

A grafia de palavras hipersegmentadas em textos produzidos nos anos finais do Ensino Fundamental

(Hypersegmentation of words in the lastest years of Elementary School's texts)

Lilian Maria da Silva¹

Instituto de Biociências, Letras e Ciências Exatas – Universidade Estadual Paulista (Ibilce/Unesp)

msilva.lilian@gmail.com

Abstract: This paper introduces the subject hypersegmentation of words by analysing texts written by students in Elementary School (6th – 9th). Based on the hypersegmentation data, we aim to show the evidence that the acquisition of the written words is complex because it mobilizes, simultaneously, linguistic information from different natures as phonological, morphosyntax/semantic and also literacy. The developed analysis establishes a dialogue with researches which deal with non-conventional segmentation of words, especially with the ones conducted by Capristano's (2003, 2007b), Chacon's (2005, 2006) and Tenani's (2009, 2010, 2011b) and with the idea of heterogeneous writing developed by Corrêa (2004).

Keywords: written word; hypersegmentation; spelling.

Resumo: Este trabalho apresenta uma análise de hipersegmentações de palavras que podem ser encontradas em textos de alunos do Ensino Fundamental II (6º a 9º anos – denominação a partir de 2009). Com base nesses dados de hipersegmentação, busca-se levantar evidências de que a aquisição da noção de palavra escrita é complexa por mobilizar, simultaneamente, informações linguísticas de diferentes naturezas, como fonológica, morfossintática/semântica e, também, letrada. A análise desenvolvida dialoga com pesquisas sobre segmentações não convencionais de palavras, especialmente com as de Capristano (2003, 2007b), Chacon (2005, 2006) e Tenani (2009, 2010, 2011b) e com a concepção de escrita heterogênea desenvolvida por Corrêa (2004).

Palavras-chave: palavra escrita; hipersegmentação; ortografia.

Introdução

Este artigo aborda os resultados parciais de uma pesquisa de mestrado em curso, a qual busca, através de um estudo longitudinal de produções escritas de alunos do Ensino Fundamental II (EF-II, doravante), analisar grafias hipersegmentadas de palavras, com o propósito específico de: (i) realizar uma descrição geral dos dados, que possibilite a identificação de regularidades linguísticas das hipersegmentações; e (ii) tratar das ocorrências que, em relação à descrição geral, não se encaixam nas regularidades encontradas, mas que são relevantes para a compreensão de uma relação mais particular do sujeito com a linguagem em seu modo de enunciação escrito. Neste texto, em específico, apresentamos os resultados já obtidos na descrição dos dados.

Desse modo, quando buscamos relação entre possíveis tendências que caracterizam as grafias hipersegmentadas no EF-II e dados que não seguem essas características, visamos a encontrar, com nossa pesquisa, indícios de hipóteses gerais e singulares a respeito do estatuto de palavra escrita convencional, hipóteses que interpretamos como registros do

¹ Fapesp (Proc. n. 2012/11869-5).

modo como os escreventes deixam na (sua) escrita marcas de seu trânsito pelas práticas de linguagem oral/falada e letrada/escrita em que se inserem (CORRÊA, 2004).

Registros não convencionais das fronteiras de palavras

A hipersegmentação ocorre quando há a presença de um limite gráfico no interior de palavras, em posições que a ortografia não prevê. Exemplos desse fenômeno são “a miga” e “mora-va” (para as palavras escritas “amiga” e “morava”). Em estudos recentes, como os de Capristano (2003), Chacon (2005, 2006), Paula (2007) e também de Tenani (2009, 2010, 2011b), as hipersegmentações de palavras têm sido interpretadas como marcas do *modo heterogêneo de constituição da escrita*. Essa expressão, formulada por Corrêa (2004), foi proposta com o intuito de “questionar a delimitação do campo da escrita apenas pela constatação óbvia de um material específico – o gráfico – que lhe serve como base semiótica” (p. 2). Segundo Corrêa (2004), a escrita é um modo de enunciação heterogêneo que se constitui dialogicamente por meio da relação entre os fatos linguísticos – falado e escrito – e as práticas sociais de linguagem – oralidade e letramento.

Além disso, o autor propõe três lugares privilegiados para a observação da heterogeneidade da escrita, os quais denunciam um imaginário partilhado socialmente a respeito desse modo de enunciação. O primeiro lugar é o eixo de representação da (suposta) gênese da escrita – momentos em que o escrevente acredita ser possível registrar tudo do oral no escrito, igualando os dois modos de enunciação. O segundo lugar é o eixo de representação do código escrito institucionalizado – tentativas de alçar formas escritas que o escrevente supõe ser a mais próxima do que lhe é exigido pelas instituições com as quais dialoga. Finalmente, o terceiro lugar é o eixo de representação da dialogia com o já falado/escrito e ouvido/lido – trata-se do fato de, no texto que se escreve, estarem sempre recuperados aspectos já vividos na linguagem. A identificação desses eixos, durante a análise de textos escritos, é o que permitiria ao pesquisador, ainda que hipoteticamente, reconstruir o processo subjacente à produção escrita.

Com base nesse arcabouço teórico sobre escrita, Capristano (2003), Chacon (2005, 2006), Paula (2007) e Tenani (2009, 2010, 2011b) argumentam que, na estrutura das segmentações não convencionais de palavras, seria possível recuperar fatos relativos às práticas sociais de linguagem orais/faladas e letradas/escritas. A partir desse argumento maior, os pesquisadores demonstram que, em diferentes graus, as hipo e hipersegmentações sempre obedecem aos princípios reguladores “estabelecidos para a definição de constituintes prosódicos, tais como aqueles propostos por Nespor e Vogel (1986)” (CHACON, 2004, p. 223), e a determinadas características possíveis dentro da escrita convencional do português. Porém, como ressalva Chacon (2005), a análise do fenômeno das segmentações de palavras que se distanciam da ortografia permitiria a recuperação de muitos outros aspectos linguísticos-discursivos das práticas orais e letradas. Nesse sentido, quando aqueles pesquisadores privilegiam os aspectos prosódicos e ortográficos não implica que eles, por conseguinte, estejam desconsiderando outros aspectos possíveis. Trata-se, pois, de uma delimitação metodológica que busca observar mais atentamente a atuação da prosódia e da aquisição da convenção ortográfica no funcionamento do texto escrito, por meio das marcas não convencionais das fronteiras de palavra. É nessa linha de raciocínio que desenvolvemos a análise deste trabalho.

Passando aos resultados de dois dos trabalhos citados, os quais refletem mais de perto nossa análise, sintetizamos que: (i) Chacon (2005) analisou hipersegmentações de palavras trissílabas produzidas por crianças e observou, quanto à prosódia, que a maneira como os espaços foram inseridos não convencionalmente obedeceu a princípios recorrentes de estruturação da oralidade, como “é o caso do *ritmo*, que, já estruturado na língua, marca-se na oralidade, por exemplo, por meio de contrastes entre sílabas acentuadas e não-acentuadas” (CHACON, 2005, p. 81, nosso grifo). Segundo o autor, tal característica prosódica foi marcante nos dados analisados, já que grande parte das ocorrências indicou fronteiras de constituintes prosódicos, especialmente da sílaba e do pé métrico. Chacon (2005) destacou, ainda, que as crianças, ao romperem os trissílabos em limites prosódicos, deixam pistas de informações da própria escrita, como é o caso, por exemplo, de uma parte da palavra que foi hipersegmentada ser reconhecida como uma possível palavra escrita independente; isso é o que exemplifica a sílaba à esquerda na hipersegmentação da palavra *porquinho* > *por quinho*. A partir de seus resultados, Chacon (2005) conclui ser impossível dissociar das grafias de hipersegmentação os vínculos com informações orais/faladas e letradas/escritas; (ii) Tenani (2011b) abordou as marcas de segmentação não convencional de palavras produzidas por escreventes do EF-II e argumentou que os dados que são encontrados nos anos que encerram o EF caracterizam-se por colocar em evidência o modo de organização linguística dos elementos átonos. Ou seja, a dúvida que pareceu rondar as hipóteses dos escreventes foi a de quando os elementos inacentuados seriam sílabas (pretônicas ou postônicas) de uma palavra fonológica e quando esses seriam unidades (clíticos) que se comportam junto com uma palavra formando um domínio maior que, na interpretação de Tenani (2011), seria o grupo clítico. Para a autora, a presença de um limite gráfico, no caso das de hipersegmentações, foi uma pista de que uma só sequência fônica foi analisada pelo escrevente como duas unidades lexicais independentes. Em termos prosódicos, uma palavra fonológica foi interpretada como um grupo clítico. Com esse resultado, Tenani (2011b) concluiu que há uma flutuação entre esses dois constituintes prosódicos, em virtude da imprecisão nos estatutos de sílabas átonas e clíticos fonológicos.

Apresentadas essas breves considerações teóricas sobre o objeto de investigação, passamos, a seguir, às informações metodológicas do trabalho.

O material analisado

O cópuz do trabalho é constituído por produções escritas pertencentes a um Banco de Dados inédito de textos do EF-II, o qual está disponível para pesquisas na Unesp, câmpus de São José do Rio Preto. O Banco de Dados, constituído de 2008 a 2011, é resultado do Projeto de Extensão Universitária “Oficinas de Leitura, Interpretação e Produção Textual”, coordenado pelas professoras doutoras Luciani Tenani e Sanderléia Longhin-Thomazi (ambas do Ibilce/Unesp).

Buscando atender aos nossos objetivos, selecionamos, do Banco, uma amostra longitudinal composta por 266 textos produzidos por 12 escreventes ao longo dos quatro anos finais do EF (6º a 9º anos – antigas 5ª a 8ª séries). É importante esclarecer que a opção deste estudo pelos dados de hipersegmentação que são encontrados em textos do EF-II justifica-se em razão de assumirmos com Capristano (2007b) uma perspectiva de

aquisição da escrita que, enquanto processo, não deve ser abordada como um período de aprendizagens lineares e cumulativas, embora possam ser notadas regularidades em seu percurso. A partir dessa abordagem, entendemos que os dados de hipersegmentação que analisamos deixam em evidência a complexidade do registro gráfico da palavra, noção que não se edifica, por exemplo, com o fim de uma etapa de escolarização formal e a hipótese que formulamos é a de que, no momento da escritura, o escrevente está em constante relação com informações linguísticas de naturezas fonológicas, morfossintáticas, morfossemânticas e letradas, por exemplo, que podem convergir em grafias não convencionais, como as hipersegmentações.

Conforme adiantamos na introdução do trabalho, neste artigo trazemos os resultados já alcançados em relação à descrição geral dos dados de hipersegmentação. Para a execução dessa etapa de análise, os dados de hipersegmentação foram reunidos a partir da organização prosódica das palavras convencionais, ou seja, uma vez que, em termos estruturais, a principal característica das hipersegmentações é ocorrer no nível da palavra, consideramos a forma como se organizaram as sílabas dentro das palavras convencionais. Por exemplo, a palavra cujo limite convencional é “embora” tem suas sílabas organizadas metricamente da seguinte maneira: (\bullet * \bullet);² em particular, essa palavra sempre foi hipersegmentada no corpus como “em bora”, a partir da qual houve a reorganização das sílabas da palavra como: (\bullet) σ (* \bullet) Σ , ou seja, como o registro de uma sílaba e um pé métrico isolados, de acordo com as formulações teóricas do modelo de Fonologia Prosódica de Nespor e Vogel (1986). A hipótese que rondou essa proposta de organização dos dados foi a de que informações rítmicas, que se formam no interior da palavra por meio de proeminências prosódicas, podem ser pontos de ancoragem para o tipo de limite gráfico não convencional de palavras na escrita.

Do referido modelo, foi pertinente para a análise os constituintes sílaba, pé métrico e palavra fonológica (daqui em diante: σ , Σ e ω , respectivamente). Com base em Nespor e Vogel (1986), a σ é o menor domínio prosódico, o qual possui um elemento de valor forte (chamado de *cabeça* pelas autoras) e outro(s) de valor fraco. No caso do PB (português brasileiro), o *cabeça* de uma σ é sempre uma vogal e os elementos fracos são as consoantes e/ou glides. O Σ estabelece relação entre sílabas, de modo que, no seu interior, uma será a dominante e as demais dominadas. Essa relação entre as sílabas é fundamental para a marcação do ritmo e a identificação dos acentos (primário e rítmico) das palavras. Por fim, a ω é definida pela união de um ou mais Σ s e é o primeiro constituinte prosódico a estabelecer interação com o componente morfológico da gramática. No entanto, a relação entre a ω e o constituinte morfológico correspondente não é necessariamente isomórfica. Quanto ao elemento mais proeminente, segundo Nespor e Vogel (1986), uma ω porta um único acento primário, o qual recai sobre uma de suas sílabas.

Ainda em relação à descrição geral dos dados, buscamos, também, examinar como se deram as distribuições das ocorrências de hipersegmentação pelos anos letivos, recursos gráficos que indicam palavra (espaço em branco e hífen) e escreventes investigados. No que se refere aos escreventes, informamos que optamos por atribuir-lhes nomes fictícios eleitos de forma aleatória. Com essa decisão, procuramos preservar a identidade dos alunos e, principalmente, garantir, em nosso estudo, a imagem de que os escreventes produtores dos textos se tratam de sujeitos sócio-historicamente constituídos. Assim,

² Os símbolos indicam: (\bullet) sílaba fraca; (*) sílaba forte.

a referência aos escreventes é realizada pelos seguintes pseudônimos: 1-*Ana*; 2- *Julia*; 3-*Bruno*; 4-*Henrique*; 5-*Camila*; 6- *Fernando*; 7-*André*; 8-*Pedro*; 9-*Viviane*; 10-*Lucas*; 11-*Maria*; 12-*Mateus*.³

Os dados de hipersegmentação: uma descrição geral

Iniciamos a análise considerando, na Tabela 1, a demonstração da distribuição do número de textos e dados que foram produzidos em cada ano escolar (segue junto aos valores absolutos uma correspondência percentual).

Tabela 1: Distribuição dos textos e das hipersegmentações nos anos escolares

Ano Escolar	Textos		Hipersegmentações	
	Valor absoluto	Valor percentual	Valor absoluto	Valor percentual
6º	67	25,1%	40	35,3%
7º	74	27,9%	24	21,3%
8º	61	23,0%	27	24,0%
9º	64	24,0%	22	19,4%
Total	266	100%	113	100%

De acordo com os dados da Tabela 1, constatamos uma queda significativa, quanto ao número de hipersegmentações, do sexto ano (40 dados, 35,3%) – início do EF-II – para o nono ano (22 dados, 19,4%) – fim do EF-II. Entretanto, é possível ver que o declínio não se marcou de modo linear, pelos quatro anos, visto um pequeno aumento no valor de dados que pôde ser observado no oitavo ano (27 dados, 24%). Esse resultado é importante, pois corrobora uma de nossas hipóteses de pesquisa a respeito da complexidade da noção de palavra que, além de não ter sua aquisição encerrada junto com o período escolar do EF-I, também não se mostra como tendo um tipo de aprendizado unidirecional, à semelhança do que defende Capristano (2007) sobre o processo de aquisição da escrita. Tal resultado ressalta, também, a necessidade de uma investigação mais atenta sobre esse comportamento peculiar do oitavo ano, já que outros estudos (cf. TENANI, 2010; 2011), sobre esse mesmo objeto de investigação e etapa escolar, chegaram a conclusões parecidas a nossa. Assim, em função dessa aparente sistematicidade de resultados, direcionamos para as próximas etapas da pesquisa que desenvolvemos a busca por explicações que possam justificar porque o aumento dos dados ocorre sempre no oitavo ano e não em outro, levantando, neste primeiro momento, a hipótese de que tal alteração poder ter alguma motivação no tipo de atividade letrada/escrita que se passa a desenvolver na prática escolar no referido ano.

A seguir, trazemos, na Tabela 2, informações sobre a distribuição dos textos e dados produzidos por cada um dos escreventes investigados. Um primeiro ponto que gostaríamos de destacar é em relação às diferenças que houve no número de ocorrências produzidas por cada escrevente. A esse respeito, notamos que o total de hipersegmentações, produzidas por cada um dos estudantes, variou muito entre si, oscilando desde a quantidade máxima de 37 (32,8%) ocorrências (identificadas nas produções textuais do escrevente Bruno) até a quantidade mínima de 2 (1,8%) dados (extraídos dos textos dos estudantes

³ Os números de 1 a 12 foram organizados por ordem alfabética dos nomes reais dos alunos.

André e Viviane), uma variação que, nesse caso, é de 31%, em termos percentuais. Uma hipótese explicativa que atribuímos a essas diferenças é dada a partir da relação com os modos de participação dos estudantes em práticas sociais de linguagem, as quais os levam a elaborar representações sobre a escrita, sobretudo, aquela privilegiada pela instituição escolar, que guiam os escreventes a empregarem, em seus textos, certos tipos de recursos (os recursos de fronteiras de palavras, particularmente em nosso trabalho) que demonstram, por exemplo, as hipóteses que eles têm acerca do que seja escrever a palavra gráfica segundo a convenção. Nesse sentido, cada escrevente, por meio da história de linguagem que o constitui, chega mais, ou menos, próximo daquilo que lhe é exigido em relação ao modo de registro da escrita convencional.

Tabela 2: Distribuição do número de textos e hipersegmentações produzidos por cada escrevente

Escreventes	Textos		Hipersegmentações	
	Valor absoluto	Valor percentual	Valor absoluto	Valor percentual
André	25	9,4%	2	1,8%
Viviane	23	8,7%	2	1,8%
Ana	20	7,6%	4	3,6%
Henrique	22	8,2%	4	3,6%
Camila	25	9,4%	4	3,6%
Fernando	20	7,6%	5	4,4%
Julia	25	9,4%	7	6,1%
Mateus	25	9,4%	7	6,1%
Maria	21	7,9%	9	8,0%
Lucas	24	9,0%	10	8,8%
Pedro	17	6,3%	22	19,4%
Bruno	19	7,1%	37	32,8%
Total	266	100%	113	100%

Um segundo fato a ser comentado, ainda a partir da Tabela 2, diz respeito à correlação do número de textos que o aluno produziu e o respectivo número de ocorrências. Notamos que os dois escreventes que mais produziram marcas de hipersegmentação de palavras (Bruno: 37 ocorrências e Pedro: 22 ocorrências) foram os que redigiram a menor quantidade de textos em todos os anos escolares (19 e 17 textos, respectivamente), enquanto que os dois alunos que apresentaram menos dados (André: 2 ocorrências e Viviane: 2 ocorrências) consistiram naqueles escreventes que produziram a maioria das propostas de textos que foram aplicadas ao longo dos quatro anos investigados (respectivamente, 25 e 23 textos). Isso nos faz vislumbrar, em um primeiro momento, um possível perfil dos escreventes analisados: alunos que produzem mais textos são os que apresentam menos grafias de hipersegmentação. No entanto, a fim de confirmarmos ou refutarmos essa hipótese, nas próximas etapas da pesquisa, realizaremos o cotejamento da extensão dos textos produzidos (pelo número de palavras que foram redigidas), para vermos se o fato de um escrevente ter feito mais grafias hipersegmentadas não se deva em função, por exemplo, dele ter escrito textos com mais palavras, como também, o motivo que tenha levado outro escrevente a produzir poucas ocorrências não esteja ligado com o pequeno tamanho de seus textos.

Vejam, na Tabela 3, como o total de dados produzidos por escreventes se distribuiu no decorrer dos quatro anos escolares do EF-II. Analisando a Tabela, identificamos três tipos de trajetórias dos escreventes, em relação aos registros não convencionais, por presença, das fronteiras de palavras no que diz respeito aos anos letivos. A *primeira* é a de alunos que produziram hipersegmentações em todos os anos. A *segunda* trajetória mostra alunos que tiveram ocorrências em três dos quatro anos. A *terceira*, por fim, aponta alunos que apresentaram dados em dois de quatro anos escolares. Em relação à distribuição dos escreventes no interior dessas trajetórias, identificamos: (i) cinco alunos que se encaixaram na primeira trajetória, sendo, também, os que mais produziram hipersegmentações; (ii) dois que se enquadram na segunda trajetória, sendo: um dos alunos com ocorrências no 6º, 8º e 9º anos e o outro no 6º, 7º e 9º anos; e (iii) cinco alunos restantes se enquadram na terceira trajetória, sendo: dois dos alunos com ocorrências no 6º e 7º anos e os outros três no 6º e 9º anos. De modo geral, o desenho das trajetórias nos revelou que apenas dois escreventes concluíram o EF sem apresentar hipersegmentações em seus textos e boa parte dos escreventes analisados terminou a etapa fundamental de formação escolar não escrevendo convencionalmente os limites de palavra escrita, pois embora as quantidades de ocorrências, que foram sendo produzidas ao longo dos anos, tenham diminuído na produção escrita de todos os escreventes, elas não cessaram de aparecer no último ano letivo. Tal resultado é um indício da complexidade da noção de palavra escrita, visto seu limite ser ainda um desafio para escreventes com vários anos de escolarização. Mostra, também, que a aquisição da escrita não é um processo linear de aprendizagens cumulativas (cf. CAPRISTANO, 2007).

Tabela 3: Número de hipersegmentações produzidas por cada escrevente por ano escolar

Escreventes	Nº de hipersegmentações por ano				Total
	6º	7º	8º	9º	
André	1	0	0	1	2
Viviane	1	1	0	0	2
Ana	2	2	0	0	4
Henrique	1	2	0	1	4
Camila	2	0	0	2	4
Fernando	4	0	0	1	5
Julia	2	0	1	4	7
Mateus	1	1	4	1	7
Maria	3	1	3	2	9
Lucas	3	2	3	2	10
Pedro	6	1	11	4	22
Bruno	14	14	5	4	37
Total	40	24	27	22	113

Baseamos a observação dos dados a partir da proposta de Tenani (2011b), a respeito da classificação das segmentações não convencionais de palavras quanto ao tipo de recurso gráfico empregado. Desse modo, optamos por analisar separadamente as ocorrências que se caracterizam pela presença do espaço em branco e as que se caracterizam pela presença do hífen. A frequência das hipersegmentações, nos anos escolares, pelos dois recursos gráficos é dada na Tabela 4.

Tabela 4: Classificação das hipersegmentações quanto ao recurso gráfico nos anos escolares

Ano Escolar	Recurso gráfico				Total	
	Espaço em Branco		Hífen			
	Valor absoluto	Valor percentual	Valor absoluto	Valor percentual	Valor absoluto	Valor percentual
6º	38	33,7%	2	1,8%	40	35,3%
7º	22	19,4%	2	1,8%	24	21,3%
8º	25	22,1%	2	1,8%	27	24,0%
9º	19	16,8%	3	2,6%	22	19,4%
Total	104	92%	9	8%	113	100%

Considerando, pois, a divisão das ocorrências pelo tipo de recurso gráfico, verificamos que, no *córpus* analisado, as presenças não convencionais dos limites gráficos de palavra ocorreram, em todos os anos, tanto pelo espaço em branco quanto pelo hífen (ainda que a quantidade de aparecimento dos recursos não ocorreu de forma homogênea). Por meio desse resultado, notamos que os escreventes estiveram atentos aos recursos convencionais possíveis de delimitação das palavras na escrita, bem como quais são as estruturas que se demarcam por cada um dos recursos. Sobre este último aspecto discorreremos mais adiante.

Antes, porém, comentaremos sobre as diferenças na distribuição quantitativa das ocorrências. Do total das marcas de hipersegmentação, constatamos que 92% tiveram por base o uso não convencional do espaço em branco, enquanto que 8% corresponderam à colocação não convencional do hífen. Embora, nos dados do *córpus*, a diferença numérica entre o uso de um ou outro recurso gráfico tenha sido bastante elevada, não descartamos a relevância do estudo das ocorrências de hífen separadamente em relação aos usos do branco, uma vez que esses usos não convencionais revelam aspectos linguísticos importantes sobre o processo de letramento dos escreventes. A esse respeito, notamos que todos os 9 dados de hífen se concentraram na tentativa de grafia de estruturas do tipo: *verbo+pronome enclítico*, como em “ganha-se” – quando a escrita convencional deveria ser “ganhasse”. Desse resultado, é importante destacar que não foram quaisquer tipos de palavras que os escreventes hipersegmentaram com o hífen, ou seja, foram aquelas que remeteram a um tipo de estrutura cujo funcionamento é ensinado e, especialmente, valorizado na escola.

Junto à perspectiva de escrita assumida (CORRÊA, 2004), defendemos que os usos não convencionais do hífen, em estruturas como as que analisamos aqui, podem ser interpretados como resultado de um certo imaginário social do escrevente sobre a escrita, o qual é construído pela inserção desse escrevente nos mais diversos usos da linguagem. De modo mais específico, acreditamos que a relação entre o uso do hífen em possíveis estruturas de *verbo+enclítico* se aproxima mais explicitamente daqueles momentos em que o escrevente projeta, em seu texto escrito, representações sobre o código escrito institucionalizado. Com isso, porém, não excluimos, da análise que realizamos, a ação dos outros dois momentos de circulação dialógica do escrevente (o da gênese da escrita e o da dialogia sobre o já falado/ouvido e escrito/lido), pois, ao adotarmos as formulações de Corrêa (2004), acreditamos, 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 descrição realizada até o momento foi na direção de darmos um quadro geral a respeito da distribuição dos dados em relação aos anos escolares, aos escreventes selecionados e aos tipos de recursos gráficos delimitadores de palavras, no corpus investigado. Prosseguindo com a análise para a investigação das características linguísticas das hipersegmentações, trazemos a Tabela 5 com os dados organizados em função da configuração rítmica das palavras convencionais e grafias hipersegmentadas. Segue, também, na referida tabela, o número de dados identificado em cada estrutura, em cada ano escolar do EF-II e um exemplo de ocorrência de hipersegmentação.

Tabela 5: Estruturas rítmicas de palavras convencionais e grafias hipersegmentadas e o número de dados de hipersegmentação em cada ano escolar

Estruturas				Exemplos	Nº de dados				Total
Tipos	Palavra convencional	Tipos	Hiper		6º	7º	8º	9º	
(1)	(•*)ω >	(1.1)	(•)σ (*)σ	“então” > “em tão”	8	9	9	6	32
		(1.2)	(*)σ (•)σ	“nenhum” > “nem um”	0	0	1	0	1
(2)	(*•)ω >	(2.0)	(*)σ (•)σ	“desse” > “de se”	6	2	0	0	8
(3)	(•*•)ω >	(3.1)	(•)σ (*•)Σ	“enquanto” > “em quanto”	16	6	7	6	35
		(3.2)	(•*)Σ (•)σ	“quisesse” > “quise-se”	2	2	2	0	6
		(3.3)	(*)σ (*•)Σ	“sozinho” > “so zinho”	1	0	0	0	1
(4)	(*•*)ω >	(4.0)	(*)σ (•*)Σ	“apanhou” > “a panho”	1	0	4	1	6
(5)	(*•*•)ω >	(5.1)	(*)σ (•*)Σ	“recompensa” > “recom pensa”	0	1	1	2	4
		(5.2)	(*)σ (•*•)Σ	“acabando” > “a cabano”	2	0	0	0	2
		(5.3)	(*•*)Σ (•)σ	“conversasse” > “conversa-se”	0	0	0	3	3
(6)	(*••*)ω >	(6.0)	(*)σ (••*)Σ	“aparecer” > a parecer”	1	0	1	0	2
(7)	(*••*•)ω >	(7.0)	(*)σ (••*•)Σ	“aparecido” > “a paresido”	0	1	1	0	2
(8)	Outros			“adicionar” > “adisio o na”	3	3	1	4	11
Total					40	24	27	22	113

Por meio da organização proposta, identificamos, na Tabela 5, 7 tipos de estruturas rítmicas de palavras convencionais as quais geraram 14 tipos de estruturas rítmicas de grafias hipersegmentadas. Considerando ainda as fronteiras de palavra convencional, observamos mais um tipo que denominamos “outros”, no qual incluímos hipersegmentações cujas fronteiras não convencionais não seguiram, em termos estruturais, as tendências observadas.⁴ Para ser possível observar, longitudinalmente, regularidades de distribuição dos dados, nos tipos de estrutura rítmica, consideramos todos os tipos na análise de todos os anos letivos. Em virtude da extensão deste artigo, não comentaremos a fundo os resultados do funcionamento de cada estrutura rítmica em cada um dos anos escolares. A esse respeito, apenas sintetizamos que: (i) não houve uniformidade na distribuição de cada

⁴ Informamos que, em virtude de os dados de hipersegmentação agrupados sobre o tipo “Outros” não se encaixarem nas regularidades observadas para o restante do corpus, trataremos de analisá-los separadamente, em etapa da pesquisa a ser desenvolvida.

tipo de estrutura rítmica das hipersegmentações nos anos escolares, ou seja, nem todos os tipos ocorreram sempre em todos os anos; e (ii) os tipos (1) e (3), os quais geraram hipersegmentações que remeteram à combinação prosódica de duas sílabas e de uma sílaba e um pé métrico, respectivamente, foram os tipos mais recorrentes nos quatro anos do EF-II.

Passando a tratar de algumas questões que acreditamos se evidenciarem a partir da organização dos dados em estruturas rítmicas, destacamos, primeiramente, o fato de os limites das marcas de hipersegmentação ocorrerem em pontos os quais, fonologicamente, condizem com fronteiras de constituintes prosódicos, especialmente os de sílaba e pé métrico. No entanto, esclarecemos que a interpretação prosódica realizada é apenas uma possível perante os princípios de formação do modelo prosódico adotado, visto que dados atribuídos como a combinação de $\sigma+\Sigma$, por exemplo, podem ser entendidos, também, como sendo resultado de uma combinação entre $\sigma+\omega$, já que, segundo Nespor e Vogel (1986), a palavra fonológica se forma pelas sílabas agrupadas no interior de um pé, o qual carrega o acento primário, seu principal definidor. A partir desse resultado mais geral, a respeito de as fronteiras de hipersegmentação corresponderem com fronteiras de constituintes prosódicos, podemos dizer que as hipersegmentações de palavra são uma *pista gráfica* desses elementos da língua. Conforme anunciamos, do primeiro ao último ano do EF-II, as estruturas rítmicas (1) e (3) foram as que mais geraram ocorrências de hipersegmentação. Sobre esses tipos, comentamos que as palavras convencionais abrangidas em (1) tinham a estrutura organizacional de um Σ iambo (alternância fraco/forte) que, quando hipersegmentada, passaram a corresponder à combinação de duas sílabas. No caso do tipo (3), por sua vez, a partir da estrutura trissílaba da unidade gráfica convencional, os escreventes realizaram uma escrita baseada na relação entre uma σ e um Σ , predominantemente quando correspondia sua organização a de um troqueu (cf. subtipo (3.1)).

É possível depreender nos tipos (1.1), (2.0), (3.1), (3.2), (5.1) e (5.3), das grafias hipersegmentadas, uma informação prosódica importante que diz respeito ao fato de a alocação não convencional do limite gráfico ter ocorrido, nesses casos, na fronteira a qual está localizada o acento primário (ponto de maior proeminência prosódica dentro da palavra). Levantamos a hipótese explicativa, a partir desse achado, de que a identificação de uma proeminência fônica, em determinado ponto da palavra, leva o escrevente a supor de que há ali uma fronteira a ser registrada graficamente por um limite que é dado ou pelo espaço em branco ou pelo hífen. Parece que nesses momentos, os escreventes tentam marcar, através dos limites que propõem às palavras, características dos enunciados orais/falados na produção escrita e, assim, deixam evidenciar a imagem que têm da gênese da escrita (CORRÊA, 2004). No entanto, como as práticas orais/faladas e letradas/escrita estão sempre em diálogo, além das proeminências rítmicas no interior das palavras, as hipersegmentações apresentaram, conjuntamente, fatos que não são alheios aos aspectos circundantes da própria escrita. É interessante notar que nos tipos (3.2) e (5.3) todos os dados de hipersegmentação são referentes ao uso não convencional do hífen.⁵ A interpretação dos dados de hífen como uma marca característica de enunciados escritos, especialmente daqueles mais próximos da perspectiva escolar, já foi explicitada em parágrafos anteriores, mas ainda gostaríamos de acrescentar mais um argumento a favor do nosso posicionamento, o qual encontra base nos resultados de Abaurre e Galves (1996). Em estudo sobre pronomes clíticos do PB, as autoras chegaram à conclusão de que, em enunciados falados,

⁵ Nesse conjunto, há apenas uma ocorrência que é por espaço em branco: “falam do” (falando).

as estruturas enclíticas são marcadas, observando, pois, uma nítida preferência dessa língua pela próclise. Nota-se, então, que estruturas de verbo + pronome enclítico ainda estão preservadas na língua pelas práticas de escrita convencional, fato que nos leva a fortalecer a premissa de que a presença do hífen em fronteira de palavra, nos dados analisados, traz fortes indícios de que, ao realizá-la em seu texto, o escrevente tenha buscado alçar aspectos mais institucionalizados da escrita.

Por meio da observação das ocorrências de espaço em branco também não descartamos a ação de informações gráficas nas decisões de como segmentar. Ao contrário, em todos os tipos de estruturas rítmicas pudemos identificar grafias hipersegmentadas em que alguma parte tenha correspondido com palavras que, graficamente, se apresentam espaçadas de forma independente. Por exemplo, foi frequente nos dados a sílaba pretônica da palavra convencional ter sido isolada entre brancos, sílaba a qual linguisticamente pode corresponder a um *clítico fonológico* que, por sua vez, corresponde a palavras funcionais. Um exemplo dessa correspondência é dado a partir da hipersegmentação da palavra “demais” > “de mais” (“demais”), em que entendemos que a sílaba pretônica “de” tenha sido interpretada pelo escrevente como uma possível palavra independente em termos gráficos, a qual pode ser correlacionada a uma preposição. Encontramos, também, dados em que parte resultante da hipersegmentação correspondeu com palavras de categoria como a dos verbos, como, por exemplo: “vou tando” (voltando). Buscando explicitar uma possível relação das hipersegmentações do corpus com fatos de natureza morfossintática, apresentamos, na Tabela 6, o número de ocorrências que, em alguma parte originada a partir da hipersegmentação, correspondeu graficamente com alguma classe de palavra morfossintática.

Tabela 6: Hipersegmentações de espaço em branco e correspondências gráficas⁶

Correspondência	Hipersegmentações	
	Valor absoluto	Valor percentual
Relacionada à classe de palavras funcionais	80	70,9%
Relacionada à classe de palavras lexicais	4	3,5%
Total	84	74,4%

Conforme mostra a Tabela 6, há nos textos analisados um alto percentual de ocorrências de hipersegmentação que, além de darem indícios de questões referentes à organização prosódica da língua, põem em evidência a relevância em se considerar outras motivações na interpretação desse tipo dado. Observando, pois, que as hipersegmentações de palavras, que encontramos nas produções escritas analisadas, foram condicionadas fortemente por características advindas de informações letradas/escritas, as quais acreditamos se mostrarem, em nosso trabalho, por meio da correlação que propomos entre os limites não convencionais de palavras e as classes de palavras morfossintáticas que, por sua vez, em termos gráficos, estão sempre registradas por meio de recursos (branco/hífen) delimitadores de palavras, procuramos investigar quais unidades das classes de palavras funcionais e lexicais puderam ser identificadas nos dados. Para isso, em um primeiro momento, averiguamos quais categorias de palavras funcionais puderam ser identificadas

6 O cálculo percentual desta tabela foi realizado a partir do quanto o total de hipersegmentações com correspondência gráfica era representativo do total de dados de hipersegmentação identificado (ou seja, 113 dados).

e, posteriormente, qual era a categoria de palavra lexical. Os resultados dessa etapa da análise são ilustrados nas Tabelas 7 e 8, respectivamente. Os cálculos percentuais levaram em consideração o valor que cada tipo de correspondência apresentou na Tabela 10.

Tabela 7: Correspondência gráfica e hipersegmentações: os tipos de classes de palavras funcionais

Tipos de classes	Número de hipersegmentações	
	Valor absoluto	Valor percentual
Artigo	30	26,6%
Conjunção	6	5,3%
Preposição	41	36,3%
Pronome	3	2,7%
Total	80	70,9%

Tabela 8: Correspondência gráfica e hipersegmentações: tipo de classe de palavra lexical

Tipos de classes	Número de hipersegmentações	
	Valor absoluto	Valor percentual
Verbo	4	3,5%
Total	4	3,5%

Em ordem de frequência, as correspondências gráficas com palavras funcionais nos dados indiciam que a tentativa dos escreventes foi na direção de reconhecimento e delimitação de preposições (36,3%) – “em quanto” (“enquanto”), seguida dos artigos (26,6%) – “a gora” (“agora”), depois das conjunções (5,3%) – “com migo” (comigo) e, por último, dos pronomes (2,7%) – “qual do” (“quando”). Quanto às palavras lexicais, somente uma classe se mostrou instável ao reconhecimento dos escreventes: essa foi a dos verbos (3,5%) – “recom pensa” (“recompensa”). Por meio desses resultados chegamos à evidência de que no final do EF, os escreventes têm mais dúvidas em estabelecer limites para aqueles tipos de palavras cujo funcionamento linguístico é caracteristicamente gramatical, como preposições, artigos e conjunções. Isso confirma as conclusões de Tenani (2011) a partir de dados transversais de segmentação não convencional de palavras do EF-II, a respeito da imprecisão observada de se estabelecer limites aos elementos átonos, sejam eles sílabas pretônicas sejam eles clíticos.

Destaques finais

Por meio da análise realizada, procuramos expor uma descrição geral dos dados de hipersegmentação que podem ser encontrados nos textos escritos de escreventes do EF-II. Lembramos que quanto à descrição dos dados por anos escolares, escreventes e tipos de recursos gráficos, os resultados mostraram que (i) as ocorrências se distribuíram irregularmente durante os anos, de modo que há um aumento na quantidade de dados no 8º ano; (ii) a maioria dos escreventes terminou o EF-II ainda produzindo hipersegmentações, mas foram apenas daquelas palavras que colocaram em jogo a relação entre sílabas pretônicas e possíveis classes gramaticais; e (iii) a trajetória que caracterizou a distribuição das ocorrências nos escreventes ao longo dos anos não foi linear (na direção de diminuição

gradativa). Todos esses resultados levam a sustentar a nossa opção teórica por concepções que veem a escrita como processo (CORRÊA, 2004) e a sua aquisição não segue de um período de aprendizagens predeterminadas em um tempo (CAPRISTANO, 2007).

Referente à descrição das características gerais das hipersegmentações, por meio da organização das ocorrências pela estrutura das palavras convencionais, observamos que informações de natureza prosódica como a identificação do ponto de maior proeminência dentro da palavra, ou seja, na sílaba em que recai o acento primário, constituiu-se um lugar privilegiado para que um limite gráfico não convencional fosse inserido naquela fronteira. Para nós, isso, em alguma medida, mostra que as hipóteses dos escreventes analisados são permeadas por aspectos da noção de palavra fonológica. Aliado a essa constatação, os dados revelaram a ação de informações linguísticas que podem ser recuperadas em práticas letradas/escritas, como, por exemplo, a correlação do tempo verbal subjuntivo com a inserção de um hífen (“ganhasse” > “ganha-se”) e a tentativa de correspondência gráfica de sílabas pretônicas com classes de palavras funcionais (“comigo” > “com migo”). A partir desse resultado, acreditamos, junto com Chacon (2005, p. 84) que “por qualquer lado que se olhe para as segmentações não-convencionais (tanto a partir de seu vínculo com práticas de oralidade quanto com práticas de letramento), sempre a sua outra contraparte imediatamente se mostra”.

Encerramos este artigo, apontando para o fato de que algumas questões que foram aqui tratadas merecem mais investigação e que essa será realizada em etapas da pesquisa que ainda estão a se cumprir.

REFERÊNCIAS

ABAURRE, Maria Bernadete Marques; GALVES, Charlotte. Os clíticos no português brasileiro: uma abordagem sintático-fonológica. In: CASTILHO, Ataliba; BASÍLIO, Margarida (Org.) *Gramática do português falado: estudos descritivos*. v. IV. Campinas: Unicamp, 1996. p. 273-320.

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

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

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

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

_____. Constituintes prosódicos e letramento em segmentações não-convencionais. *Letras de hoje*, Porto Alegre, v. 39, n. 3, p. 223-232, 2004.

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

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

PAULA, Isis Fernanda Vicente de. *Movimentos na escrita inicial de crianças: um estudo longitudinal de hipersegmentações*. 154f. 2007. Dissertação (Mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Unesp, São José do Rio Preto, 2007.

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

_____. *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. ms.

_____. Entre o grupo clítico e a palavra fonológica: os erros de segmentação não-convencional de palavras. Apresentação realizada no VI CONGRESSO INTERNACIONAL DA ABRALIN. Universidade Federal da Paraíba, João Pessoa, 2009.