras escritas e hipersegmentações por escrevente/ano escolar

Escreventes	Nº de palavras/ano escolar				Nº de hipersegmentações/ano escolar			
	6º ano	7º ano	8º ano	9º ano	6º ano	7º ano	8º ano	9º ano
André	555	500	444	623	1	0	0	1
Viviane	773	839	780	804	1	1	0	0
Ana	571	1.078	578	1.230	2	2	0	0
Henrique	478	853	637	978	1	2	0	1
Camila	520	980	681	1.080	2	0	0	2
Fernando	470	360	338	591	4	0	0	1
Julia	941	850	684	1.144	2	0	1	4
Mateus	512	902	1.026	1.029	1	1	4	1
Maria	867	914	984	1.009	3	1	3	2
Lucas	759	781	747	944	3	2	3	2
Pedro	319	323	336	610	6	1	11	4
Bruno	591	349	447	460	14	14	5	4
<b>Total</b>	<b>7.356</b>	<b>8.729</b>	<b>7.682</b>	<b>10.502</b>	<b>40</b>	<b>24</b>	<b>27</b>	<b>22</b>

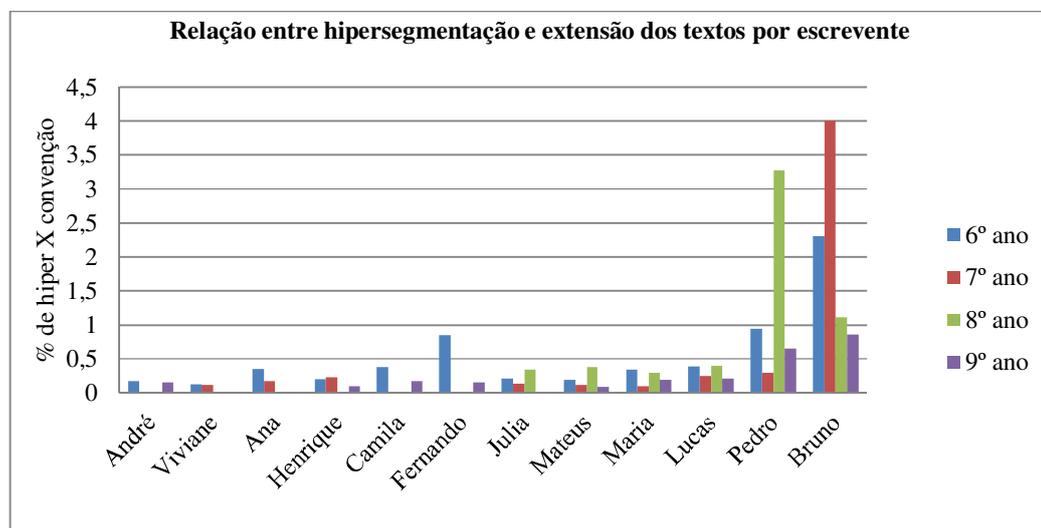
Fonte: Elaboração própria

Analisando a Tabela 5, no que diz respeito aos anos letivos, identificamos três trajetórias dos escreventes quanto à produção dos registros não-convencionais, por presença, das fronteiras de palavras. A *primeira* é a de alunos que produzem hipersegmentações em todos os anos. A *segunda* trajetória mostra alunos que têm ocorrências em três dos quatro anos. E, por fim, a *terceira* aponta alunos que apresentam dados em dois dos quatro anos escolares. Sobre a divisão dos escreventes no interior dessas trajetórias, constatamos que:

- (i) cinco alunos se encaixam na primeira delas, sendo, também, os que mais produzem hipersegmentações;
- (ii) dois alunos se enquadram na segunda trajetória, sendo: um dos alunos com ocorrências no sexto, oitavo e nono anos e o outro no sexto, sétimo e nono anos, correspondendo ao segundo conjunto de alunos que mais apresentam ocorrências; e
- (iii) cinco alunos se identificam na terceira trajetória, sendo: dois dos alunos com ocorrências no sexto e sétimo anos e os outros três no sexto e nono anos, configurando-se ao grupo de escreventes que produzem o menor conjunto de dados.

Em síntese, o desenho das trajetórias revela que apenas dois escreventes concluíram o EF sem apresentar mais grafias de hipersegmentação em seus textos, enquanto a outra parte dos escreventes terminou a etapa do EF oscilando na escrita convencional/não-convencional das fronteiras de palavra (particularmente daqueles limites marcados além do previsto). Entretanto, mesmo que as hipersegmentações não tenham cessado de aparecer na produção escrita da maioria dos escreventes deste estudo, muitos desses, ao fim, se aproximam da convenção ortográfica. Essa afirmação fica mais bem fundamentada quando se põe em análise a relação entre as grafias de hipersegmentação e a extensão dos textos elaborados no transcorrer do EF-II. Essa relação é visualizada por meio do gráfico a seguir, cujos percentuais foram obtidos a partir dos números da Tabela 5:

**Gráfico 1:** Relação entre hipersegmentação e extensão dos textos (pelo número de palavras), por escrevente



Fonte: Elaboração própria

Do gráfico apresentado, destacamos que:

- i. independentemente do número de dados produzidos, a mesma quantidade de grafias não-convencionais ganha diferentes correspondências percentuais, ao serem relacionadas com o número total de palavras que cada escrevente escreve individualmente em cada ano escolar;
- ii. na trajetória pelo EF-II, o valor das ocorrências dos escreventes André, Viviane, Ana, Camila, Fernando e Bruno (metade dos sujeitos investigados) diminui quando cotejadas com a extensão dos textos, apontando que a escrita desses escreventes foi se encaminhando em direção às normas ortográficas que regem o estabelecimento dos limites de palavra escrita;
- iii. no caso em que o percentual de dados não decaiu com o passar dos anos escolares, encontramos duas possíveis hipóteses. A primeira parece permear a escrita de Henrique e Julia, e pode ter relação com o também aumento no *número de palavras escritas* nos anos em que a quantidade de hipersegmentações foi maior. Já a segunda hipótese é a de que, a elevação do número de dados se explica em função dos tipos de palavra escrita que originam as hipersegmentações, e é o que se mostra na escrita de Mateus, Maria, Lucas e Pedro.

Buscando explicitar o tipo de palavra mobilizada durante a proposição da fronteira não-convencional, apresentamos, nos Quadros 2 a 13, a seguir, todos os dados de cada escrevente separados por ano escolar:

**Quadro 2:** Grafias de hipersegmentação produzidas por André

ANDRÉ			
6º ano	7º ano	8º ano	9º ano
e norme ( <i>enorme</i> )			amasa-se ( <i>amansasse</i> )

Fonte: Elaboração própria

**Quadro 3:** Grafias de hipersegmentação produzidas por Viviane

VIVIANE			
6º ano	7º ano	8º ano	9º ano
se nários ( <i>cenários</i> )	com migo ( <i>comigo</i> )		

Fonte: Elaboração própria

**Quadro 4:** Grafias de hipersegmentação produzidas por Ana

ANA			
6º ano	7º ano	8º ano	9º ano
quise-se ( <i>quisesse</i> ) de mais ( <i>demais</i> )	de mais ( <i>demais</i> ) mata-do ( <i>matado</i> )		

Fonte: Elaboração própria

**Quadro 5:** Grafias de hipersegmentação produzidas por Henrique

HENRIQUE			
6º ano	7º ano	8º ano	9º ano
ganha-se ( <i>ganhasse</i> )	estava-mos ( <i>estávamos</i> ) de ele ( <i>dele</i> )		a onde ( <i>aonde</i> )

Fonte: Elaboração própria

**Quadro 6:** Grafias de hipersegmentação produzidas por Camila

CAMILA			
6º ano	7º ano	8º ano	9º ano
de ele ( <i>dele</i> ) deis de ( <i>desde</i> )			conversa-se ( <i>conversasse</i> ) arranja-se ( <i>arranjasse</i> )

Fonte: Elaboração própria

**Quadro 7:** Grafias de hipersegmentação produzidas por Fernando

FERNANDO			
6º ano	7º ano	8º ano	9º ano
em fachado ( <i>enfaixado</i> ) em tão ( <i>então</i> ) na quele ( <i>naquele</i> ) vou tando ( <i>voltando</i> )			de mais ( <i>demais</i> )

Fonte: Elaboração própria

**Quadro 8:** Grafias de hipersegmentação produzidas por Julia

JULIA			
6º ano	7º ano	8º ano	9º ano
a caba ( <i>acabar</i> ) anti penúltimo ( <i>antepenúltimo</i> )		vira-se ( <i>virasse</i> )	de mais ( <i>demais</i> ) a tirar ( <i>atirar</i> ) de pois ( <i>depois</i> ) a cender ( <i>acender</i> )

Fonte: Elaboração própria

**Quadro 9:** Grafias de hipersegmentação produzidas por Mateus

MATEUS			
6º ano	7º ano	8º ano	9º ano
a quele ( <i>aquela</i> )	da quele ( <i>daquela</i> )	loca mente ( <i>loucamente</i> ) de mais ( <i>demais</i> ) de mais ( <i>demais</i> ) e moção ( <i>emoção</i> )	de mais ( <i>demais</i> )

Fonte: Elaboração própria

**Quadro 10:** Grafias de hipersegmentação produzidas por Maria

MARIA			
6º ano	7º ano	8º ano	9º ano
de se ( <i>dessa</i> ) a panho ( <i>apanhou</i> ) a gora ( <i>agora</i> )	de mais ( <i>demais</i> )	su por ( <i>supor</i> ) nem um ( <i>nenhum</i> ) de mais ( <i>demais</i> )	contra bando ( <i>contrabando</i> ) contra bando ( <i>contrabando</i> )

Fonte: Elaboração própria

**Quadro 11:** Grafias de hipersegmentação produzidas por Lucas

LUCAS			
6º ano	7º ano	8º ano	9º ano
em quanto ( <i>enquanto</i> ) em cando ( <i>enquanto</i> ) des de ( <i>desde</i> )	da qui ( <i>daqui</i> ) da quele ( <i>daquela</i> )	em bora ( <i>embora</i> ) em tão ( <i>então</i> ) fala-sem ( <i>falassem</i> )	na quele ( <i>naquela</i> ) a quele ( <i>aquela</i> )

Fonte: Elaboração própria

**Quadro 12:** Grafias de hipersegmentação produzidas por Pedro

PEDRO			
6º ano	7º ano	8º ano	9º ano
de ele ( <i>dele</i> ) a cabano ( <i>acabando</i> ) a quele ( <i>aquela</i> ) a perecer ( <i>aparecer</i> ) na viu ( <i>navio</i> ) de sero ( <i>desceram</i> )	foute i ( <i>voltei</i> )	ir motal ( <i>imortal</i> ) a traís ( <i>atrás</i> ) a liversario ( <i>aniversário</i> ) a manha ( <i>amanhã</i> ) em comtreí ( <i>encontrei</i> ) na quéla ( <i>naquela</i> ) de mais ( <i>demais</i> ) a migo ( <i>amigo</i> ) em Bora ( <i>embora</i> ) a pertou ( <i>apertou</i> ) a paichonei ( <i>apaixonei</i> )	de pois ( <i>depois</i> ) a doró ( <i>adoramos</i> ) o vino ( <i>ouvindo</i> ) a fogo ( <i>afogo</i> )

Fonte: Elaboração própria

**Quadro 13:** Grafias de hipersegmentação produzidas por Bruno

BRUNO			
6º ano	7º ano	8º ano	9º ano
so zinho ( <i>sozinho</i> ) que ria ( <i>queria</i> ) que ria ( <i>queria</i> ) a caba ( <i>acabar</i> ) com pra ( <i>comprar</i> ) com versa ( <i>conversa</i> ) a i ( <i>aí</i> ) a i ( <i>aí</i> ) adísio o na ( <i>adiciona</i> ) se não ( <i>senão</i> ) es quito ( <i>escrito</i> ) de mais ( <i>demais</i> ) de ele ( <i>dele</i> ) com panhate ( <i>acompanhante</i> )	a vó ( <i>avó</i> ) falam do ( <i>falando</i> ) que ria ( <i>queria</i> ) vou to ( <i>voltou</i> ) a trás ( <i>atrás</i> ) a trás ( <i>atrás</i> ) recom pensa ( <i>recompensa</i> ) a traz ( <i>atrás</i> ) a paresido ( <i>aparecido</i> ) qual do ( <i>quando</i> ) a te ( <i>até</i> ) em teiro ( <i>inteiro</i> ) a gora ( <i>agora</i> ) mini cidade ( <i>minicidade</i> )	de mais ( <i>demais</i> ) com migo ( <i>comigo</i> ) a lugar ( <i>alugar</i> ) de pois ( <i>depois</i> )	com ver sando ( <i>conversando</i> ) com versando ( <i>conversando</i> ) com ver sanos ( <i>conversamos</i> ) a sim ( <i>assim</i> )

Fonte: Elaboração própria

Nos Quadros ora apresentados, verificamos as ocorrências de hipersegmentação, bem como as palavras escritas que lhes deram origem. A partir dessas duas informações, fica clara a oscilação dos escreventes em grafar os mesmos tipos de palavras e, em particular, de palavras que, entre si, se assemelham quanto à distribuição das sílabas. Tratando, pois, da relação entre tipo de palavra e hipersegmentação na trajetória escrita de cada escrevente, averiguamos que, somente André e Viviane realizam duas hipersegmentações de duas

diferentes palavras. Os dados desses escreventes, como observado anteriormente, apresentam pouca representatividade em relação a toda sua produção no decorrer do EF-II; no entanto, entendemos a presença dessas grafias não-convencionais como uma pista relevante de momentos pontuais da relação desses escreventes com a busca por “descobrir o que vem a ser uma palavra na língua” (ABAURRE, 1991c, p. 209), já que, tanto nos dados de André quanto nos de Viviane não há regularidade no registro nem de uma mesma palavra nem de palavras que de alguma forma se aproximam.

Cabe salientar que a mesma conclusão não se aplica aos outros escreventes. Em termos gerais, primeiramente, constatamos que a presença dos limites não-convencionais ocorrem, sobretudo, em palavras cujas sílabas pretônicas se identificam com elementos escritos da língua, como: “a” (30 hipersegmentações a exemplo de “a gora”), “de” (23 dados como em “de pois”), “em” (nove ocorrências do tipo “em quanto”) e “com” (oito hipersegmentações à semelhança de “com pra”). No caso específico das separações não-convencionais de “de”, destacamos a palavra *demais* como a mais hipersegmentada, não apenas para os dados desse item, mas como predominante em todo o corpus; o que equivale a dizer que o limite da palavra *demais* é o que se mostra mais instável na escrita de nossos escreventes. Por exemplo, dos sete dados produzidos por Mateus, três são dessa palavra, significando que somente cinco palavras escritas geram dúvida a esse escrevente. A mesma palavra ocasiona, ainda, mais de uma grafia não-convencional, nos textos de Ana, Maria e Bruno. Especificamente, na escrita de Ana, a hipersegmentação “de mais” foi a única que ocorreu por espaço em branco, enquanto que as outras ocorreram por uma dificuldade em distinguir sequências de verbo+pronome e formas verbais do subjuntivo. Quanto ao hífen, delimitar palavras por meio desse recurso mostra ser uma dificuldade, também, para Henrique e Camila, pois metade dos dados produzidos por esses escreventes são pelo emprego do hífen.

Retornando aos tipos de palavras em que identificamos a separação não-convencional pelo espaço em branco, as hipersegmentações com “a” foram, além das mais frequentes, as que, também, demonstram mais diversidade de palavras escritas.<sup>61</sup> Dados, em que o “a” esteve grafado entre limites, foram encontrados nos textos de Henrique, Julia, Mateus, Maria, Lucas, Pedro e Bruno, isto é, em mais da metade dos escreventes do corpus. Especialmente sobre as hipersegmentações de Pedro e Bruno, a atribuição do “a” como uma palavra autônoma representou 50% e 29,7%, respectivamente, do total de dados desses escreventes.

As hipersegmentações em que identificamos a partícula “em” se marcaram na produção dos escreventes Fernando, Lucas e Pedro. Em nossa interpretação, esses registros ressaltam a identificação dos escreventes com palavras próprias da escrita (e que, portanto, deve ser demarcada entre limites), na medida em que todas são grafadas com o grafema <m> quando, as palavras que dão origem a essas hipersegmentações, são, convencionalmente, grafadas com <n> (com exceção da palavra *embora*). O mesmo comentário estende-se aos dados que demarcam fronteira para a possível palavra “com”. Das oito ocorrências não-convencionais desse elemento, sete são localizadas na escrita de Bruno, as quais se dividem em toda a etapa do EF-II, sinalizando, prioritariamente, a imprecisão desse escrevente quanto ao registro gráfico de “com” ou como sílaba ou como palavra independente.

Seguimos com a análise das características gerais das hipersegmentações, passando a tratar, neste momento, da distribuição dos recursos gráficos. Por adotarmos a proposta de Tenani (2011b) sobre a classificação das segmentações não-convencionais de palavras quanto ao tipo de recurso gráfico empregado, optamos por analisar separadamente as ocorrências que se caracterizam pela presença do espaço em branco e as que se caracterizam pela presença do hífen. A frequência das hipersegmentações, nos anos escolares, pelos dois recursos gráficos é dada na Tabela 6:

---

<sup>61</sup> Nota-se, sobre essa variedade, a ocorrência de mais de uma grafia não-convencional apenas das palavras *atrás*, *aí* (ver Bruno e Pedro) e *aquela* (ver Mateus e Pedro).

**Tabela 6:** Classificação das hipersegmentações quanto ao recurso gráfico nos anos escolares

Ano escolar	Recurso gráfico				Total	
	Espaço em branco		Hífen		Nº	%
6º ano	38	33,7%	2	1,8%	40	35,3%
7º ano	22	19,4%	2	1,8%	24	21,3%
8º ano	25	22,1%	2	1,8%	27	24,0%
9º ano	19	16,8%	3	2,6%	22	19,4%
<b>Total</b>	<b>104</b>	<b>92%</b>	<b>9</b>	<b>8%</b>	<b>113</b>	<b>100%</b>

Fonte: Elaboração própria

Considerando, então, a organização das ocorrências pelo tipo de recurso gráfico, constatamos que as presenças não-convencionais dos limites gráficos de palavra ocorrem, em todos os anos, tanto pelo espaço em branco quanto pelo hífen. Esse resultado sinaliza que os escreventes estiveram atentos aos modos convencionais possíveis de delimitação das palavras na escrita, bem como quais são as estruturas que se demarcam por cada um dos recursos. Em outras palavras, os dados, por presença do hífen e do espaço em branco, dão pistas de quais tipos de palavras devem ser separadas por um tipo ou outro de recurso na escrita convencional – mas, sobre esse aspecto, discorreremos mais adiante.

Antes, porém, comentamos as diferenças, da Tabela 6, sobre a distribuição quantitativa das ocorrências. Do total das marcas de hipersegmentação, 92% teve por base o uso não-convencional do espaço em branco, enquanto que 8% correspondem à colocação não-convencional do hífen. Embora, nos dados do cópuz, a diferença numérica entre o uso de um ou outro recurso gráfico tenha sido bastante elevada (como já era previsto), não descartamos a relevância do estudo do emprego do hífen separadamente em relação ao realizado pelo espaço em branco, uma vez que aqueles usos não-convencionais revelam aspectos linguísticos importantes sobre o processo de letramento dos escreventes (TENANI, 2011b). Comentamos, quanto ao emprego do hífen ser índice de uma característica letrada que perpassa a escrita dos escreventes, o fato de a delimitação não-convencional das palavras realizada por meio do

espaço em branco (recurso predominante de indicação de limites de palavras) ter diminuído com o passar dos anos escolares do EF-II, enquanto que os usos não-convencionais do hífen (recurso utilizado delimitar poucos tipos de palavras<sup>62</sup>) se mantém inalterada por quase todo o período investigado (notando uma pequena elevação do oitavo para o nono ano). Entendemos esse resultado como um indicativo de que a eleição do hífen para propor limites às palavras é um problema que caracteriza a escrita de escreventes com mais anos de escolarização formal, tanto quando se analisa o emprego desse recurso por um período longitudinal (como nesta pesquisa), quanto quando se observa esse mesmo uso em uma amostra transversal, conforme constatado por Tenani (2011b).

Todos os usos de hífen levantados nesta pesquisa se concentram na tentativa de grafias com a estrutura de *verbo+pronome enclítico*. Sabendo que, nos enunciados falados, o PB privilegia a próclise de elementos (*te quero* ao invés de *quero-te* – cf. BISOL, 1996, 2000, 2005), é interessante notar que o registro do hífen se deu em tipos de palavras que remetem a uma estrutura cujo funcionamento é ensinado na escola e, sobretudo, por ela valorizado. Mas essa afirmação ganha contornos mais precisos se considerarmos que, das nove ocorrências de hífen, seis envolvem o uso da partícula *se*: “amasa-se” (*amansasse*), “quise-se” (*quisesse*), “ganha-se” (*ganhasse*), “conversa-se” (*conversasse*), “arranja-se” (*arranjasse*), “vira-se” (*virasse*); formas, em sua grande maioria, existentes no funcionamento da escrita convencional, mas que, nesses casos, se distanciam da convenção por estarem ocupando a função sintática de formas verbais do tempo subjuntivo. Se, uma vez mais, tomarmos para reflexão os usos orais/falados da linguagem, chegamos à observação de que o uso do subjuntivo já é, nos dias atuais, praticamente inexistente, especialmente nas formas mais cotidianas dessas práticas. Quanto a estruturas como “ganha-se”, então, o que parece mobilizá-las é o forte impacto de práticas letradas/escritas, mais institucionalizadas (e nesse

---

<sup>62</sup> Enquanto recurso indicador de fronteira de palavra, o hífen sinaliza, na escrita, palavras compostas (*meio-dia*) e colocação enclítica do pronome em relação ao verbo (*diga-lhe*).

grupo incluímos a prática escolar), que requerem o uso de um tipo de escrita baseada em uma norma culta padrão, e, também, como as que, com frequência, encontramos veiculadas em anúncios como os de “*vendem-se*”, de “*alugam-se*”, dentre outros, que, certamente, geram grande impacto.

Assim, junto à perspectiva de escrita assumida nesta dissertação (CORRÊA, 2004), defendemos que os usos não-convencionais do hífen, em estruturas como as que estamos analisando, são resultado do imaginário social do escrevente sobre a escrita, construído pela sua inserção nos mais diversos usos (formais/informais) da linguagem. Acreditamos, de modo mais específico, a aproximação desses momentos com uma projeção do que seja a representação do código escrito institucionalizado. Com essa hipótese, porém, não excluimos, da análise que realizamos, a ação dos outros dois momentos de circulação dialógica do escrevente (o da gênese da escrita e o da dialogia sobre o já falado/ouvido e escrito/lido), pois, ao adotarmos as formulações de Corrêa, assumimos, junto com o autor, ser possível apenas uma separação metodológica dos eixos, visto que um texto escrito é sempre constituído a partir da relação dialógica entre os três.

Em síntese, a descrição realizada até esta parte foi na direção de darmos um entendimento geral a respeito da distribuição dos dados em relação aos anos escolares, aos escreventes selecionados e aos tipos de recursos gráficos delimitadores de palavras. Prosseguindo para a segunda etapa descritiva da análise, passamos a investigar as características gerais das hipersegmentações no que diz respeito às suas configurações rítmicas, adotando como ponto de partida, para tanto, as estruturas das palavras escritas convencionalmente. Nossa opção por essa organização é dada em função da premissa de que as informações rítmicas, que se formam no interior da palavra por meio de proeminências prosódicas, parecem ser pontos de ancoragem para o tipo de limite gráfico não-convencional de palavras registrado na escrita.

A partir da Tabela 7,<sup>63</sup> verificam-se, com base na organização proposta, sete Tipos de estruturas rítmicas de palavras convencionais que, ao todo, geraram 14 Subtipos de estruturas rítmicas das grafias hipersegmentadas. Há, também, mais um Tipo que nomeamos “outros” e sob o qual incluímos hipersegmentações cujas fronteiras não se assemelham, em termos estruturais, às tendências observadas.

Para que fosse possível observarmos, longitudinalmente, regularidades de distribuição dos dados, quanto ao funcionamento de cada estrutura rítmica, consideramos os mesmos Tipos em todos os anos escolares. Cabe informar o critério de definição dos Tipos: as estruturas rítmicas foram agrupadas tendo em conta o número de sílabas da palavra convencional. Desse modo: os tipos (1) e (2) consistem em palavras dissílabas; os tipos (3) e (4) dizem respeito às palavras trissílabas; e os tipos (5), (6) e (7) correspondem às palavras polissílabas.

---

<sup>63</sup> A interpretação que fizemos dos dados como a combinação entre  $\sigma+\Sigma$  pode ser entendida, também, como sendo resultado de uma combinação entre  $\sigma+\omega$ , uma vez que, no modelo prosódico de Nespor e Vogel (1986), aqui utilizado, a palavra fonológica se forma pelas sílabas agrupadas no interior de um pé, o qual carrega o acento primário, seu principal definidor. Nessa linha interpretativa, poderíamos considerar que uma  $\omega$  teria sido registrada como um  $C$  (cf. sobre essa análise Tenani (2009a, 2011b), vistos na primeira Seção do trabalho e, também, Paranhos e Tenani (2011)). Nesse sentido, nossos dados não se restringem a uma só possibilidade de interpretação com base em uma informação de natureza prosódica.

Tabela 7: Estruturas rítmicas de palavras e de hipersegmentações por ano escolar<sup>64</sup>

Estruturas rítmicas				Exemplo de hiper	Hiper/ ano escolar				Total	
Tipos	Palavra	Subtipos	Hiper		6º	7º	8º	9º	Nº	%
(1)	(•*)	(1.1)	(•)σ (*)σ	“em tão”	8	9	9	6	32	28,3
		(1.2)	(*)σ (•)σ	“nem um”	0	0	1	0	1	0,8
(2)	(*•)	(2.0)	(*)σ (•)σ	“de se”	6	2	0	0	8	7,0
(3)	(•*•)	(3.1)	(•)σ (*•)Σ	“em quanto”	16	6	7	6	35	31,0
		(3.2)	(•*)Σ (•)σ	“quise-se”	2	2	2	0	6	5,3
		(3.3)	(*)σ (*•)Σ	“so zinho”	1	0	0	0	1	0,8
(4)	(*•*)	(4.0)	(*)σ (•*)Σ	“a panho” <sup>65</sup>	1	0	4	1	6	5,3
(5)	(*•••)	(5.1)	(*•)σ (*•)Σ	“recom pensa”	0	1	1	2	4	3,6
		(5.2)	(*)σ (•*•)Σ	“em fachado”	2	0	0	0	2	1,8
		(5.3)	(*•*)Σ (•)σ	“conversa-se”	0	0	0	3	3	2,7
(6)	(*••*) (•*•*)	(6.0)	(*)σ (••*)Σ	“a paichonei”	1	0	1	0	2	1,8
			(•)σ (*•*)Σ							
(7)	(*••••) (•*•*•)	(7.0)	(*)σ (••••)Σ	“a parecido”	0	1	1	0	2	1,8
			(•)σ (*•••)Σ							
(8)	Outros			“adísio o na”	3	3	1	4	11	9,8
<b>Total</b>					<b>40</b>	<b>24</b>	<b>27</b>	<b>22</b>	<b>113</b>	<b>100</b>

Leia-se: \*: sílaba forte; •: sílaba fraca; σ: sílaba; Σ: pé métrico;

Fonte: Elaboração própria

De acordo com a Tabela 7, de modo geral:

- i. um mesmo Tipo de estrutura rítmica de palavra possibilita diferentes grafias não-convencionais, com destaque para os Tipos (1) e (3) que, juntos, somam um percentual de 66,2% do total de hipersegmentação;

<sup>64</sup> Sobre a organização da Tabela 7, cabem, ainda, duas explicações: (i) a respeito de diferentes atribuições de sílabas forte ou fraca para as grafias de hipersegmentação tendo em consideração essa mesma organização da palavra convencional. Destacamos, a esse respeito, os Subtipos (1.2) e (3.3). Quanto ao primeiro, há, como se pode observar, uma inversão na posição da sílaba forte (localizada à esquerda) em relação à posição dessa sílaba na palavra ortográfica (localizada à direita). Em se tratando do Subtipo (3.3), ocorre, em nossa interpretação, o fato de que a sílaba pretônica *so*, da palavra *sozinho*, possa ter sido projetada como uma possível palavra escrita monossilábica tônica, ou seja, portadora de acento, o que justifica termos atribuído um acento à sílaba *so* da grafia hipersegmentada. Em ambos os casos, compreendemos que houve uma reorganização, nas hipersegmentações, na posição das proeminências de uma sequência fônica. Na análise dessa organização dos dados proposta na Tabela 7, procuramos mostrar que outras informações linguísticas, que não apenas as prosódicas, possam ter contribuído para essa reorganização rítmica notada nas hipersegmentações; e (ii) mais de uma organização rítmica da palavra escrita, em virtude das possibilidades de atribuição de acento secundário. São os casos dos Tipos (6) e (7), os quais nos levaram a supor, também, mais de uma configuração para o registro da hipersegmentação.

<sup>65</sup> “a panho” para “apanhou”.

- ii. não houve uniformidade na distribuição de cada Tipo e Subtipo de estrutura nos anos escolares, o que equivale a dizer que nem todos os que foram levantados ocorreram sempre;
- iii. os Tipos (1) e (3), além de terem estado presentes nos quatro anos do EF-II, correspondem, também, aos mais frequentes no sexto, sétimo, oitavo e nono anos. Também, em todos os anos, são identificadas grafias do Tipo “Outros”. Relembramos que, sob esse rótulo, agrupamos aquelas grafias que, além da diferença estrutural com a maioria dos dados, parecem seguir motivações distintas daquelas que extraímos a partir da Tabela 7. As características desse Tipo demandam um estudo pormenorizado de cada grafia, o que não foi realizado neste momento do trabalho, que visa uma descrição de regularidades estruturais. Do Tipo “Outros” tratamos na subseção seguinte.
- iv. no sexto ano são encontrados os Tipos (2), (4), (5) e (6) de palavras convencionais e os Subtipos (2.0), (4.0), (5.2) e (6.0) das hipersegmentações. Em relação ao conjunto proposto na Tabela 7, apenas a estrutura (7) não ocorre nesse ano escolar. Já, no sétimo ano, identificamos os Tipos e Subtipos (2)/(2.0), (5)/(5.1) e (7)/(7.0). Portanto, em relação ao ano anterior, houve redução da diversidade de todas as estruturas rítmicas. Referente ao oitavo ano, constatamos os Tipos (4), (5), (6) e (7) na configuração dos Subtipos (4.0), (5.1), (6.0) e (7.0). Por fim, o nono ano apresenta a menor variedade de estruturas, sendo as do Tipo (4) e (5) e as do Subtipo (4.0), (5.1) e (5.3).<sup>66</sup>

---

<sup>66</sup> Em anexo, constam, para cada um dos anos, Tabelas sobre informações prosódicas das estruturas rítmicas mobilizadas. Essas Tabelas apresentam a quantidade de ocorrências de hipersegmentação e um exemplo de grafia representativo para cada Tipo e Subtipo encontrados.

Expostas as características estruturais das hipersegmentações, damos início à apresentação de questões que se evidenciam a partir do modo de organização proposto. O primeiro fato a ser destacado é a presença dos recursos demarcadores de palavra fora da convenção em fronteiras de constituintes prosódicos, as quais foram interpretadas como sendo entre a  $\sigma$  e o  $\Sigma$ , os dois domínios mais baixos do modelo prosódico assumido. Desse resultado, podemos dizer que as hipersegmentações de palavras são uma *pista gráfica* de limites de constituintes prosódicos.

No que diz respeito aos tipos de estruturas rítmicas, o resultado obtido converge para um mesmo ponto: do primeiro ao último ano do EF-II, as estruturas que mais levam a grafias não-convencionais e as que resultam em maior número de possibilidades de registros são, respectivamente, palavras trissílabas do Tipo (3), como *ouvindo*, e palavras dissílabas do Tipo (1), como *assim*. Foi frequente, também, ainda que em menor número, grafias não-convencionais de palavras polissilábicas do Tipo (5), como *loucamente e recompensa*.

Passando a tratar dos tipos mais comuns, as palavras convencionais abrangidas em (1) têm a estrutura organizacional de um  $\Sigma$  iambo (alternância rítmica fraco/forte) que, quando hipersegmentada, passam a corresponder à combinação de duas sílabas. No caso do Tipo (3), é a partir da estrutura trissílaba da unidade convencional que os escreventes realizam uma escrita baseada na relação entre uma  $\sigma$  e um  $\Sigma$ , predominantemente quando correspondia sua organização à de um troqueu (cf. subtipo (3.1)). No Subtipo (5.1), determinadas palavras formadas por quatro sílabas são grafadas como duas unidades dissílabas, as quais também correspondem à organização de pés trocaicos. Esse resultado segue as constatações de Abaurre (1991c) em dados de escrita infantil: sobre as crianças parecerem preferir a grafia de palavras dissílabas cujo padrão acentual é o paroxítono (o qual carrega uma formação calcada em um  $\Sigma$  troqueu). Assim, ao fazermos relação dos nossos resultados com os de Abaurre (1991c), encontramos evidências de que, em alguma medida, informações sobre as formas

canônicas de palavras do PB ainda são pontos de ancoragem nas hipóteses dos escreventes concluintes do EF.

A partir do Tipo (1) e dos Subtipos (3.1), (3.2), (5.1) e (5.3) localizamos uma informação prosódica importante, que diz respeito ao fato de a alocação não-convencional do limite gráfico ter ocorrido, nesses casos, na fronteira em que está localizado o acento primário (ponto de maior proeminência prosódica dentro da palavra). Levantamos a hipótese explicativa, a partir desse achado, de que a identificação de uma proeminência fônica, em determinado ponto da palavra, leva o escrevente a supor que há ali uma fronteira a ser registrada graficamente por um limite dado, ou pelo espaço em branco ou pelo hífen. Parece que, nesses momentos, os escreventes tentam marcar, por meio dos limites que propõem às palavras, características dos enunciados orais/falados na produção escrita e, assim, deixam evidenciar a imagem que têm da gênese da escrita (CORRÊA, 2004). No entanto, como as práticas orais/faladas e letradas/escritas estão sempre em diálogo, além das proeminências rítmicas no interior das palavras, as hipersegmentações apresentam, conjuntamente, fatos não alheios aos aspectos circundantes da própria escrita. É interessante notar que nos subtipos (3.2) e (5.3) todos os dados de hipersegmentação são referentes ao uso não-convencional do hífen.<sup>67</sup> A interpretação dos dados de hífen, como marca característica de enunciados escritos, já foi explicitada em parágrafos precedentes; ainda gostaríamos, porém, de acrescentar mais um argumento a favor do nosso posicionamento, o qual encontra base em resultados de Abaurre e Galves (1996).

Em estudo sobre pronomes clíticos do PB, as autoras chegam à conclusão de que, em enunciados falados, as estruturas enclíticas são marcadas, observando-se, pois, uma nítida preferência dessa língua, segundo Abaurre e Galves, pela próclise. Constata-se, então, que estruturas de verbo + pronome enclítico ainda estão preservadas na língua pelas práticas de

---

<sup>67</sup> Nesse conjunto, há apenas uma ocorrência por espaço em branco: “falam do” (falando).

escrita convencional, fato que nos leva a fortalecer a premissa de que a presença do hífen em fronteira de palavra, nos dados analisados, traz fortes indícios de que, ao realizá-la em seu texto, o escrevente tenha buscado alçar aspectos mais institucionalizados da escrita.

Por meio da observação das ocorrências de espaço em branco, também não descartamos a ação de informações gráficas ligadas às decisões de como segmentar. Ao contrário, em todos os Tipos de estruturas rítmicas conseguimos identificar grafias hipersegmentadas em que alguma parte tenha correspondido com palavras que, ortograficamente, são espaçadas entre limites. Por exemplo, foi frequente nos dados a sílaba pretônica da palavra convencional ter sido isolada entre brancos, sílaba a qual linguisticamente pode corresponder tanto a um *clítico fonológico* quanto a classes de palavras funcionais. Exemplos dessas correspondências são: “de mais” (*demais*) – com preposição; “a migo” (*amigo*) – com artigo; e “qual do” (*quando*) – com pronome e formas contraídas de preposição e artigo (contração de *de + o*). Encontramos, também, dados em que parte resultante da hipersegmentação se aproxima graficamente de formas verbais, como, por exemplo: “you tando” (*voltando*). Buscando, assim, explicitar uma possível relação das hipersegmentações com fatos de natureza morfossintática, apresentamos, na Tabela 8, o número de ocorrências que, em alguma parte originada a partir da hipersegmentação, corresponde com classes de palavra morfossintática:

**Tabela 8:** Hipersegmentações de espaço em branco e correspondências gráficas

Correspondência	Hipersegmentação	
	Nº	%
Relacionada à classe de palavras funcionais	80	70,9%
Relacionada à classe de palavras lexicais	4	3,5%
<b>Total</b>	<b>84</b>	<b>74,4%</b>

Fonte: Elaboração própria

A Tabela 8 mostra que, nos textos analisados, há alto percentual de ocorrências de hipersegmentação que, além de darem indícios de questões referentes à organização prosódica da língua, põem em evidência a relevância em se considerar outras motivações na interpretação desse tipo dado. Observando, portanto, que as hipersegmentações são condicionadas, também, por características advindas de informações letradas/escritas, as quais acreditamos ficarem explícitas por meio da relação que propomos entre limites não-convencionais de palavras e classes de palavras morfossintáticas (que, por sua vez, em termos ortográficos, estão sempre divididas pelos recursos sinalizadores de palavras), investigamos quais unidades das classes de palavras funcionais e lexicais podem ser identificadas nos dados. Para o desenvolvimento dessa análise, averiguamos quais categorias de palavras funcionais e lexical foram identificadas. Os resultados dessa etapa da análise estão ilustrados na Tabela 9:

**Tabela 9:** Correspondência gráfica em hipersegmentações: tipos de classes de palavras

Tipos de classes		Hipersegmentação	
		Nº	%
Lexical	Verbo	4	3,5%
	Preposição	47	41,6%
Funcional	Artigo	30	26,6%
	Pronome	3	2,7%
<b>Total</b>		<b>84</b>	<b>74,4%</b>

Fonte: Elaboração própria

Sobre as categorias de palavras funcionais, podemos dizer que a sílaba pretônica separada entre limites é, predominantemente, analisada como preposições (41,6%), artigos (26,6%) e pronomes (2,7%), especialmente em formas como “em, na, de, com, a, qual, se”, a exemplo de dados como “em fachado” (*enfaixado*), “na quela” (*naquela*), “de pois” (*depois*), “com versa” (*conversa*), “a manhã” (*amanhã*), “qual do” (*quando*), “ganha-se” (*ganhasse*). Quanto às classes de palavras lexicais, somente uma se mostra como dúvida aos escreventes:

a dos verbos (3,5%) – “recom pensa” (*recompensa*), “vou tando” (*voltando*). Chegamos à evidência de que, no final do EF, os escreventes têm mais dúvidas em estabelecer limites para aqueles tipos de palavras cujo funcionamento linguístico é caracteristicamente gramatical resultado semelhante ao que foi descrito por Tenani (2011a) quando foram considerados dados transversais de segmentação não-convencional de palavras.

Determinados dados como, por exemplo, “de ele” (*dele*), “vou tando” (*voltando*), “falam do” (*falando*), “com migo” (*comigo*) e todas as grafias de *em* como, por exemplo, “em fachado” (*enfaixado*) e “em teiro” (*inteiro*) indicam a atuação de outra informação da escrita convencional, referente “à escolha de letras para representar certos segmentos que estão nos limites ou fronteiras dos pontos de corte do *continuum* fônico-gráfico” (TENANI, 2009b, p. 115). Nas ocorrências encontradas, há: (i) o predomínio da escolha de <m> para representar a coda nasal, quando a escrita convencional da palavra prevê <n>: “falam do” (*falando*), “em tão” (*então*), o mesmo verifica-se para o <l> em “vou tanto” (*voltando*); e (iii) duplicação de letras: “de ele” (*dele*), “com migo”. Essas constatações refletem a projeção dos escreventes sobre o que sejam palavras escritas da língua e, mais ainda, sobre quais pistas, presentes no funcionamento do próprio código escrito institucionalizado (CORRÊA, 2004), auxiliam em sua identificação. Citamos, por exemplo, a regra ortográfica de, em fim de palavra, a coda nasal ser registrada pela letra <m> e não <n> que, quando generalizada pelo escrevente em sua tarefa metalinguística de tentar reconhecer palavras, origina possíveis palavras escritas terminadas por esse grafema; são os casos, por exemplo, das separações não-convencionais de sílabas pretônicas como sendo os elementos *em* e *com*. Como resultado de reflexões metalinguísticas, podemos explicar, ainda, a grafia de “de ele” para *dele*, resultante da apreensão, na sequência fônica, de duas palavras escritas bastante comuns: *de* e *ele*. Em nossa interpretação, essas escolhas de letras evidenciam como diferentes informações linguísticas ancoradas nas diferentes formas de participação dos escreventes em práticas letradas/escritas

são mobilizadas pelo escrevente, quando diante de si tem a exigência escolar (e, especialmente, social) de grafar e delimitar palavras segundo as convenções ortográficas. Esses dados, uma vez mais, corroboram, a compreensão de que os escreventes do EF-II continuam em processo de aquisição, marcado pela relação dos sujeitos com as práticas sociais de linguagem que os constituem enquanto escreventes e que se revela pelas hipóteses que são registradas no material escrito.

As ocorrências de hipersegmentação em que a sílaba originada pela presença do limite gráfico indica o reconhecimento de alguma palavra escrita revelam, também, outras marcas explícitas de uma escrita que se constitui heterogeneamente, na medida em que essas mesmas sílabas permitem observar o funcionamento de elementos que, fonologicamente, são classificados como clíticos devido a sua ausência de acentuação própria.

Na discussão que realizamos, na primeira Seção, sobre os estudos de Bisol (2000, 2005) e Simioni (2008), que versam a respeito do comportamento prosódico dos clíticos do PB, constatamos diferentes representações da inserção desses elementos no interior da hierarquia prosódica, em função de arcabouços teóricos de naturezas distintas. Bisol defende a existência do grupo clítico, já que determinados processos fonológicos se aplicam na exata extensão das sequências de clítico+palavra. Já Simioni propõe a consideração do clítico junto a uma frase fonológica, sob o argumento de regras de restrições prosódicas e de alinhamento. Sem ignorarmos as filiações teóricas de cada um desses estudos, trazemos, para a discussão desta dissertação, pontos de convergência que interpretamos como subjacentes às análises de Bisol e Simioni: na formalização da inserção prosódica dos clíticos do PB, esses elementos são situados no domínio de constituintes que estabelecem forte relação com o componente sintático, mostrando que os clíticos mantêm uma relação que é mais independente em relação à palavra que lhe é hospedeira. Assumindo, pois, ser essa uma característica do clítico do PB, visto que em outras línguas os clíticos indicam outro funcionamento (ver VIGÁRIO, 2001,

por exemplo), acreditamos que as grafias de palavras hipersegmentadas, na medida em que representam autonomia gráfica de elementos gramaticais átonos, parecem indicar um funcionamento mais independente desses elementos também em termos fonológicos.

Na sequência, trazemos dois textos em que foram registradas diferentes grafias para os clíticos *a* e *em*, as quais julgamos corroborar com a hipótese exposta. Se observarmos a organização, na Figura 15, a seguir, da partícula *a*, podemos verificar que, em praticamente todas as ocorrências desse elemento (convencionais e não-convencionais), ele foi separado por limites gráficos, com exceção de quando o possível isolamento do *a* daria origem a uma estrutura, na palavra seguinte, que, em termos de sílaba, não é gramaticalmente possível para o PB; nos referimos à grafia de “astonasta” (*astronauta*) e a possível sílaba –sto (caso tivesse ocorrido a hipersegmentação). Na Figura 16, a seguir, as grafias de quando o escrevente tenta sinalizar o seu reconhecimento da palavra *em* são realizadas com o grafema <m>, no entanto, quando esse reconhecimento é referente à partícula como uma sílaba integrada à palavra, tal registro se dá por meio da letra <n>. Defendemos, assim, que essas flutuações gráficas ora exemplificadas demonstram a oscilação dos escreventes por fatores linguísticos que se relacionam com o funcionamento prosódico de elementos não-acentuados em convergência (de modo a se complementarem) com aspectos de práticas letradas/escritas que permeiam a noção do que seja uma palavra na escrita convencional. Acreditamos, ainda, junto com a defesa a respeito de as hipersegmentações demonstrarem uma relação complexa entre enunciados falados e escritos, que o contato que os escreventes passam a ter com o funcionamento da escrita é um dos fatores cruciais que os levam a observar esses elementos clíticos da língua como uma categoria linguística que, em determinadas situações podem ser palavras escritas:

Figura 15: Exemplo de grafias do clítico “a” hipersegmentados

Um dia fui viajar com o meu pai, para o mundo.  
 E estava a cabano a gamali no meu pai deseu o fogito.  
 E eu falei para meu pai olha a quele bicho ele e olherudo e ele e ve de.  
 E somero a pareser mais antonasta e com o siquenso, e cataam o meu pai e ca-tam em.

Fonte: Banco de Dados de Escrita do Ensino Fundamental  
 (texto: Z08\_5A\_20M\_05)

Figura 16: Exemplo de grafias do clítico “em” hipersegmentados

No dia 11 de agosto de 2006 quando eu tinha 10 anos na quarta série, eu conheci uma garota chamada Bruna da minha amiga, ela era muito bonita e eu fiquei apaixonado por ela. Mas no dia seguinte quando eu ia pra escola, ela me chamou e me falou que ela queria que eu fosse com ela pra casa dela. Na verdade tentei encontrar a Bruna mas não a encontrei; então eu fui para a casa de minha amiga e falei pra ela que eu queria ir com ela pra casa dela. Chegando lá a gente fez o trabalho de casa, comemos um sanduiche e fomos jogar video game na que tempo a play station 2 era muito cara e meus colegas tinham a play station 1, então eles queriam negociar; então eu falei que eu não vendia por R\$50,00 mas se eles falarem com meu pai eu vou vender, mas aí toda mundo falou que gostavam de uma game station e nome é aí era uma loquita e eu falei pra Bruna: garota dos meus sonhos. Mas depois de um tempo a gente esqueceu o negocio do video game e fizemos um trato de fazer uma festa com a game station e não ficou o negocio com a Bruna. No dia seguinte eu fui com meu pai pra a cantina de Bruna e a empicada, então a Bruna não durou muito tempo ela mudou de escola e nunca.

Fonte: Banco de Dados de Escrita do Ensino Fundamental  
 (texto: Z10\_7B\_26M\_05)

Resta-nos comentar dados que, no interior dos tipos de estruturas rítmicas, chamaram-nos a atenção para questões como a formação de palavras com afixos acentuados.<sup>68</sup> Para fazermos a interpretação de ocorrências que suscitaram essa questão, tomamos por base Schwindt (2001) e Quadros e Schwindt (2008). Segundo esses autores: (i) palavras derivadas de afixos acentuados apresentam dois acentos, um localizado na base e outro no afixo (por exemplo: loucamente); (ii) afixos dissílabos, por formarem  $\Sigma$ s isoladamente, são candidatos a receberem acento primário característico de  $\omega$  (por exemplo: sozinho – formam um pé binário do tipo troqueu, nesse caso); (iii) os afixos acentuados (principalmente prefixos) estão sujeitos às regras fonológicas que se aplicam entre palavras (por exemplo: ant[e]penúltimo > ant[i]penúltimo – neutralização da átona final); e (iv) tais afixos apresentam, por vezes, certa autonomia morfossintática em relação à sua base (por exemplo: “José está fazendo uma pós” (pós-graduação)). A partir desses resultados, analisamos os dados que envolvem afixos acentuados e entendemos que os escreventes tenham se ancorado, sobretudo, nas características linguísticas (i), (ii) e (iii) desses afixos, pois, em dados como “so zinho” e “loca mente”, os sufixos configuram-se como  $\Sigma$ s troqueus, formando uma fronteira acentual entre sufixo e base, fronteira a qual já constatamos por meio de outros dados (que não necessariamente envolveram afixos) ser uma informação relevante na tentativa de registro dos limites gráficos de palavra. Já na ocorrência “mini cidade”, parece haver uma independência morfossintática (e, talvez, morfossemântica) do prefixo em relação à sua base.<sup>69</sup>

A sequência desta Seção é dedicada à análise dos dados “particulares”.

<sup>68</sup> Os dados de que constam afixos acentuados são: “anti penúltimo” (antepenúltimo), “so zinho” (sozinho), “mini cidade” (minicidade), “contra bando” (contrabando) e “loca mente” (loucamente).

<sup>69</sup> Um possível enunciado em que essa relativa autonomia se mostraria seria: “Maria mora em uma grande cidade ou em uma minicidade? Ela mora em uma mini”.

## 4.2. Aspectos particulares das hipersegmentações

Na parte anterior de nossa análise, privilegamos a discussão das ocorrências de hipersegmentação que, de algum modo, permitiram o apontamento de regularidades linguísticas, as quais contribuíram para que atingíssemos o primeiro objetivo específico desta pesquisa. Agora, o nosso intuito é o de realizar uma interpretação que se diferenciam qualitativamente da maioria das hipersegmentações. Afirmamos que essas diferenças de mostram em dois planos: naquele da estrutura do dado e naquele da motivação subjacente.

Todavia, durante a análise das marcas “particulares”<sup>70</sup> de hipersegmentação, não deixamos de abordar questões linguísticas já antes discutidas, uma vez que, mesmo a partir de ocorrências grande parte das vezes diferentes das tendências encontradas, é possível, em alguma medida, estabelecer vínculos com aspectos mais gerais dos dados analisados, os quais nos permitiram tratar de generalizações. Desse modo, percebemos que hipóteses mais gerais e particulares dialogaram entre si, visto não serem excludentes, na medida em que se constituem de hipóteses sobre uma só língua.

A identificação das grafias “particulares” se deu, exclusivamente, nas produções escritas de dois dos escreventes do corpus. Em nossa compreensão, esse resultado é ainda mais interessante se considerarmos que as hipersegmentações são produzidas por aqueles escreventes que, também em termos quantitativos, se diferenciam bastante em comparação aos outros alunos. Esses escreventes são Bruno e Pedro e, juntos, somam 59 dados, representativos de 53% de todas as grafias não-convencionais identificadas. No entanto, para a discussão pretendida, selecionamos sete ocorrências de hipersegmentação mais os dois dados híbridos. Ainda, os aspectos que particularizam esses dados merecem comentário: embora sejam grafias produzidas pelos mesmos escreventes, essas ocorrem uma única vez,

---

<sup>70</sup> O termo “particular” busca registrar a diferença de alguns dados não apenas em relação ao que identificamos como “geral” do corpus, mas, especialmente, em função de os interpretarmos como pistas das diferentes formas como os escreventes são capturados pelo sistema simbólico da escrita (LEMOS, 1998).

mesmo quando, no interior de um só texto, tenham sido registradas em mais de um momento; tal fato reforça a nossa interpretação de essas marcas serem fruto de hipóteses episódicas (ABAURRE, 1988a), mobilizadas no texto não só por o escrevente visar atender a uma exigência ortográfica de segmentação em palavras, mas, também, por revelar o modo como o “*escrevente lida com sentidos que perpassam o texto ao segmentar os enunciados em porções cujos limites não coincidem com aqueles da palavra convencional*” (TENANI, 2008, p. 242, destaques nosso). Em função, também, desse possível aspecto, não é somente o dado (enquanto estrutura) que, determinadas vezes, nos interessa observar, mas, sim, o modo como a grafia se apresenta no funcionamento do enunciado em que ocorre.

No Quadro 14, abaixo, destacamos as ocorrências produzidas por Bruno e por Pedro. Nota-se, que os dados híbridos e seis dentre todas as hipersegmentações foram identificadas nos textos de Bruno, enquanto que apenas um dado esteve presente na escrita de Pedro:

**Quadro 14:** Grafias “particulares” (hipersegmentação e híbrido) produzidas de Bruno e Pedro

<b>BRUNO</b>			
<b>6º ano</b>	<b>7º ano</b>	<b>8º ano</b>	<b>9º ano</b>
a i ( <i>aí</i> ) adisi o na ( <i>adicionar</i> ) es quito ( <i>escrito</i> ) com panhate ( <i>acompanhante</i> )		come cheia ( <i>comecei a</i> ) a olado ( <i>ao lado</i> )	com ver sando ( <i>conversando</i> ) com versando com ver samos ( <i>conversamos</i> )
<b>PEDRO</b>			
<b>6º ano</b>	<b>7º ano</b>	<b>8º ano</b>	<b>9º ano</b>
	foute i ( <i>voltei</i> )		

Fonte: Elaboração própria

Um olhar mais geral já nos permite perceber as diferenças dessas formas no que tange à maior parte das hipersegmentações analisadas e, sobretudo, referente ao modo de escrita das

palavras da língua, segundo o que se espera de escreventes que concluíram, pelo menos, quatro anos de sua escolarização formal.<sup>71</sup> Dessas diferenças, gostaríamos de destacar:

- i. a distribuição da fronteira gráfica que, em alguns dos dados, ocorre em mais de um ponto da palavra – por exemplo, “adisio o na”, “com ver sando” e “com ver samos”;
- ii. a escolha e organização de grafemas que compõem a palavra – por exemplo, a troca de <v> por <f> na grafia “foute i” e o apagamento de <a>, no ataque da primeira sílaba, e de <n>, na coda da quarta sílaba, em “com panhate”.

No entanto, em pontos daquelas rupturas apresentadas (CHACON, 2005), podemos recuperar aspectos de práticas orais/faladas e letradas/escritas que caracterizam o conjunto geral das hipersegmentações do córpus: nos referimos às informações prosódicas (em termos da configuração rítmica da palavra) e gráficas (em relação à correspondência de partes da hipersegmentação com palavras escritas da língua). Vejamos, por exemplo, o dado “es quito”, produzido por Bruno. Em se tratando da estrutura prosódica da palavra escrita convencional (isto é, *escrito*), notamos ser essa idêntica àquela a partir da qual foi proposta grande parte das hipersegmentações levantadas, ou seja, um trissílabo com as sílabas organizadas em fraca/forte/fraca (cf. tipo (3) – Tabela 7). No que se refere à hipersegmentação em si, o dado “es crito” à semelhança de outros como “na quela” apresenta a configuração prosódica mais frequente para as hipersegmentações: a de uma  $\sigma$  seguida de um  $\Sigma$  troqueu (cf. subtipo (3.1) – Tabela 7). Podemos dizer, com base nessa constatação, que, apesar de a grafia “es quito” não se assemelhar com as demais ocorrências do EF-II que apresentam a mesma estrutura prosódica (motivo a ser exposto, na sequência), esse dado, por sua vez, não foge às características do PB, em relação ao tipo de estrutura preferencial de palavra da língua: as que apresentam acento à esquerda (ver BISOL, 1996).

---

<sup>71</sup> Lembrando que os escreventes analisados ingressaram no EF quando esse ainda era constituído por quatro anos em sua primeira etapa.

No entanto, conforme concluímos nos resultados anteriormente apresentados e, também, nos já mostrados por Tenani (2011a) e Paranhos e Tenani (2011), a categoria gramatical a qual a sílaba hipersegmentada à esquerda pode ser correlacionada é uma informação relevante “não só por mostrar os tipos de categorias gramaticais com as quais os alunos ainda apresentam dificuldade na escrita convencional, mas também, por permitir observar uma diferença qualitativa quando comparamos os dados de escreventes [...] com os dados de alunos em fase inicial de aquisição da escrita infantil” (PARANHOS E TENANI, 2011, p. 496). Em Silva, L. (2011), realizamos um estudo comparativo entre dados de segmentação não-convencional encontrados em textos de alunos de *quinto* (EF-I) e *nono* (EF-II) anos, com o intuito de identificarmos possíveis semelhanças (em busca de aproximações) e diferenças (em busca de caracterizações) nesses dois importantes anos do EF, correspondentes ao encerramento, por um lado, do período da alfabetização, e, por outro, da etapa fundamental de formação escolar. Deixando de lado as semelhanças e abordando, aqui, as diferenças observadas na configuração das hipersegmentações, constatamos que as sílabas pretônicas separadas entre brancos pelos escreventes do quinto ano não apresentaram, necessariamente, relação com categorias gramaticais da língua, enquanto que a grafia da sílaba desse mesmo tipo de dado encontrado no nono ano, sempre, ligou-se a algum item gramatical, conforme podemos observar no Quadro 15:

**Quadro 15:** Exemplos de hipersegmentações frequentes no EF-I e EF-II

EF-I	EF-II
es pero ( <i>espero</i> )	a quele ( <i>aquele</i> )
is tafa ( <i>estava</i> )	em bora ( <i>embora</i> )
a migo ( <i>amigo</i> )	com migo ( <i>comigo</i> )
di so ( <i>disso</i> )	a final ( <i>afinal</i> )
pro meto ( <i>prometo</i> )	de veres (deveres)
mi nuto ( <i>minuto</i> )	em vergonhada ( <i>envergonhada</i> )

Fonte: Silva, L. (2011, p. 41-42)

Comparando, por exemplo, o dado “em vergonhada”, presente no Quadro acima, com a ocorrência “es quito”, notamos, claramente, na hipersegmentação do EF-II, uma relação da sílaba da grafia não-convencional com a preposição *em*, de maneira que podemos entender que a relação entre essas duas unidades linguísticas (sílabas pretônicas e item gramatical), tanto em termos sonoros quanto gráficos, tenha despertado, no escrevente, a identificação metalinguística das possibilidades de sentido (TENANI, 2004) que o *em*, grafado entre limites, desempenha no funcionamento de determinados enunciados escritos. Como se pode compreender, essa mesma interpretação não se estende ao “es” em “es quito”; desse modo, entendemos por em evidência a particularidade desse dado em relação às outras ocorrências do corpus que apresentaram a mesma estrutura de palavra. Portanto, se, como notamos em Silva, L. (2011),<sup>72</sup> na primeira etapa do EF as hipersegmentações ainda estão, mais explicitamente, ancoradas em aspectos fonético-fonológicos da unidade *palavra* – mas não exclusivamente, segundo Chacon (2005) –, podemos considerar, então, que Bruno continua a mobilizar, em sua escrita, aspectos prosódicos que, provavelmente, lhe são salientes desde os seus primeiros anos de escolarização. Mas, sem negar essa primeira hipótese, podemos, também, interpretar a grafia hipersegmentada de “es” como resultado de uma possível correlação dessa sílaba com o prefixo “ex”, o qual apresenta estatus de palavra escrita independente em determinados enunciados.

Próximo de outras características que remetem à escrita inicial está o dado “foute i” produzido pelo escrevente Pedro. Em particular, essa ocorrência tende a dialogar em dois pontos com aspectos do início da aquisição da escrita. O primeiro aspecto já foi sinalizado e

---

<sup>72</sup> As pesquisas de Paula (2007) e Cunha (2004), por exemplo, que analisaram, também, a escrita inicial de crianças, identificaram hipersegmentações muito semelhantes às que levantamos em nossa pesquisa de iniciação científica.

diz respeito à escolha da letra que registra o início da palavra escrita *voltei*: a troca <f>/<v> (vice-versa) é bastante comum em produções escritas infantis (cf. MIRANDA E MATZENAUER, 2010) e, nesse aspecto, o dado “foute i” demonstra que essa dificuldade de seleção das letras parece ter perdurado nas hipóteses gráficas de Pedro ainda no EF-II. Defendemos que essa troca de <f> por <v> não se atribuiria ao aspecto de categorização gráfica (MASSINI-CAGLIARI E CAGLIARI, 2004) dessas letras, isto é, advinda de uma dificuldade do escrevente em reconhecer a forma de grafar as duas letras, mas, sim, por um aspecto mais fonético-fonológico que caracteriza, ao mesmo tempo, o [f] e o [v]. A esse respeito, no sistema fonológico do PB, essas consoantes são classificadas quanto ao modo e ponto de articulação e ao grau de vozeamento como: (i) [f]: fricativa labiodental desvozeada; e (ii) [v]: fricativa labiodental vozeada. Desse modo, [f] e [v] só se diferenciam em relação ao vozeamento, fato que põe em evidência a proximidade sonora entre os sons representados por essas letras e, por consequência, a complexidade encontrada pelos aprendizes de escrita em registrar “os sons representados por algumas letras [que] são sonoramente semelhantes” (SIMIONI E RAUPP, 2012, p. 111).

O segundo aspecto a que nos referimos trata, especificamente, da segmentação não-convencional proposta entre a pseudopalavra (CUNHA, 2004)<sup>73</sup> “foute” e o grafema <i>, parte da sílaba final da palavra. Um fato importante a ser ressaltado é o de que, em comparação com a ocorrência “es quito”, essa hipersegmentação é ainda mais rara em relação aos dados encontrados no EF-II e muito se aproxima das soluções escritas infantis, pois, nesse caso, “o constituinte prosódico basilar, a sílaba, foi rompido em algum ponto de sua participação na palavra hipersegmentada” (CHACON, 2005, p. 81). Na análise de Capristano (2007b), dados com essa natureza só ocorreram na primeira série/segundo ano, ou seja, em

---

<sup>73</sup> Por *pseudopalavra* Cunha (2004) entende os resultados das hipersegmentações que não demonstram relação explícita com o reconhecimento de alguma possível palavra escrita da língua, mas que, por apresentar acento, pode ser considerada palavra do ponto de vista fonológico. Exemplos de pseudopalavra, nos termos da autora, são as partes destacadas nos dados selecionados do corpus analisado: “na quela”, “com ver sando”, “adisio no na”.

período bastante inicial de aquisição da escrita. No entanto, Capristano (2007b, p. 136) salienta que, nas segmentações em que os limites silábicos não são preservados, estão subjacentes “a atuação prioritária da escrita institucionalizada que determina a distribuição feita pela criança” já que, em muitos casos, “a ruptura resulta em unidades que podem corresponder homonimicamente a monossílabos da língua”, o que poderia explicar, também, a hipersegmentação da palavra *aí* (a i), em que o escrevente possa ter buscado estabelecer estatuto de palavra escrita ao *a* e a um possível *e*, cuja realização em práticas orais/faladas mais informais pode ser [i], o que teria motivado a grafia com o <i>.

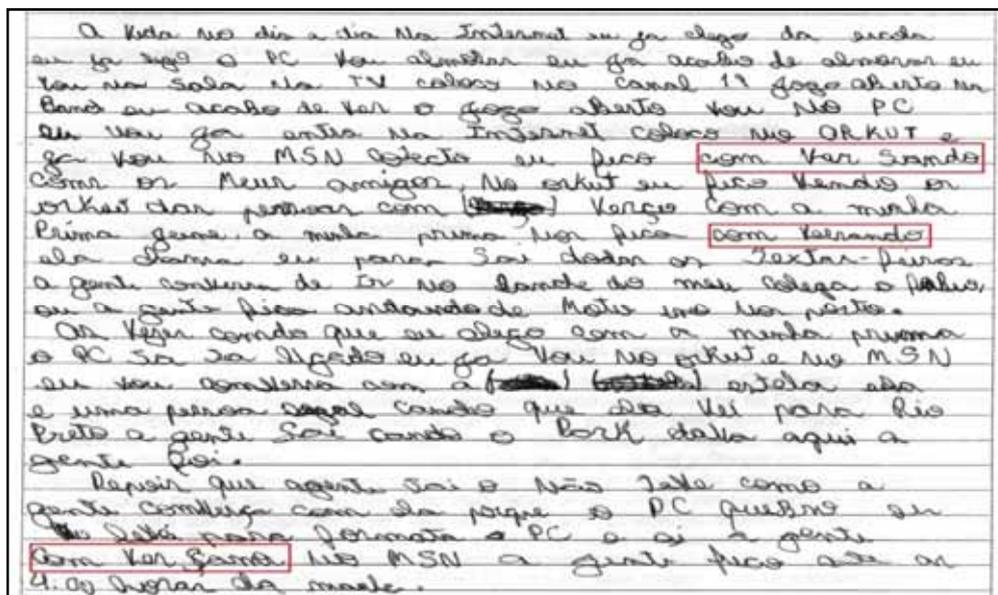
*A atuação prioritária da escrita institucionalizada*, como sugeriu Capristano (2007b, p. 136), parece rondar outras grafias de Bruno, como “com panhate” e as variações de “com ver sando”, “com ver samos” e “com versando”. Em “com panhate” (*acompanhante*) o apagamento do *a* pode ter se dado, justamente, em função da identificação do *com* em relação a forma como esse escrevente institucionalizou para a (sua) escrita a maneira convencional de registro da sequência fônica *com*. Vale acrescentar que interpretamos o *com* como uma unidade já institucionalizada para Bruno, principalmente pela observação das outras grafias da forma verbal *conversar*, proposta por esse escrevente. Soma-se a essa observação, o fato de o verbo *conversar* admitir como regência, a preposição *com*, o que, potencialmente, torna mais saliente ao escrevente a grafia de *com* como sendo uma palavra. Em alguma medida, isso está ilustrado no texto de Bruno no qual as hipersegmentações de *conversando* ocorrem, como em, por exemplo: “eu fico com ver sando coms meus amigos” (ver Figura 17, linha 6-7). Ainda sobre as três grafias não-convencionais (“com ver sando”, “com ver samos” e “com versando”), podemos interpretar, também, que a porção hipersegmentada “ver”, em dois dos dados, possa ter sido reconhecida, pelo escrevente, como a forma infinitiva do verbo “ver”. Desse modo, o escrevente, ao inserir os espaços em branco delimitadores de palavra, o faz entre os limites silábicos possíveis de projeção acerca de palavras da língua (“com” e “ver”,

nos exemplos), o que explica, por hipótese, o fato de não segmentar as sílabas “san”, “sa” e “mos”, já que essas, isoladamente, não correspondem a nenhuma palavra que, na escrita convencional, figura entre brancos. Também, em “com panhate”, “com ver sando”, “com ver samos” e “com versando” as fronteiras gráficas propostas apontam limites prosódicos dos domínios  $\sigma$  e  $\Sigma$ , como visto na tendência dos dados que já discutimos:

- (1) [com] $\sigma$  [panhate] $\Sigma$ ;
- (2) [com] $\sigma$  [ver] $\sigma$  [sando] $\Sigma$ ;
- (3) [com] $\sigma$  [ver] $\sigma$  [samos] $\Sigma$ ;
- (4) [com] $\sigma$  [versando] $\Sigma$ .

Outra hipótese explicativa, na qual nos baseamos em Tenani (2008), diz respeito ao tratamento das grafias não-convencionais em que se considera, também, para análise, “o fluxo narrativo do texto, pois dessa maneira podemos formular a hipótese de que a flutuação/oscilação das fronteiras de palavra na escrita pode ser motivada pela maneira como o branco é usado para indicar as relações de sentido que se constroem no texto, que também estão permeadas pela organização prosódica dada ao texto” (TENANI, 2008, p. 240). Por vermos que, em alguma medida, esse fato foi mobilizado por Bruno nas grafias que propôs para as diferentes formas do verbo *conversar*, recuperamos a produção escrita em que essas ocorrem e a trazemos na Figura 17, a seguir:

Figura 17: Dados “com ver sando”, “com versando” (*conversando*) e “com ver samos” (*conversamos*)



Fonte: Banco de Dados de Escrita do Ensino Fundamental  
(Texto: Z11\_8B\_13M\_01)<sup>74</sup>

A proposta de redação solicitava ao escrevente a produção de um texto narrativo em que relatasse as suas relações de amizade, bem como se, para a manutenção dessas amizades, fazia uso (ou não) da internet. Vejamos, para ficar mais claro, a instrução da proposta na íntegra: “*Provavelmente, você deve ter um grande amigo(a) com quem gosta de passear, conversar e até estudar juntos(as). Escreva um texto, relatando como surgiu essa amizade, se usa a internet ou não para manter contato com o(a) seu(sua) amigo (a), e se houver algum bom momento que vocês viveram juntos(as)*”.

<sup>74</sup> *Leitura atribuída:* A VIDA NO DIA A DIA NA INTERNET. A vida no dia a dia na internet: eu já chego da escola; eu ligo o PC; vou almoçar; eu acabo de almoçar; eu vou para a sala; na TV eu coloco no canal 19, Jogo Aberto, na Band; eu acabo de ver o Jogo Aberto; vou ao PC; eu já vou entrar na Internet; coloco no ORKUT; já vou no MSN; eu conecto; eu fico conversando com os meus amigos. No ORKUT, eu fico vendo os ORKUTs das pessoas. Eu converso com a minha prima Jeine, nós ficamos conversando e ela me chama para sair todas as sextas-feiras. Nós conversamos de ir ao lanche do meu colega, o Pablo, ou nós ficamos andando de moto, indo nos postos. Às vezes, quando eu chego com a minha prima, o PC já está ligado e eu já vou ao ORKUT e no MSN. Eu vou conversar com a Estela, ela é uma pessoa legal. Quando ela vem para Rio Preto, nós saímos, quando o parque estava aqui, nós fomos. Depois que nós saímos, não teve mais como nós conversarmos, porque o PC quebrou. Eu vou levá-lo para formatar. Aí, nós conversamos no MSN. Nós ficamos até às 04h00min da manhã.

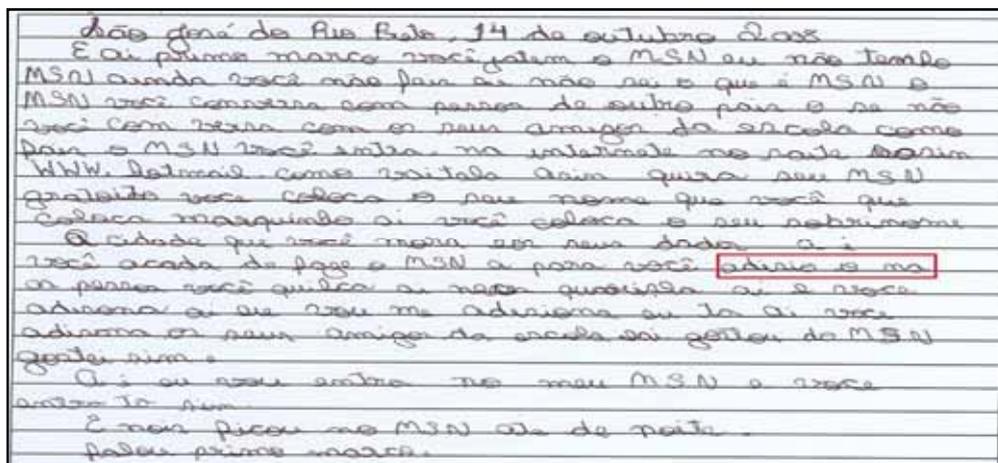
Localizando as hipersegmentações a partir da proposta temática, identificamos que a primeira segmentação não-convencional da palavra *conversando* se dá quando o escrevente passa a introduzir o que ele intitula como o seu “dia a dia na internet” (que ocorre na sexta linha do texto, seguida do verbo *conecto*). A estratégia do escrevente, ao se “conectar”, é conversar com seus amigos virtuais. Nesse aspecto, é interessante a ocorrência de diferentes formas de segmentação desse verbo: convencionais<sup>75</sup> (“conversa”, “comversa”, “comverça”); não-convencionais (“com ver sando”, “com versando”, “com ver samos”); rasura (“com verço”). A variedade de grafias relacionadas à palavra *conversar* fez com que investigássemos a razão que levou o escrevente a produzir várias grafias para formas verbais do mesmo verbo *conversar*. A esse respeito, se considerarmos o enunciado da proposta de texto, a qual abre possibilidades de construção da narrativa, quando, em sua instrução, afirma sobre os tipos de atividades que se desenvolvem entre amigos, como *passear, conversar e estudar*. Interpretamos, então, que o escrevente elege a atividade *conversar* como o eixo central da construção de sua narrativa, da qual se originam as interações que ele estabelece com seus amigos fora da rede. A marcação não-convencional das palavras, nesse sentido, é utilizado pelo escrevente como um recurso para demonstrar ao interlocutor os sentidos privilegiados em seu textos. E a forma como se dão essas segmentações não são aleatórias, evidenciando um conhecimento metalinguístico do escrevente em identificar palavras escritas no interior de uma única sequência ao mesmo tempo fônica e ortográfica.

E Bruno parece seguir nessa mesma ancoragem, quando propõe outras segmentações não-convencionais, como “adisio o na”, por exemplo. Observemos, então, a produção textual em que a identificamos.

---

<sup>75</sup> Convencional do ponto de vista do limite gráfico.

**Figura 18:** Dado “adisião na” (*adicionar*)



Fonte: Banco de Dados de Escrita do Ensino Fundamental  
 (Texto: Z08\_5C\_09M\_04)<sup>76</sup>

Bruno, ao escrever uma carta para seu primo Marco, busca fornecer instruções para ensiná-lo a se comunicar pela internet por meio do MSN. Dada essa proposta, chamamos a atenção para o fato de a forma hipersegmentada em discussão introduzir uma prática característica do uso do MSN: adicionar os amigos a sua lista de contatos. Nota-se, ainda, que a informação sobre a adição dos amigos é a última instrução dada por Bruno ao primo a respeito de como se faz para criar uma conta no MSN. E essa informação parece quase ter sido esquecida pelo escrevente, quando, ao registrá-la, Bruno inicia o enunciado com a expressão “Ah, para você adicionar as pessoas” (linha 11 do texto), no sentido de: “não se esqueça de adicionar as pessoas, do contrário não poderá falar com elas”. Assim, inferimos uma possível ênfase naquela informação que recai sobre o verbo *adicionar*, já que ela é fundamental para a comunicação pelo MSN. Somam-se, nesse aspecto, sentidos privilegiados

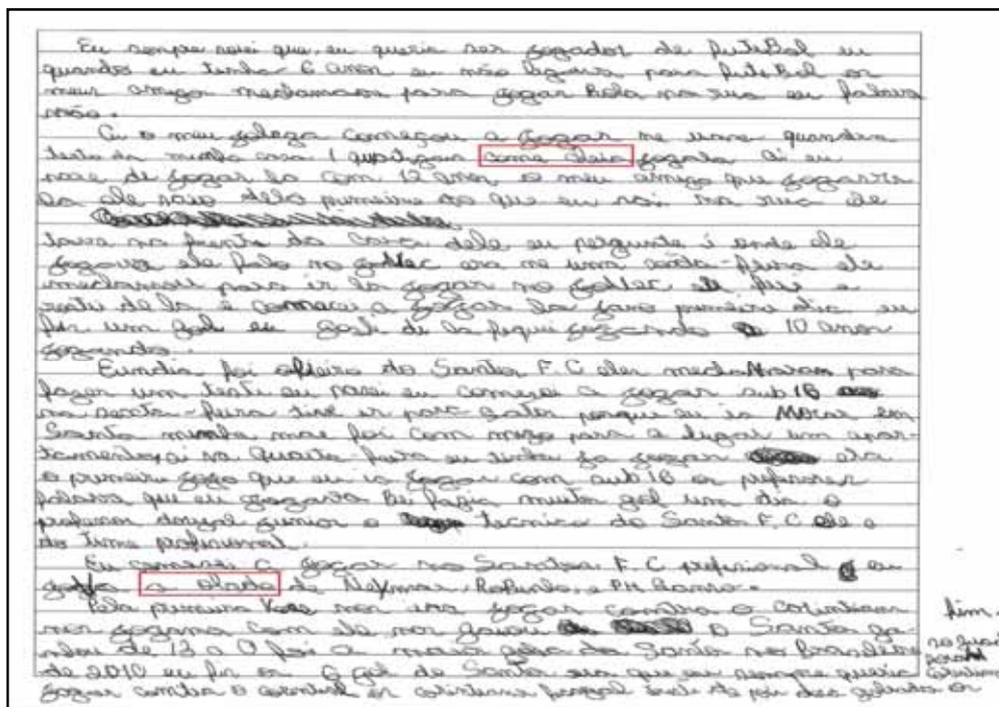
<sup>76</sup> Leitura atribuída: São José do Rio Preto, 14 de outubro de 2008. E aí, primo Marco, você já tem MSN? Eu não tenho MSN ainda. Você não fez? Ai, eu não sei o que é MSN. O MSN, você conversa com pessoas de outro país ou, senão, você conversa com os seus amigos da escola. Como faz o MSN? Você entra na internet, no site www.hotmail.com, vai estar lá assim: “crie seu MSN gratuito”, você coloca o nome que você quiser. Coloque Marquinho. Aí você coloca o seu sobrenome, a cidade que você mora e os seus dados. Aí você acaba de fazer o MSN! Ah, para você adicionar as pessoas, você clica aí nesse quadrinho. Aí você me adiciona, aí eu vou me adicionar. Aí você adiciona os seus amigos da escola. E aí, gostou do MSN? Gostei, sim! Aí eu vou entrar no meu MSN e você entra, tá! Sim. E nós ficamos no MSN até de noite. Falô, primo Marco.

do uso desse verbo em práticas letradas digitais, como o MSN, tema sobre o qual foi feita a proposta de produção escrita. Em termos rítmicos, na grafia de “adísio o na”, o escrevente parece querer plasmar um possível ritmo silabado para colocar a palavra em ênfase, e a duplicação e separação do *o* se dá entre as sílabas pretônicas “adísio”, “o” e a tônica “nar” (grafada sem a marca de infinitivo do verbo), como um possível indício de representação de um ritmo que desacelera e culmina na sílaba tônica da palavra enfatizada.

Com base na análise apresentada dos dois textos, compreendemos com Tenani (2008) que, em uma tentativa de atender ao que lhe é solicitado, o escrevente se baseia em certas estruturas e lança mão de registros alternativos de enunciados de seu texto, como nesse caso, por exemplo, do modo de segmentação das palavras. Nesse sentido, a segmentação de palavras ganha uma dimensão textual na produção desse escrevente, visto que a maneira de propor limites às palavras não ocorre de qualquer modo quando considerada a construção de todo o texto.

Finalizamos a discussão desenvolvida ao longo desta Seção, abordando os dados de híbridos: “come cheia” e “a olado”. Para a análise pretendida, observamos o texto em que aparecem os dados (ver Figura 19).

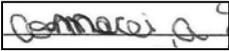
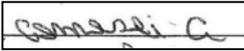
Figura 19: Dados “come cheia” (comecei a) e “a olado” (ao lado)



Fonte: Banco de Dados de Escrita do Ensino Fundamental  
(Texto: Z10\_7B\_05M\_02)<sup>77</sup>

Em se tratando da primeira ocorrência destacada, notamos, ao ampliarmos nosso olhar para a própria distribuição gráfica dos espaços em branco, ao longo do texto, pistas que nos levaram a classificá-la como um dado híbrido: é o caso de outras duas grafias da sequência

<sup>77</sup> Leitura atribuída: Eu sempre sonhei ser jogador de futebol. Quando eu tinha seis anos, não ligava para futebol. Meus amigos me chamavam para jogar bola na rua, eu falava: não! Ai, o meu colega começou a jogar em uma quadra perto da minha casa um quarteirão. Comecei a jogar lá e parei com doze anos. O meu amigo que jogava lá, saiu primeiro que eu. Eu saí na rua e ele estava na frente da sua casa e perguntei onde ele jogava. Ele falou no JOVEC. Em uma sexta-feira ele me chamou para ir jogar no JOVEC e eu fui. Gostei de lá e comecei a jogar e, já no primeiro dia, fiz um gol. Fiquei jogando por dez anos, e um dia o olheiro do Santos Futebol Clube me chamou para fazer um teste e eu passei! Comecei a jogar no sub 16. Na sexta-feira tive que ir para Santos, porque eu iria morar em Santos. Minha mãe foi comigo para alugar um apartamento. Na quarta-feira eu já tinha que jogar, era o meu primeiro jogo com o sub 16. Os professores falavam que eu jogava bem e fazia muitos gols. Um dia o professor Dorival Junior, técnico do Santos Futebol Clube, do time profissional. Comecei a jogar no Santos profissional. Eu jogava ao lado de Neymar, Robinho e Paulo Henrique Ganso. Pela primeira vez nós íamos jogar contra o Corinthians, nós jogamos e ganhamos. O Santos ganhou de 13 a 0, foi a maior goleada do Santos no Brasileiro de 2010. Eu fiz seis gols do Santos. Era o que eu sempre quis: jogar contra o Corinthians. Os corinthianos ficaram tristes depois dessa goleada. Eles foram zoados.

comecei a:  (linha 13),  (linha 17) – escritas de modo convencional, quanto ao espaço em branco.

Mesmo tratando-se de uma ocorrência particularizada, em relação às demais analisadas, podemos depreender que “come cheia” encontra motivação linguística semelhante às que já apontamos. Ao segmentar de modo não convencional a sequência “comecei a”, o escrevente propõe a grafia de duas novas palavras da língua: “come” e “cheia”, mesmo sendo, essa última, diferente do que se poderia esperar de uma hipersegmentação “regular” da palavra “comecei”, uma vez que o mais comum seria que o escrevente atribuísse autonomia à última sílaba da palavra, relacionando-a à palavra escrita “sei” e não à palavra “cheia”, como ocorre. Contudo, em termos prosódicos, essas sequências são palavras fonológicas, pois comportam dois acentos primários.

Estudos sobre hipossegmentações têm apontado que as junções que ocorrem entre um elemento clítico e uma palavra são, tendencialmente, realizadas à direita (como em “aboneca”, para *a boneca*), isto é, os clíticos são hipossegmentados, preferencialmente, à palavra que o seguem. Seguindo essa tendência, na sequência *comecei a*, o clítico *a* deveria estar unido à palavra seguinte, no caso o verbo *jogar* e não à palavra anterior, ou seja, o verbo *começar*. No entanto, o dado “come cheia” foge à tendência geral relatada na literatura sobre dados do EF-I e, também, em relação às características identificadas nesta pesquisa, na subseção anterior. A partir de indícios da organização do texto, levantamos as seguintes hipóteses para essa ocorrência:

- i. a primeira hipótese parte da flutuação entre <ch> e <c> para grafar o som [s]. Inicialmente cabe observar que o dígrafo <ch> representa, na ortografia do PB, o som [ʃ]; a letra <c>, o som [s], no contexto em análise. Na classificação das consoantes, quanto ao grau de vozeamento, ponto e modo de articulação, [ʃ] e [s] diferenciam-se, apenas, pelo ponto de articulação (alveolar e alveopalatal,

respectivamente). Essa diferença quanto ao ponto articulatorio das consoantes mostra a proximidade fonético-fonológica entre os sons representados por <ch> e <c>, embora <ch> não seja utilizado pela convenção ortográfica para representar o som [s]. Além dessa possível relação fonético-fonológica, que, em alguma medida, pode ter sido apreendida pelo escrevente, há possibilidade de a dimensão gráfico-visual também permear a troca entre a letra <c> e o dígrafo <ch>, pois é comum haver certa instabilidade no registro de letras e dígrafos (<c> e <ch>; <n> e <nh>, <l> e <lh>), nos registros escritos de crianças em fase inicial de alfabetização. Assim, ao grafar com <ch> para representar o som [s], por razões fonético-fonológicas e/ou gráfico-visuais, o escrevente é levado a reconhecer que a sequência que se forma “chei” não tem nenhuma correspondência gráfica, mas que, ao ser acrescida da letra *a*, formaria uma palavra escrita convencional: *cheia*;

- ii. outras hipóteses a serem levantadas (em consonância com a primeira) diz respeito à sequência seguinte após a ocorrência do dado híbrido: a referida sequência consiste, também, de uma segmentação não-convencional, mais especificamente, uma hipossegmentação: “jogala”. Pela leitura do texto é possível depreendermos, para essa passagem, o sentido de “jogar lá”, formada de duas palavras fonológicas. Contudo, a maneira como o escrevente registra essa sequência, deixa pistas de que a palavra “lá” tenha sido interpretada como o possível clítico “la”, muito frequente em finais de verbos. Uma vez que o elemento “la” formava com o verbo “jogar” uma única palavra, para o escrevente, o elemento, também átono, que antecedia a palavra, ou seja, a partícula “a”, não faria parte dessa palavra. Em relação, ainda, aos aspectos prosódicos dessa ocorrência, destacamos que a sequência *jogar lá* forma uma

frase fonológica reestruturada (cf. NESPOR E VOGEL, 1986), sendo artigo *a*, da sequência *comecei a*, também parte dessa frase fonológica, representada como: [comecei]ϕ [a jogar lá]ϕ. Assim, a sequência não-convencional “come cheia” rompe a possibilidade prevista de organização prosódica em sintagmas.

O outro dado híbrido destacado no texto de Bruno é “a olado”. Diferentemente da ocorrência anterior, no caso de “a olado” não encontramos nenhuma grafia convencional, dessa sequência que nos permitisse contrastar com a não-convencional. Nesse caso, a identificação da ocorrência como um dado híbrido se deu, somente, pela observação do espaço em branco entre as letras *a* e *o* que se mostra claramente maior. Por meio do enunciado em que ocorre o dado, ou seja: “eu jogava **a olado** de Neymar, Robinho e P. H. Ganso”, podemos interpretar que a grafia não-convencional tenha sido motivada por uma forma de atribuição do estatuto de palavra escrita ao *a*, separando-o, por essa razão, do elemento *o* que a acompanhava, unindo esse, por sua vez, à próxima palavra do enunciado. Soma-se a essa hipótese o fato de o uso do *ao*, tanto em enunciados falados quanto em enunciados escritos, ser pouco frequente, dificultando, assim, a interpretação gráfica dessa sequência como uma palavra escrita. Talvez, uma segunda motivação possa residir no destaque que, por hipótese, observamos pelos usos não-convencionais do espaço em branco, sobre a informação de que o escrevente tenha trabalhado com jogadores importantes do futebol brasileiro; assim, na parte hipossegmentada “olado”, o clítico *o*, ao estar funcionando como uma só unidade de sentido junto à palavra “lado”, põe em destaque a presença do escrevente na partida de futebol como um jogador tão importante como aqueles que ele destaca: estar *ao lado* dos seus ídolos na mesma condição deles e não apenas como um admirador.

Gostaríamos de destacar, ao fim, que, embora dados de híbrido tenham ocorrido uma só vez e na produção de um só aluno, revelando não ser uma característica dos dados de

segmentação não-convencional produzidos nos anos mais avançados de escolarização, esses dados se diferenciam em relação aos tipos de híbrido encontrados no EF-I, conforme é possível observar em Cunha (2004). Essa autora, apesar de também não ter identificado muitos dados desse tipo, percebeu, no entanto, que, quando as crianças os propunham, elas primeiro hipossegmentavam duas palavras para depois hipersegmentar parte da segunda palavra como em: “pofa vor” (para, *por favor*) – Cunha, (2004, p. 112). Como vimos nos dados do cópús desta pesquisa: “come cheia” e “a olado”, o escrevente primeiramente hipersegmenta para depois hipossegmentar uma dada sequência de palavras. A comparação entre esses dois tipos de híbrido mostra, novamente, particularidades do EF-II em relação ao EF-I, fato que endossa a importância deste estudo.

### ***Resumo da Seção***

No decorrer desta Seção, expusemos os resultados obtidos a partir da análise das hipersegmentações, os quais demonstram características gerais e particulares das grafias não-convencionais ainda encontradas nos textos de escreventes dos anos finais do EF. Ao realizarmos a descrição dos dados, notamos que: (i) houve diminuição das ocorrências de hipersegmentação no decorrer dos anos escolares; (ii) grande parte dos escreventes investigados terminou o EF-II atribuindo, às palavras, fronteiras gráficas além da prevista pela convenção, embora as hipersegmentações tenham se revelado pouco frequentes com o passar dos anos, quando analisadas em função do tipo de palavra escrita hipersegmentada e da extensão das produções de textos; (iii) as palavras que seguiram sendo hipersegmentadas foram aquelas em que foi possível relacionar a sílaba pretônica com possíveis classes gramaticais (por exemplo, preposição, conjunção, pronome); e (iv) informações linguísticas

de natureza prosódica e ortográfica, indícios da inserção do escrevente em práticas orais/faladas e letradas/escritas, foram pontos de ancoragem para a determinação de um limite gráfico não-convencional em alguma posição da palavra.

Referente à análise das hipersegmentações que denominamos “particulares”, a trajetória dessas grafias mostra, na escrita dos escreventes (Bruno e Pedro), um movimento de aproximação com a convenção ortográfica, na medida em que a sua ocorrência passa a ser mais esporádica e localizada em uma só produção escrita por ano escolar, por exemplo. Quanto à motivação, se, por um lado, as grafias “particulares” produzidas, em especial, pelo escrevente Bruno demonstram, ao longo do percurso escolar, identificação com as hipóteses gerais que observamos para o EF-II, por outro lado, as grafias desse escrevente não deixam de guardar relação com um processo de delimitação gráfica das palavras semelhante, por exemplo, àqueles observados por estudiosos da escrita inicial. Essa última característica observamos, também, na grafia hipersegmentada produzida por Pedro.

Embora as hipersegmentações produzidas por Bruno e Pedro tenham nos chamado atenção em relação à quantidade e à configuração dessas grafias, observamos, pela análise dos textos, características comuns dos dados, que vão ao encontro das regularidades observadas para os demais escreventes da mesma etapa escolar, de modo que, em alguma medida, podemos supor que os alunos do EF-II partilham hipóteses linguísticas muito próximas acerca da noção de palavra escrita. Destacamos, então, que as particularidades identificadas na produção de Bruno e Pedro dão evidências de que o processo de aquisição da escrita não é igual mesmo entre escreventes que partilham características comuns como faixa etária e tempo de escolarização, já que a aquisição da escrita é, sempre, um processo marcado pela relação do sujeito com a linguagem (CAPRISTANO, 2007b). Desse modo, talvez, aquilo que

viemos conceituando por “particular” seja, ao fim, marca da história de linguagem que constitui individualmente cada um dos escreventes.<sup>78</sup>

---

<sup>78</sup> Certamente uma análise indiciária da produção escrita (como fez Capristano (2007b), por exemplo) teria mais argumentos para sustentar as diferenças intersujeitos explicadas pela consideração da história de linguagem. Como essa questão extrapola os objetivos deste estudo, limitamo-nos a sinalizá-la, acenando para futuras pesquisas.

## SEÇÃO 5:

### **Considerações finais**

---

Nesta dissertação, buscamos compreender, por meio de grafias de hipersegmentação, a maneira como alunos regularmente matriculados na etapa final do EF lidam com a noção de palavra escrita. Em função das hipersegmentações de palavras serem um fenômeno pouco investigado em produções escritas de escreventes com mais anos de escolarização, objetivamos, por um lado, a realização de uma descrição mais geral das características linguísticas das hipersegmentações presentes em textos do EF-II, e, por outro, a análise de ocorrências que, em grande parte das vezes, foram únicas em todo o corpus e não seguiram regularidades encontradas.

Ao observarmos a distribuição das grafias não-convencionais, chegamos ao resultado de que as hipersegmentações caracterizaram-se por apresentar uma diminuição ao longo dos anos escolares do EF-II. Esse fato ficou mais evidente quando colocamos em análise não só os dados de hipersegmentação, mas, conjuntamente, as palavras escritas que propiciaram as fronteiras gráficas não-convencionais. Notamos, como consequência dessa decisão, que os escreventes tiveram dúvida quanto aos limites das mesmas palavras e que, em decorrência, surgiram maiores números de dados entre um ano escolar e outro. Assim, não houve somente o decréscimo do número de dados, mas, também, os tipos de palavras que geram dúvidas quanto aos seus limites. Quanto à trajetória longitudinal de cada aluno, identificamos que grande parte deles iniciou e terminou o EF-II produzindo hipersegmentações. No entanto, as ocorrências se tornaram menos frequentes no decorrer dos anos, quando cotejada a quantidade de dados em relação à extensão dos textos produzidos (dada pelo número de palavras escritas)

e às dificuldades em grafar tipos específicos de palavra (como *demais*, por exemplo) ou tipos de palavras que apresentaram sílabas correspondentes (como, *amigo*, *apareceu*, *apaixonado*, por exemplo).

Buscamos compreender, também, os critérios linguísticos subjacentes às grafias de hipersegmentação. Por meio da organização dos dados a partir da estrutura prosódica das palavras convencionais, identificamos pistas de informações prosódicas e (orto)gráficas. Sobre o aspecto prosódico, as fronteiras gráficas não-convencionais puderam ser relacionadas a fronteiras de constituintes prosódicos, com destaque, para os constituintes sílaba e pé métrico. Na maioria das hipersegmentações, apreendemos informações fonológicas importantes ligadas a esses constituintes, como a presença do acento prosódico, que se constituiu em um lugar privilegiado para a inserção de um limite gráfico não-convencional na fronteira em que esteve presente essa proeminência, como em, por exemplo, “em bora” (*embora*). Com efeito, essas pistas prosódicas, inferidas por meio das hipersegmentações, permitiram-nos observar os momentos em que a escrita dos escreventes foi capturada pelo imaginário de uma suposta gênese da escrita (CORRÊA, 2004), como tentativa de plasmar “o registro integral dessa prosódia no escrito” (CORRÊA, 2004, p. 137). Compreendemos que a ação da prosódia sobre o tipo de dado analisado, sinaliza a participação dos escreventes em práticas de linguagem orais/faladas e deixa em evidência um imaginário o qual “põe às claras a falsa pureza da escrita” (CORRÊA, 2004, 83).

Interligadamente, vimos, nas hipersegmentações, o atravessamento de informações linguísticas (orto)gráficas, as quais estiveram relacionadas à inserção dos escreventes em práticas letradas/escritas, mais especificamente naqueles momentos em que essa inserção se revela na imagem do código escrito institucionalizado (CORRÊA, 2004). Para ilustrar esse aspecto, destacamos grafias como “com migo” e “em tão”, por exemplo, a partir das quais inferimos, com base em uma sequência ao mesmo tempo fônica e gráfica, possíveis palavras

escritas aliadas a generalizações de regras ortográficas que, ao buscarem organizar e distribuir os grafemas nas palavras escritas, dão pistas dos limites de palavra. Outras hipersegmentações, como as derivadas da presença não-convencional do hífen, deram pistas de informações letradas/escritas, na medida em que se basearam na tentativa de registrar formas verbais subjuntivas (por exemplo, *ganhasse*) como se fossem duas unidades relacionadas sintaticamente (por exemplo, “ganha-se”). Foi possível observarmos que, em todas as situações descritas, foram informações relativas à própria escrita que estiveram subjacentes às hipóteses acerca dos limites de palavras, fundamentadas no que o escrevente “supõe ser – a partir não só do que aprendeu na escola, mas, em grande parte, do que assimilou fora dela – a visão escolarizada do código institucionalmente reconhecido” (CORRÊA, 1997, p. 271), já que demonstram estarem envolvidas pelo “caráter reprodutor de uma prática instituída, [pela] tentativa de alçamento à escrita culta formal e [pelos] discursos estabilizados da instituição escolar” (CORRÊA, 2004, p. 167).

A respeito das hipersegmentações que não se encaixaram na organização geral dos dados, observamos essas grafias localizadas nas produções de apenas dois dos escreventes investigados (Bruno e Pedro), indicando que essas ocorrências podem ser compreendidas como hipóteses particulares sobre a noção de palavra escrita “reveladoras [por sua vez] das singularidades dos sujeitos e da relação por eles estabelecida com a linguagem” (ABAURRE, 1997, p. 24). Vimos, ainda, por meio do processo interpretativo desses dados, semelhanças com processos de segmentação em palavras observados em enunciados escritos produzidos no início do processo formal de aquisição da escrita e, ao mesmo tempo, semelhanças com momentos relacionados aos aspectos mais gerais observados em todo o corpus. Além disso, a observação de muitas das grafias particulares, vinculadas ao texto de que foram extraídas, permitiu-nos observar que o modo não-convencional de segmentação das palavras possibilitou, ao escrevente, reconhecer e demarcar, para seu interlocutor, elementos centrais

em que se edificaram as informações essenciais e sob as quais se construíram o texto. Nesse sentido, concordamos com Abaurre (1988a) acerca da não-categoricidade das hipóteses que formulam os aprendizes de escritas e, assim, não é estranho que, muitas vezes, tivemos encontrado “soluções escritas diferentes para um mesmo problema de segmentação em um mesmo texto” (ABAURRE, 1991c, p. 205).

Com base nos resultados alcançados, compreendemos que a utilização de recursos gráficos, com o intuito de dividir um enunciado escrito em palavras, continua, em alguma medida, sendo um desafio para escreventes que já completaram o período da alfabetização. Entendemos, ainda, que reconhecer os limites de palavra implica, ao escrevente da língua, lidar com a complexidade da própria definição do conceito de palavra, a qual não se restringe à forma de registro escrito e perpassa várias esferas da linguagem. Desse modo, para “acertar” onde começa e termina uma palavra escrita, o escrevente deve estar atento a diferentes critérios linguísticos; mais ainda: deve saber como convergi-los de modo a alçar às exigências da ortografia, valorizadas, sobretudo, socialmente.

Chegamos, neste momento final, vislumbrando a relevância de nossos resultados para a compreensão dos processos de segmentação não-convencional de palavras que ainda ocorrem no período do EF-II, na medida em que, por meio deles, podem ser traçados paralelos que permitem observar semelhanças e diferenças quanto ao modo de delimitação nas palavras em etapas distintas do processo escolar. Acreditamos, por tudo o que foi discutido, que o eixo central de quaisquer comparações que possam vir a ser realizadas entre nossos resultados e os dados de escrita infantil, por exemplo, devem levar em conta, sobretudo, a natureza linguística dos dados, a qual guarda relação com diferentes formas de mobilizar o relacionamento entre o oral/falado e o letrado/escrito que capturam os escreventes em momentos da sua escolarização, deixando à mostra as imagens que têm sobre a

constituição da escrita, a qual, junto à Corrêa (2004), defendemos, neste trabalho, ser heterogênea.

me despede!!!

## Referências bibliográficas

---

ABAURRE, Maria Bernadete Marques. O que revelam os textos espontâneos sobre a representação que faz a criança do objeto escrito? In: KATO, Mary (Org.). *A concepção da escrita pela criança*. 2. ed. Campinas: Pontes Editores, 1988a, p. 135-142.

\_\_\_\_\_. A relevância dos critérios prosódicos e semânticos na elaboração de hipóteses sobre segmentação na escrita inicial. *Boletim da ABRALIN*, Campinas, v. 11, 1991c, p. 203-17.

\_\_\_\_\_. (Re)escrevendo: o que muda? In: ABAURRE, Maria Bernadete Marques; FIAD, Raquel Salek; MAYRINK-SABINSON, Maria Laura. *Cenas de aquisição da escrita: o sujeito e o trabalho com o texto*. Campinas: Mercado das Letras, 1997.

\_\_\_\_\_. Posfácio: A aquisição da escrita do português – considerações sobre diferentes perspectivas de análise. In: ROJO, Roxane. *Alfabetização e letramento: perspectivas linguísticas*. São Paulo: Mercado das Letras, 1998, p. 205-232.

\_\_\_\_\_; CAGLIARI, Luiz Carlos. Textos espontâneos na primeira série: evidências da utilização, pela criança, de sua percepção fonética da fala para representar e segmentar a escrita. *Cadernos CEDES: Antropologia e Educação Interfaces do Ensino e da Pesquisa*, São Paulo, v. 14, 1985, p. 25-29.

\_\_\_\_\_; GALVES, Charlotte. Os clíticos no português brasileiro: uma abordagem sintático-fonológica. In: CASTILHO, Ataliba Teixeira de; BASÍLIO, Margarida. (Org.) *Gramática do português falado: estudos descritivos*. Vol. IV. Campinas: UNICAMP, 1996, p. 273-320.

\_\_\_\_\_; FIAD, Raquel Salek; MAYRINK-SABINSON, Maria Laura. Em busca de pistas. In: \_\_\_\_\_. *Cenas de aquisição da escrita: o sujeito e o trabalho com o texto*. Campinas: Mercado das Letras, 1997, p. 13-36.

BISOL, Leda. Constituintes prosódicos. In: \_\_\_\_\_. *Introdução a estudos de fonologia do português brasileiro*. Porto Alegre: EDIPUCRS, 1996, p. 247-261.

\_\_\_\_\_. Mattoso Câmara Jr. e a palavra prosódica. *DELTA*, São Paulo, v. 20, n. especial, 2004, p. 59-70.

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

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

BOOP DA SILVA, Taís. *Formação de palavras compostas em português brasileiro: uma análise de interfaces*. 195f. Porto Alegre: Tese (Doutorado em Letras) Universidade Federal do Rio Grande do Sul, 2010.

BRASIL. Secretaria de Educação Fundamental. *Parâmetros Curriculares Nacionais: língua portuguesa*. Brasília: 1997.

BRISOLARA, Luciene Brassols. *Os clíticos pronominais do português brasileiro e sua prosodização*. Porto Alegre. 189f. Tese (Doutorado em Linguística Aplicada). PUCRS, 2008a.

CÂMARA JR., Joaquim Mattoso. *Estrutura da Língua Portuguesa*. Petrópolis, Rio de Janeiro: Vozes, 1970.

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

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

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

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

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

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

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

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

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

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

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

\_\_\_\_\_. *O modo heterogêneo de constituição da escrita*. São Paulo: Martins Fontes, 2004.

\_\_\_\_\_. Heterogeneidade da escrita: a novidade da adequação e a experiência do acontecimento. *Filologia e Linguística Portuguesa*, São Paulo, v. 8, 2007, p. 269-286.

\_\_\_\_\_. Heterogeneidade da escrita no ensino: das modalidades às relações intergenéricas. In: PEREIRA, L.A; CARDOSO, I. (Org.). *Reflexão sobre a escrita: ensino de diferentes gêneros de textos*. 1 ed. v. 1. Aveiro (PT): Editora da Universidade de Aveiro (PT), 2009.

COX, Maria Inês Pagliarini; ASSIS-PETERSON, Ana Antônia de. A palavra: uma história de dissonâncias entre professores e aprendizes de escrita. In: \_\_\_\_ (Org.). *Cenas de sala de aula*. Campinas: Mercado das Letras, 2001, p. 51-79.

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

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

\_\_\_\_\_; MIRANDA, Ana Ruth Moresco. Índícios de reestruturação do conhecimento fonológico da criança em dados de reparo na escrita inicial. Porto Alegre. *Letras de hoje*, v. 48, n. 3, 2013, p. 343-354.

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

GOODY, Jack. *Domesticação do pensamento selvagem*. Tradução de Nuno Luis Madureira. Lisboa: Presença, 1988/[1977].

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

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

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

LONGHIN-THOMAZI; Sanderléia TENANI, Luciani. *Oficinas de leitura, interpretação e produção textual no Ensino Fundamental*. Inédito. 15f., 2013.

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

MASSINI-CAGLIARI, Gladis; CAGLIARI, Luiz Carlos. Categorização gráfica e funcional na aquisição da escrita e da leitura em língua materna. *Calidoscópico (UNISINOS)*. São Leopoldo, v. 2, n.1, 2004, p. 89-94.

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

\_\_\_\_\_; MATZENAUER, Carmen. Aquisição da fala e da escrita: relações com a fonologia. *Cadernos de educação (UFPel)*. Pelotas, v. 35, 2010, p. 359-405.

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

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

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

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

QUADROS, Emanuel Souza de; SCHWINDT, Luiz Carlos. Um estudo sobre a relação entre palavra morfológica e palavra fonológica em vocábulos complexos do português brasileiro. *Anais do CELSUL*, Porto Alegre, 2008, p. 1-18.

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

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

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

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

SILVA, Lilian Maria da. A relação entre dados de segmentação não-convencional de palavras em textos de quinto e nono anos do Ensino Fundamental. *Relatório final de pesquisa de iniciação científica*. Inédito. FAPESP: Proc. 2010/17720-8. 2011.

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

\_\_\_\_\_; RAUPP, Julhiana da Silva. Trocas entre letras que representam consoantes com oposição surda/sonora. *Verba Volant (UFPel)*. Pelotas, v. 3, n. 2, 2010, 107-123.

STREET, Brian. *Literacy in Theory and Practice*. Cambridge: Cambridge University Press, 1984.

TENANI, Luciani. Rindo das piadas, manipulando a língua. *Alfa*. São Paulo, v. 45, 2001, 115-127.

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

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

\_\_\_\_\_. Notas sobre a relação entre constituintes prosódicos e a ortografia. *Revista de estudos linguísticos*, Belo Horizonte, v. 16, n. 1, jan./jun. 2008, p. 231-245.

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

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

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

\_\_\_\_\_. Aspectos segmentais e prosódicos da escrita de crianças e adolescentes: evidências de relações entre enunciados falados e escritos. *Relatório final de pesquisa*. 44f. Inédito. FAPESP: Proc. 2009/14848-6. 2011.

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

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

\_\_\_\_\_. Hipersegmentação de palavras: análise de aspectos prosódicos e discursivos. *Linguagem & Ensino*, Pelotas, v.16, n.2, jul./dez. 2013, p. 305-326.

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

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

ZORZI, Jaime Luiz. *A apropriação do sistema ortográfico nas 4 primeiras séries do primeiro grau*. Campinas. 124f. Tese (Doutorado em Educação). Universidade Estadual de Campinas, 1997.

## Anexos

### Anexo 1:

**Tabela 10:** Estruturas rítmicas de palavras mobilizadas no 6º ano

Estruturas rítmicas				Exemplo de hiper	Total	
Tipos	Palavra	Subtipos	Hipersegmentação		Nº	%
(1)	(●*)	(1.1)	(●)σ (*)σ	“então” > “em tão”	8	20
(2)	(*●)	(2.0)	(*)σ (●)σ	“desse” > “de se”	6	15
(3)	(●*●)	(3.1)	(●)σ (*●)Σ	“enquanto” > “em quanto”	16	40
		(3.2)	(●*)Σ (●)σ	“quisesse” > “quise-se”	2	5
		(3.3)	(*)σ (*●)Σ	“sozinho” > “so zinho”	1	2,5
(4)	(*●*)	(4.0)	(*)σ (●*)Σ	“apanhou” > “a panho”	1	2,5
(5)	(*●●●)	(5.2)	(*)σ (●*●)Σ	“acabando” > “a cabano”	2	5
(6)	(*●●*)	(6.0)	(*)σ (●●*)Σ	“aparecer” > “a perecer”	1	2,5
	(●*●*)		(●)σ (*●*)Σ			
(8)	Outros			“adicionar” > “adisi o na”	3	7,5
<b>Total</b>					<b>40</b>	<b>100</b>

Leia-se: \*: sílaba forte; ●: sílaba fraca; σ: sílaba; Σ: pé métrico.

Fonte: Elaboração própria

### Anexo 2:

**Tabela 11:** Estruturas rítmicas de palavras mobilizadas no 7º ano

Estruturas rítmicas				Exemplo de hiper	Total	
Tipos	Palavra	Subtipos	Hiper		Nº	%
(1)	(●*)	(1.1)	(●)σ (*)σ	“demais” > “de mais” <sup>79</sup>	9	37
(2)	(*●)	(2.0)	(*)σ (●)σ	“quando” > “qual do”	2	8,4
(3)	(●*●)	(3.1)	(●)σ (*●)Σ	“daquela” > “da quela”	6	25
		(3.2)	(●*)Σ (●)σ	“falando” > “falam do”	2	8,4
(5)	(*●●●)	(5.1)	(*)σ (●*)Σ	“recompensa” > “recom pensa”	1	4,1
(7)	(*●●*●)	(7.0)	(*)σ (●●*●)Σ	“aparecido” > “a paresido”	1	4,1
	(●*●*●)		(●)σ (*●*●)Σ			
(8)	Outros			“voltei” > “foute i”	3	12,5
<b>Total</b>					<b>24</b>	<b>100</b>

Leia-se: \*: sílaba forte; ●: sílaba fraca; σ: sílaba; Σ: pé métrico.

Fonte: Elaboração própria

<sup>79</sup> No levantamento dos dados, identificamos várias hipersegmentações com a configuração semelhante à da palavra “demais”, ou seja, formadas por um ditongo crescente, cuja característica nos enunciados falados, é a de poder se realizar, também, como hiato. Em nossa análise, porém, consideramos tais palavras como constituídas por um ditongo, por isso a nossa interpretação como sendo hipersegmentações de duas sílabas.

## Anexo 3:

**Tabela 12:** Estruturas rítmicas de palavras mobilizadas no 8º ano

Tipos	Estruturas rítmicas			Exemplo de hiper	Total	
	Palavra	Subtipos	Hipersegmentação		Nº	%
(1)	(●*)	(1.1)	(●)σ (*)σ	“supor” > “su por”	9	33,5
		(1.2)	(*)σ (●)σ	“nenhum” > “nem um”	1	3,6
(3)	(●*●)	(3.1)	(●)σ (*●)Σ	“amanhã” > “a manha”	7	26,0
		(3.2)	(●*)Σ (●)σ	“falassem” > “fala-sem”	2	7,5
(4)	(*●*)	(4.0)	(*)σ (●*)Σ	“alugar” > “a lugar”	4	15
(5)	(*●*●)	(5.1)	(*●)σ (*●)Σ	“loucamente” > “loca mente”	1	3,6
(6)	(*●●*) (●*●*)	(6.0)	(*)σ (●●*)Σ	“apaixonei” > “a paichonei”	1	3,6
			(●)σ (*●*)Σ			
(7)	(*●●*●) (●*●*●)	(7.0)	(*)σ (●●*●)Σ	“aniversário” > “a liversario”	1	3,6
			(●)σ (*●*●)Σ			
(8)	Outros			“voltei” > “foute i”	1	3,6
<b>Total</b>					<b>27</b>	<b>100</b>

Leia-se: \*: sílaba forte; ●: sílaba fraca; σ: sílaba; Σ: pé métrico.

Fonte: Elaboração própria

## Anexo 4:

**Tabela 13:** Estruturas rítmicas de palavras mobilizadas no 9º ano

Tipos	Estruturas rítmicas			Exemplo de Hipersegmentação	Total	
	Palavra	Subtipos	Hipersegmentação		Nº	%
(1)	(●*)	(1.1)	(●)σ (*)σ	“assim” > “a sim”	6	27,3
(3)	(●*●)	(3.1)	(●)σ (*●)Σ	“ouvindo” > “o vino”	6	27,3
(4)	(*●*)	(4.0)	(*)σ (●*)Σ	“acender” > “a cender”	1	4,5
(5)	(*●*●)	(5.1)	(*●)σ (*●)Σ	“contra bando” > “contra bando”	2	9,1
		(5.3)	(*●*)Σ (●)σ	“conversasse” > “conversa-se”	3	13,6
(8)	Outros			“conversando” > “com ver sando”	4	18,2
<b>Total</b>					<b>22</b>	<b>100</b>

Leia-se: \*: sílaba forte; ●: sílaba fraca; σ: sílaba; Σ: pé métrico.

Fonte: Elaboração própria