

# **HIPERSEGMENTAÇÕES DE PALAVRAS NO ENSINO FUNDAMENTAL**

**LILIAN MARIA DA SILVA  
E LUCIANI TENANI**

**HIPERSEGMENTAÇÕES  
DE PALAVRAS NO  
ENSINO FUNDAMENTAL**

CONSELHO EDITORIAL ACADÊMICO  
Responsável pela publicação desta obra

Erotilde Goreti Pezatti  
Fabiana Cristina Komesu  
Douglas Altamiro Consolo  
Maria Cristina Parreira da Silva  
Mircia Hermenegildo Salomão (representante discente)

LILIAN MARIA DA SILVA  
LUCIANI TENANI

# HIPERSEGMENTAÇÕES DE PALAVRAS NO ENSINO FUNDAMENTAL

**CULTURA  
ACADÊMICA**   
*Editora*

© 2014 Editora Unesp

**Cultura Acadêmica**

Praça da Sé, 108

01001-900 – São Paulo – SP

Tel.: (0xx11) 3242-7171

Fax: (0xx11) 3242-7172

www.editoraunesp.com.br

www.livrariaunesp.com.br

feu@editora.unesp.br

CIP – BRASIL. Catalogação na publicação  
Sindicato Nacional dos Editores de Livros, RJ

---

S581h

Silva, Lilian Maria da

Hipersegmentações de palavras no ensino fundamental [recurso eletrônico] / Lilian Maria da Silva, Luciani Tenani. – 1. ed. – São Paulo: Cultura Acadêmica, 2014.

Recurso digital

Formato: ePub

Requisitos do sistema: Adobe Digital Editions

Modo de acesso: World Wide Web

ISBN 978-85-7983-583-4 (recurso eletrônico)

1. Língua portuguesa – Ortografia. 2. Língua portuguesa (Ensino fundamental). 3. Livros eletrônicos. I. Tenani, Luciani. II. Título.

14-18126

CDD: 469.152

CDU: 811.134.3'354

---

Este livro é publicado pelo Programa de Publicações Digitais da Pró-Reitoria de Pós-Graduação da Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp)

Editora afiliada:



Asociación de Editoriales Universitarias  
de América Latina y el Caribe



Associação Brasileira de  
Editoras Universitárias

## AGRADECIMENTOS

Pelos comentários e sugestões valiosas, agradecemos a Ana Ruth Moresco Miranda, Cristiane Carneiro Capristano e Lourenço Chacon.

Pelo financiamento concedido para o desenvolvimento da pesquisa, agradecemos à Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp).



# SUMÁRIO

Apresentação 9

Primeiras considerações 11

1 Noção de palavra 17

2 Um olhar para a escrita 67

3 Procedimentos metodológicos e constituição do *corpus* 87

4 Análise das hipersegmentações 107

Últimas considerações 153

Referências bibliográficas 157

Sobre as autoras 163





# APRESENTAÇÃO

Esta obra é resultado da dissertação de mestrado *Um estudo longitudinal sobre as hipersegmentações de palavras escritas nos anos finais do Ensino Fundamental*, vinculada ao Programa de Pós-Graduação em Estudos Linguísticos do Instituto de Biociências, Letras e Ciências Exatas (Unesp – São José do Rio Preto), defendida em 2014. Nela é apresentada uma descrição de grafias de hipersegmentações de palavras (por exemplo, “em bora” – embora; “mora-va” – morava) encontradas na escrita de estudantes dos quatro últimos anos do Ensino Fundamental. Na literatura sobre o tema, o objeto de estudo circunscreve-se às segmentações não convencionais de palavras encontradas nos anos iniciais dessa etapa escolar. Portanto, aqui são analisados dados de escrita sobre os quais não foram encontrados estudos anteriores.

A análise desenvolvida fundamenta-se na busca por pistas linguísticas que possam explicar a ocorrência de grafias não convencionais de fronteiras de palavras no período posterior ao da alfabetização, ou seja, 6º ao 9º anos do Ensino Fundamental, quando essas grafias não seriam esperadas. Essas ocorrências são fortemente motivadas, do nosso ponto de vista, pela complexidade linguístico-discursiva da palavra. A reflexão realizada neste livro está alicerçada na premissa de que a noção de palavra escrita

é marcada por essa complexidade, a qual se revela não apenas no registro escrito da linguagem, mas se apresenta como um problema linguístico.

Além dos resultados obtidos em relação às ocorrências de hipersgmentação, este livro traz reflexões teóricas sobre noções de palavra, de escrita, e sobre a relação entre fala e escrita. Essas reflexões permitem não só observar a complexidade linguístico-discursiva do objeto investigado, bem como fornecem subsídios para o tratamento de dados não convencionais de escrita para além de uma postura teórico-metodológica que os concebe como simples erros ortográficos.

As discussões apresentadas nesta obra contribuem tanto para os estudos linguísticos sobre a unidade básica palavra quanto para os estudos da área de Educação e Fonoaudiologia que lidam com os chamados erros ortográficos. Acreditamos que atende aos interesses de professores, fonoaudiólogos e pesquisadores da linguagem interessados em entender o funcionamento da escrita, notadamente da palavra escrita, e das relações entre fala e escrita.

*As autoras*

## PRIMEIRAS CONSIDERAÇÕES

Hipersegmentações são registros não convencionais de fronteiras de palavras, caracterizadas pela presença de um recurso gráfico (espaço em branco ou hífen) no interior de palavras ortográficas. No bojo de uma reflexão mais ampla, elas compõem uma categoria maior de dados, a qual abriga outras possibilidades de distribuição dos limites da palavra na escrita distantes da convenção. 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> Trazemos, a seguir (figuras 1, 2 e 3), exemplos de cada um dos tipos de segmentação não convencional mencionados.

---

1 Todas as referências feitas aos termos “segmentação” e/ou “segmentação não convencional” correspondem, neste livro, 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).

2 Cunha (2004, 2010) identificou dados híbridos em que primeiro ocorreu a hipossegmentação e depois a hipersegmentação. Porém, em nosso *corpus*, as ocorrências híbridas apresentaram primeiro a hipersegmentação e depois a hipossegmentação. Desse modo, consideramos a possibilidade de haver dados

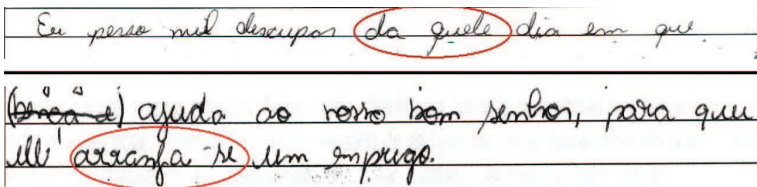


Figura 1 – Fonte: Banco de dados de escrita do EF II: Z09\_6B\_26M\_07 e Z11\_8B\_15F\_04.

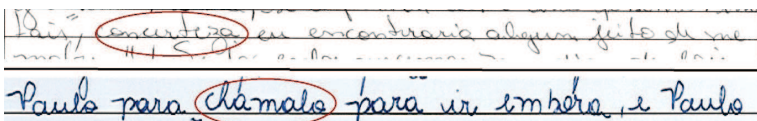


Figura 2 – Fonte: Banco de dados de escrita do EF II: Z08\_5C\_39F\_03 e Z08\_8A\_07F\_01.

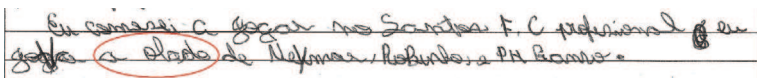


Figura 3 – Fonte: Banco de dados de escrita do EF II. Z10\_7B\_05M\_02.

híbridos que envolvam tanto um caso como o outro. É importante lembrar que, a partir de um ponto de vista distinto, Chacon (2004) denominou de “mesclas” as 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 hipossegmentação ou hipersegmentação, uma vez que a ausência/presença do limite gráfico mostra-se “em partes nas quais deveria (ou se suporia) haver um *limite* ao mesmo tempo ortográfico e prosódico” (p.226; destaque 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 – p.229) toda a sequência sugeriria a ação de um constituinte frasal, como a frase fonológica, e as partes separadas entre limites, constituintes menores do que a palavra fonológica, como o pé métrico (com a configuração de um iambo), no caso do exemplo apresentado. Não é nosso propósito 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 o mesmo critério da autora para a definição de hipersegmentação de palavra.

É pertinente destacar que, enquanto objeto de investigação, as segmentações não convencionais têm despertado o interesse de diferentes áreas do saber, como a Fonoaudiologia (cf. Zorzi, 1997), a Educação (cf. Cox; Assis-Peterson, 2001), a Linguística (cf. Chacon, 2006; Tenani, 2011b) e, ainda, estudos na interface Linguística–Educação (cf. Cunha, 2010; Cunha; Miranda, 2013). Por sua vez, cada um dos campos do conhecimento tem questões específicas de pesquisa e aborda o fenômeno aqui investigado segundo propósitos de análise distintos. Com relação à Linguística (em particular, no viés norteador de reflexão adotado), as marcas de segmentação não convencional têm auxiliado na argumentação de discussões referentes à organização da fala em constituintes prosódicos; à relação constitutiva entre fala e escrita e oralidade e letramento; às formulações construídas pelos escreventes sobre os limites (orto)gráficos de palavras.

No que concerne ao tipo de material em que se identificam segmentações não convencionais, são bastante frequentes estudos realizados com base em produções de textos infantis (cf. Abaurre, 1988a, 1991c; Silva, A., 1991; Cunha, 2004; Capristano, 2003; Chacon, 2005, 2006).<sup>3</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 dos anos finais do Ensino Fundamental (Ensino Fundamental II).

As reflexões contidas neste livro, nesse aspecto, pretendem contribuir para a área de pesquisas sobre essa etapa escolar, à semelhança do que, em épocas recentes, vem sendo realizado por Tenani (2009a, 2010, 2011b) e Paranhos e Tenani (2011). A importância de investigar dados do Ensino Fundamental II justifica-se não apenas pelo fato de serem poucos os estudos realizados, mas principalmente por já se esperar do escrevente, nesse período de escolarização, a produção de textos “*com domínio da separação em palavras*, estabi-

---

3 Ainda em contexto inicial de aquisição da escrita, existem trabalhos que têm como objeto a escrita de jovens e adultos em período de alfabetização (ver Tenani, 2008, e Ferreira, 2011, por exemplo).

lidade 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” (PCNs, 1997, p.80; destaque nosso).

Por meio de nossos dados, mostramos que escreventes que já concluíram, em termos formais, o período dos anos iniciais da alfabetização em ambiente escolar continuam a ter dificuldades para reconhecer os limites da palavra escrita. Assim, nosso objetivo principal diz respeito a investigar hipóteses linguísticas que permanecem e/ou passam a ser mobilizadas nos anos mais avançados do Fundamental, os atuais 6º a 9º anos.<sup>4</sup> Com o objetivo de buscar fatores linguísticos subjacentes às grafias de hipersegmentação, consideramos que a noção de palavra escrita é complexa e o domínio da convenção ortográfica exige do escrevente conhecimentos linguísticos de distintas naturezas (fonológico, morfossintático/semântico e também letrado).

No que diz respeito aos objetivos específicos desta obra, com base em uma amostra longitudinal de produções escritas, almejamos realizar:

- uma descrição das características gerais das hipersegmentações, primeiro para observar a distribuição dos dados ao longo dos anos escolares do Ensino Fundamental II e os escreventes investigados, segundo para que possamos apontar possíveis regularidades linguísticas das grafias não convencionais encontradas;
- uma análise de dados de hipersegmentação que, em relação à descrição geral deles, não se assemelhe às tendências identificadas, constituindo-se, em grande parte das vezes, em grafias únicas do *corpus*.

---

4 A Lei nº 11.274 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, passou a contar com um ano a mais. As antigas 1ª a 4ª séries passaram a denominar-se 1º a 5º anos, e as antigas 5ª a 8ª séries, 6º a 9º anos.

Por fim, cabe explicitar que, ao assumirmos a complexidade da noção de palavra escrita e do domínio da ortografia, 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>5</sup> enquanto *sujeito* escrevente” (Capristano, 2007b, p.160; destaques no original).

---

5 A palavra “criança”, nesse texto, possui significado mais abrangente, referindo-se aos “aprendizes de escrita”.





# 1

## NOÇÃO DE PALAVRA

O tratamento das grafias de hipersegmentação perpassa por uma questão: o que é palavra? No campo dos estudos linguísticos, essa não é uma questão nova. São bastante antigas as reflexões sobre a palavra como objeto de pesquisa. Ainda que diferentes níveis de análise, como a morfologia, a fonologia, a sintaxe, a semântica, a lexicologia, a lexicografia, tenham se dedicado a estudá-la, não há definição consensual a respeito do que é a palavra, e uma das razões para isso está relacionada aos interesses específicos de cada área. Uma das consequências da divergência com relação aos aspectos considerados na formalização do conceito de palavra é que os seus limites não são, em grande parte das vezes, coincidentes entre si, conforme apresentaremos aqui, ao considerar a noção de palavra nos níveis morfológico e fonológico.

Neste capítulo, discutimos a noção de palavra, tomando como ponto de partida estudos sobre dados de segmentação não convencional, a fim de tratar de registros da noção de palavras em enunciados escritos. Os trabalhos apresentados no capítulo, ainda que não tenham se preocupado especificamente em discutir o conceito de palavra, apontam, em suas análises, para fatos linguísticos subjacentes à sua conceituação, os quais teriam mobilizado, de modo mais proeminente, as grafias discordantes da convenção

ortográfica. Para além de uma reflexão mais ampla acerca da concepção de palavra escrita, esses trabalhos permitem a identificação de diferentes formas de análise do objeto que investigamos. Nesse sentido, mostrou-se fundamental revisitar pesquisas sobre os tipos de segmentação não convencional, pois isso permitiu posicionar-nos perante uma posição teórico-metodológica de análise mais coerente com nossos propósitos de investigação e concepção de escrita assumida (tratada no Capítulo 2), e demonstrar a relevância das segmentações não convencionais para os estudos da linguagem.

Em seguida, abordamos a noção de palavra fonológica. A partir da discussão sobre os limites de palavra fonológica, demonstramos que a definição de palavra é, antes de tudo, um desafio linguístico, pois é difícil identificar, também nos enunciados falados, as fronteiras de palavras, por serem instáveis e/ou difíceis de identificar. Consideramos que, ao serem registradas palavras na escrita, essa instabilidade de fronteiras é marcada por meio de grafias como as hipersegmentações, por exemplo. Dito de outro modo, a reflexão proposta sobre a relação dos estudos fonológicos sobre o conceito de palavra e os tipos de dados analisados parte, neste livro, 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.

## **Registros escritos de palavras: grafias de segmentação não convencional**

O eixo central das discussões, neste livro, está respaldado na compreensão das hipersegmentações como fruto 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 as considere como simples desvios da escrita padrão.

Assim, apresentamos, a seguir, considerações que permitem perceber a importância de dados de segmentação não convencional para os estudos da linguagem.

## **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 no início da aquisição da escrita. Analisar esse tipo de material mostrou-se importante especialmente porque o modo como os aprendizes analisariam e representariam a escrita 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. Para esses autores, nesse tipo de texto as crianças revelariam suas hipóteses em relação ao novo objeto sobre o qual estão refletindo, ou seja, a escrita.<sup>1</sup> Ao tomar o texto infantil como objeto de análise, uma das questões investigadas passou a ser a segmentação em palavras. O interesse por esse aspecto gráfico voltou-se predominantemente para a des-

---

1 É interessante esclarecer a ideia de texto espontâneo proposta por Abaurre e Cagliari (1985), 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 1ª 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 as distanciam progressivamente da noção de *texto*, para elas tão natural em termos de produção oral, quando ingressam na escola” (p.25; destaque no original).

coberta dos critérios linguísticos que guiariam os aprendizes no reconhecimento dos limites das palavras escritas.

Interessada nos dados de segmentação não convencional, Abaurre (1988a) sinaliza a importância da relação entre fala e escrita para uma melhor compreensão das hipóteses infantis acerca dos limites das palavras. A autora aponta a crença comum de que, nas primeiras produções escritas, os aprendizes tendem a escrever do modo como falam (visão que parece perdurar até os dias atuais) e afirma que essa consideração é infundada, pois, como ela própria 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” (p.137, destaques nossos).

Com efeito, essas considerações de Abaurre deixam entrever que a relação entre fala e escrita, a qual pode ser observada por meio dos registros escritos não convencionais dos mais variados tipos, é bem mais complexa do que pressupõe a ideia de que “a criança escreve como fala”. Na aceitação de tal complexidade, trabalhos mais recentes defendem 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). Na verdade, esses trabalhos consideram que a escrita forma-se heterogeneamente e, por isso, aspectos do falado presentes no produto escrito não são marcas de interferência, mas pontos de emergência daquilo que, acima de tudo, a constitui.<sup>2</sup> Assim, podemos considerar 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.

---

2 A proposta de considerar a relação fala/escrita como constitutiva do modo de enunciação escrito será abordada no capítulo seguinte, em virtude de se tratar do eixo central de organização da noção de escrita que embasou a pesquisa que originou este livro.

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 investigativo tem sido desenvolvida grande parte dos estudos do Grupo de Pesquisa Estudos sobre a Linguagem (GPEL/CNPq)<sup>3</sup> e Grupo de Estudos sobre a Aquisição da Linguagem Escrita (Geale)<sup>4</sup>. Os diversos trabalhos realizados por esses grupos têm mostrado pontos de intersecção entre dados não convencionais de escrita e fatos fonológicos. As conclusões de tais pesquisas serão apresentadas adiante, ao delinearmos os diferentes tipos de hipóteses explicativas a respeito dos modos de análise dos dados de segmentação não convencional.<sup>5</sup>

Outro ponto relevante para a discussão linguística refere-se aos limites da palavra escrita. Concordamos com Abaurre (1991c), que propõe como problema trivial da escrita a necessidade de se conhecer os critérios convencionais que definem palavra, para que assim as palavras possam ser separadas entre limites. A transformação simbólica operada pela escrita faz que, muitas vezes, negligenciemos “a relação 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

---

3 Grupo de pesquisa coordenado pelos professores Lourenço Chacon (FFC/Unesp) e Luciani Tenani (Ibilce/Unesp).

4 Grupo de pesquisa coordenado pela professora Ana Ruth Moresco Miranda (FaE/UFPeL).

5 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).

maneira, ‘surdos’ para as características mais contínuas dos enunciados orais” (Abaurre, 1991c, p.203) e ouvimos 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 é possível 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; Cagliari, 1985, p.27).

Como discutiremos a seguir, 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 palavras são um indício dessa instabilidade pouco notada (e tantas vezes esquecida) por parte dos usuários da escrita, sobretudo quando esses tipos de dados permitem recuperar justamente as faltas de correspondência entre as diferentes formas de palavras, formas essas que, por sua vez, em seus respectivos níveis de análise, também constituem um problema.

Sintetizamos os pontos de estudos relevantes abordados nesta seção: questões referentes à relação fala/escrita; comprovação de pressupostos de teorias fonológicas; aspectos acerca da noção de palavra. É fundamental destacar que esses pontos foram abordados em separado por razões puramente expositivas, com o intuito de chamar a atenção para a importância linguística do fenômeno das segmentações não convencionais, visto que, em outros estudos sobre o assunto (sobretudo os mais recentes), todos os aspectos estão considerados de modo interligado.

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

## Perspectivas de análise das hipersegmentações

Ao entender as hipersegmentações como fenômeno de linguagem, buscamos, em pesquisas de orientação linguística, hipóteses explicativas que oferecessem interpretações para tais ocorrências. Foi possível identificar, em todos os trabalhos consultados, a relação entre a distribuição não convencional dos limites de palavras e fatos linguísticos de natureza prosódica. Identificamos nos estudos analisados um desdobramento de tendências prototípicas para a compreensão das motivações que guiam os escreventes na proposição de segmentações distantes das convenções ortográficas.

Uma primeira tendência tem apoio em argumentos baseados no aspecto fonético da prosódia e a relação desse aspecto com fatos de natureza semântica (daqui em diante, denominaremos essa tendência de orientação explicativa fonético-semântica), encontrada, por exemplo, nos trabalhos de Abaurre e Cagliari (1985), Abaurre (1988a), 1991c) e A. Silva (1991). Uma segunda tendência está respaldada em fatos fonológicos da prosódia, em conexão com informações letradas (que denominaremos de tendência fonológico-letrada), e pode ser identificada, por exemplo, nas pesquisas de Capristano (2003, 2004), Chacon (2004, 2005, 2006), Tenani (2004, 2009a, 2010, 2011b) e Paula (2007). Consideramos 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 relação a questões pedagógicas.

A partir deste ponto trataremos de cada uma das tendências de modo mais detalhado.

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 da palavra escrita. Para os autores, essas hipóteses não são categóricas e/ou estáveis, porque se, por um lado, em alguns momentos, as grafias não convencionais parecem guiadas por aspectos fonéticos da linguagem – em especial quando, ao tomar como eixo norteador a pronúncia das palavras, produzem escritas baseadas em grupos



tonais,<sup>6</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 em seus textos, à medida que entram em contato com outros textos escritos. Com base nessa consideração, os autores apontam 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” (p.28), como eles destacam, por exemplo, em relação à grafia “ador mesida” (adormecida).

Abaurre reforça a argumentação de que as hipóteses infantis não se mostram de forma decisiva para todos os problemas de escrita. Mais ainda, nesse trabalho, a autora parece ampliar a rede de possibilidades que poderiam explicar as segmentações para 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.135-6).

Na sua análise, contudo, a autora continua a privilegiar os critérios fonético e semântico, ao explicar os motivos das ausências

---

6 Grupo tonal é uma unidade fonética “maior do que o pé (por isso pode ser composto de um ou mais pés) e é delimitado por um padrão entoacional específico chamado tom” (Silva, A., 1991, p.40).

e presenças não convencionais dos limites gráficos. Em relação ao primeiro critério, aponta a correspondência entre grafias de hipossegmentação e unidades fonéticas, como em “tocuazimoreino” (estou quase morrendo) e “guenãovazia” (que não fazia), em que são observados, respectivamente, limites de grupo tonal e de força.<sup>7</sup> Com relação ao critério semântico, Abaurre (1988a) considera que existe uma ação da escrita sobre as decisões infantis não só nas propostas de segmentação, mas também 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” (p.140).

Em publicação posterior, Abaurre (1991c) aborda a ideia de representações canônicas da língua, ao discutir 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 a autora, muitas vezes, ao delimitarem as palavras, as crianças, em suas hipóteses, pareceram combinar o padrão 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 “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” (p.204).

Ainda, para 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 trocaicos (configuração que alterna

---

7 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).

uma sílaba forte com uma fraca), unidades nas quais se fundamenta um padrão rítmico preferencial do português brasileiro, como bem sinaliza a autora. Note-se que, a partir dessa interpretação, ela já assinala aspectos fonológicos subjacentes às hipóteses linguísticas sobre as fronteiras de palavras escritas.

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

Ao observar a organização dos textos infantis, a autora sinaliza que determinados aspectos da produção escrita, relacionados à disposição dos elementos da narrativa, estão mais propensos à ocorrência de segmentações não convencionais. Abaurre observa, por exemplo, que as crianças escrevem de modo mais convencional, no tocante à segmentação em palavras, quando, ao desenvolverem o enredo, colocam-se como narradoras de terceira pessoa, e parecem segmentar menos quando demarcam a fala das personagens por meio do discurso direto. Nesses momentos, surge maior número de hipossegmentações, o que pode ser explicado pela aproximação dos diálogos com os enunciados falados, que caracterizam-se pela continuidade; por esse aspecto, os diálogos seriam registrados em porções gráficas maiores. No entanto, a autora salienta: essa possível hipótese explicativa não significa afirmar que as palavras escritas restantes do texto são segmentadas, pelas crianças, ao encontro da convenção.

Grande parte das considerações formuladas por Abaurre (1988a, 1991c) estão reunidas e sistematizadas em A. Silva (1991). Este autor desenvolveu um estudo baseado na escrita de crianças da antiga 1ª série (atual 2º ano), no qual buscou observar a relação da fala com as ocorrências de hipo e de hipersegmentação. Nota-se,

contudo, que os resultados que obteve 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, o autor ressalta a atuação de grupos tonais e de força, em especial nos casos em que o espaço em branco não é empregado. Ele 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, identifica, nesse tipo de segmentação, uma busca da criança em registrar aspectos da escrita convencional, a exemplo de: casos em que as crianças inferem, a partir de unidades como “o”, “a”, “um”, “da”, “de”, “em”, “com”, possíveis palavras escritas de modo independente, como ocorre em “da cela” (daquela); casos em que a interrupção da escritura, para acrescentar determinado símbolo à palavra (como um acento gráfico), 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” (roncou); casos em que a delimitação das fronteiras não convencionais se dá no reconhecimento 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, A. Silva (1991) destaca hipersegmentações como “com tente” (contente). Mas essa nova hipótese não implica o abandono das outras já apontadas, no sentido de que as hipóteses explicativas podem conviver.

No entanto, essa convivência não aconteceria em um mesmo dado, visto que é possível “prever em que circunstâncias ocorrem, ou qual é a mais preponderante” (p.61). Não adotaremos, neste livro, a visão desse autor, pois, do nosso ponto de vista, não há evidências que sustentem que o aprendiz, em um ou outro mo-

mento, dê maior ênfase a um critério, em detrimento de outro. Entendemos que diferentes hipóteses atuam de modo simultâneo, porém, no momento, não nos ateremos a essa reflexão, pois teremos mais subsídios que a sustentam quando analisarmos os trabalhos de orientação analítica que definimos como fonológico-letrada.

Reconhecemos a relevância dos estudos de Abaurre e Cagliari (1985), Abaurre (1988a, 1991c) e A. Silva (1991), não apenas por serem os primeiros estudos brasileiros a abordarem o tema das segmentações não convencionais do ponto de vista linguístico, mas, fundamentalmente, porque suas formulações explicativas mostram-se distantes de uma perspectiva de análise essencialmente normativa.

Notamos que, ao se distanciarem dessa visão, os pesquisadores mencionados, 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, lugar de observação do trabalho do sujeito sobre a linguagem. As contribuições trazidas por esses trabalhos revelam sua importância, ainda, pelo fato de suas discussões terem se constituído como ponto de partida para as reflexões mais recentes sobre o tema, ampliando até mesmo o interesse de estudiosos sobre esse aspecto em produções escritas de escrevantes com mais anos de escolarização e também de adultos em processo de alfabetização. Estamos nos referindo aos estudos que denominamos como de orientação explicativa fonológico-letrada e que retomaremos na sequência.

Os trabalhos de Capristano (2003, 2004), Chacon (2004, 2005, 2006), Tenani (2004, 2009a, 2010, 2011b) e Paula (2007) adotam a concepção do modo heterogêneo de constituição da escrita, conforme Corrêa (1997, 2004). Como resultado dessa escolha, esses autores têm identificado, nos tipos de segmentação não convencional, informações linguísticas relativas às práticas sociais de linguagem orais/faladas e letradas/escritas. Para fins explicativos, eles demonstram que, em diferentes graus, as hiposegmentações e as hipersegmentações obedecem aos princípios reguladores “estabelecidos para a definição de constituintes prosódicos, tais como aque-

les propostos por Nespor e Vogel (1986)”<sup>8</sup> (Chacon, 2004, p.223), e também a estruturas possíveis dentro da convenção ortográfica. No entanto, Chacon (2005) faz uma observação importante: ao considerar que a análise das segmentações não convencionais permitiria ao investigador identificar vários outros aspectos linguístico-discursivos advindos de práticas sociais orais/letradas, privilegiar os aspectos prosódicos e ortográficos constitui apenas uma decisão metodológica que busca analisar mais atentamente a atuação da prosódia e da própria convenção ortográfica nas hipóteses construídas por escreventes em processo de aquisição do que constitui o limite de palavra escrita.

Um dos nossos intuitos, ao revisitar as linhas de análise dispensadas ao nosso objeto de estudo, como já mencionado, foi adotar aquela que mais dialogasse com os nossos objetivos e as nossas escolhas teóricas. Nesse cenário, esta pesquisa filia-se aos trabalhos desenvolvidos com base em análises fonológico-letradas, cujas contribuições apresentamos agora. Iniciamos com o estudo de Capristano (2003). Na análise proposta, com base em textos de crianças da 1ª série (atual 1º ano), a autora reúne as marcas de segmentação não convencional (hipossegmentação e hipersegmentação) em tipos de funcionamentos que apreenderiam, do seu ponto de vista, propriedades dos dados que permitiriam “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; destaque no 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

---

8 Todos os trabalhos apresentados nesta seção valeram-se de informações fonológicas para explicar estruturas de segmentação não convencional e basearam-se no modelo de fonologia prosódica de Nespor e Vogel (1986).

abaixo do domínio da palavra fonológica da hierarquia prosódica (sílabas e pé) e informações sobre o código escrito institucionalizado; 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. Ressalte-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 outras publicações, 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 a autora, o uso dos espaços gráficos não convencionais permite, de modo geral, concluir que as grafias são sempre resultado de uma ação simultânea e, em diferentes graus, 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 hipossegmentada e hipersegmentada no funcionamento analisado e que lhe permitiram propor a sua conclusão mais geral. O primeiro fator funda-se em aspectos prosódicos, mais especificamente, em fronteiras de constituintes prosódicos que, ao assumirem certo destaque para as crianças, levariam-nas a 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 segundo funcionamento (segmentações não convencionais resultantes da oscilação entre diferentes

trânsitos por constituintes prosódicos e informações sobre o código escrito institucionalizado), relacionando os textos infantis com as produções de adultos em processo de 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; destaque nosso). A esse respeito, a autora argumenta que a possibilidade de sentido não é garantida na escrita apenas pela delimitação das palavras, mas também por meio da disposição das sílabas proeminentes nos enunciados. Ela lembra que é desse mecanismo de compreensão dos sentidos que se valem, por exemplo, textos chistosos,<sup>9</sup> porque

entender a piada pede a descoberta de possíveis similaridades na cadeia fônica que estão “estrategicamente” articuladas com as semelhanças 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)

Ao traçar um paralelo entre textos chistosos e aqueles produzidos durante o processo de aquisição da escrita (por crianças e adultos), Tenani assinala que, na tarefa de delimitar palavras escritas – em especial no caso de escreventes que se supõe que não dominam 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 que não se trata de coincidência que, com grande frequência, ocorram diferentes segmentações para uma mesma palavra, no interior de um mesmo texto, ou em textos diferentes de um mesmo

---

9 Em trabalho anterior, Tenani (2001) dedicou-se exclusivamente à análise de mecanismos fonológicos na construção de textos chistosos.



escrevente. Ela enfatiza a não estabilidade e homogeneidade da língua, entendendo que os dados não convencionais de escrita são relevantes por permitirem mostrar “‘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): 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.

Chacon (2004) 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 que não parecem explicar-se “com base em algoritmos de formação como aqueles que definem constituintes da hierarquia prosódica, tal como proposta por Nespor e Vogel (1986)” (p.225). No entanto, ele não percebe, por essas 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, o que poderia ser explicado pela ação de práticas de letramento sobre as crianças, tanto ligadas diretamente à escolarização como desenvolvidas fora desse contexto, mas que do mesmo modo envolvem 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” (a chorar) e “ficoa paxonado” (ficou apaixonado) apenas enquanto estruturas 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 negar por completo o funcionamento da prosódia no enunciado escrito, na medida em que em uma afirmação dessa natureza estaria subjacente a postura de 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 existiria, antes de tudo, um funcionamento simultâneo (e não hierárquico) entre os constituintes prosódicos. A observação dos dados demonstra, de modo mais amplo, que toda a sequência não convencional indicia a ação de constituintes maiores (como aqueles localizados acima da palavra fonológica), enquanto, de modo mais específico, as partes demarcadas entre limites indiciam fronteiras de constituintes menores (como aqueles formados abaixo da palavra fonológica).

Para Chacon, a simultaneidade de funcionamento seria desencadeada por uma não sincronização entre limites **de** constituintes e de proeminências **em** constituintes; 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; por uma dissensão entre os dados de linguagem (as segmentações não convencionais, nesse caso) e os padrões dos modelos de análise; e/ou, ainda, 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 de modo mais detalhado o aspecto da ação mais latente do ritmo e menor de outros fatores prosódicos. No primeiro trabalho, o autor analisa grafias de trissílabos hipersegmentados produzidos por crianças em início de aquisição da escrita, por perceber na estrutura de palavras compostas por três sílabas fatos mais interessantes, relacionados a questões prosódicas, do que em palavras polissílabas e dissílabas, porque: dados a partir de palavras polissílabas apareceram com baixa frequência no *corpus* analisado pelo autor; as ocorrências de palavras dissílabas, embora em termos numéricos tenham acontecido em quantidade maior do que as de 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, revelando que apenas o constituinte prosódico sílaba foi relevante nessas ocorrências.

O autor parte da hipótese de que qualquer ponto de um trissílabo pode oferecer contexto a um possível lugar de separação não convencional. Porém, as rupturas não ocorreram em quaisquer lugares. Pelo contrário, o modo como os espaços em branco foram inseridos sugere 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; destaque nosso). Esses contrastes aparecem na língua, segundo o autor, em decorrência 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. Dentre as 136 ocorrências levantadas, 111 (81,62%) apresentaram a ruptura entre os limites de sílabas e de pés, sendo que 92 (82,88%) promoveram a combinação de uma sílaba e um pé, como em “e Rita” (irrita), e 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” (saude).

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 sugerem 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, em boa parte das vezes, corresponde a categorias gramaticais, como mostra a sílaba à esquerda na grafia hipersegmentada da palavra “porquinho”: “**por** quinho”. Mas, quando não é identificado um monossílabo, o reconhecimento pode ocorrer com o dissílabo, como na palavra “cola” a partir da hipersegmentação de “escola”: “s **cola**”.

Como último destaque sobre as considerações do autor, apon-tamos o refinamento ortográfico de alguns dados. Para ilustrar esse aspecto, retomamos a 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 gra-far com a letra <e> um segmento que, em enunciados falados mais informais (esperados 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 o nome Rita.<sup>10</sup>

Posteriormente, Chacon (2006) aborda as pistas linguísticas que podem ser observadas nas hipersegmentações e que marcam a pre-sença 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 chega nessa análise são muito semelhantes àquelas do estudo anterior (Chacon, 2005). Com relação à aquisição da noção de palavra, o autor afirma que ela não apresenta apenas questões estruturais para a reflexão do es-crevente. “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 ver-

---

10 A escolha de letras para registrar os limites das segmentações (tanto conven-cionais quanto não convencionais) foi alvo da discussão de Tenani (2009b). A autora mostra, ao analisar textos de alunos dos dois primeiros anos do Ensino Fundamental II, que a maneira como o escrevente seleciona as letras do alfa-beto é um indício do modo como informações de práticas orais e letradas constituem hipóteses acerca dos limites da palavra. Com esse estudo, Tenani (2009b) conclui que o escrevente flutua tanto em relação às possibilidades de segmentação (convencional ou não) quanto no que se refere às possibilidades de preenchimento de dada posição da palavra por 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).

bal: o falado e o escrito” (Chacon, 2006, p.165). Ainda sobre esse aspecto, 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 e fora do contexto escolar), as quais tornam passível o acesso do sujeito escrevente não só à noção de palavra como também aos diferentes fatos de língua, tanto em seu modo de enunciação falado, quanto em seu modo de enunciação escrito. (Chacon, 2005, p.165)

No seu trabalho, Paula (2007) também analisa apenas dados de hipersegmentação. A autora realiza um estudo longitudinal de textos de oito crianças dos três primeiros anos do Ensino Fundamental. Com base nesse material, 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 pé, os dois domínios prosódicos localizados abaixo do nível da palavra. Sobre esses constituintes, foi mais frequente o uso do espaço em branco entre duas sílabas: “que ro” (quero) e entre uma sílaba e um pé métrico: “es tava” (estava), em particular quando esse último apresentou a configuração de um troqueu. A autora, 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 passou a escrita de cada escrevente.

O primeiro movimento (que envolveu a escrita de cinco crianças investigadas) caracteriza-se por apresentar, de início, uma forma de escrita comprometida com relação à atribuição de sentido às palavras, em especial pelo fato 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 permitir atribuir algum sentido às palavras, dificultou a identificação dos espaços em branco, por causa da caligrafia. Ao longo dos anos escolares, esse primeiro movimento desapareceu das produções escritas dos escreventes, visto que a distribuição dos espaços em branco e a atribuição de sentido passaram a se mostrar de modo mais claro.

A autora verifica, em relação ao segundo movimento (presente na escrita de duas crianças), a convivência entre uma escrita na qual os blocos de letras comprometem a identificação do sentido e outra em que os espaços em branco organizam-se irregularmente. A detecção desse movimento foi se tornando esporádica, sendo mais frequente nas produções escritas da 1ª série.

Em relação ao terceiro movimento, a autora observa que apenas uma criança apresentou-o. Em termos das suas características, esse movimento delinea-se a partir de uma escrita que poderia aproximar-se 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 nele agrupadas.

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)

Tenani (2009a, 2010, 2011a, 2011b), seguindo o mesmo eixo teórico-metodológico de análise dos trabalhos de Capristano (2003, 2004), Chacon (2004, 2005, 2006) e Paula (2007), desenvolve reflexões acerca dos tipos de segmentação não convencional que podem ser encontrados na produção escrita de alunos dos anos finais do Ensino Fundamental. Em suas pesquisas, a autora procura mostrar que a dificuldade para identificar os limites de palavra escrita não está circunscrita aos escreventes em início de aquisição da escrita (adultos ou crianças), mas é percebida também na escrita de escreventes com mais tempo de escolarização. Desse modo, em seus trabalhos, Tenani vem procurando investigar, juntamente com outras pesquisadoras que receberam orientação acadêmica da autora (cf. Paranhos, 2010, e Silva, L., 2011), quais soluções linguísticas escreventes já saídos do período de alfabetização utilizam para delimitar as palavras. A partir 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 Ensino Fundamental II convergem e/ou divergem em relação aos produzidos no início da aquisição da escrita. Com relação ao aspecto prosódico, a autora tem se baseado nos constituintes prosódicos – grupo clítico, palavra fonológica, pé métrico e sílaba – para a 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 fono-

lógicos do português brasileiro. Ao analisar marcas de hipossegmentação e hipersegmentação, Tenani argumenta que as grafias encontradas nos anos que encerram o Ensino Fundamental caracterizam-se por colocar em evidência o modo de organização dos elementos átonos da língua. Em outras palavras, a dúvida que se mostra frequente nas hipóteses dos escreventes é saber quando os elementos átonos são sílabas (pretônicas ou postônicas) de uma palavra fonológica e 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 interpretaram a cadeia fônica como uma só unidade lexical. Prosodicamente, esse resultado corresponde a um grupo clítico compreendido como uma palavra fonológica. Nas ocorrências de hipersegmentação, a presença do espaço em branco é uma pista de que apenas uma 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. Diante dessas análises, Tenani conclui que há flutuação entre os dois constituintes prosódicos (palavra fonológica e grupo clítico), explicada, em parte, em função da imprecisão dos estatutos de sílabas átonas e de clíticos fonológicos.<sup>11</sup>

No final do seu trabalho, a autora reúne aquilo que considera as principais contribuições dos resultados que obteve, aqui sintetizadas:

- indícios de hipóteses do que seja palavra: para Tenani, as hipóteses dos escreventes acerca dos limites de palavras são ancoradas tanto em informações de natureza prosódica, cuja evidência relaciona-se à organização de elementos átonos

---

11 Paranhos e Tenani (2011) 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 Ensino Fundamental II.



(sílabas ou clíticos) em diferentes constituintes, quanto em informações de natureza letrada, a qual parece mostrar-se especialmente no uso do espaço em branco para indicar possíveis palavras escritas, em particular no caso das hipersegmentações;

- pistas da reflexão dos escreventes sobre classes gramaticais: a configuração dos dados evidencia, segundo a autora, que os escreventes ainda encontram dificuldade em identificar itens gramaticais como preposição, pronome e conjunção. A dúvida consiste exatamente em identificar quando os elementos átonos são itens gramaticais e quando são sílabas de uma palavra;
- evidências da organização prosódica da língua: com relação a este aspecto, Tenani enfatiza a flutuação dos escreventes em organizar prosodicamente as sílabas átonas, ora como parte de uma palavra fonológica, ora como parte de um grupo clítico. Como já demonstramos, nas hipossegmentações essa flutuação relaciona-se ao fato de o clítico ser interpretado como sílaba da palavra. De modo inverso, nas hipersegmentações a flutuação revela-se pelo fato de a sílaba da palavra ser registrada como um clítico. Com relação a esse último tipo de segmentação, Tenani observa que as hipersegmentações evidenciam não só a leitura de uma sílaba pretônica como um clítico, correlacionado, por sua vez, a uma possível apreensão desses elementos como palavras independentes na escrita, mas também porque a sua estrutura evidencia 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 anteriormente (Tenani, 2009a). Nesse trabalho, a autora aproxima-se das interpretações de Chacon (2005) e de Paula (2007), centradas nos 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 nos quatro últimos anos do Ensino Fundamental.

Os resultados obtidos pela 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. Com relação à configuração prosódica dos dados, foi interessante observar que a ruptura ocorreu no ponto adjacente à sílaba de maior proeminência sonora da palavra, ou seja, no limite do acento primário (destaque-se que, das cinco estruturas descritas, quatro apresentaram essa motivação). A autora aponta, ainda, que, dentre todas as combinações, a sequência 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, o que já foi observado por Abaurre (1991c), Chacon (2005) e Paula (2007) para dados de escrita inicial.

Cabe ressaltar 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 sugerem a tentativa de autonomia gráfica de partes de palavras, como 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 o Ensino Fundamental I e o Ensino Fundamental II, no que tange à segmentação de palavras, fundam-se em informações de natureza letrada. Segundo a autora, essa afirmação parece ficar mais evidente 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 Fundamental II). Verifica-se que, no Fundamental I, nem sempre a sílaba pretônica, grafada à esquerda do espaço

em branco, coincide com algum clítico correspondente, por sua vez, a alguma classe gramatical, característica que sempre foi observada nos dados de hipersegmentação do Fundamental II, conforme reforça Tenani. Desse modo, para a autora, 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 esses 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 Fundamental II, ao serem analisadas sistematicamente nos diversos estudos de Tenani, despertaram a atenção da 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 uso de hífen relacionam-se, no caso das hipossegmentações (a maioria dos dados encontrados), a ocorrências em estruturas verbo + clítico, como em “despistalos” (despistá-los); no caso das hipersegmentações, dois resultados convivem, o primeiro ligado a estruturas de palavras compostas, por exemplo, “mau-humor”, e o segundo relacionado a 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 nos anos finais do Ensino Fundamental.

Finalizamos com a apresentação 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 apresentados anteriormente

a concepção de escrita fundamentada no modo de enunciação heterogêneo de constituição da escrita. Para além dessa diferença teórica, os trabalhos ora referidos, embora de cunho linguístico, ao fim de suas análises buscam estabelecer relação entre os resultados que obtiveram com práticas pedagógicas de ensino e aprendizagem da língua materna, a propósito do que encontramos em Cunha:

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. (Cunha, 2010, p.15)

Inseridas nas reflexões desenvolvidas pelo 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>12</sup> é elemento chave para a descoberta

---

12 Nos trabalhos do Geale, a ideia de erro está respaldada na teoria piagetiana, para a qual o erro, de qualquer natureza, é parte integrante do processo de aprendizagem. Assim, ele é relevante, na medida em que demonstra as hipóteses construídas pelos aprendizes na sua relação com o objeto a aprender. Ressalte-se que essa noção é distinta da qual se vincula, por exemplo, o trabalho de Zorzi (1997), que abordaremos no próximo capítulo. Para esse autor, o erro, na aquisição da escrita, pode ser indicio de um problema ligado a uma patologia de aprendizagem, em especial 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).

de hipóteses das crianças sobre o sistema de escrita que estão a adquirir. (Miranda, 2009b, 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>13</sup> tonicidade<sup>14</sup> e tipo de sílaba<sup>15</sup>) e extralinguísticas (série e tipo de escola), Cunha observa, em relação à variável tipo de palavra, uma tendência comum entre as hipossegmentações e as hipersegmentações: em ambas, a ausência/presença do espaço em branco ocorre entre uma palavra gramatical e uma palavra fonológica, como nos exemplos: “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 escrita, terem dificuldades em reconhecer palavras formadas por uma ou duas letras como unidades autônomas da língua, sobretudo quando se trata de palavras desprovidas de acentuação fonológica.

Já nos dados relativos à hipersegmentação, a tendência descrita mostra o contrário, ou seja, a criança, ao identificar a sílaba inicial de uma palavra como uma possível palavra gramatical, procura dar-lhe certa autonomia gráfica. Com relação às demais variáveis linguísti-

---

13 Cunha aponta as possíveis combinações com essa variável: palavra gramatical + palavra fonológica; palavra fonológica + palavra gramatical; palavra gramatical + palavra gramatical; palavra fonológica + palavra fonológica. Por palavra gramatical, a autora entende “aquela que engloba segmentos que possuem apenas significado gramatical, como os clíticos” (Cunha, 2004, p.62), e, por palavra fonológica, 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” (p.62-3).

14 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.

15 A autora observa dois aspectos relacionados a essa variável: ocorrências de possíveis processos de ressilabificação (a partir da aplicação de processos de sândi); preservação ou não da estrutura da sílaba.

cas, Cunha chega a este resultado: 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); 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.

Ao analisar os dados em função das variáveis extralinguísticas, Cunha (2004) 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” (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 a autora, os resultados que obteve são relevantes para a prática pedagógica que se desenvolve no período da alfabetização, pois auxiliam os professores a compreender as hipóteses que as crianças formulam sobre a escrita, as quais fundamentam-se em conhecimentos linguísticos, e não constituem simples problemas de ortografia. Para a autora, a criança é um sujeito ativo no processo de aprendizagem e 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 português brasileiro e do 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” (p.17). Com vistas a atingir o propósito principal de sua pesquisa, a autora adota estes procedimentos:

- Descrever os tipos de segmentação não convencional de palavras produzidas 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 “escondeuce” (escondeu-se). Já nas hipersegmentações, a separação não convencional se dá com a formação de uma sequência de palavra gramatical e palavra fonológica, por exemplo, “da quela” (daquela). Com relação aos 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.
- Cotejar os resultados encontrados em Cunha (2004), com relação 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 português brasileiro e do português europeu no que se refere às ocorrências que envolvem a combinação entre palavra fonológica e gramatical. No brasileiro, as hipossegmentações decorrem, em sua maioria, da união entre uma palavra gramatical e uma palavra fonológica, enquanto no europeu a junção ocorre entre palavra fonológica e palavra gramatical, revelando uma mudança importante, que é a da direcionalidade em que o clítico associa-se à palavra. Em relação às hipersegmentações, tanto o português brasileiro quanto o europeu se assemelham, visto que as separações marcam uma combinação de palavra gramatical e fonológica.
- Buscar, nas grafias de segmentação não convencional do português brasileiro e do europeu, evidências de processos fonológicos. Em ambas as variedades, a autora verificou que as ocorrências de hipossegmentação e híbrido proporcionam contextos favoráveis à aplicação de processos fonológicos

como: 1) vozeamento da fricativa – no português brasileiro: quando as junções ocorreram entre uma palavra gramatical e uma fonológica; no português europeu: apenas um dado registrou o processo, nos limites de duas palavras fonológicas; 2) *tapping* – em ambas, o contexto favorável foram hipossegmentações em que, à esquerda, ocorreu a palavra “por”; 3) processos de sândi vocálico – a degeminação e a ditongação, no português brasileiro, foram aplicadas em hipossegmentações que uniram um pronome a um verbo (ou seja, uma palavra gramatical a uma fonológica); a elisão ocorreu em poucos casos, não tendo sido identificada no português europeu, e a degeminação aplicou-se mesmo quando a sílaba alvo do processo era a portadora do acento primário; a ditongação ocorreu em poucos casos, talvez pelo fato de a realização fonética da vogal /i/, em contexto final de palavra, ser sempre /e/ no português europeu; 4) haplologia – este processo não aconteceu no português europeu e foi bastante raro no brasileiro, ficando restrito aos casos de reestruturação silábica; 5) epêntese e monotongação – ambos os processos foram incomuns nas fronteiras de palavras e, quando ocorreram, favoreceram a formação da estrutura silábica CV, tanto no português brasileiro quanto no europeu.

- 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 mostram-se mais suscetíveis ao acento secundário, enquanto os dados de hipersegmentação e híbridos pareceram encontrar motivação no pé portador do acento primário, em especial da configuração do troqueu.

A partir 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, apud Cunha, 2010) sobre o estudo das diferenças rítmicas entre o português brasileiro e o europeu. Para essas autoras, o ritmo, nas duas variedades, dis-



tingue-se em relação à hierarquização de três restrições: integridade da palavra fonológica, binariedade do pé e pé trocaico. No português europeu, o ritmo caracteriza-se pela hierarquia: 1) pé trocaico, 2) integridade da palavra fonológica, 3) binariedade do pé; no português brasileiro: 1) integridade da palavra fonológica, 2) binariedade do pé, 3) pé trocaico (cf. Abaurre; Galves, 1998, p.393-4, apud Cunha, 2010, p.72-3).

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 português europeu, último lugar no português brasileiro), pois, em relação à integridade da palavra fonológica, tanto no português brasileiro quanto no europeu ela precede a posição da binariedade do pé. Desse modo, entende-se que “o ritmo em português europeu é baseado no troqueu, enquanto o ritmo brasileiro se constrói respeitando, antes de mais nada, as fronteiras de palavras fonológicas” (Abaurre; 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 une-se à palavra adjacente, por exemplo, a palavra “te” nas hipossegmentações “para tener melhor” (no português brasileiro) e “parate ver melhor” (no europeu). Como já observamos, as hipossegmentações, no português brasileiro, seguiram a tendência de união entre uma palavra gramatical e uma palavra fonológica e, no europeu, mostraram preferência pela união entre uma palavra fonológica e uma palavra gramatical, como ilustram os exemplos anteriores. No entanto, o que diferencia a hipossegmentação “parate ver melhor”, no caso do português europeu, é o fato de o monossílabo “te” ter sido unido a uma preposição e não a um verbo; a união com essa categoria constitui a maior parte das ocorrências identificadas nessa variedade, a qual reflete a sintaxe da fala.

De acordo com Cunha, o que explica a união não convencional de “para tener melhor” e “parate ver melhor” é o ritmo (as barras incli-

nadas indicam limites de grupos rítmicos): para // *tever* // *milhor* (português brasileiro) e *parate* // *ver me* // *lhor* (europeu). Com base na divisão em grupos rítmicos sugerida por Abaurre e Galves, Cunha considera que, em “*parate ver melhor*”, o monossílabo “*te*”, em função da sua atonicidade, insere-se no grupo anterior, formando com ele um pé trocaico,<sup>16</sup> o que justifica, ainda, a presença da sílaba “*me*” (também átona) junto com o grupo rítmico a que pertence “*ver*”. Mas, segundo Cunha (2010), não se espera a grafia de “*me*” hipossegmentada a “*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” (p.145).

Sobre o grupo representado em “*para tener melhor*”, a autora entende que a sua formação, além de ser esperada, pela preferência do português brasileiro 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), “a hierarquia proposta pelas autoras [Abaurre; Galves, 1998], para a análise do ritmo no português brasileiro e no português europeu, é igualmente adequada à análise rítmica das segmentações não convencionais” (p.174).

Em seu estudo, Ferreira (2011) trata de processos de segmentação não convencional no contexto da Educação de Jovens e Adultos (EJA). A autora analisa a escrita de três sujeitos que, na época da coleta dos dados, tinham 44, 22 e 16 anos de idade e cursavam a segunda etapa do EJA.<sup>17</sup> As grafias de hipossegmentação, hipersegmentação e híbridos são analisadas tendo por base as variáveis linguísticas, definidas por Cunha (2004), para a análise dos mesmos tipos de dados em produções escritas de crianças. Essa decisão metodológica foi tomada a fim de observar, em um primeiro mo-

16 Cunha (2010) 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” (p.144).

17 A segunda etapa do EJA corresponde ao 3º ano do Fundamental (antiga 2ª série).

mento, a relação entre os dados produzidos por adultos e crianças em processo de alfabetização e, em um segundo momento, para verificar a adequação das categorias formuladas por Cunha na análise de textos de adultos sendo alfabetizados.

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

Na avaliação da precisão da categoria “tipos de palavras” (Cunha, 2004) na análise dos dados de segmentação não convencional de EJA, Ferreira chegou à conclusão de que essa categoria é totalmente pertinente para a compreensão dos processos de segmentação que ocorrem na escrita de adultos em processo de alfabetização. Em termos quantitativos, os dados de Ferreira (2011) mostram, com relação 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 (p.141). Também em função dos resultados quantitativos, no caso das hipersegmentações, 48% do total foi resultado 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, com relação às ocorrências híbridas, a autora observa que tanto na escrita inicial de adultos quanto na de crianças, esses tipos de dados são pouco frequentes.

## Abordagens linguísticas da noção de palavra

Como se sabe, a ortografia do português brasileiro faz uso de noções morfossintáticas/semânticas para definir a convenção de palavra escrita. Porém, com base no que foi discutido anteriormente, observamos que não são apenas os critérios convencionais que os escreventes levam em conta ao defrontarem com a tarefa de registrar (orto)graficamente os limites das palavras. Nesse sentido, defendemos a relevância de noções fonológicas (na medida em que as palavras apresentam uma organização fônica que, pela observação de nossos dados, mostra-se relevante para interpretar as grafias não convencionais) e de noções letradas (que rondam o imaginário 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, apresentamos a seguir 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 registrados, em alguma medida (lembramos que outros aspectos, como o morfossemântico e o letrado, estão perpassados na análise dos dados).

Um dos problemas a enfrentar na caracterização de palavra diz respeito aos critérios fonológicos e morfológicos mobilizados para a sua definição. Bisol propõe uma distinção entre palavra morfológica e fonológica:

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 clíticos. (Bisol, 2004, p.59; destaques nossos)

A partir das definições da autora, abordamos a diferença entre as duas noções de palavras, observando a configuração de palavras compostas. Tomemos 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 palavra morfológica e fonológica: morfológicamente, as duas bases – [mesa]β e [redonda]β – funcionam juntas na sentença, como uma só unidade lexical, portanto, com um só significado; fonologicamente, no entanto, a palavra em análise compõe-se de duas unidades independentes – [mesa]ω e [redonda]ω –, cada uma portadora de um acento primário, presente nas sílabas destacadas: **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, p.61; destaques nossos).

Contudo, há casos em que os limites da palavra fonológica e da morfológica coincidem, como ocorre em: “comprei uma mesa redonda para minha casa”. Ao contrário da primeira frase, nessa ambos os constituintes linguísticos são isomórficos, visto que o modo como “mesa” e “redonda” estão empregadas sintaticamente faz que essas palavras tenham, em termos morfológicos, significados distintos, ou seja, são duas unidades lexicais independentes. Também, nos critérios fonológicos, as duas palavras têm dois acentos independentes. Existem, portanto, duas palavras morfológicas e duas fonológicas, nesse caso.

Um dos preceitos que rege a organização dos constituintes prosódicos<sup>18</sup> é o uso de informações não específicas, isto é, não fonológicas, na construção de seus domínios. A palavra fonológica, em

---

18 Essa afirmação está respaldada no modelo de fonologia prosódica de Nespor e Vogel (1986).

especial, é o primeiro constituinte prosódico a estabelecer interação com outro componente que não o fonológico, no caso, o morfológico (Nespor; Vogel, 1986, p.109). Mas, como exemplificamos, essa interação não é sempre de natureza isomórfica. Desse modo, os limites da palavra fonológica não precisam coincidir sempre com os limites do constituinte morfológico correspondente, embora, na maioria das vezes, isso aconteça.

O funcionamento de determinadas palavras derivadas é outro aspecto 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 palavra 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), o que sinaliza a mudança de acento para a sílaba seguinte (beleza).<sup>19</sup> É fundamental observar que o deslocamento do acento presente na 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”. A regra geral de distribuição do acento no português brasileiro determina 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 receber o acento primário:

---

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

“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, 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 português brasileiro. Nesse aspecto, a manutenção da vogal média-baixa na base da palavra é evidência da presença de mais de um acento (cf. Quadros; Schwindt, 2008).

Então, no que diz respeito ao estatuto fonológico, a palavra “*belamente*” é uma palavra composta, de modo que não apresenta coincidência com sua contraparte do componente morfológico, visto que, para a morfologia, trata-se 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 no caso 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 ocorre com os prefixos acentuados, conforme análise de Schwindt (2001). Em seu estudo, o autor procurou, dentre outros objetivos, categorizar os prefixos do português brasileiro quanto ao seu comportamento fonológico. Para tanto, submeteu-os a comparação a partir de três critérios: 1) presença/

não presença de acento; 2) oposição forma livre/forma presa; 3) tipos de processos fonológicos a que estiveram propensos.

Ao fim da comparação, Schwindt chegou a duas classes de prefixos. Uma delas é a dos prefixos composicionais (PCs),<sup>20</sup> formada por prefixos que 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>21</sup> em determinados contextos, conseguem instanciar-se como formas livres (João reencontrou sua **ex** – ex-mulher); 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 outra classe é a dos prefixos legítimos (PLs),<sup>22</sup> os quais, diferentemente daqueles do primeiro grupo, demonstram comportamento de sílaba átona interior à palavra (**inesquecível**); não se sustentam sozinhos na sentença, igualando-se às formas presas (João desfez as malas, antes que Maria \*des as dela); e, por fim, são atingidos apenas por processos fonológicos que ocorrem no interior de palavras, como a neutralização da pretônica (n[e]ologismo ~ n[e]ologismo), a harmonização vocálica (r[e]fiz ~ r[i]fiz) e a assimilação da nasal (aN+alfabeto → a[n]alfabeto).

Para evidenciar a importância de tais considerações para o estudo que originou este livro, antecipamos dados de hipersegmentação do *corpus* que apresentam ancoragem nessa característica morfofonológica referente à formação de palavras. Na Figura 4, é apresentado um exemplo de hipersegmentação a partir de uma palavra formada por prefixo dissílabo e, na Figura 5, um exemplo de hipersegmentação motivada por um sufixo dissílabo acentuado.

20 A denominação “composicional” é dada porque, em virtude de a palavra fonológica não poder possuir mais de um acento, os prefixos acentuados ganham contornos semelhantes aos dos elementos de um composto.

21 Todos os exemplos citados foram extraídos do próprio autor.

22 Legítimos, pois, pela não acentuação, caracterizam-se como sílabas átonas que são afixadas à esquerda de uma base.



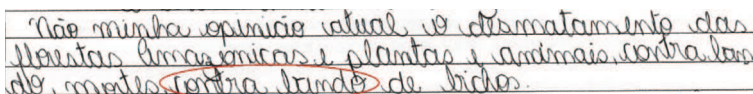


Figura 4 – Fonte: Banco de dados de escrita do EF II: Z11\_8D\_31F\_05.

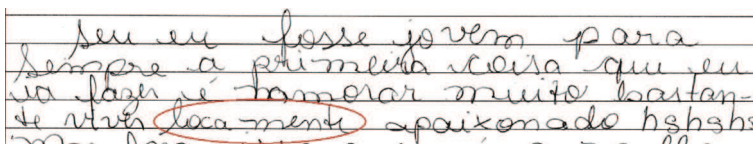


Figura 5 – Fonte: Banco de dados de escrita do EF II: Z10\_7A\_31M\_01.

Outro problema a enfrentar na identificação de palavra diz respeito à caracterização de elementos que não têm acento lexical e, portanto, não constituem palavra fonológica, mas são dependentes dela, configurando o conjunto dos clíticos fonológicos.

Os clíticos fonológicos, de acordo com Bisol (2005, p.164), apresentam as propriedades universais de atonicidade, comportamento de formas dependentes e não pertencimento a uma classe morfológica específica. A primeira propriedade os diferencia das 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>23</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 palavras funcionais (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**.<sup>24</sup>

23 Essas três possibilidades de união do clítico ao seu hospedeiro remetem a formas de integração prosódica distintas, como mostraremos adiante.

24 Há palavras funcionais dissílabas que, em funcionamento sintático, perdem informações segmentais e tornam-se clíticas (para > pra > pa, por exemplo). Nesses casos, a atonicidade não é estável, tal como ocorre com os clíticos monossílabos.

Determinar o funcionamento prosódico dos clíticos, apesar de suas propriedades universais, não é tarefa simples, pois, dependendo da língua em discussão, são notadas instabilidades em seu comportamento. Segundo Nespor e Vogel (1986, p.145), os clíticos podem apresentar: comportamento independente, comportamento de sílaba e comportamento específico. Essa natureza híbrida tem gerado dissensões teóricas quanto ao seu nível de inserção na estrutura fonológica. Assim, abarcando os tipos de comportamento mencionados, os clíticos podem ser integrados nos domínios da frase fonológica, da palavra fonológica e do grupo clítico.<sup>25</sup>

De acordo com essas autoras,<sup>26</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) + palavra. Os trabalhos de Bisol (2000, 2005) seguem esse mesmo posicionamento teórico e defendem, para o português brasileiro, a prosodização do clítico com a palavra adjacente no domínio do grupo clítico.

Antes de apresentarmos os argumentos usados por Bisol (2000, 2005), gostaríamos de observar que, embora essa estudiosa, por um lado, assuma a existência de um domínio prosódico “que se compõe de uma palavra fonológica e de um ou mais clíticos” (p.24), junto com Nespor e Vogel (1986), por outro lado ela não avalia como pertinente o modo como isso é feito na análise dessas autoras: como uma palavra fonológica.<sup>27</sup> Bisol (2005) respalda o seu posicio-

---

25 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.

26 Maiores detalhes sobre o modelo formulado por essas autoras serão apresentados no Capítulo 3.

27 Essa medida das autoras é para atender a *Strict Layer Hypothesis*, a qual exige que “uma unidade de um dado nível da hierarquia está exaustivamente 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; Vogel, 1986, p.7; tradução nossa).

namento contrário nas três propriedades da palavra fonológica: ser portadora de relações de proeminência; ser domínio de aplicação de regras lexicais;<sup>28</sup> 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 (presente em um constituinte maior, a proeminência recai sobre a palavra à qual o clítico está relacionado). Em segundo lugar, seqüências de clítico + hospedeiro mostram-se insensíveis a determinados processos que se aplicam no domínio da palavra fonológica, a exemplo da harmonia vocálica, como mostrou Brisolara (2008), para o português brasileiro. Em último lugar, no português brasileiro há a restrição fonotática de palavras que não são iniciadas por consoantes palatais; porém, na língua, há o clítico “lhe”.

O não atendimento a essas propriedades indica que os clíticos, no português brasileiro, não podem receber o estatuto de palavra fonológica. Nesse sentido, Bisol (2000, 2005) considera que eles 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” (id., 2005, p.171).

Segundo Bisol (2000, 2005), há evidências de que, no português brasileiro, os clíticos são adjungidos “a uma palavra fonológica pronta, sem integrá-la, emergindo daí o primeiro constituinte pós-lexical” (id., 2000, p.19). Com efeito, a defesa de seu posicionamento se faz em torno de três argumentos principais. O primeiro diz respeito ao fato de os clíticos serem insensíveis à restrição das três janelas, a qual impede que o acento de uma palavra recaia na quarta sílaba ou em outras posteriores (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” (id., 2000, p.24). O segundo argumento diz respeito à característica de mobilidade posicional dos

---

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

clíticos (em particular os pronominais), que é típica de constituintes frasais: **Te** dei um presente / Dei-**te** um presente. Por fim, o terceiro argumento refere-se ao fato 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]. De acordo com 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 se aplica aos clíticos é a nasalização, sobretudo “a nasalização vocálica, que não precisa de informação morfológica e que não tem exceções” (id., 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~ijj. A sonorização da fricativa, outra regra elencada pela autora, ocorre dentro de palavras: pa[z]mo e, fora desse domínio, casa[z] bonitas, afetando também os clíticos: o[z] meninos (id., 2005, p.170). Outra regra importante diz respeito à palatalização de /t/ e /d/ diante de /i/: ([ˈtʃimi]) em palavras; ([ˈtʃi vi]) em clíticos.

Mas a autora acredita que as regras de sândi vocálico constituem a principal evidência do grupo clítico. Ao observar o comportamento dos clíticos no processo de elisão, Bisol (2000, p.26-7) observa, em relação à elisão da vogal /a/, que esse processo não se aplica no interior de uma palavra fonológica (maometano \*mometano), porém apresenta-se 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/, a autora (2005) conclui que essa regra só se aplica na presença de um clítico; a exceção ocorre quando diante do clítico há um elemento terminal designado (DTE), “elemento de proeminência relativa que projeta na grade uma posição forte, pronta para carregar o acento preponderante” (id., 2005, p.174), como ocorre em: de amor \*damor. A elisão de /e/ também não ocorre em frases fonológicas (cidade antiga \*cida-dantiga), 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, com referência ao processo de elisão, a autora considera que o fato de

não se aplicar dentro de uma palavra é um indicio de que o *clítico* + *hospedeiro* não 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), considera que o clítico é anexado diretamente a um nó de frase fonológica, uma vez que, segundo a pesquisadora, não há processos, no português brasileiro, que confirmem a existência do grupo clítico. Na construção de sua argumentação, com base em Selkirk (2004, apud Simioni, 2008, p.440), a autora apresenta e discute quatro tipos de possibilidades de inserção prosódica dos clíticos: **a)** (cl (lex)ω)φ; **b)** ((cl (lex)ω)ω)φ; **c)** ((cl lex)ω)φ; **d)** ((cl)ω (lex)ω)φ.

A justificativa apresentada pela autora é a possibilidade de integração representada em **a**, que é a do clítico anexado a uma frase fonológica; em consequência, as demais possibilidades são rejeitadas para o português brasileiro, 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 englobam determinados prefixos legítimos do português brasileiro (cf. Schwindt, 2001), a autora considera que apresentam, por exemplo, processos de ressilabação que não ocorrem na combinação clítico + palavra. Como exemplo, 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)ω)ω > (a[na]eróbico)ω)ω; isso não se aplica aos clíticos: (sem(esperança)ω)φ \*(se([ne]sperança)ω)φ (cf. Simioni, 2008, p.438). A partir dessas comparações, a autora considera que “essa diferença é explicada pelo fato de que prefixos se adjungem a uma palavra prosódica, enquanto clíticos se anexam a uma frase fonológica” (ibid.). A configuração em **c** representa a prosodização do clítico como integrado à palavra fonológica, e, nesse sentido, ele 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 eles sofrem processos característicos de sílabas átonas finais de palavra, como é o caso da neutralização da átona final, processo também discutido 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][i]fante, com exceção dos 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/, /o/) é seguida de um enclítico, ela tem a sua vogal final alçada: (ped[e])ω-se ~ (ped[i])ω-se.

Para a autora, esses processos que acontecem em contexto final indiciam a existência de uma fronteira entre o clítico e o hospedeiro e, 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 português brasileiro 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 vale-se de um arcabouço teórico baseado em restrições, sobretudo aquelas de dominância prosódica (*layeredness*;

*headedness*; exaustividade; não recursividade)<sup>29</sup> e de alinhamento de fronteiras.<sup>30</sup> Segundo Simioni, a configuração pela qual o clítico se anexa diretamente a uma frase fonológica é a estrutura que, além de se mostrar a mais coerente para o português brasileiro, a partir da observação de processos fonológicos, é também a possibilidade que viola menor número de restrições, ou seja, não atende apenas a exaustividade (em virtude do fato 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 a autora.

Por fim, conclui que o grupo clítico não possui relevância para explicar fatos fonológicos do português brasileiro, 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 precisam de um constituinte prosódico específico para isso, contrariando a análise da autora para essa mesma variedade.

Com base em diferentes abordagens teóricas, Bisol (2000, 2005) e Simioni (2008) fazem propostas distintas de representação da estrutura prosódica do clítico em português brasileiro. Bisol considera que existem fatos dessa língua que só se explicam na sequência

---

29 *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.466-7, apud Simioni, 2008, p.439).

30 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.468-9, apud Simioni, 2008, p.441).

exata de um clítico mais o seu hospedeiro, daí a necessidade de considerar um domínio como o grupo clítico. 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 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, parece encaminhar-se para uma mesma direção: a de que, no português brasileiro, os clíticos tendem a ser mais independentes com relação à palavra hospedeira, pois estão sempre ganhando *status* prosódico nos domínios que já se encontram em interação com o componente sintático.<sup>31</sup>

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 português brasileiro. Retomaremos esse tema ao realizar a análise

---

31 Ressalte-se que essas posturas notadas para o português brasileiro são diferentes de outras propostas teóricas correntes na literatura sobre integração prosódica de clíticos, a exemplo da que foi desenvolvida por Vigário (2003) sobre os clíticos do português europeu. Afirmando haver assimetria na inserção prosódica de proclíticos e enclíticos, Vigário argumenta que os primeiros unem-se ao hospedeiro por adjunção (neste caso, seriam duas unidades); já os últimos são incorporados à palavra fonológica precedente (neste caso, clítico e hospedeiro seriam uma só unidade). Para chegar a tal avaliação, a autora observou que regras fonológicas pós-lexicais que se aplicam à borda direita final das palavras também se aplicam à sequência de palavra + enclítico, indicando que o clítico, nessa posição, apresenta comportamento idêntico ao de segmento interno de palavra. Já a presença de um proclítico constitui contexto favorável para a ocorrência de regras pós-lexicais na margem esquerda inicial da palavra, sinalizando, segundo a autora, que entre o proclítico e o hospedeiro há a presença de uma fronteira prosódica e, 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 os seguem também no nível pós-lexical.



dos dados, no Capítulo 4. Por ora, apresentamos apenas dois registros escritos que estariam, de acordo com a interpretação que faremos mais à frente, em diálogo com essa discussão, mostrados nas figuras 6 e 7.

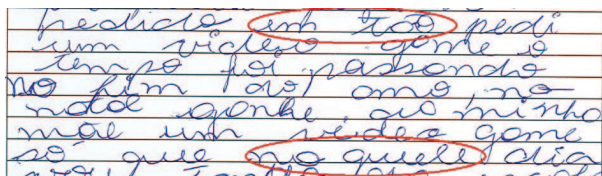


Figura 6 – Fonte: Banco de dados de escrita do EF II: Z08\_5C\_12M\_03.

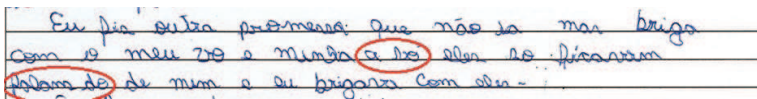


Figura 7 – Fonte: Banco de dados de escrita do EF II: Z09\_6B\_04M\_01.

## Uma síntese

Neste capítulo, procuramos delinear as reflexões que compreendem a primeira parte dos subsídios teóricos deste livro. Inicialmente, lançamos mão, para abordar a noção de palavra, de trabalhos sobre as segmentações não convencionais. Os trabalhos consultados 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 para crianças e adultos que iniciam o processo de escrita, mas também para escreventes com mais anos de escolarização. Essas constatações auxiliam a reforçar nossa hipótese de que a palavra escrita constitui uma noção complexa, a qual põe em jogo a decisão entre diferentes informações linguísticas.

Com relação ao estatuto dos registros das fronteiras de palavras, é importante destacar que, em todos os estudos, identificamos um posicionamento comum entre os estudiosos, que os consideraram

como ocorrências representativas do funcionamento da linguagem e da sua relação com os seus sujeitos. Assim, os dados de hipossegmentação e hipersegmentação deixam de ter relevância apenas para interesses pedagógicos de ensino e aprendizagem (como 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 adotamos, para o trabalho que originou este livro, a que definimos como tendência fonológico-letrada, por referir-se 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 perfilar tais reflexões sobre a noção de palavra fonológica com as hipersegmentações justifica-se pelo fato de que as dificuldades em definir os limites de palavra, quando considerados os enunciados falados, refletem-se, 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” (contrabando) e de elementos clíticos em grafias como “na quele” (naquele).

No próximo capítulo, abordaremos a concepção de escrita proposta por Corrêa (1997, 2004).



## 2

# UM OLHAR PARA A ESCRITA

Neste capítulo, explicitamos a forma como as hipersegmentações são observadas em relação ao seu funcionamento no interior do texto escrito. Compreendemos 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) e defendemos a importância de as noções de fala e escrita serem claramente definidas, assim como a natureza da relação entre ambas. A justificativa para a ênfase nessas noções está baseada no fato de que, em especial, certa concepção da relação fala/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.

Seguem essa linha argumentativa 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. Ao abordarmos essas diferenças, iniciamos discutindo, 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 classificaria como diferentes erros ortográficos (incluin-

do, nessa classificação, as grafias de segmentação não convencional). Os autores, ao conceberem a relação entre fala/escrita por meio da perspectiva da interferência, assumem, em consequência, que a fala e a escrita constituem modalidades linguísticas homogêneas, no sentido de que as características que constituiriam cada uma delas seriam, por princípio, puras e opostas à da outra modalidade.

No trabalho que originou este livro, adotamos, para a análise das hipersegmentações, outra visão teórica sobre a relação fala/escrita, ancorada na tese do modo heterogêneo de constituição da escrita, formulada por Corrêa (1997, 2004). Ao discutir e comparar a concepção de Corrêa com a proposta de Zorzi, assumimos posição contrária 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 neles já seriam evidências de patologias da aprendizagem. A partir da contraposição do trabalho de Corrêa à abordagem em que se insere Koch, lançamos luz sobre posicionamentos distintos no interior da Linguística concernentes à relação fala/escrita. Nesse aspecto, no trabalho dessa autora, identificamos um tipo análise que, ao perpassar o nosso objeto de estudo, distancia-se fortemente daquela proposta desenvolvida na mesma área do conhecimento e com a qual nos filiamos.

Ainda, esclarecemos que não nos aproximamos 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 investigação da relação fala/escrita, como também não adotam, em termos teóricos, nenhuma filiação com discussões aprofundadas sobre o tema que tenham como ponto de partida a definição dos conceitos de fala e escrita. Assim, por negligenciarem aspectos dessa natureza, consideramos que as análises de Zorzi e Koch deixam de contemplar a complexidade que identificamos nas grafias de hipersegmentação.

Pelas razões explicitadas, nosso trabalho fundamentou-se na proposta teórica desenvolvida por Corrêa (1997, 2004), a qual se sustenta na ideia de não oposição entre fatos linguísticos e práticas

sociais. 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 (neste caso, 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 ilustrarem tipos de análise muito frequentes acerca de objetos de investigação como o do estudo que originou este livro (sobretudo em diferentes correntes da Linguística nas quais ele se insere, 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, pelo ganho em adotar uma abordagem teórica dedicada ao tratamento exclusivo das questões que tocam a relação fala/escrita como campo específico de investigação. A partir dessa perspectiva, privilegiamos a observação de processos/hipóteses que mobilizam as grafias não convencionais e deixam marcas da “trajetória desses escreventes em sua inserção na escrita convencional” (Capristano, 2010, p.175).

## **Perspectivas de análise da relação entre fala e 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 escolar 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” (ibid., p.21). Desse modo, “a escolha destas escolas teve a pretensão de evitar que problemas relativos ao desenvolvimento de uma nova linguagem pu-

dessem ser atribuídos exclusivamente a fatores sociais, econômicos ou pedagógicos” (p.21).

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

Os erros identificados são organizados e classificados pelo autor em dez categorias (cf. Zorzi, 1997, p.25-34):<sup>1</sup> 1) alterações ou erros devido à possibilidade de representações múltiplas – casos em que a ortografia oferece a 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: *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*”, *quatrocentos* – “*quatro sentos*”; 5) alterações decorrentes de confusão entre as terminações *am* e *ão* – casos em que a ortografia prevê *am*, mas a grafia foi *ão* e vice-versa, por exemplo: *falaram* – “*falarão*”; 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* –

---

1 Exemplos retirados do próprio autor.

“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”; 10) inversão de letras – casos que a letra encontra-se em posição invertida dentro da palavra, por exemplo: fraquinho – “farquinho”.

A partir dessas categorias, Zorzi (1997) 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 em separado, eles convergem em relação a sua motivação, pois, “nestes três casos, há uma *forte influência de padrões de linguagem oral* determinando a ocorrências das alterações” (p.87; destaque nosso).

Para o autor, esse resultado pode ser bastante compreensível, pois “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” (p.10). Assim, na medida em que fala e escrita constituem duas línguas diferentes, a criança, por dominar apenas as características da língua falada, tenta, nos 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” (p.88).<sup>2</sup>

De acordo com o autor, 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,

---

2 Vale lembrar a postura de Abaurre (1988a), já mencionada neste trabalho, para quem a afirmação de que a criança baseia-se 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.136-7).



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*. (Zorzi, 1997, p.10; destaques nossos)

Observamos que a relação oralidade/escrita, na qual se fundamenta o autor, está baseada em uma concepção de interferência mútua, que se constrói de uma para outra (isto é, da fala para a escrita e 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 apreendidas 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 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 para a escrita inicial – menos calcada nos preceitos convencionais da escrita padrão – está reservado o lugar de manifestação de características da fala, uma vez que, ao atender os critérios que regem a convenção ortográfica, os textos escritos estariam “blindados” contra qualquer interferência da fala. A partir dessas constatações, a escolarização desempenharia papel essencial na transição de interferência da fala na escrita para interferência da escrita na fala, pois, como salienta Capristano (2010), 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” (p.178).

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 ela irá basear suas primeiras produções escritas e demais textos por muitos anos. De posse do que considera como as diferenças do texto falado em relação ao texto escrito, a autora elenca os principais pontos de interferência encontrados nos textos infantis, dentre os quais se identificam: 1) mecanismos de referência – correspondem à utilização indevida de recursos anafóricos e catafóricos; 2) repetições – caso comum no texto falado, “podendo mesmo ser considerada um dos mecanismos organizadores dessa modalidade oral” (ibid., 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 o uso de nenhum 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 como se realizam foneticamente; 8) correções comuns nos textos falados – casos em 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 resultado da apreensão das crianças de vocábulos fonológicos (ver Câmara Jr., 1970). Porém, nas hipersegmentações, como “na quela” (naquela), as crianças buscam efetuar uma segmentação convencional, por isso é que elas estariam, segundo Koch (1997), “caindo no extremo oposto, isto é, ‘picando’ demais a palavra” (p.37).

Percebe-se que a explicação da autora não sustenta a sua própria interpretação, calcada no argumento da interferência do texto falado no texto escrito, pelo fato de ser o único modelo de texto que a criança conhece ao ingressar na escola, visto que, como ela própria afirma, em determinados momentos as crianças estão tentando realizar uma segmentação adequada aos padrões ortográficos. Como o uso de recursos gráficos para dividir os enunciados em palavras é exigência exclusiva da escrita, é bastante compreensível que, quando os aprendizes arriscam segmentar as palavras de modo

convencional, eles não estão lidando apenas com informações que organizariam o texto falado. Como mostram outros pesquisadores do tema, as crianças estão inseridas também em práticas letradas, por isso mobilizam conjuntamente informações dessas práticas para compor os textos que escrevem.

Diante do 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 identifica o que esse estudioso definiu como modelo de letramento autônomo, em 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.47-8)

Assim, no modelo autônomo, a escrita é concebida 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. A partir 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 circula a escrita. Além do mais, pela apropriação da escrita tem-se a abstração do pensamento e o desenvolvimento cognitivo, os quais, em consequência, trazem o progresso e a mobilidade social dos indivíduos.<sup>3</sup> Nessa abordagem,

---

3 A tese da relação entre aquisição da escrita e desenvolvimento cognitivo ganhou força sobretudo em investigações de orientação etnográfica, as quais

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 entre fala e escrita, à qual ele atribui um caráter rigidamente formal, capaz apenas de alçar “bons resultados na descrição estritamente empírica, [enquanto] manifesta enorme insensibilidade para os fenômenos dialógicos e discursivos” (p.28), isto é, não se sustenta na análise de fatos linguísticos concretos.

Buscando avançar em relação a essa abordagem, o autor considera que fala e escrita são modalidades linguísticas que se realizam em práticas de linguagem (orais e letradas), por meio de gêneros textuais. Por compreender a existência de práticas sociais na realização dos usos da linguagem, ele chama a atenção para o fato de que as relações entre modalidades linguísticas e práticas sociais não são iguais e precisam ser distinguidas. Assim, segundo o autor, 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). Adiante apresentaremos o questionamento a essa 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) 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 polos opostos” (p.37). A sua posição é elucidada na Figura 8.

---

buscaram comparar as estratégias de resolução de problemas por parte de grupos utilitários da escrita e grupos ágrafos (cf. Goody, 1988/[1977]). Argumenta-se, ao realizar tal comparação, que os grupos que conhecem a escrita têm mais habilidade para realizar as tarefas propostas. Street (1984) questiona esse argumento, declarando ser impossível mensurar 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.

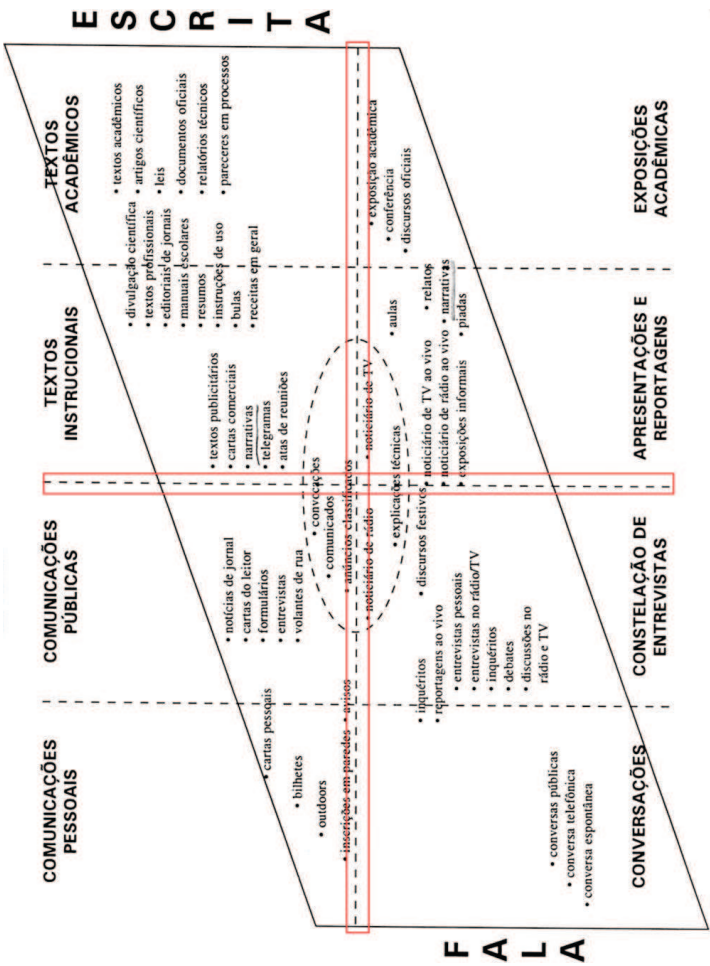


Figura 8 – Fonte: Marcuschi (2001, p.41, modificada)

O esquema apresentado na figura explicita o *continuum* de gêneros textuais, alicerçado em duas diferenciações, representadas pelas linhas pontilhadas centrais (marcadas em vermelho) que o separam. A linha horizontal indica a distinção entre meio de produção (sonoro – gráfico), e o traço vertical, a separação pela concepção discursiva (oral – escrita). Com base nesse princípio organizador, fala e escrita aparecem nos extremos do *continuum* e os gêneros textuais são distribuídos de modo gradativo, o que permite a observação de gêneros falados e escritos mais homogêneos, bem como a identificação de gêneros mistos dessas modalidades (centrados em um círculo no meio do *continuum*).

Diante dessa disposição, que ilustra as diferenças e semelhanças da relação entre fala e escrita, Marcuschi (2001) 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” (p.42).

Apesar de reconhecermos o ganho explicativo da perspectiva do autor, quando comparada à tese da separação radical entre fatos de fala e fatos de escrita, não deixamos de concordar com o pensamento de Corrêa (1997, 2001, 2004) de que a proposta do *continuum* de gêneros textuais recupera a separação entre fala e escrita, ao pressupor certa homogeneidade nos seus extremos. Desse modo, consideramos que o tipo de relação entre fala e escrita, e também entre oralidade e letramento (a qual se revela por meio do registro escrito), é de ordem distinta das apresentadas até então. A seguir expomos a proposta de Corrêa (1997, 2004), a qual nos ajuda a entender as hipersegmentações como dado linguístico que possibilita tematizar sobre a relação falado/escrito.

## A escrita como modo de enunciação

Dentre as várias possibilidades teóricas de 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>4</sup> firma-se como contraposição ao ideal de oposição radical entre fala e escrita e também ao de dicotomização como recurso metodológico.<sup>5</sup> Para Corrêa, a presença de marcas da enunciação falada na enunciação escrita não se explica nem como uma questão de interferência direta, nem pela ocorrência de gêneros textuais mistos. Mais do que isso, a presença do falado no escrito é evidência de que a escrita (e a linguagem de modo geral) não se caracteriza pela homogeneidade, e sim por uma heterogeneidade que lhe é constitutiva.

O cerne dessa teorização, portanto, é que a heterogeneidade é própria da escrita (e não está 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; destaque nosso). Desse modo, toda e qualquer produção de escrita é, por excelência, heterogênea. Ao assumir essa perspectiva, Corrêa (1997) distancia-se 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 comum com que se tomam as produções escritas consideradas como menos integradas a um padrão tido como legítimo” (p.86).

Por termos buscado, no tratamento das hipersegmentações, distanciamento de algumas interpretações que as classificam tão

---

4 A proposta do modo heterogêneo de constituição da escrita foi resultado da tese de doutoramento do autor (ver Corrêa, 1997). Em trabalho posterior (id., 2004) essas ideias foram retomadas e organizadas em forma de livro.

5 A proposta teórica do contínuo de gêneros textuais foi apresentada por Corrêa (1997, 2004) como perspectiva de dicotomização metodológica.

somente como erros ortográficos<sup>6</sup> (por exemplo, Koch, 1997) ou como indícios de algum problema de aprendizagem (por exemplo, Zorzi, 1997), assumimos a premissa da heterogeneidade da escrita. A partir desse lugar teórico, identificamos a oportunidade de levantar outras hipóteses explicativas, em especial aquelas 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 que os fatos da língua (falada/escrita) só se realizam por constituírem modos de enunciação concretizados por meio de práticas de linguagem (oral/letrada). Dessa forma, o autor entende que há uma indissociabilidade entre ambos (modos enunciativos e práticas sociais) e, por essa razão, apresenta uma notação conjunta – oral/falado e letrado/escrito – ao tratar os usos da linguagem.<sup>7</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 com o já falado/escrito

---

6 Como conceituamos ao longo de todo o trabalho, as hipersegmentações são vistas aqui como fatos não convencionais de escrita, e não como erros, seja na aceção de construção de hipóteses (conforme Cunha, 2010), seja como desvios de escrita (conforme Zorzi, 1997). Entendemos as hipersegmentações como possibilidades gráficas no interior da língua e, neste aspecto, elas se distanciam apenas em relação ao que está convencionalizado para a escrita.

7 Veja-se que, nesse ponto, Corrêa (1997, 2004) diferencia-se de Marcuschi (2001), para quem fala e escrita são modalidades da língua pensadas paralelamente às práticas sociais de oralidade e letramento. Em estudo mais recente, Corrêa (2009) também fez uma crítica à ideia que subjaz ao termo “modalidade linguística”. Segundo o autor, “modalidade” relaciona-se a um juízo de código abstrato, o que significa, em outros termos, 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” (p.2).



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. Seu ponto de partida, então, é a concepção de que toda produção de linguagem é sempre resultado de um já dito/ouvido e escrito/lido.

Para compreender a não distintividade entre as práticas sociais orais/faladas e letradas/escritas, Corrêa 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 registro de sociedades de oralidade primária. Conhecer um fato, como afirma o autor, independe do contato empírico com uma materialidade física, figurada em um tipo de escrita como a alfabética, mas pode se dar por modalidades 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; destaque nosso). Assim, as práticas do relato configu-

ram-se 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” (ibid., p.139).

No desenvolvimento do trabalho que originou este livro, participamos de discussões realizadas no âmbito do curso de extensão “Introdução aos estudos de letramentos”, ministrado entre os meses de abril e 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).

Neste livro, adotamos o conceito de letramento amplo definido por Corrêa (2004), pois compreendemos que essa concepção demonstra a importância de 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 de que consideram apenas o tipo de escrita alfabética e as convenções ortográficas que a regem. Mas é importante ressaltar que o material e o dado investigados, bem como a perspectiva teórica adotada, nos levam a identificar aspectos da inserção do escrevente em práticas letradas/escritas e acabam por encaminhar-se na direção do imaginário em relação a um tipo de escrita esperado da escola, como lugar institucionalizado que valoriza sobremaneira textos escritos em pleno acordo com as convenções ortográficas. Nossos apontamentos, pois, sobre os tipos de informações letradas/escritas que parecem circundar o aparecimento de hipersegmentações, não excluem o fato de que compreendemos o letramento como uma prática social que não se reduz a um só tipo de informação e a um só tipo de escrita.

Note-se, ainda, outra questão fundamental em relação à caracterização da teorização do modo heterogêneo de constituição da escrita: 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, ao enunciar, circula dialogicamente por elas, deixando pistas linguísticas do que julga ser a constituição da sua escrita, de si mesmo como escrevente e de seus interlocutores.

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 verificação desse fato levou Corrêa a postular o que denominou de 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 o autor se valeu para propor esses lugares metodológicos, ele lembra que esses “processos de constituição da escrita podem ser retomados em qualquer época, na escrita de qualquer pessoa, em qualquer texto” (id., 2004, p.89).

Os eixos de circulação dialógica são os que seguem:

- Representação da (suposta) gênese da escrita.

Com base em Corrêa, esse eixo diz respeito à atribuição de um lugar específico ao 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, são necessárias duas ressalvas de ordem teórica a respeito desse eixo. A primeira está relacionada à negação do termo “gênese” como sinônimo de ato fundador. Se-

gundo Corrêa, o uso desse termo está 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” (ibid., 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 do escrevente em relação ao 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*” (ibid., p.89; destaque nosso).

A busca por marcas da gênese da escrita, portanto, não tem ligação, no trabalho de Corrêa (nem 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 não separação 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.

- 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.10-1)

Para Corrêa (2004), as marcas produzidas sob a imagem desse eixo reproduzem “a dinâmica social de institucionalização de valores para as diversas formas linguísticas” (p.165). No material de análise de que ele se vale – e também no material deste trabalho –, tal dinâmica aparece marcada por uma relação tríade, a qual vincula a língua com a escrita, e esta com a escola. Nesse sentido, o escrevente guia a sua escrita pelo que acredita 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*” (ibid., p.166; destaque nosso).

Pelas 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” (ibid., p.166) e, nesse sentido, a forma como o escrevente representa a si mesmo na escrita, por meio da circulação dialógica nesse eixo, realça: o seu lugar de reprodutor de uma prática de linguagem já edificada; a sua busca por atender às expectativas da instituição escolar em relação à escrita culta; a atribuição a si próprio de um lugar único enquanto escrevente, no interior de práticas letradas/escritas institucionalmente reconhecidas e em função da projeção de um interlocutor.

- Dialogia com o já falado/escrito e ouvido/lido.

Com esse último eixo, a intenção do autor é mostrar que todo e qualquer texto é sempre produto de um já dito. Mais do que pistas

intertextuais que poderiam marcar (de modo mais direto) a relação do escrevente com o já falado/ouvido e o escrito/lido, o autor 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 (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 lado, que a apropriação de determinado discurso deu-se de um modo que não permite mais recuperar explicitamente. Assim, mesmo quando o escrevente tenta 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 dão indícios da sua circulação por outras práticas de linguagem e que o constituem como escrevente.

## Uma síntese

Iniciamos as discussões tecendo observações sobre aquilo que consideramos diferentes formas teóricas de entendimento da relação fala/escrita, com as quais não dialogamos. Examinamos, com base nesse propósito, trabalhos que permitem entrever os tipos de entendimento e a consequência deles no tocante à análise de marcas linguísticas, como as de hipersegmentação. 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 discussão, lançamos luz sobre a concepção de escrita adotada para a compreensão das produções textuais de nosso *corpus*, das quais extraímos as marcas de hipersegmentação. Emprestamos dos trabalhos de Corrêa (1997, 2004) a noção do modo

heterogêneo de constituição da escrita e, a partir dela, entendemos a escrita como 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, entendemos 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 das suas inserções em práticas orais/faladas e letradas/escritas de uso da linguagem.

# 3

## PROCEDIMENTOS METODOLÓGICOS E CONSTITUIÇÃO DO *CORPUS*

O propósito deste capítulo é explicitar 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. A apresentação dessas escolhas será feita em três seções.

### Características do material analisado

O material de investigação é composto por produções escritas realizadas em ambiente escolar por alunos dos quatro últimos anos do Fundamental (6º, 7º, 8º e 9º anos). Os textos pertencem ao Banco de Dados de Escrita do Ensino Fundamental II,<sup>1</sup> o qual é resultado do projeto de extensão universitária Desenvolvimento de Oficinas de Leitura, Interpretação e Produção Textual. O projeto contou com a coordenação das professoras doutoras Luciani Tenani e Sanderléia Longhin-Thomazi, do Ibilce-Unesp, e desenvolveu-se no âmbito da escola pública Zulmira da Silva Salles, localizada na área urbana do município de São José do Rio Preto, SP. Também

---

1 O banco de dados é vinculado ao Grupo de Pesquisa Estudos sobre a Linguagem (GPEL/CNPq).



participaram 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. Tenani e Longhin-Thomazi, 2014), a parceria universidade–escola buscou aprimorar atividades de leitura e de escrita oferecidas aos alunos do Ensino Fundamental II e também propiciar experiência e reflexão sobre a prática pedagógica da língua materna aos graduandos e pós-graduandos envolvidos.

O projeto foi desenvolvido em quatro anos (2008, 2009, 2010 e 2011).<sup>2</sup> Nesse período, foram realizadas atividades de dois tipos distintos, porém interligadas. 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. Neste tempo, os alunos deviam, junto com o aplicador responsável, ler e discutir a proposta de produção textual e, em seguida, redigir um texto sem possibilidade de reescrita.

A análise dos textos coletados nas oficinas revelou as dificuldades dos estudantes do Fundamental II em relação ao uso da escrita. Por essa razão, outra atividade do projeto foi a de oferecer minicursos facultativos aos alunos, em horários alternados aos das aulas, visando a 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 Tenani e Longhin-Thomazi (2014, p.22), 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,

---

2 Em 2008, o projeto recebeu financiamento da Fundação para o Desenvolvimento da Pesquisa de São José do Rio Preto (Faperp). 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.

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, apresentamos um exemplo de proposta.<sup>3</sup>

**Exemplo de proposta de produção textual do projeto de extensão** (cf. Reis; Tenani, 2011, p.133)

- A tirinha abaixo foi criada por Maurício de Sousa, um dos mais conhecidos cartunistas infantojuvenis brasileiro. Suas principais personagens (Mônica, Cascão, Cebolinha e Magali) foram inspiradas na vida real. Observe-a com atenção e, em seguida, discuta com seus colegas as características de cada uma das personagens, tendo em vista os desejos de cada uma.

(Na proposta, foi apresentada a tirinha da Turma da Mônica sobre o “Poço dos Desejos”, disponível em: <<http://www.monica.com.br/comics/tirinhas/tira119.htm>>. Trata-se de uma tirinha com três quadros. O primeiro retrata Mônica pensando em um ursinho de pelúcia, ao jogar uma moeda no “poço dos “desejos”. No segundo quadro, de modo semelhante, Cebolinha pensa em um carrinho de brinquedo e repete a atitude, jogando uma moeda no mesmo poço. Por fim, no terceiro quadro, é retratada Magali pensando em pirulito, sorvete, bombom, ao sair de dentro do poço com moedas na mão.)

- Imagine que agora é a sua vez de depositar uma moedinha no poço dos desejos e produza um texto contando seus maiores sonhos: que objeto gostaria de comprar, que lugares gostaria de conhecer, que profissão você pretende exercer etc.
- Seu texto deve conter de 15 a 20 linhas e deve ser escrito a tinta. Não deve ultrapassar os limites designados para a escritura.

Em 2008, primeiro ano do projeto, foram aplicadas propostas de redação junto a todas as turmas dos quatro anos do Fundamental II. A partir de 2009, assim como nos demais anos, as oficinas de produ-

3 A proposta textual apresentada é a terceira aplicada em 2008, no 6º ano.

ção 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 constituiu-se de modo a possibilitar a execução de pesquisas tanto por amostras transversais quanto por amostras longitudinais, como a nossa.

O banco de dados é composto por 5.468 textos. Destes, 2.752 estão destinados a estudos transversais e 3.614, a investigações longitudinais.<sup>4</sup> Quanto à amostra longitudinal do banco, 124 alunos participaram integralmente do projeto de extensão; desse total, 119 escreveram textos em todos os anos, somando um número final de 2.496 produções (611 textos no 6º ano, 726 no 7º ano, 477 no 8º ano e 682 no 9º ano). Com relação à quantidade de textos escritos por aluno em cada ano e ao total das coletas, a maioria (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” (Tenani; Longhin-Thomazi, 2014, p.28).

Por fim, faz-se necessário registrar que, além de todos os textos do banco terem sido digitalizados e digitados, encontra-se em desenvolvimento, sob financiamento da Fapesp (2013/14.546-5), um sistema de acesso informatizado aos textos escritos, a ser disponibilizado gratuitamente.

Tendo em vista nossa escolha por um tipo de estudo longitudinal, *a priori* já possuíamos um recorte dos textos para a composição do material, já que, entre as produções disponíveis do banco, 2.496

---

4 As produções de textos do banco são identificadas por um código, com estas informações: 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, 5ª. A, estudante 1, sexo feminino, proposta 1). Cabe lembrar que o início do projeto de extensão aconteceu quando o Ensino Fundamental ainda era composto por oito anos, o que explica as nomenclaturas antigas (5ª, 6ª, 7ª, 8ª séries) no código de identificação dos textos. No entanto, embora os escreventes do *corpus* tenham cursado oito, e não nove anos, no Fundamental, usaremos a nomeação atual dos anos letivos, porque este estudo foi realizado quando a implementação dos nove anos já se encontrava em vigor.

foram realizadas por alunos que participaram das oficinas do projeto de extensão. Após a eleição do número de sujeitos e de textos (ou seja, 119/2.496), estabelecemos mais dois critérios de seleção do material, uma vez que a grande extensão da amostra longitudinal do banco de dados inviabilizaria a organização e interpretação dos dados a partir de um cunho qualitativo. Em consequência, os dois critérios adicionais de seleção tiveram por horizonte um recorte que contemplasse, ao mesmo tempo, um conjunto relevante tanto em termos quantitativos, de modo que possibilitasse a identificação de regularidades dos dados de hipersegmentação, quanto em termos qualitativos, permitindo a análise da trajetória de cada escrevente no decorrer do Ensino Fundamental II em relação às grafias não convencionais das fronteiras de palavras.

Apresentamos, a seguir, os dois critérios que embasaram a composição do material: 1) produção, no 6º ano, de cinco e seis textos por um mesmo sujeito; 2) produção, no 7º, 8º e 9º anos, de pelo menos três textos por sujeito identificado pelo primeiro critério.

O primeiro critério foi ancorado nos resultados obtidos em pesquisa de iniciação científica<sup>5</sup> (Silva, L., 2011). Nesse estudo, dados de segmentação não convencional de palavras foram analisados em textos de 5º ano do Ensino Fundamental, a partir dos quais pudemos concluir que os escreventes encerraram essa etapa escolar sem conhecer 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.

Diante dessas constatações, adotamos o mesmo critério de seleção desses autores, ao eleger para análise os textos de alunos que haviam produzido, no 6º ano, entre cinco e seis textos (85% e 100% das coletas realizadas),<sup>6</sup> quantidade relativamente alta em relação ao número total de propostas de produção de textos aplicadas na-

---

5 Fapesp/Proc. 2010/17.720-8.

6 Nos anos de 2008 e 2010 (6º e 8º anos) foram aplicadas, em cada ano letivo, seis propostas textuais. Já nos anos de 2009 e 2011 (7º e 9º anos), foram sete propostas.

quele 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 do trabalho que originou este livro.

Definido o critério e identificados os alunos que atenderam a ele, procuramos os textos produzidos pelos alunos nos demais anos escolares do Fundamental II. Lemos todos os textos e pudemos constatar que, para o 7º, 8º e 9º anos, o primeiro critério não era pertinente aos objetivos do nosso estudo, uma vez que dados de hipersegmentação eram praticamente inexistentes na produção desses alunos. Esse cenário levou-nos a adotar mais um critério de seleção: alunos que haviam produzido pelo menos 50% das propostas aplicadas. Com o estabelecimento desse critério, buscamos ampliar o número de textos a analisar e os possíveis números de ocorrências de hipersegmentação. Formalizados os critérios, identificamos um total de 565 textos para constituir o material de pesquisa.

## **Seleção do *corpus* e identificação dos dados**

Dos 565 textos que compõem o material da nossa pesquisa, foram eleitas, para a formação do *corpus*, as produções escritas de alunos que apresentaram: 1) grafia de hipersegmentação no 6º ano; e 2) grafia de hipersegmentação em outros anos escolares. O primeiro requisito permitiu verificar se aqueles que produziam hipersegmentação no início do Ensino Fundamental II continuavam a fazê-lo ou apreendiam a convenção da palavra escrita ao longo dos anos subsequentes. Já o segundo requisito foi estabelecido com vistas a observar como se caracterizariam as hipersegmentações durante o restante do Fundamental II, tanto em relação à sua estrutura quanto às pistas de uma relação mais particular do escrevente com a escrita. Ao final do levantamento, chegamos ao *corpus* deste trabalho, composto por doze escreventes (cinco do sexo feminino e sete do sexo masculino) que, ao longo dos quatro últimos anos do Fundamental, produziram 266 produções textuais.

Na análise dos dados, fazemos referência aos escreventes do *corpus* por meio de nomes fictícios elegidos aleatoriamente, buscando preservar a identidade dos alunos e garantir aos escreventes produtores dos textos a imagem de sujeitos constituídos sócio-historicamente.<sup>7</sup> Os pseudônimos utilizados foram os seguintes: 1) André, 2) Viviane, 3) Ana, 4) Henrique, 5) Camila, 6) Fernando, 7) Júlia, 8) Mateus, 9) Maria, 10) Lucas, 11) Pedro, 12) Bruno (os números de 1 a 12 indicam ordem crescente do número de grafias de hipersegmentação produzidas).

Uma vez que o nosso objeto de estudo caracteriza-se pela presença de um limite gráfico em posição não prevista pela convenção ortográfica, uma questão metodológica importante diz respeito a estabelecer quando a presença de um recurso gráfico pode ser considerada um caso de hipersegmentação. Tal questão é pertinente sobretudo pelo fato de nosso *corpus* ser constituído por textos manuscritos, nos quais, muitas vezes, a caligrafia dos escreventes mostra irregularidades.

Para a identificação das hipersegmentações presentes em nosso *corpus*, tomamos como eixo norteador a proposta de Tenani (2011b) (Fapesp/Proc. 2009/14848-6). 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 (homônimas: “de mais”, quando a grafia é “demais”). Segundo a autora, 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 explicitar novas questões linguísticas subjacentes no processo de delimitação gráfica das palavras.

Ao adotar essa metodologia de levantamento de dados, tratamos, em um primeiro momento, da categorização do espaço em

---

7 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.

branco e percebemos quatro problemas referentes à identificação das hipersegmentações: dificuldades em interpretar a caligrafia do escrevente, casos de translineação, ocorrências de palavras estrangeiras e de nome próprio.

Com relação ao primeiro problema, tomamos a decisão de excluir do *corpus* grafias nas quais a caligrafia não permitiu definir se a palavra em análise tratava-se de uma hipersegmentação ou se consistia apenas de uma característica da forma como o escrevente registra as letras (sem ligaduras) na folha de papel e organiza o tamanho dos espaços em branco entre as palavras. A Figura 9 ilustra esse problema. À primeira vista, as duas ocorrências da palavra “provavelmente” foram entendidas como um caso de hipersegmentação, pois parece haver um limite gráfico maior entre “prova” e “velmente”. Entretanto, através de uma análise mais detalhada da distribuição dos espaçamentos entre as palavras em todo o texto, consideramos que ambas as grafias não poderiam ser classificadas como hipersegmentações, pois outras palavras do texto também pareceram apresentar a presença de um espaço: “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.

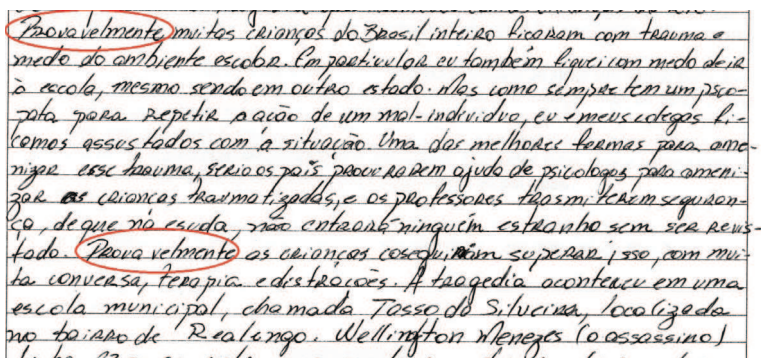


Figura 9 – Fonte: Banco de dados de escrita do EF II: Z11\_8B\_04F\_03.

O segundo problema encontrado na categorização do espaço em branco foi a translineação. Na escrita, ela ocorre devido à necessidade de separar uma palavra em duas linhas, de modo que uma parte da palavra fica no final de uma linha, e o restante, 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 separação.

Na Figura 10, apresentamos um exemplo de translineação identificada no *corpus*. Ao segmentar em duas linhas a palavra “naquele”, o escrevente não usou o hífen e, na linha em que foi grafada a sílaba “na”, havia espaço suficiente para o registro de toda a palavra, levando em conta o tamanho da letra dele. Também a parte da palavra localizada na linha inferior apresenta uma fronteira entre as sílabas “que” e “le”. No entanto, apesar das ocorrências relativas à palavra “naquele”, as quais poderiam permitir a caracterização como hipersegmentação, não a classificamos dessa maneira, pois optamos por descartar todas as grafias de palavras que estivessem no contexto da translineação.

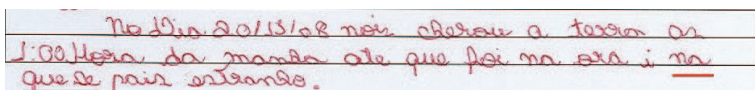


Figura 10 – Fonte: Banco de dados de escrita do EF II: Z08\_5C\_09M\_05.

A seguir são apresentados 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. As ocorrências que se encaixaram nessas categorias (nome próprio e palavra estrangeira) foram excluídas do *corpus*, já que palavras estrangeiras seguem uma ortografia distinta daquela do português e os nomes próprios não precisam, necessariamente, atender à convenção ortográfica. Extrapolam, pois, nesses aspectos, os limites da análise almejada para este trabalho quanto a grafias que se distanciam das normas ortográficas do português brasileiro.

Na hipersegmentação “Came Kase” (Figura 11), o escrevente fez referência a um brinquedo comumente encontrado em parques de





Em outro momento da seleção dos dados, tratamos da categorização do hífen. Dentre os usos desse recurso gráfico em nosso *corpus*, dois foram entendidos como não convencionais. Um uso foi excluído da análise pelo fato de o emprego do hífen nos casos levantados não configurar uma hipersegmentação, já que o critério principal que a define diz respeito ao emprego, para além do 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, pois o recurso foi utilizado de modo a unir/separar duas unidades morfosintáticas. Na verdade, essas ocorrências indicam um tipo de relação entre palavras semelhante àquela observada em palavras compostas. Destacamos, ainda, sobre o emprego do hífen em nosso *corpus*, que as grafias ilustradas na Figura 14 distinguem-se, por exemplo, do dado “arranja-se”, mostrado na Figura 1, na parte inicial deste livro, pois no caso dessa grafia não convencional o hífen hipersegmenta a palavra escrita “arranjasse”.

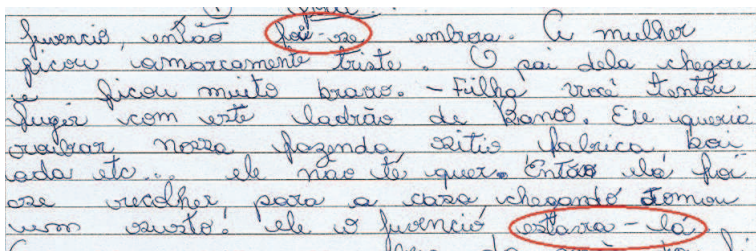


Figura 14 – Fonte: Banco de dados de escrita do EF II: Z08\_5A\_23F\_01.

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, em especial em materiais manuscritos. Nas pesquisas de Paula (2007) e Capristano (2003), por exemplo, há uma descrição de critérios que elegem dados de segmentação não convencional em textos infantis. Como

esse trabalho utiliza textos escritos do Ensino Fundamental II, consideramos necessário estabelecer critérios distintos daqueles adotados pelos estudos de dados do Fundamental I. A análise das produções manuscritas com as quais trabalhamos revelou características (quanto à forma de escrita das letras e distribuição dos recursos gráficos) muito distintas das que foram observadas nos textos de escrita inicial. Portanto, julgamos importante, para nosso estudo, um procedimento atento e particularizado para identificação dos dados, com vistas a garantir o levantamento das hipersegmentações com base em critérios explícitos e justificados, obtendo um *corpus* de hipersegmentação que caracteriza a escrita de escritores com mais anos de escolarização.

## Procedimentos de análise dos dados

As hipersegmentações identificadas nos textos do Fundamental II foram submetidas às análises quantitativa e qualitativa. O intuito foi descrever as características gerais dos dados, examinando a distribuição das ocorrências em cada ano letivo e para cada sujeito escrevente. Para verificar a distribuição das grafias de hipersegmentação, observamos a quantidade de dados considerando: a comparação entre o número de palavras convencionais e de hipersegmentação dessas palavras; a relação entre a extensão dos textos (determinada pelo número de palavras escritas) e o número de dados encontrados.

Ao organizar os dados, identificamos palavras que continuam a ser hipersegmentadas nos anos finais do Ensino Fundamental. A partir desse resultado geral, foi possível verificar: se os tipos de palavras convencionais e de hipersegmentações diminuíram com o passar dos anos, como um possível efeito das práticas letradas/ escritas; se essa possível diminuição aplicou-se também quando foi analisada a trajetória de cada escrevente; se as palavras hipersegmentadas foram sempre as mesmas em todo o Fundamental 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, considerando que a principal característica desse tipo de dado é ocorrer no nível da palavra (diferindo, nesse aspecto, das hipossegmentações, que sempre acontecem entre duas unidades), analisamos a forma como se organizam as sílabas no interior das palavras convencionais e das grafias não convencionais. Por exemplo, a palavra cujo limite convencional é “embora” tem suas sílabas organizadas metricamente da seguinte maneira: (• \* •). Essa palavra sempre foi hipersegmentada no *corpus* como “em bora”, e houve a reorganização das sílabas da palavra como: (•)  $\sigma$  (\* •)  $\Sigma$ , 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>8</sup> dos enunciados falados. Restringimos nossa investigação a 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

---

8 Na apresentação do livro *Estudos de prosódia*, Scarpa (1999) afirma 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 correlatos acústicos dos sistemas de acento, entoação e ritmo, no âmbito das línguas. Ao segundo tipo de estudo (mais fonológico) interessa a 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.

alinha-se também com a escolha pelo viés de análise proposto nos trabalhos de Tenani (2010, 2011b), Chacon (2004, 2005) e Capristano (2003, 2004), entre outros, sobre a relação entre fronteiras não convencionais de palavras e padrões rítmico-entoacionais que se formalizam em constituintes prosódicos da língua. Acreditamos, ainda, em concordância com os referidos pesquisadores, que fatos prosódicos não são exclusivos dos enunciados falados e, por isso, também podem ser recuperados na escrita.

Para subsidiar a descrição dos dados, sobretudo no que se refere à identificação de aspectos do oral/falado no texto escrito, utilizamos as formulações do modelo de fonologia prosódica de Nespor e Vogel (1986).<sup>9</sup>

Em sua proposta, essas autoras 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” (p.6).<sup>10</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 a constatação de diferentes processos fonológicos em diferentes línguas, como o turco, o latim e o grego, Nespor e Vogel chegam

---

9 A escolha da proposta de Nespor e Vogel (1986), com relação 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 essas duas autoras para a explicação de fenômenos do português brasileiro. Nesse sentido, é importante esclarecer que os objetivos deste livro não são discutir, a partir das hipersegmentações de palavras, a pertinência dos algoritmos de formação dos constituintes prosódicos para a explicação de dados da escrita. Nosso interesse é, antes de tudo, observar, a partir do arcabouço teórico da fonologia prosódica, como marcas ortograficamente discordantes podem constituir evidências de estruturas linguísticas, ou seja, como grafias de palavras refletem possíveis fatos da língua.

10 “That organizes a given strings of language into a series hierarchically arranged phonological constituents that in turn form the contexts within phonological rules apply” (tradução nossa).

à identificação de sete domínios prosódicos: sílaba ( $\sigma$ ), pé métrico ( $\Sigma$ ), palavra fonológica ( $\omega$ ), grupo clítico (C), frase fonológica ( $\varphi$ ), frase entoacional (I) e enunciado fonológico (U), organizados, nessa apresentação, em ordem crescente.

Os constituintes prosódicos baseiam-se em diferentes informações linguísticas e, desse modo, levam em consideração informações fonológicas e também de outros componentes da gramática, como o morfológico, o sintático e o semântico. Embora, em alguns casos, os constituintes prosódicos venham a coincidir com outros constituintes linguísticos, isso não acontece obrigatoriamente, 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 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>11</sup>

---

11 “**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 supordinate 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)” (tradução nossa).

Dentre os sete constituintes prosódicos, quatro mostraram-se relevantes para a análise das hipersegmentações de palavras: a sílaba, o pé métrico, a palavra fonológica e o grupo clítico.

Segundo Nespor e Vogel (1986), a sílaba e o pé métrico são os menores constituintes prosódicos e se definem, exclusivamente, por informações fonológicas. No seu interior, a sílaba 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 português brasileiro, uma vogal e, como elementos dominados (nós de valor fraco), as consoantes e/ou glides. A sílaba é dominada pela palavra fonológica, após ser organizada no interior de pés métricos. Essa última categoria prosódica caracteriza-se pela relação instituída entre duas ou mais sílabas, de modo que, à semelhança do constituinte anterior, apenas uma delas será a portadora da proeminência (o acento, nesse caso).

Ao analisar a organização das palavras no português brasileiro, Bisol (1996) considera que a maioria delas é constituída por pés métricos binários de proeminência à esquerda, um troqueu, como é a organização rítmica de palavras como “casa” e “mesa”. Existem também palavras cuja configuração é a de pés métricos binários de cabeça à direita, formando um iambo, a exemplo de “sofá” e “café”. Em palavras como “dócil” e “órgão”, há a formação de um pé espondeu, formado por duas sílabas longas, uma pelo acento, outra em razão da coda silábica; palavras com essa configuração são encontradas em baixa ocorrência no português brasileiro. Por fim, existem palavras constituídas por pé métrico 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, por exemplo, em “jacaré”, “coração”.

Seguindo os princípios da hierarquia prosódica, os pés métricos de uma sequência são dominados pela palavra fonológica. Nesse domínio, o pé métrico mais proeminente será aquele no qual sobre uma das sílabas recair o acento primário. A palavra fonológica caracteriza-se por ser portadora de um só acento, fato que faz que,

muitas vezes, não coincida com os limites da palavra morfológica. A esse respeito, Nespor e Vogel salientam que a palavra fonológica é o primeiro constituinte prosódico a manter relação com noções de outros componentes gramaticais, mais especificamente, nesse caso, o da morfologia. Por fim, o grupo clítico 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 palavra fonológica (que serve de hospedeira fonológica), o grupo clítico tem como nó mais proeminente essa palavra, já que internamente um clítico é desprovido de proeminência e, assim, não carrega a proeminência do constituinte.

Na análise que realizamos, demos maior destaque aos tipos de informações linguísticas que podem ser detectadas a partir de uma organização prosódica dos dados. Com essa escolha, não descartamos 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 trata-se apenas de aspectos das práticas orais/faladas subjacentes às fronteiras não convencionais de palavras, mas sempre há um trânsito entre essas práticas e as práticas letradas/escritas que constituem todo e qualquer material escrito. Dessa maneira, a nossa identificação das marcas letradas segue duas direções: a primeira trata do aspecto visual, ou seja, da atenção sobre as formas de delimitação das palavras, com o uso de espaço em branco ou hífen; a segunda 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 ocorrências muitas vezes únicas da escrita dos escreventes. Nesse aspecto, tratamos também os dados de hipersegmentação que, estruturalmente, não seguiram todas as regularidades observadas na descrição geral dos dados, mas 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 ocorreu hipersegmentação. Considerar o texto pode ser relevante para a discussão e explicação dos dados, na medida em que alguns dos limites não convencionais podem oferecer pistas de que tenham sido utilizados como um modo de articular os sentidos do texto e garantir uma prosódia ao dizer, como 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 pudéssemos contemplar uma investigação tanto do geral quanto do particular das grafias de hipersegmentação.

## Uma síntese

Inicialmente, abordamos a caracterização do Banco de Dados de Escrita do Ensino Fundamental, um banco inédito de textos do Fundamental II que está disponível para pesquisas na Unesp, campus de São José do Rio Preto. Apresentamos também os critérios para a seleção do material utilizado, que levaram em conta um recorte já preestabelecido em relação à amostra longitudinal do banco de dados. Em seguida, tratamos da definição do *corpus* e dos aspectos gráficos e morfossemânticos relevantes para a atenta seleção dos dados de hipersegmentação. Com relação a essa última tarefa, durante a leitura das produções escritas do *corpus*, excluímos da análise aquelas grafias de espaço em branco que se caracterizaram como casos de translineação, palavra estrangeira, nome próprio, e casos em que a caligrafia do escrevente deixou em dúvida se havia a presença de espaço ou tratava-se apenas 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, desconsideramos o uso desse recurso entre duas palavras, por não termos identificado, nesses casos, a ocorrência da não convencionalidade dos limites morfossintáticos da palavra.

Por último, dedicamos nossa atenção às formas de análise dos resultados e aos aspectos metodológicos que a embasaram. Esbo-

çamos então os pressupostos do modelo de fonologia prosódica de Nespor e Vogel (1986), o qual trata da organização prosódica das línguas, por meio de sete domínios. Foram trazidas maiores informações acerca dos algoritmos de formação dos domínios que se mostraram pertinentes para a análise e a interpretação das hipersegmentações utilizadas no trabalho que originou este livro.



## 4

# ANÁLISE DAS HIPERSEGMENTAÇÕES

Neste capítulo, expomos a análise dos dados de hipersegmentação. Por meio dessa análise, discutimos: aspectos referentes à distribuição das ocorrências, os quais indicam hipóteses mais gerais dos escreventes em relação à noção de palavra escrita; e aspectos mais particulares das grafias não convencionais, tendo em conta hipóteses individuais de alguns dos escreventes analisados. A cada um desses aspectos é dedicada uma seção deste capítulo.

## Características gerais das hipersegmentações

Dos 266 textos produzidos no decorrer dos quatro anos do Ensino Fundamental II foi possível extrair 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>1</sup>

---

1 Os dados de híbrido são: “a olado” (ao lado) e “come cheia” (comecei a), ambos produzidos no 8º ano pelo escrevente Bruno. Esses dados serão analisados a seguir, neste capítulo.

A produção do número de textos e dados, em cada um dos anos escolares, é apresentada na Tabela 1, na qual aparecem os números absolutos e os números percentuais correspondentes.

Tabela 1 – Distribuição dos textos e das hipersegmentações nos anos escolares do Ensino Fundamental II

Ano escolar	Total de textos		Total de hipersegmentação	
	Número	%	Número	%
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%
Total	266	100%	113	100%

Pela tabela, é possível observar a distribuição das grafias de hipersegmentação. Destaque-se o menor número de dados do 6º ano (40 dados, 35,3%) para o 9º ano (22 dados, 19,4%). No entanto, ao comparar esses dados com o total de dados encontrados nos outros dois anos escolares (7º e 8º anos), notamos que a queda parece não se dar da mesma forma, ou seja, a diminuição no número de dados não acontece de modo regular em todo o Fundamental II, pois, do 7º para o 8º ano, há um aumento na quantidade de hipersegmentações (de 24 dados – 21,3% – para 27 dados – 24%). Por isso, entendemos necessário investigar mais atentamente o 8º ano do Fundamental II, já que também outros estudos (ver Tenani, 2010, 2011b) sobre o mesmo objeto e a mesma etapa escolar obtiveram resultados semelhantes aos do nosso estudo. Quais fatores mobilizariam essa diferença que, aparentemente, caracteriza os anos escolares do Fundamental II *no corpus* investigado?

Para avaliar a confiabilidade desse primeiro resultado geral sobre a distribuição das hipersegmentações para, com base nele, fazer generalizações e comparações com outras pesquisas que analisaram o Fundamental II, realizamos o cotejamento do número total de dados e do tipo de palavra escrita que deu origem à hipersegmentação em cada ano. Ou seja, a comparação deu-se entre a palavra na qual foi usado o limite não convencional – por exemplo, a palavra

escrita “atrás” – e o número de hipersegmentações nela realizadas – por exemplo, as ocorrências “a traz”, “a tras”, “a trais”.<sup>2</sup>

Ao considerarmos essa nova variável, supusemos que o maior número de dados, no 8º ano, pudesse estar relacionado a grafias não convencionais de um mesmo tipo de palavra. A partir dessa hipótese, acreditamos poder demonstrar que os escreventes analisados, embora produzam muitas hipersegmentações, fazem-nas nas mesmas palavras. Portanto, as dificuldades com relação aos limites da palavra escrita que permanecem nessa etapa escolar tendem a acontecer em certas palavras, e não em qualquer palavra. A Tabela 2 ilustra essa relação entre número de dados e palavra escrita convencional.<sup>3</sup>

Tabela 2 – Relação entre número de palavras escritas convencionais e número de hipersegmentações, por ano escolar

Ano escolar	Total de palavras convencionais		Total de hipersegmentações	
	Número	%	Número	%
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>

Os resultados da tabela permitem afirmar que o número de palavras que geraram dúvidas, quanto ao limite gráfico, tende a diminuir ao longo do Fundamental II. Tanto no que se refere ao tipo de palavra escrita quanto às hipersegmentações, fica evidente a diferença quantitativa entre o 6º ano e os demais anos dessa etapa escolar. A elevação no número de grafias não convencionais no 8º ano é muito menor quando comparada ao número de dados do 6º,

2 Exemplos retirados do 7º e do 8º anos.

3 Os números referentes às palavras escritas convencionais foram obtidos por meio da contagem, em cada ano escolar, dos tipos de palavras nas quais 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, ela foi contabilizada apenas uma vez.

7º e 8º anos. Isso nos leva a supor que, em alguma medida, novas informações linguísticas introduzidas na nova etapa do Ensino Fundamental podem ser um “gatilho” para que os escreventes formulem novas hipóteses sobre a escrita.

Como observamos na Tabela 2, talvez, ao longo dos anos escolares, essas hipóteses sejam sistematizadas e se aproximem da convenção. Uma investigação exclusiva sobre dados de segmentação não convencional presentes em textos do 6º ano já foi realizada por Paranhos (2010) e sintetizada em Paranhos e Tenani (2011). No nosso estudo, buscamos indícios que nos auxiliem na caracterização das hipersegmentações dos anos do Fundamental II como um todo.

Voltando à diferença expressa na Tabela 2 sobre a quantidade do tipo de palavra e do número de hipersegmentações correspondente em cada ano, há palavras que se repetem e possibilitam maior número de dados. No *corpus*, 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”. A palavra “demais”, em todos os anos, deu origem a doze grafias não convencionais (equivalentes a 10,61% do total de dados). Foram dois dados no 6º ano, dois no 7º ano, cinco no 8º ano e três no 9º ano. Ao excluir 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 relativas ao número de dados ao longo do ano escolar antes observadas desaparecem quando se excluem da análise os dados que sempre se repetem. Isso nos leva a apontar, mais uma vez, a necessidade de, na análise desse tipo de dado, levar em conta não só a quantidade das ocorrências não convencionais, mas, fundamentalmente, os tipos dessas ocorrências na relação com a palavra convencional que as origina.

Verificamos que as grafias de hipersegmentação, no final do Ensino Fundamental (pelo menos no *corpus* analisado), diminuem ao longo dos anos que o constituem, mas não só elas: o número de palavras cujos limites mostram-se como um desafio aos escreventes também reduz. Nesse aspecto, podemos inferir que a ideia dos escreventes de que, na escrita, as palavras são marcadas por limites

(ora com o uso do branco, ora do hífen) altera-se, aproximando-se mais da convenção ortográfica.

Acreditamos que esse resultado se deva às atividades letradas/ escritas das quais os escreventes participam no decorrer de todo o Fundamental II nas aulas de Língua Portuguesa, sobretudo. Isso não significa que estamos negando a relevância da inserção dos escreventes em outras práticas sociais de linguagem para a representação que construíram para a (sua) escrita, pelo contrário, uma vez que concordamos com o fato de que todo texto sempre é 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).

Na Tabela 3 constam informações sobre a distribuição dos textos e sobre os dados produzidos pelos escreventes investigados.

Tabela 3 – Distribuição do número de textos e hipersegmentações produzidas, por escrevente

Escreventes	Total de textos		Total de hipersegmentações	
	Número	%	Número	%
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%
Júlia	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>



Na tabela, apresentamos as diferenças na quantidade de hipersegmentações entre os escreventes. O total de dados na relação intersujeitos apresenta grande variação. As diferenças 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%.

Ressalte-se que os dois escreventes que mais produziram marcas de hipersegmentação de palavras (Bruno: 37 ocorrências; Pedro: 22) foram os que elaboraram a menor quantidade de textos em todos os anos escolares (Bruno: 19 textos; Pedro: 17), enquanto os dois alunos que produziram menor número de hipersegmentações (André: duas ocorrências; Viviane: duas ocorrências) foram os que elaboraram a maioria das propostas de textos aplicadas ao longo dos quatro anos do Fundamental II (André: 25 textos; Viviane: 23).

Com base nesses fatos, duas hipóteses podem ser formuladas, e uma não exclui a outra. A primeira é que essas diferenças podem estar ancoradas nos modos de participação dos alunos em práticas sociais de linguagem, as quais os levam a elaborar representações sobre a escrita (sobretudo aquela privilegiada pela instituição escolar) que os guiam no uso das convenções ortográficas em seus textos (em particular, o registro de fronteiras de palavras, foco do nosso trabalho). A segunda hipótese tem relação com a extensão dos textos produzidos. O fato de um escrevente ter produzido maior número de grafias hipersegmentadas em alguma medida guarda relação com o fato de ele ter escrito textos com mais palavras, o que amplia a possibilidade de aparecimento de grafias não convencionais. Do mesmo modo, o fato de outro escrevente produzir poucas ocorrências de hipersegmentação pode estar ligado ao menor tamanho dos seus textos, o que restringe as possibilidades de produção de hipersegmentações.

Para verificar a validade dessa última hipótese, comparamos a quantidade total de palavras redigidas nos quatro anos escolares com o total de hipersegmentações. O resultado obtido é apresentado na Tabela 4.

Tabela 4 – Número de hipersegmentações em relação ao número de palavras por escrevente

Escreventes	Total de			Hipersegmentações em relação ao número de palavras (%)
	Textos	Palavras	Hipersegmentações	
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%
Júlia	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>

A tabela permite observar a relação entre os escreventes que produzem mais/menos grafias de hipersegmentação, a partir da extensão dos textos, em função do número de palavras neles utilizadas. Em termos gerais, as marcas de hipersegmentação representam 0,32% de tudo o que foi escrito pelos doze escreventes, o que indica que as fronteiras de palavras marcadas além do esperado pela ortografia, embora persistentes nos anos que encerram o Fundamental, aparecem em pequena quantidade nessa etapa escolar.

Com relação à trajetória escrita de cada escrevente, o levantamento do número de palavras dos textos ressalta que as diferenças entre os sujeitos estão diretamente relacionadas a quanto cada um escreveu, pois foram identificados:

- escreventes que grafam quantidade idêntica de hipersegmentações e próximas de textos (André e Viviane: dois dados em 25 e 23 textos; Ana, Henrique e Camila: quatro dados em 20, 22 e 25 textos; Júlia e Mateus: sete dados em 25 textos cada

um), mas apresentam percentuais distintos de representatividade dos dados em relação a toda a sua produção ao longo do Fundamental II (André e Viviane: 0,09% e 0,06%; Ana, Henrique e Camila: 0,11%, 0,13% e 0,12%; Júlia e Mateus: 0,19% e 0,20%);

- escreventes que produzem mais dados em função do maior número de palavras redigidas (Júlia, 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 (André e Viviane: dois dados cada um em 2.122 e 3.196 palavras);
- escreventes que se destacam pela alta representatividade das hipersegmentações em comparação com o tanto que escrevem em todo o Fundamental II (Fernando: 0,28%; Lucas: 0,30%; Pedro: 1,38% e Bruno: 2%).

Corroborar para que afirmemos que esses escreventes são os que têm mais dúvidas quanto aos limites das palavras o fato de que o número de palavras escritas, em relação ao total de textos produzidos, foi baixo, em comparação com outros alunos que produziram uma quantidade de textos idêntica ou com pouca variação. A título de exemplificação, Ana (que escreveu, ao todo, vinte textos) grafou 1.698 palavras a mais do que Fernando (que produziu vinte textos) e 1.610 palavras a mais do que Bruno (que elaborou dezenove textos). Embora Fernando, Lucas, Pedro e Bruno tenham participado com grande frequência das atividades de elaboração dos textos, suas produções, por terem curta extensão, evidenciam a dúvida em relação a quando as palavras devem ter autonomia gráfica. Mais especificamente, ao observarmos a distribuição das hipersegmentações nas escritas 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, apresentaremos evidências que auxiliem na compreensão dos tipos de fatos linguísticos que são mobilizados

nas hipóteses desses dois escreventes e que fazem que eles se particularizem em relação aos demais.

Na Tabela 5 apresentamos a distribuição dos totais de palavras escritas e de dados produzidos por escrevente em cada ano escolar.

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

Escreventes	Número de palavras/ano escolar				Número 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
Júlia	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>

Analisando a tabela, no que diz respeito aos anos letivos, identificamos três trajetórias dos escreventes quanto à produção dos registros não convencionais, com a presença das fronteiras de palavras. A primeira é a de alunos que produziram hipersegmentações em todos os anos; a segunda, de alunos que produziram ocorrências em três dos quatro anos; e, por fim, a terceira, de alunos que apresentaram dados em dois dos quatro anos escolares. Constatamos que:

- cinco alunos encaixam-se na primeira trajetória, sendo também os que mais produziram hipersegmentações;

- dois alunos enquadram-se na segunda trajetória. Um deles apresentou ocorrências no 6º, 8º e 9º anos, e o outro, no 6º, 7º e 9º anos, correspondendo ao segundo conjunto de alunos que mais produziram ocorrências;
- cinco alunos encaixam-se na terceira trajetória. Dois produziram ocorrências no 6º e no 7º anos e os outros três, no 6º e no 9º anos, configurando-se como o grupo de escreventes que produziu o menor conjunto de dados.

Em síntese, o desenho das trajetórias revela que apenas dois escreventes concluíram o Ensino Fundamental sem produzir mais grafias de hipersegmentação em seus textos, enquanto a outra parte dos escreventes terminou essa etapa escolar oscilando entre a escrita convencional e a não convencional das fronteiras de palavras (em particular daqueles limites marcados além do previsto). Entretanto, mesmo que as hipersegmentações tenham continuado a aparecer na produção escrita da maioria dos escreventes do nosso estudo, muitos deles, no fim, aproximaram-se da convenção ortográfica. Esta afirmação encontra fundamento na análise da relação entre as grafias de hipersegmentação e a extensão dos textos elaborados no transcorrer do Fundamental II. Essa relação é visualizada por meio do Gráfico 1, cujos percentuais foram obtidos a partir dos números da Tabela 5.

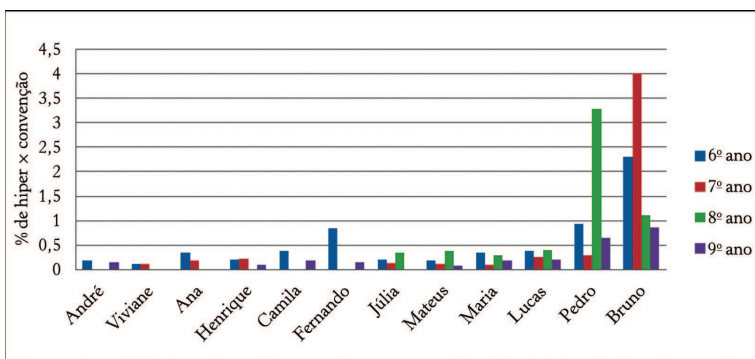


Gráfico 1 – Relação entre hipersegmentação e extensão dos textos (pelo número de palavras), por escrevente

Sobre o gráfico, destacamos que:

- independente 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 registra individualmente em cada ano escolar;
- na trajetória pelo Fundamental II, o número das ocorrências dos escreventes André, Viviane, Ana, Camila, Fernando e Bruno (metade dos sujeitos investigados) diminui quando elas são cotejadas com a extensão dos textos, o que indica que a escrita desses escreventes foi se encaminhando em direção às normas ortográficas que regem o estabelecimento dos limites da palavra escrita;
- no caso em que o percentual de dados não decaiu com o passar dos anos escolares, levantamos duas possíveis hipóteses. A primeira parece permear a escrita de Henrique e Júlia e pode ter relação também com o 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 palavras escritas que originam as hipersegmentações, e é o que aparece na escrita de Mateus, Maria, Lucas e Pedro.

Com vistas a explicitar os tipos de palavras em que ocorreu o uso da fronteira não convencional, apresentamos, nos quadros 1 a 12 a seguir, todos os dados de cada escrevente, separados por ano escolar.

Quadro 1 – Grafias de hipersegmentação produzidas por André

André			
6º ano	7º ano	8º ano	9º ano
e norme (enorme)			amasa-se (amansasse)

Quadro 2 – Grafias de hipersegmentação produzidas por Viviane

Viviane			
6º ano	7º ano	8º ano	9º ano
se nários (cenários)	com migo (comigo)		

Quadro 3 – Grafias de hipersegmentação produzidas por Ana

Ana			
6º ano	7º ano	8º ano	9º ano
quise-se (quisesse) de mais (demais)	de mais (demais) mata-do (matado)		

Quadro 4 – Grafias de hipersegmentação produzidas por Henrique

Henrique			
6º ano	7º ano	8º ano	9º ano
ganha-se (ganhasse)	estava-mos (estávamos) de ele (dele)		a onde (aonde)

Quadro 5 – Grafias de hipersegmentação produzidas por Camila

Camila			
6º ano	7º ano	8º ano	9º ano
de ele (dele) deis de (desde)			conversa-se (conversasse) arranja-se (arranjasse)

Quadro 6 – Grafias de hipersegmentação produzidas por Fernando

Fernando			
6º ano	7º ano	8º ano	9º ano
em fachado (enfaixado) em tão (então) na quele (naquele) vou tando (voltando)			de mais (demais)

Quadro 7 – Grafias de hipersegmentação produzidas por Julia

Júlia			
6º ano	7º ano	8º ano	9º ano
a caba (acabar) anti penúltimo (antepenúltimo)		Vira-se (virasse)	de mais (demais) a tirar (atirar) de pois (depois) a cender (acender)

Quadro 8 – Grafias de hipersegmentação produzidas por Mateus

Mateus			
6º ano	7º ano	8º ano	9º ano
a quele (aquele)	da quela (daquela)	loca mente (loucamente) de mais (demais) de mais (demais) e moção (emoção)	de mais (demais)

Quadro 9 – Grafias de hipersegmentação produzidas por Maria

Maria			
6º ano	7º ano	8º ano	9º ano
de se (desse) a panho (apanhou) a gora (agora)	de mais (demais)	su por (supor) nem um (nenhum) de mais (demais)	contra bando (contrabando) contra bando (contrabando)

Quadro 10 – Grafias de hipersegmentação produzidas por Lucas

Lucas			
6º ano	7º ano	8º ano	9º ano
em quanto (enquanto) em cando (enquanto) des de (desde)	da qui (daqui) da quele (daquela)	em bora (embora) em tão (então) fala-sem (falassem)	na quele (naquele) a quela (aquela)

Quadro 11 – Grafias de hipersegmentação produzidas por Pedro

Pedro			
6º ano	7º ano	8º ano	9º ano
de ele (dele) a cabano (acabando) a quele (aquele) a perecer (aparecer) na viu (navio) de sero (desceram)	foute i (voltei)	ir motal (imortal) a trais (atrás) a liversario (aniversário) a manha (amanhã) em comtrei (encontrei) na quêla (naquela) de mais (demais) a migo (amigo) em Bora (embora) a pertou (apertou) a paichonei (apaixonei)	de pois (depois) a doro (adoramos) o vino (ouvindo) a fogo (afogo)



Quadro 12 – Grafias de hipersegmentação produzidas por Bruno

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

Nos quadros, aparecem os casos de hipersegmentação e as palavras escritas em que ocorreram. Fica clara a dúvida dos escreventes em grafar os mesmos tipos de palavras e, em particular, palavras que se assemelham quanto à distribuição das sílabas. No que se refere à relação entre tipo de palavra e hipersegmentação na trajetória escrita de cada escrevente, averiguamos que somente André e Viviane realizaram duas hipersegmentações de duas palavras diferentes. Os dados desses escreventes, como observado anteriormente, apresentam pouca representatividade em relação a toda a sua produção no decorrer do Fundamental 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 essa observação não se aplica aos outros escreventes. Em termos gerais, primeiramente, constatamos que os limites não convencionais ocorreram sobretudo em palavras cujas sílabas pretônicas identificam-se com elementos escritos da língua,

como: “a” (trinta 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”, a palavra “demais” foi a mais hipersegmentada, não apenas para os dados desse item, mas em todo o *corpus*, o que permite dizer que o limite dessa palavra é o que se mostra mais instável na escrita de nossos sujeitos. Por exemplo, dos sete dados produzidos por Mateus, três são dessa palavra; então, apenas cinco palavras escritas geraram dúvida para esse escrevente. A mesma palavra ocasionou, 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; as outras aconteceram devido à dificuldade em distinguir sequências de verbo + pronome e formas verbais do subjuntivo. Quanto ao hífen, delimitar palavras por meio desse recurso constituiu uma dificuldade também para Henrique e Camila, pois metade dos dados produzidos por esses escreventes está relacionada ao emprego do hífen.

Retomando os tipos de palavras em que identificamos a separação não convencional pelo espaço em branco, as hipersegmentações com “a” foram as mais frequentes e também as que originaram maior diversidade de palavras escritas.<sup>4</sup> Dados em que o “a” foi grafado entre limites foram encontrados nos textos de Henrique, Júlia, Mateus, Maria, Lucas, Pedro e Bruno, isto é, em mais da metade dos escreventes do *corpus*. Especialmente em relação às hipersegmentações de Pedro e Bruno, o “a” como palavra autônoma representou 50% e 29,7%, respectivamente, do total de dados desses escreventes.

As hipersegmentações em que identificamos a partícula “em” marcaram a produção dos escreventes Fernando, Lucas e Pedro.

---

4 Note-se, sobre essa variedade, a ocorrência de mais de uma grafia não convencional apenas das palavras “atrás”, “ai” (ver Bruno e Pedro) e “aquele” (ver Mateus e Pedro).

Esses registros ressaltam a identificação com palavras próprias da escrita (e que, portanto, devem ser demarcadas entre limites), pois todas foram grafadas com o grafema <m>, mas as palavras que dão origem a essas hipersegmentações são convencionalmente grafadas com <n> (com exceção da palavra “embora”). A mesma observação estende-se aos dados que demarcam fronteira para a palavra “com”; das oito ocorrências não convencionais dessa palavra, sete foram encontradas na escrita de Bruno, espalhadas ao longo de todo o Fundamental II, sinalizando, prioritariamente, o fato de esse escrevente não ter clareza quanto ao registro gráfico de “com” como sílaba ou como palavra independente.

Seguindo a análise das características gerais das hipersegmentações, passamos a tratar dos recursos gráficos. Por adotarmos a proposta de Tenani (2011b) sobre a classificação das segmentações não convencionais de palavras de acordo com o tipo de recurso gráfico empregado, optamos por analisar em separado as ocorrências caracterizadas pela presença do espaço em branco e do hífen. A frequência das hipersegmentações com o uso dos dois recursos gráficos, ao longo dos anos escolares do Fundamental II, é apresentada na Tabela 6.

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

Ano escolar	Recurso gráfico				Total	
	Espaço em branco		Hífen		Número	%
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>

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 de delimitação das palavras na escrita, bem como às estruturas demarcadas por cada

um dos recursos. Em outras palavras, os dados, pela 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.

A Tabela 6 mostra que, do total das marcas de hipersegmentação, 92% está relacionado ao uso não convencional do espaço em branco e 8%, ao emprego não convencional do hífen. Embora, nos dados do *corpus*, 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 em separado em relação ao do espaço em branco, uma vez que esses usos não convencionais revelam aspectos linguísticos importantes sobre o processo de letramento dos escreventes (Tenani, 2011b).

A delimitação não convencional das palavras realizada por meio do espaço em branco (recurso predominante) diminuiu ao longo do Ensino Fundamental II, enquanto os usos não convencionais do hífen (utilizado para delimitar poucos tipos de palavras<sup>5</sup>) mantiveram-se inalterados por quase toda essa etapa escolar, com pequena elevação do 8º para o 9º ano. Entendemos que esse resultado indica que a escolha do hífen para delimitar 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 em uma amostra transversal, conforme constatado por Tenani (2011b).

Todos os usos de hífen identificados concentram-se na tentativa de grafias com a estrutura verbo + pronome enclítico. Nos enunciados falados, o português brasileiro privilegia a próclise (“te quero” ao invés de “quero-te” – cf. Bisol, 1996, 2000, 2005), e é interessante notar que o registro do hífen ocorreu em palavras que remetem a uma estrutura cujo funcionamento é ensinado na escola e, sobretudo, por ela valorizado. Mas essa afirmação ganha contor-

---

5 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).

nos mais precisos se considerarmos que, das nove ocorrências de hífen, seis envolvem o uso da partícula “se”: “amasa-se” (amansas-se), “quise-se” (quisesse), “ganha-se” (ganhasse), “conversa-se” (conversasse), “arranja-se” (arranjas-se), “vira-se” (virasse). Essas formas, em sua grande maioria, existem no funcionamento da escrita convencional, mas, nesses casos, distanciam-se da convenção, por ocuparem a função sintática de formas verbais do tempo subjuntivo.

Se, uma vez mais, considerarmos os usos orais/falados da linguagem, perceberemos que o uso do subjuntivo já é, nos dias atuais, praticamente inexistente, em especial no uso cotidiano da língua. Com relação a estruturas como “ganha-se”, o que parece fazer que sejam usadas é o forte impacto de práticas letradas/escritas (incluindo a escolar) mais institucionalizadas, que requerem o emprego de um tipo de escrita baseada em uma norma culta padrão. Também, com frequência, elas aparecem veiculadas em anúncios – “vendem-se”, “alugam-se”, dentre outros exemplos –, que certamente produzem grande impacto.

Assim, da perspectiva de escrita assumida neste livro (Corrêa, 2004), consideramos 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 que revela a aproximação com uma projeção do que seja a representação do código escrito institucionalizado. Não excluimos da nossa análise a ação da gênese da escrita e da dialogia sobre o já falado/ouvido e escrito/lido (Corrêa, 2004). 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.

A análise realizada até este momento buscou oferecer um entendimento geral a respeito da distribuição dos dados ao longo dos anos escolares relativos aos escreventes selecionados e aos tipos de recursos gráficos delimitadores de palavras. Prosseguindo, passamos a investigar as características gerais das hipersegmentações no que diz respeito às suas configurações rítmicas, iniciando pelas

estruturas das palavras escritas convencionalmente. Partimos do pressuposto de que as informações rítmicas, que se formam no interior da palavra, por meio de proeminências prosódicas, constituem pontos de ancoragem para o tipo de limite gráfico não convencional de palavras registrado na escrita.

Na Tabela 7 a seguir,<sup>6</sup> verificam-se, com base na organização proposta, sete tipos de estruturas rítmicas de palavras convencionais que, ao todo, geraram catorze subtipos de estruturas rítmicas de grafias hipersegmentadas. Há também outro tipo, que nomeamos “outros”, no qual incluímos hipersegmentações cujas fronteiras não se assemelham, em termos estruturais, às tendências observadas.<sup>7</sup>

---

6 A interpretação que fizemos dos dados, como a combinação entre  $\sigma + \Sigma$ , pode ser entendida também como resultado de uma combinação entre  $\sigma + \omega$ , uma vez que, no modelo prosódico de Nespor e Vogel (1986), aqui utilizado, a palavra fonológica forma-se pelas sílabas agrupadas no interior de um pé, o qual carrega o acento primário, seu principal definidor. Nessa linha interpretativa, poderíamos considerar que uma palavra fonológica teria sido registrada como um grupo clítico (cf., sobre essa análise, Tenani, 2009a, 2011b, e 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.

7 Sobre a organização da Tabela 7, cabem, ainda, duas explicações: 1) a respeito de diferentes atribuições de sílaba forte ou fraca para as grafias de hipersegmentação, levando em consideração essa mesma organização da palavra convencional, destacamos os subtipos (1.2) e (3.3). Com relação 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). No caso do subtipo (3.3), entendemos que a sílaba pretônica *so*, da palavra “sozinho”, pode 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, consideramos 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, e não apenas as prosódicas, podem ter contribuído para essa reorganização rítmica notada nas hipersegmentações; 2) mais de uma organização rítmica da palavra escrita, em virtude das possibilidades de atribuição de acento secundário, como é o caso dos tipos (6) e (7), os quais nos levaram a supor também mais de uma configuração para o registro da hipersegmentação.

Para observar longitudinalmente regularidades na distribuição dos dados relativos ao funcionamento de cada estrutura rítmica, consideramos os mesmos tipos em todos os anos escolares. Para a definição dos tipos, usamos o seguinte critério: as estruturas rítmicas foram agrupadas de acordo com o número de sílabas da palavra convencional. Desse modo: os tipos (1) e (2) consistem em palavras dissílabas; os tipos (3) e (4), em palavras trissílabas; e os tipos (5), (6) e (7), em palavras polissílabas.

Tabela 7 – Estruturas rítmicas de palavras e de hipersegmentações por ano escolar

Estruturas rítmicas				Exemplo de hipersegmentação	Hiper/ano escolar				Total	
Tipos	Palavra	Subtipos	Hipersegmentação		6ª	7ª	8ª	9ª	Número	%
(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>(1)</sup>	1	0		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			“adisio o na”	3	3	1	4	11	9,8
Total					40	24	27	22	113	100

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

(1) Uso de “a panho” no lugar de “apanhou”.

De acordo com a Tabela 7:

- 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 hipersegmentações.

- Não houve uniformidade na distribuição de cada tipo e sub-tipo de estrutura nos anos escolares, o que equivale a dizer que nem todos os que foram identificados ocorreram sempre.
- Os tipos (1) e (3), além de presentes nos quatro anos do Fundamental II, foram também os mais frequentes no 6º, 7º, 8º e 9º anos. Ainda, em todos os anos, foram identificadas grafias do tipo “outros”, as quais, 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 serão tratadas na parte seguinte deste capítulo.
- No 6º ano foram encontrados os tipos (2), (4), (5) e (6) de palavras convencionais e os subtipos (2.0), (4.0), (5.2) e (6.0) de hipersegmentações. Em relação ao conjunto proposto na Tabela 7, apenas a estrutura (7) não ocorreu nesse ano escolar. Já no 7º ano identificamos os tipos e subtipos (2)/(2.0), (5)/(5.1) e (7)/(7.0). Portanto, em relação ao ano anterior, ocorreu redução da diversidade de todas as estruturas rítmicas. Com relação ao 8º 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 9º ano apresentou a menor variedade de estruturas, dos tipos (4) e (5) e dos subtipos (4.0), (5.1) e (5.3).

O primeiro fato a destacar, pelo exposto, é a presença dos recursos demarcadores de palavra fora da convenção em fronteiras de constituintes prosódicos, as quais foram interpretadas como sendo entre a sílaba e o pé métrico, os dois domínios mais baixos do modelo prosódico assumido. Podemos considerar 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 Fundamental II, as estruturas que mais levam a grafias não convencionais e 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”.



Foram frequentes, também, ainda que em menor número, grafias não convencionais de palavras polissilábicas do tipo (5), como “loucamente” e “recompensa”.

Com relação aos tipos mais comuns, as palavras convencionais abrangidas em (1) têm a estrutura organizacional de um pé métrico iambo (alternância rítmica fraco/forte), a qual, quando hipersegmentada, passa a corresponder à combinação de duas sílabas. No caso do tipo (3), foi a partir da estrutura trissílaba da unidade convencional que os escreventes realizaram uma escrita baseada na relação entre uma sílaba e um pé métrico, predominantemente quando sua organização correspondia à de um troqueu (cf. subtipo (3.1)). No subtipo (5.1), determinadas palavras formadas por quatro sílabas foram grafadas como duas unidades dissílabas, as quais também correspondem à organização de pés trocaicos. Esse resultado segue as constatações de Abaurre (1991c) em dados sobre escrita infantil: as crianças parecem preferir a grafia de palavras dissílabas cujo padrão acentual é o paroxítono (o qual carrega uma formação calcada em um pé métrico troqueu). Assim, ao compararmos os nossos resultados com os de Abaurre (1991c), encontramos evidências de que, em alguma medida, informações sobre as formas canônicas de palavras do português brasileiro ainda são pontos de ancoragem nas hipóteses dos escreventes que concluem o Fundamental.

A partir do tipo (1) e dos subtipos (3.1), (3.2), (5.1) e (5.3), identificamos 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, 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, pelo espaço em branco ou pelo hífen. Parece que, nesses momentos, os escreventes tentam marcar, por meio dos limites nas 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 referem-se ao uso não convencional do hífen.<sup>8</sup>

A interpretação dos dados de hífen, como marca característica de enunciados escritos, já foi explicitada em parágrafos precedentes. Gostaríamos, porém, de acrescentar mais um argumento a favor do nosso posicionamento, o qual encontra base nos resultados encontrados por Abaurre e Galves (1996).

Em estudo sobre pronomes clíticos do português brasileiro, essas autoras chegaram à conclusão de que, em enunciados falados, as estruturas enclíticas são marcadas, observando-se, pois, uma nítida preferência dessa língua pela próclise. Constata-se, então, que estruturas formadas por verbo + pronome enclítico ainda estão preservadas na língua, pelas práticas da escrita convencional, fato que fortalece a premissa de que a presença do hífen em fronteira de palavra, nos dados analisados, é um significativo indício de que, ao realizá-la em seu texto, o escrevente tenha buscado aproximar-se de aspectos mais institucionalizados da escrita.

A partir da observação das ocorrências de espaço em branco, também não descartamos a ação de informações gráficas ligadas às decisões sobre como segmentar. Ao contrário, em todos os tipos de estruturas rítmicas conseguimos identificar grafias hipersegmentadas nas quais alguma parte tenha correspondido a palavras que, ortograficamente, são espaçadas entre limites. Por exemplo, nos dados, a sílaba pretônica da palavra convencional com frequência foi isolada entre brancos, sílaba que, linguisticamente, pode cor-

---

8 Nesse conjunto, há apenas uma ocorrência de espaço em branco: “falam do” (falando).

responder 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 casos em que parte resultante da hipersegmentação aproxima-se graficamente de formas verbais, como “vou tando” (voltando). Para explicitar uma possível relação das hipersegmentações com fatos de natureza morfossintática, apresentamos, na Tabela 8, o número de ocorrências que, em alguma parte originadas da hipersegmentação, correspondem a classes de palavras morfossintáticas.

Tabela 8 – Hipersegmentações de espaço em branco e correspondências gráficas

Correspondência	Hipersegmentação	
	Número	%
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>

A tabela mostra alto percentual de ocorrências de hipersegmentação nos textos analisados que, além de darem indícios de questões referentes à organização prosódica da língua, evidenciam a relevância de considerar outras motivações ao interpretar esse tipo de dado. As hipersegmentações são condicionadas também por características advindas de informações letradas/escritas, as quais procuramos explicitar por meio da relação estabelecida entre limites não convencionais de palavras e classes de palavras morfossintáticas (as quais, 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 esta análise, verificamos quais categorias de palavras funcionais e lexicais foram identificadas. Os resultados são apresentados na Tabela 9.

Com relação às categorias de palavras funcionais, podemos dizer que a sílaba pretônica separada entre limites foi predomi-

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

Tipos de classes de palavras		Hipersegmentações	
		Número	%
Lexical	Verbo	4	3,5%
Funcional	Preposição	47	41,6%
	Artigo	30	26,6%
	Pronome	3	2,7%
Total		84	74,4%

nantemente analisada como preposição (41,6%), artigo (26,6%) e pronome (2,7%), em especial 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 revelou-se como dúvida dos escreventes, a dos verbos (3,5%), como em “recom pensa” (recompensa), “vou tando” (voltando).

Concluímos que, no final do Ensino Fundamental, os escreventes têm mais dúvidas em estabelecer limites para aqueles tipos de palavras cujo funcionamento linguístico é caracteristicamente gramatical, resultado semelhante ao descrito por Tenani (2011a) quando foram considerados dados transversais de segmentação não convencional de palavras.

Determinados dados, como “de ele” (dele), “vou tando” (voltando), “falam do” (falando), “com migo” (comigo) e todas as grafias de “em”, como “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: 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 que se verifica também para o <l> em “vou tanto” (voltando); a duplicação de letras: “de ele” (dele), “com migo”.

Essas ocorrências refletem a noção dos escreventes sobre o que são palavras escritas da língua e, mais ainda, sobre as pistas, presentes no funcionamento do próprio código escrito institucionalizado (Corrêa, 2004), que auxiliam na sua identificação. Citamos, por exemplo, a regra ortográfica pela qual, em final de palavra, a coda nasal é registrada pela letra <m> e não <n>. Generalizada pelo escrevente ao tentar reconhecer palavras, a regra origina palavras escritas terminadas por esse grafema. É o caso das separações não convencionais de sílabas pretônicas como sendo os elementos “em” e “com”. Podemos ainda associar a grafia de “de ele” para “dele” como resultado da apreensão, na sequência fônica, de duas palavras escritas bastante comuns: “de” e “ele”.

A nosso ver, essas escolhas de letras evidenciam diferentes informações linguísticas ancoradas nas diferentes formas de participação dos escreventes em práticas letradas/escritas, mobilizadas por eles quando têm diante de si a exigência escolar (e, especialmente, social) de grafar e delimitar palavras segundo as convenções ortográficas. Esses dados, uma vez mais, indicam que os escreventes do Fundamental II continuam em processo de aquisição, marcado pela relação deles com as práticas sociais de linguagem que os constituem enquanto escreventes e revelado pelas hipóteses registradas no material escrito.

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

Na discussão realizada no primeiro capítulo sobre os estudos de Bisol (2000, 2005) e Simioni (2008) a respeito do comportamento prosódico dos clíticos do português brasileiro, 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 aplicam-se na exata extensão

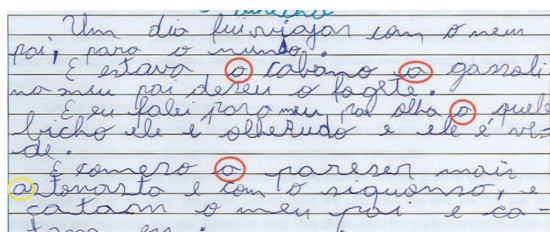
das sequências formadas por clítico + palavra. Já Simioni propõe a consideração do clítico junto a uma frase fonológica, com o argumento de regras de restrições prosódicas e de alinhamento.

Sem desconsiderar as filiações teóricas de cada um desses estudos, trazemos para discussão pontos de convergência subjacentes às análises de Bisol e Simioni: na formalização da inserção prosódica dos clíticos do português brasileiro, esses elementos estão situados no domínio de constituintes que estabelecem forte relação com o componente sintático, mostrando que os clíticos mantêm uma relação de maior independência em relação à palavra que os hospeda. Considerando que essa é uma característica do clítico do português brasileiro, visto que em outras línguas os clíticos indicam outro funcionamento (ver, por exemplo, Vigário, 2003), acreditamos que as grafias de palavras hipersegmentadas, na medida em que representam autonomia gráfica de elementos gramaticais átonos, parecem indiciar um funcionamento mais independente desses elementos também em termos fonológicos.

Apresentamos, a seguir, dois textos em que foram registradas diferentes grafias para os clíticos *a* e *em*, as quais julgamos corroborar a hipótese exposta. Ao observar 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 dos casos em que o possível isolamento do “a” daria origem a uma estrutura, na palavra seguinte, que, em termos de sílaba, não é gramaticalmente possível para o português brasileiro; referimo-nos à grafia de “astonasta” (astronauta) e à possível sílaba -sto (caso tivesse ocorrido a hipersegmentação).

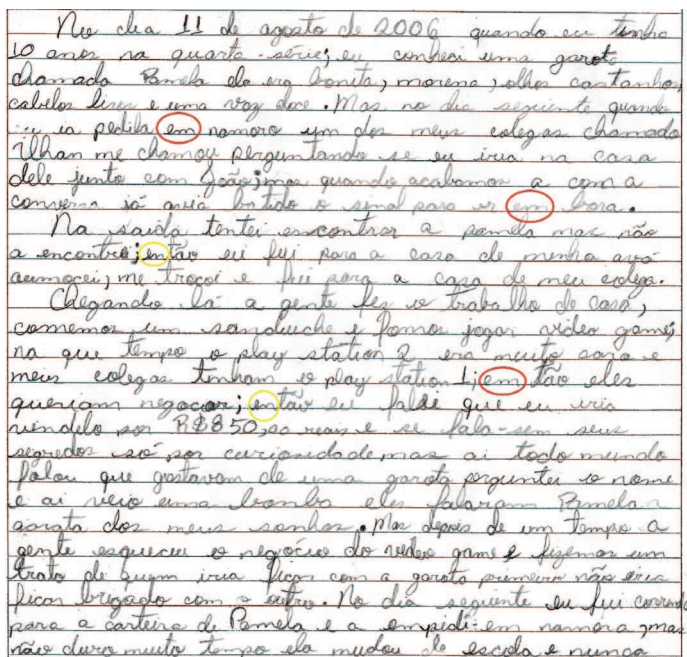
Na Figura 16, as grafias em que o escrevente tenta sinalizar o reconhecimento da palavra “em” são realizadas com o grafema <m>; no entanto, quando esse reconhecimento refere-se à partícula como uma sílaba da palavra, tal registro se dá por meio da letra <n>. Essas flutuações gráficas demonstram a dúvida dos escreventes com relação a fatores linguísticos relacionados com o funcionamento prosódico de elementos não acentuados em convergência (de modo a se complementarem) a aspectos de práticas letradas/escritas que

permeiam a noção do que é uma palavra na escrita convencional. Acreditamos ainda que, além de as hipersegmentações demonstrarem uma relação complexa entre enunciados falados e escritos, o contato que os escreventes passam a ter com o funcionamento da escrita é fator crucial que os leva a observar esses elementos clíticos da língua como uma categoria linguística que, em determinadas situações, podem ser palavras escritas.



Uma das férias que eu tive  
foi, para o mundo.  
E estava o sábado o ganho  
no meu pai deu o fagete.  
E eu falei para meu pai que o  
ficheio ele é o fagete e ele é  
de.  
E com o o pariser mais  
antigo e com o signonzo e  
catao o meu pai e co-  
tam em.

Figura 15 – Fonte: Banco de dados de escrita do EF II: Z08\_5A\_20M\_05.



No dia 11 de agosto de 2006 quando eu tinha  
10 anos na quarta série eu conheci uma garota  
chamada Pamela da eu bonita, morena, ela cantava  
cabelo lizo e uma voz doce. Mas no dia seguinte quando  
eu ia pra escola em nome de um dos meus colegas chamada  
Tham me chamou perguntando se eu iria na casa  
dele junto com Josimara quando acabamos a com a  
conversa se quis o fagete e se eu queria o fagete.  
Na saída tentei encontrar a Pamela mas não  
a encontrei; então eu fui para a casa de minha avó  
acompanhada de Josimara e fui para a casa de meus colegas.  
Chegando lá a gente fez os trabalhos de casa,  
comemos um sanduíche e fomos jogar video game,  
naquele tempo o play station 2 era muito raro e  
meus colegas tinham o play station 1; então eles  
queriam negociar; então eu falei que eu seria  
vendo por R\$850,00 reais e se fala sem  
seguidor só por curiosidade mas o todo mundo  
falou que gostavam de uma garota perguntar o nome  
e aí veio uma criança e ela falou Pamela a  
garota dos meus sonhos. Mas depois de um tempo a  
gente esqueceu o negócio do video game e fizemos um  
trato de quem iria ficar com a garota por um mês  
ficou brigado com o outro. No dia seguinte eu fui com  
para a casa de Pamela e a encontrei em nome de uma  
não durou muito tempo ela mudou de escola e nunca

Figura 16 – Fonte: Banco de dados de escrita do EF II: Z10\_7B\_26M\_05.

Passamos a comentar dados que, no interior dos tipos de estruturas rítmicas, chamaram a nossa atenção para questões como a formação de palavras com afixos acentuados.<sup>9</sup> Para interpretar essas ocorrências, tomamos por base Schwindt (2001) e Quadros e Schwindt (2008). Segundo esses autores: palavras derivadas de afixos acentuados apresentam dois acentos, um localizado na base, o outro no afixo (por exemplo: **loucamente**); afixos dissílabos, por formarem pés métricos isoladamente, são candidatos a receber acento primário característico de sílaba (por exemplo: **sozinho** – forma-se um pé binário do tipo troqueu, nesse caso); 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 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 desses afixos, pois, em dados como “so zinho” e “loca mente”, os sufixos configuram-se como pés métricos troqueus, formando uma fronteira acentual entre sufixo e base; já constatamos, por meio de outros dados (que não envolveram necessariamente afixos), que ela constitui uma informação relevante na tentativa de registro dos limites gráficos de palavras. Já na ocorrência “mini cidade”, parece haver uma independência morfossintática (e, talvez, morfossemântica) do prefixo em relação à sua base.<sup>10</sup>

9 Os dados de que constam afixos acentuados são: “anti penúltimo” (antepe-núltimo), “so zinho” (sozinho), “mini cidade” (minicidade), “contra bando” (contrabando) e “loca mente” (loucamente).

10 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”.



## Características particulares das hipersegmentações

Na parte anterior da nossa análise, privilegiamos a discussão das ocorrências de hipersegmentação que, de algum modo, permitiram o apontamento de regularidades linguísticas. Agora, o nosso intuito é realizar uma interpretação das ocorrências que se diferenciam qualitativamente da maioria das hipersegmentações, diferenças que se mostram em dois planos: no da estrutura do dado e no da motivação subjacente.

Ao realizar a análise das marcas “particulares”<sup>11</sup> de hipersegmentação, não deixaremos de abordar questões linguísticas já antes discutidas, uma vez que, mesmo a partir de ocorrências em boa parte diferentes das tendências encontradas, é possível, em alguma medida, estabelecer vínculos com aspectos mais gerais dos dados analisados, os quais permitem tratar de generalizações. Hipóteses mais gerais e particulares dialogaram entre si, visto não serem excludentes, na medida em que se constituem em hipóteses sobre uma só língua.

A identificação das grafias “particulares” restringiu-se às produções escritas de dois escreventes do *corpus*. Esse resultado parece-nos ainda mais interessante se considerarmos que as hipersegmentações foram produzidas por dois escreventes que, também em termos quantitativos, diferenciam-se bastante dos outros alunos. Esses escreventes são Bruno e Pedro, os quais, juntos, somam 59 dados, representativos de 53% de todas as grafias não convencionais identificadas. No entanto, para a nossa discussão, selecionamos sete ocorrências de hipersegmentação e os dois dados híbridos.

Os aspectos que particularizam esses dados merecem comentário: embora se trate de grafias produzidas pelos mesmos escreventes, ocorreram uma única vez, mesmo quando, no interior de um

---

11 O termo “particular” busca registrar a diferença de alguns dados não apenas em relação ao que identificamos como “geral” do *corpus*, mas, em especial, pelo fato de os interpretarmos como pistas das diferentes maneiras como os escreventes são capturados pelo sistema simbólico da escrita (Lemos, 1998).

mesmo texto, tenham sido registradas em mais de um momento; tal fato reforça a suposição de que essas marcas são fruto de hipóteses episódicas (Abaurre, 1988a), empregadas no texto não só para atender a uma exigência ortográfica de segmentação em palavras, mas também por revelarem 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; destaque nosso). Também em função desse aspecto, não é apenas o dado (enquanto estrutura) que, determinadas vezes, interessa observar, mas o modo como a grafia apresenta-se no funcionamento do enunciado em que ocorre.

No Quadro 13 destacamos as ocorrências produzidas por Bruno e Pedro. Os dados híbridos e seis, dentre todas as hipersegmentações, foram identificados nos textos de Bruno; apenas um dado apareceu na escrita de Pedro.

Quadro 13 – Grafias “particulares” (hipersegmentação e híbrido) produzidas por Bruno e Pedro

Bruno			
6º ano	7º ano	8º ano	9º ano
a i (ai) adisi o na (adicionar) es quitto (escrito) com panhate (acompanhante)		come cheia (comecei a) a olado (ao lado)	com ver sando (conversando) com versando com ver samos (conversamos)
Pedro			
6º ano	7º ano	8º ano	9º ano
	foute i (voltei)		

É possível perceber as diferenças dessas formas na maior parte das hipersegmentações analisadas e, sobretudo, no modo de escrita das palavras da língua, considerando o que se espera de escreventes que concluíram pelo menos quatro anos de escolarização formal.<sup>12</sup> Dessas diferenças, gostaríamos de destacar:

12 Lembre-se que os escreventes analisados ingressaram no Ensino Fundamental quando ainda era constituído por quatro anos em sua etapa inicial.

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

No entanto, em certos pontos das rupturas apresentadas (Chacon, 2005), é possível recuperar aspectos de práticas orais/faladas e letradas/escritas que caracterizam o conjunto geral das hipersegmentações do *corpus*. Referimo-nos 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, a ocorrência “es quito”, produzida por Bruno. No que se refere à estrutura prosódica da palavra escrita convencional (isto é, “escrito”), notamos que é idêntica àquela a partir da qual foi realizada grande parte das hipersegmentações identificadas, 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 sílaba seguida de um pé métrico 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 Fundamental II que apresentam a mesma estrutura prosódica, esse dado não foge às características do português brasileiro, em relação ao tipo de estrutura preferencial de palavras da língua: as que apresentam acento à esquerda (ver Bisol, 1996).

No entanto, conforme concluímos com base nos resultados anteriormente apresentados e também naqueles de Tenani (2011a) e Paranhos e Tenani (2011), a categoria gramatical à 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; Tenani, 2011, p.496). Realizamos um estudo comparativo entre dados de segmentação não convencional encontrados em textos de alunos de 5ª e 9ª anos do Fundamental, com o intuito de identificar possíveis semelhanças (em busca de aproximações) e diferenças (em busca de caracterizações) nesses dois importantes anos dessa etapa escolar, um (o 5º ano) correspondente ao encerramento do período da alfabetização e o outro (9º ano), à etapa final do Fundamental (Silva, L., 2011). Deixando de lado as semelhanças e abordando as diferenças observadas na configuração das hipersegmentações, constatamos que as sílabas pretônicas separadas entre brancos pelos escreventes do 5º ano não apresentaram, necessariamente, relação com categorias gramaticais da língua, enquanto, com relação ao 9º ano, sempre se relacionaram a algum item gramatical, como podemos observar no Quadro 14.

Quadro 14 – Exemplos de hipersegmentações frequentes no Ensino Fundamental I e II (conforme Silva, L., 2011, p.41-2)

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

Comparando, por exemplo, o dado “em vergonhada”, presente no quadro, com a ocorrência “es quito”, notamos claramente, na hipersegmentação do Fundamental II, uma relação da sílaba da grafia não convencional com a preposição “em”. Podemos entender que a relação entre essas duas unidades linguísticas (sílabas pretônicas e item gramatical), tanto em termos sonoros quanto gráficos, permitiu ao escrevente a identificação metalinguística das possibilidades de sentido (Tenani, 2004) que o “em”, grafado entre limites, desempenha no funcionamento de determinados enunciados escritos. Essa mesma interpretação não se estende ao “es” em “es quito”. Desse modo,

evidenciamos a particularidade desse dado em relação às outras ocorrências do *corpus* que apresentaram a mesma estrutura de palavra.

Se, como notamos (Silva, L., 2011),<sup>13</sup> na primeira etapa do Fundamental as hipersegmentações ainda estão de modo mais explícito ancoradas em aspectos fonético-fonológicos da unidade palavra – mas não exclusivamente, segundo Chacon (2005) –, podemos considerar que Bruno continua a mobilizar, na sua escrita, aspectos prosódicos que, provavelmente, lhe são salientes desde os primeiros anos de escolarização. Mas, sem negar essa hipótese, podemos também interpretar a grafia hipersegmentada de “es” como resultado de uma possível correlação dessa sílaba com o prefixo “ex”, o qual apresenta *status* 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 sinalizado, 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; 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 Fundamental II.

Consideramos que a troca de <f> por <v> não tenha relação com o aspecto de categorização gráfica (Massini-Cagliari; Cagliari, 2004) dessas letras, decorrente de uma dificuldade do escrevente em reconhecer a forma de grafar as duas letras, mas sim com um aspecto mais fonético-fonológico que caracteriza, ao mesmo tempo, o [f] e o [v]. A esse respeito, no sistema fonológico do português brasileiro, essas consoantes são classificadas, quanto ao modo e ponto de articulação e ao grau de vozeamento, como: [f]: fricativa labiodental desvozeada; [v]: fricativa labiodental vozeada. Desse modo, [f] e [v] só se diferenciam em relação ao vozeamento, fato que põe em evi-

13 Os trabalhos de Paula (2007) e Cunha (2004), por exemplo, que também analisaram a escrita inicial de crianças, identificaram hipersegmentações muito semelhantes às que levantamos em nossa pesquisa de iniciação científica.

dência a proximidade sonora entre os sons representados por essas letras e, em consequência, a complexidade encontrada pelos aprendizes da escrita em registrar “os sons representados por algumas letras [que] são sonoramente semelhantes” (Simioni; Raupp, 2012).

O segundo aspecto refere-se, especificamente, à segmentação não convencional proposta entre a pseudopalavra (Cunha, 2004)<sup>14</sup> “foute” e o grafema <i>, parte da sílaba final da palavra. Um fato importante a ressaltar é que, em comparação com a ocorrência “esquito”, essa hipersegmentação é ainda mais rara em relação aos dados encontrados no Fundamental II e aproxima-se muito 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 dessa natureza só ocorreram na 1ª série/2º ano, ou seja, em período inicial de aquisição da escrita. No entanto, a autora salienta que, nas segmentações em que os limites silábicos não são preservados, está subjacente “a atuação prioritária da escrita institucionalizada que determina a distribuição feita pela criança” (p. 136), 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 “ai” (a i), em que o escrevente pode 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 <i>.

A atuação prioritária da escrita institucionalizada, como sugere Capristano (2007b, p. 136), parece rondar outras grafias de Bruno, como “com panhate” e as variações “com ver sando”, “com ver samos” e “com versando”. Em “com panhate” (acompanhante), a supressão do “a” pode ter ocorrido justamente em função da

14 Por “pseudopalavra” Cunha (2004) entende os resultados das hipersegmentações que não revelam 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 pseudopalavras, nos termos da autora, são as partes destacadas nos dados selecionados do *corpus* analisado: “na **quela**”, “com ver **sando**”, “**adisio** o na”.

identificação do “com” em relação à 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”, usadas por esse escrevente. Soma-se a essa observação o fato de o verbo “conversar” admitir como regência a preposição “com”, o que torna mais saliente para o escrevente a grafia de “com” como sendo uma palavra. Em alguma medida, isso aparece ilustrado no texto de Bruno, no qual as hipersegmentações de “conversando” ocorrem, por exemplo, em: “eu fico **com** ver sando **coms** meus amigos” (ver Figura 17 a seguir, linhas 6-7).

Ainda sobre as três grafias não convencionais (“com ver sando”, “com ver samos” e “com versando”), podemos interpretar que a porção hipersegmentada “ver”, em dois dados, pode ter sido reconhecida pelo escrevente como a forma infinitiva do verbo “ver”. Desse modo, ao inserir os espaços em branco delimitadores de palavra, o escrevente o faz entre os limites silábicos possíveis de projeção acerca de palavras da língua (“com” e “ver”, nos exemplos), o que explica o fato de não segmentar as sílabas “san”, “sa” e “mos”, já que estas, 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 sílaba ( $\sigma$ ) e pé métrico ( $\Sigma$ ), como visto na tendência dos dados já discutidos:

- (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 também se considera 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 sen-

tido que se constroem no texto, que também estão permeadas pela organização prosódica dada ao texto” (Tenani, 2008, p.240). Em alguma medida, percebemos que esse fato foi mobilizado por Bruno nas grafias que propôs para as diferentes formas do verbo “conversar”, em que recuperamos a produção escrita<sup>15</sup> em que ocorreram essas grafias, apresentada na Figura 17.

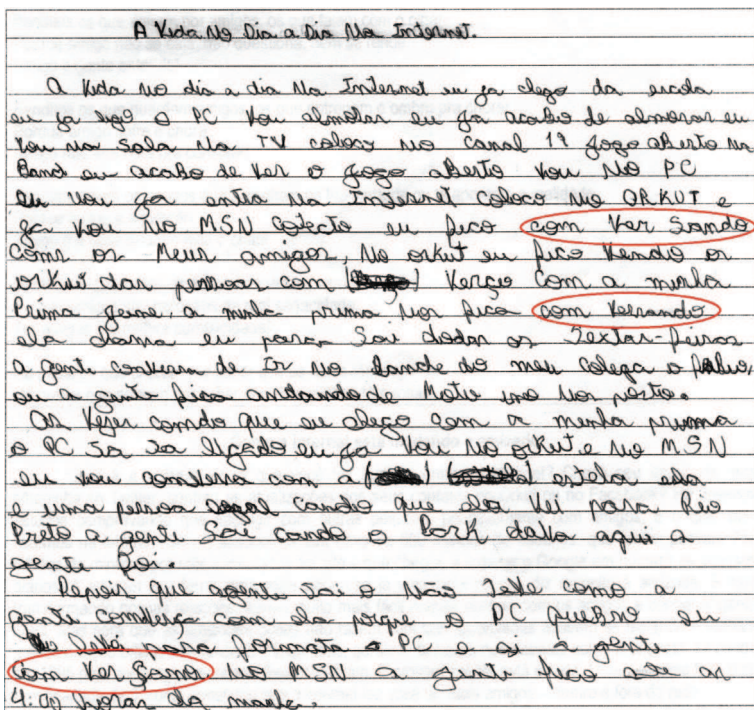


Figura 17 – Fonte: Banco de dados de escrita do EF II: Z11\_8B\_13M\_01

- 15 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 para o 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



A proposta de redação solicitava ao escrevente a produção de um texto narrativo em que relatasse as suas relações de amizade e se faria uso (ou não) da internet para mantê-las. Vejamos 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 houve algum bom momento que vocês viveram juntos(as)”.

Nas hipersegmentações a partir da proposta temática, verificamos que a primeira segmentação não convencional da palavra “conversando” se deu quando o escrevente introduziu o que intitulou como o seu “dia a dia na internet” (linha 6 do texto, seguida do verbo “conecto”). A estratégia do escrevente, ao se “conectar”, é conversar com os seus amigos virtuais. Nesse aspecto, é interessante a ocorrência de diferentes formas de segmentação desse verbo: convencionais<sup>16</sup> (“conversa”, “comversa”, “comverça”), não convencionais (“com ver sando”, “com versando”, “com ver samos”), rasura (“com verço”).

A variedade de grafias da palavra “conversar” levou-nos a investigar por que o escrevente as produziu. No enunciado da proposta de texto, a qual abre possibilidades de construção da narrativa, mencionam-se os tipos de atividades que amigos realizam, como passear, conversar e estudar. O escrevente parece ter elegido a atividade de conversar como o eixo central da construção da sua narrativa, a partir da qual ele interage com os amigos fora da rede. A marcação não convencional das palavras, nesse sentido, é utilizada pelo escrevente como um recurso para demonstrar ao interlocutor os sentidos privilegiados em seu texto. A forma como ocorrem essas segmentações não é aleatória, evidenciando o conhecimento meta-

---

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.

16 Convencional do ponto de vista do limite gráfico.

linguístico do escrevente para identificar palavras escritas no interior de uma única sequência ao mesmo tempo fônica e ortográfica.

Bruno parece ter seguido nesse mesmo rumo ao usar outras segmentações não convencionais, como “adisio o na”. A Figura 18 mostra a produção textual<sup>17</sup> em que foi identificada.

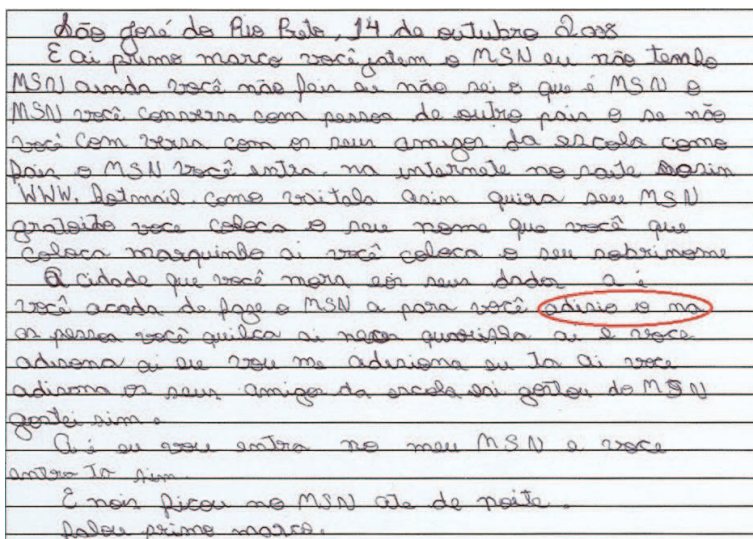


Figura 18 – Fonte: Banco de dados de escrita do EF II: Z08\_5C\_09M\_04.

Bruno, ao escrever uma carta para seu primo Marco, objetivou ensiná-lo a se comunicar pela internet por meio do MSN. Chamamos a atenção para o fato de que a forma hipersegmentada em

17 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](http://www.hotmail.com), vai estar lá assim: “crie seu MSN gratuito”, você coloca o nome que você quiser. Coloque Marquinho. Ai você coloca o seu sobrenome, a cidade que você mora e os seus dados. Ai você acaba de fazer o MSN! Ah, para você adicionar as pessoas, você clica aí nesse quadrinho. Ai você me adiciona, aí eu vou me adicionar. Ai você adiciona os seus amigos da escola. E aí, gostou do MSN? Gostei, sim! Ai eu vou entrar no meu MSN e você entra, tá! Sim. E nós ficamos no MSN até de noite. Falô, primo Marco.

discussão introduz uma prática característica do uso do MSN: adicionar os amigos à lista de contatos. Nota-se, ainda, que a informação sobre a adição dos amigos é a última instrução dada por Bruno ao primo sobre como criar uma conta no MSN. Essa informação parece quase ter sido esquecida pelo escrevente, pois Bruno iniciou a instrução 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”.

Inferimos uma possível ênfase nessa informação que recai sobre o verbo “adicionar”, fundamental para a comunicação pelo MSN. Somam-se, nesse aspecto, sentidos privilegiados 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 “adisio n a”, o escrevente parece ter tentado plasmar um possível ritmo silabado para dar ênfase à palavra, e a duplicação e separação do “o” se dá entre as sílabas pretônicas “adisio”, “o” e a sílaba tônica “nar” (grafada sem a marca de infinitivo do verbo), um possível indício de representação de um ritmo que desacelera e culmina na sílaba tônica da palavra enfatizada.

Com base na análise dos dois textos, compreendemos, com Tenani (2008), que, em uma tentativa de atender ao que lhe é solicitado, o escrevente baseia-se em certas estruturas e lança mão de registros alternativos de enunciados do seu texto, como no caso do modo de segmentação das palavras, a qual ganha uma dimensão textual na produção desse escrevente, visto que a maneira de propor limites às palavras não ocorre de qualquer modo ao se considerar a construção de todo o texto.

Finalizamos a nossa discussão abordando os dados de híbridos: “come cheia” e “a olado”, que aparecem no texto<sup>18</sup> da Figura 19.

---

18 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! Aí, 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,



Ao observar a própria distribuição gráfica dos espaços em branco, encontramos, ao longo do texto, pistas que nos levaram a classificar essa ocorrência como um dado híbrido. É o caso de outras duas grafias da sequência “comecei a”: ~~comecei a~~ (linha 13), ~~comecei a~~ (linha 17), escritas de modo convencional no que se refere ao espaço em branco.

Mesmo tratando-se de uma ocorrência particularizada, em relação às demais, 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 que esta última seja diferente do que se poderia esperar de uma hiposegmentação “regular” da palavra “comecei”, uma vez que o mais comum seria ele atribuir autonomia à última sílaba da palavra, relacionando-a à palavra escrita “sei” e não à palavra “cheia”. Contudo, em termos prosódicos, essas sequências são palavras fonológicas, pois comportam dois acentos primários.

Estudos sobre hiposegmentações têm apontado que as junções entre um elemento clítico e uma palavra em geral são realizadas à direita (como em “aboneca”, para “a boneca”), isto é, os clíticos são hiposegmentados, preferencialmente, à palavra que vem depois deles. 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, o verbo “começar”. No entanto, o dado “come cheia” foge à tendência geral relatada na literatura sobre dados do Ensino Fundamental I e também em relação às características apontadas na seção anterior. A partir de indícios da organização do texto, formulamos as seguintes hipóteses para essa ocorrência:

- A primeira hipótese baseia-se na flutuação entre <ch> e <c> para grafar o som [s]. Inicialmente cabe observar que o dígrafo <ch> representa, na ortografia do português brasileiro, o som [ʃ]; a letra <c>, o som [s], no contexto em análise. Na classificação das consoantes, quanto ao grau de vozeamento, ao ponto

e ao modo de articulação, [ʃ] e [s] diferenciam-se apenas pelo ponto de articulação (alveolar e alveopalatal, respectivamente). Essa diferença 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 a possibilidade de a dimensão gráfico-visual também permear a troca entre a letra <c> e o dígrafo <ch>, pois é comum 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 <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, ao ser acrescida da letra “a”, formaria uma palavra escrita convencional: “cheia”.

- A outra hipótese diz respeito à sequência seguinte após a ocorrência do dado híbrido: a referida sequência consiste também em uma segmentação não convencional, mais especificamente, uma hipossegmentação: “jogala”. Pela leitura do texto, é possível depreender 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” foi interpretada como o possível clítico “la”, muito frequente em finais de verbos. Uma vez que o elemento “la” formou 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 dela. 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 o 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, nessa não encontramos nenhuma grafia convencional que nos permitisse contrastar com a não convencional. Nesse caso, a identificação da ocorrência como um dado híbrido deu-se somente pela observação do espaço em branco entre as letras “a” e “o”, que aparece claramente maior.

Por meio do enunciado em que ocorre o dado, ou seja: “eu jogava **a olado** de Neymar, Robinho e P. H. Ganso”, entendemos que a grafia não convencional pode ter sido motivada por uma forma de atribuição do estatuto de palavra escrita ao “a”, separando-o, por essa razão, do elemento “o” que a acompanha e unindo-o, por sua vez, à próxima palavra do enunciado. Some-se a isso o fato de que o uso do “ao”, tanto em enunciados falados como escritos, é pouco frequente, o que dificulta a interpretação gráfica dessa sequência como uma palavra escrita. Uma segunda motivação pode residir no destaque, pelo uso não convencional do espaço em branco, na informação de que o escrevente trabalhou com jogadores importantes do futebol brasileiro. Assim, na parte hipossegmentada “olado”, o clítico “o”, ao funcionar 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 destaca: estar “ao lado” dos seus ídolos na mesma condição deles, e não apenas como um admirador.

Gostaríamos de destacar, por fim, que, embora dados de híbrido tenham ocorrido uma só vez e na produção de um só aluno, constituindo uma característica dos dados de segmentação não convencional produzidos nos anos mais avançados de escolarização, eles diferenciam-se em relação aos tipos de híbridos encontrados no Fundamental I, como é possível observar em Cunha (2004). Esta autora, apesar de também não ter identificado muitos dados desse tipo, percebeu que as crianças primeiro hipossegmentavam duas palavras para depois hipersegmentarem parte da segunda pa-



lavra, como em: “pofa vor” (para “por favor”) (ibid., p.112). Como vimos nos dados do *corpus*: “come cheia” e “a olado”, o escrevente primeiro hipersegmentou para depois hipossegmentar uma dada sequência de palavras. A comparação entre esses dois tipos de híbridos mostra, mais uma vez, particularidades do Ensino Fundamental II em relação ao Ensino Fundamental I, o que reafirma a importância deste estudo.

## Uma síntese

Este capítulo foi dedicado à exposição dos 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 Ensino Fundamental. Ao analisarmos os dados, notamos que: houve diminuição das ocorrências de hipersegmentação ao longo dos anos escolares; grande parte dos escreventes investigados terminou o Ensino Fundamental II atribuindo às palavras fronteiras gráficas além do previsto 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 texto; as palavras que continuaram sendo hipersegmentadas foram aquelas em que foi possível relacionar a sílaba pretônica com possíveis classes gramaticais (por exemplo, preposição, conjunção, pronome); 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.

Com referência 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, pois a sua ocorrência passou 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 Fundamental 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 também foi observada na grafia hipersegmentada produzida por Pedro.

Embora as hipersegmentações produzidas por Bruno e Pedro tenham chamado a nossa atenção para a quantidade e a 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. Em alguma medida, podemos supor que os alunos do Fundamental II partilham hipóteses linguísticas muito próximas acerca da noção de palavra escrita. As particularidades identificadas na produção de Bruno e Pedro evidenciam 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, pode ser que o que consideramos “particular” seja, ao fim, uma marca da história de linguagem que constitui individualmente cada um dos escreventes.<sup>19</sup>

---

19 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.

## ÚLTIMAS CONSIDERAÇÕES

Neste livro, buscamos compreender, por meio das grafias de hipersegmentação, a maneira como alunos regularmente matriculados na etapa final do Ensino Fundamental lidam com a noção de palavra escrita. Pelo fato de as hipersegmentações de palavras que ocorrem em produções escritas de escreventes com mais anos de escolarização serem um tipo de dado pouco investigado, objetivamos, por um lado, a realização de uma descrição mais geral das características linguísticas das hipersegmentações presentes em textos do Fundamental II, e, por outro, a análise de ocorrências que, em grande parte das vezes, foram únicas em todo o *corpus* e não constituíram regularidades.

Ao observar a distribuição das grafias não convencionais, verificamos que a ocorrência de hipersegmentações diminuiu ao longo dos anos escolares do Fundamental II. Esse fato ficou mais evidente quando analisamos não só os dados de hipersegmentação, mas também as palavras escritas que propiciaram as fronteiras gráficas não convencionais. Notamos que os escreventes tiveram dúvidas quanto aos limites das mesmas palavras e, em decorrência, foi maior o número de dados entre um ano escolar e outro. Não ocorreu somente o decréscimo do número de dados, mas também dos tipos de palavras que geraram dúvidas quanto aos seus limites.

Com relação à trajetória longitudinal de cada aluno, verificamos que boa parte deles iniciou e terminou o Fundamental II produzindo hipersegmentações. No entanto, as ocorrências tornaram-se menos frequentes ao longo dos anos, quando cotejada a quantidade de dados em relação à extensão dos textos produzidos (estabelecida pelo número de palavras escritas) e às dificuldades em grafar tipos específicos de palavras (como “demais”) ou tipos de palavras que apresentaram sílabas correspondentes (como “amigo”, “apareceu”, “apaixonado”).

Buscamos compreender também os critérios linguísticos subjacentes às grafias de hipersegmentação. Ao organizar os dados a partir da estrutura prosódica das palavras convencionais, identificamos pistas sobre informações prosódicas e (orto)gráficas. Com relação ao aspecto prosódico, as fronteiras gráficas não convencionais puderam ser relacionadas a fronteiras de constituintes prosódicos, com destaque para a sílaba e o 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, por exemplo, em “em **bora**” (embora). Com efeito, essas pistas prosódicas, inferidas por meio das hipersegmentações, permitiram 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” (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 coloca em evidência um imaginário que “põe às claras a falsa pureza da escrita” (ibid., p.83).

Vimos nas hipersegmentações o atravessamento de informações linguísticas (orto)gráficas, relacionadas à inserção dos escreventes em práticas letradas/escritas, mais especificamente, naqueles momentos em que essa inserção revela-se na imagem do código escrito institucionalizado (ibid.). Para ilustrar esse aspecto, destacamos

grafias como “com migo” e “em tão”, a partir das quais inferimos, com base em uma sequência ao mesmo tempo fônica e gráfica, palavras escritas relacionadas a generalizações de regras ortográficas que, ao buscarem organizar e distribuir os grafemas nas palavras escritas, dão pistas dos limites delas. Outras hipersegmentações, como aquelas derivadas da presença não convencional do hífen, forneceram pistas sobre 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 (“ganha-se”).

Foi possível observar que, em todas as situações descritas, informações relativas à própria escrita que estiveram subjacentes às hipóteses acerca dos limites das 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 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).

No caso das hipersegmentações que não se encaixaram na organização geral dos dados, elas foram encontradas apenas na produção dos escreventes Bruno e Pedro, o que indica 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 da interpretação desses dados, semelhanças com processos de segmentação de 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 perceber 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 se construiu o texto. Nesse sentido, concordamos com Abaurre (1988a) acerca da não categoricidade das hipóteses que os aprendizes da escrita formulam e, assim, não é estranho que, muitas vezes, tenhamos encontrado “soluções escritas diferentes para um mesmo problema de segmentação em um mesmo texto” (Abaurre, 1991c, p.205).

Com base nos resultados obtidos, 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 de alfabetização. Entendemos, ainda, que reconhecer os limites de palavras explique que o escrevente da língua lide 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 atender às exigências da ortografia, valorizadas sobretudo socialmente.

Neste momento final, vislumbramos a relevância dos resultados que obtivemos para a compreensão dos processos de segmentação não convencional de palavras que ainda ocorrem no período do Ensino Fundamental II. Por meio deles, podem ser traçados paralelos que permitem observar semelhanças e diferenças no modo de delimitação das 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 os resultados que reunimos e os dados de escrita infantil, por exemplo, devem levar em conta sobretudo a natureza linguística deles, a qual guarda relação com diferentes formas de mobilizar o relacionamento entre o oral/falado e o letrado/escrito que os escreventes capturam em momentos da sua escolarização, revelando as noções que possuem sobre a constituição da escrita, a qual consideramos, como Corrêa (2004), heterogênea.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ABAURRE, M. B. M. O que revelam os textos espontâneos sobre a representação que faz a criança do objeto escrito? In: KATO, M. (Org.). *A concepção da escrita pela criança*. 2.ed. Campinas: Pontes Editores, 1988a. p.135-42.
- \_\_\_\_\_. A relevância dos critérios prosódicos e semânticos na elaboração de hipóteses sobre segmentação na escrita inicial. *Boletim da Abralin*, Campinas, v.11, p.203-17, 1991c.
- \_\_\_\_\_. (Re)escrevendo: o que muda? In: ABAURRE, M. B. M.; FIAD, R. S.; MAYRINK-SABINSON, M. L. *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, R. (Org.). *Alfabetização e letramento: perspectivas linguísticas*. São Paulo: Mercado das Letras, 1998. p.205-32.
- \_\_\_\_\_; CAGLIARI, L. C. 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, p.25-9, 1985.
- \_\_\_\_\_; GALVES, C. Os clíticos no português brasileiro: uma abordagem sintático-fonológica. In: CASTILHO, A. T.; BASÍLIO, M. (Orgs.). *Gramática do português falado: estudos descritivos*. Campinas: Unicamp, 1996, v.IV, p.273-320.

- \_\_\_\_\_.; FIAD, R. S.; MAYRINK-SABINSON, M. L. 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, L. Constituintes prosódicos. In: *Introdução a estudos de fonologia do português brasileiro*. Porto Alegre: EDIPUC-RS, 1996, p.247-61.
- \_\_\_\_\_. Mattoso Câmara Jr. e a palavra prosódica. *Delta*, São Paulo, v.20, n.esp., p.59-70, 2004.
- \_\_\_\_\_. 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, p.163-84, 2005.
- BOOP DA SILVA, T. *Formação de palavras compostas em português brasileiro: uma análise de interfaces*. Porto Alegre, 2010, 195f. Tese (doutorado em Letras) – Universidade Federal do Rio Grande do Sul.
- BRASIL. Secretaria de Educação Fundamental. *Parâmetros Curriculares Nacionais: língua portuguesa*. Brasília: 1997.
- BRISOLARA, L. B. *Os clíticos pronominais do português brasileiro e sua prosodização*. Porto Alegre, 2008, 189f. Tese (doutorado em Linguística Aplicada) – PUC-RS.
- CAMARA JR., J. M. *Estrutura da língua portuguesa*. Petrópolis: Vozes, 1970.
- CAPRISTANO, C. C. *Aspectos de segmentação na escrita infantil*. São José do Rio Preto, 2003. 213f. Dissertação (mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista.
- \_\_\_\_\_. A propósito da escrita infantil: uma reflexão sobre segmentações não convencionais. *Letras de Hoje*, Porto Alegre, v.39, n.3, p.245-60, 2004.
- \_\_\_\_\_. *Mudanças na trajetória da criança em direção à palavra escrita*. Campinas, 2007b. 253f. Tese (doutorado em Linguística Aplicada) – Universidade Estadual de Campinas.
- \_\_\_\_\_. Por uma concepção heterogênea da escrita que se produz e se ensina na escola. *Cadernos de Educação (UFPel)*, Pelotas, v.35, p.171-93, 2010.
- CHACON, L. *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, n.3, p.223-32, 2004.

- \_\_\_\_\_. Hipersegmentações na escrita infantil: entrelaçamentos de práticas de oralidade e letramento. *Estudos Linguísticos*, Campinas, v.34, p.77-86, 2005.
- \_\_\_\_\_. Prosódia e letramento em hipersegmentações: reflexões sobre a aquisição da noção de palavra. In: CORRÊA, M. L. G. (Org.). *Ensino de língua: representação e letramento*. Campinas: Mercado de Letras, 2006. p.155-67.
- COLLISCHONN, G. Acento secundário em português. *Letras de Hoje*, Porto Alegre, v.29, n.4, p.43-53, 1994.
- CORRÊA, M. L. G. *O modo heterogêneo de constituição da escrita*. Campinas, 1997. 435f. Tese (doutorado em Linguística) – Universidade Estadual de Campinas.
- \_\_\_\_\_. Letramento e heterogeneidade da escrita no ensino de português. In: SIGNORINI, I. (Org.). *Investigando a relação oral/escrito e as teorias do letramento*. Campinas: Mercado das Letras, 2001. p.135-66.
- \_\_\_\_\_. *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, p.269-86, 2007.
- \_\_\_\_\_. Heterogeneidade da escrita no ensino: das modalidades às relações intergenéricas. In: PEREIRA, L. A.; CARDOSO, I. (Orgs.). *Reflexão sobre a escrita: ensino de diferentes gêneros de textos*. 1.ed. Aveiro (PT): Editora da Universidade de Aveiro, 2009, v.1.
- COX, M. I. P.; ASSIS-PETERSON, A. A. A palavra: uma história de dissonâncias entre professores e aprendizes de escrita. In: *Cenas de sala de aula*. Campinas: Mercado das Letras, 2001. p.51-79.
- CUNHA, A. P. N. *A hipo e a hipersegmentação nos dados de aquisição da escrita: um estudo sobre a influência da prosódia*. Pelotas, 2004. 132f. Dissertação (mestrado em Educação) – Universidade Federal de Pelotas.
- \_\_\_\_\_. *As segmentações não convencionais da escrita inicial: uma discussão sobre o ritmo linguístico do português brasileiro e europeu*. Pelotas, 2010. 188f. Tese (doutorado em Educação) – Universidade Federal de Pelotas.
- \_\_\_\_\_; MIRANDA, A. R. M. Indícios de reestruturação do conhecimento fonológico da criança em dados de reparo na escrita inicial. *Letras de Hoje*, Porto Alegre, v.48, n.3, p.343-54, 2013.



- FERREIRA, C. R. G. *Um estudo sobre a segmentação não convencional na aquisição da escrita de alunos de EJA*. Pelotas, 2011. 163f. Dissertação (mestrado em Educação) – Universidade Federal de Pelotas.
- GOODY, J. *Domesticação do pensamento selvagem*. Trad. Nuno Luis Madureira. Lisboa: Presença, 1988/[1977].
- KLEIMAN, A. Modelos de letramento e as práticas de alfabetização na escola. In: *Os significados do letramento: uma nova perspectiva sobre a prática social da escrita*. Campinas: Mercado de Letras, 1995. p.15-64.
- KOCH, I. Interferências da oralidade na aquisição da escrita. *Trabalhos em Linguística Aplicada*, Campinas, v.30, p.31-8, 1997.
- LE MOS, C. Sobre a aquisição da escrita: algumas questões. In: ROJO, R. (Org.). *Alfabetização e letramento: perspectivas linguísticas*. São Paulo: Mercado de Letras, 1998. p.13-32.
- MARCUSCHI, L. A. Oralidade e letramento. In: *Da fala para a escrita: atividades de retextualização*. São Paulo: Cortez, 2001. p.15-43.
- MASSINI-CAGLIARI, G.; CAGLIARI, L. C. Categorização gráfica e funcional na aquisição da escrita e da leitura em língua materna. *Calidoscópio (Unisinos)*, São Leopoldo, v.2, n.1, p.89-94, 2004.
- MIRANDA, A. R. M. Informação fonológica na aquisição da escrita. In: DEL RÉ, A. et. al. (Org.). *Estudos linguísticos contemporâneos: diferentes olhares*. São Paulo: Cultura Acadêmica, 2013, p.11-36.
- \_\_\_\_\_; MATZENAUER, C. Aquisição da fala e da escrita: relações com a fonologia. *Cadernos de Educação (UFPEL)*, Pelotas, v.35, p.359-405, 2010.
- NESPOR, M.; VOGEL, I. *Prosodic phonology*. Dordrecht: Foris Publications, 1986.
- PAULA, I. F. V. *Movimentos na escrita inicial de crianças: um estudo longitudinal de hipersegmentações*. São José do Rio Preto, 2007. 154f. Dissertação (mestrado em Estudos Linguísticos) – Universidade Estadual Paulista.
- PARANHOS, F. C. *A segmentação não convencional de palavras em textos de alunos de 5ª série do Ensino Fundamental*. Relatório final de pesquisa de iniciação científica. Inédito. Fapesp: Proc. 2009/14491-0, 2010.
- \_\_\_\_\_; TENANI, L. Análise prosódica de segmentações não convencionais de palavras em textos do 6º ano do Ensino Fundamental. *Filologia e Linguística Portuguesa*, v.13, n.2, p.477-504, 2011.
- QUADROS, E. S.; SCHWINDT, L. C. 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.

- REIS, M. C.; TENANI, L. E. *Registros da heterogeneidade da escrita: um olhar para as grafias não convencionais de vogais pretônicas*. São Paulo: Cultura Acadêmica, 2011.
- SCARPA, E. M. Apresentação. In: *Estudo de prosódia*. Campinas: Editora da Unicamp, 1999. p.7-20.
- SCHWINDT, L. C. O prefixo no português brasileiro: análise prosódica e lexical. *Delta*, São Paulo, v.17, n.2, p.175-207, 2001.
- SELKIRK, E. On derived domains in sentence phonology. *Phonology Yearbook*, n.3, 1986.
- SILVA, A. *Alfabetização: a escrita espontânea*. São Paulo: Contexto, 1991.
- SILVA, L. M. *A relação entre dados de segmentação não convencional de palavras em textos de 5ª e 9ª anos do Ensino Fundamental*. Relatório final de pesquisa de iniciação científica. Inédito. Fapesp: Proc. 2010/17720-8, 2011.
- SIMIONI, T. O clítico e seu lugar na estrutura prosódica em português brasileiro. *Alfa*, São Paulo, 52(2), p.431-46, 2008.
- \_\_\_\_\_; RAUPP, J. S. Trocas entre letras que representam consoantes com oposição surda/sonora. *Verba Volant (UFPEL)*, Pelotas, v.3, n.2, p.107-23, 2012.
- STREET, B. *Literacy in theory and practice*. Cambridge: Cambridge University Press, 1984.
- TENANI, L. Rindo das piadas, manipulando a língua. *Alfa*, São Paulo, v.45, p.15-27, 2001.
- \_\_\_\_\_. *Domínios prosódicos no português do Brasil: implicações para a prosódia e para a aplicação de processos fonológicos*. Campinas, 2002. 331f. Tese (doutorado em Linguística) – Universidade Estadual de Campinas.
- \_\_\_\_\_. Segmentações não convencionais e teorias fonológicas. *Letras de Hoje*, Porto Alegre, v.39, n.3, p.233-44, 2004.
- \_\_\_\_\_. Notas sobre a relação entre constituintes prosódicos e a ortografia. *Revista de Estudos Linguísticos*, Belo Horizonte, v.16, n.1, p.231-45, jan./jun. 2008.
- \_\_\_\_\_. *Entre o grupo clítico e a palavra fonológica: os erros de segmentação não convencional de palavras*. Comunicação apresentada no VI Congresso Internacional da Abralín, 2009a.
- \_\_\_\_\_; LONGHIN-THOMAZI, S. Oficinas de leitura, interpretação e produção textual no Ensino Fundamental. *Em Extensão (UFU)*, Uberlândia, v.13, n.1, p.20-34, 2014.

- \_\_\_\_\_. A segmentação não convencional de palavras: evidências de representações de relações entre enunciados falados e escritos. In: II SIMELP: LÍNGUA PORTUGUESA: ULTRAPASSAR FRONTEIRAS, JUNTAS CULTURAS. *Anais...* Évora, 2009b, p.107-27.
- \_\_\_\_\_. *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, L. *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 Abralín*, v.10, n.2, p.91-119, jul./dez. 2011.
- \_\_\_\_\_. Hipersegmentação de palavras: análise de aspectos prosódicos e discursivos. *Linguagem & Ensino*, Pelotas, v.16, n.2, p.305-26, jul./dez. 2013.
- TFOUNI, L. V. *Letramento e alfabetização*. São Paulo: Cortez, 1995.
- VIGÁRIO, M. *The prosodic word in european portuguese*. Berlim/New York: Mouton de Gruyter, 2003.
- ZORZI, J. L. *A apropriação do sistema ortográfico nas quatro primeiras séries do primeiro grau*. Campinas, 1997. 124f. Tese (doutorado em Educação) – Universidade Estadual de Campinas.

## SOBRE AS AUTORAS

**Lilian Maria da Silva.** Graduação em Licenciatura em Pedagogia (2008-2011) pela Unesp (São José do Rio Preto). Mestrado em Estudos Linguísticos (2012-2014) pela mesma universidade. Em 2014, ingresso no curso de doutorado. Pesquisa sobre segmentações não convencionais de palavras, enfatizando a relação entre aspectos prosódicos da língua e convenção ortográfica.

**Luciani Tenani.** Graduação em Linguística pela Unicamp (1990-1993). Mestrado (1994-1996) e doutorado (1996-2002) em Linguística na mesma universidade. Docente, desde 1997, da Unesp e bolsista de produtividade em Pesquisa do CNPq desde 2010. Pesquisa e orienta dissertações e teses sobre os temas: fonologia da língua portuguesa, fonologia e escrita, prosódia e pontuação.

SOBRE O LIVRO

*Formato:* 14 x 21 cm

*Mancha:* 23,7 x 42,5 paicas

*Tipologia:* Horley Old Style 10,5/14

EQUIPE DE REALIZAÇÃO

*Coordenação Geral*

Maria Luiza Favret

