

Luciani Ester Tenani

**Prosódia e escrita: uma análise a partir de (hiper)segmentações de
palavra**

SÃO JOSÉ DO RIO PRETO
2016

Luciani Ester Tenani

**Prosódia e escrita: uma análise a partir de (hiper)segmentações de
palavra**

Tese apresentada como parte dos requisitos para obtenção do título de Livre-Docente em Fonologia, junto ao Departamento de Estudos Linguísticos e Literários, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto.

São José do Rio Preto
2016

Ao Gabriel e à Raquel

AGRADECIMENTOS

O desenvolvimento da pesquisa que resultou nesta tese teve contribuições de várias pessoas que, embora aqui anônimas, foram importantes em diferentes etapas do processo, como os escreventes dos textos que serviram de material para a pesquisa e os bolsistas TT-FAPESP que colaboraram na organização dos dados analisados.

Também contribuíram gentilmente colegas dos Departamentos de Estudos Linguísticos e Literários, de Letras Modernas e de Educação que se dispuseram a sugerir bibliografia sobre questões que são tratadas nesta tese.

Dentre aqueles que contribuíram mais diretamente para a elaboração da tese, nominalmente agradeço:

ao companheiro Fábio Villela pelo diálogo sobre questões da área de Educação e por seu apoio afetivo ao longo do processo de desenvolvimento da pesquisa e, especialmente, no processo de escrita deste texto;

às professoras Dras. Fabiana Komesu (UNESP/SJRP) e Larissa Berti (UNESP/Marília) pelos incentivos, sugestões e discussões sobre pontos específicos de questões tratadas nesta tese;

ao professor Dr. Lourenço Chacon (UNESP/Marília), líder do grupo de pesquisa “Estudos sobre a linguagem”, pelas muitas oportunidades de diálogo e reflexão sobre escrita, prosódia e teorias da linguagem;

ao professor Dr. Manoel Corrêa (USP), supervisor de estágio de pós-doutorado, pelos comentários e sugestões sobre escrita que embasam este trabalho.

Agradeço ainda:

aos professores doutores Raquel Fiad (UNICAMP), Ester Scarpa (UNICAMP), Manoel Corrêa (USP), Gladis Massini-Cagliari (UNESP/Araraquara) e Roberto Camacho (UNESP/São José do Rio Preto), pelas observações e sugestões feitas quando da defesa desta tese;

ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq –, pelo apoio financeiro concedido, na forma de Bolsa de Produtividade em Pesquisa, aos projetos: *Aspectos segmentais e prosódicos da escrita de crianças e adolescentes: evidências de relações entre enunciados falados e escritos* (Processo: 306.471/2009-4) e *Segmentação de palavras escritas: evidências de relações entre enunciados falados e escritos* (Processo: 309872/2012-0);

à Fundação de Auxílio à Pesquisa do Estado de São Paulo – FAPESP –, pelo apoio financeiro, na forma de auxílio regular à pesquisa, concedido aos projetos: *Aspectos segmentais e prosódicos da escrita de crianças e adolescentes: evidências de relações entre enunciados falados e escritos* (Processo: 2009/14.848-6) e *Hipersegmentação de palavras escritas: evidências de relações entre enunciados falados e escritos* (Processo: 2013/14546-5).

Resumo

Esta tese trata de segmentações não-convencionais de palavra, de uma perspectiva linguística, e as define como *palavras escritas cujas fronteiras gráficas não seguem as convenções ortográficas do português*, seja pela ausência de fronteiras entre palavras, como em “concerteza”, “ajudime” – exemplos de hipossegmentações –, seja pela presença de fronteiras dentro da palavra, como em “na quela”, “morava-mos” – exemplos de hipersegmentações. O objetivo geral é tratar da relação entre prosódia e escrita a partir de análises quantitativa de hipo e hipersegmentações e qualitativa de hipersegmentação que ocorrem em textos produzidos durante o ciclo II do Ensino Fundamental (EF II) no Brasil. É feita descrição da amostra longitudinal dos textos analisados, por se tratar de tipo de material pouco investigado em relação aos textos do ciclo I do Ensino Fundamental (EF I). Também são discutidos problemas metodológicos sobre identificação de segmentações não-convencionais em textos manuscritos, uma vez que não há consenso sobre essa metodologia. A partir da explicitação da metodologia proposta, destaca-se o papel de práticas letradas de leitura e escrita a respeito da categorização do espaço em branco entre palavras escritas. Da análise quantitativa, destacam-se os seguintes resultados: (i) hipo e hipersegmentações são mais frequentes em textos dos alunos do que nos das alunas, uma característica do EF II semelhante a do EF I; e (ii) as hipersegmentações, mas não as hipossegmentações, diminuem de frequência em função do aumento dos anos letivos, uma característica do EF II distinta do EF I. É feita, ainda, análise de características prosódicas, morfossintáticas e semânticas de palavras e dos enunciados em que hipersegmentações ocorreram e são identificadas características linguísticas das hipersegmentação no EF II. Dos resultados obtidos nessa análise, destacam-se: (i) a tendência em haver hipersegmentação de sílabas átonas quando essas potencialmente correspondem a clíticos prosódicos monossilábicos; (ii) o uso dos espaços entre sílabas em certas hipersegmentações para expressar silabação, um recurso para destacar palavras de um enunciado. Na interpretação desses resultados, argumenta-se que a correlação entre frequência de segmentação não-convencional de palavras e sexo/gêneros dos estudantes e, também, as diferenças entre as características da segmentação do EF II em relação às descritas do EF I sejam efeitos de sobreposições de representações sociais de sexo/gênero e de infância/adolescência, por um lado, com representações de práticas letradas/escritas e orais/faladas, por outro lado, que constituem os sujeitos da linguagem. Por fim, é feita análise discursiva de hipersegmentações pouco frequentes no EF II, considerando-se a organização prosódica do enunciado e as produções textuais em que essas ocorreram. No desenvolvimento da análise, é defendida a tese de que haja relação entre prosódia e escrita que se atualiza em cada produção textual e essa relação, observável por meio das hipersegmentações, é constituída de tensões entre possibilidades de representação do verbal e do verbo-visual da língua(gem). A principal contribuição desta tese é a caracterização da natureza complexa do processo de segmentação do texto em palavras, consideradas relações entre fala e escrita, oralidade e letramento, temas que interessam não apenas a pesquisadores das áreas de Linguística e Linguística Aplicada, mas também de áreas de Educação e Saúde.

Palavras-chave: palavra, escrita, ortografia, fala, prosódia, Língua Portuguesa, Ensino Fundamental.

Abstract

This thesis provides a linguistic perspective of unconventional word segmentation, which is defined as written words whose graphical boundaries do not follow the orthographic conventions of Portuguese. Two types are considered: (i) hyposegmentations, defined by the lack of boundaries between words, such as “concerteza” (surely), “ajudime” (help me); and (ii) hypersegmentations, which are characterized by the presence of boundaries within a word, such as “na quela” (those), “morava-mos” (we lived). The main objective is to address the relationship between prosody and writing based on a quantitative analysis of hypo- and hypersegmentations and a qualitative analysis of hypersegmentations which occur in texts written by students from Junior High School (EF II) in Brazil. A description of the longitudinal sample of analyzed texts is made, since it is a little-known type of texts if compared to texts from Elementary School (EF I). In the sequence, methodological problems of identifying unconventional segmentations in writings are discussed, since there is no consensus about this methodology. Following the explanation of the proposed methodology, we highlight the role of literacy practices of reading and writing in regard to the categorization of white space between written words. Based on the quantitative analysis, the main results are as follows: (i) hypo- and hypersegmentation occur more frequently in texts of male students than in texts of female students, a trace of EF II similar to EF I; and (ii) hypersegmentation, but not hyposegmentation, decreases in frequency throughout the school years, a characteristic which distinguishes EF II from EF I. Analysis of prosodic, morphosyntactic and semantic features of words and utterances in which hypersegmentations occur has been made and linguistic characteristics of hypersegmentations from EF II have been identified. The main results of this analysis are: (i) the tendency of targeting unstressed syllables in case they could correspond to monosyllabic prosodic clitics; (ii) the use of spaces between syllables in certain hypersegmentations to express syllabification, as a means to stress words in an utterance. The interpretation of these results leads us to the conclusion that the correlation between the frequency of unconventional word segmentation and sex / gender of the students and the differences between the segmentation characteristics of EF II compared to those of EF I are effects of social representations of sex / gender and childhood / adolescence, on one hand, overlapping with social representations of literacy / written practices and oral / spoken practices, on the other hand, which compose the subjects of language. Finally, a discourse analysis of uncommon hypersegmentation in EF II is made, considering the prosodic organization of the utterance and the textual production in which they occur. Throughout the development of analysis, it is argued that there is a relation between prosody and writing which is updated in every textual production and this relationship, evidenced by hypersegmentations, is established by tensions between possibilities of verbal and verbal-visual representations of language. The main contribution of this thesis is the characterization of the complex process of segmenting texts into words, considering relations between speech and writing, orality and literacy, a matter of interest for researchers not only from Linguistics and Applied Linguistics, but also from areas such as Education and Health.

Keywords: word, writing, spelling, speech, prosody, Portuguese, Junior High School.

Lista de Figuras

Figura 2.1: Grafias de “dificuldade” e “a minhas”	48
Figura 2.2: Grafia de translineação de “en contra”	49
Figura 2.3: Grafias “de mais”, “estuda-se” e “forma-se”	49
Figura 2.4: Grafias “denovo” e “persiguihos, perseguido”	50
Figura 2.5: Grafia de homonímia “por que”	52
Figura 2.6: Grafias “tem de” e “tem de”	53
Figura 3.1: Hierarquia prosódica segundo Nespor e Vogel (1986)	79
Figura 3.2: Representação da prosodização do clítico	83
Figura 3.3: Representação da configuração binária do pé métrico	84
Figura 3.4: Usos não-convencionais de hífen em formas verbais	110
Figura 4.1: Representação do contínuo de gêneros textuais na fala e na escrita segundo Marcuschi (2001)	118
Figura 4.2: Texto em que ocorrem “su seço”, “o portunidade”, “o bistracos”	125
Figura 4.3: Enunciado da proposta “Uma grande conquista”	127
Figura 4.4: Texto em que ocorrem “com ver sando”, “com versando”, “com ver samo”	132
Figura 4.5: Enunciado da proposta “Amizade e internet”	138

Lista de Gráficos

Gráfico 2.1: Índice de segmentação não-convencional de palavras por sujeito.....	58
Gráfico 2.2: Relação entre sexo/gênero e faixas de índice de segmentação	60
Gráfico 2.3: Número de palavras escritas em função de ano letivo e sexo/gênero	61
Gráfico 2.4: Índice de hipersegmentação em função de ano e sexo/gênero.....	63
Gráfico 2.5: Índice de hipossegmentação em função de ano letivo e sexo/gênero	64
Gráfico 3.1: Frequência de oxítonos, paroxítonos e proparoxítonos hipersegmentados <i>versus</i> número de sílabas das palavras	90

Lista de Quadros

Quadro 2.1: Usos não-convencionais do hífen.....	51
Quadro 3.1: Conjunto de fatores e variáveis considerados na análise prosódica.....	88

Lista de Tabelas

Tabela 2.1: Médias obtidas e metas do IDESP para alunos do 9º ano do EF.....	43
Tabela 2.2: Médias obtidas e metas do IDEB para alunos do 9º ano do EF	44
Tabela 2.3: Faixas de índices de segmentação não-convencional de palavras por sujeito	59
Tabela 3.1: Estruturas métricas e hipersegmentação.....	92
Tabela 3.2: Pés métricos e hipersegmentação de polissílabos.....	95
Tabela 3.3: Classe gramatical das palavras hipersegmentadas.....	101

Lista de Abreviaturas, Siglas e Símbolos

EF	Ensino Fundamental
EF I	Ensino Fundamental Ciclo I
EF II	Ensino Fundamental Ciclo II
PB	Português Brasileiro
PE	Português Europeu
σ	sílaba
Σ	pé métrico
ω	palavra prosódica
CG	grupo clítico
ϕ	frase ou sintagma fonológico
<i>I</i>	frase ou sintagma entoacional
<i>U</i>	enunciado fonológico

SUMÁRIO

Resumo	4
Abstract.....	5
Lista de Figuras	6
Lista de Gráficos.....	7
Lista de Quadros	8
Lista de Tabelas	9
Lista de Abreviaturas, Siglas e Símbolos	10
Apresentação	12
1 Abordagens sobre segmentações não-convencionais de palavra	17
1.1 Segmentação de palavras e perspectivas sobre erro ortográfico.....	18
1.2 Segmentação não-convencional de palavra como dado de escrita	27
2 Segmentações não-convencionais de palavra na escrita do EF II.....	39
2.1 Caracterização do material analisado.....	40
2.2 Constituição do objeto de análise	47
2.3 Caracterização de segmentações de palavra na escrita do EF II.....	56
3 Hipersegmentação e organização prosódica dos enunciados	74
3.1 Organização prosódica dos enunciados	74
3.2 Hipersegmentação e organização métrica da palavra	86
3.3 Hipersegmentação e organização morfossintática dos enunciados	99
4 Hipersegmentação, prosódia e sentidos dos enunciados	114
4.1 Abordagens sobre a relação entre fala e escrita	115
4.2 Hipersegmentação e relações textuais-discursivas	124
5 Considerações finais.....	143
Referências:	157
Anexos.....	169

Apresentação

Esta tese trata de segmentações não-convencionais de palavra que são aqui definidas por *palavras escritas cujas fronteiras gráficas não seguem as convenções ortográficas do português*. As fronteiras gráficas são de dois tipos: *espaço em branco*, caracterizado por ausência de marca gráfica entre palavras, e *hífen*, caracterizado por um traço entre palavras. As segmentações não-convencionais de palavra podem ser classificadas em dois tipos principais: *hipossegmentação*, quando estão ausentes as fronteiras gráficas entre palavras, como em “concerteza”, “ajudime”; e *hipersegmentação*, quando estão presentes as fronteiras gráficas dentro de uma palavra ortográfica, como em “na quela”, “morava-mos”. Serão analisadas quantitativamente as hipo e hipersegmentações em textos do Ensino Fundamental II (EF II), mas será aprofundada a caracterização da hipersegmentação por esse tipo de registro gráfico permitir tratar, de modo privilegiado, da relação entre prosódia¹ e escrita.

As hipo e hipersegmentações são predominantemente nomeadas “erros de segmentação de palavra”, o que implica, na maioria dos trabalhos, posicionamento normativo, o qual não é adotado nesta tese. Outra posição teórica é explicitada quando assumida a denominação “segmentação não-convencional de palavra”, pois envolve diálogo com as convenções ortográficas sem compactuar com a perspectiva que classifica, necessariamente, os registros escritos dos alunos em acertos e erros. Busca-se, de partida, explicitar a relevância desse diálogo com as convenções ortográficas, pois o material de análise são textos de crianças e adolescentes em ambiente escolar e que são avaliados por atenderem ou não a essas convenções ortográficas.

Antes de discorrer sobre como a relação entre prosódia e escrita será investigada nesta tese, é importante explicitar que a relevância da investigação das segmentações não-convencionais de palavra de uma perspectiva linguística está, principalmente, em sua natureza complexa enquanto dado de linguagem que interessa a especialistas de diferentes

¹ Prosódia é um termo, segundo Scarpa (1999, p.8), que recobre, no interior dos estudos linguísticos, “uma gama variada de fenômenos que abarcam parâmetros de altura, intensidade, duração, pausa, velocidade de fala, bem como o estudo dos sistemas de tom, entoação, acento, ritmo das línguas naturais”. A autora sugere, ainda, a existência de dois grandes polos de interesse no que se tem chamado de estudos prosódicos. O primeiro polo se interessaria por estudos voltados para o tratamento acústico, mensurável, instrumental de altura, intensidade e quantidade, correlatos perceptuais de frequência, volume e duração. De outro lado, existiria o interesse pelos estudos voltados para a consideração fonológica das organizações e representações dos sistemas de acento, ritmo, entoação nas línguas e suas interfaces com os demais componentes linguísticos. Nesta tese, a abordagem da prosódia filia-se ao segundo grupo de interesse, como será detalhado no capítulo 3.

subáreas da Linguística e da Linguística Aplicada, tais como Fonética, Fonologia, Aquisição da Linguagem e Psicolinguística, mais diretamente. As segmentações não-convencionais de palavra também se constituem em objeto de interesse de especialistas das áreas de Educação, como os da Psicopedagogia, e da Saúde, como os da Fonoaudiologia. Nesta tese, dessas áreas extraem-se abordagens dadas ao tema a fim de alargar o leque de interlocutores e, assim, as contribuições que esta tese possa alcançar junto a essas áreas de pesquisa que também têm a linguagem, particularmente os usos da escrita, como objeto de investigação.

Na área de Educação, o interesse nas segmentações não-convencionais de palavra decorre do fato de esses registros gráficos poderem ser interpretados, de certa perspectiva, como um tipo de “erro ortográfico”. Esses “erros” podem ser vistos como constitutivo do chamado processo de alfabetização, quando presentes em textos dos anos iniciais do EF, mas como um tipo de problema de escrita, quando atestados em textos dos anos finais do EF. Em Educação, o domínio da segmentação convencional de palavras é parte de um conjunto de habilidades da área de língua portuguesa que serve de índice da qualidade do processo de escolarização de crianças (no EF I) e de adolescentes (no EF II).

No primeiro capítulo desta tese, essa perspectiva da Educação será comentada a fim de ser demonstrado que, nesse campo de investigação, identificam-se abordagens distintas entre si que tomam esses registros ou como índice de déficit/defasagem do aluno ou como dado constitutivo do processo de alfabetização. A noção de “erro” é diferente em cada abordagem: em uma perspectiva, “erro ortográfico” compreende grafias que estão fora da norma ortográfica; noutra perspectiva “erro ortográfico” é evidência de um processo mais amplo de desenvolvimento cognitivo do sujeito. Em relação a essas abordagens, esta tese se distancia de ambas, embora dialogue de modo mais próximo com a segunda abordagem, como será explicitado na primeira seção do capítulo 1.

Ainda nesse primeiro capítulo, serão discutidas perspectivas de análise dadas aos erros ortográficos, de modo amplo, e às segmentações não-convencionais de palavra, em particular, no campo da Saúde. Nesse campo, tal como no campo da Educação, encontram-se perspectivas distintas que ou tomam essas grafias de palavra como índices de *disortografia*, um distúrbio específico da escrita, ou como evidências que sustentam um processo mais amplo de patologização da Educação. Essa caracterização quanto a abordagens dadas no campo da Saúde é especialmente relevante, porque permite problematizar a abordagem que interpreta a ocorrência dos chamados erros ortográficos no ciclo II do EF como índice de que os sujeitos que os produzem tenham problemas de

desenvolvimento cognitivo, por exemplo. Ao serem caracterizadas linguisticamente as segmentações não-convencionais de palavra, será argumentado que grande parte desses registros gráficos é motivada muito mais pela complexa relação entre prosódia e escrita do que pela suposta incapacidade cognitiva dos alunos do EF.

Ainda na seção 1.1, serão tratadas abordagens feitas no âmbito dos estudos linguísticos que tomam os chamados erros ortográficos (incluídas as segmentações não-convencionais de palavra) como efeitos da interferência da fala na escrita, motivados, segundo essas abordagens, pelas dificuldades dos alunos em compreenderem, explicitamente, variação linguística presente na fala, por um lado, e convenções ortográficas e normas gramaticais que devem ser observadas na produção de textos escritos, por outro lado. Desse modo, problematizam-se certas perspectivas linguísticas encontradas notadamente nas subáreas da Psicolinguística e da Sociolinguística, que, por sua vez, partilham de abordagens encontradas nos campos de Educação e Saúde sobre os chamados erros ortográficos.

Na segunda parte do capítulo 1, são identificadas abordagens que assumem os chamados erros ortográficos como indícios de processo de aquisição da escrita decorrentes ou do processo de aprendizagem, numa perspectiva construtivista, ou do processo de constituição do sujeito da linguagem, numa perspectiva discursiva. Explicitam-se, ainda, questões teóricas comuns a essas perspectivas no que tange, por exemplo, à relevância da noção de palavra e à organização prosódica dos enunciados para a caracterização das segmentações não-convencionais de palavra. Portanto, no percurso, trilhado no capítulo 1, por diferentes abordagens dadas aos chamados erros ortográficos, serão destacadas perspectivas com as quais esta tese não compactua e outras das quais se aproxima.

Uma vez traçado o cenário em que se insere esta tese, no capítulo 2, será feita a caracterização do material analisado, da constituição do objeto de análise e das segmentações não-convencionais de palavra encontradas no EF II. Na seção 2.1, é caracterizado o banco de textos escritos por alunos dos quatro últimos anos do EF a partir dos quais foram identificadas ocorrências de segmentação não-convencional de palavras em textos do EF II. A descrição desse material se faz importante por se tratar de tipo de material pouco estudado em relação aos materiais do EF I que, predominantemente, são investigados. Em seguida, na seção 2.2, são discutidos problemas metodológicos relevantes para a constituição das segmentações não-convencionais de palavra como dado de escrita, uma vez que não há consenso estabelecido na literatura sobre o conjunto de

passos metodológicos a serem observados nesse processo. Na seção 2.3, última desse capítulo, são apresentados resultados de natureza quantitativa que visam não apenas descrever as frequências de hipo e hipersegmentações de palavra na amostra longitudinal de textos do EF II, mas principalmente caracterizá-las frente às frequências de segmentações identificadas na chamada escrita infantil.

Após descrever e discutir os resultados quantitativos em termos de frequência na amostra de escrita analisada, será feita, no capítulo 3, descrição de características linguística das hipersegmentações, em particular. A restrição da análise desse tipo de segmentação não-convencional de palavras é embasada nos resultados quantitativos descritos no capítulo anterior a respeito de haver entre as hipersegmentações, mas não entre as hipossegmentações, diminuição de frequência em função do aumento dos anos letivos. Nesse capítulo, após apresentadas as bases teóricas que fundamentam a concepção fonológica da organização prosódica dos enunciados (seção 3.1), são feitas análises de características prosódicas da palavra (seção 3.2) e morfossintáticas dos enunciados (na seção 3.3) que possibilitam responder uma questão central desta tese: *quais características linguísticas particularizariam dados de hipersegmentação de palavra no EF II em relação aos dados produzidos no EF I?*

No capítulo 3, os resultados e as discussões permitem delinear que traços da escrita do EF I continuam a ser observados na escrita do EF II e que traços se constituíram como mais prototípicos daquilo que poderia ser denominado de *escrita infantojuvenil* que os distinguiriam da *escrita infantil*. Haveria, assim, características da escrita do EF II, ao menos quanto aos seus aspectos ortográficos, produzida por adolescentes (pertencentes à faixa etária de 11 a 15 anos) ao longo dos quatro últimos anos do EF que se particularizariam em relação à escrita do EF I, produzida por crianças (pertencentes à faixa etária de 6 a 10 anos) ao longo dos primeiros cinco anos do EF, no sistema vigente de Educação Básica no Brasil, à época da coleta dos textos (2008 a 2011). Uma proposta de distinção entre as produções escritas em cada um dos dois ciclos do EF não poderia, porém, ser sustentada em recortes de faixas etárias (que se constituíam em função de esperadas etapas do desenvolvimento cognitivo de crianças e adolescentes) ou em ciclos do EF (que se constituem a partir de organização curricular de conteúdos da Educação Básica previstos nos documentos governamentais, cf. BRASIL, 1998, 1997). A análise dos dados feita nesse capítulo visa sustentar a *hipótese de que características que particularizam a escrita do EF II em relação à escrita do EF I sejam efeitos dos diferentes*

modos pelos quais crianças e adolescentes são atravessados por práticas sociais letradas/escritas e, também, por práticas sociais orais/faladas.

A essa hipótese subjaz certa abordagem da noção de escrita e da relação entre fala e escrita. Nesta tese, *escrita é concebida como modo de enunciação*, conceituação proposta por Corrêa (2004), e não como modalidade da língua (ao lado da modalidade falada), como inicialmente defendeu Saussure (1973). Em consequência da assunção desse conceito de escrita, a relação entre fala e escrita é abordada em outras bases que não aquela que subjaz à abordagem que parte da distinção entre as materialidades fônica (da fala) e gráfica (da escrita) de realização da língua. Assume-se, com Corrêa (2004), a concepção de que a fala não é exterior à escrita, mas a constitui, seja considerada a materialidade da escrita alfabética, sejam consideradas práticas orais e letradas que constituem fala e escrita, como será argumentado no início do capítulo 4, na seção 4.1.

Dessa perspectiva da escrita, dados de segmentação não-convencional de palavras – incluídas hiper e hipossegmentações – serão tomados como objeto de análise que permitem responder à *segunda questão central desta tese: que evidências segmentações não-convencionais de palavras fornecem sobre como se constituem relações entre prosódia e escrita?* A hipótese que guiará a construção da resposta a essa questão, a ser apresentada na seção 4.2, *é a de que há relação entre prosódia e escrita, presente na constituição da escrita alfabética, que se atualiza em cada produção escrita e atravessa os sujeitos da linguagem de modo a revelar como se instauram tensões entre possibilidades de representação do verbal e do verbo-visual da língua(gem).*

Com base nas descrições e análises dos dados e dos textos em que esses dados ocorrem, conclui-se esta tese explicitando em que medida as reflexões desenvolvidas contribuem, por um lado, com a literatura linguística que trata das relações entre prosódia e escrita e, de modo mais amplo, das relações entre fala e escrita e, por outro lado, com os estudos das áreas de Educação e Saúde, que abordam o chamado erro ortográfico, especificamente ao ser caracterizada a escrita de crianças e adolescentes. Portanto, esta tese visa contribuir não apenas com estudos linguísticos acerca de possibilidades de abordagens sobre a relação entre fala e escrita, mas também com estudos das áreas de Educação e Saúde com base nas análises de dados e nas discussões teóricas desenvolvidas a partir das segmentações não-convencionais de palavra.

1 Abordagens sobre segmentações não-convencionais de palavra

Neste capítulo, é apresentado um quadro acerca das diferentes abordagens dadas aos chamados erros ortográficos, uma vez que segmentações não-convencionais de palavra são predominantemente vistas como um tipo de erro ortográfico. Sem pretensão de oferecer uma retrospectiva exaustiva dos estudos sobre o tema, opto por delinear os matizes das pesquisas que se encontram nas áreas de Educação, Saúde e Linguística, na seção 1.1, caracterizando o cenário em que esta tese se insere.

Por meio dessa caracterização de abordagens que configuram o terreno conflituoso em torno dos chamados erros ortográficos, é construída a perspectiva adotada nesta tese sobre as segmentações não-convencionais de palavra no EF II. Demarco distanciamento em relação a certos estudos da área de Saúde, especificamente da área de Fonoaudiologia Educacional, e da Educação, especialmente da área de Psicopedagogia, pois não partilho dessas áreas a abordagem predominante dada aos aspectos da ortografia da escrita do EF I.

Em relação à literatura da área de Linguística, inicialmente explico posicionamento crítico em relação à interpretação de erros ortográficos (incluídas as segmentações não-convencionais de palavra) como decorrentes de problemas de desenvolvimento cognitivo que afetariam a capacidade dos alunos em lidar com as convenções ortográficas (invariantes) frente à variação linguística, característica da fala, que engloba fatores fonético-fonológicos, morfossintáticos, semânticos e textuais.

Em seguida, na seção 1.2, passo a caracterizar o viés teórico adotado por meio da apresentação de questões e conceitos centrais para as pesquisas que mobilizam conceitos da área da Fonética e Fonologia na análise de segmentações não-convencionais de palavra. Explico, também, o lugar desta tese em relação aos estudos já realizados sobre segmentação não-convencional de palavra no EF II.

1.1 Segmentação de palavras e perspectivas sobre erro ortográfico

Nesta seção, os objetivos são caracterizar e sistematizar as principais abordagens para o estudo dos chamados erros ortográficos nas produções escritas nos anos iniciais do EF no Brasil, situando tensões entre perspectivas que permeiam e constituem o objeto desta investigação. Na configuração desse cenário, lanço luz especialmente a pesquisas das áreas de Saúde e Educação cujas perspectivas sobre erros ortográficos são compartilhadas por esta tese.

Inicialmente, cabe observar que, de uma perspectiva escolar, especificamente para os professores envolvidos com o ensino de português e com o processo de alfabetização, segmentar e juntar palavras é um tipo de erro ortográfico que merece atenção. Essas segmentações e junções podem ser, quando considerados textos produzidos nos anos iniciais do EF, índices de reflexão sobre possibilidades de registros das palavras (FERREIRO et al., 1996) ou, quando considerados textos produzidos nos anos finais do EF, índices de que o aprendiz apresenta problemas de alfabetização ou, mais particularmente, “distúrbio de aprendizagem específico de escrita” (ZORZI, 2006, 1997). Não tendo sido sanados no ciclo I do EF, os chamados erros ortográficos tendem a permanecer no ciclo II do EF. Essa constatação leva profissionais envolvidos com o processo de ensino e aprendizagem de escrita em ambiente escolar (como professores, coordenadores de escola) a diagnosticar que o aluno do ciclo II necessita, por exemplo, de aulas de reforço (para aprender o que não conseguiu nas aulas de língua portuguesa) e/ou de acompanhamento fonoaudiológico (para sanar problemas de distúrbios de aprendizagem).

De uma perspectiva clínica, particularmente da área da Psicopedagogia e da Fonoaudiologia Educacional, as junções e segmentações de palavras que fogem à convenção ortográfica podem ser tomadas como um dos “sintomas” de *disortografia*, conforme afirmam Fernández et al. (2010, p. 501):²

A caracterização da disortografia se dá pela dificuldade em fixar as formas ortográficas das palavras, apresentando como sintomas típicos a substituição, omissão e inversão de grafemas, alteração na segmentação de palavras, persistência do apoio da oralidade na escrita e dificuldade na produção de textos. (grifos acrescentados)

² Os autores tratam da disortografia “enquanto manifestação de alteração no processamento fonológico e ortográfico decorrente de condições determinadas genética e neurologicamente, como os transtornos de aprendizagem específico (dislexia do desenvolvimento) e global (distúrbio de aprendizagem)” (FERNÁNDEZ et al., 2010, p. 500)

A segmentação não-convencional de palavra, objeto deste estudo, é, de acordo com essa perspectiva clínica, um dos “sintomas” de um distúrbio específico da escrita, um tipo de manifestação de alteração no processamento fonológico que decorre de condições determinadas genética e neurologicamente. Portanto, investigar as hipo e hipersegmentações dialoga com estudos que tratam, de modo amplo, da relação entre mente e linguagem e, particularmente, desse distúrbio de aprendizagem, o qual é definido, por fonoaudiólogos, como “um padrão de escrita que foge às regras ortográficas estabelecidas convencionalmente, que regem determinada língua” (FERNÁNDEZ et al., 2010, p. 501).

Para essa perspectiva clínica, a interpretação de grafias não-convencionais de palavra tem por base a hipótese de que o “aprendiz” junta palavras ao escrever porque fala tudo junto, na medida em que se apoia na oralidade para produção do enunciado escrito. O resultado, na escrita, seria ocorrência de hipossegmentações, um tipo de transcrição fonética, como “concerteza” (“com certeza”). Depois de corrigido, o aluno passaria a separar demais as palavras por aplicar regras para segmentar o fluxo da fala em palavras escritas em contexto em que isso não é pertinente. Nesse caso, o resultado, na escrita, seria a ocorrência de hipersegmentações, um tipo de hipercorreção, isto é, um registro de aplicação de uma regra de ortografia em uma palavra em que não é pertinente, como “com versa” (“conversa”).

Nota-se que haveria um percurso de aprendizado, dessa perspectiva, que se caracterizaria pela construção da escrita a partir da fala, havendo uma interferência inicial esperada a qual deveria ser superada, na medida em que o aluno, inicialmente, tomaria a fala como base para escrever um enunciado e, em seguida, dominaria as regras da ortografia. Se o aluno não alcançar essa etapa final do processo de aprendizagem, continuando a produzir textos com hipo e hipersegmentações, consequentemente, essas grafias são classificadas como “erros ortográficos”, “sintomas” de que o aluno apresenta *disortografia*.

Esse distúrbio de aprendizagem, nessa perspectiva, é interpretado como resultado da “persistência do apoio da oralidade na escrita” (FERNÁNDEZ et al., 2010, p. 500). Fica pressuposto, nessa afirmação, que a escrita convencional é totalmente desvinculada da fala. Filia-se essa visão à concepção de que fala e escrita distinguem-se dicotomicamente, uma vez que fala realiza-se pelo meio fônico, isto é, fala compreende os sons de uma dada língua; e escrita se materializa pelo meio gráfico, isto é, escrita é a

representação gráfica de uma dada língua. Vista a escrita como forma de representação “imperfeita” da fala, a conclusão lógica é que a permanência no apoio na fala é a fonte dos erros ortográficos e de dificuldades de produção de textos. A solução para esses “problemas” estaria, ao que parece, na completa desvinculação da escrita da fala. Se tomada a escrita como autônoma da fala, estariam excluídas relações entre fala e escrita que os estudantes fazem equivocadamente.

Essa visão não se restringe ao campo da Fonoaudiologia, mas é partilhada por (ao menos parte de) profissionais da área da Psicopedagogia (NOBILE; BARRERA, 2009, dentre muitos outros), que veem nessas grafias elementos não apenas para encaminhamento dos alunos para terapia com fonoaudiólogos, como também para elaboração de programas de reforço ou aceleração destinados a atender alunos que escrevem textos com características da oralidade.

Verificado, dessa perspectiva escolar, que o aluno não domina a ortografia de palavras comuns, apresentando dúvidas para separar ou juntar palavras, conclui-se que, mais do que não saber escrever, o aluno não sabe português, sua língua materna. Ainda, a depender da frequência e do ano escolar em que o aluno se encontra, essas dúvidas ortográficas são concebidas como sintomas de disortografia, como problematizado anteriormente (podendo ser, até, interpretadas como evidências de problemas de desenvolvimento cognitivo (porque mobilizadas relações entre mente e linguagem), quando frequentes em textos de alunos dos anos finais do EF). Por meio dessa interpretação dada ao chamado “erro de ortografia”, identifica-se, também, certa concepção da relação entre fala e escrita, a qual guia, muito fortemente, o trabalho pedagógico do professor, o trabalho clínico do fonoaudiólogo e, ainda, determinadas políticas públicas de ensino e aprendizagem de língua portuguesa no país.

Portanto, de certa perspectiva da área da Educação e da Saúde, separar uma palavra em duas ou juntar duas palavras fora das convenções ortográficas tem sido interpretado: (a) como um erro de escrita decorrente de uma indesejada interferência da fala na escrita, (b) como um indício de que o aluno não segue o percurso esperado de desenvolvimento da linguagem escrita e (c) como um índice de distúrbio, que pode ser medicado e/ou submetido a procedimentos terapêuticos a fim de que seja superado/sanado.

Quanto à primeira interpretação, é predominante não apenas nas abordagens encontradas nas áreas de Educação e Saúde, mas também nas abordagens da área da Linguística, sobre as quais tratarei mais à frente. Nesta seção, interessa trazer à cena

inicialmente interpretações que envolvem discussões pedagógicas sobre o ensino da escrita que cruzam com abordagens clínicas dadas aos sintomas do distúrbio de aprendizagem.

Particularmente quanto à última interpretação clínica que é dada à segmentação não-convencional de palavra, vale trazer à discussão o arcabouço mais amplo em que está inserida. No cruzamento dos campos da Educação (especialmente, a Psicopedagogia) e da Saúde (principalmente, a Pediatria e a Fonoaudiologia Educacional), discussões são feitas sobre o processo de patologização da educação que diz respeito ao processo de transformação do espaço pedagógico em espaço clínico. Collares e Moysés (1994, p. 25), por exemplo, debatem criticamente o processo de medicalização que se define por “transformar questões não-médicas, eminentemente de origem social e política, em questões médicas, isto é, tentar encontrar no campo médico as causas e soluções para problemas dessa natureza”.

Nesse processo de medicalização da educação, patologias como disortografia e dislexia, que se configuram como distúrbios de aprendizagem que provocariam o fracasso escolar, são atribuídas às crianças que frequentam o EF I. Essa patologização tem levado “de um lado, à rotulação de crianças absolutamente normais e, de outro, a uma desvalorização crescente do professor, cada vez menos apto a lidar com tantas ‘patologias’ e ‘distúrbios’” (COLLARES; MOYSÉS, 1994, p. 29). Partilho dessa posição crítica das referidas autoras e dela me valho para assumir uma perspectiva que não toma como sintomas de disortografia as segmentações não-convencionais de palavras nos textos do EF II, objeto de análise.

Esta tese, portanto, ao tratar de grafias que servem de índices para diagnosticar alunos do EF II como portadores de distúrbios de aprendizagem, põe-se em diálogo não apenas com profissionais de Educação, mas também com profissionais da Saúde que são mobilizados para diagnosticar os problemas de escrita de crianças e adolescentes. Nesse diálogo, duas posições se confrontam: por um lado, há educadores, médicos e fonoaudiólogos que defendem que os alarmantes índices de déficit da educação básica (como os dados nacionais do IDEB – Índice de Desenvolvimento da Educação Básica) sejam decorrentes dos distúrbios de aprendizagem (provocados ou não por subnutrição infantil ou por Transtorno de Déficit de Atenção e Hiperatividade,³ por exemplo) e, por

³ Divulgação ao grande público acerca dos preocupantes índices de diagnóstico de Transtorno de Déficit de Atenção e Hiperatividade (TDAH) e de consumo de cloridrato de metilfenidato (conhecido no mercado como Ritalina) tem sido feita por reportagens em jornais de grande circulação nacional, como *Folha de São*

outro lado, há profissionais dessas mesmas áreas que assumem postura crítica diante dos crescentes índices⁴ de diagnósticos desses distúrbios de aprendizagem que levam à constatação de que o “espaço escolar, voltado para a aprendizagem, para a normalidade, para o saudável, transforma-se em espaço clínico, voltado para os erros e distúrbios.” (COLLARES; MOYSÉS, 1994, p. 31). Nesse cenário, essa tese partilha dessa postura crítica frente à interpretação dos chamados erros ortográficos como sintomas de distúrbios de aprendizagem. Argumentos que sustentam este posicionamento serão dados, mais detalhadamente, nos capítulos 3 e 4, a partir da análise de segmentações não-convencionais de palavra em amostra longitudinal de textos do EF II.

Outra importante problematização relacionada ao tema de desempenho dos alunos na escola, diz respeito aos índices de mensuração de incidência dos erros ortográficos como parte das habilidades de escrita, avaliadas, por exemplo, pelo Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (SARESP).⁵ No Estado de São Paulo, resultados do SARESP compõem o IDESP, Índice de Desempenho da Educação do Estado de São Paulo. Esse índice, tido pela Secretaria da Educação como “um dos principais indicadores da qualidade do ensino paulista”, “estabelece metas que as escolas devem alcançar ano a ano” e, ainda, “compõe o bônus por desempenho pago aos servidores da Educação do estado” (cf. <http://www.educacao.sp.gov.br/saresp>).

O IDESP, em âmbito estadual, ao lado do IDEB, em âmbito nacional, são exemplos de índices gerados a partir de avaliações de desempenho dos alunos para uso nas políticas públicas, com o objetivo de regular os sistemas de ensino, segundo Almeida, Dalben e Freitas (2013, p. 1155). Para esses pedagogos, o IDEB, por exemplo, não contempla o nível socioeconômico dos alunos em sua composição, um fator que afeta sobremaneira os resultados de eficácia escolar. Os autores exploram dados longitudinais, oriundos de 312 escolas de cinco grandes cidades brasileiras, relativos a medidas de

Paulo (<http://www1.folha.uol.com.br/cotidiano/2015/10/1696475-governo-quer-protocolo-para-conter-uso-de-ritalina-por-criancas.shtml>), ou de circulação regional, como o *Diário da Região*, de São José do Rio Preto (SP), cidade onde residiam os alunos cujos textos foram analisados nesta pesquisa (<http://www.diariodaregiao.com.br/cidades/cresce-o-uso-de-medicamento-para-domar-criancas-e-adolescentes-1.352713>). Essas reportagens igualmente divulgam medidas do governo federal para melhorar o diagnóstico de TDAH e alertam para problemas de saúde decorrentes do uso abusivo de medicação por crianças e adolescentes possivelmente saudáveis.

⁴ Na página web da Associação Brasileira de Dislexia (<http://dislexia.org.br/v1/index.php/health-living-c/49-page-break-example>) é dada informação de que o maior número de crianças atendidas decorre de distúrbio de aprendizagem.

⁵ O SARESP é aplicado pela Secretaria da Educação do Estado de São Paulo com a finalidade de “produzir um diagnóstico da situação da escolaridade básica paulista”, conforme informado na página oficial da secretaria: <http://www.educacao.sp.gov.br/saresp> (Acesso em: 10 jan. 2016).

proficiências em leitura e matemática de mais de 35 mil alunos. A partir dessas medidas, argumentam haver “relação explícita entre desempenho médio, nível de entrada de alunos e seu NSE [nível socioeconômico]” (ALMEIDA; DALBEN; FREITAS, 2013, p. 1165) e demonstram que, “mesmo em situações mais adversas, em muitas escolas públicas, os alunos se desenvolvem (observando o valor agregado e não o desempenho médio) tanto quanto em algumas escolas privadas” (p. 1165).

Em síntese, os dados analisados por Almeida, Dalben e Freitas (2013) corroboram certa posição crítica de pedagogos em relação ao uso de indicadores baseados em testes (como o IDEB e o IDESP) de desempenho de alunos como únicos parâmetros de avaliação escolar, desconsiderando-se variáveis sociais que influenciam o trabalho da escola e problemas relativos à constituição do próprio instrumento de avaliação. Esta tese partilha também dessa posição crítica acerca da política educacional (nacional e estadual) que usa de indicadores de desempenho de alunos como parâmetros de avaliação escolar, não apenas pelos argumentos apresentados pelos pedagogos citados, mas principalmente por não comungar com a abordagem da relação entre fala e escrita que subjaz a esses testes.

Nos testes de avaliação nacional ou estadual (como será demonstrado no próximo capítulo), a não observação das convenções ortográficas por parte dos alunos (ou seja, os chamados erros ortográficos) é mensurada como parte das habilidades, esperadas para alunos do EF II, em estabelecer relação entre fala e escrita (cf. seção 2.3). A forma de avaliação dessa habilidade, bem como a conceituação da relação entre fala e escrita que subjaz aos testes, tem como traço desconsiderar relações entre fala e escrita, bem como práticas orais e letradas, relevantes para se interpretar a presença dos chamados erros ortográficos nos textos do EF II, como será demonstrado detalhadamente no capítulo 4.

As abordagens ora sumarizadas da área de Educação acerca dos índices de desempenho dos alunos em ambiente escolar e da área da Saúde acerca dos diagnósticos de distúrbio de aprendizagem vinculados aos desempenhos desses alunos permitem caracterizar um cenário de tensões entre abordagens em cada uma das áreas consideradas. Essas tensões decorrem, muito fortemente, de posicionamentos teóricos que consideram ou não aspectos socioeconômicos e/ou sócio-históricos na caracterização de problemas de leitura e escrita de modo amplo ou, de modo específico, na caracterização dos

chamados erros ortográficos, que abrangem, dentre outros erros, as segmentações não-convencionais de palavras, objeto desta tese.⁶

No âmbito dessa problematização, feita no campo da Educação acerca de instrumentos de avaliação de desempenho dos estudantes em língua portuguesa (e matemática) como medidas objetivas e eficazes, interessa construir outra problematização que põe em questão abordagens feitas no campo da Linguística, como a de Scliar-Cabral (2003) e seguidores, que interpretam erros ortográficos (incluídas as segmentações não-convencionais de palavras) em textos do EF como evidências (seguras, porque facilmente identificáveis) de baixo índice de aprendizagem, subsidiando, inclusive, generalizações negativas quer sobre o desenvolvimento cognitivo (e linguístico) de estudantes, quer sobre o sistema educacional brasileiro.

Nesse cenário no campo da Linguística, o objeto desta investigação pode ser considerado um dado de fracasso escolar (pois produzidos por alunos de fraco desempenho nos testes de avaliação) e, também, um dado de patologia (pois produzidos por adolescentes com baixo desenvolvimento cognitivo). No entanto, nesta tese, outra perspectiva será defendida, sustentada em argumentos a serem apresentados nos capítulos 3 e 4. Nesses capítulos, por meio da análise de segmentações não-convencionais de palavras, argumento que essas segmentações se constituem em dado de língua e de discurso, não sendo, pois, dado “marginal”, porque produzidos por alunos que supostamente “ficam à margem” das convenções escritas e do desenvolvimento cognitivo.

Ainda sobre os instrumentos de avaliação de desempenho dos alunos no EF, faz-se importante observar que, no bojo de políticas públicas (em âmbito estadual, nacional e internacional), os governos estadual e nacional têm se valido desses instrumentos para mensurar a qualidade da Educação (da escola, do estado, do país), e para legitimar as conhecidas desigualdades socioeconômicas sob o viés da objetividade do índice de avaliação.

A consideração dessa problematização de natureza política raramente é mencionada nos estudos linguísticos que analisam fatores motivadores dos erros ortográficos. Os registros não-convencionais considerados nas pesquisas de cunho linguístico são tomados, predominantemente, como “dados de língua escrita” que têm sua

⁶ Nesses trabalhos, no entanto, não se problematiza a concepção de escrita que subjaz à definição de erro ortográfico. Discussão a esse respeito será feita no capítulo 4.

motivação nas não coincidências entre fala e escrita. Parte dessas não coincidências decorre da variedade linguística falada pelos alunos oriundos de classes sociais cujas práticas orais e letradas não são aquelas valorizadas socialmente. Essa abordagem subjaz a várias pesquisas que ajustam o foco de investigação à caracterização de fatores linguísticos (relacionados à fala) e ortográficos (relacionados ao sistema de escrita alfabético) como principais motivadores dos chamados erros ortográficos.

Exemplo dessa perspectiva se encontra em trabalhos feitos no âmbito da Sociolinguística, como o de Freitag et al. (2014), que se caracterizam por interpretar distúrbios de aprendizagem, incluídos os relativos à leitura e à escrita, como tendo sido motivados por dificuldades dos alunos em lidar com a complexidade intrínseca à estrutura linguística, de natureza variável, e com a tecnologia do sistema de escrita alfabética, que representa, por meio de convenções ortográficas, apenas parte do que é enunciado.

Os “distúrbios de aprendizagem” têm sido mais “recorrentes” em crianças em processo de alfabetização. As dificuldades de leitura apresentadas pelas crianças que compõem o ensino fundamental menor são enormes: elas não entendem, por exemplo, por que falam de um jeito e escrevem de outro. Como falante da língua que é, essa criança já percebeu que existem diferenças entre a fala e a escrita e também que existem modos diferentes de se falar a mesma coisa, ou seja, há variação linguística. Por conta disso, cabe ao professor estar atento à variedade linguística que seu aluno traz de casa, uma vez que a distância entre o oral e o escrito será algumas vezes muito grande, cabendo explicações específicas, pois, embora o sistema escrito seja um só para todos, a diversidade reina na fala. (FREITAG et al., 2014, p. 49-50, grifos acrescentados)

A descrição de fatores linguísticos motivadores dos erros ortográficos contribui para explicitar a complexidade das grafias analisadas enquanto dado linguístico. Uma das soluções propostas para levar ao desejado decréscimo de incidência desses erros ortográficos é proporcionar aos alunos e professores reflexão sobre variação linguística, em contraste com a fixidez da ortografia do português. Outra solução está na explicitação das relações entre realização fonética e representação ortográfica, formalizada pela chamada consciência fonológica (SCLIAR-CABRAL, 2003; CAPOVILLA e CAPOVILLA, 2000), como explicitam Freitag et al. (2014, p. 50):

Esses aspectos [linguísticos] demonstram a necessidade de se trabalhar a consciência fonológica no ciclo infantil. Consciência fonológica, de acordo com Scliar-Cabral (2009), não é um método, mas uma estratégia de trabalho que facilita a aprendizagem da leitura pela criança. O aluno precisa aprender que “as palavras são formadas por pedacinhos representados por letras. Mudando uma pela outra, muda o seu significado” (SCLIAR-CABRAL, 2003, p. 40). Nessa perspectiva, a consciência fonológica incide na estratégia de trabalho consciente que leve a criança (alfabetizando) a desenvolver a capacidade de desmanchar a sílaba em pedacinhos menores. É a capacidade de fazer o alfabetizando

perceber os sons (valor) de cada letra/grafema, em cada composição silábica, através do reconhecimento dos traços que diferenciam as letras entre si.

Dessa perspectiva, a consciência fonológica promoveria reflexão sobre a relação entre fala e escrita que levaria o alfabetizando, na etapa inicial do EF, a apreender as convenções ortográficas. A consciência fonológica somada à chamada consciência morfológica (MOTA et al., 2009) e sintática (CAPOVILLA; CAPOVILLA, 2006), a serem desenvolvidas ainda no EF I, promoveriam reflexão metalinguística, sobre relações entre fala e escrita, especificamente, que sanariam dificuldades ortográficas típicas do EF I, além de evitar a recorrência de erros ortográficos no EF II, incluídos os analisados nesta tese. Essa perspectiva, amplamente difundida no país, ancora-se em bases biopsicológicas da linguagem, como explicita Scliar-Cabral (2013, p. 277) já na primeira afirmação que abre seu texto: “O fundamento epistemológico deste capítulo reside em reconhecer a base biopsicológica da cultura e, portanto, da linguagem”.

Esse posicionamento teórico, desenvolvido no âmbito da Psicolinguística, dialoga mais fortemente com as abordagens promovidas pelas pesquisas na área de Saúde, tais como as referenciadas no início desta seção. Em comum, essas abordagens sobre os erros ortográficos, de modo específico, e da relação entre fala e escrita, de modo amplo, apostam na identificação de explicações advindas das neurociências para mapear processamentos neurológicos envolvidos na alfabetização e leitura, entendida a escrita como código e a leitura como decodificação.⁷ O domínio desses processamentos, por parte dos alfabetizadores, por exemplo, implica adotar estratégias, em sala de aula, que visam à superação de dificuldades (cognitivas) envolvidas no processo de leitura e escrita identificadas no EF.

Esta tese não partilha dessa perspectiva, porque elege como relevantes (potenciais) relações entre características (linguísticas) que sejam motivadoras dos chamados erros ortográficos e modos pelos quais falantes/ouvintes, escreventes/leitores usam fala e escrita imersos em práticas sociais orais e letradas que os constituem enquanto sujeitos da e na linguagem. Na próxima seção, explicito esse viés de análise a ser dado às chamadas grafias não-convencionais de palavras, situando esta proposta em relação aos estudos que tratam da segmentação não-convencional de palavra de modo específico.

⁷ Na seção 4.1, será discutida essa concepção de escrita e de leitura.

1.2 Segmentação não-convencional de palavra como dado de escrita

Nesta seção, caracterizo o conjunto de estudos realizados sobre segmentação não-convencional de palavra segundo abordagens linguísticas com o objetivo de situar esta tese nesse conjunto e, desse modo, explicito as bases comuns a pesquisas feitas no campo da Linguística, particularmente, na área de Fonética e Fonologia, que tomam essas grafias como evidências de que unidades ritmo-entoacionais da fala é, em alguma medida, plasmada na escrita (ABAURRE, 1988, 1991a, b; ABAURRE; SILVA, 1993; SILVA, 1991) e, também, como pistas do trabalho do sujeito com a linguagem (CAPRISTANO, 2007a, b, 2004; CHACON, 2005, 2006).

Essa abordagem linguística dada às segmentações não-convencionais de palavra de que esta tese partilha nasce de reflexões sobre temas mais amplos acerca da alfabetização, feitas na década de 1980 no Brasil,⁸ período marcado pela proliferação de propostas alternativas ao ensino da escrita feito, até então, com base em cartilhas.⁹ As várias propostas buscavam enfrentar, principalmente, os baixos índices de alfabetização, somados a altos índices de reprovação no EF, verificados em todo o país.

Nessa mesma década, outro marco importante para os estudos sobre alfabetização, em particular, e o ensino da escrita, de modo amplo, foi a publicação, no Brasil, de *Psicogênese da Língua Escrita*, por Emília Ferreiro e Ana Teberosky. A pesquisa, feita no campo da Psicologia, de linha cognitivista, baseada em Jean Piaget, consiste na explicitação do processo de aquisição da escrita, dando-se ênfase à capacidade cognitiva de aprendizagem da criança.¹⁰ Muitas pesquisas, desde então, são desenvolvidas (no Brasil, na Argentina, no México, principalmente) com o objetivo de determinar o que a

⁸ Exemplos dessas reflexões encontram-se nos Anais do Seminário Multidisciplinar de Alfabetização (1984), realizado pela Pontifícia Universidade Católica de São Paulo em parceria com o Instituto Nacional de Estudos e Pesquisas Educacionais, do Ministério de Educação e Cultura, em 1983.

⁹ Vale mencionar que essa proliferação de propostas decorre, fortemente, de esse período ser caracterizado pela chamada redemocratização da educação, após o fim da ditadura militar, fato que tem como consequência a volta de exilados, dentre eles Paulo Freire, o que fomenta não apenas reflexões sobre os rumos da educação no país, mas chega à implantação de propostas em redes públicas de ensino. À frente da secretaria da Educação na cidade de São Paulo, o pedagogo teve oportunidade de implementar ações baseadas em sua proposta de educação que coloca o aluno – seja adulto ou criança – como sujeito participante do processo de alfabetização (cf. FREIRE, 1991).

¹⁰ Dentre outros aspectos importantes, essa tese teve como efeito deslocar a questão central da alfabetização dos métodos de ensino (sintético, analítico, global etc.), por meio dos quais o professor desempenhava papel principal, para o processo de desenvolvimento cognitivo que leva à aprendizagem (cf. FERREIRO; TEBEROSKY; 1999, p. VIII), em que a criança passa a ter papel principal. Embora não fosse um método de ensino, os conceitos e as reflexões propostas por Ferreiro e Teberosky (1999) deram origem ao método construtivista.

criança já sabe sobre escrita quando entra na escola (na época, obrigatoriamente no primeiro ano do EF) e que estratégias (cognitivas) ela usa para desenvolver suas “habilidades” de leitura e escrita.

Pesquisas interessadas no processo de aquisição da escrita proliferaram não apenas na área da Psicologia ou da Educação, mas também na área da Linguística, notadamente, na subárea da Aquisição da Linguagem, uma vez que são temas centrais nessa subárea investigar a aquisição da língua materna e a aquisição da escrita (cf. SCARPA, 2001).¹¹ Também se verificaram pesquisas sobre a mesma temática na subárea da Psicolinguística, que se particulariza por investigar relações entre organização de sistemas linguísticos e organização de pensamento e, naturalmente, caracteriza-se como um campo interdisciplinar, entre Psicologia e Linguística (cf. BALIEIRO JR., 2001).

Os estudos sobre aquisição da escrita, em particular, feitos na segunda metade da década de 1980 no Brasil, têm traços dessa interdisciplinariedade, além de traços de diversidade de abordagens teóricas advindas tanto da Linguística, quanto da Psicologia. Dentre a diversidade de abordagens, identificam-se, por exemplo, pesquisas ancoradas nos princípios da Teoria Gerativista, que define como questão central para os estudos linguísticos explicar a competência linguística (que compreende investigar relação entre mente e linguagem, cf. CHOMSKY, 2009), ao lado de pesquisas que mobilizam fundamentos da abordagem Sócio-interacionista¹² ou da Teoria Histórico-Cultural, baseadas nos estudos de Vygotsky (1988) sobre a relação entre pensamento e linguagem (cf. SCARPA, 2001).

Em meio a essa diversidade de pesquisas desenvolvidas sob várias vertentes teóricas interessadas em explicar a aquisição da linguagem, a escrita produzida na fase inicial do Ensino Fundamental ganhou o interesse de linguistas que investigavam aspectos fonéticos e fonológicos do português, por exemplo, pois as produções escritas iniciais se mostraram como fonte de dados relevantes para as teorias linguísticas, de modo amplo, e para as teorias fonológicas, de modo específico, como argumenta Abaurre (1988a, 1996).

¹¹ Mais especificamente, dados da aquisição da linguagem são fonte empírica importante para validar teorias linguísticas ou fazer avançar modelos teóricos que formalizam como crianças aprendem a falar sua língua materna e/ou uma segunda língua.

¹² Lemos (1984) explicita o que entende como avanços da abordagem piagetiana sobre alfabetização feita por Emília Ferreiro e seguidores e critica essa proposta, tomando por base noções advindas dos estudos conduzidos por Vygotsky, para defender uma abordagem sócio-interacionista da aquisição da escrita.

Nesse cenário, explicitam Abaurre et al. (1984) e Abaurre (1988a) a relevância da produção escrita realizada espontaneamente, pois, nesse tipo de texto, as crianças registrariam suas hipóteses sobre a escrita, a partir de relações que estabelece entre fala e escrita.¹³ Ao tomar o texto espontâneo como objeto de análise, uma das questões de investigação passou a ser a segmentação do texto em palavras. O interesse por esse aspecto gráfico voltou-se, predominantemente, para a descoberta de quais critérios linguísticos guiariam as crianças no reconhecimento dos limites de palavras escritas.

Interessada nos dados de segmentação não-convencional, Abaurre (1988b) sinaliza a importância da relação entre fala e escrita, para compreensão das hipóteses infantis acerca dos limites de palavra. A autora posiciona-se contrariamente à crença de que, nas primeiras produções escritas, as crianças tenderiam a escrever do modo como falam (visão que não foi completamente superada nos dias atuais) e afirma que “*ninguém se programa para simplesmente ‘escrever como fala’*”, seja em termos de uso das letras e das suas relações com o som, seja em termos de *segmentação*, seja em termos das próprias estruturas linguísticas” (ABAURRE, 1988b, p. 137, destaques acrescentados).

Comungar dessas considerações de Abaurre (1988b) sobre a relevância de dados não-convencionais de escrita, como os de segmentação de palavra aqui investigados, implica mobilizar teorias linguísticas que tratem da complexa relação entre fala e escrita. Abordagens sobre essa relação serão discutidas no capítulo 4 (seção 4.1), visando não apenas traçar características de cada uma, mas, especialmente, argumentar a favor da abordagem sobre a relação entre fala e escrita, adotada nesta tese, proposta por Corrêa (2004) e discutida no capítulo 4.

Essa abordagem sobre a relação entre fala e escrita embasa as pesquisas sobre segmentação não-convencional de palavras com as quais esta tese se filia, desenvolvidas (e orientadas) por Chacon (2004, 2005, 2006) e Capristano (2004, 2007a,b), a partir de textos produzidos no EF I, e Tenani (2010, 2011a,b, 2013), a partir de textos produzidos no EF II. Essas pesquisas, bem como esta tese, interpretam as segmentações não-

¹³ Vale destacar que Abaurre et al. (1984) defendem a importância da produção espontânea da criança em oposição à produção de ditados ou cópias (predominantes no início da alfabetização em razão do uso de cartilhas): “As crianças da primeira série normalmente não produzem textos espontâneos. A relação que elas estabelecem com a escrita é, via de regra, extremamente artificial, porque se pressupõe que, uma vez que ainda não dominam a convenção ortográfica, elas não são ainda capazes de desenvolver, com a escrita, atividades significativas. Seus exercícios costumam ser, assim, absolutamente controlados pela professora, que reproduz a orientação pedagógica vigente. [...] Pode-se dizer, portanto, que a partir do seu primeiro contato com a escrita as crianças são submetidas de forma sistemática a uma série de exercícios que a distanciam progressivamente da noção de *texto*, para elas tão natural em termos de produção oral, quando ingressam na escola” (ABAURRE et al., 1984, p. 25, destaque no original).

convencionais de palavra como evidências do *modo heterogêneo de constituição da escrita* (Corrêa, 2004), como será demonstrado no capítulo 4 (seção 4.2).

As segmentações não-convencionais de palavra são também importantes para tratar de questões que interessam às teorias fonológicas. Segundo Abaurre (1991a), as grafias das fronteiras de palavras podem ser indícios importantes na “validação das unidades prosódicas propostas nos modelos fonológicos não-lineares como constitutivas de uma hierarquia que reconhece vários domínios prosódicos como significativos em termos das representações fonológicas subjacentes” (p. 204). Sob esse viés investigativo, tem sido desenvolvido grande parte dos estudos dos grupos de pesquisa *Estudos sobre a Linguagem* (GPEL/CNPq) e *Estudos sobre a Aquisição da Linguagem Escrita* (GEALE/CNPq).¹⁴

Destaca-se que os trabalhos realizados por ambos os grupos têm demonstrado – com base em bancos de textos escritos por crianças, adolescentes e adultos – em que medida grafias não-convencionais de palavra¹⁵ são reguladas, de modo complexo, por estruturas fonológicas que não se restringem aos constituintes sílaba e palavra (prosódica). Em Silva e Tenani (2014, p. 23-51), encontra-se resenha detalhada do conjunto desses trabalhos de modo que ficam evidenciados os avanços alcançados sobre a caracterização de estruturas fonológicas relevantes para compreender o processo de aquisição da escrita, notadamente, na fase delimitada pelos anos iniciais do EF.

Dos estudos realizados, cabe retomar aqueles em que são traçadas relações entre características da segmentação não-convencional no EF I e as do EF II, tendo essas últimas como ponto de partida para comparação com as primeiras. Em Tenani (2011a,b), encontram-se análises de segmentações não-convencionais de palavra identificadas no EF II seguidas de comparação entre os resultados sobre segmentações não-convencionais de palavra no EF I descritos, principalmente, por Capristano (2007a,b) e Chacon (2005, 2006), por esses autores partilharem dos mesmos pressupostos teóricos, quer quanto ao modelo da Fonologia Prosódica (a ser detalhado na seção 3.1) mobilizado na análise da

¹⁴ Sobre segmentação não-convencional de palavras e sua relação com características fonológicas, encontram-se, no âmbito do GEALE, os seguintes trabalhos: Cunha (2004, 2010), Ferreira (2011), Cunha e Miranda (2013). No âmbito do GPEL, além dos trabalhos de Capristano (2004, 2007a, b), Chacon (2004, 2005, 2006), Tenani (2004, 2009, 2010, 2011a,b) já citados, foram produzidos artigos e dissertações: Tenani e Longhin (2015), Capristano e Ticianel (2014), Silva e Tenani (2014), Paranhos (2014), Tenani e Paranhos (2011), Silva (2011), Paula (2007).

¹⁵ Cabe explicitar que, no âmbito do GEALE, é mobilizado o conceito de *erro* definido com base na teoria piagetiana, segundo a qual o erro (de fala ou de escrita) é parte do processo de aprendizagem. Nessa perspectiva, erros evidenciam hipóteses construídas pelos aprendizes na sua relação com o objeto a ser aprendido. Essa noção de erro é distinta daquela mobilizada por Zorzi (1997), para quem erro é sintoma de distúrbio de aprendizagem, como já discutido na seção anterior.

organização prosódica dos enunciados, quer quanto à concepção da relação entre fala e escrita enquanto modos de enunciação (abordagem caracterizada na seção 4.1).

Quando da análise de textos transversais produzidos no sexto ano do EF, foi identificada, por Tenani (2011a), predominância, dentre os casos de hipersegmentação, de a sílaba pretônica de uma palavra ortográfica ser analisada como um clítico, como “na quele”. Já dentre os casos de hipossegmentação, foram identificadas duas estruturas principais, a saber: (i) um clítico foi grafado junto à palavra fonológica que lhe seguia, como “meajuda” e (ii) um clítico foi grafado junto à palavra fonológica que lhe precedia, como “ajudime”. Em comum, as hipo e hipersegmentações envolvem, predominantemente, grafia de clíticos fonológicos.¹⁶ No conjunto dos dados analisados, foram identificados como mais relevantes os constituintes ω e *CG*,¹⁷ sendo a flutuação entre um e outro constituinte traço característico das segmentações não-convencionais: quando houve hipersegmentação, a ω foi analisada como se se tratasse de *CG*, como em “da quele”, “em quanto”; quando houve hipossegmentação, um *CG* foi analisado como se fosse uma ω , como em “porfavor”, “pegalo”.

Essa flutuação entre hipo e hipersegmentação em textos de alunos do sexto ano do EF foi interpretada como tendo sido motivada por dificuldades dos alunos em grafar monossílabos átonos, que correspondem – no caso das hipossegmentações – e que os alunos parecem supor corresponderem – no caso das hipersegmentações – às chamadas palavras gramaticais, como preposições, conjunções e pronomes, as quais pertencem a categorias gramaticais que não têm seus usos sistematizados nos anos iniciais do EF, mas nos anos finais do EF.

A identificação da existência de correlação entre tipo de palavras mais frequentemente grafadas não-convencionalmente e categorias gramaticais de palavras que não fazem parte do conteúdo programático do EF I levou Tenani (2011a) a interpretar que essa correlação era relevante não somente por dar visibilidade aos tipos de categorias gramaticais que motivariam “dificuldades ortográficas”, mas também por permitirem observar uma diferença qualitativa entre esses dados do EF II e os do EF I. A partir da

¹⁶ Clíticos fonológicos compreendem, predominantemente, palavras gramaticais – ou funcionais – (como preposições, conjunções, pronomes, artigos) que são desprovidos de acentos. Uma caracterização mais detalhada será feita na seção 3.1.

¹⁷ Descrição detalhada dos conceitos de palavra prosódica e grupo clítico será feita na seção 3.1. Por ora, é suficiente informar que *palavra prosódica* é constituída de informações morfológicas sobre formação de palavra e tem o acento (lexical) como característica fonológica central; *grupo clítico* é a unidade que compreende um elemento clítico (isto é, elemento que não corresponde a afixos, nem a palavras, desprovido de acento) e seu hospedeiro (isto é, uma palavra prosódica, portadora de acento).

consideração de hipersegmentações produzidas por alunos do EF I, descritos por Paula (2007), e comparando-os aos encontrados no EF II, Tenani (2011a) observa que, diferentemente das hipersegmentações do EF II, em que sílabas pretônicas correspondem – predominantemente – à monossílabos átonos (que são palavras gramaticais) do português, as sílabas pretônicas de palavras hipersegmentadas no EF I não necessariamente correspondem a monossílabos átonos, como a sílaba pretônica “ma” em “ma telo” (“martelo”). Naquele trabalho, interpretou-se que a diferença quanto às características das hipersegmentações de sílabas pretônicas de palavra estava ancorada em diferentes modos pelos quais se revelam a atuação de informações letradas (acerca da colocação de espaços em branco para delimitar palavras em textos escritos), constituídas por práticas sociais de leitura e de escrita com as quais os alunos teriam tido ou não contato em ambiente escolar.

Já em Tenani (2011b), é feita descrição de segmentações não-convencionais de palavras identificadas em amostra transversal de textos coletados dos quatro últimos anos do EF. De modo geral, constatou-se que dados de hipossegmentação predominam sobre os de hipersegmentação, característica comum aos dados de escrita do EF I, o que pode ser visto como característico do processo de aquisição da escrita.

Ao tratar de evidências da organização prosódica do português dos enunciados escritos, foi defendida a interpretação de que as segmentações não-convencionais de palavra do EF II seriam motivadas em, principalmente, três estruturas prosódicas da língua, a saber: Σ ,¹⁸ ω e *CG*. Mais especificamente, nas hipersegmentações, houve evidência da atuação do pé troqueu dissílabo (como em “dele gato” [delegado] e “em quanto” [enquanto]) e, nas hipossegmentações, do *CG* (com predomínio de próclise, como em “concerteza”).

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, pode ser observada a atuação de todos os constituintes da hierarquia prosódica¹⁹ como demonstram Capristano (2007a) e Chacon (2011).²⁰ Exemplifica-se, em (1.1), uma hipossegmentação

¹⁸ O *pé métrico* consiste em uma cadeia de sílabas que constituem uma palavra prosódica, sendo atribuído a uma sílaba o valor forte e às demais valor fraco. O *pé troqueu* configura-se quando uma sílaba forte é seguida por outra átona (como em “cama, mesa, banho”).

¹⁹ Apresentação detalhada da hierarquia prosódica é feita no capítulo 3 (seção 3.1).

²⁰ Os dados de (1.1) a (1.3) foram selecionados de Chacon (2011) e os dados de (1.4) a (1.6) de Capristano (2007a).

de palavras que correspondem a um enunciado fonológico (U);²¹ em (1.2), hipossegmentação entre palavras que correspondem a uma frase entoacional (I);²² em (1.3), hipossegmentação de palavras que constituem uma frase fonológica (ϕ);²³ em (1.4), hipossegmentação de palavra mais um clítico formando um GC; em (1.5), a hipersegmentação de palavra que revela palavras prosódicas (ω); em (1.6), a hipersegmentação de palavra delimitando pés métricos (Σ).

1.1. “ELEEARTERO” (“ELE É ARTEIRO”)

1.2. “TEUMAPEMASO” ([O SACI] “TEM UMA PERNA SÓ”)

1.3. “SUMACASISO” (“FUMA CACHIMBO”)

1.4. “FONO” (“FOI NO” [DIA DO CASAMENTO])

1.5. “EM TÃO” (“ENTÃO”)

1.6. “FELIZ BINA” (“FELIZBINA”)

Já no EF II, houve evidências de segmentações não-convencionais de palavras relacionadas a fronteiras de Σ , ω e CG, como anteriormente exemplificado, e não houve dados de escrita que envolvessem fronteiras de constituintes mais altos da hierarquia prosódica, como I e U . Portanto, em comum, verificou-se a atuação do acento (predominantemente o acento lexical no EF I e II, além de acento rítmico, em menor frequência no EF II), como ancoragem prosódica para delimitação de fronteira de palavra no processo de representar na escrita a segmentação do texto em unidades.

Uma importante diferença entre EF I e EF II diz respeito às hipersegmentações. Somente no EF I, encontram-se hipersegmentações como “a j u d as”, “aj u da” e “pe di g rre” (respectivamente, “ajudas”, “ajuda”, “pedigre”), que não correspondem a constituintes prosódicos, como argumenta Chacon (2005), mas que – possivelmente – são motivadas por aspectos da disposição gráfico-visual (das letras na mancha do papel) revelando o que o aluno supõe ser a escrita. Outra diferença diz respeito à presença de grafias em que hipo e hipersegmentação se verificam em uma mesma sequência de palavras como “e levive” (“ele vive”): há hipersegmentação de “ele” e hipossegmentação

²¹ *Enunciado fonológico* é o maior constituinte prosódico e abrange frases entoacionais. Seus limites são definidos a partir dos limites de sentença.

²² *Frase entoacional* é constituinte definido pelas fronteiras de contornos entoacionais, podendo corresponder a uma sentença sintaticamente complexa (como em “A Marina escreveu perguntando se a encomenda havia chegado”) ou a uma única palavra (como em “Socorro!”).

²³ *Frase fonológica* caracteriza-se a partir da configuração de sintagmas definidos pela sintaxe, por exemplo, “Saci fuma cachimbo” configura-se em três frases fonológicas: [saci] [fuma] [cachimbo].

entre “ele” e “vive”. Essas grafias comuns em textos do EF I são raras nos textos do EF II, como será detalhado no capítulo 3.

Para além de serem identificadas semelhanças e diferenças, em termos de constituintes prosódicos, entre os dados de escrita produzida em diferentes momentos de EF, encontraram-se pistas da natureza desses dados e de como são estabelecidas relações entre fala e escrita por crianças e adolescentes. Por meio das análises empreendidas por Tenani (2011a, b) a partir da amostra transversal de textos, foi explicitado em que medida e quais características dos enunciados falados são passíveis de serem observadas por meio das segmentações não-convencionais de palavras nos enunciados escritos. Argumentou-se que essas segmentações: i) se constituem em dados linguísticos que evidenciam aspectos do processo de aquisição da escrita, e ii) podem ser vistas como marcas que resultam da junção das práticas orais/faladas e letradas/escritas e que permitem flagrar a presença do oral/falado no letrado/escrito.

Outra questão relacionada à segmentação não-convencional de palavras relevante para os estudos desenvolvidos a partir da consideração de teorias fonológicas diz respeito justamente à noção de palavra. Discussão sobre a complexidade da noção de palavra é feita por Silva e Tenani (2014, p.51-64) a partir das considerações de Abaurre (1991a) sobre tratar como problema trivial da escrita a necessidade de se conhecerem os critérios convencionais que definem palavra, para que, assim, textos possam ser segmentados em palavras. Sobre essa questão, vale lembrar as considerações de Lemos (1998, p.17), segundo a qual a transformação simbólica operada pela escrita faz com que, muitas vezes, seja negligenciada a relação que aquele que não sabe ler tem com os sinais gráficos que se apresentam transparentes para os que sabem ler. Essa condição dos considerados “alfabetizados” torna-os “de certa maneira, ‘surdos’ para as características mais contínuas dos enunciados orais” (ABAURRE, 1991a, p. 203) e, desse modo, ouvem palavras como unidades morfossintáticas organizadas sintagmaticamente conforme estabelecido pelas convenções ortográficas.

Ainda que seja inegável o fato de que, em uma sociedade predominantemente letrada como a nossa, os sujeitos (alfabetizados ou não) já possuam uma representação de que haja na escrita unidades gráficas, não se pode afirmar que os alunos já “compartilhem dos critérios morfossintáticos e semânticos utilizados [...] na identificação das palavras” (ABAURRE, 1991a, p. 204), até mesmo porque não estão “habituaados à reflexão metalinguística que tal segmentação pressupõe” (ABAURRE et al., 1984, p. 27).

Cabe retomar, ainda de Silva e Tenani (2014, p. 22), a consideração de que “palavra não é um objeto de investigação facilmente definível, já que suas funções e seus limites não são definidos *a priori*”. Os desafios envolvidos no estudo de palavra são explicitados por Veloso (2016, p.14) nos seguintes termos:

confrontamo-nos, em linguística, com uma contradição entre, por um lado, as intuições dos falantes – que geralmente possuem uma noção implícita bastante nítida do que entendem por palavra e se mostram intuitivamente capazes de identificar e de isolar palavras num *continuum* verbal - e, por outro, a dificuldade de encontrarmos critérios estritamente linguísticos [...] que permitam uma definição satisfatória da noção de palavra e sua identificação/delimitação em *continua* linguísticos mais extensos.

Dessas considerações, enfatizo a complexidade linguística da noção de palavra e, especialmente, sua identificação/delimitação em um enunciado (seja falado ou escrito). Nesta tese, grafias não-convencionais das fronteiras de palavra são interpretadas como indícios dessa complexidade da unidade linguística “palavra”, característica pouco notada (e tantas vezes esquecidas) por parte daqueles que escrevem segundo as convenções ortográficas. De modo mais explícito: segmentações não-convencionais de palavra tornam visíveis o fato de que fronteiras de constituintes morfossintáticos e semânticos, que embasam (em parte) as convenções de fronteiras de palavra na escrita,²⁴ não correspondem a fronteiras de constituintes prosódicos (especialmente ω), que se configuram com base em critérios rítmicos e entoacionais e que guiam (em certa medida) os alunos do EF a segmentarem seus textos em unidades distintas daquelas adotadas pelas convenções ortográficas.²⁵

Interessa exemplificar três lócus de não correspondências entre fronteiras de palavras definidas pelas convenções ortográficas e fronteiras de constituintes prosódicos: o primeiro diz respeito às chamadas palavras compostas, exemplificadas em (1.7); o segundo, a um conjunto de palavras derivadas, exemplificadas em (1.8); o terceiro, às chamadas palavras gramaticais não acentuadas, exemplificadas em (1.9).

1.7 i. [espaçonave]_w = [espaço]_ω [nave]_ω

ii. [mesa-redonda]_w = [mesa]_ω [redonda]_ω

²⁴ Vale lembrar que convenções ortográficas são também constituídas sócio-historicamente a partir de práticas letradas (de leitura e escrita) tal como será discutido no capítulo 2.

²⁵ Nesta tese, também são consideradas relevantes noções sobre palavras mobilizadas pelos alunos advindas de suas práticas letradas/escritas. No capítulo 4, essa temática será tratada detalhadamente.

- 1.8. i. [simplesmente]_w = [simples]_ω [mente]_ω
 ii. [autodidata]_w = [auto]_ω [ditada]_ω
- 1.9. i. [de]_w [novo]_w = [de novo]_{GC}
 ii. [jogá]_w [lo]_w = [jogá-lo]_{GC}

Em (1.7), as palavras “espaçonave” e “mesa-redonda” são compostas por duas palavras lexicais, constituindo, cada uma delas, uma palavra morfológica (w). Quanto às convenções ortográficas, a fronteira entre as palavras lexicais é marcada por hífen apenas em “mesa-redonda”, não ficando delimitada em “espaçonave”. Nos dois casos, a ortografia registra os limites da palavra morfológica. Fonologicamente, nos dois exemplos, há duas palavras prosódicas, pois são preservados os acentos de cada uma das palavras lexicais. Por meio desses exemplos, demonstra-se a relevância da interação entre critérios fonológicos e morfológicos para a caracterização da noção de palavra. Faz-se importante, portanto, distinguir entre palavra prosódica (ou fonológica) e morfológica, tal como proposto por Bisol (2004, p.59):

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

Em (1.8), “simplesmente” é uma palavra morfológica constituída da palavra lexical “simples” e o sufixo derivacional “mente”. Fonologicamente, “simples” e “mente” mantêm seus acentos, constituindo-se em duas palavras prosódicas, embora uma morfológica. Já “autodidata” é uma palavra morfológica constituída do prefixo “auto” e da palavra lexical “didata”. Fonologicamente, “auto” e “didata” preservam seus acentos, constituindo, cada um, uma ω. Portanto, há palavras morfológicas (derivadas por certos sufixos ou por certos prefixos) que correspondem a duas palavras prosódicas.²⁶

Em (1.9), as palavras morfológicas “de” e “lo” compreendem palavras gramaticais que dependem sintaticamente das palavras lexicais que lhe são adjacentes, ou seja, a preposição “de” depende do adjetivo “novo” que lhe segue; a forma pronominal “lo”, da forma verbal “joga” que lhe precede. Fonologicamente, essas palavras gramaticais não têm acento e, por essa razão, são clíticos fonológicos, ou seja, são palavras morfológicas

²⁶ Detalhes sobre o funcionamento de palavra compostas, veja-se Boop da Silva (2010). Sobre o funcionamento dos prefixos acentuados, veja-se discussão feita por Schwindt (2000).

que não têm status de ω . Essa relação morfossintática entre palavras morfológicas em que uma delas é um clítico fonológico é interpretada como característica do CG, constituinte prosódico assim denominado na teoria fonológica adotada, como será mais bem detalhado no capítulo 3 (seção 3.1).

Por meio dessas considerações, ficam explicitados um fato e uma hipótese. O fato é que as fronteiras de palavra definidas pelas convenções ortográficas (palavras ortográficas) são definidas, em grande parte, com base em fronteiras de palavra morfológicas e essas fronteiras são distintas daquelas dos constituintes prosódicos (palavras prosódicas, grupos clíticos). A hipótese explicativa para as ocorrências de segmentações não-convencionais de palavra estaria na não correspondência entre fronteiras de palavra morfológica e de ω /CG. Portanto, segmentar um texto em palavras segundo as convenções ortográficas implicaria mobilizar fronteiras de constituintes linguísticos que são definidos por informações de natureza prosódica, sintática e semântica, como será discutido no capítulo 3.²⁷ No entanto, será demonstrado, no capítulo 4, que essa hipótese não é suficiente para explicitar as segmentações não-convencionais de palavra, sendo necessário ampliar a análise ao enunciado em que as segmentações ocorrem.

Em suma, as pesquisas realizadas sobre segmentação não-convencional de palavra permitem identificar ao menos três aspectos relevantes para seu estudo no campo da Linguística, que dizem respeito à: (i) compreensão sobre relações entre fala e escrita mobilizadas pelos alunos no processo de produção escrita, (ii) comprovação de pressupostos de teorias fonológicas relevantes para descrever regularidades linguísticas observadas nos textos escritos e (iii) caracterização da noção de palavra, notadamente, consideradas características de sua constituição enquanto unidade fonológica, morfossintática e semântica. Esses três aspectos serão abordados no decorrer da análise de hipersegmentações, a ser apresentada nas seções 3.2 e 3.3, visando demonstrar em que medida essas são resultados de hipóteses conflitantes que os escreventes elaboram a partir de sua multifacetada relação com a língua(gem) (TENANI, 2009, p. 109).

Por fim, à luz das considerações feitas acerca das noções de palavra (ortográfica, morfológica, prosódica), retomam-se as considerações que abriram este capítulo. Esta

²⁷ Vale observar que segmentação de palavras está inserido em tema mais amplo sobre o processo de segmentar enunciados em unidades linguísticas, relevante para a aquisição da linguagem, por exemplo. Freitas et al. (2006) demonstram que há efeitos de frequência de ocorrências de unidades e padrões prosódicos relevantes para a identificação, pelas crianças, de regularidades silábicas do PE, por exemplo.

tese apresentará argumentos contrários à interpretação das segmentações não-convencionais de palavras, quer como simples desvios das convenções ortográficas (decorrentes de falta de atenção do aluno ou de dificuldades de alfabetização vividas principalmente em escolas públicas ao longo do EF I), quer como sintomas de disortografia (decorrentes de dificuldades de processamento cognitivo de alunos do EF II quanto a estabelecer relação entre fala e escrita), para defender abordagem que privilegia a investigação de aspectos linguísticos (e também discursivos) que serão identificados como característicos de grafias não-convencionais de palavra que ocorrem em textos produzidos por alunos do EF II. Serão apresentados argumentos a favor da hipótese, apresentada na introdução, de que *segmentações não-convencionais de palavras são efeitos dos modos pelos quais os sujeitos escreventes (crianças e adolescentes) são atravessados por práticas sociais letradas/escritas e, também, por práticas sociais orais/faladas*.

No próximo capítulo, é feita uma caracterização das segmentações não-convencionais de palavra no EF II quanto à frequência de ocorrência em uma amostra longitudinal de textos produzidos por alunos de uma escola pública paulista. Essa caracterização é precedida por uma descrição do banco de dados, bem como do perfil de desempenho escolar dos alunos que produziram os textos analisados, e por uma discussão sobre a metodologia desenvolvida para identificação e análise das segmentações não-convencionais de palavra nos textos manuscritos.

2 Segmentações não-convencionais de palavra na escrita do EF II

Neste capítulo, caracterizo as segmentações não-convencionais de palavra identificadas em uma amostra longitudinal de textos do EF II. O propósito dessa caracterização é o de circunstanciar a contribuição inédita desta tese para os estudos em Linguística principalmente, mas também para os estudos em Educação, haja vista tratar-se de textos escritos em ambiente escolar. Essa caracterização é feita em três seções.

Na seção 2.1, a amostra longitudinal de textos escritos por alunos do EF II que foram analisados é descrita a partir de três eixos. Primeiramente, são dadas informações gerais sobre o Banco de Dados ao qual pertencem os textos selecionados com o propósito de expor como eles foram coletados. Em seguida, são caracterizados os alunos que produziram os textos da amostra a partir de índices governamentais de desempenho escolar da educação básica, visando situar, no cenário estadual e nacional, a produção escrita desses alunos. Ainda nessa seção, são dadas informações sobre o desenvolvimento humano do município em que residiam os alunos cujos textos foram analisados, com destaque para os índices sobre Educação do governo federal. Por meio dessas caracterizações, pretendo dimensionar o alcance deste estudo sobre escrita do EF II.

Na seção 2.2, são descritas e problematizadas as decisões metodológicas tomadas para a constituição da segmentação não-convencional de palavra enquanto dado de escrita e que ganha traços de dados de leitura. Além disso, explico os critérios em que se assentam a identificação e a classificação dos dados em hipossegmentação ou hipersegmentação. A proposta é sistematizar procedimentos metodológicos para a constituição do dado de escrita enquanto objeto de análise linguística.

Na seção 2.3, é feita a caracterização da amostra quanto à frequência de segmentações não-convencionais de palavra nos textos, tendo a palavra ortografia como unidade de análise. Foi feita análise da distribuição dos dados ao longo dos anos letivos, pois a amostra longitudinal de textos produzidos pelos mesmos alunos durante quatro anos permitiu esse tipo de investigação. Também foi investigada a incidência de tipos de segmentação não-convencional de palavras por sujeito a fim de identificar tendências gerais e de detectar trajetórias particulares quanto às características das segmentações de

palavra. Com base nos resultados apresentados, é justificada a seleção das hipersegmentações para análise prosódica e morfossintática a ser feita no capítulo 3.

2.1 Caracterização do material analisado

Como já anunciado, o conjunto de segmentações não-convencionais de palavra analisado nesta tese foi extraído de uma amostra longitudinal de textos que integra o “Banco de Dados de Escrita do Ensino Fundamental II” (TENANI, 2015). Esse banco é constituído por 5.519 textos escritos por 662 alunos dos quatro últimos anos do Ensino Fundamental. Esses textos compõem uma amostra transversal formada por 2.759 textos e uma amostra longitudinal do banco formada por 3.645 textos.²⁸ Nesta tese, foram analisados 2.457 textos (ou 67,4%) da amostra longitudinal, os quais foram produzidos por 116 alunos, sendo 55 (47,4%) do sexo/gênero²⁹ feminino e 61 (52,5%) do sexo/gênero masculino. Esse subconjunto de textos se caracteriza por todos esses 116 alunos (dos 182 que compõem a amostra longitudinal) terem produzidos textos em todos os anos letivos. Esse critério de seleção de sujeitos se deve ao objetivo de realizar pesquisa de natureza longitudinal dos textos do EF II, sendo, pois, necessário haver textos dos mesmos sujeitos em todos os quatro anos letivos considerados.

A coleta dos textos do banco foi feita por meio do projeto de extensão universitária “Desenvolvimento de oficinas pedagógicas de leitura, interpretação e produção textual no ensino fundamental”, desenvolvido pelas professoras doutoras Luciani Tenani e Sanderléia Longhin-Thomazi (IBILCE/UNESP), em parceria com uma escola pública no município de São José do Rio Preto, noroeste paulista. O projeto, submetido e aprovado pelo Comitê de Ética em Pesquisa do IBILCE/UNESP,³⁰ foi desenvolvido por quatro anos, de 2008 a 2011, com o objetivo geral de aprimorar práticas de leitura e escrita dos

²⁸ Cabe observar que parte dos dados da amostra transversal compõe a amostra longitudinal. A amostra transversal (coletada em 2008) compreende textos de alunos dos quatro anos finais do EF. A amostra longitudinal (coletada de 2008 a 2011) compreende textos dos alunos da então quinta série que permaneceram na escola nos anos letivos subsequentes até concluírem o EF.

²⁹ Quando da coleta dos textos, apenas foi identificado o sexo de cada aluno, não tendo sido feito levantamento quanto à identidade de gênero.

³⁰ 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) em 2008, pelo Comitê de Ética em Pesquisa do IBILCE/UNESP.

alunos matriculados na escola.³¹ No período de desenvolvimento desse projeto, houve ampliação de oito para nove anos na duração do EF, de modo que o ciclo I passou a ter cinco anos de duração e o ciclo II manteve-se com quatro anos.³² Em termos de planejamento curricular do ciclo II, não houve mudanças de modo que foi estabelecida relação entre as então quinta, sexta, sétima e oitavas séries aos sexto, sétimo, oitavo e nono anos do EF, respectivamente. Nesta tese, empregou-se a denominação atual dos anos letivos, mas registra-se que os textos analisados foram produzidos por alunos que cursaram oito anos de EF.

A partir de oficinas de leitura e escrita de textos de gêneros variados, foram apresentados (em folha impressa) excertos de textos selecionados sobre temas diversos os quais eram lidos e discutidos coletivamente em aulas conduzidas por monitores do projeto, que eram licenciandos em Letras ou pós-graduandos em Estudos Linguísticos do IBILCE/UNESP. Após leitura e interpretação dos textos que introduziram pontos de vistas diversos sobre temas relativamente polêmicos, os alunos eram convidados a produzirem um texto manuscrito sobre o tema discutido, segundo um conjunto de orientações impressas em folha dada a cada um em sala de aula.

Essas orientações consistiam de definição de (i) tema, (ii) gênero e tipologia textual e (iii) forma de usar a folha destinada à elaboração do texto.³³ O tema e o gênero foram definidos pela coordenação do projeto de extensão em reuniões com os monitores integrantes do projeto de extensão que atuavam na escola. A tipologia textual solicitada em cada proposta era previamente estabelecida pela escola em função do planejamento escolar, desenvolvido conforme a Proposta Curricular do Estado de São Paulo (2008).³⁴ Aos alunos não era permitido reelaborar o texto produzido, sendo necessário entregá-lo ao responsável pela oficina ao final da aula de língua portuguesa. Essas condições de produção definidas pela coordenação do projeto de extensão não eram diferentes das práticas de escrita adotadas naquela escola, de modo que os alunos não apresentaram dificuldades em seguir essas orientações. Sem possibilidade de reelaboração (por não ser disponibilizado tempo extra ou em aula para esse fim) e sem possibilidade de uso de lápis

³¹ Uma descrição detalhada das atividades desenvolvidas no projeto de extensão, as quais motivaram a produção de textos manuscritos é feita por Tenani e Longhin-Thomazi (2014).

³² Por meio da Lei Federal Nº 11.274, de 2009, instituiu-se o Ensino Fundamental de nove anos, sendo que o ciclo II de quinta a oitava séries passou a corresponder aos sexto a nono anos.

³³ Em todas as propostas, eram dadas orientações quanto a usar caneta preta ou azul; escrever no espaço de folha de papel A4, delimitado por margens e contendo 25 linhas.

³⁴ Um quadro sinóptico das 44 propostas de produção textual quanto a tema, gênero e tipologia textual encontra-se disponível em Tenani (2015).

grafite (o que levou a rasuras do texto escrito), os alunos deixaram preciosos registros do processo de elaboração da produção textual.

É relevante destacar que essa metodologia empregada na coleta dos textos proporciona ao analista identificar pistas, indícios (ABAURRE; FIAD; MAYRINK-SABINSON, 1997) de hipóteses do sujeito sobre amplo conjunto de elementos linguísticos que constituem o texto escrito, dentre os quais estão as formas ortográficas das palavras. Cada texto analisado é visto aqui como fotografia que captura um momento de um processo. É, pois, a partir da informação verbo-visual inscrita no texto que o analista elabora hipóteses sobre o trabalho epilinguístico (ABAURRE, 1988b) do sujeito sobre a linguagem, particularmente, no que interessa analisar, sobre relações entre fala e escrita.

Explicitada a metodologia adotada na coleta dos textos que compõem o material investigado, passa-se a tratar do perfil dos alunos quanto ao rendimento escolar, mensurado por meio dos índices estadual (IDESP) e nacional (IDEB). Vale lembrar que essa caracterização visa situar, no cenário estadual e nacional, a escola em que os textos do banco foram produzidos, permitindo-nos dimensionar o quão esses textos são representativos de características da escrita do EF II no Brasil.

Inicialmente, faz-se necessário explicitar que o IDESP de cada ano escolar é obtido a partir das notas dos alunos obtidas nos Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (SARESP)³⁵ – que compreende provas que avaliam habilidades em várias disciplinas –, e do fluxo escolar, determinado pela taxa de aprovação média em cada ciclo.³⁶ Nos quatro anos em que os textos foram coletados na escola, houve variação nas disciplinas avaliadas, sendo, porém, constantes avaliações dos conteúdos de língua portuguesa e matemática.³⁷ Na tabela a seguir, são apresentadas as

³⁵ Os desempenhos dos alunos em cada disciplina são classificados em *insuficiente*, *suficiente* e *avançado*, com base em quatro níveis de proficiência, conforme a seguinte correspondência: *insuficiente* para médias de nível abaixo do básico; *suficiente* para médias consideradas de nível básico e adequado; *avançado* para o nível avançado. Alunos de sétimo e nono anos que têm médias *insuficientes* são retidos no ano escolar.

³⁶ Cabe observar que em 2014, houve mudança na composição do IDESP quando foi incluído o INSE (índice de nível socioeconômico do aluno), obtido a partir de formulário preenchido com dados sobre bens e renda familiar de cada aluno.

³⁷ Em 2008, foi feita avaliação do desempenho dos alunos de sexta e oitava séries quanto a conteúdos de Língua Portuguesa, Matemática, Ciências e Ciências da Natureza e Redação. Em 2009, alunos das mesmas séries foram avaliados quanto a conteúdos de Língua Portuguesa, Matemática, História, Geografia e Redação. Em 2010, alunos dos sétimo e nono anos foram avaliados quanto aos conteúdos das mesmas disciplinas em 2008, exceto Redação. Em 2011, os conteúdos avaliados foram os mesmos de 2009. Em todos esses anos, foram divulgados boletins de médias da escola comparadas com as médias de cinco instâncias administrativas, a saber: escolas estaduais do município, da diretoria de ensino, das coordenadorias do interior (CEI) e da capital e da Grande São Paulo (COGSP), e do estado. Atualmente, a comparação restringe-se ao IDESP da diretoria de ensino, do município e do estado.

médias em língua portuguesa, da escola e do estado no IDESP para os anos de 2008 a 2011, período em que foram coletados os textos analisados. Acrescenta-se ainda informação sobre as metas de médias no IDESP que deveriam ser alcançadas pela escola no ano escolar seguinte à aplicação do SARESP.

Tabela 2.1: Médias obtidas e metas do IDESP para alunos do 9º ano do EF

ANO LETIVO	MÉDIA EM LG. PORT.	MÉDIA DA ESCOLA	MÉDIA DO ESTADO	META A SER ALCANÇADA (ANO)
2008	3,7900	3,43	2,60	3,53 (2009)
2009	3,9680	3,70	2,84	3,81 (2010)
2010	3,9537	3,37	2,52	3,54 (2011)
2011	3,9070	3,40	2,57	3,58 (2012)

Fonte: Secretaria da Educação do Estado de São Paulo, IDESP, 2015.

De modo geral, as médias em língua portuguesa são relativamente mais altas do que a média geral do IDESP da escola, o qual também compreende resultados no SARESP em matemática (e que são sistematicamente inferiores aos de língua portuguesa, uma tendência geral do IDESP). Entre as metas estabelecidas para a escola pela Secretaria da Educação,³⁸ somente em 2009 a meta foi alcançada, tendo sido o IDESP apurado maior do que o estabelecido na meta. Esses resultados sugerem que, nesse período considerado, não houve a desejada melhoria do ensino nesse ciclo da Educação Básica no estado.³⁹

Quanto às médias da escola ao longo dos quatro anos considerados, constata-se aumento de 0,27 pontos no índice de 2009 em relação ao de 2008, mas decréscimo de 0,33 pontos em 2010 seguido de ligeiro aumento de 0,03 pontos. Essa variação também se verifica entre as médias do IDESP para o Estado no período.⁴⁰ Uma vez que a variação

³⁸ A Secretaria de Educação do Estado de São Paulo estabeleceu metas a serem alcançadas até 2030 para cada ciclo do ensino básico e ensino médio. A meta para o estado é alcançar índice 6 para o ciclo II do EF.

³⁹ O IDESP da escola manteve-se em queda em 2012 (3,22) e 2013 (2,71), mas teve aumento de 0,46 pontos em 2014 (3,17). Esse aumento coincide com mudança na fórmula de cálculo do IDESP que passou a incluir Índice de Nível Socioeconômico do aluno (INSE). Em 2015, novo aumento do IDESP foi registrado para o ciclo II do EF, uma tendência geral para todo o estado. Esse aumento levou o índice ao mesmo patamar registrado em 2009. Não podendo, pois, ser considerado índice de melhoria da educação básica, como alertam pesquisadores da área de Educação críticos do modelo implantado pelo governo estadual paulista (FREITAS, 2016a). Pondera-se, ainda, que os resultados de 2015 sejam efeito de uma possível “calibração” da prova para um patamar de menor exigência, uma suspeita que põe em alerta professores e pesquisadores paulistas (FREITAS, 2016b).

⁴⁰ Cabe levantar a hipótese de que essa variação do IDESP esteja relacionada a uma possível mudança na forma de obtenção do IDESP no período considerado, uma vez que se verifica mudança nas fundações responsáveis pela aplicação e correção das provas do SARESP: CESGRANRIO (Rio de Janeiro) em 2008, CAED (Juiz de Fora) em 2009 e VUNESP (São Paulo) em 2010 e 2011. No entanto, nenhuma informação

no IDESP da escola é semelhante à verificada para o Estado, pode-se assumir que os dados da escola analisada representam características gerais do desempenho dos alunos no final do EF no Estado, sendo, porém, uma escola com índices sistematicamente acima da média geral do Estado.

Descritos os índices de avaliação da escola em relação ao Estado de São Paulo, passo a tratar dos resultados da escola no IDEB, notadamente para os anos escolares de 2009 a 2011, a fim de ser caracterizado o perfil da escola analisada em relação à educação básica no país. Vale observar que, diferentemente do IDESP (cujos índices são calculados a partir das provas do SARESP), o IDEB é calculado com base na Prova Brasil, implantada pelo Ministério da Educação. A Prova Brasil mede proficiência em leitura, como parte das, assim denominadas, habilidades de língua portuguesa, enquanto as provas do SARESP medem vasto conjunto de habilidades em língua portuguesa, que visam mensurar os níveis de proficiência em leitura e escrita no estado paulista.⁴¹ Portanto, os dois índices, embora sejam definidos em escala de 0 a 10, não avaliam as mesmas habilidades de língua portuguesa. Na Tabela 2.2, são apresentados resultados do IDEB observados para a escola parceira em comparação com a média dos índices das redes estaduais do Estado de São Paulo e do Brasil, além das metas estabelecidas para essas instâncias a cada ano para alunos do nono ano do EF.

Tabela 2.2: Médias obtidas e metas do IDEB para alunos do 9º ano do EF

ANO LETIVO	ESCOLA	META	ESTADO	META	PAÍS	META
2009	4,7	4,8	4,3	4,0	3,8	3,5
2011	5,1	5,0	4,3	4,2	3,9	3,8

Fonte: Ministério da Educação, INEP (2015).

A comparação dos resultados do IDEB apresentados na Tabela 2.2 permite constatar que, embora em 2009 o índice obtido de 4,7 estivesse abaixo da meta de 4,8 que fora estabelecida pelo Ministério da Educação, a escola alcançou índices acima dos verificados para o Estado de São Paulo e o país, tanto em 2009, quanto em 2011. Esses índices foram acima das metas que haviam sido previstas para as escolas das redes

sobre mudança na concepção das provas ou nos critérios de correção de avaliação é dada no sítio oficial da Secretaria de Educação.

⁴¹A Secretaria da Educação do Estado de São Paulo (2009) estabeleceu matrizes de referências para avaliação no SARESP. Nesse documento, são explicitadas as matrizes para cada disciplina avaliada em função de cada ciclo do EF e do Ensino Médio.

estaduais paulista e brasileira, o que subsidia considerar o desempenho dos alunos da escola parceira do projeto de extensão universitária como representativo de características gerais do Estado e do país no período analisado, sendo, porém, uma escola com índices ligeiramente acima das médias apuradas para a educação básica no estado e no país entre escolas de rede estadual.

Outra informação relevante para a caracterização dos alunos que produziram os textos analisados diz respeito ao perfil do município quanto aos índices de desenvolvimento humano. A análise de Almeida, Dalben e Freitas (2013), apresentada na seção 1.1, aponta que dados socioeconômicos dos alunos sejam relevantes para avaliar o desenvolvimento dos estudantes em leitura. Assumi a premissa de que também são relevantes os Índices de Desenvolvimento Humano do Município (IDHM)⁴² não apenas para caracterizar o município, mas especialmente para ter parâmetros abrangentes sobre características sociais de desenvolvimento humano do local onde residiam os alunos cujos textos são analisados. Por meio desses índices, são descritas características socioeconômicas e educacionais que contribuem para a configuração, em última instância, da representatividade das análises sobre as produções escritas consideradas.

A escola estadual onde foram coletados os textos analisados está situada no município de São José do Rio Preto, a 440 km da capital paulista, sendo sede administrativa da região do noroeste do estado. Quanto aos dados do IDHM, divulgados pelo Programa das Nações Unidas para o Desenvolvimento (PNUD) em conjunto com o Instituto de Pesquisa Econômica Aplicada (IPEA) e a Fundação João Pinheiro, São José do Rio Preto é classificado na 50ª posição entre os 5.565 municípios brasileiros,⁴³ com índice de 0,797 em 2010, considerado *alto* (faixa de 0,7 a 0,799 pontos) entre as cinco faixas de desenvolvimento. Dentre as três dimensões que compõem esse índice, *Educação* apresenta o pior índice com 0,748 pontos, enquanto as dimensões *Longevidade* e *Renda* alcançam, respectivamente, os índices de 0,846 e 0,801, patamares classificáveis na faixa *muito alto*. Esses índices sugerem que a dimensão Educação carece de mais

⁴² O Índice de Desenvolvimento Humano (IDH) é uma medida do progresso a longo prazo em três dimensões consideradas básicas do desenvolvimento humano: renda, educação e saúde. Essas três dimensões compreendem indicadores de população, educação, habitação, saúde, trabalho, renda e vulnerabilidade social, extraídos de dados dos censos demográficos de 1991, 2000 e 2010. Segundo informado no sítio oficial do Programa das Nações Unidas para o Desenvolvimento (2016), “o objetivo da criação do IDH foi o de oferecer um contraponto a outro indicador muito utilizado, o Produto Interno Bruto (PIB) per capita, que considera apenas a dimensão econômica do desenvolvimento”.

⁴³ Entre os 50 municípios mais bem classificados, 44% são paulistas e, exceto Brasília, todos os demais pertencem às regiões sul e sudeste do país. O estado de São Paulo com índice de 0,783 é, após o Distrito Federal (único com índice *muito alto* de 0,824), a unidade da federação mais bem classificada.

investimento no município a fim de proporcionar aos cidadãos melhores condições de desenvolvimento.

Quando considerados os índices dessas três dimensões do IDHM no período de 1991 a 2010 (sobre os quais há dados de três Censos Demográficos), constata-se que a dimensão Educação foi a que mais cresceu (na ordem de 0,339 pontos), seguida por Longevidade e Renda.⁴⁴ Essa tendência é semelhante à observada no país, uma vez que a dimensão Educação também foi a que mais cresceu em relação às demais dimensões avaliadas, porém o índice nacional teve taxa de crescimento de 0,358, um pouco maior do que a do município que apresentava índices maiores do que os nacionais em 1991.⁴⁵

Outro conjunto de dados do IDHM Educação que vale considerar são os que dizem respeito à situação da educação entre a população em idade escolar, consideradas as proporções de crianças e jovens frequentando ou tendo completado ciclos de formação escolar. No município de São José do Rio Preto, a proporção de crianças de 11 a 13 anos frequentando os anos finais do EF foi de 87,98% em 2010.⁴⁶ Em relação à taxa apurada em 1991, houve um acréscimo de 24,54%, e, em relação à taxa em 2000, apenas 1,37% de aumento. Essa taxa de crescimento de crianças no EF II foi, no período, a menor em relação às taxas verificadas para os demais ciclos de formação escolar.⁴⁷ Em relação aos índices do Estado de São Paulo e do país, a taxa de 87,98% de porcentagem de crianças de 11 a 13 anos no EF II (ensino regular seriado) do município é praticamente semelhante à taxa do Estado (89,04%) e ligeiramente maior do que a taxa do país (84,86%).⁴⁸

Os índices de desenvolvimento humano ora reportados para o município onde residiam os adolescentes cujos textos foram analisados podem ser considerados representativos de características nacionais da proporção de crianças de 11 a 13 anos no

⁴⁴ Para o município, as taxas nos anos de 1991, 2000 e 2010 para Educação são, respectivamente: 0,409; 0,655 e 0,748. Para Longevidade, as taxas nos três anos são: 0,748; 0,813 e 0,846. Para Renda, as taxas nos três anos são: 0,741; 0,777 e 0,801. Esses índices permitem visualizar que a dimensão Educação apresentou taxas inferiores em todo o período considerado.

⁴⁵ De modo geral, no período de 1991 a 2010, a taxa do IDHM de São José do Rio Preto cresceu 30,66% (enquanto a taxa foi de 47% no país) e houve uma redução do hiato de desenvolvimento humano de 52,05% (enquanto a taxa foi de 53,66% na federação).

⁴⁶ No censo demográfico de 2010 realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE), a população residente no município totalizou 408.258 habitantes, sendo 48,01% homens e 51,98% mulheres. Na faixa etária de 10 a 14 anos de idade, havia 27.304 habitantes, sendo 50,79% do sexo masculino e 49,04% do sexo feminino (Dados disponíveis em: <http://cidades.ibge.gov.br/xtras/temas.php?lang=&codmun=354980&idtema=1&search=sao-paulo|sao-jose-do-rio-preto|censo-demografico-2010;-sinopse-> Acesso em 10 de setembro de 2016).

⁴⁷ No período de 1991 a 2010, a taxa aumentou em 55,36% para crianças entre 5 a 6 anos frequentando o EF I; em 37,81% para jovens entre 15 a 17 anos frequentando o EM.

⁴⁸ Para essa variável, o estado de Santa Catarina detém o melhor índice com 91,51%; e o estado do Pará, o pior, com 73,13%.

EF II e do aumento do acesso à Educação da população brasileira de modo amplo, verificado nas últimas duas décadas (de 1991 a 2010).

Por meio dessas informações, descrevem-se (i) contextos de como foram produzidos os textos coletados, (ii) características do material selecionado e (iii) características da escola e do município onde os textos foram produzidos que possivelmente representam características mais gerais da escrita do EF II produzida no país.

2.2 Constituição do objeto de análise

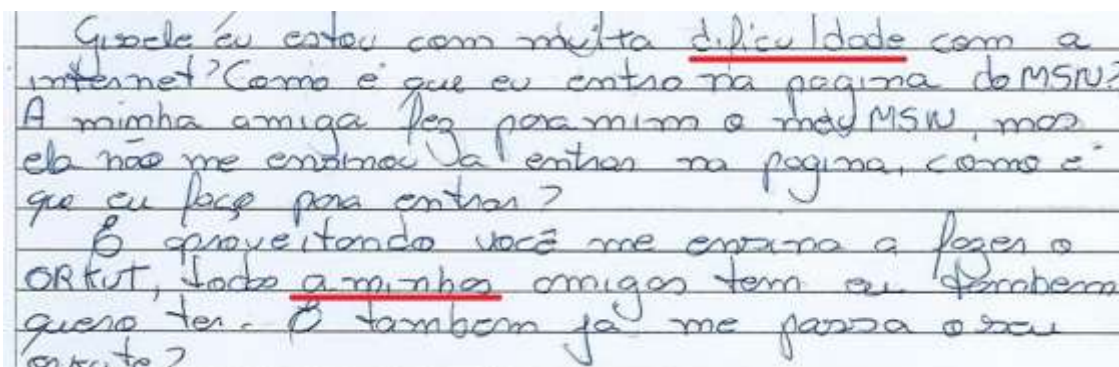
Para a constituição do objeto de análise desta tese, faz-se necessário tratar das decisões metodológicas quanto à identificação do dado de segmentação não-convencional de palavra. Por meio de apresentação dessas decisões, pretendo: (i) problematizar aspectos metodológicos quanto à constituição do dado de escrita; e (ii) caracterizar a ocorrência dos tipos de segmentação de palavra nos textos do EF II. Além disso, busco responder às seguintes questões: (i) o que é um dado de segmentação de palavra?; e (ii) como é identificada a segmentação de palavra não-convencional no EF II?

Vale lembrar que, na literatura sobre escrita do EF I, a segmentação não-convencional de palavra gráfica pode ser classificada em dois tipos principais: (i) hipossegmentação, quando há ausência de espaço em branco entre palavras, como em “ajudime”, “porfavor”, e (ii) hipersegmentação, quando há presença do espaço em branco dentro de palavra, como em “na quela”, “cava lo”. Uma questão metodológica importante relacionada a essa distinção é, em textos manuscritos, categorizar o branco entre palavras, identificando quando há ou não o espaço de fronteira de palavras empregado fora das convenções ortográficas. As Figuras 2.1 a 2.3, apresentadas mais à frente, são exemplos de diferentes tipos de grafias que geram dúvidas na identificação de dados.

Na Figura 2.1, a grafia de “dificuldade” apresenta um espaçamento entre letras, proporcionalmente maior entre “dificu” e “ldade”. Essa grafia foi vista em Tenani (2011b) como resultado de uma flutuação na forma de grafar as ligaduras entre letras em uma mesma palavra. O manuscrito se caracteriza pela alternância entre letras manuscritas e impressas de modo que a ausência das ligaduras entre as letras dentro de uma palavra, por vezes, ocorre. 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 2.1, destaca-se a grafia de “a minhas” que poderia ser um dado de hipossegmentação, 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 leva a afirmar que a grafia em questão envolve aspecto caligráfico e não ortográfico da palavra.

Figura 2.1: Grafias de “dificuldade” e “a minhas”



Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z08_6A_19F_04)⁴⁹

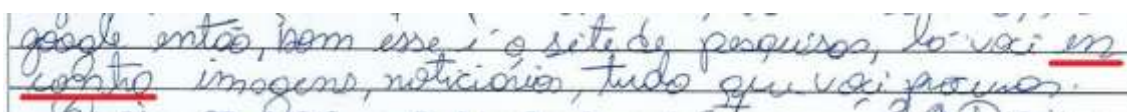
A descrição ora realizada exemplifica dificuldades quanto à categorização do espaço em branco entre palavras e, simultaneamente, permite avaliar o tipo de olhar necessário para a identificação dos dados. Na literatura sobre escrita do EF I, Paula (2007) trata detalhadamente de aspectos metodológicos relevantes para identificação de ocorrências de segmentação não-convencional de palavras. Diferentemente daquele trabalho que analisou textos grafados principalmente com letras em caixa alta, 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. Explicito, pois, os critérios adotados para a identificação dos dados, uma vez consideradas as características do material de pesquisa: (1) comparação da proporção dos espaços entre

⁴⁹ Leitura possível: “Gisele eu estou com muita dificuldade com a internet? Como é que eu entro na página do MSN? A minha amiga fez para mim o meu MSN, mas ela não me ensinou a entrar na página, como é que eu faço para entrar? / E aproveitando você me ensina a fazer o ORKUT, todas as minhas amigas têm e também quero ter. E também já me passa o seu orkute?”

palavras ao longo do texto; e (2) comparação entre grafias das mesmas letras em palavras semelhantes que ocorram no mesmo texto.

Faz-se necessário explicitar também o tratamento dado à translineação, ilustrada no registro de “encontro”, na Figura 2.2. Segundo as convenções para manuscritos, um hífen deve ser colocado onde a palavra ficar dividida na translineação. A ausência desse hífen, como ocorre em “en contra”, foi interpretada como somente ausência de registro gráfico da translineação e não foi considerada ocorrência de hipersegmentação.

Figura 2.2: Grafia de translineação de “en contra”

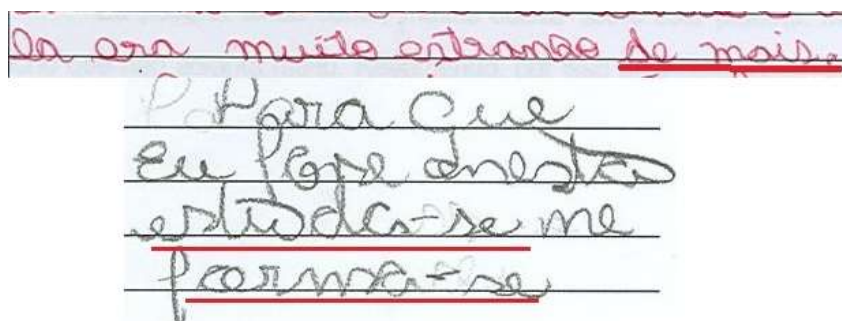


Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z08_6A_29F_04)⁵⁰

Além desses critérios para identificação do dado, são adotados dois critérios para a classificação das segmentações de palavras: (i) critério gráfico, que diz respeito ao tipo de informação verbo-visual que delimita a fronteira da palavra escrita, e (ii) critério morfossemântico, que diz respeito à identificação de palavras homônimas.

Quanto ao critério gráfico, é importante distinguir, nas hipersegmentações, a presença não-convencional do branco (“de mais”) ou do hífen (“estuda-se”, “forma-se”), ilustrada na Figura 2.3, e, nas hipossegmentações, a ausência não-convencional do branco (“denovo”) ou do hífen “perseguido, perseguidos”, ilustrada na Figura 2.4.

Figura 2.3: Grafias “de mais”, “estuda-se” e “forma-se”

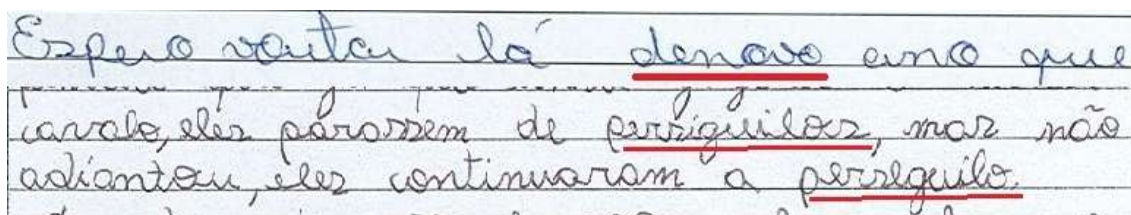


Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z08_5C_09M_05 e Z08_5A_26F_02)⁵¹

⁵⁰ Leitura possível: “google então, bom esse é o site de pesquisas, lá você encontra imagens, noticiários, tudo que você procurar”.

⁵¹ Leitura possível: “lá era muito estranho demais.” (trecho 1). “Para que eu fosse honesta, estudasse me formasse.” (trecho 2).

Figura 2.4: Grafias “denovo” e “persiguiolos, perseguido”



Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z08_5B_35F_06 e Z08_5A_30M_01)⁵²

Vale observar que indicar fronteira entre palavras por branco “só se tornou prática geral na sociedade ocidental no século XIX” (FISCHER, 2009, p. 229). Portanto, a representação de fronteira de palavras é prática tardia na história da escrita ocidental.⁵³ A prática escrita dos primeiros escribas medievais caracterizava-se pela junção entre as palavras e, ainda segundo Fischer (2009, p. 229), “só quando as minúsculas se tornaram caligrafia literária, os escribas perceberam a necessidade de representar a palavra como unidade independente – talvez por causa de leitura silenciosa ter se tornado mais frequente nos *scriptoria* medievais”.

Nota-se, nessa afirmação, que Fischer (2009) lança mão de hipótese que correlaciona a prática escrita de uso do branco que delimita palavras à prática de leitura silenciosa do texto.⁵⁴ Haveria, assim, uma mudança na representação de palavra na escrita relacionada à outra que diz respeito à relação entre fala e escrita, na medida em que a oralização do texto escrito dá lugar à prática de leitura ancorada, também, na organização verbo-visual da escrita que se instaura na folha de papel. Conclui-se, dessas considerações, que o processo histórico de segmentar o enunciado em palavras, delimitadas por branco, parecer ter decorrido de processos mais amplos que englobaram práticas de leitura e escrita. Com base nessa interpretação, aventa-se a hipótese de que ao processo de segmentar o enunciado em palavras, a que estão sujeitos alunos do EF II,

⁵² Leitura possível: “Espero voltar lá de novo ano que” (trecho 1). “cavalo, eles parassem de persegui-los, mas não adiantou, eles continuaram a persegui-lo” (trecho 2)

⁵³ Fischer (2009, p.229) afirma que “os primeiros escribas gregos e latinos também não deixavam espaço entre as palavras, embora alguns colocassem pequenos pontos configurando as separações, em inscrições de monumentos e papiros literários”. Apenas os parágrafos eram separados em unidades autônomas de texto, indicados por meio de traços horizontais. A forma de representar parágrafos se modifica no século V e se mantém durante a Idade Média. Nesse período, “a primeira letra da primeira linha de um novo parágrafo era colocada junto à margem e ampliada” (p. 229). O recuo de início de parágrafo, um espaçamento definido a partir da margem esquerda do texto, é datado a partir dos anos 1600.

⁵⁴ Mudanças em práticas de leitura denunciam, de modo amplo, mudanças sociais. No caso das sociedades ocidentais, práticas de leitura silenciosa são frequentemente associadas à expansão da imprensa na Europa (o que implica dominar técnicas de impressão e de produção do papel em larga escala) e, principalmente, à expansão de leitores (decorrente da ampliação de população alfabetizada).

subjaz outros processos mais amplos de práticas de leitura e escrita, especialmente práticas letradas promovidas pela escola nessa etapa da educação básica.

Quanto ao registro de fronteira de palavra por hífen, faz-se importante chamar atenção para sua natureza “ambígua”, na medida em que o hífen tanto delimita, quando une palavras, como pode ser atestado a partir da caracterização desse sinal gráfico dada em dicionários on-line de língua portuguesa, como os exemplificados abaixo.⁵⁵

Sinal gráfico que une: os elementos de um composto; um verbo a pronomes oblíquos enclíticos e mesoclíticos; e indica partição de sílabas de um vocábulo; traço de união, tirete. (Michaelis Moderno Dicionário da Língua Portuguesa)

Sinal gráfico (-) de separação de elementos de um composto, de alguns prefixos, de sílabas em fim de linha e de ligações enclíticas, mesoclíticas ou proclíticas. (Dicionário Priberam da Língua Portuguesa)

Em relação à fronteira de palavra delimitada por hífen, constata-se que o contexto é restrito a estruturas constituídas de palavras compostas e de verbo e pronome. Portanto, o hífen sinaliza certos tipos de relações morfossintáticas entre palavras ortográficas, o que o particulariza em relação ao branco que delimita fronteira entre palavras de modo mais geral.

No material analisado, há quatro tipos de registros não-convencionais do hífen, sendo que apenas dois deles implicam segmentação não-convencional de palavras, como descrito no Quadro 2.1. Outros dois tipos de registros se caracterizam por troca do hífen por branco ou vice-versa e esses não afetam a identificação de fronteira de palavra, como pode ser atestado nos exemplos do quadro que segue.

Quadro 2.1: Usos não-convencionais do hífen

Tipos de registro da fronteira	Classificação do registro	Exemplos
Presença do hífen dentro de palavra	Hipersegmentação	“estava-mos”, “tive-se”
Ausência do hífen entre palavras	Hiposegmentação	“chamalo”, “guarda roupa”
Uso do hífen quando previsto branco	Troca de branco por hífen	“chegando-lá”, “certo-dia”
Uso do branco quando previsto hífen	Troca de hífen por branco	“guarda roupa”, “casando se”

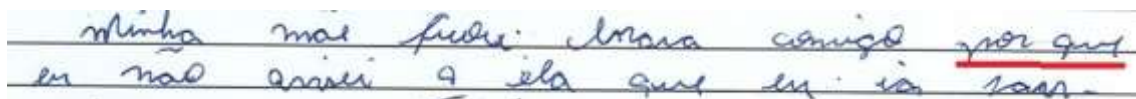
Nesta tese, apenas os usos não-convencionais do hífen que levam à hipo e à hipersegmentação de palavra serão considerados, pois os demais tipos envolve a seleção

⁵⁵ A referência a definições encontradas em dicionários on-line parte da suposição de que alunos e professores do EF II poderiam ter facilmente acesso a essas informações.

do registro gráfico para representar a fronteira de palavras, mas não altera a identificação do *locus* onde a fronteira é registrada. Os registros de hífen que envolvem alteração da localização de fronteira de palavra evidenciam, por hipótese, pistas do processo de segmentar o enunciado em palavra ortográfica. Particularmente, os dados do Quadro 2.1 dão pistas do processo de representação da informação verbo-visual que particulariza o hífen frente ao espaço em branco na delimitação de fronteira de palavra.

Por fim, quanto à consideração de critério morfossemântico, é importante identificar grafia não-convencional de formas homônimas, como “por que” quando previsto “porque”, cujas grafias implicam identificação de palavra e/ou locução. No enunciado exemplificado na Figura 2.5, constata-se o registro de “porque” hipersegmentado. A oração “eu não avisei a ela que eu iria sair” é apresentada como causa de “minha mãe ficou brava comigo”. Essa relação causal entre orações, expressa pela conjunção “porque” hipersegmentado, é bastante frequente no *cópus*, como será demonstrado na próxima seção.

Figura 2.5: Grafia de homonímia “por que”



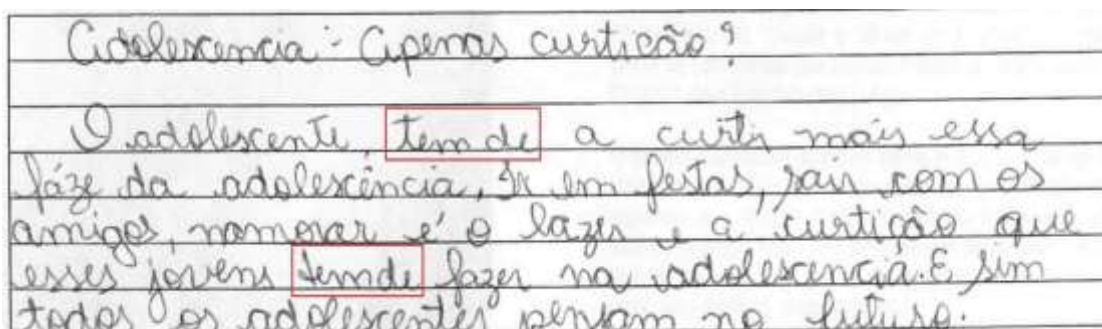
Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z11_8C_31M_07)⁵⁶

Ainda relacionado aos critérios de identificação de grafias não-convencionais, enfatiza-se a importância de se considerar o significado da palavra ou locução, bem como a construção sintática do enunciado em que a grafia ocorreu. Na Figura 2.6, a ocorrência “tem de” pode ser, em um primeiro momento, considerada grafia convencional, já que parece se tratar da forma verbal “tem” seguida da preposição “de”. Porém, verifica-se que se trata da forma flexionada “tende” do verbo “tender”, ao ser analisada a estrutura sintática e o significado do enunciado: “o adolescente tende a curtir”. Portanto, a grafia “tem de” foi considerada uma hipersegmentação da palavra ortográfica “tende”. No mesmo texto, em seguida, verifica-se a grafia “temde”, que poderia ser – a princípio – apenas exemplo de uso não-convencional da letra <m> para representar a coda nasal em “tende”. No entanto, a consideração da estrutura sintática e do significado das palavras

⁵⁶ Leitura possível: “minha mãe ficou brava comigo porque eu não avisei a ela que eu ia sair”.

no enunciado “e a curtição que esses jovens temde fazer na adolescência” leva à constatação de que se trata de hipossegmentação entre a forma verbal “tem” e a preposição “de”.

Figura 2.6: Grafias “tem de” e “temde”



Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z11_8C_31M_07)⁵⁷

Os critérios metodológicos adotados na identificação de segmentação não-convencional de palavras na escrita do EF II ora sistematizados são embasados em decisões tomadas no desenvolvimento de pesquisas feitas com base no mesmo banco de textos as quais foram tratadas em trabalhos anteriores: Tenani (2011a, b), Tenani e Paranhos (2011), Silva e Tenani (2014), Tenani e Fiel (2016). Em relação a esses trabalhos, esta tese avança ao explicitar que a complexidade encontrada principalmente na categorização do espaço em branco entre palavras é motivada pela natureza verbo-visual da escrita alfabética. Mais especificamente, a decisão do investigador implica classificar uma dada grafia como resultado de um padrão caligráfico do escrevente (portanto, relacionado à categorização gráfica apreensível na dimensão verbo-visual) ou como resultado de hipótese sobre a escrita convencional (portanto, relacionado à convenção ortográfica, também apreensível na dimensão verbo-visual). Essa classificação demanda confrontar práticas letradas/escritas conforme abonadas pelas convenções ortográficas da língua portuguesa a práticas letradas/escritas do escrevente em processo de inserção naquelas práticas letradas/escritas, muitas vezes, tomadas como “novidade”, porque são cobradas do estudante num contexto institucional específico, o da escola. Esse conflito entre diferentes práticas sociais letradas/escritas expõe, de nosso ponto de vista, a complexidade enunciativa da qual o sujeito é (e)feito: no chamado “erro”

⁵⁷ Leitura possível: “O adolescente tende a curtir mais essa fase da adolescência. Ir em festas, sair com os amigos, namorar; é o lazer e a curtição que esses jovens tem de fazer na adolescência. E “sim”, todos os adolescentes pensam no futuro”.

ortográfico, por hipo e/ou por hipersegmentação, o pesquisador pode levantar hipóteses interessantes e relevantes sobre a constituição dos sujeitos como escreventes em formação.

É salutar lembrar que essa problematização não se circunscreve ao dado de escrita do EF I ou EF II. De nossa perspectiva, nos estudos paleográficos e filológicos de textos manuscritos, é relevante que o investigador mobilize informações sobre aspectos verbo-visuais relacionados a formas gráficas de letras, a usos de diacríticos e espaços, por exemplo, que tenham sido estabelecidas por práticas sociais letradas/escritas privilegiadas em dado momento histórico.

De uma perspectiva paleográfica, por exemplo, Acioli (2003)⁵⁸ trata de critérios para leitura de documentos manuscritos do Brasil Colônia (séculos XVI e XVII) e lista conjunto de 18 normas para transcrição e edição de documentos brasileiros. Dentre essas normas (adaptadas para documentos brasileiros com base nas normas publicadas pela Escola de Estudos Medievais em Madri, no ano de 1944), uma trata da junção e separação de palavras: “As palavras ligadas entre si foram separadas, e as letras ou sílabas de uma mesma palavra quando separadas no texto, foram ligadas. Os apóstrofes foram respeitados” (ACIOLI, 2003, p. 60). Como se nota, a transcrição paleográfica de textos implica interpretação de limites de palavras manuscritas com vistas à leitura de historiadores.

Sob o viés que visa à constituição da história do PB, Oliveira (2006) analisa documentos da Sociedade Protetora dos Desvalidos do século XIX, encontrados em Salvador, e problematiza a dificuldade em identificar “se a escrita de vocábulos juntos ou a inscrição do branco em uma palavra resulta de manifestação voluntária das mãos de que saiu” (OLIVEIRA, 2006, p. 239).⁵⁹ O autor discorre sobre práticas de segmentação em textos antigos (fazendo referência inclusive a estudos sobre a escrita da Carta de Pero Vaz de Caminha) e afirma que hiper e hipossegmentações “são manifestações condicionadas, longe de testemunharem sobre mãos pouco diligentes” (OLIVEIRA, 2006, p. 241).

Em resumo, a problematização de critérios metodológicos adotados na identificação de segmentação não-convencional de palavra escrita é relevante não apenas pela necessidade de rigor metodológico, mas especialmente por colocar em destaque

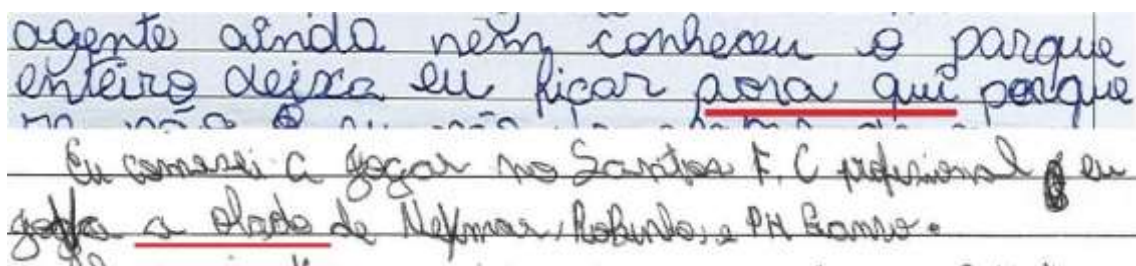
⁵⁸ Cabe aqui um agradecimento à professora Dra. Gladis Massini-Cagliari (UNESP/Araraquara) pela indicação dessa referência bibliográfica para tratar dos aspectos metodológicos envolvidos na identificação de segmentação não-convencional de palavra.

⁵⁹ Outro agradecimento deve ser dirigido ao professor Dr. Márcelo Módulo (USP) pela indicação dessa referência bibliográfica.

práticas letradas de leitura e escrita que dizem respeito à dimensão verbo-visual inscrita na categorização do espaço em branco da escrita alfabética.

Explicitados os critérios para a identificação de segmentação não-convencional de palavra, passa-se a tratar brevemente de um terceiro tipo de segmentação não-convencional de palavras que foi nomeado *mesclas* por Chacon (2006) e *híbridos* por Cunha (2004, 2010).⁶⁰ Na amostra analisada de escrita do EF II, apenas 12 ocorrências desse tipo foram identificadas, como as exemplificadas na Figura 2.7: “pora qui” (“por aqui”), produzida por um aluno de sexto ano, e “a olado”, produzida por um aluno de oitavo ano.

Figura 2.7: Grafias de “pora qui” e “a olado”



Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z08_5C_19F_06 e Z10_7B_05M_02)⁶¹

Na ocorrência “pora qui”, verifica-se hipossegmentação entre as fronteiras de “por” e “aqui”, seguida de hipersegmentação em “aqui”, separando a sílaba pretônica “a” de “qui”. Cunha (2004) descreve como regular nos dados de escrita do EF I a ocorrência de hipossegmentação seguida de hipersegmentação. No entanto, na amostra de escrita do EF II investigada, também foi encontrado registro de hipersegmentação seguida de hipossegmentação, como em “a olado” (“ao lado”). A descrição feita em Silva e Tenani (2014, p. 163) sobre esse último registro demonstrou a relevância em aprofundar a análise desse tipo de segmentação não-convencional de palavras, haja vista a complexidade de fatores linguísticos potencialmente envolvidos. Nesta tese, é feita opção

⁶⁰ A diferença terminológica entre *mesclas* e *híbridos* decorre da diferença de abordagens assumidas pelos autores. Como explicitado em Silva e Tenani (2014, p.7), Cunha (2004, 2010) classifica como híbridos dados em que primeiro ocorreu hipossegmentação e depois hipersegmentação, enquanto Chacon (2004) denominou *mesclas* grafias que envolvem os dois tipos de segmentação não-convencional e, também, grafias que, dada a sua estrutura, não podem ser explicadas apenas em função da classificação em hipo ou hipersegmentação, uma vez que a ausência/presença do limite gráfico está presente “em partes nas quais deveria (ou se suporia) haver um *limite* ao mesmo tempo ortográfico e prosódico” (CHACON, 2004, p. 226, grifo no original).

⁶¹ Leitura possível: “a gente ainda nem conheceu o parque inteiro deixa eu fiar por aqui porque” (trecho 1). “Eu comecei a jogar no Santos F. C. profissional eu jogava ao lado de Neymar, Robinho e PH Ganso”.

por excluir da análise linguística dados dessa natureza em função de a complexidade que lhes é característica demandar discussão aprofundada de temas que escapam dos objetivos inicialmente definidos. Enfatizo, no entanto, que essas grafias carecem de estudo específico dada a complexidade que as particularizam em relação aos dois outros tipos de segmentação não-convencional de palavras. A baixa ocorrência desses dados na escrita do EF II constitui diferença em relação à escrita do EF I, que se caracteriza por presença de mesclas/híbridos principalmente nos anos iniciais (cf. TORQUETTE, 2016; CUNHA, 2004).

Por meio das considerações acerca da constituição da segmentação não-convencional de palavra feitas nesta seção, aspectos metodológicos quanto à constituição do dado de escrita podem ser problematizados e os tipos de segmentação de palavra nos textos do EF II podem ser caracterizados. Simultaneamente, argumento que decisões metodológicas relativas à identificação e classificação do dado de escrita implicam mobilizar práticas letradas de leitura e escrita que dizem respeito à dimensão verbo-visual inscrita na categorização do espaço em branco da escrita alfabética. Por meio dessa argumentação, mais do que responder as questões (explicitadas no início desta seção) relativas à constituição da segmentação de palavra não-convencional enquanto dado de escrita, defendo ser esse um dado de língua e de discurso, como será demonstrado nos capítulos 3 e 4.

Antes, porém, de a análise dos dados avançar, faz-se necessário caracterizar as segmentações não-convencionais de palavra na amostra longitudinal de escrita do EF II, o que será feito na próxima seção.

2.3 Caracterização de segmentações de palavra na escrita do EF II

Nesta seção, inicia-se a caracterização da escrita do EF II a partir da descrição da frequência de ocorrência de hipossegmentações e hipersegmentações. Busca-se responder às seguintes questões: (i) qual é a frequência de segmentação não-convencional de palavra ao longo do EF II?; e (ii) há características que particularizam essas segmentações de palavra da escrita do EF II em relação às aquelas identificadas na escrita do EF I?

A caracterização quantitativa das ocorrências dos tipos de segmentação de palavra nos textos do EF II que constituem a amostra longitudinal analisada pode ser feita com

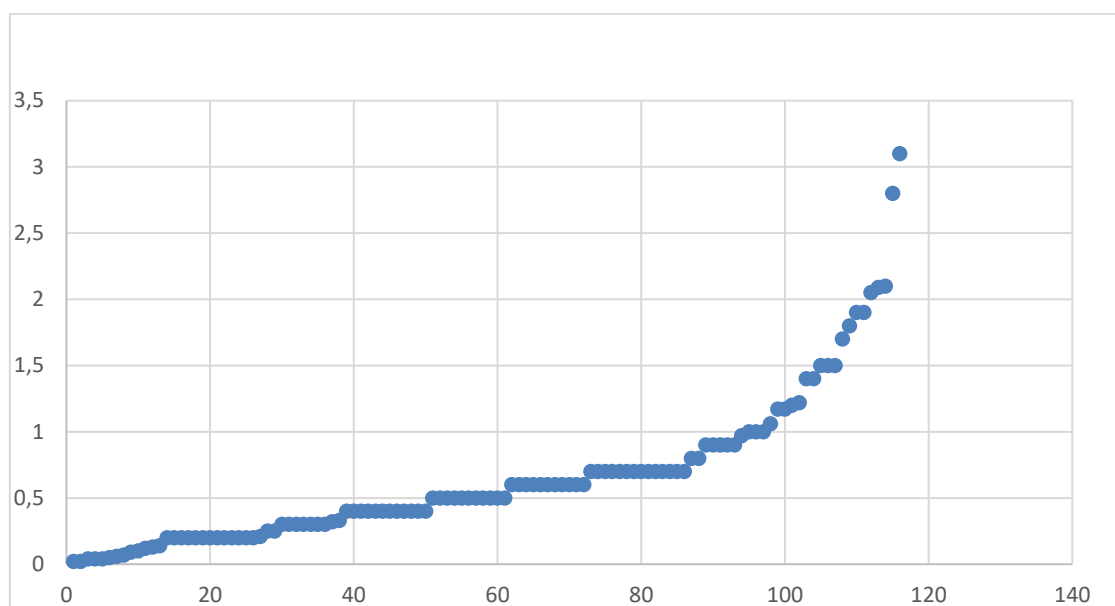
os seguintes dados. Os 2.457 textos dessa amostra totalizam 344.917 palavras, tendo sido identificadas 2.095 segmentações não-convencionais de palavras (índice de 0,6 segmentações não-convencionais/palavra escrita) que podem ser classificadas em: 1.002 hipersegmentações (47,82%) e 1.093 hipossegmentações (52,17%). Esse resultado geral de segmentação não-convencional de palavras na amostra de escrita do EF II reitera a tendência em haver, de modo geral, menos hipersegmentações do que hipossegmentações em textos de escrita do EF I, tal como observado, por exemplo, em textos de crianças do interior paulista por Capristano (2007a) e Torquette (2016) e do interior gaúcho por Cunha (2004).

Considerando-se que as segmentações não-convencionais de palavras (i) já não são esperadas no EF II, dado o tempo de escolarização (cf. habilidades ortográficas definidas para o EF II pela Secretaria Estadual da Educação de São Paulo, 2009) e (ii) são interpretadas, de certa perspectiva das áreas de Educação e Saúde, como indícios de *disortografia* (discutido na seção 1.1) quando muito frequentes na produção escrita de um mesmo sujeito, investigou-se como se deu a distribuição da frequência da segmentação não-convencional de palavras para cada sujeito a partir de três índices de “erros”. Os índices são calculados a partir da razão entre: (i) hipersegmentação X total de palavras escritas, (ii) hipossegmentação X total de palavras escritas e (iii) soma das hipossegmentação e hipersegmentação X total de palavras escritas. Esses índices foram calculados para cada um dos 116 sujeitos e em cada ano letivo.

A seguir, visualiza-se o gráfico de índice obtido a partir da razão entre o total de segmentação não-convencional de palavra (hiper e hipossegmentação) e o total de palavras escritas ao longo dos quatro anos do EF para cada sujeito.⁶²

⁶² Exemplo do cálculo realizado: o sujeito 114 teve 63 dados de segmentação não-convencional de palavras em um total de 2.228 palavras escritas ao longo dos quatro anos do EF. Portanto, o cálculo é: $63/2228 = 0,02827$. O índice é obtido multiplicando-se o resultado da divisão por 100. Para o exemplo: $0,02827 \times 100 = 2,827$. Na elaboração do gráfico 1, foram desprezadas as duas últimas casas decimais.

Gráfico 2.1: Índice de segmentação não-convencional de palavras por sujeito



Onde: no eixo horizontal, são indicados os 116 sujeitos da pesquisa; no eixo vertical, o índice de segmentação não-convencional de palavras.

De modo geral, o Gráfico 2.1 permite visualizar que a maioria dos sujeitos tem baixos índices de ocorrência de segmentação não-convencional de palavras. No conjunto desses 116 sujeitos, a média geral é de 0,6 dados/sujeito, havendo somente dois meninos com índices muito mais altos do que o conjunto de sujeitos analisados (índices de 2,8 e 3,1 dados/sujeito). Porém, há apenas três meninos (2,58%) e 14 meninas (12%) que concluem o EF II sem qualquer segmentação não-convencional de palavras.⁶³ A distribuição desse índice entre os sujeitos permite estabelecer faixas com intervalo de 0,5 pontos entre si, de modo a identificar tendências para o grupo de sujeitos. Na Tabela 2.3, visualizam-se cinco faixas de índices e o número de sujeitos identificados para cada faixa, bem como as respectivas porcentagens.

⁶³ Dos 116 alunos, 28 meninas (24,1%) e 14 meninos (12%) concluem o EF II sem qualquer hipersegmentação; 23 meninas (19,8%) e 11 meninos (9,5%), sem hipossegmentação.

Tabela 2.3: Faixas de índices de segmentação não-convencional de palavras por sujeito

Faixas de índices	Nº de Sujeitos	%
0,02 a 0,5	61	52,59
0,60 a 1,0	36	31,03
1,06 a 1,5	10	8,62
1,70 a 2,1	07	6,03
acima de 2,1	02	1,72

Os dados da Tabela 2.3 permitem constatar que: (i) a maioria dos sujeitos se concentra nas duas primeiras faixas de índices de dados, totalizando 83,62% e sendo predominante para esse grupo (52,59%) os índices inferiores a 0,5 dados/sujeito; (ii) uma minoria dos sujeitos, que corresponde a 6,03%, apresenta índice três vezes maior do que 0,5 dados/sujeito, índice da maioria dos alunos, havendo apenas dois sujeitos com índices que chegam a ser seis vezes maiores do que 0,5 dados/sujeito.

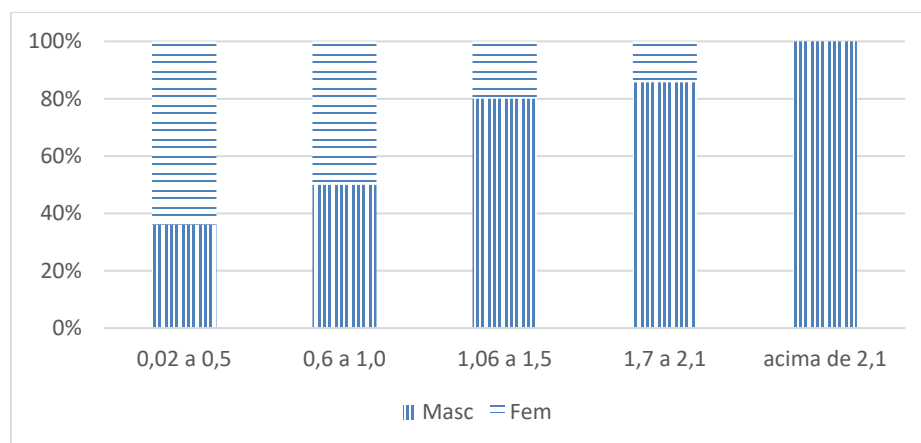
Por meio deste resultado, confirma-se que: (i) a maioria dos alunos do EF II tende a produzir poucas segmentações não-convencionais de palavras, embora seja esperado que nenhuma dessas grafias ocorra nessa etapa escolar; (ii) alunos que apresentam muitos registros não-convencionais de palavras devem receber atenção particular do professor,⁶⁴ pois não é essa a característica da escrita no EF II. Esses últimos alunos potencialmente seriam diagnosticados como portadores de disortografia. Porém, esta tese se afasta dessa abordagem que atribui ao aluno diagnóstico de distúrbio de escrita e busca caracterizar – do campo dos estudos linguísticos – a frequência de segmentação não-convencional de palavra na amostra longitudinal de escrita do EF II com o objetivo de identificar características gerais dessa escrita.

Dentre esse universo de 116 sujeitos, saltam aos olhos, por meio do gráfico abaixo, que há concentração na distribuição das meninas em relação aos meninos nas duas primeiras faixas de índices, os quais se caracterizam por serem os mais baixos em relação ao verificado para o grupo com base na tabela anterior. As meninas totalizam 49,14%, enquanto os meninos, 34,48% dos 83,62% dos sujeitos que pertencem às duas primeiras faixas de índices de dados/sujeito. Essas porcentagens de meninos e meninas com

⁶⁴ Vale observar que o domínio das convenções ortográficas é avaliado pela Secretaria da Educação paulista, via SARESP, como parte das habilidades vinculadas ao tema “Reflexões para os usos da língua falada e escrita” (SÃO PAULO, 2009).

melhores desempenhos não decorrem dos totais de meninos e meninas no conjunto de sujeitos, pois se verifica que as meninas constituem 51,71% da amostra de sujeitos e alcançam a taxa de 58,76% dentre os sujeitos com os menores índices de dados/sujeito. Outro fato a ser destacado está na concentração de meninos nas faixas dos mais altos índices de dados/sujeito.

Gráfico 2.2: Relação entre sexo/gênero e faixas de índice de segmentação



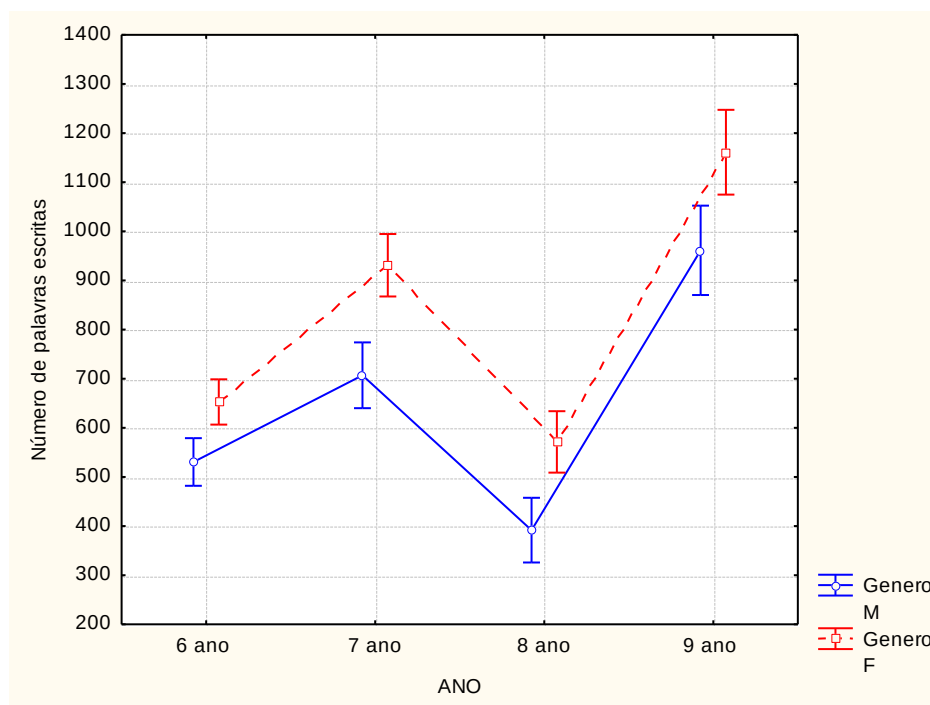
Esses resultados quanto ao índice de segmentação não-convencional de palavras ao longo do EF II conduziram à investigação do efeito do sexo/gênero⁶⁵ dos sujeitos na frequência de segmentação não-convencional de palavra.

Para obter resultados que pudessem confirmar ou refutar esse efeito, parti do levantamento de frequência de palavras na amostra longitudinal em função das variáveis ano letivo e sexo/gênero dos sujeitos, os quais foram submetidos, posteriormente, à análise estatística.⁶⁶ Os resultados são apresentados no Gráfico 2.3. Em seguida, foram investigadas as frequências de hipossegmentações e de hipersegmentações separadamente, em função do número de palavras escritas por anos letivos e por sexo/gênero dos sujeitos.

⁶⁵ Optou-se por “sexo/gênero”, termo usado predominantemente em Sociolinguística, em detrimento de “gênero”, termo usado nos estudos sobre gênero e sexualidade, a fim de evitar confusão com a noção de “gênero” textual/discursivo. Além disso, foi considerada apenas a informação quanto ao sexo dos sujeitos, não tendo sido feito levantamento (ou entrevista, por exemplo) quanto à identidade de gênero dos alunos à época da coleta dos textos.

⁶⁶ O tratamento inferencial dos dados foi feito com o uso do *software Statistica* versão 7.0. Foi adotado 0,05 como nível de significância e foi utilizado o teste estatístico ANOVA de Medidas Repetidas, considerando-se como variável intergrupo o sexo/gênero (masculino e feminino), como variável intra-grupo o ano escolar (do 6º ao 9º ano) e sua respectiva interação sexo/gênero *versus* ano escolar. Essas análises foram conduzidas pela professora Dra. Larissa Berti (UNESP/Marília) a quem muito agradeço.

Gráfico 2.3: Número de palavras escritas em função de ano letivo e sexo/gênero



Do Gráfico 2.3, visualizam-se dois importantes resultados: (i) quanto ao ano letivo, constata-se haver aumento de palavras escritas do sexto ano para o nono ano, mas com queda significativa no oitavo ano; (ii) quanto ao sexo/gênero dos sujeitos, sistematicamente se verifica que meninas escrevem mais do que meninos em todos os anos letivos. Do tratamento inferencial dos dados e da análise estatística de medidas repetidas, atesta-se que houve diferença significativa para o sexo/gênero ($F(1,114)=26,667$, $p<0,00$) e ano escolar ($F(3,342)=176,770$, $p<0,00$), não havendo interação entre sexo/gênero e ano escolar. Confirma-se, estatisticamente, que meninas escreveram mais palavras do que meninos, independente do ano escolar. O teste *post hoc* de Scheffe (feito para assegurar a relevância dos resultados) mostrou que: (i) o número de palavras escritas pelos sujeitos, independentemente do sexo/gênero, foi diferente, considerando a comparação dos quatro anos entre si (todos com valor de $p<0,00$); e (ii) tanto para meninos quanto para meninas, o número de palavras escritas aumenta em função do ano escolar, exceto no oitavo ano (o número de palavras escritas em ordem decrescente é: $9^o>7^o>6^o>8^o$).

Cabe aqui uma breve observação quanto ao fato de a queda do número de palavras escritas no oitavo ano (em 2010) ter sido identificada na amostra longitudinal analisada. Essa queda de palavras coincide com o início do último ciclo do EF, quando conteúdos

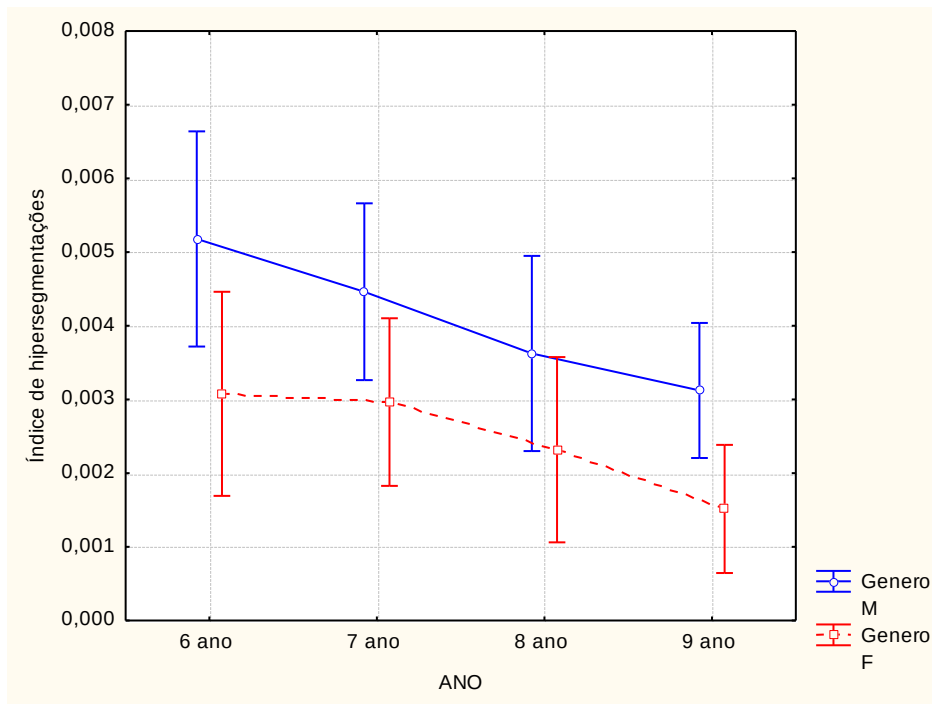
de língua portuguesa passam a ser trabalhados nesse ano letivo. Objetivamente, não foram encontrados na amostra dados que pudessem justificar a queda no número de palavras escritas no material coletado. A investigação de fatores que tenham motivado esses resultados quanto à produção escrita dos alunos no oitavo ano escapa dos objetivos desta tese.⁶⁷ De todo modo, da comparação entre início e fim do EF II, pode-se afirmar que: (i) os alunos passam a escrever mais em função do aumento do tempo de escolarização, (ii) meninas escrevem mais em relação aos meninos. Esses resultados, inéditos para o EF II, sugerem ser relevante considerar a variável sexo/gênero na caracterização da escrita do EF II.

Avanço na caracterização da amostra investigada apresentando resultados de frequência de segmentação não-convencional de palavras em relação aos anos letivos e ao sexo/gênero dos sujeitos. Os resultados de análise atestam que: (i) há diferença estatística significativa quanto ao sexo/gênero ($F(1,112)=10,9$; $p=0,001$), (ii) há diferença significativa entre os anos letivos quanto à frequência de segmentação não-convencional de palavras ($F(3,3)=6,49$; $p=0,000$) e (iii) não há interação entre sexo/gênero e anos letivos. A análise *post hoc*, feita por meio do teste Scheffe, revela que há diferença estatisticamente relevante somente no nono ano quando comparado aos anos sexto ($p=0,002$) e sétimo ($p=0,007$), mas esse ano não se diferencia em relação ao oitavo ($p=0,409$).

Faz-se relevante refinar essa análise de frequência de segmentação não-convencional de palavras por meio da consideração dos tipos de segmentação não-convencional de palavra em função das variáveis anos letivos e sexo/gênero dos sujeitos. Os resultados dessas análises são ilustrados por meio dos Gráficos 2.4 e 2.5.

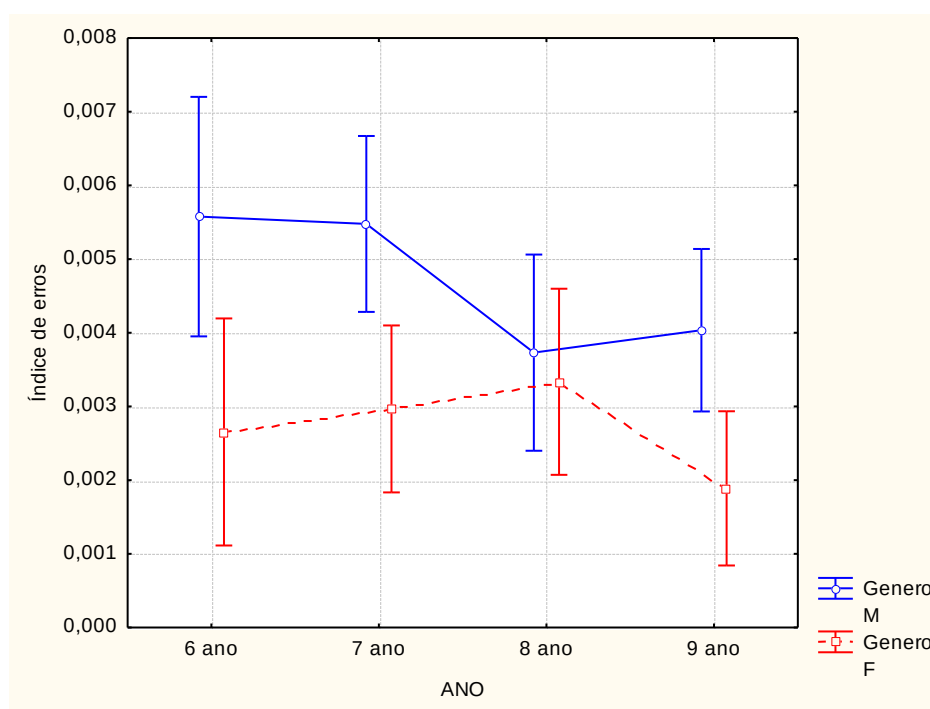
⁶⁷ Entretanto, vale observar que, em 2010, houve mudanças importantes na escola parceira decorrentes da mudança na direção da escola. Por 25 anos (até 2009), o diretor da escola, em conjunto com equipe pedagógica, adotava conjunto rígido de regras de convivência escolar, penalizando alunos que chegavam atrasados para o início da aula, por exemplo. A nova diretora alterou em boa parte as regras então vigentes, reorganizando, inclusive, a distribuição dos alunos em função do ano letivo, de sorte que todas as turmas de oitavo e nono anos passaram a ser oferecidas apenas no período matutino, enquanto as turmas de sexto e sétimo anos, no período vespertino.

Gráfico 2.4: Índice de hipersegmentação em função de ano e sexo/gênero



Do Gráfico 2.4, verifica-se que: (i) meninos produzem mais hipersegmentação do que meninas, independente do ano escolar, tendo sido atestada diferença estatisticamente significativa ($F(1,112)=6,97$, $p=0,009$) entre sexo/gênero dos sujeitos em relação à ocorrência de hipersegmentação; (ii) há queda da hipersegmentação em função do aumento do tempo de escolarização, tendo sido verificada diferença estatisticamente significativa entre os anos letivos. Para identificar em qual(uais) ano(s) é(são) diferente(s) entre si, foi aplicado o teste *post hoc* Scheffe. A análise mostra que somente no nono ano há diferença estatisticamente relevante em relação aos sexto ($p=0,004$) e sétimo ($p=0,046$) anos, mas esse ano não se diferencia em relação ao oitavo ($p=0,6115$).

Gráfico 2.5: Índice de hipossegmentação em função de ano letivo e sexo/gênero



Do Gráfico 2.5, visualiza-se que: (i) meninos produzem mais hipossegmentação do que meninas, independentemente do ano escolar, tendo sido atestada diferença significativa ($F(1,112)=9,96$, $p=0,002$); (ii) há decréscimo numérico de hipossegmentações ao longo dos anos letivos, mas não foi atestada diferença estatística significativa entre os anos de escolarização.

A comparação entre os Gráficos 2.4 e 2.5 permite verificar que: (i) há frequências que são estatisticamente significativas tanto de hipersegmentação quanto de hipossegmentação em relação ao sexo/gênero dos sujeitos analisados; (ii) há frequência de hipersegmentação, mas não de hipossegmentação, estatisticamente relevante quanto aos anos escolares. Em resumo, pode-se afirmar, com base na análise da amostra longitudinal de textos do EF II, quanto às duas variáveis consideradas, que: (i) meninas escrevem mais palavras em seus textos e tendem a seguir a convenção ortográfica quanto aos critérios de segmentação de palavras e (ii) o aumento de anos de escolarização contribui significativamente para o declínio da hipersegmentação, mas não significativamente para a queda da hipossegmentação.

Os resultados descritos não apenas demonstram haver diferença relacionada a sexo/gênero do sujeito na escrita do EF II, mas também sugerem que essas diferenças possam ser observadas em período anterior, na escrita do EF I, e período posterior, na

escrita de jovens e adolescentes. De fato, Torquette (2016) descreve, para dados de escrita do EF I oriundos da rede pública e privada do município de Marília (SP), resultados quantitativos que confirmam que: (i) meninos escrevem menos palavras do que meninas, seja em escola pública ou privada; (ii) tanto meninos quanto meninas produzem menos segmentações não-convencionais de palavras⁶⁸ à medida que aumenta o tempo de escolarização, seja em escola pública ou privada e (iii) meninos produzem mais segmentações não-convencionais de palavra do que meninas, mas a diferença de sexo/gênero apenas determina a porcentagem de segmentações não-convencionais em escola pública. Ao discutir esses resultados, a autora afirma que:

Nossos resultados indiciam não haver, pois, simplesmente uma diferença de gênero/sexo, mas uma diferença socialmente construída, a partir de discursos que geram expectativas e diferentes construções do gênero/sexo nos diferentes extratos sociais. [...] A diferença de desempenho intergêneros em escolas públicas pode ser explicada, portanto, pelo fato de que as expectativas projetadas nas meninas desse contexto envolvem trabalhos que mais exigem fala e escrita na norma padrão, o que as incentivam a estudarem mais, enquanto os meninos, para exercerem melhor suas tarefas (menos verbais) fazem cursos profissionalizantes. (TORQUETTE, 2016, p. 124).

Essa projeção de expectativas quanto aos papéis sociais a serem desempenhados por meninos e meninas varia em função da classe social, segundo a autora. Desse modo, nas escolas privadas (frequentada por crianças de classes sociais mais altas em relação àquelas de escolas públicas), “as projeções feitas para as crianças já supõem que suas futuras tarefas terão a escrita como um elemento de destaque – ou de bastante importância –, como as profissões com formação universitária privilegiadas socialmente” (TORQUETTE, 2016, p. 127). Ainda supõe que práticas de escrita teriam mesmo grau de importância para meninos e meninas de escolas particulares, mas importância diferenciada para meninos e meninas de escolas públicas. Com base nessa hipótese, práticas letradas relacionadas aos (supostos) papéis sociais a serem desempenhados por meninos (pertencentes a classes sociais mais baixas) seriam menos valorizadas em relação àquelas práticas letradas que se supõem desempenhadas por meninas (pertencentes também a classes sociais mais baixas) que teriam diferentes papéis sociais em relação aos meninos.

⁶⁸ Não foi investigada a frequência de hipossegmentação e de hipersegmentação na amostra de textos infantis analisada por Torquette (2016), de modo que não é possível comparar os resultados da escrita do EF I em relação às frequências de hipossegmentação e de hipersegmentação na escrita do EF II aqui analisadas.

Os resultados de haver um efeito do sexo/gênero na frequência de segmentação não-convencional de palavras em textos de crianças seriam, pois, efeitos de projeções sociais de práticas de escrita associadas aos papéis sociais que potencialmente atravessariam meninos e meninas de modos distintos em função das classes sociais constituídas sócio-historicamente no país. Essa interpretação dada aos resultados sobre o efeito do sexo/gênero na frequência de segmentação não-convencional de palavras em textos de escrita do EF I (produzidos por crianças do centro-oeste paulista) pode ser estendida aos resultados sobre esse mesmo efeito que foram descritos a partir da análise de textos de escrita do EF II (produzidos por adolescentes do noroeste paulista), com base na hipótese de que não haveria diferenças substanciais entre os municípios paulistas onde foram desenvolvidas as pesquisas no que diz respeito às potenciais relações entre escrita, sexo/gênero e classes ou papéis sociais.

Ainda vale observar que esses resultados quanto a haver diferença entre sexo/gênero do aluno e características de sua escrita produzida em ambiente escolar seguem na direção dos resultados divulgados, em 2012, pela Organização para Cooperação e Desenvolvimento Econômico (OCDE) no âmbito do Programa Internacional de Avaliação de Estudantes (PISA). Dados do PISA 2012 demonstram haver diferença entre sexo/gênero relacionadas a desempenho de jovens com idade acima de 15 anos em testes de avaliação de competências de leitura, matemática e ciências, sendo melhor o desempenho de meninas nos testes de competência em leitura,⁶⁹ mas pior nos testes de competência em matemática e ciências em todos os países participantes do PISA, incluído o Brasil.⁷⁰

Em âmbito nacional, não há informações governamentais⁷¹ sobre relação entre sexo/gênero e desempenho escolar dos alunos. Observa-se que o Ministério da Educação (por meio do SAEB – Sistema Nacional de Avaliação da Educação Básica) divulga o IDEB que é produto da avaliação do desempenho geral de escolas do país em língua portuguesa e matemática (como já descrito na seção anterior), mas não disponibiliza

⁶⁹ Diferenças de desempenho nas avaliações de competência de leitura entre meninos e meninas são discutidas pela OCDE no volume 01/2012 do *PISA em foco*, intitulado “Meninos e meninas estão preparados para a era digital?”.

⁷⁰ Essa diferença entre sexos/gêneros tem motivado, por exemplo, políticas de inclusão de meninas/mulheres nas áreas de exatas (como física, química, engenharias) que envolvem competência matemática como forma de estimular a inserção e o desenvolvimento de mulheres nessas áreas. Também tem sido despertada a atenção para dificuldades que meninos/homens apresentam em atividades que envolvem práticas de leitura, a fim de serem investigadas motivações para dificuldades dessa natureza.

⁷¹ Fazer referências a índices governamentais de desempenho dos estudantes não implica compactuar com os pressupostos sobre escrita e ortografia, por exemplo, que são subjacentes a esses instrumentos de mensuração de qualidade de educação.

índices sobre desempenho dessas competências em função do sexo/gênero dos alunos. Igualmente no âmbito do estado paulista, a Secretaria da Educação, por meio do SARESP, coleta dados socioeconômicos dos alunos, mas não disponibiliza informações do IDESP relacionando desempenho dos alunos em função do sexo/gênero.⁷²

Pesquisas na área de Educação problematizam os tímidos avanços nas políticas públicas nacionais quanto à igualdade de sexo/gênero (veja-se, por exemplo, Vianna e Unbehaun, 2006; 2004) e as práticas pedagógicas que implicam construção de representações sociais assimétricas entre meninos e meninas na escola (cf. LOURO, 2008; FELIPE, 2000). Essas temáticas sobre representações sociais e sexo/gênero e educação têm características do campo da Psicologia Social (cf. AMÂNCIO, 1993), que nasce a partir do movimento feminista dos anos 1960, principalmente, mas ganha, contemporaneamente, contornos amplos e relevantes para reflexões sobre representações simbólicas gestadas em sociedades ocidentais que reiteram assimetrias supostamente dadas pela natureza, como aquelas que estariam vinculadas ao sexo e à raça dos indivíduos, por exemplo. No âmbito dessas discussões de viés social e educacional, Stromquist (2001) estabelece relações entre sexo/gênero e conceitos de letramento (formais e informais), buscando argumentar que práticas de leitura e escrita carecem de serem abordadas de uma perspectiva que levem à conquista da cidadania por parte, especialmente, de mulheres pobres.

Nesse cenário em que práticas de leitura e escrita são discutidas como parte dos requisitos necessários para ampliar o acesso à cidadania, os resultados descritos nesta seção sobre a existência de diferença de incidência de segmentações não-convencionais de palavras relacionada ao sexo/gênero dos escreventes ganham contornos que transcendem a descrição de características da escrita do EF II em particular. Esses resultados inéditos são bases para aprofundar estudos (nas áreas da Educação, Ciências Sociais e Linguística, por exemplo) que venham a considerar sexo/gênero como característica social dos sujeitos relevante para descrição e análise da escrita produzida/ensinada em ambiente escolar.

No âmbito dos estudos linguísticos, mais especificamente da Sociolinguística, é dada importância à consideração de sexo/gênero dos informantes na descrição de usos da língua, notadamente na descrição de aplicação de regras variáveis da fala. Exemplos de discussão atual no país sobre a relação entre sexo/gênero e usos da língua no campo da

⁷² Os resultados do SARESP, divulgados por meio de boletins consolidados, restringem-se a médias obtidas pelos alunos de uma dada escola para cada disciplina avaliada em uma escala de competência.

linguística são encontrados em Freitag e Severo (2015). De modo específico, Freitag (2015) parte da abordagem dada pela Sociolinguística Laboviana à variável sexo/gênero com o propósito de argumentar a favor da necessidade de se considerarem mudanças dos papéis sociais das mulheres, uma vez que esse campo de investigação prima por investigar relação entre língua e sociedade. A autora tece essa problematização “em três movimentos”, a saber: “1) a contextualização da variável sexo/gênero; 2) a sua apropriação no domínio da Sociolinguística; e 3) a revisão das abordagens para a variável sexo/gênero” (FREITAG, 2015, p.18) e conclui que encontra “uma mudança da correlação entre uso da língua e sexo para as dimensões simbólicas e ideológicas da linguagem” (FREITAG, 2015, p.67), ao identificar que estudos recentes dão atenção às ideologias da masculinidade e da feminilidade diferentemente da maioria dos estudos da área que interpretam diferenças linguísticas relacionadas à sexo/gênero como expressão do valor econômico de variedades de prestígio no mercado de trabalho. A autora defende revisão da concepção de sexo/gênero no campo da Sociolinguística a fim de alçar compreensão e explanação mais amplas dos resultados que descrevem relações entre sociedade e usos da língua.

No que diz respeito aos usos da escrita, esta tese – de outra perspectiva que não a da Sociolinguística – contribui com resultados que sinalizam como um campo fértil aos estudos linguísticos a investigação dos modos pelos quais diferenças atestadas na escrita de meninos e meninas relacionadas ao chamado desempenho ortográfico possam estar mais amplamente correlacionadas a representações sociais da escrita e de sexo/gênero dos escreventes.

Com base na perspectiva de escrita assumida nesta tese, a ser detalhada na seção 4.1, interpreta-se que os resultados ora descritos são efeitos de representações sociais de práticas letradas/escritas e orais/faladas que atravessam meninas/mulheres e meninos/homens de maneira distinta nas sociedades ocidentais. Argumentos que sustentarão essa interpretação serão dados no último capítulo desta tese.

No entanto, a descrição feita até o momento sobre a frequência de segmentação não-convencional de palavra ao longo do EF II permitiu responder a questão inicial desta tese quanto à caracterização desse tipo de grafia não-convencional de palavras na escrita do EF II, tendo sido identificadas características comuns à escrita do EF I: (i) decréscimo de frequência de segmentação não-convencional de palavras em função do aumento do tempo de escolarização e (ii) diferenças quanto à frequência de segmentação não-convencional de palavra em função do sexo/gênero dos alunos (sendo menor a frequência

nos textos de meninas do que nos textos de meninos). Quanto a características que particularizam a escrita do EF II em relação à escrita do EF I, destacamos a baixa ocorrência de mesclas/híbridos. Outras características da escrita do EF II serão apresentadas no próximo capítulo.

Dos resultados descritos, destaco o fato de ter sido atestada relevância estatística apenas para o efeito do ano letivo na ocorrência de hipersegmentação, tanto em textos de meninos quanto de meninas na escrita do EF II. Com base nesse resultado, o desenvolvimento da análise de dados se restringirá às hipersegmentações de palavra, pois o objetivo é descrever característica da escrita do EF II também em função dos anos letivos. Embasa-se esta decisão não apenas nos resultados quantitativos descritos, mas também na interpretação de que há diferença qualitativa entre hipersegmentação e hipossegmentação quanto às características linguísticas que englobam práticas letradas/escritas mobilizadas em um e outro dado na linha do que Capristano (2007a) sinalizou para dados de escrita do EF I.

Antes de passar à análise de características das hipersegmentações, no capítulo 3, retomo aqui informações dos Gráficos 2.1 e 2.4 a fim de problematizar o índice de “erro” de hipersegmentação do sujeito, tendo em vista as diferenças descritas entre meninos e meninas quanto às características de suas escritas. Do Gráfico 2.1, recupero o fato de haver variação intersujeitos quanto à taxa de segmentação não-convencional de palavras e, do Gráfico 2.4, a constatação de que mais tempo de escolarização leva à menor incidência de hipersegmentações no EF II tanto para meninos quanto para meninas.

Embora descritas essas duas tendências importantes das hipersegmentações no EF II, cabe chamar a atenção para o comportamento variável intersujeitos quanto aos registros não-convencionais de palavras em seus textos manuscritos. Nos Gráficos 2.6 e 2.7, é exemplificada a variação do índice de hipersegmentação para um conjunto de cinco meninas e de cinco meninos, selecionados por representarem as possibilidades observadas no conjunto dos sujeitos analisados que não grafaram convencionalmente fronteiras gráficas de palavras em sua língua materna.

Gráfico 2.6: Índices de hipersegmentação por sujeitos fem./anos letivos

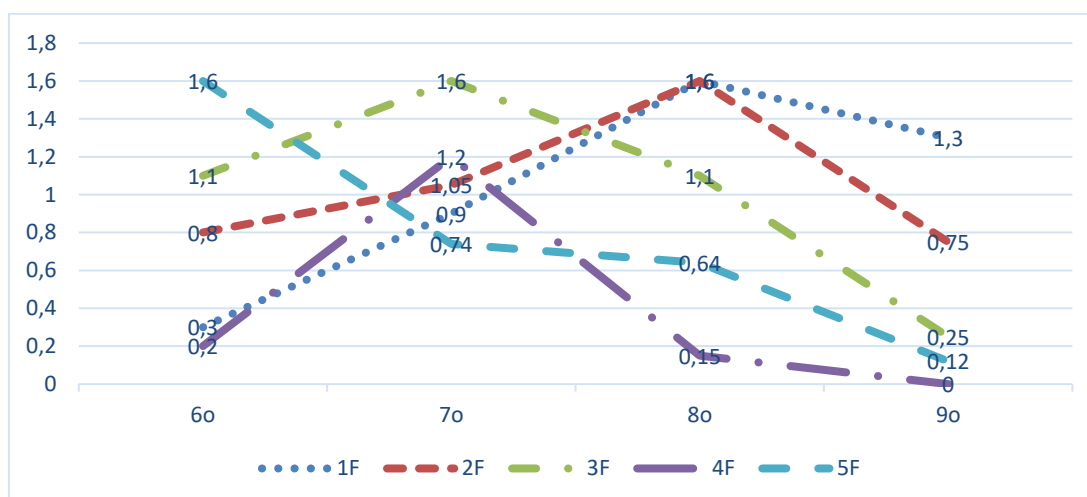
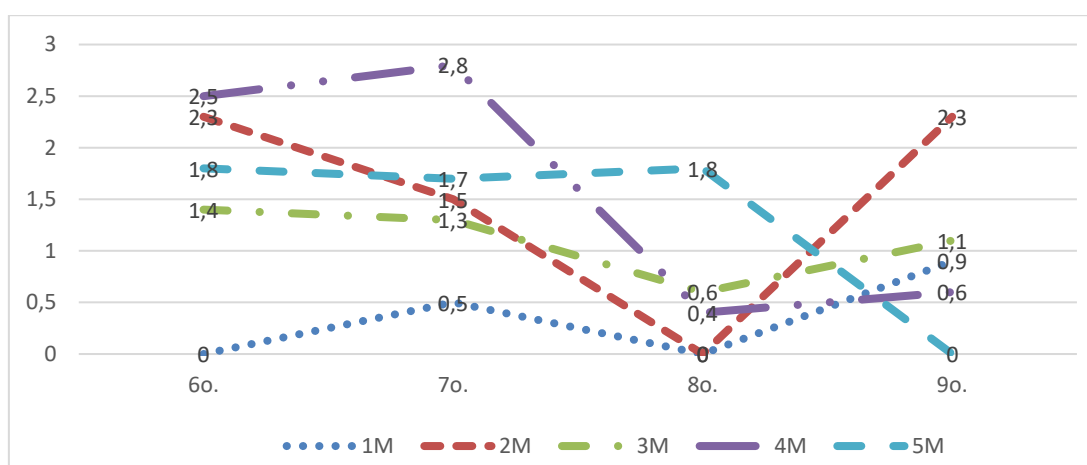


Gráfico 2.7: Índices de hipersegmentação por sujeitos masc./anos letivos



Nos Gráficos 2.6 e 2.7, visualiza-se como se dá a variação de hipersegmentação para cada sujeito em função dos anos letivos. Verifica-se que há, ao longo do EF II, tanto entre meninas quanto entre meninos, três trajetórias: (i) aumento dos índices de hipersegmentação nos anos finais do EF II (ex: sujeito 1F e 1M), (ii) manutenção dos índices de hipersegmentação nos mesmos patamares quando comparados aos anos inicial e final do EF II (ex: sujeito 2F, 2M e 3M) e (iii) decréscimo dos índices, havendo trajetória de decréscimo abrupto (cf. sujeito 5M) ou decréscimo progressivo, tendendo a poucos registros não-convencionais entre os sujeitos com índices iniciais altos (ex: sujeitos 3F, 5F, 4M) e à convenção ortográfica entre sujeitos que iniciam o EF com poucos registros não-convencionais (ex: sujeito 4F).

O fato de haver sujeitos que continuam ou que passam a grafar não-convencionalmente fronteiras de palavras ao final do EF, motivou-nos a investigar quais palavras continuam ou passar a ser grafadas não-convencionalmente. A fim de obter respostas a essa questão, foi feito levantamento dos tipos de hipersegmentação identificados nas produções textuais e, em seguida, obtido índice da relação entre tipo de hipersegmentação (*types*) e número de ocorrências (*tokens*) de cada tipo. Visualizam-se possibilidades de trajetórias dos sujeitos femininos, no Gráfico 2.8, e de sujeitos masculinos, no Gráfico 2.9.

Gráfico 2.8: Relação entre types/tokens de hipersegmentação para meninas

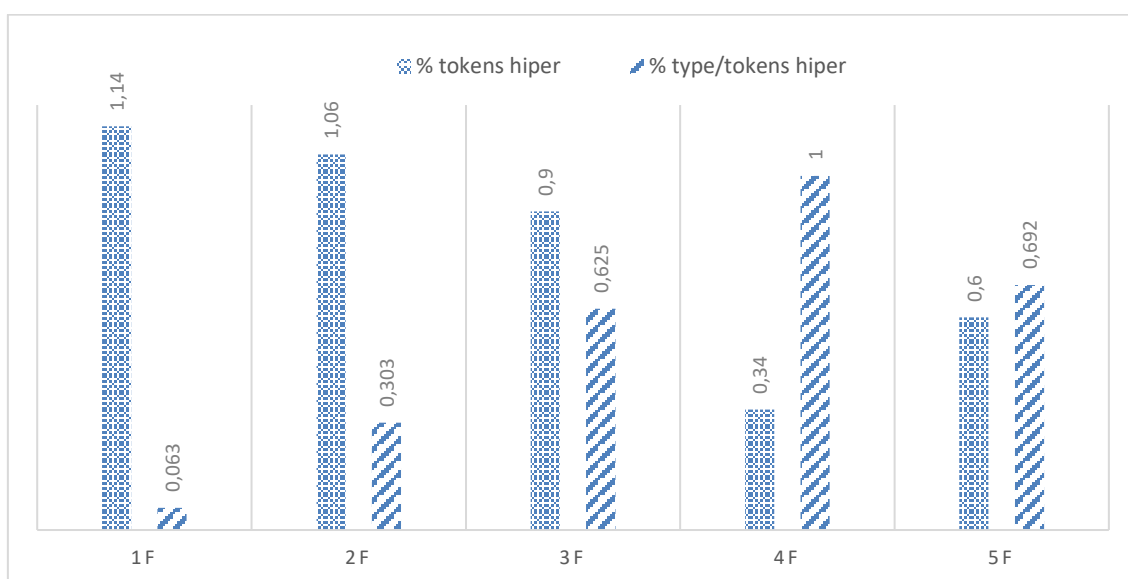
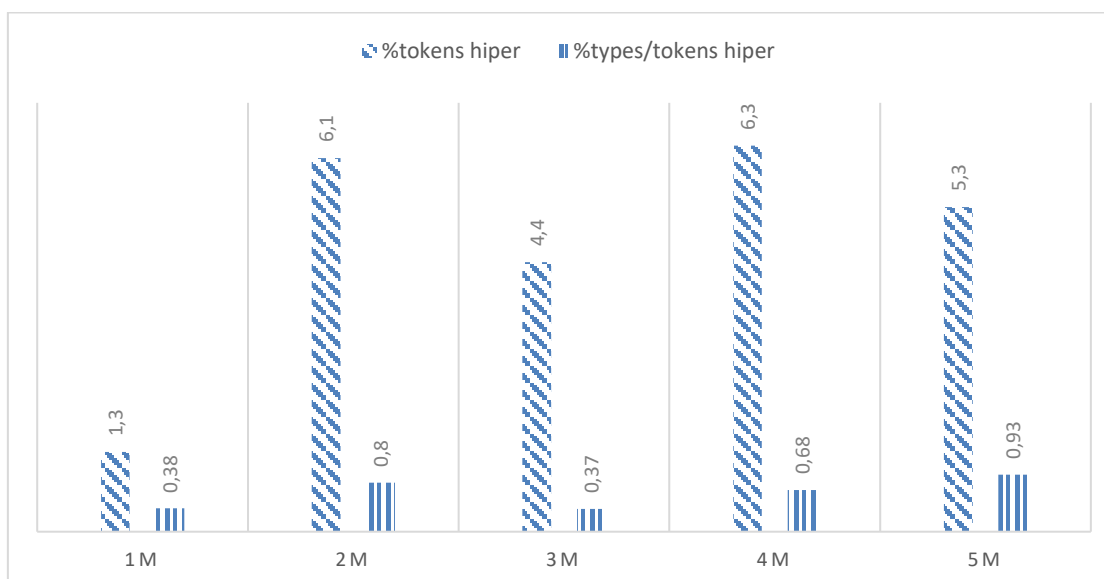


Gráfico 2.9: Relação entre types/tokens de hipersegmentação para meninos



Do Gráfico 2.8, constata-se que as cinco meninas que apresentam os piores índices de hipersegmentação, informação indicada por meio da porcentagem de *tokens* de hipersegmentação para cada sujeito (calculada a partir da relação entre número de hipersegmentação e número de palavras escritas), têm, no entanto, índices bastantes distintos entre si quando considerados os tipos de palavras hipersegmentadas, como é atestado por meio da comparação entre os sujeitos 1F e 4F. Observa-se que o sujeito 1F tem o maior índice de hipersegmentação (1,14%) no conjunto de sujeitos femininos por ter produzido 47 hipersegmentações, mas essas compreendem apenas três palavras distintas (a saber: “porque”, “senão”, “videogame”), enquanto o sujeito 4F tem índice de hipersegmentação (0,34%) próximo da média do grupo investigado, mas todas as dez palavras hipersegmentadas são diferentes entre si (resultando em índice 1).

Entre os meninos selecionados, observa-se ser predominante o fato de grafarem várias vezes a mesma palavra hipersegmentada. Dentre o grupo selecionado, o sujeito 1M tem o menor índice de hipersegmentação e grava apenas cinco palavras hipersegmentadas, apresentando características semelhantes ao sujeito 1F (menina com o mais alto índice de hipersegmentação). Por outro lado, o sujeito 5M apresenta trajetória semelhante ao sujeito 4F, na medida em que produz várias hipersegmentações de palavras distintas (15 *types*/16 *tokens*). Essa comparação permite verificar que, independentemente do sexo/gênero: (i) há sujeitos que hipersegmentam muitas vezes as mesmas palavras, apresentando dúvidas

ortográficas sobre um conjunto pequeno de palavras e (ii) há sujeitos com dificuldades ortográficas sobre palavras que apresentam características linguísticas comuns, não sinalizando dúvidas quanto à grafia de palavras em particular.

No próximo capítulo, esses dados e as características linguísticas comuns às hipersegmentações mais frequentes na escrita do EF II serão tratados. Quando pertinentes, serão feitas observações quanto à distribuição dos tipos de dados em função do ano letivo e do sexo/gênero do sujeito, mas não será o objetivo fazer cruzamento entre essas variáveis e estruturas linguísticas das hipersegmentações. No entanto, não se descarta a possibilidade dessa análise em trabalhos futuros à luz dos resultados apresentados neste capítulo.

3 Hipersegmentação e organização prosódica dos enunciados

Neste capítulo, são apresentados, na seção 3.1, os fundamentos da perspectiva teórica que embasa a concepção de que a prosódica é parte da estrutura linguística que organiza os enunciados falados e escritos. Em seguida, são descritas, na seção 3.2, regularidades quanto à organização métrica das palavras que atuam favorecendo as hipersegmentações e, na seção 3.3, regularidades quanto à organização do enunciado em que a palavra hipersegmentada ocorre, demonstrando-se em que medida as hipersegmentações são efeitos da atuação de aspectos morfossintáticos e semânticos dos enunciados. Por meio dessas análises, visa-se alcançar os objetivos de: (i) descrever quais características prosódicas se mostram de modo privilegiado para esse tipo de segmentação não-convencional de palavra e (ii) caracterizar a escrita do EF II em relação à escrita do EF I quanto à organização prosódica dos enunciados.

3.1 Organização prosódica dos enunciados

Nesta tese, prosódia é vista de uma perspectiva fonológica que tem como premissa haver constituintes prosódicos que estruturam os enunciados de cada língua (SELKIRK, 1984; NESPOR; VOGEL, 1986). Em outras palavras, assume-se que os enunciados são organizados em estruturas prosódicas que abrangem o ritmo e a entoação das línguas. Portanto, a assunção central desta tese é que a estrutura prosódica (fonológica) determina possibilidades de realizações (fonéticas) dos elementos prosódicos dos enunciados falados.⁷³ Desta assunção, nasce a hipótese em que está embasada a análise de dados de escrita, segundo a qual a estrutura prosódica subjaz, também, a possibilidades de registros gráficos dos enunciados. Argumentos que sustentam essa hipótese serão apresentados na

⁷³ Essa concepção fonológica difere de uma caracterização fonética da prosódia. Para Cagliari (2007), de uma perspectiva fonética, a prosódia compreende conjunto de elementos da dinâmica, da melodia da fala e da qualidade de voz. Investigação de evidências fonéticas da estrutura prosódica é feita por Tenani (2002), Fernandes (2007), por exemplo.

seção 3.2. Antes, porém, faz-se necessário apresentar os fundamentos teóricos que sustentam uma visão fonológica da prosódia.

O estudo da prosódia de uma perspectiva fonológica que embasa esta tese é aquele desenvolvido segundo o arcabouço da chamada Fonologia Prosódica. Esse ramo dos estudos fonológicos (feitos no âmbito das chamadas Fonologias Não-lineares) compreende modelos teóricos que têm como objeto comum a investigação de fenômenos fonético-fonológicos que evidenciam a interface entre fonologia e demais componentes da gramática, especialmente a interface sintaxe-fonologia. A Fonologia Prosódica caracteriza-se como uma teoria formalista sobre estruturas prosódicas as quais são definidas a partir da identificação de informações de natureza sintática e/ou morfológica relevantes para caracterizar domínios de aplicação de regras fonológicas.

Assumir que seja a partir de informação morfossintática que constituintes prosódicos são construídos não implica que haja isomorfismo entre os constituintes morfossintáticos e prosódicos, como se verifica da comparação entre “mesa-redonda” e “mesa redonda” nos pares de sentenças em (3.1) e (3.2).⁷⁴

3.1. Vi uma mesa-redonda. $\Rightarrow$ [mesa-redonda]_w = [mesa]_ω [redonda]_ω

3.2. Vi uma mesa redonda. $\Rightarrow$ [mesa]_w [redonda]_w = [mesa]_ω [redonda]_ω

Em (3.1), a palavra composta “mesa-redonda”, um tipo de palavra morfológica (w), um substantivo, corresponde a duas palavras prosódicas (ω). Em (3.2), “mesa” e “redonda” são duas palavras morfológicas (substantivo e adjetivo) e duas palavras prosódicas, que constituem um sintagma nominal. Constata-se, pois, que em (3.2), mas não em (3.1), há isomorfismo entre os constituintes morfossintáticos e prosódicos. A convenção ortográfica do português representa essa diferença ao normatizar o registro gráfico de fronteira de palavra: em (3.1), o emprego do hífen entre “mesa” e “redonda” registra a constituição de palavra morfológica; em (3.2), a presença de espaço entre “mesa” e “redonda” delimita palavras morfológicas que têm entre si relação sintagmática.

Interessam a essa tese pares de estruturas, como em (3), que têm como traço característico a não isomorfia entre os constituintes prosódicos e os morfossintáticos. Essa

⁷⁴ O não isomorfismo entre sintaxe e fonologia também é exemplificado, na literatura, pela comparação entre localização de fronteiras sintáticas (em i) e prosódicas (em ii) para sentença com orações encaixadas: (i) [Este é [o gato que comeu [o rato que comeu o queijo]]] *versus* (ii) [[Este é o gato] [que comeu o rato] [que comeu o queijo]]. Nesses exemplos, interessa a não coincidência entre as fronteiras sintáticas e prosódicas indicadas e não exatamente os rótulos de cada uma dessas fronteiras.

não isomorfia implica não coincidências entre fronteiras de constituintes da gramática da língua, as quais são – em certa medida – registradas nas produções escritas de sujeitos em processo de aquisição da escrita. Como essa relação entre prosódia e escrita pode ser identificada nas segmentações não-convencionais de palavra será detalhado na próxima seção.

Antes de traçar essa relação entre prosódia e escrita, faz-se necessário apresentar outra premissa básica da Fonologia Prosódica. Nesse arcabouço teórico, é central conceber a gramática dos sons como organizada em constituintes hierárquicos, portanto, sujeitos a relações de dependência regidas por princípios universais comuns às gramáticas das línguas do mundo. A definição de quais são essas relações de dependência entre os constituintes prosódicos é tema de discussão entre os pesquisadores e as propostas de formalização dessas relações são encontradas em diferentes modelos de análise prosódica.

Na década de 1980, quando nascem os modelos gerativistas de análise fonológica da prosódia, registram-se dois modelos principais, distintos entre si. O modelo *end-based*, elaborado por Selkirk (1984), tem como traço principal eleger fronteira sintática como informação a partir da qual os constituintes prosódicos são construídos. O modelo *relation-based*, proposto por Nespor e Vogel (1986), assume que relações sintáticas são informações relevantes a partir das quais os constituintes prosódicos se configuram. Essa diferença entre os modelos quanto ao tipo de informação sintática relevante para a construção de estruturas prosódicas resulta na configuração de diferentes propostas de hierarquias prosódicas, além de diferentes maneiras de construção dos constituintes prosódicos.

Quanto à configuração da hierarquia prosódica, no modelo *end-based*, são propostos por Selkirk (1980, 1981) cinco domínios prosódicos, a saber (do menor para o maior): sílaba (σ), pé métrico (Ft), palavra prosódica (Pw), frase fonológica (PhP), frase entoacional (IPh). Posteriormente em Selkirk (1984), a autora afirma serem apenas a sílaba e a frase entoacional os constituintes prosódicos em inglês.⁷⁵ No modelo *relation-based*, os domínios prosódicos são, em ordem crescente: sílaba (σ), pé métrico (Σ), palavra prosódica (ω), grupo clítico (GC), frase fonológica (ϕ), frase entoacional (I) e

⁷⁵ Cf. “The particular claim we are making about the prosodic constituent structure of phonological representation in English, then, is that the phonological phrase, the prosodic word, and the foot are not units in the hierarchy, but that the syllable and the intonational phrase are.” (SELKIRK, 1984, p. 31).

enunciado fonológico (*U*).⁷⁶ Nota-se que, além de diferenças entre os modelos na forma de construir cada constituinte, há o acréscimo do domínio do *CG* e *U* no segundo modelo. A proposta de dois, cinco ou sete domínios na hierarquia prosódica implica discussão sobre como conceber a gramática que define a estrutura prosódica dos enunciados. Essa é uma questão teórica que desafia os investigadores da área.

A falta de consenso sobre, por exemplo, que tipo de informação sintática é relevante para a modelagem da gramática dos sons motiva pesquisas de fenômenos segmentais, rítmicos e entoacionais de diferentes línguas a fim de ser sustentado um ou outro modelo de análise. Encontram-se estudos filiados a distintas concepções de gramática, com base em fatos das línguas que justificam a necessidade de mais ou de menos domínios prosódicos ou, alternativamente, que evidenciam a atuação de certos domínios em uma, mas não em outra língua.⁷⁷

No cenário brasileiro, há várias pesquisas embasadas no modelo *relation-based*,⁷⁸ que apostam ser fundamental a relação de adjacência entre os constituintes sintáticos para se compreender a configuração prosódica da língua portuguesa.⁷⁹ Essa premissa teórica do modelo é adotada por esta tese, uma vez que, como será demonstrado, a relação de adjacência entre constituintes da língua é conhecimento linguístico fundamental para a interpretação dos registros escritos investigados sejam compreendidos. Os princípios que norteiam a hierarquia prosódica de Nespor e Vogel (1986, 2007) serão mobilizados para que os objetivos deste capítulo sejam alcançados: (i) descrever características prosódicas

⁷⁶ Observa-se que há variação na tradução dos termos do inglês para o português. Para “prosodic word”, a preferência é por “palavra prosódica”, podendo ser encontrada a denominação “palavra fonológica”. Para “phonological phrase”, registram-se “frase fonológica” e “sintagma fonológico”. Para “intonational phrase”, encontram-se “frase entoacional” e “sintagma entoacional”.

⁷⁷ Uma retrospectiva detalhada sobre as diferentes abordagens na literatura da Fonologia Prosódica é encontrada em Gayer (2015).

⁷⁸ Sobre o PB, Tenani (2002), Fernandes (2007), Serra (2009), por exemplo.

⁷⁹ Há também pesquisa sobre a prosódia do português desenvolvida no arcabouço da Teoria da Otimalidade (cf. GUZZO, 2015). A diferença na abordagem entre os modelos pioneiros para o modelo em desenvolvimento por Selkirk (2000, 2004) reside no modo de conceber a formalização da gramática. Enquanto os modelos *end-based* e *relation-based* são desenvolvidos a partir de uma proposta teórica da gramática formal gerativa baseada em princípios universais e parâmetros definidos para línguas particulares (modelo vigente na década de 1980), a proposta de análise prosódica do modelo segundo a abordagem otimalista é desenvolvida com base em uma proposta teórica também formal, mas conexcionista, ou seja, não mais baseada em princípios e parâmetros linguísticos, mas em conjunto de restrições universais, passíveis de violação, que são hierarquizadas entre si de modo a configurar a gramática de cada língua. Embora seja diferente a formalização teórica entre o modelo mais recente e os mais antigos, observa-se que se mantêm algumas questões centrais para o campo da chamada Fonologia Prosódica, a saber: (i) quais informações sintáticas são relevantes para a configuração de estruturas prosódicas das línguas do mundo?, (ii) como se configuram os constituintes prosódicos em cada língua em particular?, (iii) quais são os tipos de evidências dos constituintes em cada língua? e (iv) há características comuns entre línguas de modo a ser possível agrupá-las quanto à organização prosódica?

das hipersegmentações de palavra; e, em consequência, (ii) argumentar que estruturas prosódicas também são constitutivas dos enunciados escritos.

De modo sucinto, o modelo da Fonologia Prosódica adotado pressupõe que: (i) há uma organização prosódica dos enunciados que rege processos segmentais, rítmicos e entoacionais da língua e (ii) essa organização prosódica é constituída a partir de informações dos demais componentes da gramática, sendo mobilizadas informações morfossintáticas e semânticas, atendidas condições pragmáticas de enunciação. Essa organização prosódica é regida por quatro princípios, apresentados a seguir, que garantem boa formação dos constituintes prosódicos.

3.3. Princípios da hierarquia prosódica⁸⁰

Princípio 1. Uma dada unidade não-terminal da hierarquia prosódica, X^p , é composta de uma ou mais unidades da categoria imediatamente mais baixa, X^{p-1} .

Princípio 2. Uma unidade de um dado nível da hierarquia está exhaustivamente contida na unidade hierarquicamente superior de que faz parte.

Princípio 3. As estruturas hierárquicas da fonologia prosódica têm ramificação n-ária;

Princípio 4. A relação de proeminência relativa definida para nós irmãos é tal que a um nó é atribuído o valor forte (s) e a todos os outros nós, o valor fraco (w).

Esses princípios garantem que todos os constituintes têm a mesma configuração geométrica, o quais são construídos sempre da mesma forma segundo a regra da Gramática Universal, formulada em (3.4).

3.4. Construção de constituinte prosódico⁸¹

Junte em uma ramificação n-ária X^p todos os X^{p-1} incluídos na cadeia delimitada pela definição do domínio de X^p .

⁸⁰ “Principle 1. A given nonterminal unit of the prosodic hierarchy, X^p , is composed of one or more units of the immediately lower category, X^{p-1} .

Principle 2. A unit of a given level of the hierarchy is exhaustively contained in the superordinate unit of which it is a part.

Principle 3. The hierarchical structures of prosodic phonology are n-ary branching.

Principle 4. The relative prominence relation defined for sister nodes is such that one node is assigned the value strong (s) and all the other nodes are assigned the value weak (w).” (NESPOR; VOGEL, 1986, p. 7).

⁸¹ “Prosodic Constituent Construction: Join into an n-ary branching X^p all X^{p-1} included in a string delimited by the definition of the domain of X^p .” (NESPOR; VOGEL, 1986, p. 7).

último, Fernandes-Svartman et al. (2012) demonstram haver evidências rítmicas e segmentais da ω as quais são diferentes das discutidas por Vigário (2003) para o PE.

A esta tese não interessa detalhar evidências sobre cada um dos domínios dessa estrutura prosódica do PB, mas mobilizar, dos estudos ancorados em experimentos fonéticos, resultados sobre a estrutura prosódica relevantes para análise dos enunciados escritos, como será feito nas próximas seções deste capítulo. A seguir, de modo sucinto, são apresentados os sete constituintes prosódicos – com base em Nespor e Vogel (1986, 2007) – para, em seguida, serem apresentados detalhamentos sobre a constituição dos constituintes *CG*, Σ e ω em PB, por serem esses constituintes mais mobilizados nas análises de hipersegmentações, feitas na seção 3.2 e 3.3.

- sílaba (σ): é o menor constituinte da hierarquia prosódica, no qual segmentos se organizam, segundo princípios universais e específicos de cada língua. É o domínio de aplicação de várias regras⁸³ e a unidade básica a partir da qual uma cadeia fônica é estruturada, como exemplificado na Figura 3.1;
- pé métrico (Σ): é formado de uma cadeia de sílabas que constituem uma palavra prosódica, sendo atribuído a uma sílaba o valor forte e às demais valor fraco. Na Figura 3.1, os Σ são constituídos de duas sílabas (exemplos de pé silábico binário) em “maracujás”, mas três sílabas em “semana” (exemplo de pé silábico ternário);⁸⁴
- palavra prosódica (ω): tem como domínio elementos terminais de uma árvore sintática, o que inclui raiz, todos os afixos e os dois membros de palavras compostas (NESPOR; VOGEL, 2007, p. 110).⁸⁵ É, pois, constituída de informações morfológicas sobre formação de palavra e tem o acento (lexical) como característica fonológica central. A forma verbal “comprei” – na figura 3.1 – constitui uma ω , tal como “comprávamos”, outra forma flexionada de “comprar”, ou como “recomprei”, forma verbal flexionada de “recomprar”,

⁸³ Em PB, observa-se a inserção de [i] em “apto”, gerando [‘apito], por atuação de princípios universais e particulares de formação da sílaba (a oclusiva [p] é pouco sonora para ocupar a coda silábica e as oclusivas [pt] também formam uma sequência que não atende às condições de boa formação do ataque silábico) – cf. discussão sobre epêntese vocálica em Collischonn (2000, 2004).

⁸⁴ Na literatura sobre Fonologia Métrica (cf. HAYES, 1985), é discutida a constituição dos pés métricos, havendo evidências de que as línguas diferem entre si quanto ao número de sílabas permitidas no domínio do pé (binário ou n-ário) e quanto à sensibilidade à estrutura interna da sílaba (podendo o pé métrico ser ou não sensível à quantidade silábica – isto é, à ramificação da rima em núcleo e coda, por exemplo).

⁸⁵ Mais à frente, serão comentados os desafios de se estabelecer distinção entre os conceitos de palavra, afixo e clítico.

derivada de “comprar”. Em todos esses exemplos, há apenas um acento (portanto, há uma ω) e a raiz e os afixos constituem uma palavra morfológica;

- grupo clítico (*CG*): é a unidade que compreende um elemento clítico (isto é, elemento que não corresponde a afixos, nem a palavras, desprovido de acento) e seu hospedeiro (isto é, uma ω , portadora de acento). Na Figura 3.1, o grupo clítico “na semana” constitui-se do clítico “na” seguido do hospedeiro “semana”;⁸⁶
- frase ou sintagma fonológico (ϕ): é o domínio constituído pela junção de grupos clíticos e caracteriza-se a partir da configuração de sintagmas definidos pela sintaxe. Por exemplo, o substantivo “maracujás” e o adjetivo “doces” formam ϕ_s independentes. Porém, há reestruturação, resultando em um único ϕ , “maracujás doces”⁸⁷ – como visto na Figura 3.1 –, pois “doces” é complemento (que fica ao lado direito) de “maracujás” (que não é ϕ ramificado)⁸⁸ – cf. Nespor e Vogel, (2007, p. 173). Quando ramificado, o nó mais à direita será rotulado como o mais forte, em línguas cuja ramificação sintática é à direita, como em português; mas será o nó mais à esquerda rotulado como o mais forte quando a ramificação sintática for à esquerda, como o japonês. Em [maracujás doces] ϕ da Figura 3.1, é forte “doces” e fraco, “maracujás”;
- frase ou sintagma entoacional (*I*): compreende estruturas sintáticas que correspondem a (i) sentença raiz, (ii) elementos não anexáveis à estrutura da sentença raiz e (iii) elementos remanescentes de uma sentença raiz interrompida por elemento(s) anexado(s) a ela (cf. algoritmo de formação de *I* em Nespor e Vogel, 2007, p. 189). Formada a partir dessas bases sintáticas, mas não necessariamente isomórfica a elas, a *I* é a unidade prosódica que se define pelos limites do contorno entoacional (independentemente de o contorno ser ascendente

⁸⁶ A identificação de (tipos de) elementos clíticos, bem como a caracterização de seu funcionamento prosódico nas línguas, é tema de acaloradas discussões, pois entram em cena tanto aspectos morfológicos quanto sintáticos e semânticos em sua caracterização (cf. SPENCER; LUÍS, 2012; RIEMSDIJK, 1999).

⁸⁷ Evidência de que “maracujás doces” constitui um único ϕ é a aplicação da regra de resolução de choque acentual: o acento de “maracujás” é adjacente ao acento de “doces” e, por estarem no mesmo ϕ , o acento é movido de “jás” à sílaba precedente de modo a ser obtida uma sequência de alternância rítmica, a saber: [maracujás doces] ϕ . Quando há adjacência entre acentos de palavras que pertencem a ϕ s distintos, o processo não se aplica (cf. “Comprei maracujás, peras e maçãs ontem”).

⁸⁸ Alternativamente, se a estrutura fosse “belos maracujás doces”, o mapeamento prosódico em ϕ passa a ser: [belos maracujás] ϕ [doces] ϕ , pois “belos” se encontra do lado não-recursivo de “maracujás”. Nessa estrutura, não é possível haver reestruturação *[belos maracujás doces] ϕ , porque “belos maracujás” formam um ϕ ramificado, isto é, constituído de duas palavras prosódicas.

ou descendente).⁸⁹ Na Figura 3.1, indica-se a constituição de dois *I*s, sendo cada um constituído de contornos entoacionais: [[na semana passada]*I* [comprei maracujás doces]*I*]*U*. Como visto em 3.1, todos os ϕ s são anexados sob o nó de *I*, sendo o mais forte aquele ϕ que for semanticamente mais proeminente. Em enunciados neutros, como na Figura 3.1, é proeminente o ϕ mais à direita dentro de *I* (e cuja sílaba proeminente será portadora do acento frasal, ou *pitch accent*): “doces” em [comprei maracujás doces]*I* e “passada” em [na semana passada]*I*;

- enunciado fonológico (*U*): é a maior unidade prosódica e se constitui de um ou mais *I*s, como se visualiza na Figura 3.1. Seus limites correspondem aos limites de sentença (maior constituinte sintático), onde pausas podem ocorrer. Quando constituído de dois ou mais *I*s, é atribuído o valor forte ao *I* mais à direita da ramificação *U*.⁹⁰ No exemplo da Figura 3.1, [comprei maracujás doces]*I* é o *I* forte dentro do *U*.

Na hierarquia prosódica, exemplificada na Figura 3.1, problemas são identificados especialmente quanto ao estatuto de pé e palavra atribuído ao clítico (como exemplificado por “na”), o que implica discutir a pertinência do grupo clítico como constituinte prosódico (cf. VIGÁRIO, 2007; BISOL, 2000). Também é tema de discussão se, em PB, é pertinente aceitar a ramificação *n*-ária do Σ , como “semana” e “passada” na Figura 3.1. Esses problemas serão abordados a seguir.

No exemplo da Figura 3.1, os *status* de Σ e ω atribuídos ao clítico “na” decorrem da atuação dos princípios 1 e 2 (denominados por *Strict Layer Hypothesis* por Selkirk, 1981). Essa configuração não representa o funcionamento linguístico de clíticos monossilábicos como “na”, pois clíticos prosódicos não constituem pés métricos, nem palavras prosódicas em português. Diferentemente, o clítico fonológico é sílaba átona prosodizada a um hospedeiro que se configura como ω , constituída de, ao menos, um Σ . Diante desse e de outros fatos de várias línguas (cf. SPENCER; LUÍS, 2012), os princípios 1 e 2 que garantem condições de boa formação dos constituintes prosódicos

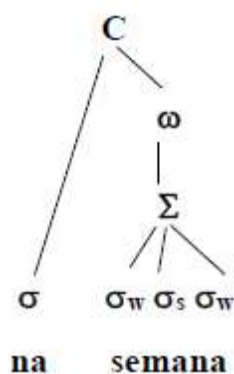
⁸⁹ Uma frase entoacional pode corresponder a diferentes estruturas sintáticas, pois essa se define em função do contorno entoacional, como em [Cuidado com o carro!]*I* ou em [Cuidado!]*I*, que têm um único contorno, embora sejam sintaticamente distintas entre si.

⁹⁰ Dois *Us* podem ser reestruturados formando um único *U*, desde que atendidas duas condições pragmáticas, a saber: (i) as duas sentenças devem ser enunciadas pelo mesmo falante e (ii) as duas sentenças devem ser endereçadas ao(s) mesmo(s) interlocutor(es); e duas condições fonológicas, a saber: (i) as duas sentenças devem ser relativamente pequenas e (ii) não deve haver pausa entre as duas sentenças. – cf. Nespor; Vogel (2007, p. 240). Exemplos de *Us* reestruturados em PB são discutidos em Tenani (2002).

são postos em questão. Dentre as possíveis soluções propostas pelos estudiosos do tema, Nespor e Vogel (2007, p. XV-XIX) admitem que, no máximo, dois níveis possam ser saltados de modo a acomodar o fato de que o clítico prosódico tem estatuto de sílaba que é prosodizada a uma ω .

Ao admitirem uma versão enfraquecida da *Strict Layer Hypothesis*, as autoras também resolvem o problema de serem analisadas como grupo clítico estruturas que não têm clíticos, como “comprei maracujás doces”. Não ficam resolvidos, no entanto, os problemas relativos à pertinência de o grupo clítico fazer parte da hierarquia da Fonologia Prosódica. Apesar de fatos e argumentos contrários ao grupo clítico,⁹¹ Nespor e Vogel (2007) reafirmam que o grupo clítico é constituinte na hierarquia prosódica, porque é domínio para aplicação de regras segmentais, como em italiano (cf. NESPOR; VOGEL, 1986 e 2007). Nesta tese, não interessa trazer à cena detalhes dessa problematização, feita com base em fenômenos segmentais, principalmente, sobre a prosodização dos clíticos, pois, para fins de análise dos dados de escrita, é suficiente assumir a proposta de os clíticos serem sílabas prosodizadas no domínio grupo clítico, como se visualiza na figura (3.2).⁹²

Figura 3.2: Representação da prosodização do clítico



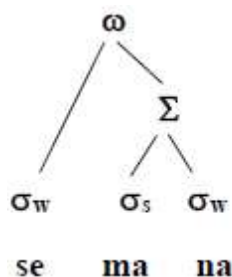
Outra questão a que se faz necessário lançar luz diz respeito ao princípio quatro, especificamente à assunção de o constituinte prosódico ter ramificação n-ária. Essa

⁹¹ Bisol (2005) admite o CG como domínio ativo em PB a partir da consideração do processo de elisão. Mas Simioni (2008) propõe outra interpretação aos mesmos dados de elisão e defende que clíticos sejam prosodizados no nível de ϕ .

⁹² Cabe ponderar, entretanto, que não se considera essa proposta como definitiva, uma vez que ainda são poucos os estudos sobre o tema e faz-se necessário, portanto, ampliar o conjunto de evidências que permitam aprofundar aspectos teóricos quanto à formalização da prosodização de clíticos e, consequentemente, quanto à pertinência (ou não) do domínio grupo clítico.

assunção é particularmente relevante para análise de enunciados do português em pés métricos. Estudos sobre pés métricos em PB admitem serem característicos dessa variedade pés binários, predominantemente, pés troqueus dissílabos, como proposto por Bisol (1992). Na abordagem n-ária de Nespor e Vogel (1986), trissílabos paroxítonos, como “semana”, da Figura 3.2, têm as três sílabas formando um pé ternário, não capturando a característica predominante de pés binários do PB. No entanto, se admitida a versão fraca da *Strict Layer Hypothesis* (NESPOR; VOGEL, 2007) também para o domínio do Σ , é possível capturar a constituição de pés binários para trissílabos paroxítonos, tal como representado na Figura 3.3.

Figura 3.3: Representação da configuração binária do pé métrico



Nessa abordagem, a sílaba pretônica do trissílabo não é analisada em Σ , porque esse será formado por duas sílabas (as sílabas tônica e postônica de “semana”), mas é prosodizada no constituinte imediatamente superior a Σ , ou seja, na ω . Por meio dessa representação, a configuração binária fica assegurada, além de ser capturada uma relação assimétrica entre sílaba pretônica – que fica fora do domínio do Σ – e sílaba postônica – que faz parte do Σ . Evidência a favor da representação apresentada na Figura 3.2 será dada, na próxima seção, por meio da análise de hipersegmentações de trissílabos paroxítonos atestadas nos textos do EF II.

Essa representação binária implica mobilizar basicamente dois tipos de pés métricos, a saber: (i) troqueu, que se configura por uma sílaba forte seguida de uma fraca (como “compra”); e (ii) iambo, que se configura por uma sílaba fraca seguida de outra forte (como “demais”). Ainda com base na tradição métrica greco-latina, será denominado de pé dátilo, quando houver a configuração de sílaba forte seguida de duas sílabas fracas (como “morávamos”) entre trissílabos e polissílabos proparoxítonos. A última sílaba do pé dátilo será analisada como extra-métrica (BISOL, 1992; 2013).

Além do Σ , a ω é outro domínio relevante para a análise, a ser feita nas próximas seções. Embora a definição do domínio de ω – apresentada anteriormente – pareça bastante clara, não é consenso na literatura fonológica critérios para identificar o que seja palavra nas diferentes línguas (cf. DIXON; AIKHENVALD, 2002), sendo desafiador estabelecer distinção entre os conceitos de palavra, afixo (elemento interno à palavra) e clítico (elemento externo à palavra).

Sobre o PE, por exemplo, Vigário (2003) discute as características da ω , considerando o funcionamento dos afixos e clíticos, por um lado, e das chamadas palavras compostas, por outro lado, e propõe um domínio – grupo de ω – visando acomodar, na hierarquia prosódica, tanto o fato de afixos terem status de palavra por possuírem acento (como “-mente” em “vagarosamente”),⁹³ embora sejam formas que sempre ocorram afixadas a uma base, quanto o fato de clíticos – que não têm status de ω por não terem acento – terem comportamento morfossintático mais independente do que os afixos em relação ao seu hospedeiro prosódico.

Lee (1995), Schwindt (2000) e Lopes (2016), por exemplo, também discutem alternativas de análise, a partir do PB, para os comportamentos de afixos que exibem autonomia acentual em relação às suas bases morfológicas. Evidências obtidas por meio de análise de processos segmentais de formação de palavras reunidas em Lopes (2016), por exemplo, levam à constatação, sob nossa perspectiva, de que não há sobreposição dos limites de ω aos de palavra morfológica, pois, argumenta a autora, há tanto ω resultante da junção de duas ou mais palavras morfológicas (como “ferrovia”, “espaçonave”, “floricultura”),⁹⁴ quanto palavra morfológica constituída de duas ou mais palavras prosódicas (como “simplesmente”, “pós-graduação” e “guarda-roupa”).⁹⁵ Interessa, dessas considerações, o fato de os limites de ω não coincidir com os de palavra morfológica, pois será argumentado, como se verá nas próximas seções deste capítulo, que as segmentações não-convencionais de palavra identificadas são efeitos dessa relação

⁹³ Abreu-Zorzi (2016) analisa advérbios em “-mente” em duas sincronias do português, o português do século XIII e o português brasileiro atual, e conclui que formas adverbiais em “-mente” têm acento e são, portanto, compostas do ponto de vista prosódico em ambas as sincronias investigadas.

⁹⁴ Cada um dos exemplos constitui uma ω , porque tem um acento lexical, mas duas palavras morfológicas, a saber: “ferro” e “via” em “ferrovia”; “espaço” e “nave” em “espaçonave”; “flor” e “cultura” em “floricultura”.

⁹⁵ Em cada exemplo, há duas palavras prosódicas porque são identificados dois acentos, a saber: no sufixo “-mente” e na base “simples”, em “simplesmente”; no prefixo “pós-” e na base “graduação”, em “pós-graduação”; e nas bases “guarda” e “roupa”, em “guarda-roupa”.

não-unívoca entre os constituintes de palavra prosódica e morfológica, definidos a partir da interface morfologia – fonologia.

Em suma, é adotada a hierarquia prosódica de Nespor e Vogel (1986, 2007), na sua versão em que é assumida uma concepção fraca da *Strict Layer Hypothesis*, embora seja admitida a pertinência de se problematizar que: (i) todos os sete domínios são universais linguísticos e (ii) a construção da hierarquia prosódica possa ser feita com base em princípios ou em restrições violáveis. Para os objetivos desta tese, no entanto, é relevante a assunção de que existe uma estrutura prosódica constituída de domínios prosódicos hierarquicamente organizados e que essa estrutura é construída a partir, principalmente, de informação sobre características morfológicas e sintáticas da língua.

Também é importante assumir que a hierarquia prosódica proporciona meios de explicitar como processos segmentais e configurações de acentos e de tons se realizam em uma dada língua. Ora, a configuração de acentos dos enunciados é guiada pela organização rítmica da língua e a configuração dos tons, pela organização entoacional dessa mesma língua. Portanto, a Fonologia Prosódica também tem no seu escopo de análise a relação entre ritmo e entoação de cada língua em particular.

Interessa a esta tese mobilizar essa formalização teórica dessa relação entre ritmo e entoação por permitir não apenas explicitar que essa relação pode estar evidenciada nos enunciados escritos (e de modo privilegiado nos dados de segmentação de palavra, como será demonstrado nas seções 3.2 e 3.3), mas – sobretudo – analisar modos pelos quais essa organização rítmico-entoacional (dimensão fônica da linguagem, nos termos de Chacon, 1998) constitui os textos escritos (dimensão textual-discursiva para o mesmo autor), como será discutido no capítulo 4.

3.2 Hipersegmentação e organização métrica da palavra

Nesta seção, serão descritas as configurações rítmicas das palavras hipersegmentadas encontradas na amostra longitudinal de textos de escrita do EF II. A hipótese que subjaz a essa análise é aquela predominante assumida nos estudos sobre esse tipo de grafia na escrita do EF I, segundo a qual a organização métrica da palavra motiva registros não-convencionais de fronteiras de palavras (CUNHA, 2010, 2004; CAPRISTANO, 2007a).

Retomo do capítulo 2 (seção 2.2), para fins de análise, dois critérios de classificação das segmentações de palavras: (i) critério gráfico, que diz respeito ao tipo de informação verbo-visual (branco ou hífen) que delimita a fronteira da palavra escrita, e (ii) critério morfossemântico, que diz respeito à identificação de palavras e sequências de palavras homônimas.

Com base no critério morfossemântico, foi constatado que 487 das hipersegmentações compreendem “porque”, “demais” e “senão” hipersegmentados (respectivamente, 410, 41 e 36 ocorrências). Esse conjunto totaliza 48,6% dos dados, constituindo alta frequência das mesmas palavras (relação *types/tokens*). Essa frequência enviesa a análise quantitativa de dados, o que levou à decisão de excluí-las da análise prosódica a ser feita nesta seção. Esse conjunto de ocorrências será analisado na próxima seção, quando da análise de características morfossintáticas envolvidas nas hipersegmentações da escrita do EF II. Em função dessa exclusão, o conjunto de dados restantes totaliza 515 hipersegmentações. Desse subconjunto, farão parte da análise prosódica 428 hipersegmentações (ou 42,71% do total de ocorrências), e as 87 hipersegmentações restantes (8,68%) serão analisadas qualitativamente no capítulo 4, pois essas hipersegmentações apresentam características linguístico-discursivas – a serem descritas – que permitem que seja demonstrado de que maneira essas características verbais se integram às verbo-visuais no processo de produção escrita dos escreventes analisados.

Para análise prosódica das palavras hipersegmentadas selecionadas, fez-se necessário mobilizar fatores de natureza fonológica, portanto verbal, ao lado de fatores de natureza verbo-visual. Quanto aos critérios de natureza verbo-visual, além do tipo de informação fronteira gráfica que delimita a palavra ortográfica (branco ou hífen), será considerada a localização da fronteira não-convencional em relação às fronteiras (direita ou esquerda) convencionais de palavra. Os critérios de natureza fonológica dizem respeito a: (i) localização do acento lexical de palavra e (ii) número de sílabas da palavra escrita. Cada um desses fatores abrange um conjunto de variáveis, conforme apresentado no Quadro 3.1.

Quadro 3.1: Conjunto de fatores e variáveis considerados na análise prosódica

Fatores	Variáveis Consideradas
Tipo de fronteira gráfica	Branco ou hífen
Localização da fronteira não-convencional dentro da palavra	Após a primeira sílaba da palavra, no meio da palavra ou antes da última sílaba da palavra
Localização do acento lexical	Oxítono, paroxítono ou proparoxítono
Número de sílabas de palavra	Dissílabo, trissílabo ou polissílabo

A consideração do tipo de fronteira gráfica de palavra faz-se relevante, porque o emprego do branco ou do hífen implica mobilizar diferentes informações verbo-visuais que dizem respeito à identificação de limites de palavra e de relação entre palavras, como discutido na seção 2.2. Vale lembrar daquela seção que o hífen tem emprego mais restrito em relação ao branco em função de sua característica “ambígua” que delimita certas fronteiras morfossintáticas entre palavras escritas, como é o caso de compostos e estruturas formadas de verbo e clítico.

A localização da fronteira não-convencional em relação às fronteiras (direita/esquerda) convencionais de palavra permite apreender regularidades da hipersegmentação quanto à projeção de fronteiras prosódicas. Esse critério é pertinente para trissílabos ou polissílabos, uma vez que dissílabos têm apenas a opção de haver segmentação entre sílabas, como em “de mais”, “com pras”. Nota-se que possibilidades de localização da fronteira não-convencional estão correlacionadas ao número de sílabas da palavra,⁹⁶ critério sobre o qual serão feitas considerações mais à frente. Antes, cabe observar que, nesta seção, serão analisadas hipersegmentações caracterizadas por apenas uma fronteira não-convencional dentro da palavra. As ocorrências de duas fronteiras não-convencionais dentro de palavra totalizaram cinco hipersegmentações, a saber: “adicio o na”, “em com trei”, “na que le”, “com ver sando”, “com ver samos”. Esse tipo de hipersegmentações será analisado no capítulo 4, uma vez que permitem tratar de modo

⁹⁶ Para trissílabos, são 12 possibilidades de hipersegmentação (2 tipos de registro X 3 tipos de acento lexical X 2 possibilidades de localização da fronteira não-convencional). Entre os polissílabos, foram encontradas palavras com quatro, cinco e seis sílabas. Para palavras com quatro sílabas, são 18 possibilidades de hipersegmentação (uma vez que são três as possibilidades de localização da fronteira não-convencional de palavra). Para palavras com cinco sílabas, são 24 possibilidades de hipersegmentação (pois são quatro possibilidades de localização da fronteira não-convencional de palavra). Para palavras com seis sílabas, são 30 possibilidades de hipersegmentação (por serem cinco as possibilidades de localização da fronteira não-convencional de palavra).

privilegiado de características linguístico-discursivas imbricadas a aspectos verbo-visuais envolvidos no processo de produção de enunciados escritos.

Quanto ao número de sílaba, as palavras são classificadas em dissílabos, trissílabos ou polissílabos segundo a tradição gramatical. Nesta tese, é controlado o número de sílabas entre os polissílabos, tendo sido identificadas hipersegmentações de palavras com quatro, cinco e seis sílabas na amostra de escrita do EF II. O número de sílabas de palavra é adotado como critério para identificar regularidades quanto à frequência da segmentação de palavra em função do tamanho da palavra dado a partir da unidade fonológica básica, a sílaba. A consideração desse tipo de informação também é importante, porque, na literatura sobre aquisição de escrita do EF I, especialmente as que seguem viés cognitivista, como as conduzidas por Emília Ferreiro e colaboradores (FERREIRO et al, 1996; FERREIRO; PONTECORVO, 1996; FERREIRO; TEBEROSKY, 1999), encontram-se resultados que revelam que crianças em fase inicial de aquisição da escrita têm mais dificuldades em grafar palavras com mais de três sílabas ou com apenas uma sílaba.⁹⁷ A investigação dessa característica da palavra permitirá, pois, caracterizar a escrita do EF II frente à escrita do EF I.

Por fim, a localização do acento lexical de palavra envolve a classificação das palavras em oxítonos, paroxítonos e proparoxítonos. Por meio dessa classificação, pretende-se investigar em que medida a organização métrica das palavras pode ser favorecedora de hipersegmentações, uma vez que a localização do acento de palavra leva à configuração de diferentes pés métricos. Oxítonos, que têm acento lexical na última sílaba da palavra, levam à identificação do Σ iambo ($\sigma\sigma$), como “voltei”. Paroxítonos, com acento na penúltima sílaba da palavra, configuram a constituição de Σ troqueu ($\sigma\sigma$), como “quero”. Proparoxítonos, com acento na antepenúltima sílaba da palavra, configuram o pé dátilo ($\sigma\sigma\sigma$), como “vítimas”.

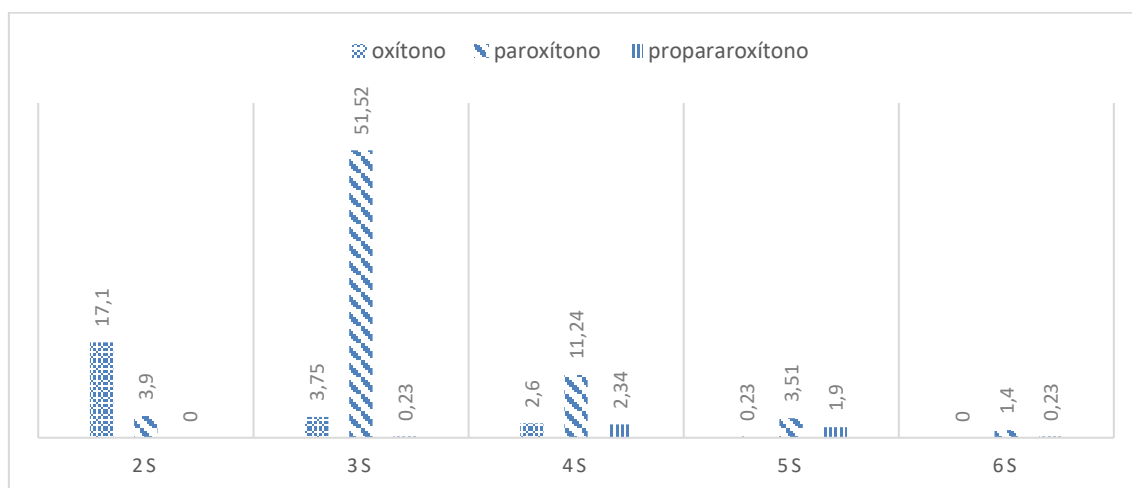
A interação entre os fatores “tipo de fronteira gráfica”, “localização da fronteira não-convencional de palavra” e “localização do acento lexical de palavra” leva a agrupar os resultados a partir do quarto fator considerado, número de sílabas da palavra convencional. Em função desse fator, as hipersegmentações apresentam a seguinte distribuição percentual: 21% dissílabos; 55,5% trissílabos; 23,4% polissílabos. É com

⁹⁷ Registra-se que, nesses trabalhos, as noções de sílaba e palavra são de natureza ortográfica, enquanto que, nesta tese, essas noções são definidas com base em teorias linguísticas.

base nesse resultado geral que será organizada a apresentação dos resultados dos demais fatores analisados.

A análise parte da relação entre número de sílabas de palavras e localização do acento lexical, a fim de serem descritas características gerais das palavras hipersegmentadas. No total, foram hipersegmentados 23,68% oxítonos, 71,57% paroxítonos e 4,5% proparoxítonos. No Gráfico 3.1, visualiza-se a distribuição dessas porcentagens de oxítonos, paroxítonos e proparoxítonos em função do número de sílabas das palavras hipersegmentadas.

Gráfico 3.1: Frequência de oxítonos, paroxítonos e proparoxítonos hipersegmentados versus número de sílabas das palavras



Onde: “2S” são dissílabos; “3S”, trissílabos; “4S”, polissílabos com quatro sílabas; “5S”, polissílabos com cinco sílabas; “6S” polissílabos com seis sílabas.

Dentre as potenciais relações que os resultados do Gráfico 3.1 permitem fazer, três cabem ser destacadas, a saber: (i) dissílabos oxítonos são mais hipersegmentados do que dissílabos paroxítonos, (ii) trissílabos paroxítonos são mais hipersegmentados do que trissílabos oxítonos e (iii) polissílabos paroxítonos são os mais hipersegmentados e polissílabos com quatro ou cinco sílabas passam a ter mais proparoxítonas hipersegmentadas do que trissílabos ou polissílabos com seis sílabas.

Essas relações entre número de sílabas da palavra e localização do acento lexical entre hipersegmentações são mais bem interpretadas se considerados dados de frequência de palavras na língua. Resultados de frequência relativa de palavras, descritos por Araújo et al. (2007, p. 55), por exemplo, com base em 150.875 palavras registradas no Dicionário Houaiss de Língua Portuguesa, demonstram que a frequência relativa de paroxítonas é de

pouco mais de 50% entre dissílabos, chega a quase 60% entre os trissílabos e se mantém em torno de 70% para polissílabos com quatro ou mais sílabas. Entre as oxítonas, a frequência é de pouco menos de 50% entre os dissílabos e decresce à medida que aumenta o número de sílabas da palavra, de modo que chega a 10% entre polissílabos com seis sílabas. Entre as proparoxítonas, a frequência exhibe padrão inverso das oxítonas, ou seja, quanto maior o número de sílabas da palavra, maior a frequência de palavra proparoxítona.

À luz desses dados de frequência de número de sílabas de palavra em relação à localização de acento lexical descritos para o PB, interpreta-se que os resultados em (ii) e (iii) obtidos para as hipersegmentações na amostra de escrita do EF II são, em parte, decorrentes da configuração de palavras na língua. Dito de outro modo, haver mais trissílabos e polissílabos paroxítonos hipersegmentados seria, por hipótese, apoiado na maior frequência de palavras com essa configuração na língua e, nessa interpretação, não necessariamente trissílabos e polissílabos paroxítonos oferecem maior dificuldade ortográfica para alunos do EF II. A mesma hipótese poderia ser proposta para interpretar a maior frequência de polissílabos proparoxítonos hipersegmentados, porém não poderia ser mobilizada para os resultados descritos em (i), uma vez que, na língua, há mais dissílabos paroxítonos do que dissílabos oxítonos, portanto, o resultado atestado para dissílabos hipersegmentados é o inverso do registrado na língua.

A interpretação desse resultado quanto à hipersegmentação de dissílabos será feita após análise dos dois outros fatores considerados nesta tese: localização da fronteira não-convencional de palavra e tipo de fronteira gráfica. A interação entre esses fatores de natureza verbo-visual e os dois de natureza verbal é a chave para identificar características prosódicas que propiciam certas possibilidades de hipersegmentação de dissílabos, trissílabos e polissílabos. Na primeira coluna da Tabela 3.1, são indicadas as sílabas átonas e tônicas das palavras convencionais seguidas pelas segmentações identificadas por meio de parênteses. Nessa notação, identificam-se tanto a localização do acento lexical, quanto a localização da segmentação não-convencional de palavra. Na segunda coluna, cada estrutura é exemplificada com hipersegmentações da escrita do EF II. Nas demais colunas, são apresentados números e porcentagens de cada tipo de estrutura analisada. Visualiza-se, ainda, informação quanto ao tipo de fronteira gráfica: por meio do sombreado das linhas, são agrupadas hipersegmentações caracterizadas pela ocorrência não-convencional do hífen e, nas linhas sem sombreado, são listadas hipersegmentações caracterizadas pela presença não-convencional do branco.

Tabela 3.1: Estruturas métricas e hipersegmentação

Nºσ	Estruturas	Segmentações	Exemplos	Nº	%	Nº	%	Nº	%	
2σ	(σσ)	(σ) (σ)	“na viu”	73	17,02	90	21,02	90	21,02	
	(σσ)	(σ) (σ)	“tem de”	17	3,97					
3σ	(σσσ)	(σ) (σσ)	“a cabar”	14	3,27	221	51,63	239	55,84	
	(σσσ)	(σ) (σ)	“colo cou”	02	0,47					
	(σσσ)	(σ) (σσ)	“a gora”	195	45,56					
	(σσσ)	(σ) (σ)	“abri se”	09	2,1					
	(σσσ)	(σ) (σ)	“viti mas”	01	0,23					
	(σσσ)	(σ) (σσ)	“sub-mundo”	02	0,47		18			4,2
	(σσσ)	(σ) (σ)	“dança-se”	16	3,73					
4σ	(σσσσ)	(σ) (σσσ)	“im perfeição”	10	2,33	44	10,28	69	16,12	
	(σσσσ)	(σ) (σσ)	“acon teceu”	01	0,23					
	(σσσσ)	(σ) (σσσ)	“ver dadeira”	07	1,63					
	(σσσσ)	(σ) (σ)	“satis feito”	22	5,14					
	(σσσσ)	(σ) (σ)	“relembra mos”	03	0,7					
	(σσσσ)	(σ) (σσ)	“currí culos”	01	0,23					
	(σσσσ)	(σ) (σ)	“anti-fumo”	02	0,47					25
	(σσσσ)	(σ) (σ)	“envolve-se”	14	3,27					
	(σσσσ)	(σ) (σ)	“esta-vamos”	01	0,23					
	(σσσσ)	(σ) (σ)	“mantinha-mos”	08	1,87					
5σ	(σσσσσ)	(σ) (σσσσ)	“homo sexual”	01	0,23	12	2,8	23	5,37	
	(σσσσσ)	(σ) (σσσσ)	“com pletamente”	06	0,7					
	(σσσσσ)	(σ) (σσσ)	“deses perado”	01	0,23					
	(σσσσσ)	(σ) (σσ)	“aero porto”	03	0,7					
	(σσσσσ)	(σ) (σσσ)	“poderia mos”	01	0,23					
	(σσσσσ)	(σ) (σσ)	“espaço-nave”	01	0,23					11
	(σσσσσ)	(σ) (σ)	“comunicar-mos”	03	0,7					
	(σσσσσ)	(σ) (σ)	“estudava-mos”	07	1,63					
6σ	(σσσσσσ)	(σ) (σσσσσ)	“a contecimento”	01	0,23	05	1,16	07	1,63	
	(σσσσσσ)	(σ) (σσσσ)	“tele transportado”	01	0,23					
	(σσσσσσ)	(σ) (σσ)	“vacarosa mente”	02	0,47					
	(σσσσσσ)	(σ) (σσσσ)	“anti penúltimo”	01	0,23					
	(σσσσσσ)	(σ) (σσσσ)	“continuava-mos”	02	0,47					02

Onde: σ indica sílaba; σ̄, sílaba acentuada; parênteses indicam fronteiras gráficas. Linhas tracejadas delimitam ocorrências quanto à localização do acento lexical. Linhas sombreadas restringem-se a dados de hipersegmentação por hífen.

De partida, constata-se que, dentre as 428 hipersegmentações, os trissílabos não apenas são as formas mais hipersegmentadas (55,84%), mas também as que apresentaram o maior conjunto de possibilidades de segmentação não-convencional de palavra em relação aos polissílabos que oferecem um maior número de possíveis hipersegmentações.⁹⁸ Para os trissílabos, predominam hipersegmentações de paroxítonos

⁹⁸ São atestadas 58% (7/12) das possibilidades de hipersegmentação para trissílabos, 55% (10/18) para polissílabos de quatro sílabas, 33% (4/24) para polissílabos de cinco sílabas e 16,6% (5/30) para polissílabos com seis sílabas (cf. nota 83 sobre o cálculo dessas possibilidades). Para dissílabos, as duas possibilidades registradas são segmentações não-convencionais por meio do branco de oxítonos e paroxítonos.

que se caracterizam por isolar a primeira sílaba da palavra entre espaços em branco (45, 56%), o que resulta na configuração de sílaba seguida de Σ troqueu (“a gora”).

O segundo tipo de configuração em número de ocorrências (3,73% do total de dados), dentre as sete configurações atestadas para os trissílabos, compreende hipersegmentação por hífen da última sílaba de paroxítonos (“dança-se”), o que configura hipersegmentação de morfemas verbais principalmente. Esse resultado leva a duas constatações. A primeira delas diz respeito à relação predominante entre as hipersegmentações quanto à localização da fronteira não-convencional e ao tipo de fronteira gráfica: sílabas pretônicas tendem a ser segmentadas por branco, enquanto sílabas postônicas tendem a ser segmentadas por hífen. A segunda constatação é feita a partir da configuração métrica da palavra como fator que motiva hipersegmentações. Na medida em que a configuração de sílaba seguida de pé troqueu é admitida para trissílabos paroxítonos, a localização de fronteira não-convencional de palavra que decorre dessa configuração somente poderia ser aquela que coincide com a fronteira prosódica entre sílaba e pé métrico, como em “com verso”, e nunca aquela que segmenta a sílaba átona do pé, como em “dança-se”. Essa constatação, somada ao fato de o hífen ser o tipo de fronteira gráfica predominante entre essas hipersegmentações, aponta para a atuação de certa estrutura morfossintática nesse subconjunto de registros: a ênclise verbal de pronomes átonos. Essa colocação do pronome enclítico ao verbo (como “ajude-me”) é prosodicamente marcada em PB, não apenas por ser preferencial a colocação proclítica ao verbo (como “me ajude”), mas porque a ênclise verbal decorre de práticas letradas fortemente permeadas pela valorização de registros escritos que não necessariamente são motivados em regras gramaticais atestadas em enunciados falados em PB.

Na próxima seção, características morfossintáticas das hipersegmentações são analisadas detalhadamente a fim de se identificar de que maneira elas atuam nas hipersegmentações da escrita do EF II. A hipótese de que hipersegmentações tenham motivações morfossintáticas não enfraquece o papel da prosódia como fator central que motivaria hipersegmentações de palavra, pois próclise ou ênclise verbal são estruturas em que é crucial a atuação da prosódia⁹⁹ à medida que pronomes pessoais átonos necessariamente devem ser prosodizados a um hospedeiro (uma ω) e a direção dessa prosodização (se enclítica ou proclítica) desses pronomes em relação ao verbo é uma

⁹⁹ Em PE, por exemplo, a ênclise ou a próclise é definida em função da estrutura sintática e do tamanho dos constituintes prosódicos (cf. FROTA; VIGÁRIO, 2001).

informação que decorre da organização rítmica preferencial da língua (cf. discussão em Galves e Abaurre, 1996).

Em resumo, a hipersegmentação da sílaba pretônica de trissílabos paroxítonos indicia a atuação da configuração métrica no domínio de palavra e a hipersegmentação de sílabas postônicas de trissílabos paroxítonos, a atuação da colocação enclítica de sílabas átonas ao verbo que se dá no domínio em que essa informação morfossintática lhe é característica, seja ou não o grupo clítico esse domínio prosódico (cf. discussão na seção anterior sobre a pertinência desse domínio).

Retomando a descrição de resultados, identifica-se que, entre os polissílabos (23,12% das hipersegmentações), há 5,6% (que corresponde a 24% dos polissílabos) de hipersegmentações da primeira sílaba da palavra por meio do branco e 7,94% (que corresponde a 34% dos polissílabos) de hipersegmentações da última sílaba da palavra por meio do hífen. Esses resultados reiteram a relação entre localização da fronteira não-convencional e o tipo de fronteira gráfica atestada entre os trissílabos, embora, em termos quantitativos, haja mais hipersegmentação por hífen do que por branco entre os polissílabos do que ocorre entre os trissílabos, resultado que será retomado mais à frente. Ainda resta tratar de 9,58% (que corresponde a 41% dos polissílabos) de hipersegmentações que se caracterizam pela presença não-convencional do branco no meio da palavra de modo que há coincidência entre essa fronteira gráfica e a fronteira de pés métricos (iambos ou troqueus). Para tratar desses dados, as 99 ocorrências de polissílabos hipersegmentados da Tabela 3.1 foram reorganizadas em função dos tipos de pés troqueu, iambo e dátilo de modo a tornar visível, na Tabela 3.2, como atuam os pés métricos nessas hipersegmentações. Foi mantida a distinção quanto ao tipo de fronteira gráfica (se branco ou hífen) e ao número de sílabas das palavras (quatro, cinco e seis sílabas) na apresentação dos dados, bem como as porcentagens calculadas em função do total de 428 hipersegmentações consideradas nesta seção.

Tabela 3.2: Pés métricos e hipersegmentação de polissílabos

	Pés métricos	Segmentações	Exemplos	Nº	%	Nº	%	Nº	%
Iambo	(σσσσ)	(σ) (σσσ)	“a creditar”	10	2,33	10	2,33		
	(σσσσσ)	(σσ) (σσσ)	“desres peitar”	01	0,23	02	0,47	12	2,80
		(σσ) (σσσ)	“homo sexual”	01	0,23				
Troqueu	(σσσσ)	(σ) (σσσ)	“de volvido”	07	1,63	14	3,27		
	(σσσσσ)	(σ) (σσσσ)	“com pletamente”	06	0,70				
	(σσσσσσ)	(σ) (σσσσσ)	“a contecimento”	01	0,23				
	(σσσσ)	(σσ) (σσ)	“conhe cido”	22	5,14	32	7,47	66	15,42
		(σσ) (σσ)	“anti-fumo”	02	0,47				
	(σσσσσ)	(σσ) (σσσ)	“deses perado”	01	0,23				
		(σσσ) (σσ)	“aero porto”	03	0,70				
		(σσσ) (σσ)	“espaço-nave”	01	0,23				
	(σσσσσσ)	(σσ) (σσσσ)	“tele transportado”	01	0,23				
		(σσσσ) (σσ)	“vacarosa mente”	02	0,47				
	(σσσσ)	(σσσ) (σ)	“começa mos”	03	0,70	20	4,67		
			“envolve-se”	14	3,27				
	(σσσσσ)	(σσσσ) (σ)	“comunicar-mos”	03	0,70				
	(σσσσ)	(σσ) (σσ)	“currí culos”	01	0,23	03	0,70		
			“esta-vamos”	01	0,23				
Dátilo	(σσσσσσ)	(σσ) (σσσσ)	“anti penúltimo”	01	0,23			21	4,90
	(σσσσ)	(σσσ) (σ)	“mantinha-mos”	08	1,87	18	4,20		
	(σσσσσ)	(σσσσ) (σ)	“poderia mos”	01	0,23				
			“estudava-mos”	07	1,63				
	(σσσσσσ)	(σσσσσ) (σ)	“continuava-mos”	02	0,47				

Onde: σ indica sílaba; σ̣, sílaba acentuada; parênteses indicam fronteiras gráficas. Linhas tracejadas delimitam ocorrências quanto ao local de ocorrência de fronteira não-convencional.

Saltam aos olhos que pés métricos dissílabos são recorrentes entre hipersegmentações em que a fronteira gráfica ocorre no meio de palavras polissílabas. Em todas as ocorrências, exceto em uma (“currículos”), a fronteira não-convencional (por branco ou por hífen) localiza-se na fronteira de *pé troqueu dissílabo*. Em todas as ocorrências (8,64% do total de 428 hipersegmentações), a porção da palavra delimitada pelo pé dissílabo corresponde a uma ω. Essa palavra prosódica pode ser: um morfema derivacional, como o sufixo “mente” (“vacarosa mente”), os prefixos “tele” (“tele transportado”) e “anti” (“anti-fumo”), ou uma das bases lexicais de uma palavra composta, como em “porto” em “aeroporto”, “nave” em “espaço-nave”. Em todos esses casos, destaca-se que a fronteira não-convencional corresponde a uma fronteira morfológica. Há ainda ocorrências em que a porção delimitada pelo pé dissílabo não corresponde a uma fronteira morfológica, mas a um possível dissílabo da língua, como “peitar” em “desres peitar”, e também em ocorrências em que essa potencial relação é

mais opaca, como reconhecer “desses” em “deses perado” ou “corri” em “currí culos”.¹⁰⁰ Novamente salienta-se que o lócus da fronteira gráfica não-convencional, nesses dados, corresponde à convergência entre fronteiras morfológicas (mesmo que potenciais) e prosódicas de pés métricos. Essa convergência quanto à localização de fronteiras (gráficas, morfológicas) não enfraquece a análise prosódica proposta, mas a fortalece à medida que é no domínio da palavra prosódica em que se dá a interface entre informações morfológicas e fonológicas.

Mesmo se desprezada essa potencial relação entre pés dissílabos e palavras morfológicas para esses dois últimos exemplos, o resultado relevante para caracterização dos polissílabos hipersegmentados permanece: a correspondência de uma porção hipersegmentada a um Σ dissilábico, predominantemente um pé troqueu. Ainda cabe destacar que, entre os polissílabos em que se identificam pés dátilos, verifica-se tendência em segmentar a sílaba postônica final (18/21 ocorrências) de verbos, o que implica reconhecer um pé dissílabo troqueu na grafia hipersegmentada. Constata-se, pois, como a organização métrica interna à palavra contribui favoravelmente à hipersegmentação de polissílabos.

Em síntese, entre os polissílabos, são três as principais possibilidades de localização da fronteira não-convencional de palavra, tendo sido verificada a seguinte distribuição percentual de hipersegmentações quanto à localização da fronteira não-convencional e o tipo de fronteira gráfica: 34% na última sílaba por hífen; 28% no meio da palavra por branco; e 24% na primeira sílaba de palavra por branco. Desprezado o tipo de fronteira gráfica, os valores desses percentuais aumentam para, respectivamente: 38%, 37% e 24%. Constata-se desses percentuais que não se altera o resultado de haver flutuação, principalmente, entre segmentar a última sílaba do polissílabo – que corresponde a morfema verbal – e segmentar um pé dissilábico dividindo o polissílabo em duas partes – sendo uma delas a que corresponde, majoritariamente, a palavras da língua.

¹⁰⁰ Reconhecer “desses” na grafia não-convencional “deses” está baseado no fato de ainda serem encontrados registros não-convencionais que envolvem o dígrafo “ss” na escrita do EF II e “esperado” na grafia “perado”, um registro da não realização da sílaba pretônica “es” de formas flexionadas do verbo “esperar” na variedade do português falado pelos escreventes cujos textos foram analisados. Reconhecer “corri” em “currí” é fundamentado em pesquisa sobre grafia de vogais pretônicas feita por Reis e Tenani (2011). Naquele estudo, identificam-se registros de “u” para “o” em sílabas pretônicas, predominantemente quando, em verbos, houver vogais altas, como em “corri”. O acento gráfico não-convencional marcaria a sílaba tônica da forma verbal “corri”, além de ser acento previsto para a grafia convencional de “currículos”.

Os principais resultados quanto à hipersegmentação de polissílabos, quando comparados com aqueles descritos para os trissílabos, permitem verificar que há não apenas relação entre tipo de fronteira gráfica e a localização não-convencional da fronteira de palavra, mas também que essa relação sofre efeito do número de sílabas da palavra, pois: (i) para polissílabos, mas não para trissílabos, a hipersegmentação da sílaba postônica final (de paroxítonos e proparo-xítonos) por hífen é mais característica; (ii) para trissílabos, mas não para polissílabos, a hipersegmentação da sílaba pretônica por branco (de paroxítonos e oxítonos) é mais característica. No primeiro caso, a sílaba hipersegmentada, um sufixo verbal, é grafada como se fosse um pronome enclítico ao verbo: como “se” fosse “-sse” em “dança-se” (“dançasse”), “nos” fosse “-mos” em “morava-mos” (“morávamos”).¹⁰¹ No segundo caso, a sílaba pretônica hipersegmentada tende a corresponder a um possível monossílabo átono, como as preposições “com, em” em “com versar”, “em bora”, seguida da porção da palavra em que se encontra o acento lexical. Em ambos os casos, a sílaba hipersegmentada corresponde a um possível clítico da língua, resultado atestado em Tenani (2011b) para amostra transversal de textos do EF II.

Constata-se que, simultaneamente à atuação do Σ , a localização não-convencional de fronteira de palavra coincide com fronteira morfossemântica de, ao menos, uma das porções hipersegmentadas. Essa regularidade quanto à atuação de critérios prosódicos e semânticos na segmentação não-convencional de palavras havia sido observada inicialmente por Abaurre (1991a) na escrita do EF I e, posteriormente, atestadas em pesquisas feitas com outros corpórea de mesma natureza (CUNHA, 2004, 2010; CAPRISTANO, 2007a). Os resultados levam à interpretação de que a escrita do EF II continua a ter essa característica da escrita do EF I, pois é possível identificar palavras sob palavras em todas as 428 hipersegmentações analisadas nesta seção. Em pesquisa anterior conduzida a partir de amostra transversal de textos do EF II (TENANI, 2011a), foram identificadas algumas poucas ocorrências em que porções hipersegmentadas da palavra não correspondem a nenhuma possível palavra da língua, como “su ceço” (“sucesso”), ocorrência que será analisada no capítulo 4.

Por fim, foram analisados os 21,02% de dissílabos hipersegmentados, lembrando que esse percentual é encontrado a partir da exclusão das ocorrências de “porque”,

¹⁰¹ Na próxima seção, serão detalhadas essas potenciais correspondências entre sílabas hipersegmentadas e palavras gramaticais.

“senão” e “demais” por essas totalizarem 48,4% dos dados identificados na amostra de escrita do EF II, o que enviesa os resultados quantitativos, e por apresentarem funcionamentos linguísticos específicos (que serão descritos na próxima seção).

Os dissílabos hipersegmentados são predominantemente oxítonos (17,02% do total de dados), havendo, proporcionalmente, poucos paroxítonos hipersegmentados (3,97%). Como já observado, na língua há mais dissílabos paroxítonos do que dissílabos oxítonos, o que faz supor que a maior porcentagem de hipersegmentação de dissílabos oxítonos não seja necessariamente decorrente de um maior número de palavras desse tipo na língua.

A análise de hipersegmentação de dissílabos oxítonos leva à constatação de que a sílaba átona corresponde a um clítico prosódico e a sílaba tônica, à palavra da língua, como em “na viu”. Nessa ocorrência, a primeira sílaba corresponde à preposição “na” e a segunda sílaba à forma verbal “viu”. Em todas as hipersegmentações identificadas na amostra de escrita do EF II analisadas, a sílaba tônica é constituída de rima complexa, o que permite interpretá-las como pés métricos bimoraicos (como propõe Bisol, 1992). Segundo essa interpretação, grafias como “de pois” passam a ser, prosodicamente, constituídas de sílaba átona seguida de Σ , resultando na configuração predominante entre os trissílabos. Haveria, nessa interpretação de hipersegmentação de oxítonos, a atuação da configuração métrica da palavra ao lado da identificação de monossílabos átonos que corresponderiam a palavras gramaticais do PB.

Quanto ao menor número de dissílabos paroxítonos hipersegmentados, como “que ro” é possível que esse fato decorra de a configuração métrica de troqueu ser característica do PB e, desse modo, a princípio, essa configuração métrica não favoreceria a emergência de hipersegmentações como as atestadas. Portanto, esse subconjunto de grafias não-convencionais (apenas 3,97% do total de hipersegmentações) não está sujeito à atuação do Σ como demonstrado para trissílabos e polissílabos.

O exame das ocorrências de todos esses dissílabos leva à constatação de que uma, se não as duas sílabas, corresponde a palavras ortográficas. À semelhança dos trissílabos e polissílabos que têm a primeira sílaba hipersegmentada, a sílaba átona do dissílabo corresponde, em sua maioria, a um monossílabo que pertence ao conjunto de palavras gramaticais (preposição, artigo, conjunção) como “de” em “des de” (“desde”). Já a sílaba tônica desses dissílabos corresponde, em sua maioria, a uma palavra lexical monossilábica, como “tem” em “tem de” (“tende”).

Portanto, a hipersegmentação de dissílabos paroxítonos tem motivação não na configuração métrica da palavra, mas na identificação de monossílabos como palavras

ortográficas. Essas grafias de dissílabos poderiam ser interpretadas como motivadas exclusivamente em dificuldades ortográficas dos alunos do EF II, herdadas do EF I. No entanto, essas hipersegmentações de dissílabos paroxítonos parecem decorrer do fato que se mostrou recorrente na análise dos trissílabos e polissílabos: a hipersegmentação de sílabas que corresponderiam a palavras gramaticais monossilábicas átonas.

Ora, palavras gramaticais monossilábicas predominantemente configuram-se em clíticos prosódicos em PB e, portanto, têm um caráter complexo: são unidades no domínio da morfossintaxe, mas são sílabas incorporadas ou anexadas a um hospedeiro, no domínio prosódico (cf. discussão feita na seção 3.1). Em outras palavras, a hipersegmentação de sílabas átonas de palavras na escrita do EF II permite verificar que clíticos prosódicos têm sua identificação construída fortemente a partir de práticas letradas escritas. Essa interpretação encontra respaldo na diminuição desses registros em função do aumento do tempo de escolarização, como visto na seção 2.3. Esta é, pois, a característica principal da hipersegmentação da escrita do EF II: a segmentação não-convencional de sílaba quando essas poderiam corresponder a clíticos prosódicos monossilábicos. Interpreto que justamente são as características morfossintáticas e fonológicas desses clíticos que contribuem para a presença de registros não-convencionais em todos os anos do EF II.

3.3 Hipersegmentação e organização morfossintática dos enunciados

Nesta seção, é feita uma caracterização de regularidades quanto à organização morfossintática dos enunciados que também motivam hipersegmentações nos textos da escrita do EF II. A hipótese que guia essa caracterização é a de que certas hipersegmentações são mais bem explicadas se consideradas não apenas características métricas que se configuram no domínio de ω (como as descritas na seção anterior), mas também, de modo mais relevante, características prosódicas dos enunciados que se constituem a partir de informação morfossintática. Será argumentado que características prosódicas dos enunciados são relevantes, porque estão em jogo características morfossintáticas e semânticas das palavras hipersegmentadas que estabelecem relações entre (partes dos) enunciados. Essa hipótese foi, em certa medida, mobilizada na análise de “senão” hipersegmentado (TENANI; LONGHIN, 2015) e será aqui explorada para interpretar, de modo mais amplo, segmentações não-convencionais de palavra.

A fim de alcançar o objetivo de desenvolver análise que demonstre a relevância da referida hipótese, serão analisados dois conjuntos de dados, a saber: (a) “senão” e “demais”, por figurarem entre as hipersegmentações, por branco, mais frequentes na amostra analisada; e (b) hipersegmentações de formas verbais que se caracterizam por usos não-convencionais do hífen, como “dança-se” e “morava-mos”. Para caracterizar esses dados, serão consideradas também as hipersegmentações analisadas na seção anterior, servindo a análise realizada como base de comparação. O ponto de partida dessa caracterização será a classe gramatical a que pertencem as palavras hipersegmentadas, uma vez que essa classificação¹⁰² – mesmo que seja passível de problematização¹⁰³ – será útil para demonstrar que características gramaticais (morfológicas, sintáticas e semânticas) das palavras hipersegmentadas não se restringem a um tipo de palavra, mas são mobilizadas no processo de construção do texto escrito de modo complexo.¹⁰⁴

No conjunto de 941 dados (93,91% da amostra), predominam hipersegmentações de chamadas palavras gramaticais (64,61%) em relação às palavras lexicais (35,38%), conforme se verifica na Tabela 3.3. Dentre as palavras gramaticais hipersegmentadas, identificam-se, em ordem decrescente de ocorrência de dados: conjunções, pronomes e preposições. Dentre as palavras lexicais, as hipersegmentações atingem, em ordem decrescente de ocorrência: verbos, advérbios, substantivos e adjetivos.

¹⁰² Para a classificação em classe gramatical das palavras hipersegmentadas, foi consultado o dicionário Houaiss (2001), além de ser considerada a estrutura sintática em que o registro não-convencional ocorreu.

¹⁰³ Uma introdução à problematização dos conceitos de palavra (fonológica, morfológica, sintática) se encontra em Villalva e Silvestre (2014). Uma discussão aprofundada é feita por Veloso (2016).

¹⁰⁴ A consideração de classe de palavra também permite estabelecer diálogo com profissionais da área de Educação, especialmente professores e coordenadores pedagógicos que atuam junto ao EF II. Na medida em que conceitos sobre classe de palavras integram o conteúdo programático do EF II de língua portuguesa previstos para o ciclo II do EF (cf. Proposta Curricular do Estado de São Paulo, 2009), os alunos são chamados a fazerem uso desses conceitos também em suas produções escritas. Nesse processo, vislumbram-se os chamados erros de segmentação de palavra que podem ser vistos como efeitos não exclusivamente motivados em defasagem de aprendizagem do aluno, mas também no próprio processo de ensino/aprendizagem de conteúdos de língua portuguesa.

Tabela 3.3: Classe gramatical das palavras hipersegmentadas

Palavra	Classe gramatical	Nº	%	Nº	%
Lexical	Verbo	146	15,51		
	Advérbio	118	12,54	333	35,38
	Substantivo	49	5,21		
	Adjetivo	20	2,12		
Gramatical	Conjunção	468	49,73		
	Pronome	134	0,64	608	64,61
	Preposição	06	14,24		

Dentre as palavras gramaticais, predomina a hipersegmentação de conjunções (com índice de 49,73%) em razão, principalmente, da alta frequência de “porque” (410/468) e “senão” (36/468). Excluídas essas ocorrências, restam apenas 4,48% de conjunções hipersegmentadas. Embora em maior número de ocorrências, não analisaremos as grafias de “porque”, mas as grafias de “senão”, pois a complexidade linguística dessa conjunção permitirá demonstrar que a hipersegmentação de palavras também é efeito de relações que se dão entre (partes dos) enunciados.

Antes de avançar na análise, observa-se que, diferentemente das grafias de “senão”, nos usos de “porque” hipersegmentado não foram encontradas evidências de que sejam motivados em função de características morfossintáticas e semânticas dos enunciados. Em estudo feito por Soncin e Longhin (2014), a partir de ocorrências selecionadas do corpus desta tese, é descrita a complexidade sintático-semântica de funcionamento de “porque” em construções causais. As autoras argumentam que gêneros de discurso são fatores condicionantes da emergência e do uso de construções que materializam relações de sentido no domínio causal mobilizado por “porque”, mas não identificam relação entre estruturas ou sentidos mobilizados nas construções e as grafias de “por que” e “porque” nos manuscritos do EF II. A flutuação entre “por que” e “porque” observada nesses textos do EF II seria, pois, decorrente do processo de *convencionalização da escrita* que diz respeito à inserção do aluno em práticas de escrita formais.

Argumento, porém, que segmentar palavras convencionalmente implica categorização do espaço em branco da escrita alfabética, portanto, segmentações não-convencionais podem ser interpretadas como rastros desse processo de categorização do espaço entre palavras que, naturalmente, engloba convencionalização, porém dele não

decorre. Ou seja, espaços não-convencionais entre palavras decorrem de relações morfossintáticas e semânticas, além das prosódicas, entre (porções dos) enunciados, como será demonstrado mais à frente com base na análise de hipersegmentações de “senão”, não sendo efeito tão somente do processo de convencionalização de registros gráficos de palavras.

Ainda no que diz respeito aos resultados da Tabela 3.3, constata-se que palavras lexicais totalizam 35,38% das hipersegmentações e, dentre esse conjunto, predominam formas verbais hipersegmentadas (146/333), se consideradas ocorrências por branco (93 dados) e por hífen (53 dados). Essas formas verbais hipersegmentadas por hífen foram selecionadas para análise não apenas por serem predominantes entre os registros de usos não-convencionais do hífen (que totalizam 91,37%), mas também por – diferentemente das hipersegmentações por branco¹⁰⁵ – passarem a ocorrer nos anos finais do EF II, fato que merece ser analisado à luz da hipótese apresentada nesta seção. Somam-se a esse resultado as considerações acerca do hífen discutidas na seção 2.2, que imprimem a esses dados características que os particularizam em relação às hipersegmentações por branco.

Se excluídas essas hipersegmentações por hífen, são os advérbios as formas mais hipersegmentadas por branco (118/333) dentre as palavras lexicais. Ainda entre as palavras lexicais, “demais” apresenta o segundo maior índice de frequência de hipersegmentação, sendo o advérbio que mais ocorrências apresenta (41/333). Acrescenta-se que os registros de “demais” hipersegmentado, à semelhança dos registros de “senão” hipersegmentado, podem ser mais bem explicados como efeitos de relações que se dão entre (partes dos) enunciados.

Em suma, nesta seção, a análise será feita a partir hipersegmentações de “demais”, “senão” e de formas verbais hipersegmentadas por hífen, assim ordenados em função do número de ocorrência na amostra considerada do EF II.

No que diz respeito às hipersegmentações de “demais”, a segunda forma mais frequentemente hipersegmentada no material analisado (4,1%) foi objeto de um estudo de natureza morfossintática feito por Tenani e Oliveira-Codinhoto (inédito). Nesse estudo, verificou-se que apenas ocorrem usos de “demais” enquanto advérbio, com duas

¹⁰⁵ Dentre as formas verbais hipersegmentadas por branco (com três ou mais sílabas), 68,3% (41/60) são sílaba pretônica hipersegmentada (como “a caba”, “com parecer” “em louqueci”) e 10% (6/60) são segmentações no meio da palavra (como “recom pensa”), características discutidas na seção anterior. Houve 21,6% (13/60) de formas verbais em que a última sílaba é hipersegmentada, como “come sou”. Nesses dados, a hipersegmentação resulta também da identificação de formas verbais

das três acepções registradas em Houaiss (2001), a saber: (i) algo “em excesso, além da conta, além da justa medida”, atuando enquanto item enfático e intensificador; (ii) “demasiadamente, de maneira muito forte”, atuando enquanto advérbio de modo. Na origem etimológica, essas acepções foram atestadas e a grafia apenas variou quanto ao registro da semivogal: no século XIII, encontram-se as grafias “demays” e “demaes”, ao lado de “demais”. Portanto, “demais” sempre teve a mesma grafia no que diz respeito à segmentação das sílabas que o constitui (diferentemente do item “senão”, que será abordado mais à frente).

Outro aspecto morfossintático relevante para a análise da hipersegmentação é a observação encontrada em Bechara (2001) sobre o fato de “demais” ser sempre grafado sem espaço entre “de” e “mais”, quando for locução – com função de advérbio ou pronome – que se contrapõe a “de menos”. Desse modo, em uma relação paradigmática, “demais” se contrapõe a “de menos”. Por outro lado, em uma relação sintagmática, “de mais” é um registro convencional da preposição “de” seguida do advérbio “mais”, como em “Preciso de mais folha de papel”. Essa construção não implica que “demais” e “de mais” sejam formas homônimas em português (cf: “Preciso demais dessa folha de papel”). Desse modo, as ocorrências de “de mais” atestadas nos textos do EF II são classificadas como casos de segmentação não-convencional do advérbio “demais”, sendo as segmentações não-convencionais decorrente tanto das possibilidades que a língua oferece na construção das sentenças (“demais” X “de mais”), quanto de relações paradigmáticas que o escrevente possa fazer entre “demais” e “de menos”.

Nesta tese, defendo, a partir de Abaurre (1991a), que noções morfossintáticas e semânticas (usadas para estabelecer as convenções ortográficas) são mobilizadas simultaneamente com as prosódicas (como pé métrico e palavra prosódica) pelos escreventes no registro gráfico dos limites da palavra. Avanço, porém, em relação às considerações dessa autora ao demonstrar que *a organização rítmica e entoacional do enunciado*, que abrange domínios maiores do que a palavra, motiva as grafias não-convencionais que se mantêm no EF II. É, pois um conjunto de motivações linguísticas que leva à hipersegmentação de “demais”, como a seguir exemplificado.

- (1) No final do ano passado eu tinha prometido não falar muito mas não deu certo porque os meus pais me falava para eu ficar queto [porque **eu falava de mais**], então eu prometi a mim mesmo que eu falava menos mas não deu certo porque eu não consigui me controla (Fonte: Z09_6D_11M_01)

(2) eu sou uma pessoa que independente do que os outros falem [**eu so feliz de mais**]_I sou meio loirra gosto de pagode (Fonte: Z11_8A_24F_02).

Na ocorrência (1), exemplifica-se a tendência de o verbo ser escopo de “demais”, que frequentemente o precede, podendo ocorrer um complemento verbal entre o verbo e “demais”, como em (2). Portanto, sintaticamente, “demais” atua sobre o predicado da oração. Semanticamente, a intensificação trazida pelo significado do advérbio “demais” não fica restrita somente ao predicado ou aos modificadores, mas, por extensão, à situação descrita pela oração ou à avaliação que se faz dessa situação descrita pela oração.

Quanto às características prosódicas, observa-se que a posição de “demais” em fronteira de *I*, como nos exemplos dados, implica que às duas sílabas de “demais” estão associados eventos tonais que configuram fim de *I* em enunciados afirmativos: tom alto (H) associado à sílaba “de”, seguido de tom baixo (L) associado à sílaba tônica “mais” (TENANI, 2002).¹⁰⁶ Em estudos sobre a configuração prosódica dos enunciados falados em PB que adotam o arcabouço da Fonologia Prosódica (SERRA, 2009; FERNANDES, 2007; TENANI, 2002), afirma-se que o evento tonal H+L*, que representa contorno descendente alto-baixo é marca prototípica de fim de enunciado afirmativo neutro. Em outras palavras, o evento tonal H+L* codifica a informação gramatical de fim de enunciado e traz o acento tonal do enunciado. Ou seja, o evento tonal associado às sílabas da palavra em fim de *I* codifica que essa palavra é a mais proeminente em relação às demais que estão no mesmo *I*.

Diferentemente das ocorrências anteriores quanto a características formais é a ocorrência em (3). Nesse exemplo, “de mais” é delimitado por sinais de pontuação e constitui um comentário por meio do qual o escrevente expressa sua avaliação em relação à descrição feita anteriormente. Semanticamente, “demais” tem como escopo todo o enunciado que lhe é precedente. Prosodicamente, “demais” constitui um *U*.

(3) E como o sonho de toda menina quero conhecer o castelo da cinderela, vejo muitas reportagem pela televisão imagina como seja e sempre tive vontade vou adorar fazer essa viagem. [**De mais.**]_U (Fonte: Z08_5A_29F_06)

¹⁰⁶ Cabe explicitar que os tons H (alto) e L (baixo) são representações fonológicas associadas às sílabas dos enunciados. Os tons representam realizações fonéticas as quais, por sua vez, são identificadas a partir da variação da frequência fundamental (F0). A configuração dos tons é dada pelas estruturas prosódicas construídas a partir de relações sintático-semânticas dos enunciados (cf. TENANI, 2002).

Em (3), verifica-se que “de mais” atua não localmente no escopo de partes do enunciado (como 1 e 2 exemplificam), mas na construção da significação do texto. Na medida em que se admite que o escopo de “demais” abrange, sintaticamente, orações ou sentenças, projeta-se que também, prosodicamente, está em jogo relação entre constituintes prosódicos, particularmente, entre U_s . Essa relação prosódica entre U_s implica haver maior saliência do enunciado constituído por “demais”, de modo que cada uma das sílabas recebe um evento tonal que as destacam. Nessa interpretação, o espaço em branco em “demais” indicia, na materialidade da escrita, a ênfase dada ao enunciado [De mais]_U, por meio do qual o escrevente expressa sua avaliação em relação aos enunciados anteriores.

Essa configuração entoacional somada à posição do item “demais” em fronteira do constituinte prosódico *I* (em 1 e 2) e *U* (em 3) confere à hipersegmentação uma forma de registrar, por meio do espaço entre as sílabas, a saliência prosódica do item, destacando partes do enunciado (em 1 e 2) ou o enunciado (em 3) em relação ao que lhes precede.

Ora, é uma prática oral destacar palavras de um enunciado por meio de ritmo silábico (CAGLIARI, 2007, p. 141). Essa prática oral é passível de ser representada na escrita convencional por meio da colocação de hífen entre as sílabas, como em: “Ela fala de-va-gar!”.¹⁰⁷ Para os dados em análise, poderia se supor como registro convencional do destaque a porções do enunciado o uso de hífen e do ponto de exclamação, por exemplo: “mas o rei era nervoso até de-mais!”. Essa hipótese explicativa para a motivação da hipersegmentação toma características rítmicas e entoacionais dos enunciados falados como base para interpretar que a hipersegmentação iconiza o destaque dado a “demais” nos enunciados escritos do EF II. Portanto, a hipersegmentação de “demais” seria motivada não pela configuração métrica da palavra, mas pela configuração rítmico-entoacional do enunciado em que a palavra ocorre.

No que diz respeito à forma “senão” (a terceira hipersegmentação mais frequente no corpus com 3,6% dos dados), serão descritas características da constituição histórica dessa conjunção e de seu funcionamento sintático-semântico nos enunciados escritos no EF II relevantes para os objetivos desta tese. Estudo detalhado do complexo funcionamento de “senão” encontra-se em Tenani e Longhin (2015).

¹⁰⁷ Ramos (2009, p. 66), ao descrever estratégias de representação da oralidade em histórias em quadrinhos, afirma que o uso do hífen para separar sílabas sugere que o personagem falou de maneira silabada. Mas não só em quadrinhos, mas também em crônicas se encontra esse uso do hífen, como por exemplo: “Comem-se os ovos e toma-se a sopa: uma de-lí-cia.” (ANDRADE, 1997). Portanto, silabação que destaca ou enfatiza uma palavra é traço da oralidade representado por meio de hífen entre sílabas.

Desse estudo, vale destacar que, na constituição histórica de “senão”, Houaiss (2001) sinaliza duas possibilidades gráficas, “senão” e “se não”, e registra, em apêndice gramatical, que a grafia “se não” é uma alternativa para um dos usos do item, aquele em que “senão”, delimitado por vírgulas, funciona como conjunção alternativa (“perdoe; se não, a vida lhe será amarga”). As autoras assumem a flutuação gráfica entre “senão” e “se não”, sendo a primeira grafia não-marcada (porque é a grafia preferencial para a conjunção) e a segunda, marcada (uma vez que a grafia “se não” é preferencialmente empregada quando houver a conjunção “se” seguida do advérbio de negação “não”, como em “Se não chegar cedo, perderemos o cinema”). A grafia marcada consiste na opção menos frequente e contextualmente restrita em relação aos usos de “senão” abonados pela ortografia do português. Dessa perspectiva, a grafia “se não”, encontrada nos textos do EF II, é convencionalmente marcada e não exatamente uma hipersegmentação. Essa característica de flutuação quanto ao registro gráfico é comum a palavras que passaram por processos de mudança linguística. Desse modo, a flutuação gráfica registrada em dicionário decorre do processo histórico de constituição da própria palavra. A forma “senão” é, pois, fruto de um processo de mudança que levou o condicional “se” e a negativa “não” a serem reanalisados como uma única palavra gramatical.

No exame das ocorrências de “senão” nos textos do EF II, as autoras identificam quase absoluta preferência pela segmentação de “se não” (opção marcada) e a análise de seus usos mostra um padrão polissêmico muito peculiar, com uma multiplicidade de sentidos. Daquela análise de Tenani e Longhin (2015), é importante retomar a possível correlação entre a expressão dos sentidos e a segmentação gráfica de “senão”, a partir da articulação entre mecanismos interpretativos de natureza semântico-pragmática e mecanismos sintáticos e fonológicos. A hipótese que guiou aquela análise é a de que, nas grafias atestadas no EF II, estaria plasmada outra flutuação: a constituição do próprio item, tanto no aspecto categorial, como na emergência dos sentidos. Argumentou-se que à constituição morfossintática e morfossemântica de “senão” estão associadas características prosódicas, as quais emergem por meio da grafia marcada “se não”. Articulada a essa hipótese central, foi formulada outra hipótese explicativa que interpretou as grafias marcadas “se não” como indícios de configuração prosódica dos enunciados, particularmente, aquelas configurações que dizem respeito às maneiras pelas quais proeminências métricas e ênfase entoacional podem ser associadas a porções dos enunciados articulados por “senão”, como exemplificado a seguir.

- (4) (i) [[Assalto,]_I [fica quieto]_I [se não eu atiro]_I]_U
 (ii) [[Assalto,]_I [fica quieto]_I [se não]_I [eu atiro]_I]_U
 (iii) [[Assalto,]_I [fica quieto se não eu atiro]_I]_U

Observa-se que, independentemente de cada uma das possibilidades de fraseamento prosódico de (4), “não” é sempre candidato a carregar o acento de *I* a que pertencer. Em todas as potenciais representações prosódicas (indicadas por meio de colchetes em (4)), “não” é o elemento enfatizado. Essa interpretação se sustenta à medida que se identifica, por meio da análise semântico-pragmática, ênfase em “não” por esse elemento trazer a informação mais saliente dos enunciados, qual seja: “não” nega a hipótese introduzida por “se”, “se não (ficar quieto)”, e, em seguida, é apresentada a consequência desfavorável dessa negação (“eu atiro”). A ênfase tem uma expressão prosódica e se realiza, conforme estudo de Fernandes (2007) sobre o PB, por meio de tom associado à sílaba mais proeminente do enunciado aliada à desacentuação tonal. Portanto, a sílaba “não” de “senão” que carrega tom associado à ênfase prosódica é a mais proeminente do enunciado. É essa sílaba da palavra enfatizada nos textos escritos analisados que aparece representada por meio da delimitação, entre espaços em branco. Dessa perspectiva, o espaço em branco é interpretado como registro que dá visibilidade, na escrita, à ênfase em “não” identificada nos enunciados em que “senão” ocorre.¹⁰⁸

As considerações sobre as grafias não-convencionais de “demais” e “senão” (com alta frequência no material analisado) permitem afirmar que, no EF II, um conjunto específico de palavras gera dúvidas ortográficas. Estão em jogo par de grafias não marcada e marcada (respectivamente: “senão” e “se não”) e par de palavra e estrutura (“demais” e “de mais”) que são apenas homófonos entre si, mas não homônimos, uma vez que a grafia depende da estrutura morfosintática. Nesses dois casos, a informação verbo-visual sobre a colocação do espaço em branco é fundamental para a representação ortográfica dessas formas linguísticas. O emprego convencional de cada uma demanda o domínio de estruturas morfosintáticas mobilizadas, predominantemente, em estruturas

¹⁰⁸ A análise mais aprofundada de “se não”, publicada em Tenani e Longhin (2015), primou por interpretar dados de escrita como representações de estruturas morfosintáticas e semântico-pragmáticas que recuperam estágios anteriores da língua e como pistas de representações prosódicas dos enunciados articulados por “senão”. Outra contribuição consistiu em postular a noção de *grafia marcada*, que é uma grafia prevista, embora não privilegiada pela convenção ortográfica, como relevante para o estudo dos chamados erros de segmentação de palavra. Ao contrastar as grafias marcada “se não” e não marcada “senão”, essa proposta contribuiu com os estudos sobre interpretação e classificação de registros de segmentação não-convencional de palavra, bem como sobre a relação entre fala e escrita que esses dados permitem investigar.

inter-oracionais que são parte dos conteúdos de língua portuguesa trabalhados no EF II (cf. Secretaria da Educação do Estado de São Paulo (2009), tema: “Relações entre usos de língua falada e escrita”). Portanto, identifica-se não apenas quais palavras geram os chamados erros ortográficos, mas também quais estruturas morfossintáticas suscitam grafias não-convencionais produzidas por estudantes do EF II.

A ancoragem em características prosódicas dos enunciados falados (que abrange diferentes constituintes prosódicos) se apresenta como fator importante que motiva as grafias não-convencionais atestadas. Essa constatação implica explicitar a relevância de reflexões sobre possíveis relações entre fala e escrita no EF II. Essas reflexões serão tratadas no capítulo final da tese.

Por fim, serão tratadas as hipersegmentações que se constituem de usos não-convencionais do hífen em estruturas verbais como se fossem sequência verbo-clítico, como se verifica em “vira-se” (“virasse”) e “brincava-mos” (“brincávamos”). De modo geral, essas grafias – que totalizam 5,6% das hipersegmentações – são evidências de tentativas de emprego de estruturas verbais enclíticas. O uso de hífen isola sufixos verbais homófonos a pronomes do português.

Dois são os casos principais de hipersegmentação por hífen, a saber: (i) em 47,17% dos dados, o sufixo de pretérito imperfeito do subjuntivo “-sse” é grafado como se fosse o pronome reflexivo de terceira pessoa “se”, em uma colocação enclítica ao verbo (“trabalha-se”); (ii) em 45,28% dos dados, o sufixo de primeira pessoa plural “-mos” é grafado como se fosse o pronome reflexivo de primeira pessoa “nos”, em uma colocação enclítica ao verbo (“colocava-mos”).¹⁰⁹ Identifica-se, pois, relação entre tipo de fronteira gráfica, local em que essa fronteira ocorre dentro da palavra e classe gramatical da palavra hipersegmentada: o hífen é empregado não-convencionalmente para segmentar a última sílaba de forma verbal flexionada que é, em 92,45% dos dados, morfemas analisados como se fossem pronomes. Os 7,55% restantes compreendem quatro formas verbais hipersegmentadas, a saber: “chama-da”, “coloca-vam”, “esta-vamos” e “deixo-u”. Em todas essas hipersegmentações em que a sílaba final é hipersegmentada, o hífen ocorre entre tema e morfemas verbais. Ainda em 83,01% (44/53) das hipersegmentações, as formas verbais pertencem ao tempo pretérito imperfeito, sendo 47,16% (25/53) das formas flexionadas no modo subjuntivo e 35,85% (19/53) no modo indicativo.

¹⁰⁹ Dentre esse conjunto de hipersegmentação, as formas verbais pertencem aos seguintes tempos verbais: pretérito imperfeito do indicativo (15/24), futuro do subjuntivo (7/24) e pretérito imperfeito do subjuntivo e presente do indicativo (uma ocorrência de cada).

Somam-se a essas tendências quanto aos usos não-convencionais do hífen em formas verbais os resultados de que essas hipersegmentações (i) ocorrem predominantemente (47,2%) no nono ano do EF,¹¹⁰ e (ii) são produzidas majoritariamente (72,72%) por meninas. Esses dois resultados vão na direção oposta aos resultados gerais, descritos na seção 2.3, quanto à hipersegmentação no EF II: há queda de frequência de hipersegmentação em função do aumento dos anos letivos e há menor frequência de hipersegmentação entre meninas. A fim de investigar motivos para esses resultados de hipersegmentação por hífen, lançou-se mão da análise das produções escritas dos sujeitos ao longo do EF II. A seguir, é descrita a trajetória de uma menina¹¹¹ quanto aos usos não-convencionais do hífen em formas verbais que apresenta aumento de hipersegmentação no último ano do EF.

A aluna Camila (sujeito 46 no banco de dados) empregou, em 11 dos 20 textos que escreveu, 71 formas verbais flexionadas na primeira pessoa do plural. Dentre essas formas, oito são grafadas não-convencionalmente, sendo todas flexionadas no pretérito imperfeito do indicativo.¹¹² Exemplo eloquente do tipo de correlação entre o uso não-convencional do hífen (tipo de recurso gráfico que delimita palavra), estrutura morfossintática (colocação enclítica ao verbo de pronomes átonos) e estrutura prosódica (direção à direita de prosodização de clíticos em relação ao hospedeiro na fala *versus* a colocação de clítico à esquerda do hospedeiro na escrita) é o texto na Figura 3.4, produzido por Camila no nono ano do EF.

¹¹⁰ No que diz respeito à distribuição de dados em relação aos anos letivos, vale observar que, no nono ano, regras de usos de hífen fazem parte do conteúdo programático previsto pela Proposta Curricular do Estado de São Paulo (2009). Essa distribuição de dados traz indícios de que a incidência de hipersegmentação por hífen são motivadas, também, no processo de ensino/aprendizagem das convenções ortográficas do português. Investigar essa potencial relação escapa aos objetivos desta tese, mas fica registrada a questão que merece investigação em pesquisa futura.

¹¹¹ Os alunos serão identificados por nomes fictícios a fim de preservar suas identidades, conforme previsto pelas normas de ética em pesquisa estabelecidas pelo CONEP – Comissão Nacional de Ética em Pesquisa.

¹¹² Das oito formas verbais grafadas não-convencionalmente, uma ocorreu no sétimo ano, uma no oitavo ano e seis no nono ano.

Figura 3.4: Usos não-convencionais de hífen em formas verbais

Esta amizade minha e da Lidiame começou em 2009 quando estava-mos na 6ª série. A primeira vez que conversamos foi no segundo dia de aula, estudava-mos à tarde, conversa-mos na quadra, foi muito bom. E ~~(daí-a)~~ deste dia à dianta começamos a nos falar mais, e também man-tinha-mos contato pela internet. Era muito legal es-tudar com ela, agente brincava, cantava, dançava, e até acabava-mos indo para diretoria. E assim foi 2009 estudando juntas, fazendo trabalhos, provas, jogando futebol com os garotos. Quando saímos de férias con-versávamos as vezes pelo telefone e pelo MSN. Aí veio 2010, estávamos na 7ª série, caímos na mesma turma, continuava-mos brincando muito, eu ia na casa dela fazer trabalho, comia-mos, agente comia briga-deiros, assistia-mos filmes, demos muita, muita, muita risada e assim também foi mais um ano, passou tão rápido ao lado dela, é claro que eu tinha outros amigos e amigas, nunca briguei com ^{nenhum} ~~(todos)~~ deles, em três anos de amizade, mas agora na 8ª série não caímos na mesma sala, infelizmen-te, mas ainda fazemos muita coisa juntas e assim vai durar muitos anos ainda “Eu te amo amiga”.

Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z11_8B_18F_01)¹¹³

No texto da Figura 3.4, a aluna usou 36 formas verbais, sendo sete delas grafadas não-convencionalmente, a saber: “estava-mos, estudava-mos, mantinha-mos, acabava-mos, conversava-mos, continuava-mos, assistia-mos”. Essas hipersegmentações por hífen presente na fronteira do sufixo “-mos” (que também é a última sílaba átona do pé dátilo), ocorreram somente em formas verbais flexionadas no pretérito imperfeito do indicativo,

¹¹³ Leitura possível: “Esta amizade minha e da Lidiame começou em 2009 quando estava-mos na 6ª série. A primeira vez que conversamos foi no segundo dia de aula, estudava-mos à tarde, conversamos na quadra, foi muito bom. E ~~(daí-a)~~ deste dia à diante começamos a nos falar mais e também mantinha-mos contato pela internet. Era muito legal estudar com ela. Agente brincava, cantava, dançava, e até acabava-mos indo para diretoria. E assim foi 2009 estudando juntas, fazendo trabalhos, provas, jogando futebol com os garotos. Quando saímos de férias conversávamos as vezes pelo telefone e pelo MSN. / Aí veio 2010, estávamos na 7ª série, caímos na mesma turma, continuava-mos brincando muito, eu ia na casa dela fazer trabalho (~~comia-mos~~) agente comia brigadeiro, assistia-mos filmes, demos muita, muita, muita risada e assim também foi mais um ano, passou tão rápido ao lado dela, é claro que eu tinha outros amigos e amigas, nunca briguei com ~~(todos)~~ nenhum deles, em três anos de amizade, mas agora na 8ª série não caímos na mesma sala, infelizmente, mas ainda fazemos muita coisa juntas e assim vai durar muitos anos ainda “Eu te amo amiga”.”

na primeira pessoa do plural. Dentre as 15 formas flexionadas na primeira pessoa do plural, apenas uma forma flexionada no pretérito imperfeito do indicativo é gravada convencionalmente (“estávamos”); as demais formas grafadas convencionalmente pertencem ao pretérito perfeito do indicativo (“conversamos, começamos, caímos (1ª ocorrência), saímos”) e ao presente do indicativo (“caímos” (2ª ocorrência) e “fazemos”). Nessas grafias, o sufixo “-mos”, sílaba átona, constitui, com a sílaba tônica da forma verbal, pé troqueu dissílabo. Ainda vale destacar que são grafadas convencionalmente todas as quatro formas verbais flexionadas no pretérito imperfeito do indicativo na terceira pessoa singular (“brincava, cantava, dançava, comia”). Portanto, as grafias não-convencionais da Figura 3.4 são restritas às formas verbais no pretérito imperfeito do indicativo na primeira pessoa plural.

Essa descrição põe em evidência a convergência de dois fatores que favorecem a hipersegmentação de formas verbais: tanto característica morfológica, quanto característica métrica das formas flexionadas. Quanto à caracterização métrica, argumentou-se – na seção anterior – que a relação de proeminência entre as sílabas que leva à configuração de pés métricos no domínio da palavra se mostra como uma das características dos enunciados que proporciona hipersegmentações. As grafias não-convencionais atestadas na Figura 3.4 permitem constatar que também características morfológicas atuam favoravelmente à hipersegmentação por hífen na medida em que preferencialmente formas flexionadas no pretérito do imperfeito na primeira pessoa plural têm o sufixo “-mos” hipersegmentado. Vale destacar que formas de expressão gramatical de primeira pessoa plural é aspecto importante a ser observado nesta análise: Camila flutua entre a expressão de primeira pessoa plural por meio de “a gente” (grafado “agente”) – que leva o verbo a ser flexionado na terceira pessoa do singular – e expressão dessa mesma noção gramatical por meio de sufixo verbal “-mos” – que resulta em forma verbal menos frequente (quando comparada com as formas verbais flexionadas na terceira pessoa do singular) no conjunto de textos por ela mesma produzidos.

Há um terceiro fator que conspira para que essas formas verbais sejam hipersegmentadas: a colocação de pronomes em relação ao verbo. Estão em jogo as duas possibilidades de colocação pronominal (a saber: próclise ou ênclise verbal) às quais são associados modos de enunciação e práticas orais e letradas. Nas práticas orais/faladas, predominam usos de próclise verbal (como “nos mantinha”), enquanto nas práticas letradas/escritas, usos de ênclise verbal (como “mantinha-nos”). A projeção da estrutura verbo-pronome se mostra nos registros não-convencionais de formas verbais

identificados no texto da Figura 3.4, na medida em que há semelhanças fonológicas e morfológicas entre a forma verbal hipersegmentada (“mantinha-mos”) e a estrutura morfossintática projetada (“mantinha-nos”), pois: (i) o pronome “nos” e o sufixo “-mos” carregam informação gramatical de primeira pessoal plural e, quanto à materialidade fônica, ambas as sílabas são semelhantes entre si, diferindo apenas quanto ao ponto de articulação da consoante nasal; (ii) a porção hipersegmentada que compreende o tema verbal, como “mantinha”, tem a mesma cadeia segmental que a forma verbal flexionada no pretérito imperfeito do indicativo em terceira pessoa do singular, como “ele mantinha”. Essas semelhanças entre formas verbais somadas às convenções de uso do hífen corroboram a projeção da estrutura verbo-clítico (“mantinha-nos”) no registro de forma verbal flexionada no pretérito imperfeito do indicativo na primeira pessoa do plural (“mantínhamos”). Essa descrição permite vislumbrar como estruturas morfossintáticas são, também, motivadoras de hipersegmentações de palavras no EF II, particularmente de hipersegmentação de formas verbais.

Por meio desta análise em que as produções escritas dos alunos ao longo dos quatro anos do EF II são consideradas, esta tese avança em relação aos estudos feitos a partir de subconjuntos de dados extraídos da amostra longitudinal de manuscritos do EF II (PARANHOS, 2014; SILVA, 2014), pois demonstra de que modo as hipersegmentações por hífen se revelam motivadas pelas estruturas morfossintáticas que os alunos são chamados a empregar em seus textos escritos (como formas verbais flexionadas na primeira pessoa plural, emprego de ênclise verbal). Notadamente, são as meninas que, ao respondem positivamente ao solicitado, deixam rastros (mais visíveis ao analista) do processo no qual estão imersas.

De modo amplo, esses usos não-convencionais do hífen em formas verbais, bem como os registros não-convencionais de “demais” e marcados de “senão”, devem ser interpretados como resultados do imaginário social do escrevente sobre a escrita, construído pela sua inserção nos mais diversos usos (formais/informais) da linguagem (CORRÊA, 2004). A tentativa de se aproximar do que seja o código escrito institucionalizado (CORRÊA, 2004) leva o aluno ao chamado erro ortográfico. É o trânsito do escrevente por práticas orais (como aquelas em que predominam formas verbais flexionadas na terceira pessoa do singular e colocação pronominal proclítica) e por práticas letradas (como aquelas que demandam emprego de formas verbais flexionadas no pretérito imperfeito e em primeira pessoa do plural e colocação

pronominal enclítica ao verbo) que fica evidenciado nos registros não-convencionais do hífen, por exemplo.

Notadamente, esse trânsito se revelou marcado pelo sexo/gênero dos sujeitos, na medida em que predominantemente meninas grafam não-convencionalmente por hífen. Esse resultado qualitativo também pode ser interpretado como evidência de que meninas/mulheres são constituídas por práticas letradas/escritas de modo distinto dos meninos/homens, como argumentado na seção 2.3, quando da discussão dos resultados quantitativos da amostra do EF II. Esses resultados quanti e qualitativos, da perspectiva desta tese, são evidências de que não apenas há entrelaçamento entre práticas orais/faladas e letradas/escritas, mas que esse entrelaçamento é constituído pelos papéis sociais ocupados pelos sujeitos da linguagem. Essas relações entre práticas orais/faladas e letradas/escritas serão detalhadas no próximo capítulo.

Ainda se faz relevante observar que as hipersegmentações analisadas nesta seção (“de mais”, “se não” e formas verbais hipersegmentadas por hífen) podem ser interpretadas como registros verbo-visuais que revelam em que medida e de que maneira relações morfossintáticas e semânticas somadas a características prosódicas dos enunciados deixam aflorar características do processo de produção escrita dos escreventes analisados. Em última instância, constata-se, mais uma vez, que características prosódicas constitutivas dos enunciados pulsam nos registros escritos não-convencionais de palavra. Em outras palavras, demonstra-se, por meio da análise de hipersegmentações, como se dá a relação entre prosódia e escrita em textos do EF II.

4 Hipersegmentação, prosódia e sentidos dos enunciados

Neste capítulo, são analisadas hipersegmentações de palavras que se particularizam em relação às características quantitativamente predominantes nos textos escritos do EF II, descritas no capítulo anterior. O objetivo é argumentar que segmentações não-convencionais de palavra também são efeito das tensões discursivas sofridas pelos escreventes ao responderem às propostas de produção textual.

A análise proposta é ancorada em certa perspectiva discursiva, a ser apresentada na seção 4.1, sobre a relação entre fala e escrita que assume a tese de heterogeneidade constitutiva da escrita. Nessa seção, são discutidas abordagens sobre a relação entre fala e escrita de modo a delinear aproximações e distanciamentos teóricos nos quais esta tese se assenta.

Na seção seguinte, 4.2, as hipersegmentações selecionadas serão caracterizadas, tendo como eixo de análise a relação entre segmentação de palavras e organização prosódica dos enunciados a partir da consideração das configurações rítmica e entoacional que se constroem no texto produzido pelo aluno. Essa análise sustenta-se na consideração do texto em que as hipersegmentações ocorrem, bem como nas propostas de produção textual que os motivaram. Em última instância, as análises visam promover reflexões sobre possíveis relações entre prosódia e escrita, explicitando-se a complexidade linguístico-discursiva que subjaz a essas relações.

Defende-se, no desenvolvimento da argumentação, que as hipersegmentações (e as segmentações não-convencionais de palavra, de modo geral) não apenas indiciam vínculos entre enunciados orais/falados e letrados/escritos, mas também revelam como se dá a espacialização da linguagem, um processo de representação simbólica que é próprio da escrita. Com base nessa argumentação, defende-se a tese de que as hipersegmentações são locus de materialização da heterogeneidade da escrita.

4.1 Abordagens sobre a relação entre fala e escrita

Nesta tese, a relação entre prosódia e escrita é o tema central, tratado a partir da análise das segmentações não-convencionais de palavra, com destaque – no capítulo 3 – para as características linguísticas das hipersegmentações. A abordagem adotada dialoga com, ao menos, três abordagens distintas dadas à relação entre fala e escrita, a seguir explicitadas.

Inicialmente, faz-se necessário trazer à cena a afirmação de que “já não se podem observar satisfatoriamente as semelhanças e diferenças entre *fala* e *escrita* [...] sem considerar a distribuição de seus usos na vida cotidiana” (MARCUSCHI, 2001, p. 15). Nessa afirmação, identifica-se que a relação entre fala e escrita carece de ser abordada, segundo Marcuschi (2001), a partir de uma perspectiva que prioriza *os usos da língua* em relação aos aspectos formais da língua ou, ainda, aos aspectos neurocognitivos envolvidos nos processamentos da fala e da escrita, concebidas como manifestações da *faculdade da linguagem*. Trata-se de perspectiva teórica que assume que formas linguísticas, como as palavras, são mais bem analisadas a partir de usos da língua. Desse lugar teórico, fala e escrita são analisadas de perspectiva que não se circunscreve à consideração da língua enquanto sistema abstrato de regras, mas abrange usos da língua em diversas e complexas práticas de linguagem nas sociedades contemporâneas. É importante explicitar que esse viés que traça distinções entre abordagens da relação entre fala e escrita prioriza os usos da fala e da escrita que se realizam por meio de *conjuntos de práticas sociais orais e letradas*.

Esta tese comunga desse viés teórico na medida em que trata de características linguísticas (prosódicas e morfossintáticas) de enunciados escritos, considerando-se características de práticas orais e letradas, e se distancia de perspectivas que definem fala e escrita com base nas diferentes semioses que as constituem: a fala pelo meio sonoro e a escrita, gráfico. Essa abordagem remonta aos conceitos saussureanos de *fala e escrita como modalidades da língua* (isto é, modos de realização da língua), enquanto a *língua* constitui-se em um sistema de signos, formados pela união do sentido à imagem acústica (SAUSSURE, 1973, p. 23).

Apoiados nas bases semióticas que caracterizariam as modalidades linguísticas, estudiosos interessados na relação entre fala e escrita, como Olson (1977) e Ochs (1979) – dentre outros marcadamente durante a década de 1970 –, estabelecem dicotomias

estritas: por um lado, a fala é dependente de contexto, redundante, imprecisa, não-planejada, fragmentária, por outro lado, a escrita é autônoma em relação ao contexto, concisa, precisa, planejada e completa (MARCUSCHI, 2001, p. 27). Essa classificação tem duas consequências importantes: (i) atribuir à fala características negativas em relação às da escrita, cujas características são predominantemente valorizadas e (ii) associar à variação identificada na fala motivo para a deterioração da língua e à permanência no tempo, atributo da escrita, a função de conservar a língua, reduzindo, por vezes, a língua à escrita.

Essa perspectiva identifica, nas supostas características negativas (e exclusivas) da fala, motivos para ocorrer os chamados erros de escrita, como as segmentações não-convencionais de palavras analisadas. Portanto, não apenas haveria características distintas entre fala e escrita, mas também haveria uma relação assimétrica entre essas modalidades linguísticas. De modo geral, essa perspectiva que valoriza a escrita em relação à fala perdura nas escolas que têm, dentre outras, a função social de “ensinar a escrita” a crianças e jovens, promovendo desenvolvimento da “faculdade da linguagem”. E não apenas em instituições da área de Educação – como em Secretarias de Educação, além de escolas (públicas e privadas) –, mas também nas da área de Saúde – predominantemente no âmbito de clínicas fonoaudiológicas – é adotada essa perspectiva que assume haver dicotomias entre fala e escrita, como já apontado na seção 1.1.

Dessa perspectiva, tanto a fala quanto a escrita de crianças e adolescentes são avaliadas negativamente, pois os parâmetros dessa avaliação são frequentemente construídos com base em escrita idealizada sustentada a partir de conjunto restrito de obras literárias (e de alguns jornais impressos de grande circulação no país) como se verifica nos exemplos propagados em Bechara (2001), Cunha e Cintra (1989) e Faraco e Moura (1999). Em função de normas gramaticais e ortográficas compiladas nas gramáticas tradicionais, são identificados erros, os quais, por sua vez, constituem-se em índices de avaliação da (in)capacidade cognitiva dos alunos/pacientes em lidarem com diferenças entre fala e escrita, quando da produção de textos escritos. Em suma, dessa perspectiva, os chamados erros ortográficos (como as segmentações não-convencionais de palavra) decorrem da indesejada inferência da fala na escrita e são índices do (não) desenvolvimento (escolar e cognitivo) de crianças e adolescentes.

Crítico à essa concepção dicotômica e às consequências para o ensino fundamental que decorrem dessa concepção, Marcuschi (2001, p. 37) defende que “*as diferenças entre fala e escrita se dão dentro do continuum tipológico das práticas sociais*

de produção textual e não na relação dicotômica de dois polos opostos”. Nessa proposição, constata-se que a mobilização dos conceitos de *oralidade* e *letramento* enquanto práticas sociais rompe com a abordagem teórica predominante nos estudos sobre linguagem que tratam de fala e escrita a partir de bases semióticas (fônico *versus* gráfico) e, alternativamente, estabelece estratégias de organização textual-discursiva como bases para análise de relações entre fala e escrita. Essa proposta bebe das fontes da perspectiva *sociointeracionista* da linguagem que concebe *língua como fenômeno interativo e dinâmico*. Nessa perspectiva, fala e escrita igualmente apresentam propriedades dialógicas e dinâmicas, marcadas por usos estratégicos e por funções interacionais. Essas propriedades são estabelecidas a partir de postulados da Análise da Conversação e da Linguística Textual, principalmente, que se encontram divulgados especialmente em Koch (1992 e 1997), dentre outros.

De modo mais detalhado, Marcuschi (2001) lança mão do meio de produção – sonoro *versus* gráfico – e da concepção discursiva – oral *versus* escrita – para capturar a complexidade das relações entre as modalidades linguísticas (fala e escrita) e as práticas sociais (orais e letradas). No entrecruzamento entre meios de produção e concepção discursiva, o autor representa a interseção entre oralidade e escrita (cf. Gráfico 2 em Marcuschi, 2001, p. 39) e, em seguida, projeta a distribuição dos gêneros textuais em relação à fala e à escrita. A proposta estabelece um *contínuo de gêneros textuais* de modo a distinguir e correlacionar os textos das modalidades faladas e escritas em um contínuo de variações (textuais-discursivas, lexicais etc). Essa proposta é representada no gráfico, a seguir reproduzido. Por meio desse gráfico, visualiza-se, na interseção entre os eixos vertical e horizontal (destacados), como a relação entre fala e escrita é estabelecida em função dos gêneros textuais.

Gráfico 3. Representação do contínuo dos gêneros textuais na fala e na escrita.

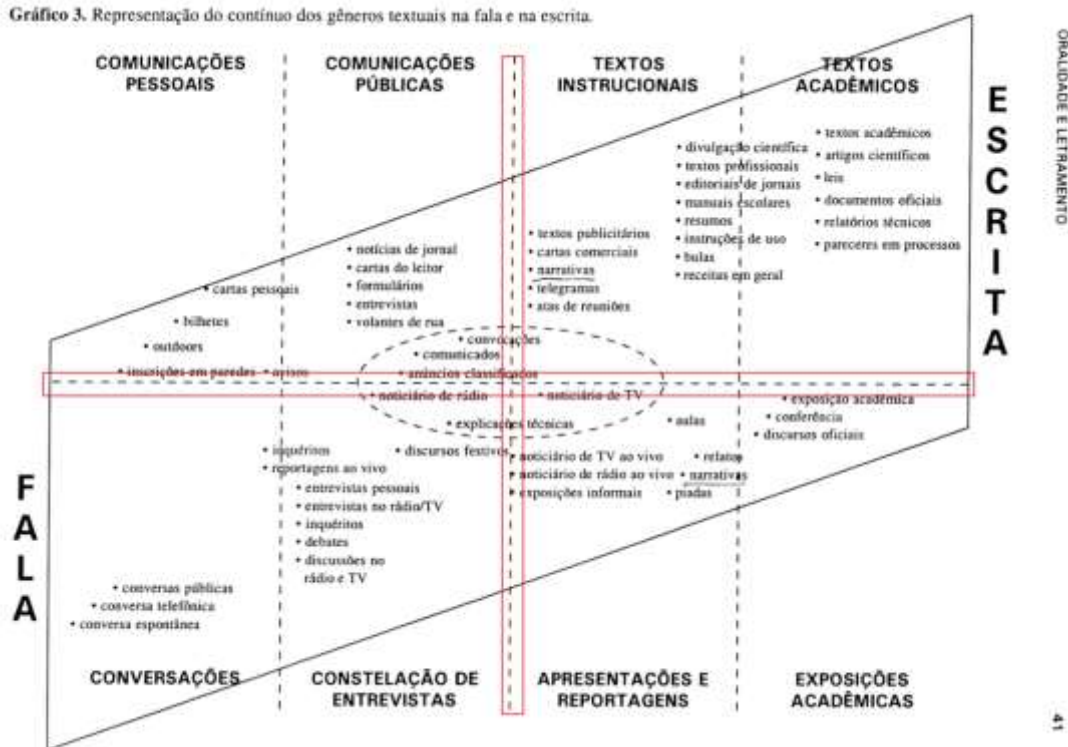


Figura 4.1: Representação do contínuo de gêneros textuais na fala e na escrita segundo Marcuschi (2001)

Cabe destacar que, no gráfico da Figura 4.1, não só se explicita como é concebido o contínuo de gêneros textuais, mas também se vislumbra como se dá o entrecruzamento dos gêneros em relação à distinção entre *meio de produção* (sonoro – gráfico) – representada pela linha horizontal – e *concepção discursiva* (oral – escrita) – representada pela linha vertical. Nessa proposta, verifica-se certa variação na proporção entre fala e escrita nos gêneros textuais, de modo que em cada um dos extremos do eixo horizontal constata-se haver predominância ou da fala ou da escrita. Em outras palavras, na representação proposta, existiriam gêneros, tanto falados quanto escritos, mais homogêneos, bem como gêneros mistos quanto às modalidades linguísticas (situados na porção central do contínuo, delimitados por um círculo no meio).

Após explicitar os fundamentos dessa representação, que ilustra as diferenças e semelhanças da relação entre fala e escrita, Marcuschi (2001) sustenta que “tanto a fala como a escrita apresentam um *continuum de variações*, ou seja, a fala varia e a escrita varia. Assim, a comparação deve tomar como critério básico de análise uma relação fundada no *continuum* dos gêneros textuais para evitar as dicotomias estritas” (MARCUSCHI, 2001, p. 42, grifos no original).

Apesar do incontestável ganho explicativo da proposta de Marcuschi (2001), quando comparada à tese da separação radical entre fala e escrita, partilhamos da crítica de Corrêa (2001, 2004), que identifica a recuperação da dicotomia entre fala e escrita, ao ser pressuposta certa homogeneidade nos extremos do contínuo. Junto com Corrêa (2001, 2004), concebo a relação entre fala e escrita e, também, entre oralidade e letramento fundamentada do caráter dialógico da linguagem, como proposto em Bakhtin (1992). A proposta de Corrêa (2001, 2004) acerca da relação fala e escrita é fundamental para que as segmentações não-convencionais de palavra, particularmente as hipersegmentações – objeto de análise –, sejam interpretadas não apenas enquanto dado de língua, mas também dado de discurso, como será demonstrado na próxima seção.

Corrêa (2004) ancora-se na proposta de Marcuschi (2001) que prima por apresentar a fala e a escrita como fenômenos de práticas sociais orais e letradas e argumenta:

Se os fenômenos de fala e escrita dados à observação podem ser vistos enquanto fatos linguísticos e enquanto práticas sociais, não se pode deixar de considerar a íntima relação entre um fato linguístico e uma prática social. [...] Da minha parte, assumo que os fatos linguísticos do falado/escrito são práticas sociais e estão ligados, portanto, às práticas orais/letradas. (CORRÊA, 2004, p. 2)

A visão de Corrêa (2004) rompe com as visões de relações dicotômica ou gradual entre fala e escrita – anteriormente apresentadas – e fundamenta-se na assunção de que “todo fato linguístico vincula-se a uma prática social” (p. 2). Segundo essa abordagem, nega-se a possibilidade de se opor fala e escrita por suas materialidades fônicas e gráficas e, alternativamente, defende-se que fala e escrita sejam constituídas pelas práticas de oralidade e letramento: o oral/falado e o letrado/escrito.

Ao assumir essa radical constituição entre fatos linguísticos e práticas sociais, o autor passa a traçar o percurso teórico que trilha para “questionar a delimitação do campo da escrita apenas pela constatação óbvia de um material específico – o gráfico – que lhe serve como base semiótica” (CORRÊA, 2004, p. 2) e propor o conceito do *modo heterogêneo de constituição da escrita* como “o encontro entre práticas sociais do oral/falado e do letrado/escrito, considerada a dialogia com o já falado/escrito e ouvido/lido” (CORRÊA, 2004, p. 9). Essa proposta está, pois, ancorada no princípio bakhtiniano de dialogicidade da linguagem, segundo o qual toda produção de linguagem é sempre resultado de um já dito/ouvido e escrito/lido. Nessa abordagem, *fala e escrita* são concebidos como *modos de enunciação* dos sujeitos da linguagem.

Sob essa conceituação, enunciação oral/falada e letrada/escrita nunca são homogêneas, mas essencialmente heterogêneas na medida em que o sujeito da linguagem é constituído por práticas orais e letradas com as quais sempre dialoga. A heterogeneidade não é concebida como propriedade da língua ou da linguagem que se faz presente **na** escrita, mas como traço **da** escrita, ou seja, a “heterogeneidade como constitutiva da escrita e não como uma característica pontual e acessória desta. Noutros termos: *a heterogeneidade é [...] interior à escrita e não exterior a ela*”. (CORRÊA, 2001, p. 144, destaques acrescentados).

Os argumentos a favor dessa tese são construídos, por Corrêa (2001), a partir de três eixos de observação da *heterogeneidade da escrita*, a saber: a representação da (suposta) gênese da escrita; a representação do código institucionalizado; a dialogia com o já falado/escrito e ouvido/lido. Em Silva e Tenani (2014, p. 89), esses eixos foram considerados na análise de hipersegmentações de modo que foi demonstrado que

os usos não-convencionais das fronteiras de palavras são *registros* da heterogeneidade da escrita, o que, em outras palavras, equivale a dizer que essas grafias são fruto da imagem que os escreventes têm do que seja a constituição da (sua) escrita, a partir de suas inserções em práticas orais/faladas e letradas/escritas de uso da linguagem.

Esta tese partilha dessa perspectiva de análise e, como fizeram as autoras naquele estudo, apresenta hipóteses explicativas para as hipersegmentações que se distanciaram de interpretações que as classificam como erros ortográficos (por exemplo, KOCH, 1997) ou como indícios de distúrbio específico de aprendizagem da escrita (por exemplo, ZORZI, 1997, 2006). Partilha, também, da tese de que hipersegmentações sejam efeitos dos encontros entre o oral/falado e o letrado/escrito. Mas diferentemente de Silva e Tenani (2014), porém, esta tese explora, por meio da análise das hipersegmentações, a proposta de Corrêa (2001, p. 5) de identificar “encontros entre as práticas [orais/faladas e letradas/escritas] a começar pelos seus materiais significantes, como o gesto articulatório e o gesto traduzido graficamente em ritmo da escrita”.

A fundamentação teórica dessa proposição é construída por Corrêa (2004) a partir da afirmação de Verón (1981, p 78-79) a propósito das características semióticas dos discursos, a seguir reproduzida:

Os discursos sociais são objetos semioticamente heterogêneos ou “mixtos”, nos quais intervêm, ao mesmo tempo, várias matérias significantes e vários códigos. O próprio discurso linguístico não é nunca monocódigo: quer se trate da escrita ou do discurso falado, há sempre regras paralinguísticas que não podem ser reduzidas apenas ao código da “língua”.

Dessa afirmação, Corrêa (2001) explora a ideia de heterogeneidade quanto à base semiótica da escrita ao considerar usos de demonstrativos, como “esses”, em que o referente está fora do texto de vestibulandos, por serem elementos dêiticos que apontam para (outros) textos escritos e para relações espaciais que se estabelecem com o que está fora da língua. Nesses casos, o escrevente constrói “dimensões conversacionais e argumentativas, bem como elementos da situação imediata de enunciação, por meio de *gestos que supõe plasmados nas projeções espaciais*, linguisticamente marcadas em seu texto” (grifos acrescentados). Nessa interpretação de Corrêa (2004), identifica-se que o gesto de apontar que traz uma dimensão visual da oralidade ganha, na projeção espacial instaurada no texto escrito, uma dimensão verbo-visual, o que resulta – de certa perspectiva – em marca de oralidade no texto escrito, ou – da perspectiva de Corrêa (2004) – em marca da heterogeneidade **da** escrita.

O autor avança em sua argumentação trazendo à baila considerações sobre o “desenvolvimento pré-histórico da escrita infantil” de Vygotsky (1988) e sobre a “escrita ritmicamente reprodutiva” de Luria (1988). Da interpretação de Corrêa (2004) acerca dessas noções para análise da escrita infantil, interessa destacar que a escrita ritmicamente reprodutiva se mostra sob dependência do plano de expressão. De modo mais preciso, o autor traz a proposta de Abaurre (1991b), desenvolvida a partir do conceito de “escrita ritmicamente reprodutiva” de Luria (1988), sobre o ritmo como um traço fundamental da escrita. A noção principal é de haver “gesto rítmico” congelado no signo gráfico da escrita. Nessa proposta de Abaurre (1991b), Corrêa (2001, p. 7) enfatiza que haveria três materiais significantes que atuariam na escrita: o gesto, o material fônico-acústico (ritmo) e o material gráfico.

Quanto à noção de *gesto*, cabe observar que são duas as noções em jogo: *os gestos articulatórios*, referente tanto a uma *ação de constrição* dos diferentes articuladores quanto à *representação* dessa manobra articulatória necessária para que essa ação se concretize, os quais são percebidos pelo aparelho auditivo (portanto, há produção e percepção de estímulos acústico-articulatórios), e *os gestos gráficos*, executados pela mão e percebidos pela visão (portanto, há produção e percepção de marcas gráfico-visuais). São, portanto, gestos que se realizam em bases semióticas distintas, porque estão mobilizadas materialidades distintas: a fônica e a gráfica. Estariam aí as bases semióticas que sustentariam a dicotomia entre fala e escrita, perspectiva anteriormente apresentada.

Porém, o conceito de “gesto rítmico” engloba ambos os gestos e, desse modo, captura a heterogeneidade da escrita quanto à materialidade que lhe constitui.

Ainda sobre os gestos articulatórios, faz-se relevante considerar que não mais têm sido concebidos como sendo de natureza exclusivamente fônica (MAN et al., 2012 dentre vários outros). Experimentos, relativamente recentes, sobre processamento perceptual da fala têm demonstrado que os gestos articulatórios são de natureza fônica e também visual e que a *fala é um evento auditivo, proprioceptivo e visual integrado* (SEKIYAMA; BURNHAM, 2004). Mais especificamente, experimentos feitos com falantes de diferentes línguas e de faixas etárias (HUYSE; LEYBAERT, 2014; DESJARDINS; ROGERS; WERKER, 1997) apresentam como resultados relevantes que a percepção de segmentos (inclusive os produzidos pelo próprio corpo, daí o traço proprioceptivo¹¹⁴ da fala) compreende processamento auditivo (de estímulos acústicos) e visual (de gestos articulatórios). Portanto, tanto na fala, quanto na escrita está presente uma dimensão visual da linguagem. A consequência dessa constatação, de nossa perspectiva, é assumir a *constituição heterogênea da materialidade da fala*, na medida em que dimensão fônica (acústico-articulatória), proprioceptiva e visual (articulatória) lhe constitui.

Por outro lado, os gestos gráficos também são impregnados da dimensão fônica da linguagem. Chacon (1998, p. 84), ao comentar sobre as contribuições de Abaurre (1991a) e Corrêa (1994) sobre relações entre ritmo da oralidade e da escrita, explicita que “o ritmo que se depreende de uma base articulatória-acústica na oralidade desenvolve-se, na escrita, sob uma base gráfico-visual”. Enquanto na oralidade, o ritmo linguístico é de ordem *temporal*, o ritmo da escrita é de ordem *espacial*. O “gesto rítmico” capturaria essa dimensão temporal dos enunciados orais/falados ao representá-lo na espacialidade da escrita. Portanto, também se faz relevante assumir a *constituição heterogênea da materialidade da escrita*, na medida em que dimensão fônica (acústico-articulatória) e visual (gráfico) lhe constitui.

Quanto à relação entre tempo (da fala) e espaço (da escrita), encontra-se desenvolvimento detalhado em Chacon (1998), com base na análise de usos de sinais de pontuação em textos de vestibulandos. Das considerações feitas por Chacon (1998, p. 199-200), interessa destacar que a pontuação indicia (i) vínculos entre escrita e oralidade, “na medida em que revela tentativas de reprodução da língua falada e na medida em que

¹¹⁴ O termo *proprioceptivo* (do inglês, “proprioceptive”) diz respeito à capacidade sensorial de receber estímulos originados do próprio corpo. Por meio dessa capacidade, localiza-se posição espacial e identificam-se movimentos de um corpo e seus membros, assim como a tensão muscular desses.

funciona como um recurso de interpretação do texto escrito”; (ii) constituição da escrita como um código particular de expressão verbal, “já que revela, no processo de representação simbólica própria da escrita, a espacialização da linguagem”, e (iii) construção do próprio estatuto simbólico da escrita, “uma vez que denuncia a transcodificação que, na atividade expressiva, as dimensões fonológica, sintática e textual operam de características da oralidade”.

Com base nessa argumentação de Chacon (1998) sobre o papel da pontuação na organização do ritmo da escrita, demonstro, na próxima seção, que a segmentação não-convencional de palavra não apenas indicia vínculos entre enunciados orais/falados e letrados/escritos, mas também revela como se dá a espacialização da linguagem, um processo de representação simbólica que é próprio da escrita. Ainda se defenderá que – assim como os sinais de pontuação – as segmentações não-convencionais de palavra, especialmente as hipersegmentações, são locus de materialização da heterogeneidade da escrita.

Em suma, esta tese assume a abordagem da escrita e, também, da relação entre fala e escrita elaborada por Corrêa (2001, 2004) por essa abordagem proporcionar lentes que ampliam a análise das segmentações não-convencionais de palavra – e particularmente as hipersegmentações – de modo a identificar como a constituição heterogênea do letrado/escrito atravessam os escreventes, como será demonstrado na próxima seção.

Por fim, faz-se relevante tratar da articulação teórica entre essa abordagem discursiva da relação entre fala e escrita e a abordagem formal dada à organização prosódica dos enunciados, explicitada no capítulo anterior, na seção 3.1 e que também será mobilizada na análise de hipersegmentações na próxima seção.

A abordagem da Fonologia Prosódica compreende, como visto na seção 3.1, a assunção de que há uma hierarquia de constituintes prosódicos que organizam os enunciados, o que implica afirmar que é a partir desses constituintes que se implementam processos segmentais, configurações de acentos e de tons em uma língua dada. Em outras palavras, nessa abordagem fonológica da prosódia, as realizações fonéticas – de natureza segmental, rítmica e entoacional – são interpretadas como evidências de estruturas que se estabelecem na interface entre fonologia e demais componentes da língua, privilegiadamente, a sintaxe. Trata-se, pois, de estruturas linguísticas que constituem os enunciados sejam eles orais/falados ou letrados/escritos. Esta tese se vale dessa abordagem formal da prosódia por proporcionar meios de explicitar a complexidade da

chamada, por Chacon (1998), dimensão fônica da linguagem e, a partir dessa complexidade linguística explicitada, argumentar como se estabelecem relações entre prosódia e hipersegmentações de palavras e, mais amplamente, relações entre fala e escrita.

4.2 Hipersegmentação e relações textuais-discursivas

Nesta seção, serão feitas análises qualitativas de segmentações não-convencionais de palavras com o objetivo de argumentar que essas segmentações também são efeito do “conflito” sofrido pelo escrevente na relação com o que lhe é solicitado pela instituição no processo de produção textual. Em outras palavras, argumenta-se que grafias não-convencionais de palavras não são apenas *dado de língua*, como demonstrado no capítulo anterior, mas especialmente *dado de discurso*.

Também será dada ênfase à explicitação de modos de entrelaçamento entre práticas orais/faladas e letradas/escritas que se mostram nos textos escritos por meio da análise das hipersegmentações. Para realizar essa análise de cunho linguístico-discursivo, serão considerados dois textos do Banco de Dados de Escrita do EF II (TENANI, 2015), selecionados com base no critério de haver hipersegmentações cujas características as particularizam em relação àquelas descritas no capítulo anterior.

No primeiro texto a ser analisado (Figura 4.2), ocorre a hipersegmentação “su ceço”, que se particulariza em relação às demais hipersegmentações descritas no capítulo 3, pois a sílaba “su” não corresponde a nenhum monossílabo átono no português, embora siga a tendência quanto à localização não-convencional do branco em trissílabos, de modo que o resultado é a sequência de sílaba seguida de pé troqueu.

No segundo texto selecionado (Figura 4.4, mais à frente), ocorrem as hipersegmentações “com ver sando” e “com ver samos” que se particularizam por apresentar duas fronteiras não-convencionais dentro da palavra (são duas sílabas pretônicas hipersegmentadas por espaços em branco), enquanto a quase totalidade das hipersegmentações no EF II têm apenas uma fronteira não-convencional (cf. descrição de dados na seção 3.2). Essas grafias particulares seguem, no entanto, a característica geral, descrita no capítulo anterior, de as sílabas hipersegmentadas por espaços em branco corresponderem a monossílabos da língua.

portunidade” (“oportunidade”), “su seço” (“sucesso”). Já as hipossegmentações têm em comum a junção de um monossílabo átono (“me, que, em”) a uma ω (“levar, eu, frente”): “melevarar” (“me levar”), “queu” (“que eu”)¹¹⁶ e “enfrente” (“em frente”). Essa característica das hipossegmentações é recorrente em textos de alunos do EF II, como demonstrado por Tenani (2011b), quando analisou textos selecionados da amostra transversal do Banco de Dados de Escrita do EF II, e por Tenani e Fiel (2016), quando analisaram parte da amostra longitudinal do mesmo banco.

Quanto às características das hipersegmentações, foi demonstrado na seção 3.2 – a partir de dados longitudinais – e também foi descrito por Tenani (2011b) – a partir de dados transversais –, que há tendência de a primeira sílaba da palavra hipersegmentada por branco corresponder a um possível monossílabo átono em PB (que corresponde, na maioria dos casos, à palavra gramatical, notadamente, preposição, conjunção, artigo), o que resulta na sequência de sílaba seguida de pé troqueu, como em: (a) σ (gora) Σ . Nesses casos, informação de natureza morfossintática (relacionada ao comportamento dos monossílabos átonos) e prosódica (relacionada ao comportamento de elementos clíticos) se somam às informações de natureza letrada/escrita, notadamente aquelas relacionadas a práticas predominantemente desenvolvidas em ambiente escolar, tais como orientações de emprego de palavras gramaticais (como preposição, conjunção, artigo) e de reconhecimento das fronteiras gráficas entre palavras (por meio da identificação da localização de branco e do hífen na palavra).

Na grafia de “su seço”, salta aos olhos que a organização métrica da palavra parece ter motivado, mais fortemente em relação a fatores morfossintáticos ou semânticos, a segmentação não-convencional do trissílabo, a saber: (su) σ (seço) Σ . Nota-se, porém, que outras duas grafias da mesma palavra não foram hipersegmentadas: “susceso” e “suceso”.¹¹⁷ Essa flutuação quanto ao registro hipersegmentado da mesma palavra

¹¹⁶ A grafia “queu”, diferentemente das demais hipossegmentações, é registro de um processo de sândi externo entre “que” [ki] e “eu” [eu]. Nesse sândi registrado, há choque entre núcleos silábicos que leva à queda da vogal átona de [ki], seguida da ressilabação da consoante [k] e [eu], que resulta em [keu], uma elisão.

¹¹⁷ Vale observar que a grafia “su seço” traz, ainda, evidências do trânsito do escrevente por práticas letradas quando considerados os usos não-convencionais dos grafemas “S” e “Ç” empregados na porção hipersegmentada “seço”. Ao ser comparada essa grafia com a convencional “sucesso”, verifica-se que os grafemas “S” (“su”) e “C” (“ce”) e o dígrafo “SS” (“sso”) representam /s/. É a convenção ortográfica que define que grafema ou dígrafo deve ser empregado em detrimento de outras possibilidades. As grafias não-convencionais de “sucesso” são exemplos de que outros grafemas podem representar /s/, segundo a ortografia do português. O grafema “Ç” empregado em “su seço” e o dígrafo “SC” empregado em “susceso” também podem representar /s/, como é o caso das grafias convencionais “maçã” e “piscina”, respectivamente. Ainda o registro de “S” em início de palavra em todas as grafias em análise revela o trânsito pelas convenções ortográficas na medida em que somente “S” é empregado quando a vogal da

enfraquece a hipótese de a estrutura métrica da palavra ser exclusivamente o fator que leva à hipersegmentação no texto de Daniel.

Essa flutuação quanto à segmentação possibilita enveredar pela hipótese de que é relevante considerar características textual-discursivas dos enunciados para explicar porque apenas uma das grafias de “sucesso” é hipersegmentada. Cabe enfatizar que considerar a dimensão textual-discursiva não implica descartar a dimensão fônica, pois ambas as dimensões constituem a organização rítmica da escrita (CHACON, 1998). Portanto, as segmentações não-convencionais de palavra, como também a ausência da segmentação da mesma palavra, seriam efeitos dessa organização rítmica da escrita. Argumentos que sustentam essa interpretação serão construídos a partir da consideração da proposta textual, a seguir reproduzida.

Figura 4.3: Enunciado da proposta “Uma grande conquista”

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

Projeto de extensão: Oficinas de Leitura, Interpretação e Produção Textual – Coordenação Luciani Tenani

E.E.
Nome: _____ Série/Turma: 7ª _____
Proposta 3 Data: _____

- Observe a foto e trechos de notícias sobre a vitória de César Cielo nos Jogos Olímpicos de Pequim e discuta com seus colegas sobre a emoção do nadador ao ganhar a medalha de ouro nas olimpíadas.



Medalhista de ouro na prova dos 50 m livre dos Jogos Olímpicos, o nadador brasileiro César Cielo não conseguiu conter a euforia pela inédita conquista da modalidade no país –primeiro ouro na história– e disse ter vivido o melhor momento de sua vida em Pequim. (fonte: <http://olimpiadas.uol.com.br/2008>)

César Cielo não conseguiu segurar as lágrimas ao conquistar o ouro histórico nos 50 m livre, o primeiro do país em Pequim-2008 e o primeiro para o país na modalidade. (fonte: <http://olimpiadas.uol.com.br/2008>)

- Escreva um texto narrativo sobre **uma grande conquista**, alguma meta que sonha em alcançar no seu futuro, que dificuldades teria que superar para alcançar seu objetivo e qual seria sua emoção ao realizar o seu sonho.
- Seu texto deve conter de 15 a 20 linhas e deve ser escrito à tinta. Seu texto não deve ultrapassar os limites designados para a escritura.

Fonte: Banco de Dados de Escrita do Ensino Fundamental II (TENANI, 2015)

sílabas for “A, O, U” (como em “saco”, “soco”, “suco”), embora mais de uma letra represente /s/ quando a vogal da sílaba em início em palavra for “I, E” (como em “sítio” e “cidade”; “seco” e “cera”). É uma regra de contexto que estabelece que “S” representa /z/ quando entre vogais, como em “asa”. É a convenção ortográfica que define o emprego do dígrafo “SS” na última sílaba de “sucesso”. Portanto, as grafias não-convencionais ora descritas compreendem “dúvidas ortográficas” quanto ao registro convencional de /s/.

A proposta que motivou a produção do texto em análise pode ser caracterizada como constituída de quatro partes. Na primeira, apresenta-se a temática “a vitória de César Cielo nos jogos olímpicos de Pequim”, em 2008, e propõe-se a discussão em grupo sobre a emoção da vitória do nadador. Na segunda parte, listam-se foto do referido nadador beijando a medalha de ouro conquistada e dois trechos de uma matéria jornalística, publicada no portal de notícias UOL, sobre a conquista inédita do ouro olímpico na prova de 50m livre de natação. Na terceira parte, são dadas orientações quanto ao gênero e ao tema para produção textual. Na última parte, as orientações se circunscrevem à extensão e formatação do texto na folha de papel. A diagramação da proposta de produção textual favorece sua identificação visual (cf. Figura 4.3) e o destaque dado, por meio de negrito, a “uma grande conquista” delimita também visualmente o tema.

Inicialmente, cabe destacar a relação que o aluno Daniel estabelece com a proposta de produção textual. Apesar de César Cielo ser dado como modelo de *sucesso* e a conquista de medalha de ouro olímpica como exemplo de uma *grande conquista*, o aluno recusa essa sugestão e busca afirmar-se a partir de si mesmo. Há um diálogo claramente conflitante entre o que é proposto como modelo de sucesso e o que é, de fato, tomado como modelo pelo escrevente. Rastros desse conflito são visíveis a partir do título “Ser eu mesmo” que anuncia uma autoafirmação a qual é reiteradamente enunciada, como se verifica em: “Meu sonho é ser eu mesmo do jeito que eu sou”. Essa afirmação de seu modo de ser implica assumir, por exemplo, (sua) identidade linguística, ou seja, (seu) modo de falar, que se materializa inclusive nas representações gráficas não-convencionais de várias palavras, como é exemplar “o bistracos”.

A grafia “o bistracos” é registro de vários processos segmentais que se aplicam a “obstáculos”. A proparoxítona “obstáculos” está sujeita ao apagamento da vogal postônica não final por meio do qual se torna paroxítona (processo atestado na variedade falada pelo aluno, conforme Ramos e Tenani, 2009). No exemplo, a queda da vogal não final “u” geraria a forma “obstáculos”. Esse apagamento da vogal tem como condição motivadora uma configuração métrica marcada, isto é, um pé dátilo, e como resultado a configuração do pé troqueu dissílabo, estrutura métrica preferencial para palavras da língua. Também ocorre metátese, processo por meio do qual a consoante lateral da sílaba “clos” (que resultou do processo de queda da vogal “u”) migra para a sílaba acentuada que lhe é precedente, resultando em “tlacos”, que, por sua vez, passa a ter a sílaba

acentuada com duas consoantes em seu início (o que configura uma sílaba relativamente mais complexa com ataque ramificado). Nessa estrutura, outro processo segmental se verifica, pois a lateral se realiza como um tepe (resultado do processo de rotacismo), gerando, desta feita, uma sílaba constituída de uma sequência consonantal /tr/ (a letra “r” representa o tepe) que é mais produtiva do que /tl/ em português. Todos esses processos segmentais resultam na sequência “tracos”, presente em “o bistracos”. Por fim, há registro do processo de epêntese vocálica (cf. COLLISCHONN, 2000, 2004) na sílaba “bis”, que é desencadeado para resolver a silabação da cadeia consonantal [bs]. Nos enunciados falados no PB, tipicamente a vogal epentética se realiza como [ɪ], resultando, nesse caso, a sílaba [bis], grafada não-convencionalmente como “bis”.

Dessa análise, cabe destacar que a grafia “o bistracos” consiste, de modo mais evidente, em representação de características fonético-fonológicas de enunciados falados por sujeitos de baixa escolaridade (especificamente, os processos de rotacismo e queda da vogal postônica não final), carregando, assim, estigma de desprestígio social. Configura-se, por meio da enunciação escrita, a projeção de um sujeito escrevente/falante que assume sua identidade (“ser do jeito que eu sou”) e sua condição: está à espera de uma oportunidade que lhe será dada pela vida ou por alguém na sociedade (“na oportunidade que poderei ter que auguem vai me compre ender”).

Diferentemente do nadador Cielo que figura como exemplo de dedicação pessoal para ganhar o ouro olímpico, um sonho pessoal e da nação, o escrevente/falante se mostra subjugado ao imponderável da vida; é como um barco a remo (sem muitos recursos) à deriva no mar que aguarda o sopro de bons ventos que o levariam a um porto, à terra firme, ou seja: a passagem para um local seguro (um “lugar” de sucesso na sociedade, por exemplo) depende de oportunidade.

O texto produzido pelo aluno pode ser tomado como unidade de uso da língua e as grafias não-convencionais como registros de características dos enunciados falados e escritos por esse mesmo sujeito. Nesses registros, observa-se o trânsito entre língua e discurso. O texto em questão materializa o discurso mobilizado predominantemente por aqueles que, mesmo estando desprovidos do amparo social, aguardam oportunidade para chegar ao sucesso pessoal (nas palavras do aluno: “do jeito que a vida me levar para oportunidade que poderá me levar ao sucesso que eu um dia eu poderei conquistar”). Digno de nota é que esse enunciado ecoa versos do pagode “Deixa a vida me levar”,¹¹⁸

¹¹⁸ A canção “Deixa a vida me levar” encontra-se disponível *on-line* gratuitamente em vários portais. Para a interpretação ora apresentada, faz-se pertinente trazer a primeira estrofe: “ Eu já passei / Por quase tudo

de Zeca Pagodinho, lançado em 2002, que fez muito sucesso também na época em que o texto em análise foi produzido. Nessa canção, o refrão é “Deixa a vida me levar / Vida leva eu!”, por meio do qual se constata o uso não-convencional do pronome “eu”, quando previsto o pronome “me” em “leva eu”. Ou seja, também na canção, com a qual há um potencial diálogo, é traçado o perfil do “pobre” (que enfrenta obstáculos) que espera sua vez para alcançar sucesso. O enunciador assume sua condição social, evidenciada por meio de usos de estruturas linguísticas que o identificam e com as quais se identifica socialmente.

Possivelmente, no texto da Figura 4.2, o sujeito escrevente traça mesmo percurso: inscreve-se na língua para se autoafirmar, pois é na língua que se constitui e por ela é atravessado. Do mesmo modo que o sujeito está inscrito nas representações escritas tanto no texto do aluno, quanto no texto do pagode, o discurso de como alcançar o sucesso, mencionado na produção textual do aluno, é o mesmo que se faz ouvir na canção “Deixa a vida me levar”. Em ambos, a ênfase está na *oportunidade* que leva ao *sucesso*, apesar de haver *obstáculos* no percurso. Não por acaso, “oportunidade”, “obstáculos” e “sucesso” são hipersegmentadas no texto analisado!

Há motivação textual-discursiva para as hipersegmentações, ou seja: os espaços entre as sílabas materializam a ênfase prosódica dada a palavras que são destacadas em função da organização textual-discursiva dos enunciados. É essa motivação que também leva a apenas uma hipersegmentação de “sucesso”, pois é no enunciado “o su seço que eu vou fazer” que o sujeito escrevente afirma, pela primeira vez no texto, que seu desejo de “sucesso” acontecerá. Diferentemente, no enunciado em que se verifica a primeira ocorrência de “sucesso” – a saber: “... para a o portunidade que poderá me levar a susceso queu um dia eu poderei conquistar...” –, observa-se a explicitação do desejo de encontrar sucesso e a ênfase na oportunidade que levará a esse sucesso: “a o portunidade” (grafia hipersegmentada) levará ao “susceso” (grafia não-segmentada). A comparação entre essas grafias nesses enunciados permite afirmar que a dimensão textual-discursiva dos enunciados motiva ocorrências de espaços não-convencionais dentro de palavras.

Vale verificar que, nas quatro hipersegmentações de polissílabos (“o portunidade”, “or portunidade” (dois registros) e “o bistracos”), a sílaba inicial da palavra

nessa vida / Em matéria de guarida / Espero ainda a minha vez / Confesso que sou / De origem pobre / Mas meu coração é nobre / Foi assim que Deus me fez...”.

(grafada “o” e “or”)¹¹⁹ fica isolada em relação à porção que resta da palavra, o que configura, também, destaque verbo-visual. Esse registro também é efeito da organização rítmica das palavras. Em termos de pés métricos, os polissílabos em questão podem ser escandidos em troqueus dissílabos da seguinte forma: (opor)Σ (tuni)Σ (dade)Σ e (obis)Σ (trácos)Σ. Portanto, a sílaba inicial de cada uma dessas palavras é portadora de acento rítmico (informação verbal), carregando, assim, saliência fônica a qual possivelmente tenha favorecido a delimitação por espaços em branco. Soma-se, a essa configuração métrica, informação morfossintática que favorece essas grafias: a sílaba pretônica “o” corresponde ao artigo definido em português e, como tal, deve ser grafado entre espaços em branco (portanto, informação verbo-visual). Identificam-se, nessas hipersegmentações, pistas de que saliências prosódicas das sílabas pretônicas de palavras e informações morfossintáticas de elementos clíticos do português, as quais estão associadas a registros escritos convencionais, também contribuem para haver registros não-convencionais no texto de Daniel.

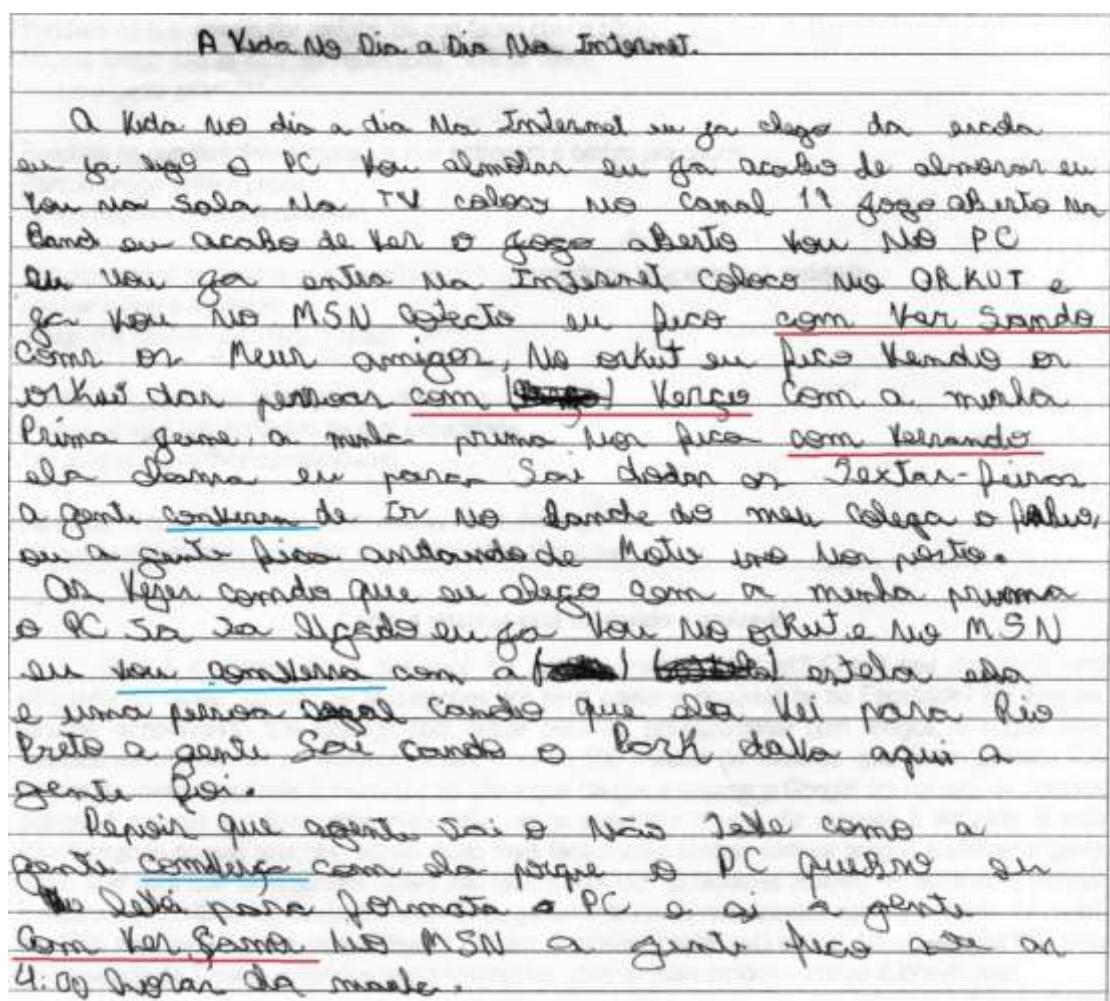
Ainda vale enfatizar a relação entre aspectos verbo-visuais da escrita e aspectos prosódicos da língua. A colocação do espaço em branco dentro da palavra leva à delimitação de porções da palavra, gerando o efeito de destaque ao que fica graficamente isolado (cf. análise de “demais” e “senão” na seção 3.3). Para essa interpretação de destaque verbo-visual também contribui o efeito da ênfase prosódica que potencialmente se verifica nos enunciados falados. Por meio da ênfase prosódica (cf. FERNANDES, 2007), parte da palavra (ou do enunciado) é destacada do que lhe é adjacente, como argumentado na seção 3.3 no capítulo precedente. Portanto, as hipersegmentações das sílabas pretônicas de “sucesso”, “oportunidade” e “obstáculos” são, também, efeito dessa ênfase prosódica – plasmada na escrita – cuja força propulsora se encontra na construção de sentidos mobilizados (e enfatizados) pelos discursos que atravessam e constituem o texto.

A descrição e análise ora realizadas do texto da Figura 4.2 – além da descrição e análise das hipersegmentações de “demais” e “senão” feitas na seção 3.4, no capítulo anterior – favorecem a conclusão de que hipersegmentações de palavras são registros escritos de ênfase (prosódica e discursiva) de (porções de) enunciados. No entanto, essa relação entre hipersegmentação e ênfase não se mostra evidente no segundo texto

¹¹⁹ Nas grafias de “oportunidade”, o registro não-convencional de “R” na primeira sílaba é, potencialmente, decorrente de antecipação do “R” da sílaba “por”, processo fonético-fonológico presente na variedade do português do interior paulista à qual pertence o aluno que produziu o texto.

selecionado, reproduzido na Figura 4.4. Como já anunciado, nesse texto, ocorrem as hipersegmentações “com ver sando” e “com ver samos” que se particularizam por apresentar duas fronteiras não-convencionais dentro da palavra. Por meio da análise dessas e das demais formas flexionadas do verbo “conversar” (hipersegmentadas ou não), será argumentado que espaços em branco grafados não-convencionalmente dentro de palavras tornam visíveis a organização prosódica dos enunciados (e não apenas a organização métrica da palavra), o que implica mobilizar configuração rítmica e entoacional dos enunciados que se constituem a partir de domínios prosódicos hierarquicamente organizados (cf. hierarquia dos constituintes prosódicos na seção 3.1). Argumento, ainda, que essa relação entre escrita e prosódia é constituída pelas dimensões fônica, morfossintática e textual-discursiva que configuram o texto (a língua).

Figura 4.4: Texto em que ocorrem “com ver sando”, “com versando”, “com ver samo”



Fonte: Banco de dados de escrita do Ensino Fundamental II. Texto: Texto: Z11_8B_13M_01 ¹²⁰

¹²⁰ Leitura possível: “A vida no dia a dia na internet. Eu já chego da escola, eu já ligo o PC. Vou almosar. Eu já acabo de almosar, eu vou na sala. Na TV, eu coloco no canal 19 Jogo aberto na Band. Eu acabo de

Inicialmente, faz-se relevante caracterizar prosodicamente as hipersegmentações das formas verbais de “conversar” e as grafias convencionais desse mesmo verbo, o que totaliza sete grafias, a saber: “com ver sando, com verço, com ver sando, conversa, comversa(r), comverça, com ver samo”.¹²¹

Para serem analisadas essas grafias, lança-se mão da configuração prosódica dos enunciados em que ocorrem. Para cada um dos sete enunciados, listados a seguir, são identificadas fronteiras de *I* e de ϕ , por esses domínios serem constituídos a partir de relações morfossintáticas que se configuram entre as palavras no enunciado em que cada forma verbal ocorre.¹²² A questão específica a ser respondida nessa etapa da análise é: como se constitui a relação entre estrutura prosódica e registros convencionais e não-convencionais de formas flexionadas de “conversar” no texto?

- 4.1. [[eu] ϕ [fico com ver sando] ϕ [coms os meus amigos] ϕ]I
- 4.2. [[com (rasura) verço] ϕ [com a minha prima Jeine] ϕ]I
- 4.3. [[nós] ϕ [fica com versando] ϕ]I
- 4.4. [[a gente] ϕ [conversa] ϕ [de ir] ϕ [no lanche] ϕ [do meu colega] ϕ [o Pablo] ϕ]I
- 4.5. [no MSN] ϕ [eu] ϕ [vou comversa] ϕ [com a Estela] ϕ]I
- 4.6. [depois que] ϕ [agente] [saiu] ϕ]I [[não teve] [como a gente] ϕ [comverça] ϕ [com ela] ϕ]I]
- 4.7. [[e ai] [a gente] ϕ [com ver çamo] ϕ [no MSN] ϕ]I

ver o Jogo aberto vou no PC. Eu ja vou entra(r) na Internet. Coloco no orkut e ja vou no MSN. Conecto. Eu fico com ver sando coms os meus amigos. No Orkut, eu fico vendo os orkut das pessoas com (rasura) verço com a minha prima Jeine, a minha prima. Nós fica com versando. Ela chama eu para sair dodas [todas] as sextas-feiras. A gente conversa de ir no lanche do meu colega, o Pablo, ou a gente fica andando de moto ino [indo] nos posto. Às vezes, condo [quando] que eu chego com a minha prima, o PC ja ta ligado. Eu ja vou no orkut e no MSN. Eu vou comversa [conversar] com a [rasura] Estela. Ela é uma pessoa legal. Cando [quando] que ela vei [vem] para Rio Preto, a gente sai. Cando [quando] o Park [parque] dava [tava], aqui a gente foi. Depois que agente sai o [saiu], não teve como a gente comverça com ela porque o PC quebro. Eu levei para formata o PC e ai a gente com ver çamo no MSN. A gente fico até às 04:00 horas da manhã”.

¹²¹ Na notação, será mantida a grafia não-convencional dos dados, a saber: emprego de “M” quando previsto “N” na sílaba pretônica; ausência de “R” na forma verbal no infinitivo; ausência de “S” na forma flexionada em primeira pessoa do plural.

¹²² A atribuição dessas fronteiras foi feita a partir dos algoritmos de formação de cada um dos domínios prosódicos identificados, tendo em conta as descrições feitas (por Serra, 2009, Fernandes, 2007 e Tenani, 2002) para o PB, como descrito anteriormente na seção 3.1.

Da comparação entre as estruturas (4.1) a (4.7), constata-se que todas as formas verbais que não foram hipersegmentadas ocorrem em ϕ_s intermediários, isto é, ϕ_s que não estão na borda direita ou esquerda de *I* (4.4, 4.5 e 4.6) e as hipersegmentações ocorrem em: (i) ϕ_s intermediários, como (4.1) e (4.7); (ii) ϕ em início de *I*, como (4.2); ϕ em final de *I*, como (4.3). Consideradas essas estruturas prosódicas e as formas hipersegmentadas em (4.1), (4.2), (4.3) e (4.7), verifica-se não haver correlação entre hipersegmentação e posição de ϕ dentro de *I*.¹²³

Avança-se na análise, identificando-se, nos sete enunciados selecionados, palavras prosódicas (ω) e configuração entoacional, definida em termos de tons (T), que são atribuídos em função da constituição de ϕ_s (e de sua localização dentro de *I*). A identificação de tons que potencialmente estariam associados a sílabas pretônicas e sílabas tônicas (indicadas em caixa alta) é feita com base nos referidos estudos sobre a prosódia do PB, que demonstraram a relevância do domínio ϕ para a caracterização entoacional dos enunciados.¹²⁴ Não está em questão se o tom poderia ser simples ou complexo, nem qual poderia ser sua configuração (se LH ou HL, por exemplo). Basta para a análise prosódica dos dados de escrita haver a possibilidade de associação de tons (T, se associado à sílaba pretônica de palavra; T*, se associado à sílaba mais proeminente de ϕ) a certas sílabas do enunciado. A descrição e análise a serem apresentadas tratarão da flutuação entre hipersegmentação e convenção ortográfica de formas do mesmo verbo, visando explicitar características da estrutura prosódica do que levariam a essa flutuação.

4.1'. [[FIco]_T] ω [COM ver SANdo] _{T*} ω] ϕ

4.2'. [[com VERso] _{T*} ω] ϕ

4. 3'. [[FIca] _T] ω [COM verSANdo] _{T*} ω] ϕ

4.4'. [[a GENte]_T] ω] ϕ [[conVERsa] ω] ϕ

¹²³ Observa-se que essa posição de ϕ dentro de *I* constitui aspecto importante para o fraseamento (*parsing*) do enunciado em domínios prosódicos (TENANI, 2002; FERNANDES, 2007).

¹²⁴ Faz-se necessário explicitar que, para fins da análise pretendida, foram consideradas as descrições feitas por Frota e Vigário (2001), Tenani (2002) e Serra (2009) acerca da configuração de eventos tonais no domínio de ϕ . Também foram mobilizadas descrições sobre atribuição de acento secundário no PB, como discutido por Collischonn (1994) e Fernandes-Svartman et al. (2012), principalmente.

$$\begin{array}{c}
T^* \qquad \qquad T^* \\
4.5'. \quad [[VOU]\omega \quad [comver \underline{SA}]\omega] \phi \\
\qquad \qquad T^* \qquad \qquad T^* \\
4.5''. \quad [[vou]\sigma \quad [COMver \underline{SA}] \omega] \phi \\
\qquad \qquad T \qquad \qquad T^* \\
4.6'. \quad [[COmo \ a \ GENte]\omega] \phi \quad [\quad [COMver \underline{SA}]\omega] \phi \\
\qquad \qquad T^* \qquad \qquad T^* \qquad \qquad T \qquad \qquad T^* \\
4.7'. \quad [[a \ GENte]\omega] \phi \quad [\quad COM \ ver \ \underline{SA}mo]\omega] \phi \\
\qquad \qquad T^* \qquad \qquad T \qquad \qquad T^*
\end{array}$$

Da configuração prosódica apresentada, constata-se que as formas verbais em (4.1'), (4.3') e (4.5') pertencem a ϕ_s formados por duas ω_s , enquanto as demais formas verbais – em (4.2'), (4.4'), (4.6') e (4.7') – pertencem a ϕ_s formados por uma única ω . Grafias hipersegmentadas ocorrem nos dois tipos de estruturas. Essa comparação leva a descartar a possibilidade de haver correlação entre hipersegmentação e número de ω_s dentro de ϕ_s .¹²⁵

Passo a analisar o papel de acento secundário e de tom associado à sílaba pretônica hipersegmentada na colocação não-convencional de branco dentro das formas verbais. A hipótese que guia essa análise é a de que a hipersegmentação se constituiria em representação gráfica, por meio do espaço entre sílabas, da representação de lócus de saliência prosódica, dada em termos de acento secundário e de tons. Essa hipótese está ancorada na constatação de haver covariância entre acentos secundários e densidade tonal no PB, o que levaria sílabas pretônicas a serem prosodicamente salientes. Sob esse viés de análise, interpreta-se que hipersegmentações são indícios de espacialização da saliência prosódica. As hipersegmentações são efeitos da organização prosódica (relações da dimensão fônica para Chacon, 1998), como visto no capítulo 3, e revelam, também, como se dá a espacialização da linguagem que, como visto na seção anterior, é um processo de representação simbólica próprio da escrita.

As hipersegmentações em (4.1'), (4.3') e (4.7') têm a sílaba pretônica hipersegmentada. A essa sílaba pretônica de cada uma dessas formas hipersegmentadas podem ser atribuídos acento secundário e tom, o que lhe confere saliência rítmica e

¹²⁵ O número de palavras prosódicas dentro de *I* é relevante para a caracterização da configuração entoacional dos enunciados (cf. TENANI, 2002).

entoacional. Essa proeminência prosódica constitui característica do enunciado que parece motivar o espaço em branco não-convencional nessas formas verbais. Em (4.4'), não há hipersegmentação e a sílaba pretônica não é potencial candidata a receber acento rítmico ou tom, o que corrobora a hipótese de haver relação entre hipersegmentação e configuração prosódica da palavra. No entanto, em (4.2'), há hipersegmentação da pretônica sem que essa sílaba seja prosodicamente saliente. Já nas grafias das formas verbais em (4.5') e (4.6') as sílabas pretônicas não são hipersegmentadas, embora essas sílabas sejam locus potencial de acento secundário e tom. Apesar de parecer enfraquecida a correlação entre hipersegmentação e sílabas pretônicas prosodicamente salientes, é importante atentar para as possíveis configurações de proeminência rítmica e de tons no enunciado em (4.5): visualiza-se em (4.5') proeminência rítmica e tom associados à sílaba pretônica “com”, enquanto o monossílabo “vou” não teria proeminência prosódica por vir a funcionar como clítico prosódico (se perde o acento lexical por passar a funcionar como item gramatical, um verbo auxiliar na locução verbal); em (4.5''), proeminência rítmica e evento tonal estão associados à sílaba “vou”, enquanto à sílaba imediatamente seguinte “com” não estaria associada à proeminência prosódica. Se admitida a possibilidade de configuração prosódica em (4.5''), a consequência é a grafia de “conversar” corroborar a hipótese explicativa que toma a proeminência prosódica como motivadora de hipersegmentações de sílabas pretônicas (isto é, não haveria segmentação, porque não haveria proeminência prosódica da sílaba pretônica).

Portanto, em cinco das sete ocorrências analisadas haveria evidências que favoreceriam a hipótese acerca do papel da configuração prosódica da palavra na hipersegmentação, sendo que, em três das quatro hipersegmentações, o espaço em branco grafado não-convencionalmente poderia ser interpretado como recurso gráfico para sinalizar proeminência prosódica (rítmica e entoacional) de sílaba pretônica de palavras.

Demonstrado que essa hipótese não é suficiente para explicar as hipersegmentações, poderia haver a suposição de que fator semântico estaria atuando, pois, nas quatro hipersegmentações, “com” e “ver”, sílabas pretônicas em “conversando” e “conversamos”, correspondem, cada uma delas, a palavras ortográficas em português, a saber: a preposição “com” e a forma verbal “ver”.¹²⁶ No caso de “com verso” parece

¹²⁶ O registro não-convencional de “M” na coda da primeira sílaba pretônica denuncia relação entre convenção ortográfica e espaço em branco na medida em que a regra de emprego das letras “M, N” para representar a nasalidade da coda silábica no português prevê: (i) “M” quando fim de sílaba coincide com fim de palavra, como regra geral, e “N”, nesse mesmo contexto, somente em algumas poucas palavras (como “hífen, pólen”), e (ii) “M” quando, em fim de sílaba interna à palavra, é seguida por consoante

evidente que a grafia da sílaba “com” é grafada como se fosse monossílabo átono (um item gramatical) e “verso”, um substantivo (um item lexical) do português.¹²⁷ A motivação da presença não-convencional de fronteira de palavra estaria, pois, na confluência de informações prosódicas e semânticas, como já observara Abaurre (1991a) para dados de escrita infantil.

Entretanto, essa potencial correspondência entre sílaba pretônica e informação morfossemântica sobre palavra também não é suficiente para explicar porque ora há, ora não há hipersegmentação das sílabas pretônicas “com” e “ver” nas formas flexionadas de “conversar”. Parto da premissa que essa flutuação não é fruto do acaso ou da simples falta de atenção do aluno (que potencialmente poderia ser diagnosticado como portador de distúrbio específico de escrita, como discutido na seção 1.1).

Antes de ser completamente descartada a hipótese de as hipersegmentações serem efeitos de proeminências prosódicas e de relações morfossemânticas, consideração importante deve ser feita sobre a ocorrência “com verso”. O fato relevante diz respeito à presença de rasura justamente na fronteira entre “com” e “verço”. O gesto de rasurar denuncia estarem em jogo aspectos outros que não apenas o linguístico, como argumentam Capristano (2013) e Felipeto (2013) a propósito de rasuras e segmentações de palavra na escrita infantil. Dentre esses aspectos, vale destacar aqueles relativos à dimensão textual-discursiva mobilizadas no processo de produção escrita.

Em outras palavras, a descrição e análise realizadas demonstram que hipersegmentações são complexos dados linguísticos que se particularizam em função de características prosódicas e morfossemânticas das palavras nos enunciados, mas há indícios de que também seja relevante considerar a constituição textual-discursiva desses enunciados em que hipersegmentações ocorrem.

Argumentos a favor dessa interpretação serão apresentados a partir da consideração do enunciado da proposta de produção textual que teve como tema “Amizade e internet”, reproduzido na Figura 4.5. Destaca-se, para fins desta análise, que o enunciado da proposta é composto por, principalmente, três partes, as quais são indicadas por meio de marcadores alinhados à margem esquerda da folha. Na primeira

homorgânica “P, B” e “N”, quando, nesse mesmo contexto silábico, é seguida por sílaba cujo ataque seja preenchido pelas demais consoantes.

¹²⁷ Vale chamar a atenção para o fato de ter sido usada predominantemente a preposição “com” na regência do verbo (“conversar com” em detrimento a “conversar sobre”); o registro de “com” parecer ter, assim, saliência também morfosintática, além de visual, como em: “com verço com”.

parte, são retomados os dois textos (de gêneros distintos)¹²⁸ que foram lidos e discutidos coletivamente em sala de aula, sob supervisão do professor responsável pela turma (cf. mais detalhes na seção 2.1), e que embasam a proposta do tema da amizade e dos usos de redes sociais para assegurar ou propiciar amizades. Na segunda parte da proposta, explicitam-se tema e gênero: relato de relações de amizade que são mantidas (ou não) por meio da internet. O destaque, por meio de negrito, a “relatando como surgiu essa amizade” aponta para a identificação de tema (“amizade”) e gênero (“relato”), requerida pela instituição do estudante. Na última parte, são dadas as orientações que se circunscrevem à extensão e formatação do texto na folha de papel.

Figura 4.5: Enunciado da proposta “Amizade e internet”

unesp  UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de São José do Rio Preto

Projeto de extensão: **Oficinas de Leitura, Interpretação e Produção Textual**

E. E. _____
Nome: _____ Série/Turma: 8ª _____ Nº _____
Proposta 1 Data: _____

- No poema, Machado de Assis declama a alegria de se ter um grande amigo. No texto da revista *Superinteressante*, a internet é apresentada como uma ferramenta que ajuda a ter mais amizades.
- Assim como o autor do poema, provavelmente, você deve ter um grande amigo(a) com quem gosta de passear, conversar e até estudar juntos(as). Escreva um texto, **relatando como surgiu essa amizade**, se usa ou não a internet para se manter contato com seu(sua) amigo(a), e se houve algum bom momento que vocês viveram juntos(as).
- Seu texto deve conter de 25 a 30 linhas e deve ser escrito à caneta. Seu texto não deve ultrapassar os limites designados para a escritura. Se quiser, dê um título a seu texto.

Fonte: Banco de Dados de Escrita do Ensino Fundamental II (TENANI, 2015)

Faz-se relevante retomar dos textos que compõem a proposta de produção textual que o poema “Bons amigos”, atribuído a Machado de Assis,¹²⁹ é uma ode à amizade (“Abençoados os que possuem amigos, os que os têm sem pedir”), enquanto “Como a internet está mudando a amizade”, extraída da revista de divulgação científica *Superinteressante*, afirma, inicialmente, que “A internet é a ferramenta mais poderosa já inventada no que diz respeito à amizade” para, em seguida, por em questão essa afirmação: “Mas será que as amizades *online* não fazem com que as pessoas acabem se

¹²⁸ Nos anexos, os textos são reproduzidos tal como foram apresentados aos alunos.

¹²⁹ O poema “Benditos”, cuja autoria é de Isabel Machado (cf. http://pensador.uol.com.br/autor/isabel_machado/), teve sua autoria atribuída à Machado de Assis pelos elaboradores da proposta textual.

isolando e tenham menos amigos *offline*, ‘de verdade’?”. A problematização do tema sobre o papel das redes sociais nas relações contemporâneas de amizade tem seu desfecho no relato da conclusão de pesquisa feita em universidade canadense: “a internet faz você ter mais amigos – dentro e fora da rede”.¹³⁰

No texto selecionado – cf. Figura 4.4 –, constata-se que o escrevente responde à proposta de relatar como mantém suas amizades por meio de redes sociais e dialoga com os textos da proposta. O título “A vida no dia a dia na internet” (que se repete no início do texto) anuncia ao suposto leitor empírico no processo escolar (privilegiadamente projetado como pesquisadores da UNESP, pois o texto foi escrito para ser enviado a essa universidade) que o relato – gênero solicitado na proposta e que estava sendo trabalhado, à época da coleta das produções textuais, em sala de aula – será sobre a vida dele *na* internet. Identifica-se, pois, que o aluno projeta-se, de partida, no relato dos usos da internet em seu cotidiano. Esse relato se inicia com a descrição de sua rotina após sua chegada da escola: almoçar, assistir jogo na televisão, entrar na internet. Da chegada da escola em casa e à entrada no ambiente digital do computador com acesso à internet, tudo é feito com muita velocidade, como indica o emprego do advérbio de tempo “já”, utilizado cinco vezes na parte inicial do relato. Essa rapidez no modo de realizar as ações prosaicas coloca em evidência, de nosso ponto de vista, tanto um ritmo veloz projetado socialmente como característico da internet, como uma ansiedade do jovem em sentir-se “completo” na relação com o outro naquele espaço (da escrita “virtual”). Na internet, especificamente no Orkut e no MSN, ferramentas digitais de forte apelo entre usuários à época, o escrevente passa o tempo “conversando com os amigos”. Ao tratar do tema proposto, ocorre a primeira hipersegmentação “com ver sando”, um registro gráfico que pode ser interpretado como efeito de destaque verbo-visual. É na consideração da dimensão textual-discursiva que se encontra ancoragem para justificar a emergência das formas verbais de “conversar” hipersegmentadas, pois o relato de como conversar na internet é central em seu texto: para manter suas amizades, o aluno *conversa* na internet.

Além de o relato tratar de como a internet e as redes sociais proporcionam “conversas” com os amigos, também há descrição de como essas conversas não se limitam a relações virtuais e duram no tempo, como se infere dos usos recorrentes de formas verbais flexionadas no gerúndio (“conversando” (2x), “vendo”). Essa duração da ação – o que não se circunscreve à forma verbal de gerúndio – é representada por meio

¹³⁰ O artigo completo encontra-se disponível em: <http://super.abril.com.br/comportamento/como-a-internet-esta-mudando-a-amizade>. Acesso em 10 de set. 2016.

de espaços em branco entre as sílabas. Espacializa-se, por meio do branco entre as sílabas das formas flexionadas do verbo “conversar”, a duração da ação no tempo que, ambigualmente, é tanto relativa ao aspecto da duração verbal (dimensão morfológica), quanto relativa à duração da enunciação (dimensão enunciativa). Explicita-se, a seguir, o aspecto em traço [+ durativo] ou [- durativo] para os sete enunciados em análise, a fim de demonstrar de que modo a dimensão textual-enunciativa se mostra relevante.

4.1”. “eu fico com ver sando com os meus amigos”
[+ durativo]

4.2”. “com (rasura) verço com a minha prima Jeine”
[+ durativo]

4.3”. “nós fica com versando”
[+ durativo]

4.4”. “a gente conversa de ir no lanche do meu colega o Pablo”
[- durativo]

4.5””. “no MSN eu vou conversa com a Estela”
[- durativo]

4.6”. “depois que agente saiu não teve como a gente conversa com ela”
[- durativo]

4.7”. “e ai a gente com ver çamo no MSN”
[+ durativo]

A identificação do aspecto [+ durativo] atribuído às formas flexionadas de “conversar” é feita não apenas com base no aspecto morfológico, inscrito na forma flexionada de gerúndio – que se verifica em (4.1”) e (4.3”) –, mas também a partir da consideração de como o relato de conversar é construído no texto. As quatro hipersegmentações coincidem com o relato de a conversa durar no tempo, enquanto as três formas verbais não hipersegmentadas coincidem com a descrição da ação de conversar. Em (4.4”), é relatado o tema da conversa (“a gente conversa de ir no lanche do meu colega o Pablo”); em (4.5””), o fato de conversar com uma amiga via internet (“no MSN eu vou conversa com a Estela”); em (4.6”), a dificuldade de conversar com a amiga por problemas no computador (“depois que agente saiu não teve como a gente conversa com ela porque o PC quebrou”). Já nos enunciados (4.1”) e (4.3”), o aspecto

durativo é expresso morfologicamente por meio do gerúndio verbal. Nos enunciados (4.2”) e (4.7”), as formas flexionadas no presente do indicativo também possibilitam interpretar o aspecto durativo: em (4.2”), a duração da ação de conversar pode ser inferida a partir da construção “no Orkut eu fico vendo os Orkut das pessoas com verço com a minha Prima Jeine”, em que há o relato do que o aluno faz na rede social Orkut; em (4.7”), a longa duração da conversa vem explicitada pelo horário em que terminou na madrugada: “e aí a gente com ver çamo no MSN a gente fico até às 04:00 horas da manhã”.

O confronto entre as hipersegmentações, por um lado, e a ausência delas, por outro, das formas flexionadas de “conversar” leva à constatação de que, nas hipersegmentações, o espaço gráfico entre as sílabas representa o tempo/aspecto da duração da própria enunciação. É a espacialização na escrita do tempo linguístico (seja o tempo organizado pelas estruturas prosódicas, seja o tempo/aspecto inscrito na enunciação). Nessa perspectiva, os espaços não-convencionais dentro da palavra também materializam o destaque de natureza textual-discursiva dado a determinadas palavras/porções do texto.

Mas não só! O espaço entre as palavras constitui espacialização do tempo histórico (aquele vivido pelo escrevente), especialmente se consideradas características do tempo atribuídas à escrita digital, tematizada pelo escrevente. A escrita digital tem como traço a sincronia entre os enunciadores, escrevente/leitor (MARCUSCHI, 2004).¹³¹ Há (certa) aproximação no tempo entre escrevente e leitor dos enunciados escritos, apesar de distantes entre si no espaço. No texto analisado, os espaços não-convencionais entre fronteiras de palavras plasmam o tempo da enunciação: o tempo histórico da interação entre os sujeitos fica capturado nos espaços que se materializam entre as sílabas das palavras, unidades linguísticas básicas.¹³²

A partir dessa constatação e com base na análise linguístico-discursiva feita sobre os textos produzidos e as propostas de produção textual, conclui-se que

¹³¹ Faz-se necessário agradecer, neste momento, a professora Dra. Fabiana Komesu pela observação sobre o conceito de escrita digital síncrona e pela indicação de referência bibliográfica.

¹³² Ainda no âmbito da escrita digital, cabe acenar à potencial relação entre a interpretação sobre tempo da enunciação e espaços verbo-visuais e a noção de *ubiquidade*, qualidade entendida como aquela do que está ou pode estar em muitos lugares ao mesmo tempo ou quase no mesmo tempo. A condição de ubiquidade seria proporcionada pelo suposto “fim de barreiras” que o emprego de tecnologias digitais permitiria (SANTAELLA, 2013). Na internet, o escrevente “conversa” (tecla) *ao mesmo tempo* com diferentes amigos (Estelas, Jeines) que estão ou podem estar fisicamente (espacialmente) distantes, mas que naquele espaço se materializam por uma *hipermobilidade* que permitiria que “individualidades” espalhadas formem um “todo” (de amizade). É um atributo (quase divino) o de uma onipresença em muitos e diferentes espaços, com diferentes interlocutores, constituindo a (quase) superação da condição humana assujeitada pelo tempo.

hipersegmentações são efeitos da construção de sentidos mobilizados (e enfatizados) pelos discursos que atravessam e constituem o texto (e o sujeito) e que se concretizam na escrita por meio de representação de características prosódicas dos enunciados (notadamente sobre configuração rítmica e entoacional) em confluência com aspectos morfológicos (seja sobre a identificação de palavras lexicais e gramaticais, seja sobre o funcionamento dessas palavras no enunciado).

Ao cabo deste capítulo, fica demonstrada haver relação entre prosódia e escrita, presente na constituição da escrita alfabética, que se atualiza em cada produção escrita e atravessa os sujeitos da linguagem e é revelado como se instauram tensões entre possibilidades de representação do verbal e do verbo-visual da língua(gem) (CORRÊA, 2006).

Essa conclusão implica “ajustar as lentes” de análise das hipersegmentações para relações textuais-discursivas que constituem a produção escrita, não sendo suficiente, embora relevante – como demonstrado – lançar mão de hipóteses que se circunscrevam a mobilizar características prosódicas e morfossintáticas das palavras e dos enunciados em que os espaços não-convencionais são registrados. Ao ser ampliada a perspectiva de análise, necessariamente entra em cena o sujeito escrevente que se revela na constituição de diferentes discursos que se materializam no texto, como busquei argumentar.

5 Considerações finais

Neste capítulo, são feitas as considerações finais visando a, por um lado, retomar e sintetizar os resultados das análises desenvolvidas e, por outro lado, estabelecer relações entre esses resultados e dois temas centrais a esta tese, a saber, abordagens sobre relações entre fala e escrita e sobre erros ortográficos. Também são explicitadas contribuições teórico-metodológicas, bem como questões que carecem de investigações futuras.

Inicialmente, faz-se importante retomar que esta tese propõe que segmentações não-convencionais de palavra sejam definidas como *palavras escritas cujas fronteiras gráficas não seguem as convenções ortográficas do português*. Nessa definição, estão inscritos dois posicionamentos teóricos. O primeiro diz respeito ao fato de se assumir a noção de *fronteira* como relevante para descrever o objeto de investigação, diferentemente do que se observa na literatura linguística que, predominantemente, define segmentações não-convencionais de palavra em função da presença e/ou ausência do branco, um dos tipos de registro da fronteira de palavra. A presença não-convencional do branco, no caso das hipersegmentações, ou a sua ausência não-convencional, no caso das hipossegmentações, são tipos de representação da fronteira de palavra na escrita. Portanto, a noção de *fronteira* é mais abrangente além de fundamental para a identificação de palavras.

Do ponto de vista linguístico, tratar de *fronteira de palavra* implica problematizar o fato de não haver, necessariamente, sobreposições entre fronteiras de palavra prosódica e morfológica, por exemplo, o que põe em evidência a complexidade da noção de palavra, unidade linguística que se constitui de características fonológicas, morfológicas, sintáticas e semânticas. Essas características foram tratadas nos capítulos 1 (seção 1.2) e 3 (seção 3.1) à medida que foram descritas as questões relevantes para análise das hipersegmentações. A abordagem dessa complexidade linguística foi feita a partir do arcabouço teórico da Fonologia Prosódica (NESPOR; VOGEL, 1986, 2007) que fundamentou a concepção de prosódia enquanto estruturas linguísticas as quais atuam na configuração do ritmo e da entoação dos enunciados. Com base nessa abordagem teórica, foi explicitado modos pelos quais a organização rítmico-entoacional dos enunciados (dimensão fônica da linguagem, nos termos de Chacon, 1998) constituem os textos

escritos (dimensão textual-discursiva para o mesmo autor) e motivam hipersegmentações, como argumentado no capítulo 4.

O segundo aspecto teórico que subjaz à denominação “segmentação não-convencional de palavra” diz respeito, como anunciado na introdução, à abordagem dada às convenções ortográficas que não compactua com as perspectivas normativas e clínicas, predominantes, respectivamente, nas áreas de Educação e Saúde, que classificam, necessariamente, os registros escritos dos alunos em acertos e erros, e esses últimos como sintomas de disortografia, como discutido no capítulo 1 (seção 1.1). A abordagem adotada nesta tese defende que investigar os chamados erros ortográficos de uma perspectiva linguística implica, necessariamente, mobilizar conceitos de língua e de escrita. Desse lugar teórico, são problematizadas as concepções de língua e de escrita grafocentricamente orientadas (CORRÊA, 2006) que se verificam, muito frequentemente, em ambiente escolar e, de nossa perspectiva, também em clínicas fonoaudiológicas.

A partir das críticas feitas à interpretação das segmentações não-convencionais de palavras quer como simples desvios das convenções ortográficas (decorrentes de falta de atenção do aluno ou de dificuldades de alfabetização vividas principalmente em escolas públicas ao longo do EF I), quer como sintomas de disortografia (decorrentes de dificuldades de processamento cognitivo de alunos do EF II quanto a estabelecer relação entre fala e escrita) e considerando-se os resultados de pesquisas feitas sobre características linguísticas relevantes para descrição e interpretação das grafias não-convencionais de palavra no EF I, é explicitada a proposta de interpretar *segmentações não-convencionais de palavras como efeitos do trânsito dos sujeitos escreventes (crianças e adolescentes) por práticas sociais letradas/escritas e, também, por práticas sociais orais/faladas*.

Essa interpretação é ancorada nos resultados obtidos a partir de análises quanti e qualitativa de segmentações não-convencionais de palavras em textos do EF II, os quais são sumarizados a seguir.

São dois os principais resultados quantitativos descritos nesta tese, obtidos a partir da investigação da frequência de ocorrência de segmentação não-convencional de palavras em uma amostra longitudinal de 2.457 textos produzidos por alunos de uma escola pública paulista. Foram identificadas características do EF II comuns à escrita do EF I que diz respeito a haver: (i) decréscimo de frequência de segmentação não-convencional de palavras em função do aumento do tempo de escolarização; (ii)

diferenças quanto à frequência de segmentação não-convencional de palavra em função do sexo/gênero dos alunos (sendo menor a frequência nos textos de meninas do que nos textos de meninos).

Ainda foi identificado resultado quantitativo que acena para o que se mostra como distinção entre do EF II em relação ao EF I: apenas no EF II, o aumento de anos de escolarização contribui significativamente para o declínio da hipersegmentação, mas não significativamente para a queda da hipossegmentação. Em outras palavras, foi observado declínio de ocorrências de hipossegmentação ao longo do EF II, mas não foi encontrado correlação estatisticamente significativa entre frequência de hipossegmentação e ano escolar. Já sobre o EF I, os estudos de Cunha (2010, 2004) e Capristano (2007a, b) apontam haver decréscimo de hipossegmentação em função do aumento de anos de escolarização. No entanto, nesses estudos não foi feito tratamento estatísticos dos dados. Já análise estatística feita por Torquette (2016) sobre segmentações não-convencionais de palavra no EF I, com base em amostra transversal de textos, não se investigou relação entre cada tipo de segmentação não-convencional de palavra e tempo de escolarização. Esse conjunto de resultados configura um cenário em que respostas carecem de ser dadas quanto a haver, ou não, efeito estatisticamente significativo do tempo de escolarização sobre hipo e hipersegmentações no EF I.

Para futuras pesquisas, portanto, uma investigação a ser feita é precisar o efeito de tempo de escolarização sobre cada tipo de segmentação não-convencional de palavra, em particular, e sobre a presença dos chamados erros ortográficos, de modo amplo. A busca por respostas a essa possível relação é especialmente relevante diante da atual política educacional brasileira que ampliou de oito para nove anos o EF e tornou obrigatória a entrada de crianças de quatro anos na pré-escola.¹³³ O aumento de 11 para 14 anos de escolarização da chamada Educação Básica visa alcançar, ao final, melhores índices de desempenho dos alunos nos sistemas (governamentais ou não) de avaliação da Educação brasileira.

Por meio da descrição da frequência de segmentação não-convencional de palavra ao longo do EF II, foi respondida à primeira das duas questões de que trata esta tese que diz respeito a caracterizar as ocorrências de hipo e hipersegmentação na escrita do EF II

¹³³ A obrigatoriedade da matrícula de crianças a partir de quatro anos na pré-escola foi estabelecida pela Lei nº 12.796/2013, publicada no D.O.U. em 05 de abril de 2013. Essa lei altera a Lei nº 9.394/1996 que estabelece as diretrizes e bases da educação nacional (LDB). Em início de 2016, encerrou-se o prazo para que fosse implantada a lei nº 12.796/2013 nas redes municipais e estaduais.

e relacioná-las àquelas características relatadas nas pesquisas realizadas sobre a escrita do EF I.

No que diz respeito à interpretação do resultado inédito de ter sido atestada relevância estatística de efeito do sexo/gênero na frequência de segmentação não-convencional de palavras em textos de adolescentes do EF II, lança-se mão, em um primeiro momento, de pesquisas feitas na área de Educação que investigam relações entre educação e sexo/gênero principalmente. Em um cenário contemporâneo em que práticas de leitura e escrita são discutidas como parte dos requisitos necessários para ampliar o acesso à cidadania (especialmente de mulheres pobres), os resultados sobre haver diferença de incidência de segmentações não-convencionais de palavras relacionada ao sexo/gênero dos escreventes ganham contornos que transcendem à descrição de características da escrita do EF II em particular, como já observado. Esse resultado quantitativo suscita o desenvolvimento de futuros estudos (nas áreas da Educação, Ciências Sociais e Linguística, por exemplo) que venham a considerar sexo/gênero característica social dos sujeitos relevante para descrição e análise da escrita produzida/ensinada em ambiente escolar.

Em um segundo momento, esses resultados quantitativos sobre os efeitos na frequência de segmentação não-convencional de palavras do fator sexo/gêneros dos estudantes e do fator tempo de escolarização (que implica diferenças entre as características da segmentação do EF II em relação às descritas do EF I) são interpretados como efeitos de sobreposições de representações sociais de sexo/gênero e de infância/adolescência, por um lado, com representações de práticas letradas/escritas e orais/faladas, por outro lado, que constituem os sujeitos da linguagem.

Com base nesse resultado quantitativo quanto à identificação de efeito dos anos letivos na ocorrência de hipersegmentação, tanto em textos de meninos quanto de meninas na escrita do EF II, a análise linguística de dados foi delimitada às hipersegmentações. Ainda quando considerada a relação entre número de ocorrências de hipersegmentação por sujeito e sexo/gênero do aluno, foi possível verificar que tanto meninos quanto meninas hipersegmentam: (i) muitas vezes as mesmas palavras, o que evidencia dúvidas ortográficas sobre um conjunto pequeno de palavras; (ii) palavras que apresentam em comum traços de mesma classe gramatical.

Esses dois resultados quantitativos gerais motivaram o desenvolvimento da análise linguística dos dados em duas etapas: inicialmente (seção 3.2), foram investigadas regularidades quanto à organização métrica das palavras que atuam favorecendo as

hipersegmentações e, posteriormente (seção 3.3), regularidades quanto à organização prosódica do enunciado em que a palavra hipersegmentada ocorre, investigando-se em que medida as hipersegmentações são efeitos da atuação de aspectos morfossintáticos e semânticos dos enunciados.

Na primeira etapa da análise linguística, foram investigados dois fatores de natureza fonológica que dizem respeito à localização do acento lexical de palavra (oxítono, paroxítono, proparoxítono) e ao número de sílabas da palavra (dissílabo, trissílabo e polissílabo) ao lado de dois fatores de natureza verbo-visual sobre tipo de informação fronteira gráfica que delimita a palavra escrita (branco ou hífen) e localização da fronteira não-convencional em relação às fronteiras convencionais de palavra (após a primeira sílaba da palavra, no meio da palavra ou antes da última sílaba da palavra).

Da análise desses fatores (feita na seção 3.2), três relações mais gerais foram identificadas, a saber: (i) dissílabos oxítonos são mais hipersegmentados do que dissílabos paroxítonos; (ii) trissílabos paroxítonos são mais hipersegmentados do que trissílabos oxítonos; (iii) polissílabos paroxítonos são os mais hipersegmentados e polissílabos com quatro ou cinco sílabas passam a ter mais proparoxítonas hipersegmentadas do que trissílabos ou polissílabos com seis sílabas.

À luz dos resultados de frequência de número de sílabas de palavra em relação à localização de acento lexical descritos para o PB, as relações descritas em (ii) e (iii) obtidas para as hipersegmentações na amostra de escrita do EF II foram interpretadas como, em parte, decorrentes de haver maior frequência de trissílabos e polissílabos paroxítonos na língua. A consequência dessa interpretação é admitir que não necessariamente trissílabos e polissílabos paroxítonos oferecerem maior dificuldade ortográfica a alunos do EF II, o que poderia ser traço semelhante ao EF I, se considerada a hipótese proposta por Abaurre (1991a) para interpretar que a estrutura métrica de trissílabos justificaria haver maior ocorrência de trissílabos hipersegmentados em textos infantis.

Também foram investigadas as estruturas métricas das palavras hipersegmentadas, considerada a localização do acento de palavra. Os resultados descritivos de configurações métricas das hipersegmentações são sumarizados, a seguir, em função da classificação das palavras em dissílabos, trissílabos e polissílabos, a começar pelos trissílabos por constituírem o conjunto com mais ocorrências de hipersegmentação e por apresentarem o maior conjunto de possibilidades de segmentação

não-convencional de palavra em comparação aos polissílabos que potencialmente oferecem maior número de possibilidades de hipersegmentação.

Para os trissílabos, predominam hipersegmentações de paroxítonos que se caracterizam por isolar a primeira sílaba da palavra entre espaços em branco (45, 56%), o que resulta na configuração de sílaba seguida de pé métrico troqueu (“a gora”). Cabe destacar também o resultado quanto ao segundo tipo de configuração em número de ocorrências (3,73%) dentre os trissílabos que compreende hipersegmentação por hífen da última sílaba de paroxítonos (“dança-se”), o que configura hipersegmentação de morfemas verbais principalmente. Em síntese, quanto à relação entre localização da fronteira não-convencional e ao tipo de fronteira gráfica, o principal resultado é: sílabas pretônicas tendem a ser segmentadas por branco, enquanto sílabas postônicas tendem a ser segmentadas por hífen, entre os trissílabos.

Verifica-se que a localização de fronteira não-convencional de palavra entre trissílabos paroxítonos decorre, em parte, da configuração métrica de sílaba seguida de pé troqueu na medida em que o espaço em branco coincide com a fronteira prosódica entre sílaba e pé métrico, como em “com verso”, e nunca segmenta a sílaba átona do pé. A hipersegmentação da sílaba postônica de trissílabos só ocorre por meio do hífen, como em “dança-se”. Esses resultados, somados ao fato de o hífen ser o tipo de fronteira gráfica predominante nessas hipersegmentações, apontam para a atuação de certa estrutura morfossintática nesse subconjunto de registros: a ênclise verbal de pronomes átonos. Essa colocação do pronome enclítico ao verbo (como “ajude-me”) é prosodicamente marcada em PB, não apenas por ser preferencial a colocação proclítica ao verbo (como “me ajude”), mas porque a ênclise verbal decorre de práticas letradas fortemente permeadas pela valorização de registros escritos que não necessariamente são motivados em regras gramaticais atestadas em enunciados falados em PB. A interpretação desses resultados de hipersegmentação de trissílabos leva à identificação de como se dá o entrecruzamento de características prosódicas e morfossintáticas de palavras com práticas letradas/escritas (especialmente as valorizadas em ambiente escolar, de modo amplo).

Entre os polissílabos, segundo conjunto de palavras mais hipersegmentadas, foi verificada flutuação, principalmente, entre segmentar a última sílaba do polissílabo – que corresponde a morfema verbal (“mantinha-mos”) – e segmentar duas sílabas do polissílabo que se configuram em pé métrico – podendo haver correspondência entre uma das duas partes hipersegmentadas a palavras da língua (“aero porto”). Esse principal resultado quanto à hipersegmentação de polissílabos, quando comparado com aqueles

descritos para os trissílabos, permitiu verificar que não apenas há relação entre tipo de fronteira gráfica e a localização não-convencional da fronteira de palavra, mas também que essa relação sofre efeito do número de sílabas da palavra, pois para polissílabos, mas não para trissílabos, a hipersegmentação da sílaba postônica final (de paroxítonos e proparoxítonos) por hífen é mais característica; enquanto para trissílabos, mas não para polissílabos, a hipersegmentação da sílaba pretônica por branco (de paroxítonos e oxítonos) é mais característica. No primeiro caso, a sílaba hipersegmentada, um sufixo verbal, é grafada como se fosse um pronome enclítico ao verbo: “-sse” como se fosse “se” em “dança-se” (“dançasse”), “-mos” como se fosse “nos” em “morava-mos” (“morávamos”). No segundo caso, a sílaba pretônica hipersegmentada tende a corresponder a um possível monossílabo átono, como as preposições “com, em” em “com versar”, “em bora”, seguida da porção da palavra em que se encontra o acento lexical. Em ambos os casos, a sílaba hipersegmentada corresponde a um possível clítico da língua, resultado atestado em Tenani (2011b) em amostra transversal de textos do EF II, fato que corrobora a interpretação de ser esse traço geral característico das hipersegmentações do EF II.

Simultaneamente à atuação do pé métrico, constatou-se que a localização não-convencional de fronteira de palavra coincide com fronteira morfossemântica de, ao menos, uma das porções hipersegmentadas. Essa regularidade quanto à atuação de critérios prosódicos e semânticos na segmentação não-convencional de palavras havia sido observada inicialmente por Abaurre (1991a) na escrita do EF I e, posteriormente, atestadas em pesquisas feitas com outros corpórea de textos do EF I (CUNHA, 2004, 2010; CAPRISTANO, 2007a). Os resultados levam, portanto, à afirmação de que a escrita do EF II continua a manifestar essa característica da escrita do EF I, pois foi identificado por meio da análise da configuração métrica de palavra que a característica principal da hipersegmentação da escrita do EF II é a segmentação não-convencional de sílaba quando essas podem corresponder a clíticos prosódicos monossilábicos.

Por fim, são sumarizados os resultados identificados para dissílabos hipersegmentados, tendo em vista a discussão prévia feita para o conjunto maior de ocorrências. Foram predominantemente hipersegmentados dissílabos oxítonos (17,02%), havendo, proporcionalmente, poucos paroxítonos hipersegmentados (3,97%). Em PB, há mais dissílabos paroxítonos do que dissílabos oxítonos, fato que leva à suposição de que a maior porcentagem de hipersegmentação de dissílabos oxítonos não seja

necessariamente decorrente de um maior número de palavras desse tipo na língua, como foi o caso para os trissílabos e polissílabos.

A análise de hipersegmentação de dissílabos oxítonos revela que a sílaba átona dessas palavras corresponde a um clítico prosódico e a sílaba tônica, a palavras da língua. A sílaba tônica, em todas as hipersegmentações de dissílabos, é constituída de rima complexa, o que permite interpretá-las como pés métricos bimoraicos (BISOL, 1992). Segundo essa interpretação, grafias como “de pois” seriam, prosodicamente, constituídas de sílaba átona seguida de pé métrico, resultando na configuração predominante entre os trissílabos. Haveria, nessa interpretação de hipersegmentação de dissílabos oxítonos, a atuação da configuração métrica da palavra ao lado da identificação de monossílabos átonos que corresponderiam a palavras gramaticais do PB. Novamente, identifica-se que hipersegmentações no EF II sejam efeitos da interação entre informação fonológica e morfológica.

Faz-se ainda relevante retomar os resultados identificados entre os dissílabos paroxítonos hipersegmentados. Embora em menor número de ocorrências, esse conjunto de dados é interessante por não terem, a princípio, uma configuração métrica que favoreceria a emergência de hipersegmentações. Portanto, esse subconjunto de grafias não-convencionais (apenas 3,97% das hipersegmentações) não estariam sujeitas à atuação do pé métrico como demonstrado para trissílabos e polissílabos. O exame das hipersegmentação de dissílabos leva à constatação de que uma, se não as duas sílabas, corresponde a palavras da língua: a sílaba átona do dissílabo tende a corresponder a um monossílabo que pertence ao conjunto de palavras gramaticais (preposição, artigo, conjunção), como “de” em “dez de” (“desde”); a sílaba tônica, a uma palavra lexical monossilábica, como “tem” em “tem de” (“tende”).

Portanto, a hipersegmentação de dissílabos paroxítonos tem motivação não somente na configuração métrica da palavra, mas na identificação de monossílabos como palavras ortográficas. Essas grafias de dissílabos poderiam ser interpretadas como motivadas exclusivamente em dificuldades ortográficas dos alunos do EF II, herdadas do EF I. No entanto, essas hipersegmentações de dissílabos paroxítonos decorrem do fato que se mostrou recorrente na análise dos trissílabos e polissílabos, portanto, mais geral: a hipersegmentação de sílabas que corresponderiam a palavras gramaticais monossilábicas.

Cabe lembrar aqui que monossílabos gramaticais predominantemente configuram-se em clíticos prosódicos em PB e, portanto, têm um caráter complexo: são unidades no domínio da morfossintaxe e, fonologicamente, sílabas prosodizadas (isto é,

incorporadas ou anexadas) a um hospedeiro. Assinala-se, pois, a necessidade de pesquisas por meio das quais sejam dadas respostas mais conclusivas quanto ao processo de prosodização de clíticos, notadamente, sobre a caracterização do(s) domínio(s) prosódico(s) relevante(s) em PB.

Da análise ora realizada sobre hipersegmentação de sílabas átonas de palavras na escrita do EF II, verificou-se que clíticos prosódicos têm sua identificação construída fortemente a partir de práticas letradas escritas (o que potencialmente afeta as representações linguísticas dos sujeitos falantes/ouvintes). Essa interpretação encontra respaldo na diminuição desses registros em função do aumento do tempo de escolarização (descrito na seção 2.3). Dados de escrita (como os analisados) possibilitam, pois, entrever (de modo detalhado, inclusive) como representações linguísticas dos enunciados são efeito de intrincada (e cambiante) relação entre práticas orais/faladas e letradas/escritas. Conclui-se que a característica principal da hipersegmentação da escrita do EF II é: a segmentação não-convencional de sílabas átonas tende a ocorrer quando essas poderiam corresponder a clíticos prosódicos monossilábicos.

Além dessa característica principal das hipersegmentações, outros resultados são identificados no que concerne a regularidades quanto à organização morfossintática dos enunciados que também motivam hipersegmentações nos textos da escrita do EF II.

Esses resultados são obtidos a partir da análise de dois conjuntos de hipersegmentações, a saber: (i) “senão” e “demais”, por figurarem entre as hipersegmentações por branco mais frequentes na amostra analisada; e (ii) hipersegmentações de formas verbais que se caracterizam por usos não-convencionais do hífen, como “dança-se” e “morava-mos”, que se particularizam por serem frequentes em textos das alunas do EF II.

Da descrição das características das hipersegmentações de formas verbais por hífen, destaque deve ser dado à análise que demonstra haver relação entre características morfológicas das formas flexionadas e características prosódicas dos enunciados. Quanto a características morfológicas, constatou-se que o hífen é colocado na junção de sufixos: nas formas flexionadas no pretérito imperfeito do subjuntivo na terceira pessoa do singular, o sufixo “-sse” é hipersegmentado (como em “dança-se”) e, nas formas flexionadas no pretérito imperfeito do indicativo na primeira pessoa plural, o sufixo “-mos” é hipersegmentado (como em “estudava-mos”). Quanto a características morfossintáticas, relevantes para a configuração prosódica dos enunciados, conspira para que essas formas verbais sejam hipersegmentadas a colocação de pronomes em relação

ao verbo. Estão em jogo as duas possibilidades de colocação pronominal (a saber: próclise ou ênclise verbal) às quais são associados modos de enunciação e práticas orais e letradas. Nas práticas orais/faladas, predominam usos de próclise verbal (como “nos mantinha”), enquanto nas práticas letradas/escritas, usos de ênclise verbal (como “mantinha-nos”). A projeção da estrutura verbo-pronome se mostra nos registros não-convencionais de formas verbais na medida em que há semelhanças fonológicas e morfológicas entre a forma verbal hipersegmentada (“mantinha-mos”) e a estrutura morfossintática projetada (“mantinha-nos”), pois: (i) o pronome “nos” e o sufixo “-mos” carregam informação gramatical de primeira pessoal plural e, quanto à materialidade fônica, ambas as sílabas são semelhantes entre si, diferindo apenas quanto ao ponto de articulação da consoante nasal; (ii) a porção hipersegmentada que compreende o tema verbal, como “mantinha”, tem a mesma cadeia segmental que a forma verbal flexionada no pretérito imperfeito do indicativo em terceira pessoa do singular, como “ele mantinha”. Essas semelhanças entre formas verbais somadas às convenções de uso do hífen corroboram a projeção da estrutura verbo-clítico (“mantinha-nos”) como visto no registro de forma verbal flexionada no pretérito imperfeito do indicativo na primeira pessoa do plural (“mantínhamos”). Essa descrição demonstra como estruturas morfossintáticas e práticas letradas/escritas são motivadoras de hipersegmentações de palavras no EF II.

Por meio desta análise em que as produções escritas dos alunos ao longo dos quatro anos do EF II são consideradas, esta tese avança em relação aos estudos longitudinais feitos anteriormente (PARANHOS, 2014; SILVA, 2014), pois demonstra de que modo as hipersegmentações por hífen se mostram motivadas pelas estruturas morfossintáticas que os alunos são chamados a empregar em seus textos escritos.

De modo amplo, esses usos não-convencionais do hífen são interpretados como resultados do imaginário social do escrevente sobre a escrita, construído pela sua inserção nos mais diversos usos (formais/informais) da linguagem (CORRÊA, 2004). A tentativa de se aproximar do que seja o código escrito institucionalizado (CORRÊA, 2004) leva os alunos ao chamado erro ortográfico. É o trânsito do escrevente por práticas orais (como aquelas em que predominam formas verbais flexionadas na terceira pessoa do singular e colocação pronominal proclítica) e por práticas letradas (como aquelas que demandam emprego de formas verbais flexionadas no pretérito imperfeito e em primeira pessoa do plural e colocação pronominal enclítica ao verbo) que fica evidenciado nos registros não-convencionais do hífen em formas verbais, por exemplo.

Notadamente, esse trânsito se revelou marcado pelo sexo/gênero dos sujeitos, na medida em que predominantemente meninas grafam não-convencionalmente por hífen. Esse resultado qualitativo também foi interpretado como evidência de que escreventes de sexo/gênero feminino são constituídas por práticas letradas/escritas de modo distinto dos escreventes de sexo/gênero masculino (como argumentado na seção 2.3).

Esses resultados quanti e qualitativos são, de nossa perspectiva, evidências de que não apenas há entrelaçamento entre práticas orais/faladas e letradas/escritas, mas que esse entrelaçamento é constituído pelos papéis sociais ocupados pelos sujeitos da linguagem. Para futuras pesquisas, ficam as tarefas de investigar novas evidências que corroborem ou refutem esses resultados e, especialmente, de ampliar/aprofundar a interpretação teórica desses resultados descritos.

Ainda no capítulo 3, encontra-se outro resultado obtido a partir da análise de “demais” e “senão” (realizada na seção 3.3) que leva à interpretação de que o espaço em branco nessas hipersegmentações indicia, na materialidade da escrita, a saliência prosódica dada a essas palavras, e, desse modo, fica expresso destaque de partes do enunciado ou do enunciado em que ocorrem em relação aos que lhes precede. A configuração entoacional somada à estruturação em constituintes prosódicos dos enunciados em que ocorrem “demais” e “senão” é analisada. Demonstra-se que o ritmo silábico (CAGLIARI, 2007) e a ênfase prosódica (FERNANDES, 2007) são projetadas na escrita por meio de espaços entre sílabas das palavras. Explicitam-se que características rítmicas e entoacionais dos enunciados falados embasam a interpretação de que a hipersegmentação iconiza o destaque dado a “demais” e a “senão” nos enunciados escritos do EF II.

Nessa interpretação, as hipersegmentações são registros verbo-visuais que permitem ao analista explicitar em que medida e de que maneira relações morfossintáticas e semânticas dos enunciados somadas a características prosódicas dos enunciados deixam aflorar características do processo de produção escrita dos escreventes analisados. Demonstra-se, por meio da análise desse conjunto de hipersegmentações, como se dá a relação entre prosódia e escrita em textos do EF II.

Em suma, no capítulo 3, foram descritas características prosódicas e morfossintáticas das hipersegmentações que particularizam a escrita do EF II em relação à escrita do EF I. Essas características das hipersegmentações são interpretadas como efeitos dos diferentes modos pelos quais crianças e adolescentes são atravessados por práticas sociais letradas/escritas e, também, por práticas sociais orais/faladas.

Essa interpretação é fundamentada em Corrêa (2004) que concebe que a fala não seja exterior à escrita, mas a constitui, seja considerada a materialidade da escrita alfabética, sejam consideradas práticas orais e letradas que constituem fala e escrita.

Da discussão teórica desenvolvida no capítulo 4, é importante explicitar o avanço teórico alcançado a partir da proposta de Corrêa (2004) que identifica, na materialidade significativa da escrita, encontros entre as práticas orais/faladas e letradas/escritas. Mais especificamente, é discutida a interpretação de Corrêa (2001) sobre a proposta de Abaurre (1991b) a respeito da noção de “gesto rítmico” congelado no signo gráfico da escrita. A noção de “gesto rítmico”, desenvolvida a partir do conceito de “escrita ritmicamente reprodutiva” de Luria (1988) sobre o ritmo como um traço fundamental da escrita, implica identificar, segundo Corrêa (2001), a atuação de três materiais significantes que atuariam na escrita: o gesto, o material fônico-acústico (ritmo) e o material gráfico.

Destaque foi dado ao fato de sob “gesto” estão em jogo duas noções: *os gestos articulatórios*, referentes tanto a uma *ação de constrição* dos diferentes articuladores quanto à *representação* dessa manobra articulatória necessária para que essa ação se concretize, os quais são percebidos pelo aparelho auditivo (portanto, há produção e percepção de estímulos acústico-articulatórios), e *os gestos gráficos*, executados pela mão e percebidos pela visão (portanto, há produção e percepção de marcas gráfico-visuais).

Problematizada, com base em estudos sobre percepção de fala, a natureza dos gestos articulatórios, argumentou-se que a *fala é um evento auditivo, proprioceptivo e visual integrado* (SEKIYAMA; BURNHAM, 2004). Portanto, quer na fala, quer na escrita está presente uma dimensão visual da linguagem e a consequência dessa constatação é assumir a *constituição heterogênea da materialidade da fala*. De nossa perspectiva, essa assunção é relevante para futuras pesquisas que investigam relações entre fala e escrita, pois se faz importante investigar em que medida características visuais dos gestos articulatórios são (ou não) mobilizadas no processo de produção textual pelo qual estudantes/escreventes estabelecem relações entre fala e escrita.

Por outro lado, também os gestos gráficos também são impregnados da dimensão fônica da linguagem. A argumentação a favor dessa conceituação de gestos gráficos é feita por Chacon (1998, p. 84) sobre as contribuições de Abaurre (1991a) e Corrêa (1994) acerca das relações entre ritmo da oralidade e da escrita. Sucintamente, Chacon (1998) argumenta sobre o ritmo na linguagem que: o ritmo da oralidade é de ordem *temporal* e o ritmo da escrita, de ordem *espacial*. O “gesto rítmico” capturaria essa dimensão temporal dos enunciados orais/falados ao representá-lo na espacialidade da escrita.

Portanto, a noção relevante para esta tese é a *constituição heterogênea da materialidade da escrita*, na medida em que dimensão fônica (acústico-articulatória) e visual (gráfico) lhe constitui, ao lado da noção da *constituição heterogênea da materialidade da fala*.

Em síntese, o conceito de “gesto rítmico” captura a *heterogeneidade da escrita* quanto à materialidade que lhe constitui. No desenvolvimento da análise das hipersegmentações no capítulo 4 (seção 4.2), defende-se ser essas grafias efeitos desses gestos rítmicos. Portanto, argumentos foram apresentados para sustentar a interpretação de que hipersegmentações sejam *lócus de materialização da heterogeneidade da escrita*.

Ainda no que diz respeito à discussão teórica, explicitou-se a proposta de articulação entre a abordagem discursiva acerca dos conceitos de escrita e da relação entre fala e escrita e a abordagem formal dada à organização prosódica dos enunciados. Partindo da premissa, que caracteriza a teoria da Fonologia Prosódica, de que haja uma hierarquia de constituintes prosódicos que organizam os enunciados, para explicitar a complexidade estrutural que subjaz a configurações de acentos e de tons da língua. Vali-me dessa abordagem formal da prosódia para explicitar o papel de estruturas linguísticas (que se constituem da interface entre fonologia e morfossintaxe) ao serem investigadas relações entre prosódia e hipersegmentações de palavras e, mais amplamente, relações entre fala e escrita. Argumentei que essas estruturas linguísticas constituem os enunciados sejam eles falados ou escritos.

Por fim, na última seção do capítulo 4, em que foram analisados hipersegmentações, textos em que essas ocorrem e propostas de produção textual, demonstrei a relação entre segmentação de palavras e organização prosódica dos enunciados a partir da consideração das configurações rítmica e entoacional que se constroem no texto produzido pelo aluno. Em última instância, por meio das análises de hipersegmentações que se particularizam quanto às suas características formais, foram promovidas reflexões sobre possíveis relações entre prosódia e escrita, explicitando-se a complexidade linguístico-discursiva que subjaz a essas relações.

Ao fim do capítulo 4, foi demonstrado haver relação entre prosódia e escrita, presente na constituição da escrita alfabética, que se atualiza em cada produção escrita e atravessa os sujeitos da linguagem. Argumentei que, por meio das hipersegmentações, identificam-se tensões entre possibilidades de representação do verbal e do verbo-visual da língua(gem) (CORRÊA, 2006).

De modo mais detalhado, é demonstrado que as hipersegmentações (e as segmentações não-convencionais de palavra, de modo geral) não apenas indiciam

vínculos entre enunciados orais/falados e letrados/escritos, mas também revelam como se dá a espacialização da linguagem, um processo de representação simbólica que é próprio da escrita. A partir dessa constatação e com base na análise linguístico-discursiva feita sobre os textos produzidos e as propostas de produção textual, conclui que hipersegmentações são efeitos da construção de sentidos mobilizados (e enfatizados) pelos discursos que atravessam e constituem o texto (e o sujeito) e que se concretizam na escrita por meio de representação de características prosódicas dos enunciados (notadamente sobre configuração rítmica e entoacional) em confluência com aspectos morfológicos (seja sobre a identificação de palavras lexicais e gramaticais, seja sobre o funcionamento dessas palavras no enunciado).

Essa conclusão implica “ajustar as lentes” de análise das hipersegmentações para relações textuais-discursivas que constituem a produção escrita, não sendo suficiente, embora seja relevante – como demonstrado – lançar mão de hipóteses que se circunscrevam a mobilizar características prosódicas e morfossintáticas das palavras e dos enunciados em que os espaços não-convencionais são registrados. Ao ser ampliada a perspectiva de análise, necessariamente entrou em cena o sujeito escrevente que se revela na constituição de diferentes discursos que se materializam no texto, como argumentado.

Para as áreas de Educação e Saúde, com as quais buscamos dialogar desde o início desta tese, essa conclusão implica considerar as produções escritas de alunos/pacientes não como produtos de relações dicotômicas entre fala e escrita, mas como registros de processos de enunciação por meio dos quais os sujeitos da linguagem (sejam professor e aluno, sejam fonoaudiólogo e paciente) se constituem nas diversas e complexas práticas orais/faladas e práticas letradas/escritas.

Referências:

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

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

ABAURRE, M. B. M. 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, 1991a.

ABAURRE, M. B. M. Ritmi dell'oralità e ritmi dela scrittura. In: ORSOLINI, M. e PONTECORVO, C. **La costruzione del texto scritto nei bambini**. Roma: La Nuova Italia, 1991b.

ABAURRE, M. B. M. Os Estudos Linguísticos e a aquisição da escrita. In: CASTRO, M. F. C. P. (Org.). **O método e o dado no estudo da linguagem**. Campinas: Editora da Unicamp, 1996, v. 1, p. 111-163.

ABAURRE, M. B. M. et alii. Leitura e escrita na vida e na escola. In: **Leitura: Teoria & Prática**, Campinas, n. 6, 1984, p. 15-26.

ABAURRE, M. B. M.; FIAD, R. S.; MAYRINK-SABINSON, M. L. T. **Cenas de Aquisição da Escrita**: o trabalho do sujeito com o texto. Campinas: Mercado de Letras, 1997.

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

ABREU-ZORZI, T. H. **O estatuto prosódico de advérbios em -mente**: um estudo comparativo entre português arcaico e português brasileiro. 2016, 228f. Tese (Doutorado em Linguística e Língua Portuguesa) - Universidade Estadual Paulista Júlio de Mesquita Filho, Araraquara, 2016.

ACIOLI, V. L. C. **A escrita no Brasil Colônia**: um guia para leitura de documentos manuscritos. Recife: Fundação João Nabuco. Editora Massangana, 2ª ed. 2003.

ALMEIDA, L. C; DALBEN, A; FREITAS, L. C. O IDEB: Limites de ilusões de uma política educacional. **Educação & Sociedade**, Campinas, v. 34, n.125, p. 1153-1174, 2013. Disponível em <http://www.cedes.unicamp.br>. Acesso em: 30 jan. 2016.

AMÂNCIO, L. Género: representações e identidades. *Sociologia: Problemas e Práticas*. n. 14, p. 127-240, 1993. Disponível em <<http://sociologiapp.iscte.pt/fichaartigo.jsp?pkid=261>> Acesso em: 16 fev. 2016.

ANDRADE, C. D. **As palavras que ninguém diz**. Rio de Janeiro: Record, 1997.

ARAÚJO, G. A. et al. As proparoxítonas e o sistema acentual do Português. In: ARAÚJO, G. A. (org.) **O acento em Português: abordagens fonológicas**. São Paulo: Parábola Editorial, 2007, p. 37-60.

BAKHTIN, M. **Estética da criação verbal**. Trad. Maria Ermantina Galvão Gomes Pereira, revisão da tradução Marina Appenseller. São Paulo: Martins Fontes, 1992 [1979].

BALIEIRO JR., A. P, Psicolinguística. In: MUSSALIN, F.; BENTES, A. C. **Introdução à linguística: domínios e fronteiras**. São Paulo: Cortez, 2001, p.171-201.

BECHARA, E. **Moderna gramática portuguesa**. Rio de Janeiro: Lucerna, 37ª ed. 2001.

BISOL, L. O acento e pé métrico binário. **Cadernos de Estudos Linguísticos**, Campinas, v. 22, p. 69-80, 1992.

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

BISOL, L. Mattoso Câmara Jr. e a Palavra Prosódica. **DELTA**, São Paulo. v. 20, n. especial, p. 59-70, 2004. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502004000300006> Acesso em: 08 jul. 2014.

BISOL, L. O clítico e o seu hospedeiro. **Letras de Hoje**, Porto Alegre, v. 40, n. 3, p. 163-184, 2005. Disponível em: <<http://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/13700>> Acesso em: 08 jul. 2014.

BISOL, L. O acento: duas alternativas de análise. **Organon**, Porto Alegre, v. 28, p. 281-231, 2013.

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

BRASIL. **Ministério da Educação**, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Brasília, 03 jun. 2015. Disponível em: <<http://ideb.inep.gov.br/resultado/>>. Acesso em: 05 fev. 2016.

BRASIL. **Parâmetros curriculares nacionais: primeiro e segundo ciclos do ensino fundamental: língua portuguesa**. Brasília: MEC/SEF. 1997. 144p. Disponível em: <<http://portal.mec.gov.br/seb/arquivos/pdf/livro02.pdf>>. Acesso em: 22 abr. 2011.

BRASIL. **Parâmetros curriculares nacionais: terceiro e quarto ciclos do ensino fundamental: língua portuguesa**. Brasília: MEC/SEF, 1998, 106p. Disponível em: <<http://portal.mec.gov.br/seb/arquivos/pdf/livro02.pdf>> Acesso em: 07 jun. 2014.

CAGLIARI, L. C. **Elementos de fonética do português brasileiro**. São Paulo: Paulista, 2007.

CAPOVILLA, A. G. S.; CAPOVILLA, F. C. **Problemas de Leitura e Escrita**: como identificar, prevenir e remediar numa abordagem fônica. (2ª ed.). São Paulo: Memnon, 2000.

CAPOVILLA, F. C.; CAPOVILLA, A.G.S. (Org.) **Prova de Consciência Sintática (PCS)** normatizada e validada para avaliar a habilidade metassintática de escolares de 1a. a 4a. séries do ensino fundamental. São Paulo: Memnon, 2006.

CAPRISTANO, C. C. A propósito da escrita do EF I: uma reflexão sobre as segmentações não-convencionais. **Letras de Hoje**, Porto Alegre, v. 39 n. 3, p. 245-260, 2004. Disponível em: <<http://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/13918>> Acesso em: 08 jul. 2014.

CAPRISTANO, C. C. Um entre outros: a emergência da rasura na aquisição da escrita. **Linguagem em (Dis)curso**. (On-Line), v. 13, p. 667-694, 2013.

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

CAPRISTANO, C. C. **Aspectos de segmentação na escrita do EF I**. São Paulo: Martins Fontes. 2007b.

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

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

CHACON, L. Constituintes prosódicos e letramento em segmentações não-convencionais. **Letras de Hoje**, Porto Alegre, v. 39, n. 3, p. 223-232, 2004. Disponível em: <<http://revistaseletronicas.pucrs.br/fale/ojs/index.php/fale/article/view/13916/9230>> Acesso em: 08 jul. 2014.

CHACON, L. Hipersegmentações na escrita do EF I: entrelaçamento de práticas de oralidade e de letramento. **Estudos Linguísticos**, Campinas, v. 34, p. 77-86, 2005. Disponível em: <http://www.marilia.unesp.br/Home/Pesquisa/Gpel/Artigos/2005_-_Chacon_-_Hipersegmentacoes.pdf> Acesso em: 08 jul. 2014.

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

CHACON, L. Segmentações não-convencionais na escrita de pré-escolares: entrecruzamentos entre convenções ortográficas e constituintes prosódicos. In: Regina Ritter Lamprecht. (Org.). **Aquisição da linguagem**: estudos recentes no Brasil. Porto Alegre: Editora da PUC-RS, 2011, p. 251-261.

CHOMSKY, N. **Linguagem e mente**. Trad. Roberto Leal Ferreira. 3.ed. São Paulo: Editora da UNESP, 2009.

COLLARES, C. A. L.; MOYSÉS, M. A. A. A transformação do espaço pedagógico em espaço clínico (a patologização da educação). **Série Ideias**, São Paulo, n. 23, p. 25- 31, 1994. Disponível em: http://www.crmariocovas.sp.gov.br/pdf/ideias_23_p025-031_c.pdf. Acesso em 10 de jan. 2016.

COLLISCHONN, G. Acento secundários em Português. **Letras de Hoje**, Porto Alegre, v. 29, n. 4, p. 43-53, 1994.

COLLISCHONN, G. A epêntese vocálica no português do sul do Brasil: análise variacionista e tratamento pela teoria da otimalidade. **Letras de Hoje**, Porto Alegre, v. 35, n. 1, p. 285-318, 2000. <<http://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/14769/9835>> Acesso em: 08 jul. 2014.

COLLISCHONN, G. Epêntese vocálica e restrições de acento no português do sul do Brasil. **Signum. Estudos Linguísticos**, Londrina, v. 7, n. 1, p. 61-78. 2004. Disponível em: <<http://www.uel.br/revistas/uel/index.php/signum/article/viewFile/3875/3113>> Acesso em: 08 jul. 2014.

CORRÊA, M. L. G. Pontuação: sobre seu ensino e concepção. **Leitura: teoria e prática**, Campinas, v. 24, p. 52-65, 1994.

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

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

CORRÊA, M. L. G. Práticas de Leitura e escrita: breve nota sobre a relação entre o verbal e o não verbal. **Contrapontos**, Itajaí, vol. 8, n. 2, p. 293-302, 2006.

CUNHA, A. P. N. **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.

CUNHA, A. P. N. **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, A. P. N.; MIRANDA, A. R. M. Índícios de reestruturação do conhecimento fonológico da criança em dados de reparo na escrita inicial. **Letras de Hoje**, Porto Alegre, v. 48, n. 3, p. 343-354, 2013.

CUNHA, C.; CINTRA, L. F. L. **Nova gramática do português contemporâneo**. 3. ed. Rio de Janeiro: Nova Fronteira, 1989.

DESJARDINS, R.; ROGERS, J.; WERKER, J. An exploration of way pre-schoolers perform differently than do adults in audiovisual speech perception tasks. In: **Journal of experimental child psychology**, Amsterdã, v. 66, p. 85-110, 1997.

DICIONÁRIO PRIBERAM DA LÍNGUA PORTUGUESA [em linha], 2008-2013. Disponível em: <<http://www.priberam.pt/dlpo/hífen>> Acesso em: 08 fev. 2016.

DIXON, R. M. W.; AIKHENVALD, A. Y. **Word: a cross-linguistic typology**. Cambridge: Cambridge University Press, 2002.

FARACO, C. E.; MOURA, F. M. **Gramática**. 12. ed. São Paulo: Ática, 1999.

FELIPE, J. Infância, gênero e sexualidade. **Educação e realidade**, Porto Alegre, v.25, n.1, p. 115-131, 2000.

FELIPETO, C. “A Deus sapinho”: quando a homonímia produz desordens e rasuras em um processo de escritura colaborativa. **Linguagem & Ensino**, Pelotas, v. 16, p. 327-342, 2013.

FERNANDES, F. R. **Ordem, focalização e preenchimento em português: sintaxe e prosódia**, 2007. 452f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas (UNICAMP), Campinas, 2007.

FERNANDES-SVARTMAN et al. Secondary stress, intensity and fundamental frequency in Brazilian Portuguese. **Journal of Portuguese Linguistics**, Lisboa, v. 11, p. 51-68, 2012.

FERNÁNDEZ, A. Y. et al.. Avaliação e intervenção da disortografia baseada na semiologia dos erros: revisão da literatura. **Revista CEFAC**, São Paulo, v. 12, n. 3, p. 499-504, 2010. Disponível em: <<http://www.revistacefac.com.br/fasciculo.php?form=edicoes/v12n3.php>> Acesso em: 05 jul. 2013.

FERREIRA, C. R. G. **Um estudo sobre a segmentação não convencional na aquisição da escrita de alunos de EJA**. 2011. 163f. Dissertação (Mestrado em Educação) – Faculdade de Educação, Universidade Federal de Pelotas, Pelotas, 2011.

FERREIRO, E. et al. **Chapeuzinho vermelho aprende a escrever**. Estudos Psicolinguísticos comparativos em três línguas. São Paulo: Ática, 1996.

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

FERREIRO, E; TEBEROSKY, A. **Psicogênese da língua escrita**. Trad. Diana Myriam Lichtenstein, Liana Di Marco e Mário Corso. Porto Alegre: Artes Médicas Sul, 1999.

FISCHER, S. R. **História da escrita**. Trad. Mirna Pinsky. São Paulo: Editora UNESP, 2009.

FREIRE, P. **A educação na cidade**. São Paulo: Cortez, 1991.

FREITAG, R. M. K. (Re)Discutindo sexo/gênero na sociolinguística. In: FREITAG, R. M. K; SEVERO, C. G. (org.) **Mulheres, Linguagem e Poder: estudos de gênero na sociolinguística brasileira**. São Paulo: Blucher, 2015, p. 17-74.

FREITAG, R. M. K; SEVERO, C. G. (org.) **Mulheres, Linguagem e Poder**: estudos de gênero na sociolinguística brasileira. São Paulo: Blucher, 2015.

FREITAG, R. M. K. et al. Culto à avaliação, patologização da alfabetização e fracasso escolar. **Revista Fórum Identidades**, Itabaiana, v. 15, p.41-59, 2014.

FREITAS et al. Efeitos prosódicos e efeitos de frequência no desenvolvimento silábico em Português Europeu. In **XX Encontro Nacional da Associação Portuguesa de Linguística. Textos Seleccionados**. Lisboa: APL/Colibri, 397-412. 2006.

FREITAS, L. C. O “solução da curva” no SARESP 15: saindo do “volume morto”. In: Avaliação Educacional – Blog do Freitas, 2016a. Disponível em: <https://avaliacaoeducacional.com/2016/02/05/o-soluco-da-curva-no-saresp-15-saindo-do-volume-morto/>. Acesso em: 05 fev. 2016.

FREITAS, L. C. SP: Saresp sob suspeita. In: Avaliação Educacional – Blog do Freitas, 2016b. Disponível em: <https://avaliacaoeducacional.com/2016/04/14/sp-saresp-sob-suspeita/2016>. Acesso em: 30 abr. 2016.

FROTA, S. **Prosody and focus in European Portuguese**. Phonological phrasing and intonation. New York: Garland Publishing, 2000.

FROTA, S.; VIGÁRIO, M. Efeitos de peso no Português europeu. In **Saberes no tempo**: homenagem a Maria Henriqueta Costa Campos. Lisboa: Colibri, 2001, p. 315-333. Disponível em: <http://labfon.lettras.ulisboa.pt/texts/efeitpe01.pdf>. Acesso em: 21 abr. 2016.

GALVES, C.; ABAURRE, M. B. M. Os clíticos no português brasileiro: elementos para uma abordagem sintático-fonológica. In: ATALIBA TEIXEIRA DE CASTILHO. (Org.). **Gramática do Português Falado IV**. Campinas: Editora da Unicamp, 1996, v. 1, p. 273-319.

GAYER, J. E. L. Breve história dos constituintes prosódicos. **Diadorim**, Rio de Janeiro, n.17, v. 2, p.149-172, 2015.

GUZZO, N. B. **A prosodização de clíticos e compostos em Português Brasileiro**. 2015. 232f. Tese (Doutorado em Letras) – Universidade Federal do Rio Grande do Sul, Instituto de Letras, Porto Alegre, 2015.

HAYES, B. **A Metrical Theory of Stress Rules**. New York: Garland, 1985.

HOUAISS, A. **Dicionário eletrônico Houaiss da Língua Portuguesa**. CD-ROOM, 2001.

HUYSE, A.; LEYBAERT, J. Effects of aging on audio-visual speech integration Effects of aging on audio-visual speech integration. **The journal of the acoustical society of America**, Nova York, n. 136, p. 1918-1931, 2014. Disponível em: <http://dx.doi.org/10.1121/1.4894685>. Acesso em 17 abr. 2016.

KOCH, I. V. **A inter-ação pela linguagem**. São Paulo, Contexto, 1992.

KOCH, I. V. **O texto e a construção de sentidos**. São Paulo, Contexto, 1997.

LEE, S. H. **Morfologia e fonologia lexical do Português do Brasil**. 1995. 201f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas (UNICAMP), Campinas, 1995.

LEMOS, C. Teorias da Diferença e Teorias do Déficit: reflexões sobre programa de intervenção na pré-escola e na alfabetização. In: **Anais do Seminário Multidisciplinar de Alfabetização**, 11 a 13 de agosto de 1983. Brasília: INEP, 1984, p. 133-145. Disponível em: http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&coobra=27629. Acesso em 10 de maio de 2016.

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

LOPES, S. M. M. **Análise do estatuto da palavra prosódica no Português Brasileiro na interação entre constituintes prosódicos e morfológicos**. 2016. 220f. Tese (Doutorado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2016.

LOURO, G. L. Gênero e sexualidade: pedagogias contemporâneas. **Pro-Posições**, Campinas, v.19, n. 56, p. 7-23, 2008. Disponível em <<http://www.scielo.br/pdf/pp/v19n2/a03v19n2>>. Acesso em: 16 fev. 2016.

LURIA, A.R. O desenvolvimento da escrita na criança. In: VYGOTSKY, L. S.; LURIA, A. R.; LEONTIEV, N. A. **Linguagem, desenvolvimento e aprendizagem**. Trad. Maria da Penha Villalobos. São Paulo, Íconen/EDUSP, 1988, p. 143-189.

MAN, K. et al. Sight and sound converge to form modality-invariant representations in temporoparietal Cortex. In: **The journal of neuroscience**, Baltimore, v. 32, n. 47, p. 16629-16636, 2012.

MARCUSCHI, L. A. Oralidade e letramento. In: _____. **Da fala para a escrita: atividades de retextualização**. São Paulo: Cortez, p. 15-43, 2001.

MARCUSCHI, L. A. Gênero textuais emergentes no contexto da tecnologia digital. In: MARCUSCHI, L. A.; XAVIER, A. C. (org.) **Hipertexto e gêneros digitais**. Rio de Janeiro: Lucerna, 2004, p. 13-67.

MICHAELIS MODERNO DICIONÁRIO DA LÍNGUA PORTUGUESA. Disponível em: <<http://michaelis.uol.com.br/moderno/portugues/index.php?lingua=portugues-portugues&palavra=hífen>> Acesso: 08 fev. 2016.

MOTA, M. M. P. E. da et al. Relação entre consciência morfológica e leitura contextual medida pelo teste de Cloze. **Psicologia: Reflexão e Crítica**, Porto Alegre, v. 22, n.2, p. 223-229, 2009.

NESPOR, M. ; VOGEL, I. **Prosodic phonology**. Dordrecht: Foris Publications, 1986.

NESPOR, M.; VOGEL, I. **Prosodic phonology: with a new foreword**. Berlin; New York: Mouton de Gruyter. 2007.

NOBILE, G. G.; BARRERA, S. D. Análise de erros ortográficos em alunos do ensino público fundamental que apresentam dificuldades na escrita. **Psicologia em Revista**, Belo Horizonte. v. 15, n. 2, p. 36-55, 2009. Disponível em: <<http://periodicos.pucminas.br/index.php/psicologiaemrevista/article/viewFile/P.1678-9563.2009v15n2p36/868>> Acesso em: 08 jul. 2014.

OCHS, E. Planned and unplanned discourse. In: GIVÓN, T. (ed.) **Discourse and syntax**. New York, Academic Press, 1979, p. 51-80.

OLIVEIRA, K. **Negros e escrita no Brasil do século XIX**: sócio-história, edição filológica de documentos e estudo linguístico. 2006. 503f. Tese (Doutorado em Letras e Linguística) – Instituto de Letras, Universidade Federal da Bahia (UFBA), Salvador, 2006.

OLSON, D. R. From utterance to text. The bias of language in speech and writing. **Harvard Educational Review**, Cambridge, v. 47, n. 3, p. 258-281, 1977.

PARANHOS, F. C. **Segmentações não-convencionais de palavra nos quatro últimos anos do ensino fundamental**: um estudo longitudinal. 2014. 169f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual Paulista Júlio de Mesquita Filho, Fundação de Amparo à Pesquisa do Estado de São Paulo, 2014.

PAULA, I. F. V. **Movimentos na escrita inicial de crianças**: um estudo longitudinal de hipersegmentações. São José do Rio Preto, 2007. 132f. Dissertação (Mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, 2007.

PISA EM FOCO. Meninos e meninas estão preparados para a era digital?, traduzido pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Organização para Cooperação e Desenvolvimento Econômico, Nº 01, 2012. Disponível em: <http://download.inep.gov.br/acoes_internacionais/pisa/pisa_em_foco/2011/pisa_em_foco_n12.pdf>. Acesso em: 10 jan. 2016.

PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO, ÍNDICE DE DESENVOLVIMENTO HUMANO. Brasília, 2016, Disponível em: <<http://www.pnud.org.br/IDH/DH.aspx>> Acesso em: 12 jan. 2016.

PROGRAMA INTERNACIONAL DE AVALIAÇÃO DE ESTUDANTES (PISA). Disponível em: <<http://portal.inep.gov.br/pisa-programa-internacional-de-avaliacao-de-alunos>>. Acesso em: 15 jan. 2016.

RAMOS, A. P.; TENANI, L. Análise métrica do apagamento das vogais postônicas não-finais no dialeto do noroeste paulista. **Estudos Linguísticos**, São Paulo, v. 38, p. 21-34, 2009. Disponível em: <http://gel.org.br/estudoslinguisticos/volumes/38/EL_V38N1_02.pdf> Acesso em: 08 jul. 2014.

RAMOS, P. **A leitura dos quadrinhos**. São Paulo: Contexto, 2009.

RIEMSDIJK, H. van (ed.) **Clitics in the languages of Europe**. Belin; New York: Mouton de Gruyter, 1999.

REIS, M. C.; TENANI, L. **Registros da heterogeneidade da escrita**: um olhar para as grafias não-convencionais de vogais pretônicas. São Paulo: Cultura Acadêmica/Editora da UNESP, 2011.

SANTAELLA, L. **Comunicação ubíqua**: repercussões na cultura e na educação. São Paulo: Paulus, 2013.

SÃO PAULO (ESTADO) SECRETARIA DA EDUCAÇÃO. **Índice de desenvolvimento da educação de São Paulo**, São Paulo, 2015. Disponível em: <<http://www.educacao.sp.gov.br/idesp>>. Acesso em: 28 jan. 2016.

SÃO PAULO (ESTADO) SECRETARIA DA EDUCAÇÃO. **Matrizes de referências para a avaliação Saresp**: documento básico. Coord. Geral: Maria Inês Fini. São Paulo: Secretaria da Educação, 2009.

SÃO PAULO (ESTADO) SECRETARIA DA EDUCAÇÃO. **Proposta Curricular do Estado de São Paulo**, São Paulo: Secretaria da Educação, 2008.

SAUSSURE, F. **Curso de linguística geral**. (Organizado por Charles Bally e Albert Sechehaye, com colaboração de Albert Riedlinger). Tradução: Antonio Chelini, José Paes e Izidoro Blikstein. 18. ed. São Paulo: Cultrix, 1973.

SCARPA, E. M. Apresentação. In: ____ (Org.) **Estudo de prosódia**. Campinas: Editora da Unicamp, 1999, p. 7-20.

SCARPA, E. Aquisição da linguagem. In: MUSSALIN, F.; BENTES, A. C. **Introdução à linguística**: domínios e fronteiras. São Paulo: Cortez, 2001, p.203-232.

SCHWINDT, L. C. **O prefixo no Português Brasileiro**: análise morfofonológica. 2000. 191f. Tese (Doutorado em Letras) – Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2000.

SCLIAR-CABRAL, L. **Princípios do sistema alfabético do português do Brasil**. São Paulo: Contexto, 2003.

SCLIAR-CABRAL, L. Avanços das neurociências para a alfabetização e leitura. **Letras de Hoje**, Porto Alegre, v.48, n.3, p.277-282, 2013.

SEKIYAMA, K.; BURNHAM, D. Issues in the development of auditory-visual speech perception: Adults, infants, and children. In: **Eighth International Conference on Spoken Language Processing**, p. 1137–1140, 2004.

SELKIRK, E. Prosodic domains in phonology: sanskrit revisited. In: ARONOFF, M.; KEAN, D. M. L. (eds) **Juncture**. Saratoga, Califórnia: Anma Libri, 1980.

SELKIRK, E. On the nature of phonological representation. In: J. Anderson, J. Laver, and T. Meyers (eds) **The cognitive representation of speech**. Amsterdam: North Holland, 1981.

SELKIRK, E. O. **Phonology and syntax**. The relation between sound and structure. Cambridge, London: The MIT Press, 1984.

SELKIRK, E. The interaction of constraints on prosodic phrasing. In: HORNE, M. (ed.). **Prosody: theory and experiment**. Dordrech, Boston, London: Kluwer Academic Publishers, 2000, p.231-261.

SELKIRK, E. The prosodic structure of function words. In: McCARTHY, J. J. **Optimality Theory in phonology: a reader**. Oxford: Blackwell, 2004. p. 464-482.

SEMINÁRIO MULTIDISCIPLINAR DE ALFABETIZAÇÃO (São Paulo, 11 a 13 de agosto de 1983). **Anais...** Brasília: INEP, 1984. 158p. Disponível em: http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&coobra=27629. Acesso em: 10 de set. 2016.

SERRA, C. R. **Realização e percepção de fronteiras prosódicas no português do Brasil: fala espontânea e leitura**. 2009. 241f. Tese (Doutorado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2009.

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

SILVA, L. M. A Palavra Fonológica, o Grupo Clítico e a Segmentação não-convencional de palavras. In: **XX Seminário do CELLIP**, 2011, Londrina. Anais do ... Seminário do Centro de Estudos Linguísticos e Literários do Paraná, p. 01-12, 2011.

SILVA, L. M. **Um estudo longitudinal das hipersegmentações de palavras escritas nos anos finais do Ensino Fundamental**. 2014. 171f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual Paulista Júlio de Mesquita Filho, Fundação de Amparo à Pesquisa do Estado de São Paulo, São José do Rio Preto, 2014.

SILVA, L. M.; TENANI, L. **Hipersegmentações de palavras no ensino fundamental**. São Paulo: Cultura Acadêmica, 2014.

SIMIONI, T. O clítico e seu lugar na estrutura prosódica em Português Brasileiro. **ALFA**, São Paulo, v. 52, n.2, p. 431-446, 2008. Disponível em: www.alfa.ibilce.unesp.br. Acesso em: 12 dez. 2015.

SONCIN, G.; LONGHIN, S. R. A causalidade de “porque” em textos escolares: domínios de atuação, gêneros de produção. **Revista Linguagem & Ensino** (On-line), v. 17, p. 525-549, 2014.

SPENCER, A.; LUÍS, A. R. **Clitics: an introduction**. Cambridge: Cambridge University Press, 2012.

STROMQUIST, N. Convergência e divergência na conexão entre gênero e letramento: novos avanços. **Educação e Pesquisa**, São Paulo, v.27, n.2, p.301-320, 2001. Disponível em <<http://www.scielo.br/pdf/ep/v27n2/a08v27n2.pdf>> Acesso em: 16 fev. 2016.

TENANI, L. **Domínios prosódicos no Português do Brasil: implicações para a prosódia e para a aplicação de processos fonológicos**. 2002. 331f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas (UNICAMP), Campinas, 2002.

TENANI, L. Segmentações não convencionais e teorias fonológicas. **Letras de hoje**. Porto Alegre, v. 39, n. 3, setembro, p. 233-244, 2004.

TENANI, L. A segmentação não convencional de palavras: evidências de representações de relações entre enunciados falados e escritos. **Anais do II SIMELP** “Língua portuguesa: ultrapassar fronteiras, juntar culturas”, p. 107-127, 2009.

TENANI, L. A grafia dos erros de segmentação não-convencional de palavras. **Cadernos de Educação**, Pelotas, v. 35, p. 247-269, 2010. Disponível em: <<http://www2.ufpel.edu.br/fae/caduc/downloads/n35/09.pdf>> Acesso em: 08 jul. 2014.

TENANI, L. Letramento e segmentações não-convencionais de palavras. In: TFOUNI, L. V. (Org.). **Letramento, escrita e leitura: Questões Contemporâneas**. Campinas: Mercado de Letras, p. 229-243, 2011a.

TENANI, L. A segmentação não-convencional de palavras em textos do ciclo II do ensino fundamental. **Revista da Associação Brasileira de Linguística**, v. 10, n. 2, p.91-119, 2011b. Disponível em: <<http://ojs.c3sl.ufpr.br/ojs/index.php/abralin/article/viewFile/32061/20400>> Acesso em: 08 jul. 2014.

TENANI, L. Hipersegmentação de palavras: análise de aspectos prosódicos e discursivos. **Linguagem & Ensino**, Pelotas, v. 16, p. 305-324, 2013. Disponível em: <<http://www.rle.ucpel.tche.br/index.php/rle/article/view/884>>. Acesso em: 08 jul. 2014.

TENANI, L. Banco de dados de escrita do Ensino Fundamental II. FAPESP/UNESP, 2015. Disponível em: <<http://www.convenios.grupogbd.com/redacoes/Login>> Acesso em: 10 nov. 2015.

TENANI, L.; FIEL, R. P. Hipersegmentação de palavras em textos do EF II: características prosódicas gerais. **Linguasagem** (São Carlos), v. 26 (2): 2016. Disponível em: <http://www.linguasagem.ufscar.br/index.php/linguasagem/issue/view/12> Acesso em: 09 de set. 2016.

TENANI, L.; LONGHIN-THOMAZI, S. Oficinas de leitura, interpretação e produção textual no ensino fundamental. **Em Extensão**. Universidade Federal de Uberlândia, Uberlândia, v. 13, n. 1, 2014.

TENANI, L.; LONGHIN, S. Flutuação gráfica entre “senão” e “se não”: considerações sobre polissemia, constituição morfossintática e prosódica. **Gragoatá**, Niterói, v.20, p.138-206, 2015.

TENANI, L.; OLIVERIRA-CODINHOTO, G. Hipersegmentações de “demais” no EF II. (Inédito).

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

TORQUETTE, A. O. **Segmentações não-convencionais de palavras em escolas públicas e privadas**, 2016. 139f. Tese. (Doutorado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2016.

VELOSO, J. **Verba manent**: a palavra como unidade pertinente para a descrição linguística do português e de outras línguas flexionais. Porto Alegre: EDIPUCRS, 2016.

VERÓN, E. **A produção de sentido**. Trad. Alceu Dias Lima. São Paulo, Cultrix/EDUSP, 1981.

VIANNA, C.; UNBEHAUM, S. O gênero nas políticas públicas de educação no Brasil: 1988-2002. **Cadernos de Pesquisa**, São Paulo, v.34, n.121, p. 77-104, 2004. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-15742004000100005> Acesso em: 14 fev. 2016.

VIANNA, C.; UNBEHAUM, S. Gênero na educação básica: quem se importa? Uma análise de documentos de políticas públicas no Brasil. **Educação e Sociedade**, Campinas, vol.27, n.95, p.407-428, 2006. Disponível em <<http://www.scielo.br/pdf/es/v27n95/a05v2795.pdf>>. Acesso em: 12 fev. 2016.

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

VIGÁRIO, M. O lugar do grupo clítico e da palavra prosódica composta na hierarquia prosódica: uma nova proposta. In: LOBO, M. & COUTINHO, M. A. (orgs). **Actas do XII Encontro Nacional da Associação Portuguesa de Linguística – Textos seleccionados**. Lisboa: Colibri Artes Gráficas, 2007, p. 673-688.

VYGOSTSKY, L. S. **A formação social da mente**: o desenvolvimento dos processos psicológicos superiores. São Paulo, Martins Fontes, 1988.

VILLALVA, A.; SILVESTRE, J. P. **Introdução ao estudo do léxico**. Descrição e análise do Português. Petrópolis: Vozes, 2014.

ZORZI, J. L. **A apropriação do sistema ortográfico nas 4 primeiras séries do primeiro grau**. Campinas, 1997. 124f. Tese (Doutorado em educação) – Faculdade de Educação, Universidade Estadual de Campinas (UNICAMP), Campinas, 1997.

ZORZI, J. L. Alterações ortográficas nos transtornos de aprendizagem. In: MALUF, M. I. (org.). **Aprendizagem**: tramas do conhecimento, do saber e da subjetividade. Rio de Janeiro: Vozes; São Paulo: ABP. p. 144-162, 2006.

Anexos

Coletânea textual e proposta “Amizade e internet” (TENANI, 2015)



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

Projeto de extensão: Oficinas de Leitura, Interpretação e Produção Textual

E.E. Professora Zulmira da Silva Salles

Nome: _____

Série/Turma: 8ª _____ Nº _____

Proposta 1

Data: _____

- Leia os textos abaixo sobre amizade.

Bons Amigos

Machado de Assis

Abençoados os que possuem amigos, os que os têm sem pedir.
Porque amigo não se pede, não se compra, nem se vende.
Amigo a gente sente!

Benditos os que sofrem por amigos, os que falam com o olhar.
Porque amigo não se cala, não questiona, nem se rende.
Amigo a gente entende!

Benditos os que guardam amigos, os que entregam o ombro pra chorar.
Porque amigo sofre e chora.
Amigo não tem hora pra consolar!

Benditos sejam os amigos que acreditam na tua verdade ou te apontam a realidade.
Porque amigo é a direção.
Amigo é a base quando falta o chão!

Benditos sejam todos os amigos de raízes, verdadeiros.
Porque amigos são herdeiros da real sagacidade.
Ter amigos é a melhor cumplicidade!

Há pessoas que choram por saber que as rosas têm espinho,
Há outras que sorriem por saber que os espinhos têm rosas!

Como a internet está mudando a amizade

Qual é a primeira coisa que você faz quando entra na internet? Checa seu *email*, dá uma olhadinha no Twitter, confere as atualizações dos seus contatos no Orkut ou no Facebook? Há diversos estudos comprovando que interagir com outras pessoas, principalmente com amigos, é o que mais fazemos na internet. Só o Facebook já tem mais de 500 milhões de usuários, que juntos passam 700 bilhões de minutos por mês conectados ao *site* – que chegou a superar o Google em número de acessos diários. A internet é a ferramenta mais poderosa já inventada no que diz respeito à amizade. E está transformando nossas relações: tornou muito mais fácil manter contato com os amigos e conhecer gente nova. Mas será que as amizades *online* não fazem com que as pessoas acabem se isolando e tenham menos amigos *offline*, “de verdade”? Essa tese, geralmente citada nos debates sobre o assunto, foi criada em 1995 pelo sociólogo americano Robert Putnam. E provavelmente está errada. Uma pesquisa feita pela Universidade de Toronto constatou que a internet faz você ter mais amigos – dentro e fora da rede.

Camila Costa

Fonte: *Superinteressante*, Edição 288, Fev/2011, p. 55.

Projeto de extensão: **Oficinas de Leitura, Interpretação e Produção Textual**

E. E.

Nome: _____

Proposta 1

Série/Turma: 8ª _____ Nº _____

Data: _____

- No poema, Machado de Assis declama a alegria de se ter um grande amigo. No texto da revista *Superinteressante*, a internet é apresentada como uma ferramenta que ajuda a ter mais amizades.
- Assim como o autor do poema, provavelmente, você deve ter um grande amigo(a) com quem gosta de passear, conversar e até estudar juntos(as). Escreva um texto, **relatando como surgiu essa amizade**, se usa ou não a internet para se manter contato com seu(sua) amigo(a), e se houve algum bom momento que vocês viveram juntos(as).
- Seu texto deve conter de 25 a 30 linhas e deve ser escrito à caneta. Seu texto não deve ultrapassar os limites designados para a escritura. Se quiser, dê um título a seu texto.