

# A SEGMENTAÇÃO NÃO-CONVENCIONAL DE PALAVRAS EM TEXTOS DO CICLO II DO ENSINO FUNDAMENTAL

Luciani TENANI

Universidade Estadual Paulista (UNESP)/São José do Rio Preto - SP

## RESUMO

*Este artigo trata das segmentações não-convencionais de palavras em textos produzidos por alunos dos quatro últimos anos do Ensino Fundamental. Parte-se da hipótese de que esses dados de escrita permitem observar características dos enunciados escritos e falados. Por meio da análise dos dados em constituintes prosódicos, argumenta-se que o escrevente opera com hipóteses (conflitantes) sobre a organização de sílabas átonas em constituintes prosódicos: pés métricos, palavra prosódica e grupo clítico. Encontra-se evidência de que as grafias não-convencionais têm sua principal motivação na dificuldade de o escrevente atribuir status de palavra escrita a itens gramaticais que se constituem em clíticos prosódicos.*

## ABSTRACT

*This paper deals with unconventional segmentations of words in texts produced by students of the last four years of elementary school. The main hypothesis is that these data allow us to observe the characteristics of written and spoken utterances. Through analysis of data on prosodic constituents, we argue that students deal with (conflicting) hypotheses on the organization of unstressed syllables into prosodic constituents: metric feet, prosodic word and clitic group. We found evidence that unconventional spellings have their main motivation in the difficulty of students to assign the status of written word to grammatical items that are prosodic clitics.*

## PALAVRAS-CHAVE

*Letramento. Oralidade. Ortografia. Palavra. Prosódia.*

## KEYWORDS

*Literacy. Orality. Orthography. Prosody. Word.*

*Chega mais perto e contempla as palavras.*

*Cada uma*

*tem mil faces secretas sob a face neutra*

*e te pergunta, sem interesse pela resposta,*

*pobre ou terrível, que lhe deres:*

*Trouxeste a chave?*

(Carlos Drummond de Andrade)

## Introdução

As pesquisas sobre a segmentação não-convencional de palavras escritas focam, predominantemente, os dados produzidos por alunos nas etapas iniciais do Ensino Fundamental (doravante, EF-I), como o fazem, dentre outros, ABAURRE (1991); SILVA (1991), ABAURRE E SILVA (1993), CHACON (2004, 2005), CUNHA (2004, 2010), CUNHA E MIRANDA (2007), CAPRISTANO (2007a, b), PAULA (2007). A partir dessas pesquisas, lançamos luz para dados de escrita que ainda não têm sido objeto de reflexão por parte dos estudiosos da linguagem, particularmente os que investigam a aquisição da escrita, ao tratarmos, aqui, das segmentações não-convencionais de palavras escritas

encontradas em textos de alunos que, à época de sua produção, cursavam os quatro últimos anos do Ensino Fundamental (doravante, EF-II). Ao encontrarmos grafias não-convencionais, como “agente” (para “a gente”) ou “disse deram” (para “decidiram”), observamos possibilidades de segmentação e de sentidos das cadeias fônicas com as quais o aluno, possivelmente, esteja lidando ao escrever. O olhar curioso que nos guia nessa investigação é aquele que vislumbra encontrar a(s) chave(s) para desvelar os segredos das palavras (escritas), tomando as segmentações não-convencionais de palavra como pistas do trabalho do sujeito com a linguagem – como, de certo modo, tematizado no trecho, acima, do poema “Procura da Poesia” de Carlos Drummond de Andrade.

O estudo que temos conduzido sobre os dados de segmentação não-convencional de palavras em textos dos últimos anos do EF (FAPESP 2009/14848-6 e CNPq 306471/2009-4) tem, entre outros objetivos, descrever suas características linguísticas, comparando-as com aquelas descritas para dados produzidos por estudantes dos primeiros anos do EF-I, como o fazem os autores anteriormente citados. Neste texto, por meio da análise de um conjunto de dados selecionados do córpus do Banco de Dados de Escrita do EF – II (a ser detalhado na próxima seção), formulamos a seguinte questão: as segmentações não-convencionais de palavras do EF-II apresentam características que evidenciam uma organização prosódica do Português Brasileiro? A premissa deste trabalho é que os dados de segmentação não-convencional da palavra escrita permitem observar tanto características dos enunciados escritos quanto as dos enunciados falados (perspectiva teórica também assumida por CHACON (2004, 2005), PAULA (2007), CAPRISTANO (2007a, b)). Nossa hipótese é: se características dos enunciados falados podem ser identificadas, então, é possível observar características fonológicas, especialmente, as prosódicas, a partir de dados de escrita que dizem respeito à colocação do espaço em branco na medida em que essa informação gráfica está relacionada, simultaneamente, à delimitação

de fronteiras de palavras e à organização sintagmática dessas palavras no texto. Dados esses pontos de partida, o objetivo principal é buscar explicitar quais características prosódicas podem ser observadas nos dados selecionados, respondendo à questão específica: de quais constituintes prosódicos há evidências em dados de segmentação não-convencional de palavras escritas por alunos do EF-II?

A resposta para essa questão será dada a partir de uma análise prosódica que toma por base as descrições dos constituintes prosódicos no Português Brasileiro (doravante, PB), feitas a partir do arcabouço teórico da Fonologia Prosódica de NESPOR E VOGEL (1986), particularmente as noções acerca dos domínios pé métrico, palavra prosódica, grupo clítico.<sup>1</sup> Antes de adentramos a fundamentação teórica sobre prosódia, na análise dos dados e na discussão dos resultados, descreveremos, na próxima seção, o material utilizado e os passos metodológicos trilhados na seleção dos dados de escrita. Ao trilhar esse percurso, almejamos contribuir com evidências para fomentar reflexões sobre a identificação de dados de escrita que possam trazer pistas das possíveis relações entre fala e a escrita.

## **1. A segmentação não-convencional de palavra: aspectos metodológicos**

A segmentação não-convencional de palavra gráfica é definida quer a partir da ausência, quer a partir da presença do espaço em branco que delimita palavra em local não previsto pelas convenções ortográficas. Quando há ausência do espaço em branco, trata-se de hiposegmentação, como em “ajudime”, “porfavor”. Quando há presença do espaço em branco, trata-se de hipersegmentação, como em “na quelá”, “cava lo”. Uma questão metodológica importante relacionada a essa distinção é, em

<sup>1</sup> É tema de discussão a pertinência em se considerar o grupo clítico como domínio prosódico para o PB. Neste texto, não será abordada essa questão teórica, ficando essa discussão para futuras publicações.

textos manuscritos, identificar quando há ou não o espaço de fronteira de palavras empregado fora das convenções ortográficas. As Figuras 1, 2, e 3, que apresentamos mais à frente, são exemplos de diferentes tipos de grafias que geram dúvidas na identificação de dados.

Na Figura 1, a grafia de “dinheiro” é destacada como um possível dado por haver um espaço entre “di” e “nheiro”. No entanto, ao serem observadas as grafias de <i> ao longo do texto, constata-se que o escrevente a produz de tal modo que, por vezes, há algum espaçamento que foge ao esperado para o traçado dessa letra. Nesse caso, para definir se essa é ou não uma ocorrência de segmentação não-convencional de palavra, foi tomada como critério a comparação dos traçados da letra <i> ao longo do texto com a distribuição do espaço em branco entre as demais palavras do texto. Feita essa comparação, definiu-se que essa não é uma ocorrência de segmentação não-convencional de palavra.

FIGURA 1: Fonte Z08\_5A\_23F\_03<sup>2</sup>

Ola, meu nome é Katherine e meu sogro  
que é de ganhar ~~na~~ na mega  
loteria. Ganharia milhares de reais de  
corrida moto casa dinheiro, e eu  
que só construi uma charca que  
tembe, comprei piscina e tratame.  
Porque eu gosto muito  
é porque desejaria ter um quanto lindo  
dentro da minha charca. E se  
desse dinheiro e ganharia High School  
músical, os artistas, e me mairio  
e os des (J) V (S). eu queria ter filhos

<sup>2</sup> A codificação apresentada segue as normas de identificação de texto e sujeito do banco de dados ao qual esse texto pertence. Os elementos separados pelo traço correspondem respectivamente a: escola/ano de coleta do texto; série/turma; sujeito/sexo; proposta de redação.

Na Figura 2, a grafia de “dificuldade” apresenta um espaçamento proporcionalmente maior entre “dificu” e “ldade”. Essa grafia foi vista como resultado de uma flutuação na forma de grafar as letras em uma mesma palavra que se caracteriza pela alternância entre letras manuscritas e impressas de modo que por vezes, ocorre a ausência das ligaduras entre as letras dentro de uma palavra. Novamente, é a comparação da distribuição do espaço em branco entre todas as palavras do texto que sustenta a desconsideração dessa grafia como sendo um dado de segmentação não-convencional de palavra. Ainda na Figura 2, destacamos as grafias de “a minhas” e “amigas”. A primeira dessas grafias poderia ser um dado de hipersegmentação, se comparada à grafia de “a fazer” na linha anterior; ou não, se comparada à grafia de “amigas”. Como já explicitado, é a comparação da distribuição do espaço em branco que delimita palavra com a forma de traçar as letras ao longo de todo o texto que nos leva a afirmar que não se trata de um dado para este tipo de estudo.

FIGURA 2: Fonte Z08\_6A\_19F\_04

The image shows two lines of handwritten Portuguese text. The first line reads: "Gostei eu entro com muita dificuldade com a internet? Como é que eu entro na pagina do MSN? A minha amiga pegou para mim o meu MSN, mas ela não me ensinou. Da entro na pagina, como é que eu faço para entrar?" The second line reads: "E aproveitando você me entra na feira o ORKUT, todos a minhas amigas tem eu também gosto ter. E também faz me fazca o seu enute?"

Na Figura 3, o traçado de <n> em “mandar” e “aprendendo”, por um lado, e em “mensagem”, por outro, gera a identificação de quatro ocorrências de hipersegmentação: “man dar” e “apren den do” (duas vezes cada). Um olhar atento, porém, constata que a ausência de ligadura observada é decorrente da letra que se segue a <n>: é a grafia de <d>

que leva ao espaço em branco nas quatro grafias consideradas. Nos demais casos de ocorrência de <n> seguida de outra letra, sempre ocorre a ligadura e a ausência do espaço em branco, como pode ser constatado na Figura 3. Portanto, nenhuma dessas ocorrências é um dado de segmentação não-convencional de palavras, sendo vistas como características de caligrafia dos escreventes.

FIGURA 3: Fonte Z08\_6C\_12F\_04

Cara amiga Gustavo eu sei que  
você sabe mandar mensagem eu estou  
aprendendo também.  
Olha eu vou falar um pouco do que eu  
sou se você tem que entrar na internet  
e acessa o MSN do seu colega ai  
você escreve o que você quer falar pra  
ele.  
O mais eu só aprendendo a ler  
e a escrever também ai eu posso  
mandar uma mensagem pra você  
meu amigo a minha (~~mãe~~) mal vai

Por fim, faz-se necessário explicitar o tratamento dado à translineação, ilustrada na Figura 4, que se caracteriza pelo registro gráfico de uma palavra quando se dá a passagem de uma linha para outra, ficando dividida. Segundo as convenções para manuscritos, um hífen deve ser colocado onde a palavra ficar divida na translineação. A ausência desse hífen, como ocorre na Figura 4, no registro de “encontro”, não foi considerada uma segmentação não-convencional de palavras, por ter sido interpretada como uma questão de translineação.

FIGURA 4: Fonte Z08\_6A\_29F\_04

Legal, tudo que você procura lá, você encontra, o  
google então, bem esse, é o site de procura, lá você en-  
contra imagens, notícias, tudo que você procura.

A descrição ora realizada exemplifica o tipo de dificuldade encontrado na categorização do espaço em branco entre palavras e, simultaneamente, permite avaliar o tipo de olhar necessário para a correta identificação dos dados. Na literatura, PAULA (2007) trata detalhadamente, para dados de escrita infantil (primeira a terceira séries do EF), dos aspectos metodológicos relevantes para identificação de ocorrências de segmentação não-convencional de palavras. Diferentemente daquele trabalho, porém, este trata de textos manuscritos por alunos do EF-II, cujas características relacionadas à grafia das letras e dos espaços em branco são distintas daquelas identificadas nos textos infantis. Por ser inédita a análise deste tipo de texto manuscrito, optamos por detalhar as decisões metodológicas tomadas. Explicitamos, pois, os critérios adotados para a identificação dos dados, uma vez consideradas as características do material de pesquisa: (1) comparação entre palavras dos espaços em branco ao longo do texto; (2) comparação entre as grafias das mesmas letras dentro de palavras semelhantes que ocorram no mesmo texto.

Definidos os critérios para a identificação de ocorrências de segmentação não-convencional de palavra, passamos a problematizar a tipologia dessas ocorrências. Antes, porém, lembramos que um terceiro tipo de dados tem sido nomeado como mesclas (CHACON, 2006) ou híbridos (CUNHA, 2004, 2010). Nos dados levantados até o momento, apenas uma ocorrência foi identificada, a saber: “*sea certa*” (“*se acerta*”), produzida por um aluno de sétima série/oitavo ano. Esse dado se caracteriza por apresentar hipo e hipersegmentação em dois vocábulos contíguos, isto é: há hiposegmentação entre as fronteiras das palavras “*se*” e “*acerta*”, seguida de hipersegmentação em “*acerta*”, separando a sílaba pretônica “*a*” de “*certa*”. Um resultado a ser confirmado, após a conclusão do levantamento de dados em todo o material da pesquisa, é o total de ocorrências desse tipo, podendo ser a sua baixa frequência um traço característico dos textos do EF-II, quando comparados com os textos produzidos no EF-I.

Os três tipos de segmentação não-convencional descritos, a saber, hipo, hipersegmentação e híbridos, têm como critério a localização não-convencional do espaço em branco que define a fronteira de palavra escrita em relação à convenção ortográfica. Na análise dos dados, encontramos ocorrências que envolvem, além do critério básico que define uma segmentação não-convencional, características linguísticas e ortográficas que, neste momento da investigação, nos levam a problematizar a pertinência (ou não) em considerá-las como tipo distinto dos demais já estabelecidos na literatura.

São recorrentes, no córpus constituído a partir de textos do EF-II, dados que envolvem a colocação do hífen. Identificamos três possibilidades, a saber: (1) presença do hífen onde não é prevista sua colocação na palavra, como “estava-mos”; (2) presença do hífen entre duas palavras, como “certo-dia”; (3) ausência do hífen seguida de hiposegmentação, como “despistalos”; (4) ausência do hífen seguida de espaço em branco, como “mata lo”.

Em “estava-mos”, a presença do hífen leva à hipersegmentação da palavra “estávamos”. Nota-se que a semelhança dessa grafia com sequências como “chamava-nos” pode ter sido a motivação para o uso não-convencional do hífen. Em “certo-dia”, a presença do hífen não altera a segmentação de uma palavra, mas estabelece relação entre duas palavras, como se fossem palavras compostas semelhantes a, por exemplo, “bem-vindo”. Este segundo caso, em relação ao primeiro considerado, se particulariza por trazer à cena a relação entre palavras. Em função dessa característica, optamos por não analisá-lo, neste texto, como um caso de segmentação não-convencional de palavra. Porém, assinalamos sua ocorrência e acreditamos que considerações acerca do estatuto desse tipo de dado poderão ser feitas tomando-se por base um conjunto maior de dados como é o nosso propósito fazer em breve. Em “despistalos” e em “mata lo”, a ausência do hífen é a característica em comum, mas a primeira é distinta da segunda na medida em que,

também, ocorre a junção<sup>3</sup>. Assim, “despistalos” se caracteriza como uma hipossegmentação, enquanto “mata lo” se caracteriza apenas pela ausência do hífen. Consideramos, em princípio, grafias semelhantes à “mata lo” como não sendo ocorrências pertinentes para o estudo das segmentações não-convencionais de palavra, por não envolver a hipossegmentação propriamente dita. No entanto, ainda observamos ocorrências como “passaranse se” que se caracteriza como uma hipossegmentação, por haver ausência do hífen e junção entre o verbo “passaram” e o pronome “se”, seguida da presença de “se”, cujo registro é semelhante ao que observamos em “mata lo” por não haver o hífen. Esse tipo de dado nos traz indícios de que o escrevente tem dúvidas quanto às possibilidades de registro dos pronomes enclíticos ao verbo, o que nos leva a não descartar, ao menos das análises qualitativas, ocorrências como essas e a propor uma tipologia baseada, também, na presença/ausência de hífen como um critério, somado à presença/ausência do espaço em branco, na categorização dos dados de segmentação não-convencional de palavras escritas.

## 2. A segmentação não-convencional de palavra em textos do EF-II

Os dados de nossa investigação são extraídos de textos produzidos por alunos que participaram do Projeto de Extensão Universitária “Desenvolvimento de Oficinas de Leitura, Interpretação e Produção Textual”, credenciado junto à Pró-Reitoria de Extensão da UNESP.<sup>4</sup> Esse material compõe a amostra transversal do Banco de Dados de Escrita

<sup>3</sup> Em uma versão anterior deste trabalho, usávamos o termo “juntura” (no sentido empregado por Cagliari, 1989). Por sugestão dos avaliadores, passamos a usar “junção”, a fim de evitar uma possível confusão com o conceito de juntura da área de fonética/fonologia.

<sup>4</sup> O projeto de extensão foi cadastrado junto ao Conselho Nacional de Saúde (FR 198751 e CAAE n. 0013.0.229.000-08) e aprovado (FR 53/08) pelo Comitê de Ética em Pesquisa/UNESP. Recebeu apoio financeiro da PROEX/UNESP em 2011, 2010 e 2009, e da Fundação de Apoio à Pesquisa de São José do Rio Preto em 2008. Agradecemos os auxílios financeiros recebidos.

do EF – II, constituída de 2469 textos. Desse banco, selecionamos 424 textos coletados no início do ano letivo entre os quatro últimos anos do EF investigados. Para cada ano, foi elaborada uma proposta a partir de uma tirinha que tinha três quadros: no primeiro, a cena era de um rapaz e uma moça, em sua garupa, montados em um cavalo galopando e a fala da moça era: “Mais depressa, Juvêncio! ... Papai está em nosso encalço!”; no segundo quadro, a cena era a mesma e a fala da moça era: “Papai, os jagunços, o delegado, o batalhão da guarda, o regimento de cavalaria, o...”; no terceiro quadro, apenas havia a cena do rapaz chutando a moça e esta, de pernas para o ar, caindo do cavalo. Cada proposta foi elaborada levando-se em conta o gênero/tipo textual trabalhado em sala de aula, conforme Proposta Curricular do Estado de São Paulo (2008)<sup>5</sup> para a disciplina de língua portuguesa.

Analisando os 424 textos selecionados, identificamos um total de 162 ocorrências de segmentações não-convencionais de palavras. Na Tabela 1, são apresentadas informações sobre: o total de textos coletados em cada ano; o total de hipo e hipersegmentação em cada ano, bem como a porcentagem dessas ocorrências em relação ao total de dados identificados.

TABELA 1: Distribuição de textos e de hipo e hipersegmentação

Ano	Textos coletados		Hipo		Hiper		Total de segmentações	
	Nº	%	Nº	%	Nº	%	Nº	%
6 <sup>a</sup>	107	25,2	24	14,8	25	15,4	49	30,2
7 <sup>a</sup>	100	23,6	20	12,3	17	10,5	37	22,8
8 <sup>a</sup>	110	26,0	33	20,4	20	12,4	53	32,8
9 <sup>a</sup>	107	25,2	16	09,9	07	04,3	23	14,2
<b>Totalis</b>	<b>424</b>	<b>100</b>	<b>93</b>	<b>57,4</b>	<b>69</b>	<b>42,6</b>	<b>162</b>	<b>100</b>

<sup>5</sup> As propostas de produção escrita, para sexto, oitavo e nono ano, foram definidas como pertencendo ao gênero conto e ao tipo narrativo; para o sétimo ano, gênero carta pessoal e tipo relato. Uma discussão sobre tipologia e gênero textual é importante para a análise dos textos, mas não é o foco deste texto.

De modo geral, constatam-se dois resultados principais. O primeiro é que diminuiu o total de ocorrências de segmentação não-convencional do sexto para o nono ano, exceto no oitavo ano, que apresenta o mais alto índice de ocorrências (32,8%) em relação aos demais anos considerados. Esse resultado demanda uma futura investigação de dois aspectos, a saber: (1) quais os motivos que levam a um comportamento particular do oitavo ano em relação aos demais? Esse resultado se mantém se considerados mais textos? (2) que correlação é possível estabelecer entre os dados encontrados e o tempo de escolarização e/ou as práticas de letramento (escolar)? O segundo resultado a comentar é que há, de modo geral, predomínio de hipossegmentação (57,4%) em relação à hipersegmentação (42,6%). Essa relação se mantém para todos os anos, exceto no sexto ano que apresenta uma relação bastante próxima entre os dois tipos de dados (14,8% de hipossegmentação e 15,4% de hipersegmentação). Essa tendência geral ao predomínio de hipossegmentação quando considerado o total de ocorrências de segmentação não-convencional de palavras também foi observada, por exemplo, por Ferreiro e Pontecorvo (1996) em pesquisa sobre esse mesmo tipo de dado em textos de crianças em fase inicial de aquisição da escrita no México, Uruguai, Brasil, Itália. Esse também é um resultado que aponta para o desenvolvimento de pesquisas que busquem interpretar essa (aparentemente, constante) relação entre hipo e hipersegmentação de palavras ao longo do processo de aquisição de escrita. Vislumbramos que uma possível motivação esteja no percurso de análise do *continuum* fônico dos enunciados falados que o escrevente é levado a fazer por meio da representação gráfica de palavras. Por hipótese, haveria o predomínio de registro de porções maiores do enunciado – resultando em hipossegmentações – em razão de os escreventes ancorarem sua escrita em (algumas) características dos enunciados falados. No entanto, a flutuação entre hipo e hipersegmentação nos leva a ponderar, como o faremos nas próximas seções, acerca da complexa relação entre

as modalidades faladas e escritas e as práticas sociais orais e letradas relacionadas a esses dados de escrita que acenam para o predomínio das hipossegmentações.

Continuando a tratar de características quantitativas dos dados do EF-II, apresentamos, no Quadro 1, todas as ocorrências de hipo e hipersegmentação consideradas.<sup>6</sup> Em seguida, apresentamos, na Tabela 2, os números (e as respectivas porcentagens) de dados cujas grafias convencionais prevêm o uso do hífen, no caso de hipossegmentação como “pegalo”, ou prevêem a sua ausência, no caso de hipersegmentação como “estava-mos”.<sup>7</sup> Seguem-se, também, para fins de comparação, os números de hipo e hipersegmentação cujas grafias não têm relação com usos do hífen.

QUADRO 1: Ocorrências de hipo e hipersegmentação em textos do EF-II

Ano	Hipossegmentação	Hipersegmentação
6 <sup>a</sup>	cercasar, porfavor, portodos, devolta, denovo, dinovo, derrepende, terrebente, dererpemte, afrente, atarde, eseu, ciesconder, meamava, perseguiło(s), rapitalo, despistalo(s), achalos, pegalo, alcançalo, socorela, ajudime,	com dinuou, com migo, com sigo, em bora, em quanto, en tão, a sim, a onde, a panho, a té, a qui, a tras, à baixo, a quele, na quele(a), da quele (a), de la, da li, ou tra, dele gato (2x), disse deram, cava lo

<sup>6</sup> Não se consideram, neste trabalho, as rasuras identificadas: foi registrado que 8 das 12 rasuras envolvem colocação de hífen. Foram computadas como ocorrências de um mesmo tipo grafias como “derrepende, terrebente, dererpemte, derrepente” por serem hipossegmentações da mesma palavra “de repente”. Uma discussão que embasa essa decisão metodológica relativa às escolhas não-convencionais de letras observadas nas grafias das hipo e hipersegmentações é feita em Tenani (2010).

<sup>7</sup> Excluímos da análise quantitativa ocorrências que se caracterizam (1) apenas pela ausência do hífen: “mata lo, eis namorado” (sétimo ano) e “trocaram se” (oitavo ano); (2) pela presença de hífen entre duas palavras: “mau-humor, sem-vergonha, certo-dia (2x), e-se (e se)” (oitavo ano). Também foi excluído “Lan-House” (oitavo ano) por se tratar de um empréstimo linguístico.

continuação Quadro 1

7 <sup>a</sup>	agente (5x), agenten, ajente, apar, oque, conquen, concerteza (2x), envez, derrepente (2x), pratas, porisso, portodos, medesbetinto, encontralo	em quando, em bora, com migo (3x), a quilo, a pesar (2x), de mais, de vagar, de presa, da quele (2x), na quela, e capar, tarde sinha, estava-mos
8 <sup>a</sup>	agente (2x), apartir, derrepente (6x), dinovo, denovo, afim (4x), anoite, oque (2x), ostros, enseguida, emcima, eporço (e por isso), semvergonha (2x), tabom (2x), pérrapado (2x), levala, pegala, pegalos, dechalo, perceguilos	em tam, em tão, em canto, em bora, e la, es pero, es que cer, ou viu, da li, da qui, de mais (2x), a trás (2x), a sim, a o (ao), a Deus (adeus), a vistos (avistou), cala afrio (calafrio), jun-tos
9 <sup>a</sup>	enfrete, embusca, derrepente (3x), denono, umdia, apouco, anoite, afim, oque, seteclarou, pedila, velo, buscala, chamalo	em bora (2x), em baixo, de presa, da quela, a quele, mora-va

TABELA 2: Ocorrências *versus* não-ocorrências de dados que envolvem hífen

Tipos	Envolvem hífen		Não envolvem hífen		Total	
	Nº	%	Nº	%	Nº	%
<b>Hipossegmentação</b>	20	12,4	73	45,0	93	57,4
<b>Hipersegmentação</b>	03	01,9	66	40,7	69	42,6
<b>Total</b>	23	14,3	139	85,7	162	100

Verifica-se que, do total de ocorrências consideradas, 14,3% correspondem a grafias de palavras que envolvem a colocação do hífen, tendo sido 12,4% resultado em hipossegmentações, como “velo”, e 1,9% em hipersegmentações, como “mora-va”. Excluídos esses dados do total de hipo e hipersegmentações, constata-se que a proporção entre hipo e hipersegmentações se aproxima: entre os dados que não envolvem

hífen, são 45% de hipossegmentações e 40,7% de hipersegmentações (uma diferença de 4,3%); mas se considerado o total de dados, a relação é de 57,4% de hipossegmentações e 42,6% de hipersegmentações (uma diferença de 14,8%). Esses resultados corroboram a proposta, feita na seção anterior, de que os dados que envolvem presença/ausência não-convencional de hífen sejam analisados separadamente, pois é possível observar, por meio desse levantamento, que: (1) no caso das hipossegmentações (a maioria dos dados), está em jogo o uso do hífen exclusivamente em estruturas VERBO+CLÍTICO, ou seja, em estruturas enclíticas que são predominantemente, empregadas em textos escritos e a partir do processo de escolarização/letramento;<sup>8</sup> (2) no caso das hipersegmentações, estão em jogo dois possíveis usos convencionais do hífen: um relacionado a estruturas de palavras compostas (como “mau-humor, certo-dia”) e outro a estruturas verbo-clítico (como “mora-va, estava-mos”); ambos os casos envolvem regras ortográficas relacionadas ao processo de letramento dos sujeitos escreventes.<sup>9</sup>

Concluímos esta descrição acerca dos tipos de dados identificados nos textos selecionados apontando para o fato de haver usos não-convencionais do hífen (quer pela presença indevida, quer pela ausência indevida) que nos levam a refletir sobre a classificação dos dados em hipo e hipersegmentação, especialmente por trazerem evidências de estarem em jogo aspectos gráficos/letrados (como, os usos convencionais do hífen e do espaço em branco delimitador de palavra) e linguísticos (como, a identificação de sequências verbo-clítico) que os caracterizam e, também, apontam para a complexa relação entre esses aspectos (aos quais o escrevente parece estar atento no processo de aquisição da escrita).

---

<sup>8</sup> Das 20 ocorrências encontradas, 10 estão concentradas no sexto ano.

<sup>9</sup> Das 8 ocorrências identificadas, 6 ocorreram nos textos de oitavo ano e nenhum dado no sexto ano.

### 3. Análise prosódica das segmentações não-convencionais de palavra

Na literatura sobre a escrita infantil, encontram-se argumentos sobre a relevância dos chamados erros de segmentação não-convencional de palavra não apenas para a caracterização do processo de aquisição da escrita, mas também para a reflexão sobre a organização prosódica da língua. Neste texto, o eixo da reflexão está ancorado nas análises de ABAURRE (1991), CAPRISTANO (2007a, b), CHACON (2004, 2005), CUNHA (2004, 2010), dentre outros, que, ao analisarem textos feitos por crianças na fase inicial de alfabetização, defendem a pertinência de tomar esses dados de escrita como evidências de constituintes prosódicos, tais como o pé métrico (doravante,  $\Sigma$ ) e a palavra fonológica (doravante,  $\omega$ ).

Argumentamos, nesta seção, que as segmentações não-convencionais estudadas se particularizam, no que diz respeito a possíveis relações com os enunciados falados, por mobilizarem, predominantemente, informações de natureza da organização de sílabas átonas ou como sílabas pretônicas/postônicas pertencentes a uma  $\omega$  ou como elementos clíticos que integram um constituinte prosódico superior a  $\omega$ .

Antes de avançarmos, faz-se necessário explicitar, resumidamente, que se assume, no PB, a noção de  $\omega$  proposta por BISOL (1996) a partir do modelo de NESPOR E VOGEL (1986). Nesse modelo, a  $\omega$  é o domínio em que se dá a interação entre os componentes fonológico e morfológico da gramática. Na hierarquia prosódica, a  $\omega$  é o constituinte imediatamente acima do  $\Sigma$ , podendo ser composta por um ou mais  $\Sigma$ s, e dominada pelo grupo clítico. Caracteriza-se por ser portadora de um acento lexical ou acento primário. Já o domínio do  $\Sigma$ , nesse mesmo arcabouço teórico, é definido pela relação de dominância que se estabelece entre duas ou mais sílabas, podendo apresentar uma configuração n-ária, em que uma sílaba é forte e as demais fracas.

A noção de grupo clítico (C) aqui assumida é a que se encontra em BISOL (2000a), segundo a qual C constitui uma unidade prosódica que domina imediatamente a palavra fonológica e é formado por uma única palavra de conteúdo acompanhada de clíticos (palavras funcionais átonas, tais como artigos, preposições, conjunções). Em um enunciado como “A casa do Carlos fica em Campinas”, os grupos clíticos são: [A casa]C [do Carlos]C [fica]C [em Campinas]C. BISOL (2000a), baseada nas ideias de NESPOR E VOGEL (1986) sobre a existência do grupo clítico como constituinte da hierarquia prosódica, propõe um estudo sobre o *status* prosódico do clítico no PB. Por meio de uma detalhada análise das relações estabelecidas entre o clítico e seu hospedeiro, BISOL (2000a: 18) apresenta “evidências que argumentam em favor da hipótese de que o clítico assume com seu hospedeiro, no pós-léxico o seu real *status* prosódico”. Por meio da comparação de regras de sândi vocálico que resulta na degeminação e na elisão, a autora encontra o argumento decisivo para a sua proposta. Em especial, observa que a elisão, no PB, não ocorre no interior de uma palavra, mas somente se aplica entre o clítico e a palavra lexical. Nos termos da investigadora:

A elisão, diferentemente da degeminação, abrange todos os domínios do pós-léxico, do menor ao maior, mas nenhum do léxico, manifestando-se exclusivamente como sândi externo. Eis um argumento seguro para sustentar a hipótese da prosodização do clítico em nível pós-lexical. (BISOL, 2000a: 21)

Dessa forma, a autora apresenta evidências de que o clítico não faz parte da palavra fonológica lexical, mas forma um constituinte prosódico pós-lexical com a palavra de conteúdo com que se relaciona.

Ponderamos que esta escolha teórica é passível de discussão, uma vez que há propostas de as mesmas estruturas serem analisadas como,

por exemplo, uma frase fonológica, como o faz SIMIONI (2008) na análise de dados de fala. Optamos por deixar para outra oportunidade a consideração de duas possíveis interpretações dos dados de escrita à luz de arcabouços teóricos distintos no que diz respeito, especificamente, a esses dois constituintes prosódicos no PB.

Limitando-nos, pois, a analisar as hipossegmentações a partir de uma proposta de hierarquia prosódica para o PB que admite o grupo clítico como um domínio prosódico, iniciamos a análise prosódica focalizando as hipossegmentações identificadas na seção anterior. A hipótese de trabalho consiste em tomar a ausência não-convencional de espaço em branco entre palavras como critério para interpretar que o registro gráfico foi, em alguma medida, motivado pela percepção de um constituinte prosódico. A análise parte da proposta que toma esse tipo de registro gráfico não-convencional como motivado pela configuração de grupo clítico nos enunciados falados. Apresentamos, por meio do Quadro 2, exemplos de cada estrutura prosódica identificada na análise das hipossegmentações e, na Tabela 3, a distribuição de ocorrências dessas estruturas prosódicas em relação aos anos em que os alunos estavam matriculados.

**QUADRO 2:** Exemplos de estruturas prosódicas para hipossegmentações

Estruturas prosódicas	cl+ω	ω+cl	ω+ω	cl+cl	cl+cl+ω
Exemplos	me amava con certeza	le vala	perra pado tabom	O que	E poriço

**TABELA 3: Tipos de estrutura prosódica para hipossegmentações**

Ano	Cl+ω	ω+cl	ω+ω	cl+cl	Cl+cl+ω
6 <sup>a</sup>	13	10	1	0	0
7 <sup>a</sup>	18	01	0	1	0
8 <sup>a</sup>	19	05	6	2	1
9 <sup>a</sup>	11	04	0	1	0
Total	61 (65,6%)	20 (21,5%)	7 (7,5%)	4 (4,3%)	1 (1%)

Onde: cl: clítico; w: palavra prosódica; +: junção; >: passa a

No Quadro 2, observa-se que as cinco estruturas prosódicas identificadas envolvem somente clíticos e palavras prosódicas que são hipossegmentados. Na Tabela 3, verifica-se que a maioria (65,6%) das hipossegmentações envolve estruturas de clíticos seguidos de palavra fonológica, como “meamava” e “concerteza” e essa estrutura também predomina em todos os anos escolares considerados. Nos estudos sobre prosodização do clítico, como em GALVES E ABAURRE (1996), encontram-se resultados que asseguram ser a próclise a estrutura típica dos enunciados falados em PB. Desse modo, as hipossegmentações aqui descritas podem ser interpretadas como motivadas, em alguma medida, por essa característica prosódica dos enunciados falados. O segundo tipo de estrutura mais recorrente (com 21,5% do total de dados) são hipossegmentações de palavra fonológica e clítico, como “levala”. Destacamos que todas as vinte ocorrências são verbos seguidos de clíticos (cf. Quadro 1). Essa colocação do pronome enclítico ao verbo é considerada uma estrutura marcada quando analisados os enunciados falados em PB (GALVES e ABAURRE, 1996: 286), sendo uma colocação pronominal fortemente associada a enunciados escritos, a qual é, predominantemente, ensinada nas escolas. Constata-se ainda que essas ocorrências se caracterizam pela ausência do hífen entre o verbo e o clítico, podendo, dessa forma, serem vistas como ocorrências que dizem respeito, predominantemente, ao processo de aquisição da estrutura verbo-clítico, cuja grafia também é trabalhada no EF-II, por meio de

práticas letradas em ambiente escolar. Somando-se as duas estruturas ora descritas, totalizam-se 87,1% dos dados, ou seja, a maioria absoluta dos dados de hipossegmentação diz respeito a grafias não-convencionais de um elemento clítico e seu hospedeiro; dentre esses se identificam quer a próclise – uma colocação pronominal mais fortemente motivada pela característica dos enunciados falados – quer a ênclise – uma colocação pronominal predominante em enunciados escritos.

Da Tabela 3, restam 12,8% de dados que podem ser analisados como envolvendo três diferentes tipos de estruturas prosódicas. Desses dados, 5,3% envolvem sequências de clíticos: “oque” (4,3%) e “eporiço” (1%). Nota-se que a partícula “que” – que não tem acento lexical – pode receber acento prosódico, passando a sequência a ser um clítico seguido de palavra prosódica, estrutura predominante nos dados. E, finalmente, 7,5% das ocorrências são, em princípio, sequências de duas palavras prosódicas, a saber: “semvergonha, tabom, pérrapado”. Porém, “semvergonha” e “tabom” podem ser interpretados como uma sequência de clítico e palavra prosódica, pois, prosodicamente, “sem” e “ta” (forma reduzida de “está”) podem ter o acento enfraquecido. Já “pérrapado” pode ter sido interpretada como uma palavra composta (como “Mirassol”), tendo em vista o emprego de <rr> na junção das palavras e, assim, não a analisamos como uma grafia que tenha relação com o domínio da frase fonológica (no qual se dão relações entre duas ou mais palavras prosódicas, como “menina bonita”). Portanto, todos esses dados não envolvem, nesta interpretação, domínios mais baixos/altos do que os que foram considerados.<sup>10</sup>

---

<sup>10</sup> Admitida a análise de formas como “tabom” como representações de palavras que possam ser vistas como constituídas de cl+ ω, resta, em nossos dados, apenas a ocorrência “pérrapado”, constituída de ω+ω, o que resultaria em uma frase fonológica, se seguida a proposta de NESPOR E VOGEL (1986), ou em uma palavra prosódica recursiva, se levadas em conta considerações feitas por VIGÁRIO (2003), sobre o Português Europeu. Uma discussão inicial sobre a pertinência da noção de palavra prosódica recursiva para o Português do Brasil, particularmente, para palavras compostas, se encontra em VIGÁRIO E FERNANDES-SVARTMAN (2010).

Concluída a descrição das características prosódicas das híposegmentações, apresentamos, por meio do Quadro 3, exemplos de cada estrutura prosódica identificada na análise das hipersegmentações e, na Tabela 4, a distribuição de ocorrências dessas estruturas prosódicas em relação aos anos letivos em que se encontravam os escreventes. A hipótese de trabalho consiste em tomar a presença não-convencional de espaço em branco nas palavras como critério para interpretar que o registro gráfico foi, em alguma medida, motivado pela percepção de fronteira de um constituinte prosódico. A análise prosódica nos leva a identificar as estruturas apresentadas no Quadro 3.

QUADRO 3: Exemplos de estruturas prosódicas para as hipersegmentações

Estruturas prosódicas	$\omega > \omega + \omega$	$\omega > cl + \omega$	$\omega > \omega + cl$	$\omega > \sigma + \sigma (+\sigma)$
Exemplos	Disse deram	em quanto; da li	cava lo; estava-mos	ou tra; es que cer

TABELA 4: Tipos de estrutura prosódica para hipersegmentações

Ano	$\omega > \omega + \omega$	$\omega > cl + \omega$	$\omega > \omega + cl$	$\omega > \sigma + \sigma (+\sigma)$
6 <sup>a</sup>	3	20	1	1
7 <sup>a</sup>	1	15	1	0
8 <sup>a</sup>	1	16	0	3
9 <sup>a</sup>	0	06	1	0
Total	5 (7,2%)	57 (82,6%)	3 (4,4%)	4 (5,8%)

Onde: cl: clítico;  $\omega$ : palavra prosódica;  $\sigma$ : sílaba; +: junção; >: passa a

Do Quadro 3, destacamos que, em duas das quatro possibilidades de estruturas prosódicas das palavras hipersegmentadas, há sílabas átonas das palavras que foram grafadas como clíticos, a saber: “em quanto” ( $cl + \omega$ ) e “estava-mos” ( $\omega + cl$ ). Na Tabela 4, constata-se que a minoria delas (4,4%) poderia ser resultado de uma tentativa de representação da estrutura do tipo verbo-cl. Porém, a colocação não-convencional do

hífen em duas das três ocorrências – em “estava-mos” e “mora-va” – e a forma do (suposto) clítico “lo” em “cava lo”<sup>11</sup> podem ser tomados muito mais como indícios de que essas ocorrências podem ser vistas como tentativas do escrevente alçar as grafias convencionais de uma estrutura típica de enunciados escritos constituídos de verbo+clítico (revelando, assim, que informações letradas também são mobilizadas nas hipersegmentações), do que como representações gráficas de uma estrutura prosódica constituída por  $\omega > \omega + \text{cl}$ . Já a maioria das ocorrências (82,6%) resulta em representações gráficas de sequências de clítico seguido de uma possível palavra fonológica (por carregarem um acento) – como “(em)cl (canto) $\omega$ ” (para “enquanto”), “(em)cl (tão)  $\omega$ ” (para “então”) –, ainda que, por vezes, essa “palavra fonológica” resultante – como “(em)cl (bora) $\omega$ ” – não corresponda a nenhuma palavra da língua, se levados em conta critérios morfossemânticos; característica essa também observada por CUNHA (2004) para dados de escrita infantil. Destacamos ainda que essa estrutura proclítica “cl+ $\omega$ ” também é a predominante em dados de todos os anos escolares, mesmo havendo um decréscimo de sua frequência do sexto para o nono ano.

Cabe aqui estabelecer uma relação entre essa característica prosódica predominante entre as hipersegmentações ( $\omega > \text{cl}+\omega$ ) e aquela que predomina entre as hipossegmentações ( $\text{cl}+\omega > \omega$ ). Nas hipersegmentações, uma palavra é grafada como uma sequência de clítico e (pseudo)palavra; enquanto, nas hipossegmentações, uma sequência de clítico e palavra é grafada como uma só palavra. Esses registros não-convencionais sinalizam uma flutuação das decisões do escrevente acerca das fronteiras de palavra, pois estão em jogo tentativas de “fraseamento” de sílaba átona que são (na hipossegmentação) ou poderiam ser (na hipersegmentação) um clítico prosódico (compreendendo palavras funcionais monossilábicas).

<sup>11</sup> Por meio desta análise, não se está afirmando que o escrevente tenha feito relação direta entre “cavalo” e “cavá-lo”, mas somente que a grafia não-convencional “cava lo” guarda relação, a partir da distribuição do espaço em branco, com uma estrutura típica de enunciados escritos em que “lo” figura como uma representação de elemento clítico.

Dado que a sequência “cl+ $\omega$ ” resulta ser a predominante entre as hipersegmentações, exploramos, brevemente, a hipótese de esses registros gráficos serem uma representação motivada pela percepção, também, de proeminências rítmicas, de modo que a grafia resulta em uma sequência de sílaba e  $\Sigma$ . Na Tabela 5, as sequências são exemplificadas e quantificadas. Entre os tipos de  $\Sigma$ s identificados, verifica-se um predomínio do troqueu dissílabo (52,7%), seguido por  $\Sigma$ s constituídos de uma sílaba (35%), a qual pode ser bimoraíca, como em “a sim”, ou não, como em “a qui”. Por fim,  $\Sigma$ s iâmbicos também ocorrem, em menor frequência (12,3%).

TABELA 5: Tipos de pés das hipersegmentações cl+ $\omega$

Tipos de pés	Exemplos	Ocorrências
(•)σ (* •)Σ	enquanto > em quanto	30 (52,7%)
(•)σ (• *)Σ	devagar > de vagar	07 (12,3%)
(•)σ (*)Σ	assim > a sim	20 (35%)

A análise métrica do tipo de hipersegmentações mais frequentes traz pistas de como a escrita não-convencional de palavras deixa entrever que a relação entre enunciados falados e escritos não é de interferência direta, mas de constituição, isto é, o falado constitui o escrito e vice-versa e essa constituição se dá, conforme CORRÊA (2004), entrelaçada pelas práticas orais e letradas. Esses dados permitem visualizar que o escrevente, em processo de aquisição da escrita, se baseia, também, na organização métrica do PB, a qual tem por característica a preferência por pés troqueus dissílabos (BISOL, 2000b).

Essa análise é corroborada pelas cinco hipersegmentações de palavras (que correspondem a 7,2% do total) que resultaram em possíveis palavras da língua (“disse deram”, “dele gato”, “tarde sinha”, “cala afrio”)<sup>12</sup> e que, por sua vez, também são constituídas de uma sequência de dois troqueus

<sup>12</sup> Vale lembrar que “tardezinha” é constituída de duas palavras fonológicas em função dos acentos em “tarde” e em “zinho(a)”.

dissílabos: (*tarde*)Σ (*sinha*)Σ. Além de registros de proeminências dos enunciados falados, destacamos que essas hipersegmentações também têm sua motivação ancorada em critérios semânticos, como observado por ABAURRE (1991) em dados de escrita infantil. Por meio desses dados, observam-se motivações quer de natureza prosódica, quer de natureza morfosemântica, além de haver motivações letradas, na medida em que é plausível supor – também – que os escreventes possam ter reconhecido registros de palavras da língua, representados por meio da distribuição do espaço em branco. Portanto, informações prosódicas e letradas simultaneamente motivam colocações não-convencionais do espaço em branco em palavras escritas.

Por fim, o último conjunto de dados é formado por quatro ocorrências (que correspondem a 5,8%) que têm em comum a segmentação da palavra em sílabas, quer pela presença não-convencional do espaço em branco – em “ou tra”, “es que cer”, “a o” (ao)<sup>13</sup> – quer pela presença não-convencional do hífen – em “jun-tos”. Exceto para essa última ocorrência, nota-se, para as demais, que ao menos uma das sílabas hipersegmentadas corresponde a um possível monossílabo átono (clítico) que é uma palavra funcional ou um item gramatical, como, nos dados coletados, a conjunção “ou” (em “ou tra”), o pronome “que” (em “es que cer”), a preposição “a” ou artigo “o” (em “ao”).

Concluímos a análise das hipersegmentações, tendo mostrado que elas oferecem simultaneamente evidências de que o escrevente mobiliza informações prosódicas (como o pé métrico ou a, palavra prosódica) e letradas (como a grafia de palavras funcionais). Essas mesmas características foram observadas em dados de escrita infantil por CHACON (2005) e CAPRISTANO (2007a). Identifica-se, assim, uma característica presente nos textos produzidos ao longo do EF.

<sup>13</sup> Observa-se que “ao” não é palavra fonológica, mas uma sequência de dois clíticos “a” (preposição) e “o” (artigo).

## Considerações finais

Neste artigo, ao tratar de evidências da organização prosódica do Português nos enunciados escritos, argumentamos que as segmentações não-convencionais de palavra podem ser motivadas por possíveis estruturas prosódicas da língua, como o pé métrico, a palavra prosódica e o grupo clítico. Mais especificamente, nas hipersegmentações, há evidência do pé troqueu dissílabo e, nas hiposegmentações, do grupo clítico (com predomínio de próclise). Responde-se, assim, à pergunta formulada na introdução deste texto: de quais constituintes prosódicos encontram-se evidências nos textos escritos.

No que diz respeito à aquisição da escrita, constatou-se que os dados de hiposegmentação predominam sobre os de hipersegmentação, característica comum aos dados de escrita infantil. Também foi mostrado que os constituintes prosódicos relevantes para caracterizar as hipo e hipersegmentações encontradas nos textos do EF-II diferem daqueles relevantes para os dados encontrados nos textos do EF-I, mais exatamente: no EF-I, podem ser observados todos os domínios prosódicos ( $U$ ,  $I$ ,  $\phi$ ,  $C$ ,  $\omega$ ,  $\Sigma$ )<sup>14</sup> e mesclas entre domínios (cf. CAPRISTANO, 2007a, e CHACON, 2005); já no EF-II, há predomínio de evidências de segmentações relacionadas a fronteiras de  $\Sigma$ ,  $\omega$  e  $C$  e não há dados de escrita que envolvam fronteiras de constituintes mais altos da hierarquia prosódica ( $U$  e  $I$ ).

Para além de serem identificadas semelhanças e diferenças, em termos de constituintes prosódicos, entre os dados de escrita produzida em diferentes momentos de EF, encontram-se pistas da natureza desses dados e de como se dão as relações entre fala e escrita. Por meio da análise empreendida, explicitamos em que medida e quais características dos enunciados falados são passíveis de serem observados por meio das segmentações não-convencionais de palavras nos enunciados escritos. Argumentamos que essas segmentações se constituem em dados linguísticos que evidenciam aspectos do processo de aquisição da escrita

---

<sup>14</sup> “U” indica enunciado fonológico; “I”, frase entoacional;  $\Phi$ , frase fonológica.

e, com base em CORRÊA (2004), podem ser vistas como marcas que resultam da conjunção das práticas orais/faladas e letradas/escritas e que permitem flagrar a presença do oral/falado no letrado/escrito. Defendemos, assim, que as características prosódicas das segmentações não-convencionais de palavra identificadas não corroboram a visão dicotômica das relações entre fala e escrita, mas sim a posição que toma a fala e a escrita como modalidades da língua que estão entrelaçadas com as práticas sociais orais e letradas de tal modo que essas relações são heterogeneamente constituídas.

## Referências

- ABAURRE, Maria Bernadete Marques. **A relevância dos critérios prosódicos e semânticos na elaboração de hipóteses sobre segmentação na escrita inicial.** *Boletim da Abralin*, Campinas, v. 11, p. 203-17. 1991.
- ABAURRE, Maria Bernadete Marques; SILVA, Ademar. **O desenvolvimento de critérios de segmentação na escrita.** Temas em psicologia, São Paulo. v.1, p. 89-102. 1993.
- ANDRADE, Carlos Drummond de. **A procura da poesia.** In: \_\_\_\_\_. *A rosa do povo*. 38 ed. Rio de Janeiro: Record, 2007 [1945]. p. 238.
- BISOL, Leda. **O clítico e o seu hospedeiro.** Letras de Hoje, Porto Alegre, v.40, n. 3, p. 163-184. 2005.
- \_\_\_\_\_. **Mattoso Câmara Jr. e a Palavra Prosódica.** DELTA, São Paulo. v. 20, n. especial, p. 59-70. 2004.
- \_\_\_\_\_. **O clítico e seu status prosódico.** Revista de Estudos da Linguagem, Belo Horizonte, v.9, n.1, p. 5-20. 2000a.
- \_\_\_\_\_. **O troqueu silábico no sistema fonológico.** DELTA, São Paulo. v.2, n.16, p. 403-413. 2000b.

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

CAGLIARI, Luiz Carlos. **Alfabetização e Linguística.** São Paulo: Scipione, 1989.

CAPRISTANO, Cristiane Carneiro. **Mudanças na trajetória da criança em direção à palavra escrita.** 2007a, 263f. Tese (Doutorado em Linguística Aplicada) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2007a.

\_\_\_\_\_. **Aspectos de segmentação na escrita infantil.** São Paulo: Martins Fontes, 2007b.

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

\_\_\_\_\_. **Hipersegmentações na escrita infantil:** entrelaçamento de práticas de oralidade e de 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, Manoel Luiz Gonçalves. (org.) Ensino de língua: representação e letramento. Campinas: Mercado de Letras, 2006. p.155-167.

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

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

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

CUNHA, Ana Paula Nobre; MIRANDA, Ana Ruth Moresco. **A influência da hierarquia prosódica em hipossegmentações da escrita de crianças de séries iniciais.** Revista Virtual de Estudos da Linguagem, Porto Alegre, v.1, p. 1-19. 2007.

FERREIRO, Emilia; PONTECORVO, Clotilde. **Os limites entre as palavras:** a segmentação em palavras gráficas. In: FERREIRO, Emilia. et alii. Chapeuzinho Vermelho aprende a escrever. São Paulo: Ática, 1996. p. 38-77.

GALVES, Charlotte Marie Chambelland; ABAURRE, Maria Bernadete Marques. **Os Clíticos no Português Brasileiro:** uma Abordagem Sintático-Fonológica. In: CASTILHO, Ataliba Teixeira de; BASÍLIO, Margarida. (Org.). Gramática do Português Falado, v. 4: Estudos Descritivos. Campinas: UNICAMP, 1996. p. 273-320.

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. 2007. 154f. Dissertação (Mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2007.

**Proposta curricular do Estado de São Paulo: língua portuguesa.** Coord. Maria Inês Fini. S.Paulo: Secretaria Estadual de Educação, 2008.

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

SIMIONI, Taise. **O clítico e seu lugar na estrutura prosódica em Português Brasileiro.** ALFA, Araraquara, v 52, n. 2, p. 431-446. 2008.

TENANI, Luciani. A grafia dos erros de segmentação não-convencional de palavras. *Cadernos de Educação*, Pelotas, v. 35, p. 247-269. 2010.

VIGÁRIO, Marina. **The prosodic word in European Portuguese.** Berlin/New York: Mounton Gruyter, 2003.

VIGÁRIO, Marina; FERNANDES-SVARTMAN, Flaviane Romani. **A atribuição de acentos tonais em compostos no português do Brasil.** In: *Atas do XXV Encontro da Associação Portuguesa de Linguística*, 2010, Porto. Porto: Tip. Nunes Ltda - Maia, v. 1. p. 769-786. 2010.