



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

**Ana Carolina Teodoro Borsato**

**Segmentações não convencionais no EF II:**

há regularidades linguísticas em dados frequentes?

São José do Rio Preto

2023

**Ana Carolina Borsato**

**Segmentações não convencionais no EF II:**

há regularidades linguísticas em dados frequentes?

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Estudos Linguísticos junto ao Programa de Pós-Graduação em Estudos Linguísticos, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto.

Financiadora: CAPES

Orientadora: Profa. Dra. Luciani Ester Tenani

São José do Rio Preto

2023

B738s      Borsato, Ana Carolina Teodoro  
Segmentações não convencionais no EF II : há regularidades  
linguísticas em dados frequentes? / Ana Carolina Teodoro Borsato. --  
São José do Rio Preto, 2023  
104 f.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp),  
Instituto de Biociências Letras e Ciências Exatas, São José do Rio  
Preto  
Orientadora: Luciani Ester Tenani

1. Linguística. 2. Ensino Fundamental. 3. Palavra (Linguística). 4.  
Segmentações. 5. Escrita. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de  
Biociências Letras e Ciências Exatas, São José do Rio Preto. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

### **Impacto potencial desta pesquisa**

Este trabalho tem impacto social, pois, além da realidade escolar nos mostrar que as dúvidas sobre segmentações não são sanadas até o fim do Ensino Fundamental I, como visam documentos oficiais da Educação, pode contribuir para que profissionais da educação repensem práticas pedagógicas e considerem questões linguísticas relevantes para o ensino.

### **Potential impact of this research**

This work has a social impact, because, in addition to the school reality showing us that doubts about segmentation are not solved until the end of Elementary School I, as official Education documents aim, it can contribute to education professionals rethinking pedagogical practices and considering language issues relevant to teaching.

**Ana Carolina Borsato**

**Segmentações não convencionais no EF II:**

há regularidades linguísticas em dados frequentes?

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Estudos Linguísticos junto ao Programa de Pós-Graduação em Estudos Linguísticos, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto.

Financiadora: CAPES

Orientadora: Profa. Dra. Luciani Ester Tenani

Comissão Examinadora

Profa. Dra. Luciani Ester Tenani, Unesp/SJRP - orientadora

Profa. Dra. Lúcia R. Lopes-Damasio, Unesp/Assis - membro interno

Prof. Dra. Cristiane Capistrano, UEM - membro externo

São José do Rio Preto

2 de março de 2023

## AGRADECIMENTOS

Esta dissertação é resultado da contribuição de diversas pessoas que foram importantes em diferentes etapas deste processo. Por isso, deixo o meu agradecimento a todos que fizeram parte desta importante etapa.

Primeiramente, agradeço imensamente a minha professora e orientadora Luciani Tenani, que é a peça fundamental aqui. Sou grata pela enorme contribuição em minha formação, profissional e pessoal, e pelos cuidados e conselhos durante todos esses anos de orientação, principalmente nesses tempos pandêmicos, que foram repletos de desafios profissionais e pessoais. Agradeço, ainda, pela compreensão e pelo acolhimento até mesmo nos momentos em que falhei como aluna, como orientanda e como pessoa. Obrigada por todo o ensinamento e conhecimento compartilhado!

Deixo meus agradecimentos também a todos os meus familiares, mas em especial a minha mãe, Maria Teodoro, e ao meu irmão, José Antônio Borsato, que me incentivam e me apoiam desde sempre e em todos os sentidos. Obrigada por todo o amor, carinho e acolhimento. Minhas conquistas são graças a vocês e também de vocês!

Agradeço aos meus avós (*in memoriam*) Benedito Teodoro, José Borsato e Elvira Borsato, que, embora, infelizmente, tenham partido durante os anos em que esta pesquisa caminhava, sempre me incentivaram e compartilharam suas sabedorias comigo, ajudando-me a tornar quem sou hoje. Gostaria de poder dividir este momento com vocês, mas, apesar dessa impossibilidade, deixo aqui registrado o meu eterno agradecimento e a minha eterna saudade. Por fim, agradeço também a minha avó Jandira Teodoro, que ainda me acompanha nesta jornada. Obrigada!

Manifesto também o meu agradecimento à professora Cristiane Carneiro Capristano e à professora Lúcia Damasio pelas inestimáveis contribuições na qualificação e na defesa. Também agradeço à professora Fabiana Komesu e à professora Geovana Soncin.

Meus sinceros agradecimentos aos meus queridos amigos e professores da pós-graduação em Estudos Linguísticos. Meu obrigada também aos meus amigos de graduação Isabela Vogliotti, Guilherme Lopes, Bianca Ribeiro, Ângela Soares, Luiz Pacheco e Guilherme Torres. Deixo um agradecimento especial também ao Fernando E. Ferreira dos Santos, que tanto me ajudou e aconselhou durante toda a minha trajetória acadêmica.

Agradeço também a minha colega de mestrado Isadora Camillo, com quem não só compartilhei esses anos de pesquisa, mas também muitos conselhos.

Agradeço ao Frederico de Luca, ao Guido Assoni, à Isabela Guiduci, à Vanessa Justiniana e à Maria Isabel Donini que, apesar de não terem participado diretamente na elaboração desta pesquisa, proporcionaram, e ainda proporcionam, bons momentos de descontração. Deixo um cumprimento especial também à Álicie Camargo e à Camila Rijandis, que compartilham a vida e os dias comigo.

Por fim, agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) [Código de Financiamento 001 - Processo: 88887.488938/2020-00] pelo financiamento desta pesquisa, o que possibilitou a minha dedicação à produção deste trabalho

## RESUMO

Nesta dissertação, analisamos as chamadas segmentações não convencionais de palavra, que podem ser de três tipos: as hipossegmentações, que ocorrem quando estão ausentes as fronteiras gráficas entre palavras, como em “porisso” e “oque”; as hipersegmentações, quando estão presentes fronteiras gráficas dentro de uma palavra ortográfica, como em “que ria” e “morava-mos”; e os híbridos, que se caracterizam pela ocorrência simultânea de dois tipos de segmentação não convencional de palavras (hipossegmentação e hipersegmentação) em uma sequência de duas palavras, como em “po risso” e “dere pente”. As grafias aqui analisadas foram identificadas em um conjunto de textos escritos por alunos do Ensino Fundamental II (EF II), categorizados, de uma perspectiva estatística, como *outliers*, segundo Fiel (2018), por produzirem uma alta quantidade de hipossegmentações em relação aos demais sujeitos do Banco de Dados de Escrita do EF II. Nesta dissertação, selecionamos para análise as segmentações não convencionais desses alunos, as quais não foram analisadas por Fiel (2018), com o objetivo central de caracterizar essas grafias quanto aos aspectos prosódicos dos enunciados falados. Para alcançar esse objetivo, nos baseamos na Fonologia Prosódica, de Nespor e Vogel (1996), que concebe a existência de sete constituintes prosódicos que estruturam os enunciados das línguas do mundo, e na abordagem da escrita como constituída de modo heterogêneo, proposta por Corrêa (2004). Dos resultados obtidos na análise, constatamos que informações linguísticas de natureza prosódica e da própria convenção ortográfica estão envolvidas nas ocorrências de um limite não convencional, o que corrobora estudos anteriores sobre o mesmo tema. Avançamos ao mostrar que as características das grafias dos sujeitos denominados *outliers* vão ao encontro do que é descrito para os demais sujeitos do Banco de Dados de Escrita do EF II. Identificamos um sujeito que, apesar de se diferenciar das tendências gerais encontradas, possui características dos primeiros anos de escolarização ao apresentar, em algumas circunstâncias, recursos gráficos que demarcam palavras como formas de destacar sentidos privilegiados do texto e de informações linguísticas que permeiam as regularidades observadas. Em suma, os dados de segmentação não convencional de palavra fornecem evidências gerais e particulares a respeito da noção de palavra escrita mobilizada pelos escreventes, evidenciando as marcas de trânsito do escrevente por práticas de uso da linguagem, oral/falada e letrada/escrita, em que se inserem (CORRÊA, 2004).

**PALAVRAS-CHAVE:** Segmentação. Não convencional. Palavra. Prosódia. Ortografia. Oralidade. Letramento.



## ABSTRACT

In this dissertation, we analyze unconventional word segmentations, which can be of three types: hyposegmentations, defined by the lack of boundaries between words, such as in “porisso” and “oque”; hypersegmentations, which are characterized by the presence of boundaries within a word, as in “que ria” and “morava-mos”; and hybrids, which are characterized by the simultaneous occurrence of two types of unconventional word segmentation in a sequence of two words (hypossegmentation and hypersegmentation), as in “por risso” and “dere pente”. The unconventional word segmentations analyzed were identified in texts written by students from Junior High School (EF II), in Brazil, that were categorized, from a statistical perspective, as *outliers*, according to Fiel (2018), because they produce a high amount of hyposegmentations in comparison to the other subjects of the EF II Writing Database. In this work, we selected for analysis the unconventional segmentations of these students, which were not analyzed by Fiel (2018), aiming to characterize these datas in terms of the prosodic aspects of the utterances. To achieve this goal, we based ourselves on Prosodic Phonology, by Nespor and Vogel (1996), which conceives the existence of seven prosodic constituents that structure the utterances of the languages of the world, and on the approach of writing as heterogeneity, proposed by Corrêa (2004). The main results of this analysis we found are that linguistic information of a prosodic nature and the orthographic convention were important for the occurrence of spaces between syllables in a certain position of the word, which corroborates previous studies on the same subject. We advanced by showing that the characteristics of the datas of the subjects called *outliers* are in line with what is described for the other subjects in the EF II corpus. We identified a subject who, despite differing from the general trends found, continues to mobilize hypotheses characteristic of the first years of schooling by making use, in some circumstances, of graphic resources that demarcate words as ways of highlighting privileged meanings of the text and linguistic information that permeate the observed regularities. In short, the unconventional word segmentation data provide general and particular clues regarding the notion of the written word mobilized by the writers, evidencing the marks of the transit of the writer through practices of language use, oral/spoken and literacy/written, in that are inserted (CORRÊA, 2004).

**KEYWORDS:** Segmentation. Unconventional. Word. Prosody. Orthography. Orality. Literacy.

## LISTA DE FIGURAS

|                                                                                             |    |
|---------------------------------------------------------------------------------------------|----|
| Figura 1. Exemplo de hipersegmentação                                                       | 17 |
| Figura 2. Exemplo de hierarquia prosódica                                                   | 19 |
| Figura 3.1. Exemplo de critério de identificação de dados de segmentação não convencional   | 39 |
| Figura 3.2. Exemplo 2 de critério de identificação de dados de segmentação não convencional | 40 |
| Figura 4. Exemplo de hipersegmentação de palavra homônima                                   | 42 |
| Figura 5. Exemplo de uso não convencional de hífen                                          | 42 |
| Figura 6. Número identificador dos escreventes                                              | 43 |
| Figura 7.1. Exemplo 1 de dificuldade de interpretação de texto                              | 44 |
| Figura 7.2. Exemplo 2 de dificuldade de interpretação de texto                              | 45 |
| Figura 8. Exemplo de texto selecionado para análise                                         | 49 |
| Figura 9. Texto 1 do sujeito 27                                                             | 75 |
| Figura 10. Enunciado de proposta de produção textual                                        | 82 |
| Figura 11. Texto 3 do sujeito 27                                                            | 83 |
| Figura 12. Texto 4 do sujeito 27                                                            | 85 |

## **LISTA DE GRÁFICOS**

|                                                            |    |
|------------------------------------------------------------|----|
| Gráfico 1. Distribuição de hipossegmentação por ano letivo | 54 |
| Gráfico 2. Distribuição de hipersegmentação por ano letivo | 54 |
| Gráfico 3. Distribuição de híbridos por ano letivo         | 55 |
| Gráfico 4. Distribuição de hífen por ano letivo            | 55 |

## LISTA DE TABELAS

|                                                                       |    |
|-----------------------------------------------------------------------|----|
| Tabela 1. Médias obtidas e metas do IDESP para alunos do 9º ano do EF | 34 |
| Tabela 2. Relação dos sujeitos considerados <i>outliers</i>           | 37 |
| Tabela 3. Sujeitos e produção textual por ano letivo                  | 38 |
| Tabela 4. Tipos de dados encontrados por ano letivo                   | 53 |
| Tabela 5. Hipossegmentação e estrutura prosódica - sujeito 9          | 59 |
| Tabela 6. Hipossegmentação e estrutura prosódica - sujeito 16         | 59 |
| Tabela 7. Hipossegmentação e estrutura prosódica - sujeito 25         | 59 |
| Tabela 8. Hipossegmentação e estrutura prosódica - sujeito 27         | 60 |
| Tabela 9. Hipossegmentação e estrutura prosódica - sujeito 60         | 60 |
| Tabela 10. Hipossegmentação e estrutura prosódica - sujeito 69        | 61 |
| Tabela 11. Hipossegmentação e estrutura prosódica - sujeito 73        | 61 |
| Tabela 12. Hipossegmentação e estrutura prosódica - sujeito 90        | 61 |
| Tabela 13. Hipossegmentação e estrutura prosódica - sujeito 93        | 62 |
| Tabela 14. Hipossegmentação e estrutura prosódica - sujeito 107       | 62 |
| Tabela 15. Hipossegmentação e estrutura prosódica - sujeito 112       | 62 |
| Tabela 16. Hipossegmentação e estrutura prosódica - sujeito 115       | 63 |
| Tabela 17. Hipersegmentação e estrutura prosódica - sujeito 9         | 64 |
| Tabela 18. Hipersegmentação e estrutura prosódica - sujeito 16        | 64 |
| Tabela 19. Hipersegmentação e estrutura prosódica - sujeito 25        | 65 |
| Tabela 20. Hipersegmentação e estrutura prosódica - sujeito 27        | 65 |
| Tabela 21. Hipersegmentação e estrutura prosódica - sujeito 60        | 65 |
| Tabela 22. Hipersegmentação e estrutura prosódica - sujeito 69        | 66 |
| Tabela 23. Hipersegmentação e estrutura prosódica - sujeito 73        | 66 |
| Tabela 24. Hipersegmentação e estrutura prosódica - sujeito 90        | 66 |
| Tabela 25. Hipersegmentação e estrutura prosódica - sujeito 93        | 67 |
| Tabela 26. Hipersegmentação e estrutura prosódica - sujeito 107       | 67 |
| Tabela 27. Hipersegmentação e estrutura prosódica - sujeito 112       | 68 |
| Tabela 28. Hipersegmentação e estrutura prosódica - sujeito 115       | 68 |
| Tabela 29. Híbridos e estrutura prosódica                             | 70 |
| Tabela 30. Hífen não convencional e estrutura prosódica               | 71 |

## LISTA DE ABREVIATURAS E SIGLAS

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

## **Sumário**

|                                                                              |    |
|------------------------------------------------------------------------------|----|
| <b>INTRODUÇÃO</b>                                                            | 12 |
| <b>1. FUNDAMENTAÇÃO TEÓRICA</b>                                              | 16 |
| 1.1 A noção de palavra                                                       | 16 |
| 1.2 Características das segmentações não convencionais de palavra            | 21 |
| 1.3 Concepção sobre escrita                                                  | 26 |
| <b>2. MATERIAL E MÉTODOS</b>                                                 | 32 |
| 2.1 Material e decisões metodológicas                                        | 32 |
| 2.2 A constituição do objeto de análise                                      | 36 |
| 2.3 Formas de análise                                                        | 47 |
| <b>3. DESCRIÇÃO E ANÁLISE DE DADOS POR SUJEITO</b>                           | 52 |
| 3.1 Segmentação de palavras: tendências quantitativas de dados longitudinais | 52 |
| 3.2 Segmentação de palavras: características prosódicas por sujeito          | 58 |
| 3.2.1 Análise de hipossegmentações                                           | 58 |
| 3.2.2 Análise de hipersegmentações                                           | 63 |
| 3.2.3 Análise de híbridos                                                    | 69 |
| 3.2.4 Análise de dados com hífen                                             | 71 |
| 3.3 Tipos de dados predominantes por sujeito                                 | 72 |
| 3.4 Um sujeito singular?                                                     | 74 |
| 3.5 Flutuação entre o convencional e o não convencional                      | 84 |
| <b>CONSIDERAÇÕES FINAIS</b>                                                  | 94 |
| <b>REFERÊNCIAS</b>                                                           | 98 |

## INTRODUÇÃO

Nesta pesquisa, analisamos as segmentações não convencionais de palavra, que dizem respeito a palavras escritas cujas fronteiras gráficas não seguem as convenções ortográficas do português (TENANI, 2016). Podemos identificar três tipos de segmentações não convencionais: as hipossegmentações, que ocorrem quando estão ausentes as fronteiras gráficas entre palavras, como em “porisso” e “oque”; as hipersegmentações, quando estão presentes fronteiras gráficas dentro de uma palavra ortográfica, como em “que ria” e “morava-mos”; e os híbridos (CUNHA, 2004), que se caracterizam pela ocorrência simultânea de dois tipos de segmentação não convencional de palavras (hipossegmentação e hipersegmentação) em uma sequência de duas palavras, como em “po risso” e “dere pente”. Em relação às fronteiras gráficas, elas podem ser de dois tipos: espaço em branco, caracterizado por uma ausência de marca gráfica entre palavras, e hífen, caracterizado por um traço entre palavras. O nosso objeto de análise são todos os três tipos de segmentações não convencionais de palavras que ocorrem em textos de sujeitos do Ensino Fundamental II (EF II) que foram considerados *outliers* estatisticamente por Fiel (2018) quanto à frequência de hipossegmentação de palavras em relação aos sujeitos do Banco de Dados de Escrita do EF II. É importante salientar, ainda, que a identificação desses sujeitos foi feita por Fiel (2018) ao investigar a frequência de hipossegmentações em amostra longitudinal de textos desse mesmo banco. A autora realizou uma caracterização geral da frequência dessas grafias para um conjunto de 92 sujeitos que produziram textos ao longo de quatro anos do EF II, e, a partir dessa quantificação geral, foram identificados, ainda, alunos em cujos textos a hipossegmentação ocorreu acima da média dos demais nos mesmos anos letivos. Em sua dissertação, Fiel (2018) classificou esses sujeitos como *outliers*, mas seus textos não foram analisados. Nesta dissertação, selecionamos justamente esses alunos não só por esse grupo não ter sido anteriormente investigado, mas, também, porque nos interessa caracterizar linguisticamente as segmentações de palavras produzidas por eles. Ademais, não analisaremos as produções escritas desses sujeitos a partir de uma perspectiva exclusivamente quantitativa, mas buscaremos entender qualitativamente como são as segmentações identificadas.

Para isso, investigamos quais aspectos prosódicos estariam envolvidos em tais segmentações, uma vez que acreditamos que a motivação para essas grafias esteja relacionada a possibilidades de organização prosódica dos enunciados falados. Assim, adotamos o

arcabouço teórico da Fonologia Prosódica (NESPOR; VOGEL, 1986), que concebe os enunciados falados como organizados hierarquicamente em constituintes prosódicos (sílabas, pé, palavra fonológica, grupo clítico, frase fonológica, frase entoacional e enunciado fonológico).

Delimitado o nosso objeto de estudo, nosso objetivo é responder à questão: quais são as características linguísticas das segmentações não convencionais de palavra em textos de um conjunto de alunos selecionados do EF II? Para responder à questão geral, traçamos três objetivos específicos: i) caracterizar as hipossegmentações nos textos produzidos por cada sujeito e comparar essas características presentes nos textos dos sujeitos considerados; ii) caracterizar as hipersegmentações nos textos produzidos por cada sujeito e comparar essas características entre os sujeitos considerados; iii) caracterizar os híbridos nos textos produzidos por cada sujeito e comparar essas características entre os sujeitos considerados. Por meio disso, conseguimos responder às seguintes perguntas específicas: i) como são as hipossegmentações ao longo dos anos nos textos analisados?; ii) como são as hipersegmentações ao longo dos anos nos textos analisados? ; iii) como são os híbridos ao longo dos anos nos textos analisados?

Dessa forma, vamos descrever as características linguísticas de segmentações não convencionais de palavras de sujeitos selecionados para, em seguida, comparar as regularidades encontradas aqui com o que é produzido, tipicamente, por alunos do EF II - tanto a produção dos alunos pertencentes ao Banco quanto as já descritas na literatura por outros trabalhos. A princípio, identificamos e classificamos os tipos de segmentações de palavras encontrados nos textos selecionados; quantificamos os dados por tipo e por sujeito, visando identificar tendências e particularidades quanto à ocorrência dos dados; analisamos cada segmentação de palavra em constituintes prosódicos, levando em conta o algoritmo de formação dos constituintes pé métrico, palavra prosódica e grupo clítico, principalmente; descrevemos regularidades linguísticas dos dados identificados; em seguida, comparamos os resultados obtidos nesta pesquisa com os de pesquisas anteriores sobre estrutura prosódica de hipo e hipersegmentações, bem como os de híbridos; e, por fim, sistematizamos características linguísticas presentes nas segmentações de palavras dos sujeitos investigado.

O caráter longitudinal se mostra importante, uma vez que permite observar o percurso das segmentações não convencionais no decorrer dos quatro últimos anos do EF II. Buscamos encontrar tendências e características particulares, além de possíveis diminuições de ocorrências. Abaurre (1991, 1992), ao tratar de dados de escrita infantil, sinalizou que estudos



com esse caráter temporal possibilitam a observação do contato do aluno com a escrita convencional, mostrando o modo como as crianças propõem segmentações para seus enunciados escritos e os critérios que nortearam essas diferentes proposições.

A relevância da investigação das segmentações não convencionais de palavra de uma perspectiva linguística está, principalmente, em sua natureza complexa enquanto dado de linguagem, como já apontado por Tenani (2016). Além disso, é um importante lugar de pesquisa que interessa a especialistas de diferentes subáreas da Linguística (ABAURRE, 1991; CAPRISTANO, 2003; CHACON, 2005; TENANI, 2011; FIEL, 2018), da Fonoaudiologia (ZORZI, 2006; OLIVEIRA, 2007), da interface Linguística-Educação (CUNHA; MIRANDA, 2007).

Este estudo também tem relevância social, pois, embora os documentos oficiais da Educação visam que as dúvidas relativas à segmentação de palavras devam estar sanadas até o 5º ano do EF (BRASIL, 1997), estudos mostram que essa não é a realidade escolar. Pesquisas como a de Silva (2014), Tenani (2011) e Fiel (2018) mostram que os alunos não só iniciam o EF II apresentando dúvidas em relação às fronteiras de palavras na escrita, como também podem terminar o EF II ainda as apresentando. Além disso, as segmentações não convencionais nos mostram indícios do processo do sujeito em direção à palavra da escrita convencional.

O mesmo ponto relevante também é ressaltado por Torquette (2016, p. 19) a respeito dos PCN:

o estudo das segmentações não convencionais de palavras se mostra especialmente importante, também, porque, em um dos documentos oficiais mais importantes do Ministério da Educação, os parâmetros Curriculares Nacionais (PCN) - norteador do ensino escolar no Brasil - , encontra-se a recomendação de que os alunos deveriam terminar o Ensino Fundamental I sendo capazes de “escrever textos com domínio da separação de palavras/”. (BRASIL, 2000, P. 80). Dito de outro modo, de acordo com os PCN, os escreventes deveriam iniciar o Ensino Fundamental II sem flutuações em relação ao registro da fronteira gráfica de palavra.

Acreditamos, ainda, que este trabalho pode servir de subsídio para que profissionais da educação possam repensar estratégias didático-pedagógicas sobre as segmentações não convencionais, bem como passarem a ter um novo olhar sob os chamados “erros” de

segmentação de escrita, uma vez que enxergá-los como meros problemas de alfabetização ou como indícios de patologia é desconsiderar a heterogeneidade da escrita e outras questões linguísticas relevantes, como as que serão apresentadas nesta dissertação.

No que diz respeito ao modo de organização deste trabalho, dedicamos, na fundamentação teórica, atenção aos subsídios teóricos. Primeiramente, apresentamos algumas das diferentes noções de palavra, como palavra ortográfica, palavra fonológica e palavra morfológica; e apresentamos, brevemente, o modelo de Fonologia Prosódica de Nespor e Vogel (1986, 2007), no qual nos baseamos para realizar as análises prosódicas, como dito anteriormente. Também resenhamos trabalhos anteriores que se dedicaram à análise de dados de segmentações em textos escritos por alunos do EF I e do EF II. Por fim, apresentamos nossa concepção de escrita, que é a da heterogeneidade constitutiva da escrita defendida por Corrêa (2001), o que possibilita a interpretação dos dados de segmentação.

No capítulo 2, descrevemos os aspectos metodológicos adotados e tratamos das características do material de investigação, apresentando o Banco de Dados de Escrita do EF II utilizado e, também, o índice de rendimento escolar da escola em que os textos do Banco foram produzidos a fim de mensurar se a amostra de textos aqui selecionada é representativa das características da escrita do EF II no Brasil. Em seguida, apresentamos os critérios empregados para a seleção do material e as formas de análises que privilegiam os nossos objetivos.

No capítulo 3, após apresentadas as bases teóricas e os aspectos metodológicos que norteiam nossa pesquisa, apresentamos a análise dos dados de segmentações não convencionais de palavras. A princípio, traçamos as tendências gerais e possíveis regularidades da escrita dos alunos investigados. Em seguida, descrevemos as características prosódicas dos dados por sujeito e discutimos possíveis motivações para a ocorrência deles. Logo após, realizamos uma análise de cunho linguístico-textual de quatro textos que, sob nosso olhar, constituem-se importantes indícios da complexidade das segmentações não convencionais. Por fim, apresentamos as considerações finais.

## **1. FUNDAMENTAÇÃO TEÓRICA**

Como adiantamos, neste capítulo, apresentamos os subsídios teóricos nos quais nos baseamos para guiar a nossa análise. A princípio, na seção 1.1, dedicamos a nossa atenção às diferentes noções de palavra, abordando, inclusive, os pressupostos teóricos da Fonologia prosódica de Nespor e Vogel (1986), que guiam grande parte da nossa análise.

Na seção 1.2, resenhamos trabalhos que se ocuparam em analisar os dados de segmentação sob um viés linguístico. Nosso intuito, a partir dessa resenha, é delinear os pontos de partida para a realização deste trabalho, bem como possíveis comparações com os resultados aqui encontrados e os já descritos na literatura sobre o tema.

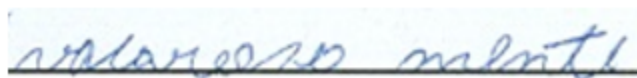
Na seção 1.3, tratamos da concepção de escrita aqui adotada. Primeiramente, abordamos as concepções que enxergam os dados de segmentação não convencional como marcas indesejadas da fala na escrita e, posteriormente, discorremos sobre a heterogeneidade constitutiva da escrita, defendida por Corrêa (2004) e da qual nos valem nesta dissertação.

### **1.1 A noção de palavra**

A definição de palavra é algo complexo, uma vez que seus limites podem variar a depender do critério considerado - critério gráfico, fonológico, sintático, lexical, entre outros. Enquanto a ortografia do português brasileiro utiliza-se predominantemente de critérios morfossintáticos para definir as fronteiras de palavras (SILVA, 2014, p. 53), trabalhos como os de Capristano (2010) e Tenani (2016) mostram que os aprendizes de escrita transitam entre critérios fonológicos, como a presença de acento lexical, para delimitar tais fronteiras, e critérios ortográficos, que mobilizam aspectos visuais, além de morfossintáticos, por exemplo.

Segundo trabalhos como os de Silva (2014) e Tenani (2016), alunos do EF II que produzem os chamados erros de segmentação de palavra segmentam o texto em unidades também com base em critérios rítmicos e entoacionais, não coincidindo, muitas vezes, com as segmentações da convenção ortográfica. Tomemos como exemplo o dado abaixo da Figura 1.

Figura 1. Exemplo de hipersegmentação



Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z09\_6D\_19M\_02)

O exemplo acima - “vacarosa mente” para a convenção “vagarosamente” - configura uma hipersegmentação, uma vez que há uma fronteira marcada pelo espaço em branco dentro de uma palavra ortográfica. Esse dado nos permite observar uma característica apresentada por Silva (2014), pois a presença do acento prosódico indica um lugar privilegiado para a inserção de um limite gráfico não convencional em que esteve presente uma proeminência (vagarosamente > vacarosa mente).

Do ponto de vista ortográfico, tal ocorrência seria vista como um erro, já que diverge da escrita esperada pela convenção. Além disso, cabe salientar que a palavra gráfica é “a apresentação do vocábulo na escrita [...] pelo critério formal. Deixa-se entre eles, obrigatoriamente, um espaço em branco, porque, mesmo quando sem pausa entre si num único grupo, cada um é considerado uma unidade mórfica de per si.” (CÂMARA JR, 2004, p. 23). Dessa forma, a noção de palavra ortográfica, segundo Donadel (2011), está relacionada à escrita. A palavra ortográfica diz respeito, assim, à palavra gráfica e, portanto, o espaço em branco é a sua delimitação/fronteira, e a convenção ortográfica define sua forma gráfica. Portanto, “a palavra ortográfica também é objeto de estudo e a grafia de uma palavra não é dada puramente por uma motivação morfológica/lexical, mas também historicamente constituída”. (TENANI, 2008, p. 234)

Com base nisso, é interessante refletir sobre o fato de que, embora “vacarosa mente” não esteja dentro da ortografia esperada, “vagarosa” e “mente” são palavras que existem na língua portuguesa e também correspondem a palavras ortográficas. Segundo Paranhos (2014), tanto a ausência de fronteira gráfica quanto a presença pode estar relacionada a conflitos entre as dimensões ortográfica e morfossintática da palavra.

Embora tal dado seja, sob esse ponto de vista, considerado um “erro”, é interessante observar que está pautado na noção de palavra fonológica, uma vez que, segundo Câmara Jr (2004), configura um vocábulo portador de um acento lexical. Assim, em “vagarosamente”, há duas palavras fonológicas, porque, segundo Silva (2014), ao considerar a regra geral de distribuição do acento no PB, o acento primário só pode localizar-se em uma das três últimas

sílabas de uma palavra (padrão proparoxítono, paroxítono, oxítono). Assim, no caso da palavra aqui considerada, dada a sua configuração polissilábica, apenas três sílabas poderiam ser candidatas a receber o acento primário: *sa*, *men* e *te*. E, dentre essas possibilidades, o acento recai sobre a primeira sílaba do sufixo *-mente* (*vagarosamente*). Ademais, a palavra “vagarosamente” apresenta quatro sílabas pretônicas e, devido a isso, quando produzida oralmente, recebe um acento secundário (COLLISCHONN, 1994). Com isso, há dois acentos e, conseqüentemente, duas palavras fonológicas.

Segundo Chacon (2021, p. 6),

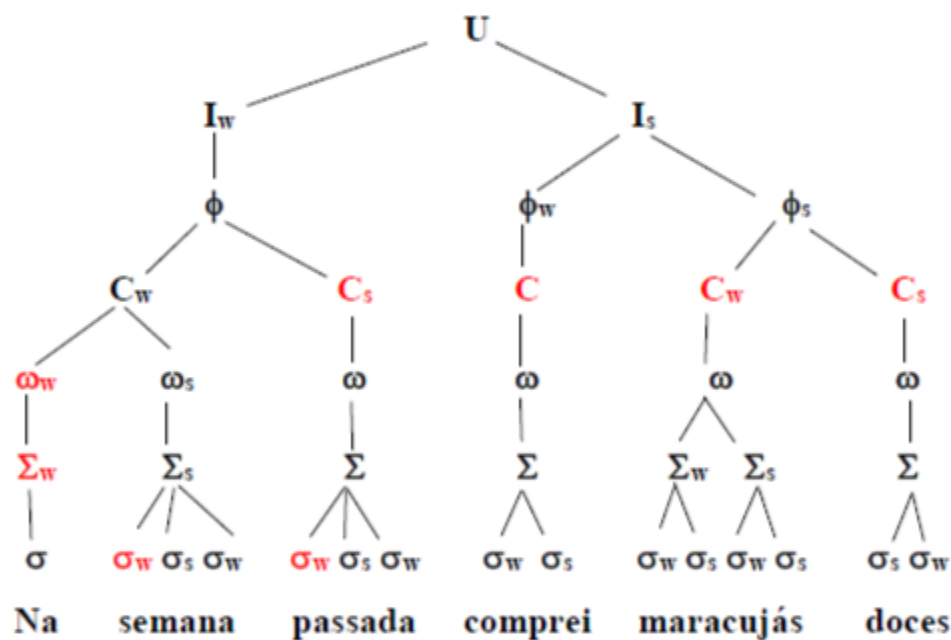
A palavra fonológica é o constituinte que coloca em relação não-isomórfica características prosódicas e morfológicas. A característica principal da palavra fonológica é a de sua proeminência coincidir com um acento da palavra – que pode ser o lexical, quando palavra fonológica e palavra morfológica coincidem, como em “porquinho”, ou outro acento da palavra, como em “guarda-chuva”, uma única palavra do ponto de vista morfológico, mas duas palavras (“guarda” e “chuva”) do ponto de vista prosódico, já que comporta dois acentos.

Para Mattoso Câmara, citado por Bisol (2004), a palavra morfológica compreende palavras lexicais, como nome, adjetivo e verbo - classes abertas -, e palavras funcionais, como preposição, conjunção e determinativos - classes fechadas; enquanto a palavra fonológica distingue palavras com acento (palavras fonológicas) e sem acento (clíticos). Assim, a distinção dessas noções pode, muitas vezes, ocasionar uma não coincidência entre essas duas palavras. Então, no que diz respeito ao estatuto fonológico, a palavra “vagarosamente” é uma palavra composta. Por outro lado, corresponde a apenas uma palavra morfológica, pois, para a morfologia, conforme destaca Boop da Silva (2010), ela se trata de uma unidade formada por derivação.

A derivação se mostra um ponto problemático, porque é outro lugar importante para se observar a não correspondência entre morfologia e fonologia em relação à noção de palavra. Ao pensarmos na palavra “vagar”, temos um acento primário e, portanto, os limites de palavra morfológica e fonológica são idênticos. Em contrapartida, ao passar pelos processos de derivação (*vago* > *vagroso* > *vagarosamente*), passamos a ter, ainda, uma única palavra morfológica, mas duas palavras fonológicas.

Já em relação à prosódia, o estatuto do grupo constituído de uma palavra com acento e outra ou outras sem acento (clíticos) que estão prosodicamente relacionadas também é um lugar problemático. Para Nespor e Vogel (1986), esse grupo é chamado de grupo clítico. Tais autoras defendem o arcabouço teórico da Fonologia Prosódica (NESPOR; VOGEL, 1986), em que os enunciados falados são concebidos como organizados hierarquicamente em constituintes prosódicos, conforme figura abaixo: sílaba ( $\sigma$ ), pé métrico ( $\Sigma$ ), palavra fonológica ( $\omega$ ), grupo clítico (CG), frase fonológica ( $\phi$ ), frase entoacional (I) e enunciado fonológico (U).

Figura 2. Exemplo de hierarquia prosódica



Fonte: Tenani (2016, p. 79) a partir da hierarquia de Nespor e Vogel (1986, 2007)

Como podemos observar no exemplo dado, o chamado grupo clítico, para Nespor e Vogel (1986), é subsequente à palavra fonológica. Entretanto, tal questão é diversa entre os fonólogos, como Selkirk (1986), que, embora apresente um modelo semelhante ao de Nespor e Vogel, defende que o domínio subsequente da palavra fonológica é a frase fonológica e não o grupo clítico. Para ela, a existência desse grupo violaria um requisito da hierarquia prosódica.

|                      |                       |
|----------------------|-----------------------|
| Selkirk (1978, 1986) | Nespor e Vogel (1986) |
| ENUNCIADO            | ENUNCIADO             |
| FRASE ENTOACIONAL    | FRASE ENTOACIONAL     |
| FRASE FONOLÓGICA     | FRASE FONOLÓGICA      |
| PALAVRA FONOLÓGICA   | <b>GRUPO CLÍTICO</b>  |
| PÉ                   | PALAVRA FONOLÓGICA    |
| SÍLABA               | PÉ                    |
|                      | SÍLABA                |

Fonte: Gayer (2015 p. 160)

Assim, como evidenciado por Bisol (2004, p. 63),

A prosodização do clítico no nível da palavra, o que corresponderia à visão de Mattoso, ou no nível subsequente da escala prosódica, proposta de Nespor e Vogel, ou ainda no nível da frase, vem produzindo interessantes artigos, mostrando que, quanto aos clíticos, da mesma forma que quanto aos compostos, as línguas particulares têm de ser ouvidas, pois há comportamentos diversificados.

Com base nessas diferentes abordagens teóricas citadas, Bisol (2000, 2005), como explicitado por Silva e Tenani (2014), defende a hipótese de que determinados fatos do português brasileiro só se explicam na sequência exata de um clítico mais o seu hospedeiro, o que corrobora a necessidade de se considerar um domínio como o grupo clítico, defendido por Nespor e Vogel. Por outro lado, Simioni (2008) “[...] apoia-se na ideia de que, por meio de restrições, o funcionamento dos clíticos pode ser englobado em um domínio prosódico superior ao da palavra (a frase fonológica), sem ter a necessidade de mais um constituinte prosódico exclusivo que abrigue o clítico e seu hospedeiro” (SILVA; TENANI, 2014, p. 63).

Apesar dos diversos posicionamentos em relação ao domínio em que os clíticos são anexados, trazemos, ainda, a reflexão de Silva (2014, p. 65) sobre o assunto:

[...] entendemos que a formalização do comportamento prosódico dos clíticos, nas duas propostas apresentadas, parece se encaminhar em uma mesma direção: a de que, no PB, os clíticos tendem a ser mais independentes com relação à palavra hospedeira, já que estão sempre ganhando status prosódico nos domínios que já estão em interação com o componente sintático. Por hipótese, acreditamos que as fronteiras de palavras que se marcam para além do esperado na escrita poderiam ser um indício importante dessa autonomia maior que os elementos clíticos apresentam no PB.

Por meio das vertentes apresentadas, podemos perceber que as noções de palavra ora privilegiam critérios gráficos, ora funcionais ou sonoros, além de não haver um consenso entre as diferentes teorias. Assim sendo, acreditamos que as segmentações não convencionais de palavras configuram tentativas do escrevente de se adequar à escrita da convenção. O escrevente parece transitar entre essas diferentes noções, em um processo inconsciente, uma vez que sua escrita ora se baseia em critérios fonéticos e fonológicos, ora se baseia em critérios morfossintáticos e até na própria convenção ortográfica. Assim, compartilhamos da ideia de Tenani (2018), que afirma que o sujeito circula por práticas de oralidade e letramento, como abordaremos no tópico 1.3, em que definiremos nossa concepção sobre escrita.

## **1.2 Características das segmentações não convencionais de palavra**

Nesta seção, resenhamos alguns trabalhos sobre as grafias não convencionais no EF I e EF II, que são descritas, principalmente, por autores que partilham dos mesmos pressupostos teóricos em relação ao modelo da Fonologia Prosódica, mobilizado na análise da organização prosódica dos enunciados. Por outro lado, nem todos os trabalhos aqui citados compartilham da mesma concepção de escrita por nós adotada, ou seja, nem todos os trabalhos enxergam as segmentações não convencionais como indícios da circulação do escrevente por práticas orais/letradas.

Dentre as possibilidades de segmentações não convencionais de palavras, já



explicitadas acima, passamos a tratar de algumas características prosódicas mais gerais encontradas em pesquisas anteriores para o EF I e EF II. Partindo das hipersegmentações do EF I, é importante salientar que esse tipo de dado pode ocorrer, segundo Abaurre (1991), pois os escreventes podem perceber que a maior parte das palavras do português brasileiro é dissilábica e constituída de sílabas compostas com a estrutura consoante-vogal (CV), como *casa*. Assim, quando alguma palavra tiver uma configuração diferente dessa, os escreventes podem acabar segmentando o enunciado em palavras não convencionais constituídas de duas sílabas. Chacon (2005), ao estudar as hipersegmentações de trissílabos, observa que a hipersegmentação não é aleatória, uma vez que nos momentos de ruptura “as crianças parecem tentar plasmar em sua escrita contrastes rítmicos que certamente detectam, sobretudo, a partir do relevo prosódico do acento da oralidade e que, por sua recorrência, padronizam-se sob forma de constituintes prosódicos da própria língua.” (CHACON, 2005, p. 82)

Já Silva e Tenani (2014), para o EF II, verificam que a maioria das hipersegmentações mostram informações fonológicas importantes ligadas aos constituintes prosódicos, como a presença de um acento, que constitui lugar privilegiado para inserção de um limite gráfico (embora > em *bora*). Tais resultados corroboram os de Chacon sobre o EF I. Tanto os dados do EF I quanto os do EF II mostram que os constituintes palavra fonológica e grupo clítico são os mais envolvidos nas hipersegmentações (cf. PARANHOS, 2014; CHACON, 2005; CAPRISTANO, 2007; SILVA e TENANI, 2014). Por outro lado, no EF II, outros constituintes também envolvidos são a sílaba e o pé métrico, como em “com verso” (converso) (SILVA, TENANI, 2014). Além disso, Tenani (2011) afirma que há tendência, dentre os casos de hipersegmentação, de a sílaba pretônica de uma palavra ortográfica ser analisada como um clítico, como em “na quela”. Apesar disso, tal característica não se fundamenta para o EF I, descrito por Paula (2007), uma vez que, no EF I, como observado por Tenani (2016), as palavras hipersegmentadas não necessariamente correspondem a monossílabos átonos, como a sílaba pretônica “ma” em “ma telo” (martelo). Por outro lado, somente no EF I têm sido encontradas hipersegmentações como “aj u da” (“ajuda”), que não correspondem a constituintes prosódicos, como argumenta Chacon (2005).

Já nos casos de hipossegmentação, Cunha e Miranda (2007) mostram que a tendência predominante nos dados de hipossegmentação foram a sequência de um clítico (correspondente a uma palavra gramatical) grafado junto a uma palavra fonológica, como em: “olobo” (“o lobo”). Isso, segundo as autoras, mostra que a criança tem dificuldade em

reconhecer como palavras conjuntos de uma ou duas letras (como os artigos “a” e “o”, ou pronomes “me” e “te”) e, por isso, ocorreria as hipossegmentações, como já apontado por Ferreiro e Teberosky (1999).

A segunda tendência predominante é de duas palavras fonológicas grafadas juntas, como em “chicobento” (“Chico Bento”). Entretanto,

Os casos de hipossegmentações entre duas palavras fonológicas ocorrem em menor quantidade e formam frases fonológicas ou frases entoacionais. Sobre esse tipo de dado, as autoras [Ferreiro e Teberosky (1999)] explicam que frases fonológicas, como “miaroupa” (“minha roupa”), podem ocorrer devido à mobilização de critérios morfossintáticos e semânticos, como a relação entre o adjetivo “minha” e o substantivo “nome” que constituem um sintagma e, podem também, formar uma frase fonológica. No caso das frases entoacionais, pode haver uma motivação devido à presença de um contorno entoacional e de pausa que delimitam porções do enunciado, como em “Chico Bento saiu correndo”, grafado como “chicobento saiu correndo”. Observa-se que “chicobento” é um nome que pode funcionar como tópico com contorno entoacional próprio, e “saiu correndo”, uma sentença que pode funcionar como comentário com seu [contorno] entoacional característico. (FIEL, 2018, p. 34 - acréscimos nossos)

Além disso, no EF I, há uma possível manutenção da integridade da sílaba e do pé métrico, já que o constituinte silábico é preservado nas ocorrências de hipossegmentação. Há envolvimento dos três constituintes de nível mais alto da hierarquia prosódica nos processos de hipossegmentação, e, além disso, o escrevente parece perceber a fala como um contínuo, apresentando, portanto, forte tendência em separar a escrita de acordo com unidades entoacionais. (CUNHA; MIRANDA, 2007)

Já dentre os casos de hipossegmentação do EF II, 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”.

Nos dados de hipossegmentação no EF I, há mobilização de todos os domínios prosódicos e mesclas entre domínios, como mostram Cunha (2004, 2010), Cunha e Miranda (2007), Capristano (2007) e Torquette (2016). No EF II, há predomínio de mobilização de fronteiras de pé métrico, palavra prosódica e grupo clítico, como mostram Tenani (2011) e Paranhos (2014), não sendo observados, dessa forma, dados que envolvessem fronteiras de

constituintes mais altos da hierarquia prosódica, como frase entoacional e enunciado fonológico.

De modo geral, Tenani (2011) constatou que há mais dados de hipossegmentação em relação aos de hipersegmentação e que os constituintes prosódicos relevantes para caracterizar as hipossegmentações e hipersegmentações encontradas nos textos do EF II diferem daqueles relevantes para os dados encontrados nos textos do EF I.

Sobre os híbridos (CUNHA, 2010), cabe explicitar, primeiramente, que também são chamados de *mesclas* (CHACON, 2005; CAPRISTANO, 2007a, 2007b) ou *mistos* (PIACENTE; QUEREJETA, 2012). São segmentações, em geral, raras e se mostram como uma combinação entre separações e junções não previstas de palavras, como em *a olado* (ao lado). A diferença entre os termos guarda uma diferença de definição do objeto: para Cunha (2004), híbridos se caracterizam pela ocorrência simultaneamente de dois tipos de segmentação não convencional de palavras em uma sequência de duas palavras (hipossegmentação e hipersegmentação); para Chacon (2005), essas grafias são *mesclas* de constituintes prosódicos de modo que se observam junção e segmentação entre fronteiras de domínios prosódicos; para Piacente et al. (2012), os *mistos* combinam características das hipo e hipersegmentações: entre duas palavras se produz uma união e uma separação não convencional.

Dessa forma, as diferentes nomeações para esse tipo de segmentação não convencional evidenciam a complexidade desse dado, além de “[...] expor a flutuação com a qual esse tipo de fenômeno tem sido classificado e estudado nas pesquisas sobre a escrita da criança, flutuação ligada a diferenças teóricas e metodológicas entre as pesquisas” (CAPRISTANO, 2021, p. 791). Além disso, trabalhos como os de Chacon (2005), sobre o EF I, e os de Tenani e Paranhos (2011), sobre o EF II, mostram que os híbridos ocorrem em quantidades muito menores do que hipo e hipersegmentações tanto no EF I quanto no EF II. Entretanto, sua ocorrência é mais comum no EF I, quase desaparecendo no EF II. Assim, apesar de praticamente não mais ocorrerem no EF II, é pouco comum que nesse tipo de dado haja ruptura de sílaba.

Para melhor caracterizar os híbridos, nós, inicialmente, comparamos as características envolvidas nesses dados com as encontradas, respectivamente, nas hipossegmentações e hipersegmentações dos dados do EF II. Assim, percebemos que clítico (monossílabos átonos) mais palavra prosódica é uma sequência que, predominantemente, gera híbridos (BORSATO, 2019), mesma estrutura que tende a gerar hipossegmentações em textos do EF II, conforme

descrito por Fiel (2018), que analisa hipossegmentações em uma amostra longitudinal do EF II. Alguns exemplos citados pela autora são “denovo”, “porai” e “anoite”. Em sua pesquisa, Fiel (2018) conclui que “a junção entre clítico e palavra prosódica foi a estrutura que mais gerou dados de hipossegmentação em todos os anos do EF II.” (p. 107). Os híbridos também se caracterizam por espaço em branco na fronteira direita de sílaba tônica, característica semelhante às de hipersegmentação no EF II, conforme caracterização feita por Silva (2014). Segundo a autora,

Na maioria das hipersegmentações, apreendemos informações fonológicas importantes ligadas a esses constituintes, como a presença do acento prosódico, que se constituiu em um lugar privilegiado para a inserção de um limite gráfico não-convencional na fronteira em que esteve presente essa proeminência, como em, por exemplo, “em **bora**” (embora). (SILVA, 2014, p 161)

Verificamos, assim, que híbridos do EF II têm características comuns à hipossegmentação e à hipersegmentação identificadas no EF II. A partir dessas considerações, ao identificar semelhanças nas motivações entre híbridos, por um lado, e hipossegmentação e hipersegmentação, por outro, passamos a comparar essas motivações de hipo e hipersegmentação do EF II com as do EF I.

Tomamos a descrição dos dados do EF I, feita por Cunha (2004), Capristano (2007), Cunha e Miranda (2007), para comparar as características encontradas nessas pesquisas com as encontradas nesta pesquisa. Partimos de Cunha e Miranda (2007) que caracterizaram híbridos por apresentarem primeiro uma hipossegmentação e, em seguida, uma hipersegmentação, como: “e levou” (“ele vou”) e “quem fim” (“que enfim”).

Ainda segundo Cunha e Miranda (2009, p. 135), “a hipossegmentação ocorre sempre entre o clítico, nesses casos um artigo definido, e uma palavra de conteúdo que começa por consoante. Essa sequência parece ser a preferência das crianças [...]”. As autoras trazem, também, alguns exemplos: “poriso” (por isso), “derepente” (de repente), “medeu” (me deu), semelhantes aos que ocorrem no EF II.

Já os dados de hipersegmentação, no EF I, caracterizam-se, predominantemente, por deixar um espaço em branco na fronteira direita da sílaba tônica, mesma característica do EF II, como descrito por Silva (2014). Sobre essa característica, vale ressaltar que Cunha e

Miranda (2009, p. 140) consideram que a maioria dos dados de hipersegmentação deixam à direita um troquele silábico, ilustrados por meio dos exemplos “a **onde**” (aonde), “em **bora**” (embora), “da **quela**” (daquela) e “com **migo**” (comigo). Citando Abaurre (1991) e Cagliari (2002), as autoras argumentam, ainda, ser possível supor que a constituição de palavras dissílabas paroxítonas é uma das variáveis que pode influenciar a decisão da criança sobre onde segmentar.

### 1.3 Concepção sobre escrita

Na seção anterior, resenhamos alguns trabalhos a respeito dos tipos e das características das segmentações não convencionais. Aqui, salientamos que, em uma perspectiva presente no ambiente escolar, apresentar muitos casos de segmentação não convencional de palavras escritas pode ser considerado um problema de alfabetização ou até mesmo um distúrbio de escrita, como na perspectiva de Zorzi (2006). Segundo esse autor, esses erros não podem ser muito frequentes e, caso sejam, por um longo período, as crianças podem ser diagnosticadas como portadoras de transtornos de aprendizagem, uma vez que tal ocorrência é considerada sintoma de *déficits* no processo de aprendizagem. Além disso, esse tipo de interpretação está ligado a uma concepção de escrita que enxerga interferências da fala na escrita, como podemos observar no trecho abaixo:

A oralidade exerce forte influência quando as crianças formulam hipóteses a respeito de como as palavras devem ser escritas, gerando uma tendência muito grande no sentido de que os padrões de fala determinem o modo de escrita. Tudo ocorre como se o sistema de representação gráfica do português fosse fonético, e não alfabético. Isto quer dizer que a criança toma como referência, para determinar que letra deve usar, aqueles sons que ela é capaz de identificar em seu padrão de pronúncia da palavra procedendo então a uma correspondência um a um: som falado / letra que escreve (ZORZI, 2006, p. 8).

Neste trabalho, faz-se relevante considerar como se dão relações entre características da fala e da escrita. Partindo de uma perspectiva dicotômica, fala e escrita são vistas como duas modalidades polarizadas da língua. De modo geral, as características próprias da fala e da escrita são descritas/prescritas por essa visão, na qual a fala é vista como imprecisa, não planejada e não normatizada, enquanto a escrita é precisa, planejada e normatizada.

Dessa forma, por meio dessa visão dicotômica, a fala acaba recebendo características negativas em relação às da escrita, que apresenta características predominantemente valorizadas. É importante ressaltar que essa valorização da escrita em relação à fala é uma visão muito forte na perspectiva escolar, uma vez que acredita-se que a escola tem a função de ensinar a norma padrão da língua portuguesa. Além disso, é comum o pensamento de que os chamados erros de escrita sejam decorrentes de uma interferência negativa da fala na escrita. Portanto, “não apenas haveria características distintas entre fala e escrita, mas também haveria uma relação assimétrica entre essas modalidades linguísticas”. (TENANI, 2016, p. 116)

Sob certo ponto de vista da abordagem clínica, os chamados erros de separação e junção de palavras podem ser interpretados como indicativos de um transtorno da escrita chamado de disortografia. Esse transtorno, como afirmam Fernández et al (2010), está diretamente relacionado aos quadros de transtorno específico de aprendizagem, como, por exemplo, o distúrbio de aprendizagem. Ainda segundo esses autores, os sintomas típicos desse transtorno são “os erros de segmentação da cadeia de fala que ocasionam segmentação indevida de palavras e união incorreta de palavras [...]” (FERNANDÉZ et al, 2010, p. 4). Essa perspectiva, similarmente à concepção adotada por Zorzi (2006), apoia-se em uma suposta interferência da fala na escrita semelhante àquela predominante em ambiente escolar, pois se entende que os casos de segmentação não convencional de palavras escritas ocorrem pelo fato de o aluno se basear na oralidade para segmentar as porções do enunciado escrito.

Fernández et al (2010, p. 2) argumentam que essa suposta interferência inicial é esperada e deve se extinguir quando os escreventes “compreenderem o que é o princípio alfabético, pois, dessa forma, eles poderão fazer uso de maneira mais eficaz da informação obtida pela decodificação letra-som”. Entretanto, se o aluno continua produzindo textos com segmentações não convencionais nos anos finais do EF, ele pode ser diagnosticado com disortografia.

Por outro lado, entendemos que compreender casos de segmentação não convencional de palavras como problemas de alfabetização ou como indícios de patologia é desconsiderar a relação heterogênea da língua, e tal questão pode ser ainda mais problemática quando

tratamos de segmentações de sujeitos considerados *outliers* em relação à quantidade alta de hipossegmentações. Por isso, distanciamos-nos desse posicionamento clínico, e até patológico, para adotarmos, neste trabalho, a concepção de heterogeneidade proposta por Corrêa (2004), na qual a enunciação oral/falada e letrada/escrita é heterogênea na medida em que o sujeito da linguagem é constituído por práticas orais e letradas.

A visão de Corrêa (2004) rompe com as visões de relações dicotômicas entre fala e escrita e fundamenta-se na ideia de que “todo fato linguístico vincula-se a uma prática social” (p. 2), ou seja, a fala e a escrita resultam do atravessamento do sujeito por práticas de oralidade e também de letramento. A heterogeneidade é, portanto, constitutiva da escrita e não uma característica acessória a ela. Segundo essa abordagem, nega-se a possibilidade de se opor fala e escrita por suas materialidades fônicas e gráficas, já que elas são constituídas pelas práticas de oralidade e letramento (o oral/falado e o letrado/escrito), ou seja, os fatos da língua só se realizam por serem modos de enunciação concretizados por meio de práticas de linguagem:

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)

Segundo Corrêa (2004), o sujeito escrevente circula por três eixos de constituição da escrita: i. o eixo da gênese da escrita; ii. o eixo do código escrito institucionalizado e iii. o eixo da relação com o falado/escrito e ouvido/lido. Ainda segundo o autor, o primeiro eixo refere-se aos momentos em que, ao apropriar-se da escrita, o escrevente faz uma representação baseada na suposição de que a escrita mantém uma relação direta com a fala, podendo representar integralmente o falado. Logo, é uma “tentativa, por parte do escrevente, de representação do planejamento conversacional e do jogo argumentativo prosodicamente marcado ou, finalmente, sua tentativa de registrar graficamente a materialidade fônico-pragmática do oral/falado” (CORRÊA, 1997, p. 185).

Já o eixo do imaginário do código escrito institucionalizado diz respeito à representação que o sujeito escrevente faz do escrito institucionalizado para a sua escrita.

Segundo Corrêa (1997, p. 271), “o escrevente lida, basicamente, com o que supõe ser – a partir não só do que aprendeu na escola, mas, em grande parte, do que assimilou fora dela – a visão escolarizada de código institucionalmente reconhecido”. Com isso, a escrita é considerada “autônoma, de tal forma que as marcas do oral/falado que eventualmente nela se pudessem reconhecer seriam vistas como desvios do instituído e, nesse sentido, deveriam ser tomadas como lhe sendo totalmente exteriores” (CORRÊA, 1997, p. 170 - 171). É possível observar, portanto, o trânsito do escrevente em direção ao que ele imagina ser a escrita escolar, ou seja, há marcas linguísticas dos entrelaçamentos do oral/falado e do letrado/escrito que, nesse eixo, “se evidenciam sempre que o escrevente leva a extremos uma tal imagem sobre o código escrito. Ou seja, nesse eixo da circulação dialógica, esses encontros se mostram pelo excesso” (CORRÊA, 1997, p. 271- 272).

O terceiro eixo de constituição da escrita diz respeito à relação que o texto do escrevente mantém com o já falado e com o já ouvido, bem como com o já escrito e com o já lido. Esse eixo se caracteriza na “ideia de que a escrita é um tipo particular de enunciação em que relações com o mundo e com o falado se dão no próprio processo de escrever, caracterizando não a representação escrita do mundo e do falado, mas a representação que o escrevente faz da relação escrita/mundo e escrita/falado” (CORRÊA, 1997, p. 341).

Outra importante contribuição dada por Corrêa (2006) é a definição dos três modos de reconhecimento da heterogeneidade da escrita, que podem ser: i. por meio de aspectos da representação gráfica, já que, no processo da escrita, o escrevente oscila entre a tentativa de representação de características fonético-fonológicas (segmentais e/ou prosódicas) detectadas em sua variedade linguística falada e a convenção ortográfica institucionalizada; ii. pela referência à heterogeneidade da língua, pois, se a admitimos como estruturalmente marcada no sistema, não há razão para não a admitirmos na escrita - com a condição de que esta última seja vista como um modo de enunciação e não apenas como uma tecnologia; iii. pela circulação dialógica que o escrevente faz ao produzir o texto escrito, uma vez que ele circula pelos três eixos de constituição da escrita já mencionados.

Portanto, assumimos que as segmentações são pistas do imaginário que os escreventes têm do que seja a constituição da (sua) escrita por meio de suas inserções em práticas orais/faladas e letradas/escritas de uso da linguagem, pois “o escrevente oscila entre a tentativa de representação de características fonético-fonológicas (segmentais e/ou prosódicas) detectadas em sua variedade linguística falada e a convenção ortográfica institucionalizada” (CORRÊA, 2007, p. 01).



Assim, a proposta de Corrêa (2004) é fundamental para interpretarmos os dados de segmentações não convencionais de palavra não apenas enquanto dado de língua, mas também como dado de discurso. Além disso, segundo Corrêa (2004), a heterogeneidade da escrita se mostra através da relação do sujeito com a linguagem, uma vez que é por meio dessa relação que se faz possível recuperar a presença das práticas sociais nas produções discursivas.

Com base nessas considerações, interpretamos que as segmentações não convencionais de palavras são indícios da heterogeneidade da escrita, como compartilhado por Capristano (2010), que, em seu trabalho, apresenta o exemplo de “ven mos” (vamos). Segundo a autora, a hipersegmentação ocorre em limites de sílaba (*ven* e *mos*), o que indicia a circulação do escrevente por práticas orais ou, em outras palavras, indicia a atuação de limites prosódicos. Por outro lado, a pesquisadora constata, ainda, a possibilidade de observar que em *ven mos* (vamos) a ruptura parece resultar

da atuação de uma restrição ortográfica da língua em sua modalidade escrita: a impossibilidade de unir sequencialmente, na escrita do português, as letras m e n ou, ainda, a atuação da proximidade gráfica da sequência mo com a configuração gráfica do monossílabo no (contração de em + o) e da proximidade gráfica da sequência vem com o verbo vir (vem). (CAPRISTANO, 2020, p. 186)

Logo, Capistrano (2010) conclui que o reconhecimento de palavras da língua ou a atuação de uma restrição ortográfica constituem indícios da circulação do escrevente por suas práticas letradas - ou seja, indiciam a atuação de informações ligadas à escrita institucionalizada -; e o reconhecimento dos limites prosódicos indicia a circulação do escrevente por práticas orais. Dessa forma, as grafias de segmentações não convencionais resultam do trânsito do escrevente por práticas orais/letradas em que ele está inserido.

Para concluir esta seção, assumimos, ancorados nos pressupostos apresentados, uma abordagem que concebe a escrita pelo encontro entre os modos de enunciação falado e escrito e as práticas sociais orais e letradas para analisar as segmentações extraídas dos textos analisados neste trabalho. Buscamos, aqui, nos distanciar de uma visão patológica (e até depreciativa) em relação às produções e aos sujeitos aqui investigados e, assim, pretendemos entender as produções desses escreventes como marcas da heterogeneidade da escrita,

observando-as como indícios do imaginário 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.

## 2. MATERIAL E MÉTODOS

Neste capítulo, apresentamos as escolhas teórico-metodológicas referentes ao material do qual selecionamos os dados para esta pesquisa e a metodologia empregada para seleção e exclusão de dados. Inicialmente, discorreremos a respeito da caracterização do Banco de Dados de Escrita do Ensino Fundamental II. Logo após, tratamos dos critérios nos quais nos baseamos para estabelecer o recorte do material e o levantamento dos dados.

Buscamos, neste capítulo, ilustrar como se deu a aplicação desses critérios a fim de evidenciar como foram fundamentais para a superação de dificuldades iniciais para a identificação das hipossegmentações. Na sequência, explicitamos as formas de análise dos resultados e os aspectos metodológicos que embasam a execução de tais formas. O propósito deste capítulo também é caracterizar as escolhas teórico-metodológicas referentes ao material e ao *cópus* de investigação, aos critérios de identificação dos dados e à forma de análise das grafias.

### 2.1 Material e decisões metodológicas

O material utilizado nesta pesquisa pertence ao Banco de Dados de Escrita do Ensino Fundamental II, que faz parte do Projeto de Extensão Universitária “Desenvolvimento de Oficinas de Leitura, Interpretação e Produção Textual” e contou com a coordenação das professoras doutoras do IBILCE-UNESP, Luciani Tenani e Sanderléia Longhin-Thomazi.

Os textos foram produzidos por alunos que cursaram os quatro anos do Ensino Fundamental II (6º, 7º, 8º e 9º anos), durante os anos de 2008 a 2011, na escola pública *Zulmira da Silva Salles*, localizada no município de São José do Rio Preto, interior de São Paulo. Além disso, cabe salientar que não era permitido aos alunos reescrever os textos, sendo necessário entregá-los ao responsável pela oficina ao final da aula de língua portuguesa. Devido à impossibilidade de reelaboração e de uso de lápis grafite (o que levou a rasuras do texto escrito), foram privilegiados os registros do processo de elaboração da produção textual.

No geral, o banco de dados é composto por 5.519 textos escritos por 662 alunos a partir de diferentes tipologias textuais (previamente estabelecidas pela escola em função do planejamento escolar) e de diferentes gêneros e temas (definidos pela coordenação do projeto de extensão em reuniões com os monitores integrantes do projeto de extensão que atuavam na escola). Dessa quantidade, 2.759 textos fazem parte da amostra transversal e 3.645 da amostra longitudinal.

Ademais, quanto a esta última, enfatizamos que 124 estudantes participaram integralmente do projeto de extensão, e, desse total, 119 alunos redigiram textos em todos os anos, somando um total final de 2.495 produções de textos. Para a nossa pesquisa, foram analisados 254 textos (10,2%) da amostra longitudinal, os quais foram produzidos por 13 alunos considerados *outliers*, sendo seis do sexo feminino e sete do sexo masculino. Esse subconjunto de textos se caracteriza pela produção textual desses 13 alunos em todos os anos letivos. Com o objetivo de melhor caracterizar o percurso das segmentações não convencionais nos textos analisados ao longo do processo de escolarização, privilegiamos uma amostra longitudinal, sendo necessário haver textos dos mesmos sujeitos em todos os quatro anos letivos considerados. Na mesma esteira de Tenani (2016), consideramos importante tratar, ainda, sobre os índices de rendimento escolar (índice estadual – IDESP – e índice nacional – IDEB), pois, segundo a autora, por meio dessas características, é possível observar se a amostra de textos aqui selecionada é representativa das características da escrita do EF II no Brasil. De acordo com Tenani (2016), ao longo dos quatro anos da coleta dos textos do banco, o IDESP da escola permaneceu um pouco acima da média geral do estado de São Paulo no mesmo período considerado.

Na tabela a seguir, são apresentadas as 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.

**Tabela 1 - Médias obtidas e metas do IDESP para alunos do 9º ano do EF**

| ano letivo | média em<br>Líng. Port. | média da<br>escola | média do<br>estado | meta a ser<br>alcançada |
|------------|-------------------------|--------------------|--------------------|-------------------------|
| 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: Tenani (2016) a partir de dados da Secretaria da Educação do Estado de São Paulo, IDESP, 2015.

Como observado por Tenani (2016), tanto os índices da escola quanto os índices do estado seguem a mesma variação de aumento e decréscimo nos anos letivos considerados. Há um aumento de 0,27 pontos no índice de 2009 em relação ao de 2008, um decréscimo de 0,33 pontos em 2010, seguido de um aumento de 0,03 pontos em 2011. Assim, ao demonstrar esses dados, a autora interpretou que os alunos que produziram os textos da amostra analisada representam as características gerais dos alunos do EF II do estado de São Paulo. Quanto ao IDEB, Tenani (2016) aponta que a escola em questão alcançou índices um pouco acima da média do país em 2009 (escola 4,7 e país 3,8) e 2011 (escola 5,1 e país 3,9), mostrando, novamente, que os dados da amostra do EF II são representativos de características gerais do Brasil. Como evidenciado até o momento, o “Banco de Dados de Escrita do Ensino Fundamental II” pode ser tomado como representativo da escrita do EF II.

Outra informação relevante, ainda segundo a pesquisadora, é em relação ao perfil do município quanto aos índices de desenvolvimento humano. Por meio disso, seria possível mensurar não apenas características da cidade, mas também características sociais de desenvolvimento humano do local onde residiam os alunos cujos textos são analisados. Isto é, “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.” (TENANI, 2016, p. 45)

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. Segundo Tenani (2016),

São José do Rio Preto é classificado na 50ª posição entre os 5.565 municípios brasileiros, com índice de 0,797 em 2010, considerado alto (faixa

de 0,7 a 0,799 pontos) entre as cinco faixas de desenvolvimento (dados 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).

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 investimento no município a fim de proporcionar aos cidadãos melhores condições de desenvolvimento. Por outro lado, Tenani (2016) ressalta que, ao considerar os índices dessas três dimensões do IDHM no período de 1991 a 2010, a dimensão Educação foi a que mais cresceu (na ordem de 0,339 pontos), seguida por Longevidade e Renda. Essa tendência é semelhante à observada no país, uma vez que a dimensão Educação também foi a que mais cresceu em relação às demais. Outros dados importantes do IDHM Educação são os que dizem respeito à situação da educação entre a população em idade 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. 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. 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) (TENANI, 2016).

Ela conclui:

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 EF II e do aumento do acesso à Educação da

população brasileira de modo amplo, verificado nas últimas duas décadas (de 1991 a 2010). (TENANI, 2016)

Por meio dessas informações, descrevemos, brevemente, o contexto em que os textos coletados foram produzidos: as características do material selecionado e as características da escola e do município em que os textos foram produzidos.

## 2.2 A constituição do objeto de análise

Após a caracterização do banco de dados, apresentamos, a seguir, os critérios de seleção para a composição do material desta pesquisa:

- i) Alunos que foram considerados *outliers* em relação à alta frequência de dados de hipossegmentação em pelo menos um ano do EF II;
- ii) Alunos que tenham participado dos quatro anos de coleta do projeto de extensão;

O *primeiro critério* de seleção foi estabelecido com base na pesquisa de Fiel (2018) que, ao realizar uma caracterização geral da frequência de hipossegmentações para o conjunto de 92 sujeitos que produziram textos ao longo de quatro anos do EF II, apresentou uma distribuição, por ano letivo, de sujeitos que produziram uma frequência elevada de hipossegmentações e que, estatisticamente, podem ser considerados *outliers*.

Fiel concluiu que houve alunos estatisticamente *outliers* em todos os anos do EF II, bem como uma média de quatro *outliers* por ano, exceto no sétimo ano, no qual se encontram seis. Dessa forma, “20% (18/92) dos sujeitos investigados apresentaram muitos dados de hipossegmentações e são considerados *outliers*” (FIEL, 2018, p. 62). Além disso, não são sempre os mesmos alunos que se comportam como *outliers* em todos os anos do EF II. Com base nisso, trazemos a Tabela 2 a seguir, no qual os números apresentados nas colunas

identificadas com “nº sujeito” dizem respeito ao número identificador atribuído ao sujeito no Banco de Dados de Escrita:

**Tabela 2 - Relação de sujeitos considerados *outliers***

| 6º ano     | 7º ano     | 8º ano     | 9º ano     |
|------------|------------|------------|------------|
| nº sujeito | nº sujeito | nº sujeito | nº sujeito |
| 16         | 90         | 93         | 69         |
| 60         | 9          | 112        | 107        |
| 115        | 77         | 27         | 9          |
| 27         | 73         | 25         | 25         |
|            | 16         |            |            |
|            | 25         |            |            |

Fonte: FIEL, 2018, p. 63

Percebemos, desse modo, que há o sujeito 25, considerado *outlier* em três anos do EF II; há três sujeitos (27, 16 e 9) considerados *outliers* em dois anos do EF II; e nenhum sujeito foi considerado *outlier* em todos os anos do EF II. Por fim, a breve descrição dos resultados quantitativos obtidos em relação à identificação de *outliers*, realizada por Fiel (2018), leva-nos a apontar a necessidade de estudo específico para esse conjunto de sujeitos. Porém, embora a classificação tenha sido realizada pela autora apenas a partir da frequência de hipo, nesta dissertação, serão analisados todos os tipos de segmentações não convencionais.

O *segundo critério* de seleção considerou a necessidade de que os alunos tenham participado de todos os anos de coleta para possibilitar o acompanhamento longitudinal das produções escritas de um mesmo sujeito.

Estabelecidos os critérios, foram identificados 18 sujeitos *outliers*, sendo quatro considerados *outliers* em mais de um ano letivo, totalizando 13 sujeitos ao todo. Para esta pesquisa, serão considerados todos os 254 textos produzidos por esses sujeitos, independentemente do número de dados encontrados em seus textos. Na tabela abaixo, trazemos a distribuição de textos por ano letivo, sendo, na coluna “Sujeito”, o número de identificação do escrevente no Banco de Dados de Escrita; nas demais colunas, indicadas pelos anos, a quantidade de textos produzidos e, por fim, o total de textos produzidos por ano e o total de textos produzidos por aluno ao longo do período:



Tabela 3 - Sujeitos e produção textual por ano letivo

| Sujeito      | 6º ano | 7º ano | 8º ano | 9º ano | Total      |
|--------------|--------|--------|--------|--------|------------|
| 9            | 03     | 06     | 04     | 05     | 18         |
| 16           | 04     | 07     | 04     | 03     | 18         |
| 25           | 04     | 06     | 05     | 07     | 22         |
| 27           | 06     | 04     | 04     | 06     | 20         |
| 60           | 03     | 04     | 01     | 01     | 09         |
| 69           | 06     | 07     | 06     | 07     | 26         |
| 73           | 05     | 06     | 06     | 05     | 22         |
| 77           | 06     | 05     | 02     | 07     | 20         |
| 90           | 05     | 06     | 02     | 03     | 16         |
| 93           | 05     | 04     | 06     | 03     | 18         |
| 107          | 06     | 06     | 04     | 07     | 23         |
| 112          | 05     | 06     | 05     | 06     | 22         |
| 115          | 06     | 06     | 03     | 05     | 20         |
| <b>Total</b> | 64     | 73     | 52     | 65     | <b>254</b> |

Fonte: Elaboração própria

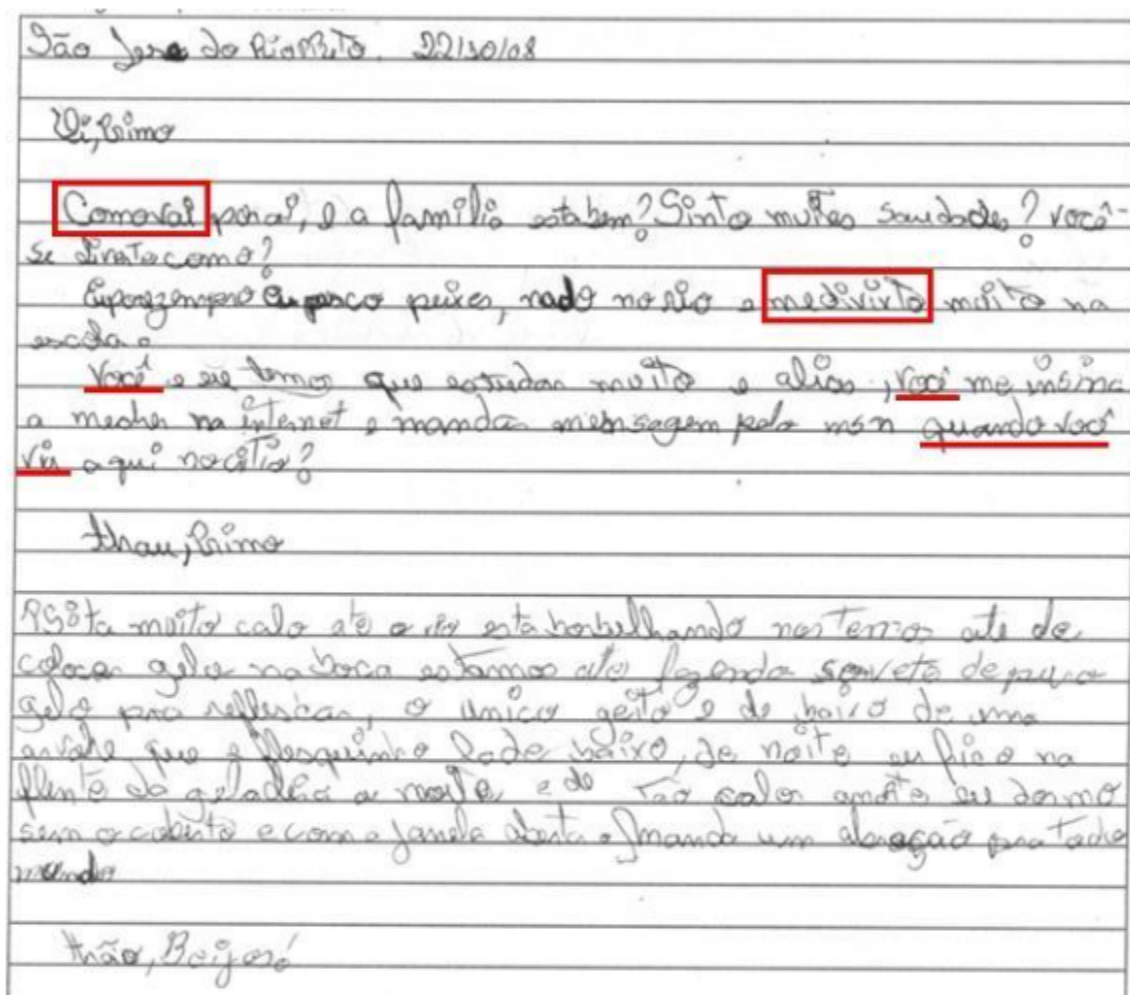
Nesta dissertação, encontramos alguns problemas metodológicos para caracterizar e classificar os dados. Como dito anteriormente, analisaremos todos os tipos de segmentações não convencionais de palavras, a saber: hipossegmentações, hipersegmentações e híbridos, além de dados envolvendo hífen. Em relação aos dados de hipossegmentação e hipersegmentação, a dificuldade esteve relacionada à categorização do espaço em branco. Já nos dados de híbridos e uso não convencional de hífen, o problema esteve mais relacionado à categorização devido à natureza complexa do dado. Para isso, adotamos alguns critérios para identificação dos dados.

Para as hipossegmentações, realizaremos uma comparação dos espaços em branco entre palavras ao longo do texto, bem como uma comparação entre as grafias das mesmas letras dentro de palavras semelhantes que ocorram no mesmo texto, metodologia descrita em Tenani (2016). Encontramos dados de difícil classificação, uma vez que a fronteira demarcada pelo espaço em branco nem sempre é tão clara, deixando dúvidas em relação às

segmentações.

A seguir, trazemos alguns exemplos:

Figura 3.1 Exemplo de critério de identificação de dados de segmentação não convencional

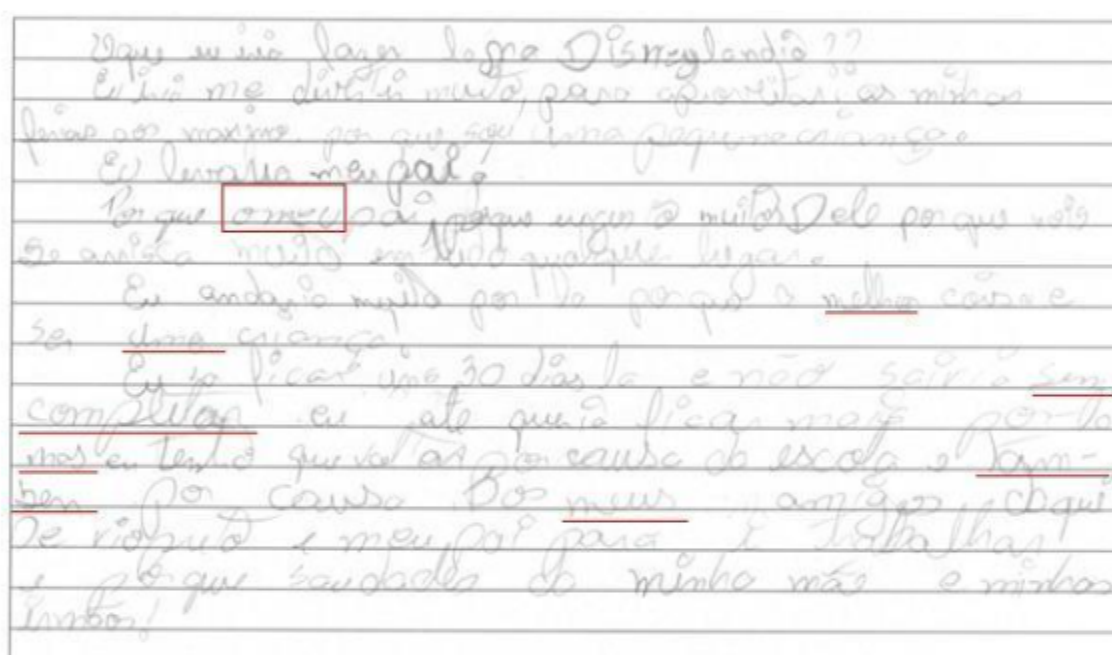


Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z08\_5D\_04M\_04)

Na Figura 3.1, trataremos do dado “Comovai”, ocorrido em um texto do sujeito 9, que parece ter uma tendência em escrever palavras com o espaço em branco de diferentes tamanhos e oscilar entre letra cursiva e letra de forma. Por conta disso, em seus textos, algumas grafias nos geraram dúvidas, e, para resolver essa questão em relação à delimitação de fronteira de palavra, realizamos uma análise na esteira de Fiel (2018), que traz o mesmo exemplo acima. Inicialmente, comparamos as grafias das mesmas letras em palavras semelhantes que ocorriam nesse texto. Observamos, a princípio, a caligrafia em “quando você”, uma vez que apresenta uma sequência parecida com “como vai” - a ligação da vogal

“o” com a letra “v”. Percebemos que, nesse exemplo, a letra “o” apresenta um traço mais baixo, enquanto a letra “v” não está inclinada, como acontece em “como vai”. Ao analisar a letra “v” nas ocorrências “você” e “viu”, percebemos a mesma característica - a ausência da inclinação. Porém, ao analisar a grafia “medivirto”, em que o “v” foi grafado junto das demais letras, também percebemos uma inclinação, o que nos fez entender que essa grafia é indício de que essas duas palavras foram grafadas juntas. Portanto, a sequência “Comovai” foi classificada como um dado de hipossegmentação.

**Figura 3.2 Exemplo 2 de critério de identificação de dados de segmentação não convencional**



Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z08\_5D\_04M\_04)

Já em outro texto do mesmo sujeito, outra grafia que causou dúvidas foi a sequência “o meu”, uma vez que o espaço entre essas duas palavras é pequeno. Porém, diferentemente do caso acima, optamos por não considerar essa grafia como um dado de segmentação ao observar que, nos exemplos grifados, a letra “m”, quando grafada no interior de uma palavra e antecedida por uma vogal, parece ter uma ligadura mais clara. Assim, ao grafar “m” após alguma vogal, o escrevente aparenta ter uma tendência de unir essas letras. Assim, nesse caso, “o meu” não foi considerado um dado de hipossegmentação.

No caso das palavras homônimas, utilizamos a estratégia de considerar todo o sentido do enunciado em que a grafia ocorreu a fim de verificarmos qual foi o significado pretendido pelo aluno. Tomemos como base o exemplo “demais” na sentença “Eu precisava *demais*

atenção.”. A palavra hipossegmentada se trata da locução “de mais” e, portanto, de acordo com a convenção ortográfica, deveria ser grafada com a presença do espaço em branco “de mais”. Por ter sido grafada sem a presença dessa fronteira gráfica, tornando-se um adjunto adverbial, consideramos a ocorrência como um dado de hipossegmentação.

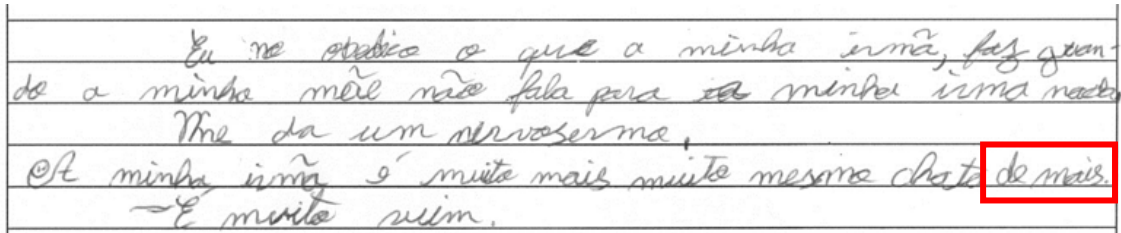
Para a identificação das hipersegmentações, tomamos como base a proposta de Tenani (2011), na qual a autora sugere uma tipologia no levantamento de dados de segmentação não convencional baseada em um critério gráfico (espaço em branco: “na que le” ou hífen: “mora-va”) e em critério morfossemântico (homonímias: “de mais”, quando a grafia é “demais”). A partir dessa metodologia de levantamento de dados, encontramos algumas dificuldades para categorização do espaço em branco, como: dificuldades em interpretar a caligrafia do escrevente, casos de translineação e ocorrências de palavras estrangeiras.

A respeito do primeiro problema, tomamos a mesma decisão em relação às hipossegmentações. Dessa forma, realizaremos não só uma comparação dos espaços em branco entre as palavras, mas também uma comparação entre as grafias das mesmas letras encontradas em palavras semelhantes que ocorram no mesmo texto, metodologia descrita em Tenani (2016).

Já o segundo problema encontrado na categorização do espaço em branco foi a translineação, em que ocorre a necessidade de dividir uma palavra em duas linhas, de modo que uma parte da palavra fique no final de uma linha e o restante dela na parte inicial da linha seguinte. A regra para marcar a translineação é a inserção de um hífen do lado direito da palavra, porém observamos que tal registro gráfico não é sempre marcado, o que causaria dúvidas em relação a considerar determinadas separações como hipersegmentações ou apenas um esquecimento do aluno, ou desconhecimento, de tal inserção do hífen. Assim, resolvemos excluir os dados em contexto de translineação. Por fim, palavras estrangeiras também serão excluídas, uma vez que seguem uma ortografia diferente da do português brasileiro.

No caso das palavras homônimas, utilizamos a estratégia de considerar todo o sentido do enunciado em que a grafia ocorreu a fim de verificarmos qual foi o significado pretendido pelo aluno. Tomemos como base o exemplo “de mais”, apresentado na Figura 4 abaixo. Na sentença: “A minha irmã é muito mais muito mesmo chata de mais” (linha 1), a palavra hipersegmentada tem a função sintática de adjunto adverbial, uma vez que está intensificando o adjetivo “chata”. Portanto, a grafia “de mais”, nesse caso, é um dado de hipersegmentação, já que a tentativa da escrevente foi a de grafar o advérbio de intensidade “demais”.

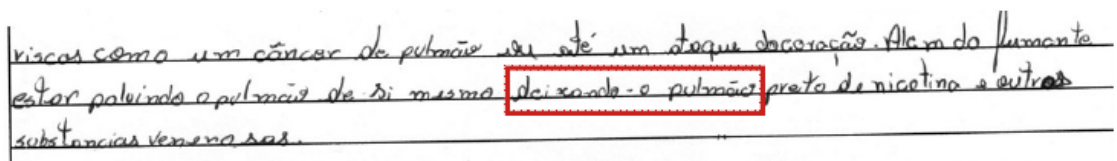
**Figura 4. Exemplo de hipersegmentação de palavra homônima**



Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z08\_5D\_04M\_04)

Tratamos, ainda, da categorização do hífen. Em relação ao uso do hífen, encontramos dados que foram entendidos como hipersegmentações, uma vez que o principal critério que define tal tipo de segmentação seria do limite gráfico no interior de palavras ortográficas, por exemplo “precisa-se” para “precisasse”, ou “morava-mos” para “morávamos”. Há também casos como o exemplo “deixando-o pulmão”, em que o aluno parece apresentar dúvidas em relação à construção enclítica, o que ocasionaria o uso não convencional do hífen. Há também dados complexos em relação a sua análise e classificação, como o exemplo “todos-a” (todos a), que não se trata, claramente, de uma grafia correspondente a um uso não convencional, mas também não é um caso de hipossegmentação, hipersegmentação ou híbrido. Apesar de sua natureza complexa, tomamos a decisão de classificá-lo como um dado de uso não convencional de hífen.

**Figura 5. Exemplo de uso não convencional de hífen**



Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z11\_8B\_03M\_06)

Por fim, em relação aos dados de híbridos, encontramos dificuldades de caracterização. Trazemos o exemplo do dado “estados zunidos” (Estados Unidos), no qual sua classificação como híbrido pode ser complexa, uma vez que a hipossegmentação e a hipersegmentação não são tão evidentes nesse dado. Apesar dessa dificuldade, classificamos grafias desse tipo como sendo ocorrências de híbridos.

Além disso, ressaltamos que, na análise dos dados, fazemos referência aos escreventes do *cópus* por meio de seus números correspondentes no Banco de Dados de Escrita do

Ensino Fundamental II. Por meio desse posicionamento, buscamos preservar a identidade dos alunos.

**Figura 6. Número identificador dos escreventes**

| Redações |        |          |          |       |                                |                                  |          |          |          |
|----------|--------|----------|----------|-------|--------------------------------|----------------------------------|----------|----------|----------|
|          | Filtro | Proposta |          |       |                                |                                  |          |          |          |
|          | Escola | Ano      | Série    | Turma | Gênero                         | Proposta                         | ID Aluno | Digital  | Texto    |
| 1        | ZSS    | 2010     | 7ª Série | B     | Relato de expectativas de vida | Benefícios da imortalidade       | 27       | Download | Download |
| 2        | ZSS    | 2010     | 7ª Série | B     | Relato de mudança de vida      | Tomar-se famoso                  | 27       | Download | Download |
| 3        | ZSS    | 2010     | 7ª Série | B     | Relato de experiência          | Vivência de mal entendido        | 27       | Download | Download |
| 4        | ZSS    | 2010     | 7ª Série | B     | Carta argumentativa            | Presentes de Natal               | 27       | Download | Download |
| 5        | ZSS    | 2011     | 8ª Série | B     | Relato de experiência          | Amizade e internet               | 27       | Download | Download |
| 6        | ZSS    | 2011     | 8ª Série | B     | Carta de opinião               | Massacre em Realengo             | 27       | Download | Download |
| 7        | ZSS    | 2011     | 8ª Série | B     | Narrativa de ficção            | Aventura em terra distante       | 27       | Download | Download |
| 8        | ZSS    | 2011     | 8ª Série | B     | Carta de opinião               | Tecnologia e impacto da Internet | 27       | Download | Download |

Fonte: Banco de Dados de Escrita do Ensino Fundamental II

Por fim, pontuamos que, nesta análise, durante a leitura dos textos, apresentamos dificuldades em atribuir sentido aos textos produzidos pelo sujeito 77, como podemos observar nas figuras a seguir.





**Figura 7.2. Exemplo 2 de dificuldade de interpretação de texto**

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. Professora Zulmira da Silva Salles

Nome: \_\_\_\_\_  
Proposta 2 \_\_\_\_\_

Série/Turma: 6ª A  
Data: 27/10/19

Um dia eu estava na rua quando um menino me chamou e falou que precisava de um livro. Eu dei o livro e ele me agradeceu. Depois ele me falou que precisava de um livro e eu dei o livro e ele me agradeceu. Depois ele me falou que precisava de um livro e eu dei o livro e ele me agradeceu.

Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z09 6A 21M 02)

A partir das figuras 7.1 e 7.2, podemos perceber que o sujeito 77, além de ter uma grafia que dificulta a compreensão de algumas palavras, apresenta produções a que não conseguimos atribuir sentido. Em 7.1, temos a seguinte transcrição a partir do que conseguimos entender:

Trecho 1. Como é a sobrevivência, os amigos e os inimigos de cada parte do mundo e de cada dia. Quando falei com a minha mãe:

- Vou ficar (não identificamos)

E não cumpri.



Em 7.2, temos:

Trecho 1. Um dia eu estava na \* quando o meu avô me chamou e falou que na realidade eu não sei jogar futebol e os meus primos queria na lua mas como não tinha transporte para ir para lá tínhamos um foguete para ir pra lá.

Sentimos peso dos homens queriam ir para a lua como não sabem ir para lá queria um piloto para pilotar cada um desses foguetes para irmos para a lua ou outros planetas que descobrimos a experiências de cada uma pessoa que consegue ir para os planetas a minha experiência foi de ficar quinze dias sozinhos na mata e eu fiquei mas superei as consequencia.

Como as pessoas ficam sem os pais em casa

\*casa de problema nas escola.

Onde as pessoas ficam com medo de gente ou de bicho para ir para a escola ou com bichos de pessoas para qualquer um que precisa de pessoas para conseguir atestado ou pais para ir con pessoas que não pode cada um está localizado em (rasura)\* estado de São Pulo em outros países localizados.

[FIM]

Assim, durante a leitura do texto, deparamo-nos com grafias que não conseguimos deduzir qual seria a palavra correspondente. Além disso, o texto, em si, não tem coerência. Devido a isso, optamos por excluí-lo de nossa análise, totalizando, portanto, 12 sujeitos ao todo e 234 textos.

## 2.3 Formas de análise

Nesta pesquisa, ao investigar a ocorrência de todos os tipos de segmentações de palavra de cada sujeito *outliers* ao longo do EF II, pretendemos caracterizar as segmentações não convencionais desses sujeitos sob uma perspectiva linguística e qualitativa. A fim de dialogar com trabalhos anteriores, analisaremos os dados de segmentações em constituintes prosódicos de modo a caracterizar possíveis relações entre as grafias encontradas e estruturas linguísticas, típicas dos enunciados falados. Em seguida, quantificaremos esses dados levando em conta a distribuição deles em relação aos sujeitos para identificar tendências e características da escrita de cada sujeito para, após, compararmos tais características com as típicas do EF II, detalhadas na seção 1.2 com os trabalhos de Fiel (2018), Silva (2014), Paranhos (2014) e Tenani (2011). Assim, poderemos observar se as produções desses sujeitos se distanciam, ou não, do que é tipicamente produzido no EF II.

Também faremos um levantamento da quantidade de dados por ano letivo, a fim de melhor caracterizar o percurso dos dados de segmentações para cada sujeito. Dessa forma, analisaremos a quantidade de dados produzidos por sujeito, bem como a relação tipo de dado/sujeito. Além disso, para uma melhor análise, consideraremos a quantidade de dados em relação à quantidade de palavras escritas por texto, uma vez que, ao considerar as segmentações em relação ao total de palavras escritas por aluno, há resultados que sinalizam, como evidenciado por Fiel (2018), um decréscimo das médias ao longo da escolarização. Assim,

A descrição ora realizada na forma de quantificar os dados demonstra que a afirmação de alto número de ocorrência de junção entre palavras não é critério suficiente para validar diagnóstico de “distúrbio específico de escrita”, uma prática bastante comum em ambiente escolar ou clínico. A consideração da relação da ocorrência da junção e o tamanho do texto escrito se mostra, pois, como abordagem mais segura para afirmações sobre o que pode ser tomado como característico da escrita. (FIEL, 2018, p. 59)

Por meio da afirmação de Fiel (2018), percebemos a relevância de se considerar o total de palavras em trabalhos sobre segmentação de palavras. Tal importância também foi apontada em trabalhos de Paranhos (2014) e de Torquette (2016).

Como já exposto, os sujeitos terão seus dados de palavras analisados de modo qualitativo. Nesta pesquisa, a análise será adotada com o propósito de descrever as características rítmicas das segmentações, de modo a apontar possíveis tendências em relação aos constituintes prosódicos mais mobilizados pelos sujeitos. Desse modo, buscamos demonstrar que as segmentações de palavras são pistas deixadas pelo sujeito em seu texto, que permitem observar a relação de representação de características dos enunciados falados nos enunciados escritos, como também a representação de seu conhecimento sobre a convenção ortográfica e sobre a noção de fronteira de palavra. Para exemplificar como as análises foram feitas, passamos ao texto, na figura que se segue, de um sujeito do material que foi considerado *outlier*.

Figura 8. Exemplo de texto selecionado para análise

Um pais estrangeiro  
 Eu tinha es quite uma carta de Rebasão e a  
 pessoa falou que ia se sortindo a carta.  
 O sorteio ia no dia 07/11/2008 ai Tarso da  
 aurna ia escale duas pessoas para i mesa pais  
 Com uma nome especial de Diego Sanchez foi  
 sortido para ir ao meu amigo Andre Campanha  
 de Jinda um grupo de saida as 10:00 da noite  
 não foi.  
 Ai não foi na janela parage o pais não,  
 rio jupiter e saturno e sol e a lua não parou  
 a lua chegou em 7:00 da noite.  
 Não deseu da aeronave o pais era, muito  
 estrangeiro es corer as danças as as comida de  
 da era muito estrangeiro de mais.  
 Ai não viu um ate ele Jinda 4 dias.  
 As manx de ele era uma Bola Jinda um  
 monte de luz.  
 As esdora delas era quadradadas.  
 As Comida de da ate que era um Jauco  
 costoso es ele era um irais astre.  
 Era o ultimo dia de viagem para volta  
 a Josa.  
 No dia 20/11/08 não chegou a terra as  
 1:00 hora da manhã ate que foi na era i na  
 que se pais estrangeiro.

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

Identificamos, no texto acima, que o aluno produziu 11 dados de segmentações não convencionais, sendo quatro dados de hipersegmentação, a saber: es quite (escrito), de mais (demais), de ele (dele), na que le (naquele); cinco dados de hipossegmentação, a saber: ea (e a), aurna (a urna), paraze (para ver), ea (e a), aotro (ao outro); e apenas dois dados de híbrido: da eronave (da aeronave) e as davano (andavam no).

O sujeito apresentou, em um único texto, vários dados de segmentação não convencional, entretanto, o dado que mais chama a atenção, neste texto, é o híbrido “as davano” (andavam no), porque se trata de um dado pouco comum nos textos que analisamos (BORSATO, 2019). Nesse dado, ocorre uma hipersegmentação de “andavam”, seguida de uma hipossegmentação entre “andavam” e “no”. Além disso, na palavra “andavam”, há uma troca de <n> por <s> (andavam – as davano). Tal troca revela características complexas do

próprio processo de escrita, sendo elas: aspecto fonético, pois os sons [s] e [n] são alveolares, sendo um traço em comum apesar de o modo e o vozeamento serem diferentes; aspecto fonológico, já que tanto <s> quanto <n> são consoantes que ocorrem em fim de sílaba, portanto, são letras que podem ocupar a mesma posição silábica; e aspecto ortográfico, uma vez que não se emprega <n> em fim de sílaba que seja fim de palavra. Desse modo, percebe-se que a troca de <n> por <s> não é aleatória, mas ancorada em aspectos fonético, fonológico e ortográfico do português. Outro fato interessante é que, na ocorrência de hipossegmentação, não há registro de <m> em fim de sílaba (davamno). É possível que haja, portanto, uma motivação devido à presença do oral no escrito, uma vez que não há consoante nasal <m> na fala, ocorrendo uma sequência de segmentos fonéticos como [vão no]. Além disso, a nasalidade está garantida com o registro do <n> adjacente a <o>. Por meio de tal análise, podemos observar que o sujeito, apesar de produzir um dado bastante diferente, utiliza-se de critérios fonológicos, fonéticos e da própria convenção ortográfica, que são indícios do que o sujeito faz em relação a sua escrita, uma vez que ele “oscila entre a tentativa de representação de características fonético-fonológicas (segmentais e/ou prosódicas) detectadas em sua variedade linguística falada e a convenção ortográfica institucionalizada (...)” (CORRÊA, 2007, p. 01). Dessa forma, temos pistas de que é a atuação simultânea de aspectos prosódicos e aspectos relativos à escrita convencional que determinaram o seu aparecimento. Segundo Chacon (2004), em casos como esse, “o reconhecimento daquilo que poderia funcionar como uma palavra [...] acompanharia o próprio movimento de sequenciação da atividade de escrever” (p. 227).

Outra observação a ser feita, em relação a esse dado, é que se trata de uma sequência de palavra prosódica mais clítico, enquanto a sequência predominante a gerar híbridos no EF II é a de clítico mais palavra prosódica. Outro dado interessante a ser observado é a grafia hipersgmentada de “de ele”, em “as naves **de ele** era uma bola...”. Observamos que há uma flutuação dessa grafia, uma vez que, na linha abaixo, percebemos que o sujeito grafia “dele” de acordo com a convenção ortográfica, em “as cadesa **deles** era quadradas”. Dessa maneira, “essa flutuação quanto à segmentação possibilita enveredar pela hipótese de que é relevante considerar características textual-discursivas dos enunciados” (TENANI, 2016, p. 127). Tal flutuação enfraquece a hipótese de que seria exclusivamente tal estrutura métrica a causadora dessa segmentação. Além disso, as características observadas nesse texto são pistas da complexa relação do sujeito com a escrita, pois evidenciam a tensão entre duas imagens que o escrevente faz da (sua) escrita: ora confere a ela “um poder quase ilimitado de representação

e de fidelidade representacional” (CORRÊA, 2004, p. 82); ora “lida, basicamente, com o que supõe ser – a partir não só do que aprendeu na escola, mas, em grande parte, do que assimilou fora dela – a visão escolarizada de código institucionalmente reconhecido.” (CORRÊA, 1997, p. 271).

### 3. DESCRIÇÃO E ANÁLISE DE DADOS POR SUJEITO

Neste capítulo, apresentamos a análise dos dados de segmentações não convencionais de palavras. Por meio dessa análise, buscamos: (i) apresentar tendências gerais e possíveis regularidades da escrita de alunos que cursaram os quatro anos do EF II; (ii) descrever características prosódicas dos dados apresentados e discutir possíveis motivações para a ocorrência desses dados; (iii) realizar análise de cunho linguístico-textual de quatro textos que, sob nosso olhar, constituem-se importantes indícios da complexidade inerente às segmentações não convencionais de palavras escritas. Para tanto, organizamos este capítulo em cinco seções. Na seção 3.1, apresentamos as análises quantitativas dos dados. Na seção subsequente, descrevemos as características prosódicas dos dados e separamos as análises por tipos de segmentações. Com isso, buscamos identificar tendências longitudinais, bem como os tipos predominantes de cada sujeito, que serão abordados na seção 3.3, a fim de melhor caracterizar o percurso das segmentações não convencionais. Na seção 3.4, realizamos uma análise qualitativa das segmentações de um sujeito que apresentou tendências particulares em relação aos demais, mostrando não apenas a estrutura linguística mobilizada, mas também apresentando uma hipótese sobre sua motivação linguística. Por fim, na seção 3.5, tratamos da flutuação entre grafias convencionais e não convencionais de segmentação de palavra realizadas pelo sujeito analisado na seção anterior.

#### 3.1 Segmentação de palavras: tendências quantitativas de dados longitudinais

Nesta pesquisa, encontramos 579 dados em 234 textos produzidos por 12 sujeitos rotulados por *outliers* no estudo de Fiel (2018). São 379 dados de hipossegmentação, 204 dados de hipersegmentação, nove dados com usos não convencionais de hífen e sete híbridos. Na tabela, a seguir, apresentamos os tipos de dados por ano letivo.

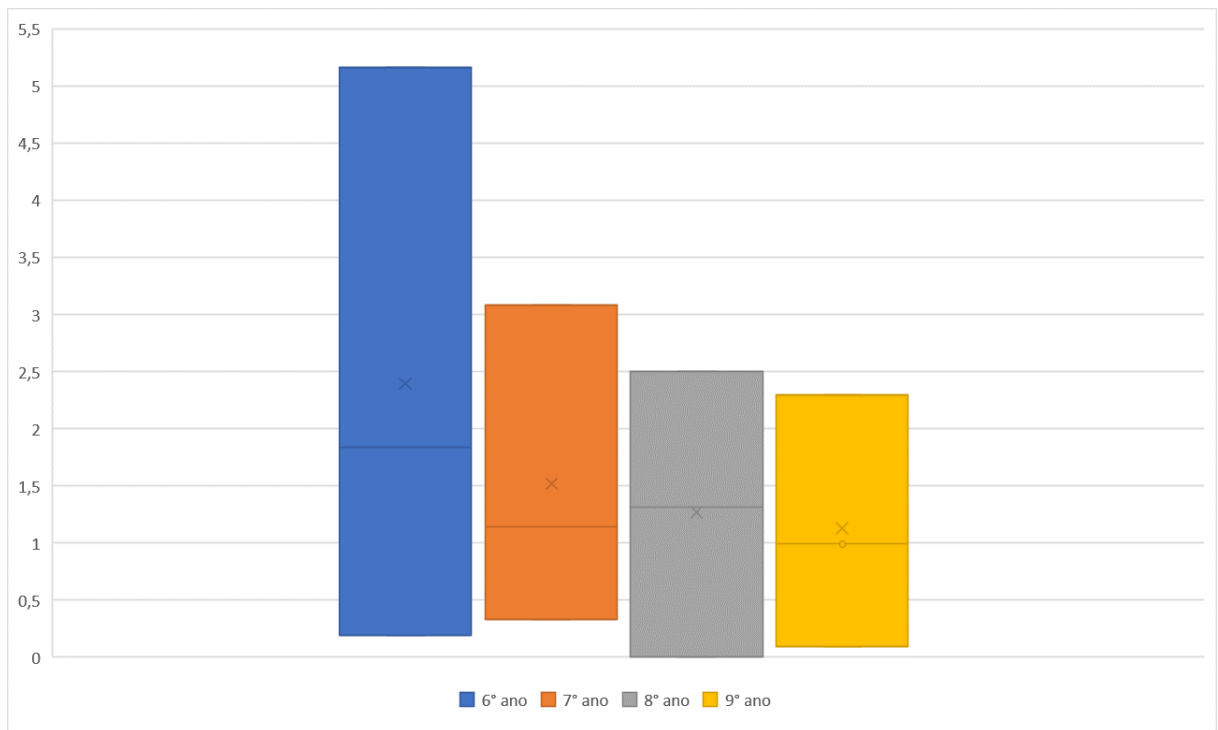
Tabela 4 – Tipos de dados encontrados por ano letivo

| <b>Ano</b>    | <b>Hipo</b> | <b>Hiper</b> | <b>Híbrido</b> | <b>Hífen</b> | <b>Total</b> |
|---------------|-------------|--------------|----------------|--------------|--------------|
| <b>6º ano</b> | 88          | 52           | 4              | 1            | 145          |
| <b>7º ano</b> | 114         | 82           | 1              | 4            | 201          |
| <b>8º ano</b> | 71          | 31           | 2              | 1            | 105          |
| <b>9º ano</b> | 86          | 39           | 0              | 3            | 128          |
| <b>Total</b>  | <b>359</b>  | <b>204</b>   | <b>7</b>       | <b>9</b>     | <b>579</b>   |

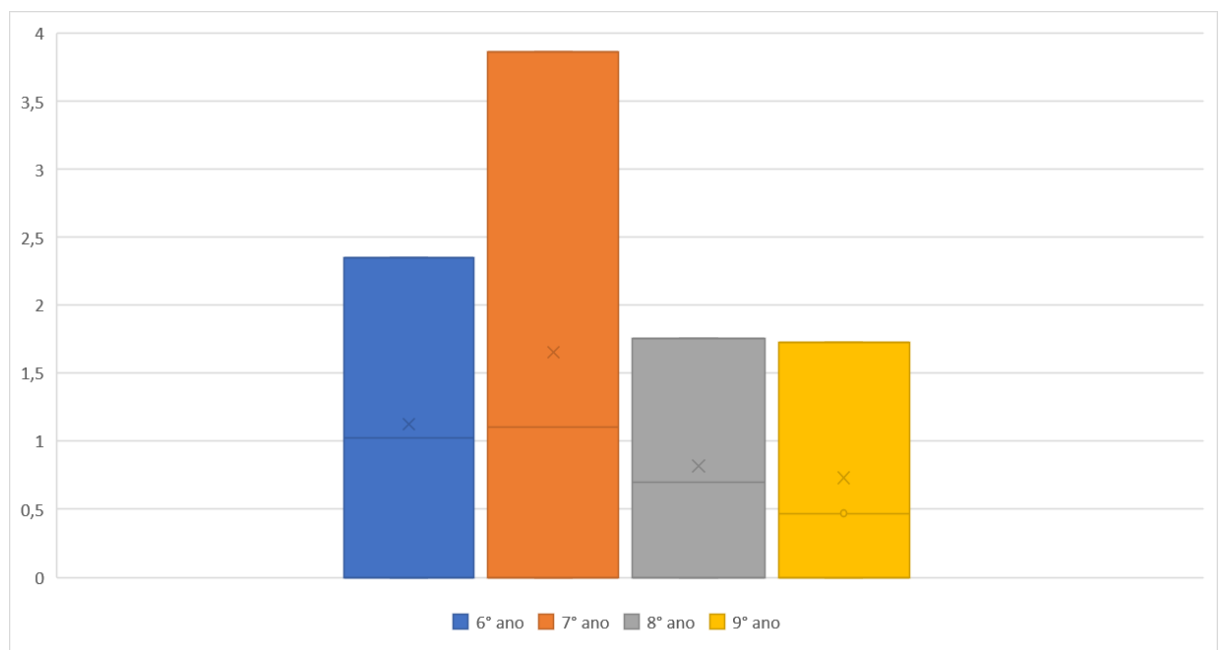
Fonte: Elaboração própria

Verificamos, por meio dessa tabela, que não houve uma redução significativa, em números absolutos, em relação aos dados produzidos por esses sujeitos e o avanço da escolarização. Com base nisso, buscamos realizar uma análise quantitativa dos dados por sujeito a fim de observar melhor o percurso de cada um durante o processo de escolarização. Além disso, levamos em conta, na esteira do que faz Tenani (2016) e Fiel (2018), o coeficiente gerado a partir da razão entre o número de segmentações de cada texto e o número de palavras escritas ao longo dos quatro anos analisados. Desse modo, realizamos, inicialmente, um levantamento das ocorrências de hiposegmentação, hipersegmentação, híbridos e uso de hífen não convencional e as separamos por sujeito e por ano letivo. Também computamos o total de palavras escritas por cada sujeito em cada texto e em cada ano letivo. Em seguida, somamos as ocorrências de cada tipo de segmentação em determinado ano letivo de cada sujeito e depois as dividimos pelo valor total de palavras produzidas nos textos - também de cada sujeito e de cada ano. Após, multiplicamos o quociente encontrado por 100 para termos a porcentagem. Traçamos uma média (sinalizada nos gráficos por meio da linha cinza) entre as porcentagens de cada segmentação entre os sujeitos para cada ano letivo. Com isso, identificamos a maior porcentagem e a menor porcentagem em cada ano. Assim, com base nessas considerações, conseguimos notar mais satisfatoriamente se houve diminuição na ocorrência dos dados. Além disso, podemos identificar melhor a amplitude de ocorrências de segmentações e o quão próximo, ou não, os alunos estão dos extremos (maior número de ocorrências e menor número de ocorrências).

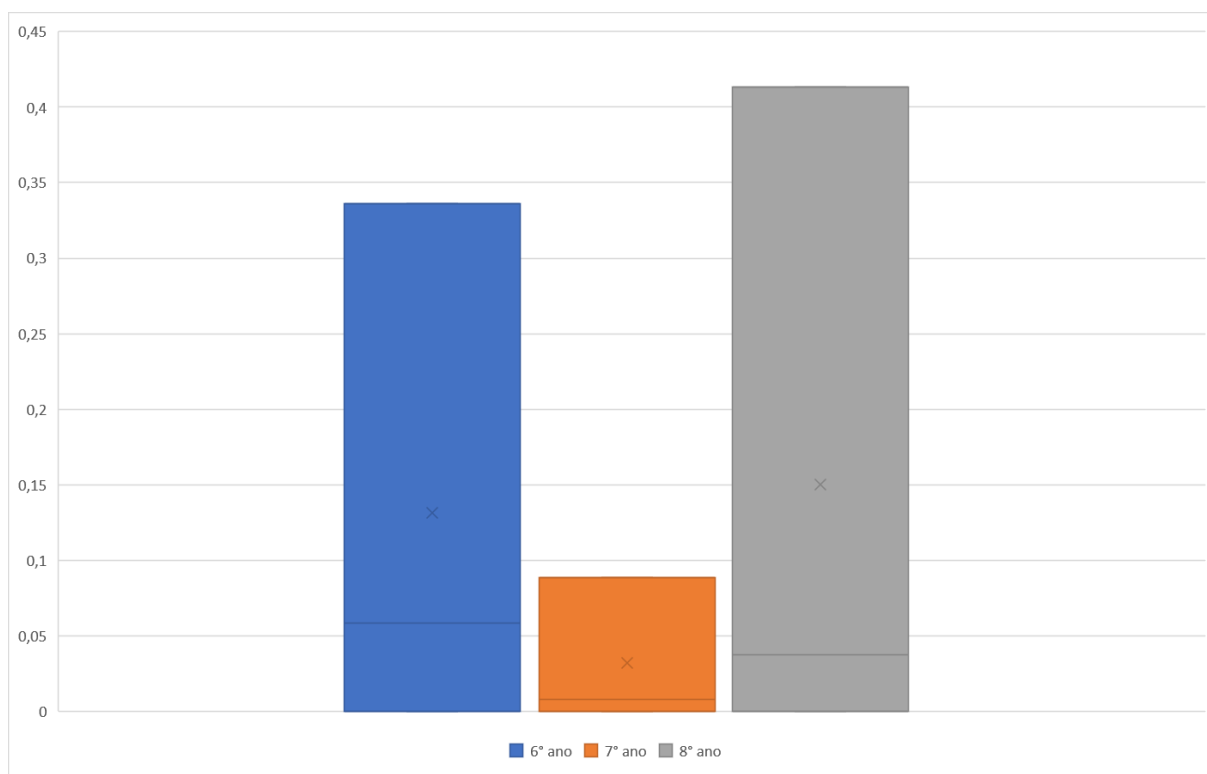


**Gráfico 1. - Distribuição de hipossegmentação por ano letivo**

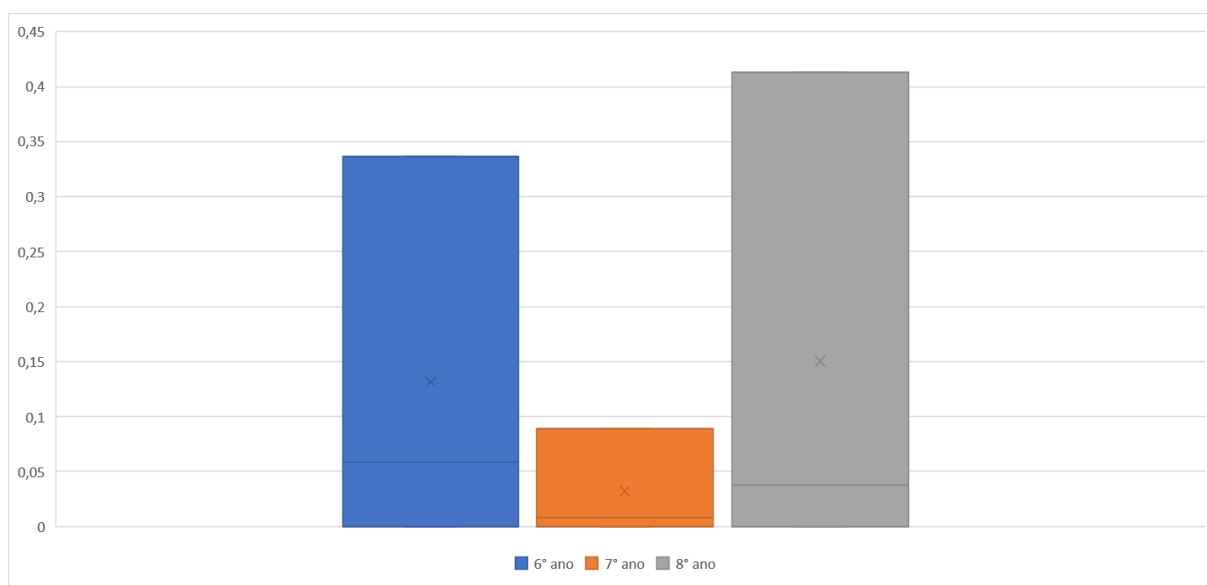
Fonte: Elaboração própria

**Gráfico 2. - Distribuição de hipersegmentação por ano letivo**

Fonte: Elaboração própria

**Gráfico 3. - Distribuição de híbridos por ano letivo**

Fonte: Elaboração própria

**Gráfico 4. - Distribuição de hífen por ano letivo**

Fonte: Elaboração própria

Por meio do levantamento apresentado acima, podemos observar, de acordo com a média, que a maioria dos sujeitos se aproximaram do extremo inferior do gráfico, ou seja, a maioria dos sujeitos produziram poucas ocorrências. Além disso, podemos observar uma diminuição no número de dados, que ocorre do sexto ano para o novo ano. No entanto, ao comparar, também, esse resultado com o total de dados encontrados nos outros dois anos escolares (sétimo e oitavo anos), notamos que a queda apontada parece não se dar de forma linear, isto é, não se dá de forma regular por todo o EF II. No caso das hipossegmentações, houve um aumento de dados no 8º ano, enquanto no caso das hipersegmentações, esse aumento ocorreu no 7º ano. Em decorrência desse resultado, entendemos a necessidade de investigar mais atentamente o oitavo e sétimo ano em relação aos demais anos do EF II, já que outros estudos (ver TENANI, 2010, 2011b), sobre a análise do mesmo tipo de objeto e etapa escolar, chegaram a resultados similares aos nossos.

Ao analisarmos, assim como fez Silva (2014), os dados por tipos, percebemos que tanto nas hipossegmentações quanto nas hipersegmentações, há um conjunto de dados bastante recorrentes no 7º e 8º ano do mesmo tipo. Para as hipossegmentações, há grandes ocorrências de “porisso”, “oque” e estruturas proclíticas e enclíticas grafadas sem separação, como “meapaixonei” (me apaixonei) e “perseguiilos” (persegui-los). Já no caso das hipersegmentações, as ocorrências mais típicas são “de mais” (demais) e em palavras cujas sílabas pretônicas se identificam com elementos escritos da língua, como: “a” ( “a gora”).

Isso demonstra que, embora não haja uma diminuição linear entre o 7º e 8º ano, os tipos de segmentações são os mesmos, evidenciando que as segmentações não convencionais diminuem, mas os tipos se mantêm. Portanto, as dificuldades enfrentadas quanto aos limites de palavra escrita que permanecem nessa etapa escolar tendem a se localizar em certas palavras e não, potencialmente, em qualquer palavra.

Em relação aos dados envolvendo o uso não convencional de hífen, constatamos que grande parte da ocorrência desses dados se concentra na tentativa de grafias com a estrutura de verbo+pronome enclítico. É interessante notar, ainda, que a estrutura enclítica não é a preferencial do português brasileiro (*te quero* ao invés de *quero-te* – cf. BISOL, 1996, 2000, 2005), sendo, portanto, um tipo de estrutura ensinada na escola e por ela valorizada. Além disso, ao considerarmos que, das nove ocorrências de hífen, seis envolvem uma tentativa de grafar tal estrutura (cheguei-la), nossa afirmação se torna mais concreta. Dessa forma, as tentativas em grafar as estruturas verbo+pronome enclítico dizem respeito ao imaginário que o escrevente tem do que seja a escrita e, de modo mais específico, a uma projeção do que seja

a representação do código escrito institucionalizado.

Tenani (2011), ao analisar também os dados não convencionais de hífen, observa que os contextos de colocação do hífen relacionam-se a: (i) no caso das hipossegmentações (a maioria dos dados encontrados), ocorrências em estruturas verbo+clítico, como em “despistalos” (despistá-los); (ii) no caso das hipersegmentações, dois resultados estão em convivência: o primeiro ligado a estruturas de palavras compostas, como, por exemplo, “mau-humor”, e o segundo relacionado, também, com estruturas verbo+clítico, como em “estava-mos” (estávamos). Em nossa pesquisa, constatamos resultados similares, tendo ocorrências de estruturas verbais enclíticas (seis dados), seguidas de palavras compostas, como “super-mercado”. A última ocorrência diz respeito à estrutura “ela-se” (para: ela se foi). Em todas as ocorrências, Tenani (2011) atribui o uso do hífen ao processo de letramento dos escreventes, pois interpreta que as estruturas em que esse recurso gráfico aparece estão relacionadas às estruturas mais trabalhadas nas atividades escolares de escrita - principalmente, no período dos anos finais do EF.

Em relação aos híbridos, é importante lembrar que das sete ocorrências, cinco eram de um único sujeito. Já as outras duas foram feitas por outros dois sujeitos (9 e 90) no sexto e no sétimo ano respectivamente. Por conta dessa característica, há uma importância em analisar mais aprofundadamente não só os híbridos, mas o sujeito 27, responsável pela maioria das ocorrências e pela alteração no decréscimo delas. No capítulo adiante, detalharemos mais sobre esses dados.

Por meio dessa análise, percebemos que é possível observar mais claramente uma diminuição das grafias não convencionais com o decorrer da escolarização, ainda mais quando observamos o sexto e o nono ano, em que a diminuição é bastante clara. Por outro lado, todos os sujeitos ainda terminam o EF II apresentando ocorrências de segmentação não convencional. Além disso, olhar o tipo de dado também é expressivo nessa análise, mas outros fatores podem estar envolvidos nas segmentações dos sujeitos *outliers*, como trataremos no capítulo 3.4.

Acreditamos, ainda, que a não regularidade na diminuição dos dados seja “natural”, uma vez que, durante o processo de escolarização, o escrevente vai se construindo por meio de sua inserção nos mais diversos usos da linguagem. Assim, durante esse processo, o escrevente aprende novas estruturas e passa a transitar em novas (e também antigas) noções que ele tem sobre o que acredita ser sua escrita e o código institucionalizado.

Assim, de acordo com a perspectiva de escrita assumida nesta dissertação (CORRÊA,

2004), acreditamos que a diminuição não regular das grafias não convencionais do hífen é decorrente da inserção do escrevente nos mais diversos usos (formais/informais) da linguagem durante a sua escolarização (tanto dentro quanto fora da escola).

### **3.2 Segmentação de palavras: características prosódicas por sujeito**

Nesta seção, nosso objetivo é apresentar as principais características prosódicas das segmentações não convencionais produzidas por sujeito. Para isso, organizamos os dados apontando as características prosódicas mobilizadas para, em seguida, encontrar regularidades ou divergências entre os sujeitos analisados e, também, entre essas características e aquelas já descritas para os demais sujeitos do EF II. Acreditamos que as junções e separações entre palavras na escrita observadas nos textos aqui analisados também são ancoradas na configuração de constituintes prosódicos. Ao identificá-las, no entanto, busca-se descrever em que medida há regularidades linguísticas de natureza fonológica para esse conjunto particular de dados.

#### **3.2.1 Análise de hipossegmentações**

Apresentamos, nas tabelas abaixo, a estrutura prosódica de grafias em que há uso não convencional de branco, resultando em hipossegmentação. Assim, nesta análise, levantamos as ocorrências de hipossegmentações e as classificamos de acordo com os constituintes prosódicos envolvidos em cada uma. Em seguida, organizamos todas as ocorrências encontradas, dispostas nas tabelas, e colocamos as quantidades daquele tipo de estrutura encontrada em cada ano letivo e para cada sujeito.

Tabela 5 – Hipossegmentação e estrutura prosódica – Sujeito 9.

| Ano    | cl + ω | cl + cl | ω + ω | ω + cl | ω + cl + ω | Total |
|--------|--------|---------|-------|--------|------------|-------|
| 6º ano | 4      | 1       | 3     |        | 1          | 9     |
| 7º ano | 6      | 1       | 1     | 1      |            | 9     |
| 8º ano | 6      |         | 1     | 1      | 1          | 9     |
| 9º ano | 3      | 6       |       |        |            | 9     |
| Total  | 19     | 8       | 5     | 2      | 2          | 36    |

Tabela 6 - Hipossegmentação e estrutura prosódica – Sujeito 16.

| Ano    | cl + ω | ω + ω | ω + cl | Total |
|--------|--------|-------|--------|-------|
| 6º ano | 10     | 2     |        | 12    |
| 7º ano | 18     | 2     | 3      | 23    |
| 8º ano | 5      | 1     | 1      | 7     |
| 9º ano | 5      | 0     | 1      | 6     |
| Total  | 38     | 5     | 5      | 48    |

Tabela 7 - Hipossegmentação e estrutura prosódica – Sujeito 25.

| Ano    | cl + ω | cl + cl | ω + ω | ω + cl | Total |
|--------|--------|---------|-------|--------|-------|
| 6º ano | 3      | 1       |       |        | 4     |
| 7º ano | 16     | 8       | 4     | 1      | 29    |
| 8º ano | 7      | 5       | 1     |        | 13    |
| 9º ano | 12     | 3       | 4     |        | 19    |
| Total  | 38     | 17      | 9     | 1      | 65    |

Tabela 8 - Hipossegmentação e estrutura prosódica – Sujeito 27.

| <b>Ano</b>    | <b>cl +<br/>ω</b> | <b>cl +<br/>cl</b> | <b>ω +<br/>ω</b> | <b>ω +<br/>cl</b> | <b>ω + ω +<br/>ω</b> | <b>cl + cl +<br/>ω</b> | <b>cl + ω + cl +<br/>ω</b> | <b>Total</b> |
|---------------|-------------------|--------------------|------------------|-------------------|----------------------|------------------------|----------------------------|--------------|
| <b>6º ano</b> | 8                 | 6                  | 3                |                   | 1                    |                        |                            | 18           |
| <b>7º ano</b> | 2                 |                    | 2                |                   |                      |                        |                            | 4            |
| <b>8º ano</b> | 5                 |                    | 1                | 1                 |                      | 1                      |                            | 8            |
| <b>9º ano</b> | 1                 | 1                  | 1                |                   |                      |                        | 1                          | 4            |
| <b>Total</b>  | 17                | 7                  | 7                | 1                 | 1                    | 1                      | 1                          | 34           |

Tabela 9 – Hipossegmentação e estrutura prosódica – Sujeito 60.

| <b>Ano</b>    | <b>cl + ω</b> | <b>cl + cl</b> | <b>ω + ω</b> | <b>Total</b> |
|---------------|---------------|----------------|--------------|--------------|
| <b>6º ano</b> | 8             | 1              | 2            | 11           |
| <b>7º ano</b> | 1             |                | 2            | 3            |
| <b>8º ano</b> |               | 1              | 1            | 2            |
| <b>9º ano</b> | 2             | 1              | 1            | 4            |
| <b>Total</b>  | 11            | 3              | 6            | 20           |

Tabela 10 - Hipossegmentação e estrutura prosódica – Sujeito 69.

| <b>Ano</b>    | <b>cl + ω</b> | <b>cl + cl</b> | <b>ω + ω</b> | <b>ω + cl</b> | <b>Total</b> |
|---------------|---------------|----------------|--------------|---------------|--------------|
| <b>6º ano</b> | 5             |                |              | 1             | 6            |
| <b>7º ano</b> | 2             | 1              | 1            |               | 4            |
| <b>8º ano</b> |               |                |              |               |              |
| <b>9º ano</b> | 1             | 9              |              |               | 10           |
| <b>Total</b>  | 8             | 10             | 1            | 1             | 20           |

Tabela 11 - Hipossegmentação e estrutura prosódica – Sujeito 73.

| <b>Ano</b>    | <b>cl + ω</b> | <b>cl + cl</b> | <b>ω + ω</b> | <b>ω + cl</b> | <b>Total</b> |
|---------------|---------------|----------------|--------------|---------------|--------------|
| <b>6º ano</b> | 3             | 1              |              |               | 4            |
| <b>7º ano</b> | 9             | 1              | 6            | 1             | 17           |
| <b>8º ano</b> | 5             | 1              | 1            |               | 7            |
| <b>9º ano</b> | 2             | 1              | 1            |               | 4            |
| <b>Total</b>  | 19            | 4              | 8            | 1             | 32           |

Tabela 12 - Hipossegmentação e estrutura prosódica – Sujeito 90.

| <b>Ano</b>    | <b>cl + ω</b> | <b>cl + cl</b> | <b>ω + ω</b> | <b>Total</b> |
|---------------|---------------|----------------|--------------|--------------|
| <b>6º ano</b> | 3             | 3              |              | 6            |
| <b>7º ano</b> | 5             | 1              | 6            | 12           |
| <b>8º ano</b> |               | 1              |              | 1            |
| <b>9º ano</b> | 6             | 1              |              | 7            |
| <b>Total</b>  | 14            | 6              | 6            | 26           |



Tabela 13 - Hipossegmentação e estrutura prosódica – Sujeito 93.

| <b>Ano</b>    | <b>cl + ω</b> | <b>cl + cl</b> | <b>ω + ω</b> | <b>Total</b> |
|---------------|---------------|----------------|--------------|--------------|
| <b>6º ano</b> |               | 5              |              | 5            |
| <b>7º ano</b> | 2             | 1              |              | 3            |
| <b>8º ano</b> |               | 7              | 1            | 8            |
| <b>9º ano</b> |               | 2              |              | 2            |
| <b>Total</b>  | 2             | 15             | 1            | 18           |

Tabela 14 - Hipossegmentação e estrutura prosódica – Sujeito 107.

| <b>Ano</b>    | <b>cl + ω</b> | <b>cl + cl</b> | <b>ω + ω</b> | <b>ω + cl</b> | <b>Total</b> |
|---------------|---------------|----------------|--------------|---------------|--------------|
| <b>6º ano</b> | 2             |                |              | 2             | 4            |
| <b>7º ano</b> | 3             |                | 2            |               | 5            |
| <b>8º ano</b> |               |                |              |               |              |
| <b>9º ano</b> | 7             | 7              |              | 2             | 16           |
| <b>Total</b>  | 12            | 7              | 2            | 4             | 25           |

Tabela 15 - Hipossegmentação e estrutura prosódica – Sujeito 112.

| <b>Ano</b>    | <b>cl + ω</b> | <b>cl + cl</b> | <b>ω + ω</b> | <b>ω + cl</b> | <b>Total</b> |
|---------------|---------------|----------------|--------------|---------------|--------------|
| <b>6º ano</b> |               |                | 1            |               | 1            |
| <b>7º ano</b> |               | 1              |              | 2             | 3            |
| <b>8º ano</b> | 15            | 1              |              |               | 16           |
| <b>9º ano</b> | 1             |                |              |               | 1            |
| <b>Total</b>  | 16            | 2              | 1            | 2             | 21           |

Tabela 16 - Hipossegmentação e estrutura prosódica – Sujeito 115.

| Ano          | cl + ω | cl + cl | ω + ω | ω + cl | Total |
|--------------|--------|---------|-------|--------|-------|
| 6º ano       | 4      | 2       | 1     | 1      | 8     |
| 7º ano       | 2      |         |       |        | 2     |
| 8º ano       |        |         |       |        |       |
| 9º ano       | 2      | 2       |       |        | 4     |
| <b>Total</b> | 8      | 4       | 1     | 1      | 14    |

Com base no produto da análise realizada, observado nas tabelas acima, encontramos quatro tipos de sequência mais predominante, a saber: Tipo 1 - Clítico + palavra prosódica (cl + w), apresentando 202 ocorrências; Tipo 2 - Clítico + Clítico (cl + cl), apresentando 83 ocorrências; Tipo 3 - Palavra prosódica + palavra prosódica (w + w), apresentando 52 ocorrências; e Tipo 4 - Palavra prosódica + clítico (w + cl), apresentando 18 ocorrências. As demais sequências, agrupamos em Tipo 5 - Outras combinações, que apresentaram quatro ocorrências, a saber: palavra prosódica + palavra prosódica + palavra prosódica (w + w + w), palavra prosódica + clítico + palavra prosódica (w + cl + w), clítico + clítico + palavra prosódica (cl + cl + w) e clítico + palavra prosódica + clítico (cl + w + cl). Dos 12 sujeitos analisados, dez produziram com maior frequência a sequência cl + w; os outros dois sujeitos (sujeito número 69 e sujeito número 93) produziram, com maior frequência, dados de cl + cl, seguidos de cl + w.

### 3.2.2 Análise de hipersegmentações

Apresentamos, nas tabelas a seguir, a estrutura prosódica de grafias em que há uso não convencional de branco, resultando em hipersegmentação. Nesta análise, realizamos o mesmo

que na anterior: levantamos as ocorrências de hipersegmentações e as classificamos de acordo com os constituintes prosódicos envolvidos em cada uma. Em seguida, organizamos todas as ocorrências encontradas, dispostas nas tabelas, e colocamos as quantidades daquele tipo de estrutura encontrada em cada ano letivo e para cada sujeito.

Tabela 17 - Hipersegmentação e estrutura prosódica – Sujeito 9

| Ano    | $\sigma + \sigma$ | $\sigma + \Sigma$ | Total |
|--------|-------------------|-------------------|-------|
| 6º ano | 3                 | 2                 | 5     |
| 7º ano | 2                 |                   | 2     |
| 8º ano |                   |                   |       |
| 9º ano | 1                 |                   | 1     |
| Total  | 6                 | 2                 | 8     |

Tabela 18 - Hipersegmentação e estrutura prosódica – Sujeito 16

| Ano    | $\sigma + \sigma$ | $\sigma + \Sigma$ | $\sigma + \Sigma + \sigma$ | $\sigma + \sigma + \sigma$ | Total |
|--------|-------------------|-------------------|----------------------------|----------------------------|-------|
| 6º ano | 4                 |                   | 1                          |                            | 5     |
| 7º ano | 14                | 7                 |                            |                            | 21    |
| 8º ano | 7                 | 1                 |                            | 1                          | 9     |
| 9º ano | 3                 | 1                 |                            |                            | 4     |
| Total  | 28                | 9                 | 1                          | 1                          | 39    |

Tabela 19 - Hipersegmentação e estrutura prosódica – Sujeito 25

| Ano    | $\sigma + \sigma$ | $\sigma + \Sigma$ | $\Sigma + \sigma$ | $\Sigma + \Sigma$ | Total |
|--------|-------------------|-------------------|-------------------|-------------------|-------|
| 6º ano | 1                 |                   |                   | 1                 | 2     |
| 7º ano | 1                 |                   |                   |                   | 1     |
| 8º ano | 1                 | 2                 | 1                 |                   | 4     |
| 9º ano |                   | 3                 |                   |                   | 3     |
| Total  | 3                 | 5                 | 1                 | 1                 | 10    |

Tabela 20 - Hipersegmentação e estrutura prosódica – Sujeito 27

| Ano    | $\sigma + \sigma$ | $\sigma + \Sigma$ | $\Sigma + \sigma$ | $\Sigma + \Sigma$ | $\Sigma + \sigma + \sigma$ | $\sigma + \sigma + \sigma$ | $\sigma + \sigma + \Sigma$ | Total |
|--------|-------------------|-------------------|-------------------|-------------------|----------------------------|----------------------------|----------------------------|-------|
| 6º ano | 6                 | 6                 |                   |                   | 1                          | 1                          |                            | 14    |
| 7º ano | 9                 | 3                 | 1                 | 1                 |                            |                            |                            | 14    |
| 8º ano | 3                 | 2                 |                   |                   |                            |                            |                            | 5     |
| 9º ano | 1                 | 2                 |                   |                   |                            |                            | 2                          | 5     |
| Total  | 19                | 13                | 1                 | 1                 | 1                          | 1                          | 2                          | 38    |

Tabela 21 - Hipersegmentação e estrutura prosódica – Sujeito 60

| Ano    | $\sigma + \sigma$ | $\sigma + \Sigma$ | $\Sigma + \Sigma$ | $\sigma + \sigma + \sigma$ | Total |
|--------|-------------------|-------------------|-------------------|----------------------------|-------|
| 6º ano |                   | 4                 |                   |                            | 4     |
| 7º ano | 4                 | 5                 | 2                 |                            | 11    |
| 8º ano |                   | 1                 |                   |                            | 1     |
| 9º ano |                   | 2                 |                   | 1                          | 3     |
| Total  | 4                 | 12                | 2                 | 1                          | 19    |

Tabela 22 - Hipersegmentação e estrutura prosódica – Sujeito 69

| <b>Ano</b>    | <b><math>\sigma + \sigma</math></b> | <b><math>\sigma + \Sigma</math></b> | <b>Total</b> |
|---------------|-------------------------------------|-------------------------------------|--------------|
| <b>6º ano</b> |                                     | 2                                   | 2            |
| <b>7º ano</b> | 1                                   | 1                                   | 2            |
| <b>8º ano</b> |                                     | 1                                   | 1            |
| <b>9º ano</b> | 2                                   | 1                                   | 3            |
| <b>Total</b>  | 3                                   | 5                                   | 8            |

Tabela 23 - Hipersegmentação e estrutura prosódica – Sujeito 73

| <b>Ano</b>    | <b><math>\sigma + \sigma</math></b> | <b><math>\sigma + \Sigma</math></b> | <b><math>\Sigma + \sigma</math></b> | <b>Total</b> |
|---------------|-------------------------------------|-------------------------------------|-------------------------------------|--------------|
| <b>6º ano</b> |                                     | 3                                   |                                     | 3            |
| <b>7º ano</b> | 4                                   | 2                                   | 1                                   | 7            |
| <b>8º ano</b> | 5                                   | 1                                   |                                     | 6            |
| <b>9º ano</b> | 2                                   | 7                                   | 1                                   | 10           |
| <b>Total</b>  | 11                                  | 13                                  | 2                                   | 26           |

Tabela 24 - Hipersegmentação e estrutura prosódica – Sujeito 90

| <b>Ano</b>    | <b><math>\sigma + \sigma</math></b> | <b><math>\sigma + \Sigma</math></b> | <b>Total</b> |
|---------------|-------------------------------------|-------------------------------------|--------------|
| <b>6º ano</b> | 1                                   | 7                                   | 8            |
| <b>7º ano</b> | 7                                   | 9                                   | 16           |
| <b>8º ano</b> |                                     | 1                                   | 1            |
| <b>9º ano</b> |                                     | 1                                   | 1            |
| <b>Total</b>  | 8                                   | 18                                  | 26           |

Tabela 25 - Hipersegmentação e estrutura prosódica – Sujeito 93

| Ano    | $\sigma + \sigma$ | Total |
|--------|-------------------|-------|
| 6º ano |                   |       |
| 7º ano | 1                 | 1     |
| 8º ano |                   |       |
| 9º ano |                   |       |
| Total  | 1                 | 1     |

Tabela 26 - Hipersegmentação e estrutura prosódica – Sujeito 107

| Ano    | $\sigma + \sigma$ | $\sigma + \Sigma$ | $\Sigma + \sigma$ | Total |
|--------|-------------------|-------------------|-------------------|-------|
| 6º ano | 1                 | 3                 | 1                 | 5     |
| 7º ano | 2                 | 1                 |                   | 3     |
| 8º ano |                   |                   |                   |       |
| 9º ano | 5                 |                   |                   | 5     |
| Total  | 8                 | 4                 | 1                 | 13    |

Tabela 27 - Hipersegmentação e estrutura prosódica – Sujeito 112

| Ano          | $\sigma + \sigma$ | $\sigma + \Sigma$ | $\Sigma + \sigma$ | $\Sigma + \Sigma$ | Total |
|--------------|-------------------|-------------------|-------------------|-------------------|-------|
| 6º ano       |                   | 1                 | 1                 | 1                 | 3     |
| 7º ano       | 1                 |                   | 1                 |                   | 2     |
| 8º ano       | 2                 | 1                 |                   |                   | 3     |
| 9º ano       |                   | 1                 |                   |                   | 1     |
| <b>Total</b> | 3                 | 3                 | 2                 | 1                 | 9     |

Tabela 28 - Hipersegmentação e estrutura prosódica – Sujeito 115

| Ano          | $\sigma + \sigma$ | $\Sigma + \sigma$ | $\sigma + \sigma + \sigma$ | Total |
|--------------|-------------------|-------------------|----------------------------|-------|
| 6º ano       | 1                 |                   |                            | 1     |
| 7º ano       | 1                 | 1                 |                            | 2     |
| 8º ano       |                   |                   | 1                          | 1     |
| 9º ano       | 1                 | 2                 |                            | 3     |
| <b>Total</b> | 3                 | 3                 | 1                          | 7     |

Com base na análise em constituintes prosódicos das hipersegmentações, percebemos que a maioria das palavras segmentadas correspondia às seguintes organizações prosódicas: Tipo 1 – Sílabas + Sílabas ( $\sigma + \sigma$ ), que apresentou 97 ocorrências; Tipo 2 – Sílabas + Pé métrico ( $\sigma + \Sigma$ ), apresentando 84 ocorrências; Tipo 3 – Pé métrico + Sílabas ( $\Sigma + \sigma$ ), com quinze ocorrências. Os demais dados foram agrupados em Tipo 5 – Outras combinações, apresentando oito ocorrências diversas, a saber: sílaba + sílaba + sílaba, sílaba + pé métrico + sílaba, pé métrico + sílaba + sílaba e sílaba + sílaba + pé métrico.

Também realizamos uma análise rítmica da convenção ortográfica, uma vez que partimos da hipótese de que as informações rítmicas possam ser pontos de ancoragem para o tipo de limite gráfico não convencional de palavras registrado na escrita. Com base nisso, encontramos as seguintes estruturas rítmicas: Estrutura fraca-forte (- \*), com 83 ocorrências,

como em “então”; Estrutura fraca-forte-fraca (- \* -), com 73 ocorrências, como em “embora”; Estrutura forte-fraca-forte-fraca (\* - \* -), com 14 ocorrências, como em “recompensa”; Estrutura forte-fraca (\* -), com 10 ocorrências, como em “desde”; e Estrutura forte-fraca-forte, com oito ocorrências, como em “saberia”. As demais estruturas foram agrupadas em Estrutura – Outros tipos, com 16 ocorrências diversas, a saber: forte-fraca-forte; forte-fraca-fraca-forte ou forte-fraca-forte-fraca, a depender da organização possível; e forte-fraca-fraca-forte-fraca ou fraca-forte-fraca-forte-fraca. Além disso, é importante constatar que a dupla possibilidade de organização rítmica da palavra escrita, aqui marcada, acontece devido às possibilidades de atribuição de acento secundário, por exemplo em “apaixonei” (apaixonei ou apaixonei).

### 3.2.3 Análise de híbridos

Apresentamos, na tabela abaixo, a estrutura prosódica de grafias de híbridos identificados nos textos da amostra. Primeiramente, levantamos as ocorrências de híbridos e as classificamos de acordo com os constituintes prosódicos envolvidos em cada uma. Dessa forma, colocamos a quantidade de dados encontrados em cada ano e para cada sujeito e, ainda, o tipo de estrutura mobilizada naquele ano e por determinado sujeito.



Tabela 29 – Híbridos e estrutura prosódica

| Ano    | Sujeito 9         | Sujeito 27           | Sujeito 90        | Total |
|--------|-------------------|----------------------|-------------------|-------|
| 6º ano | 1                 | 3                    |                   | 4     |
|        | $\omega + \omega$ | $\omega + \omega$    |                   |       |
|        |                   | $\omega + \text{cl}$ |                   |       |
|        |                   | $\text{cl} + \omega$ |                   |       |
| 7º ano |                   |                      | 1                 | 1     |
|        |                   |                      | $\omega + \omega$ |       |
| 8º ano |                   | 2                    |                   | 2     |
|        |                   | $\omega + \text{cl}$ |                   |       |
|        |                   | $\text{cl} + \omega$ |                   |       |
| 9º ano |                   |                      |                   | 0     |
| Total  | 1                 | 5                    | 1                 | 7     |

A estrutura mais recorrente nos dados de híbridos de sujeitos *outliers* foi  $\omega + \omega$  com três ocorrências, seguida de  $\omega + \text{cl}$  e  $\text{cl} + \omega$ , ambas com duas ocorrências cada. É interessante observar que a sequência que tende a gerar híbridos no EF II é a sequência  $\text{cl} + \omega$ . Por outro lado, percebemos que dos sete dados de híbridos, quatro apresentaram fronteira à direita da sílaba tônica.

### 3.2.4 Análise de dados com hífen

Nesta análise, apresentamos a estrutura prosódica de grafias em que há uso não convencional de hífen. Na tabela abaixo, colocamos o tipo de estrutura encontrada, conforme os constituintes prosódicos mobilizados, para cada sujeito e em cada ano letivo.

**Tabela 30 – Hífen não convencional e estrutura prosódica**

| <b>Ano</b>    | <b>Sujeito 9</b> | <b>Sujeito 25</b> | <b>Sujeito 69</b> | <b>Sujeito 73</b> | <b>Sujeito 115</b> | <b>Total</b> |
|---------------|------------------|-------------------|-------------------|-------------------|--------------------|--------------|
| <b>6º ano</b> | ω + cl           |                   |                   |                   |                    | 1            |
| <b>7º ano</b> | ω + ω            | ω + ω             |                   | ω + cl            | ω + ω              | 4            |
| <b>8º ano</b> | ω + ω            |                   |                   |                   |                    | 1            |
| <b>9º ano</b> | ω + cl           |                   | ω + ω             |                   |                    | 3            |
|               | ω + cl           |                   |                   |                   |                    |              |
| <b>Total</b>  | 5                | 1                 | 1                 | 1                 | 1                  | 9            |

A partir da estrutura prosódica das grafias, podemos notar que a estrutura que tende a gerar segmentações desse tipo é a sequência **ω + ω**, uma vez que, de nove dados, cinco correspondem a essa sequência; seguida de **ω + cl**, com quatro ocorrências. Apesar disso, cabe salientar que, no tipo **ω + ω**, das cinco grafias encontradas, três correspondiam a uma tentativa de grafar a estrutura verbal enclítica, mas envolvendo o advérbio “lá”, como “voltei-lá”.

### 3.3 Tipos de dados predominantes por sujeito

Por meio da análise apresentada na seção anterior, percebemos que a maioria dos sujeitos apresentaram tendências semelhantes às aquelas típicas do EF II. Dos 12 sujeitos analisados, dez produziram a maioria dos dados de hipossegmentação referente à sequência mais predominante para o EF II: cl + w (FIEL, 2018). Por exemplo, o sujeito 25 produziu 65 hipossegmentações, sendo 38 da sequência de cl + w (“anoite”) e 17 da sequência de cl + cl (“eo”). Apesar de ter sido considerado *outlier* em três dos quatro anos do EF II, as grafias produzidas correspondem ao que é típico do EF II.

Os outros dois sujeitos - 69 e 93 -, no entanto, apresentaram características ligeiramente diferentes, uma vez que apresentaram tendência à hipossegmentação entre clíticos, tendo como estrutura mais recorrente a sequência de cl + cl, que, embora não seja a mais típica, é a segunda (FIEL, 2018). Por exemplo, o sujeito 93 produziu 18 hipossegmentações, sendo 15 da sequência cl + cl e duas da sequência cl + w. É interessante observar, ainda, que dos 15 dados envolvendo clítico, 12 são ocorrências de “oque” e duas de “doque”. Notamos, com base nisso, que tal estrutura “oque”/“doque” pode ser causadora de dúvidas para esse sujeito. O mesmo foi atestado no sujeito 69, em que dos 20 dados de hipossegmentação produzidos, dez eram da sequência cl + cl e oito da sequência cl + w; e desses dez dados de cl + cl, oito correspondiam a ocorrências de “oque”.

Também observamos que o sujeito 27 apresentou características únicas, uma vez que das 77 segmentações não convencionais, 38 eram hipersegmentações e 34 eram hipossementações, assim, o sujeito produziu hipersegmentações em número maior que as hipossementações, característica que não é típica do EF II (TENANI, 2016). Além disso, ele produziu cinco dos sete dados de híbridos encontrados nesta pesquisa. Soma-se, ainda, que desses cinco dados, apenas dois correspondiam à estrutura que predominantemente gera híbridos, a saber: clítico mais palavra prosódica (BORSATO, 2019).

Em relação às hipersegmentações, percebemos que, dos 204 dados de hipersegmentação, 158 apresentaram fronteira gráfica onde está o acento gráfico. Ademais, na maioria desses dados, a sílaba pretônica isolada correspondia a um clítico fonológico e/ou a uma classe de palavras funcionais (“a migo”), características típicas do EF II (SILVA, 2014). Em relação à estrutura rítmica, notamos que a estrutura dissílaba fraca-forte (- \*) se mostrou

predominante como causadora de diferentes grafias não convencionais, seguida da estrutura trissílaba fraca-forte-fraca (- \* -), e da estrutura polissilábica forte-fraca-forte-fraca (\* - \* -).

No entanto, Silva (2014) atestou que, em sua pesquisa, as estruturas que mais levam a grafias não convencionais e as que resultam em maior número de possibilidades de registros são, respectivamente, palavras trissílabas (ouvindo), e palavras dissílabas (assim). Foi frequente, também, ainda que em menor número, grafias não convencionais de palavras polissilábicas (recompensa). Dessa forma, tais características apontadas por Silva (2014) mostram-se ligeiramente diferentes das atestadas nesta pesquisa, embora envolvam as mesmas estruturas rítmicas.

Por fim, em relação aos constituintes prosódicos, as sequências encontradas foram Sílabas + Sílabas ( $\sigma + \sigma$ ) (“de mais”); Sílabas + Pé métrico ( $\sigma + \Sigma$ ) (“em bora”); e Pé métrico + Sílabas ( $\Sigma + \sigma$ ) (“brinque do”). É interessante observar que as estruturas convencionais que correspondiam a um  $\Sigma$  iambo (alternância fraco/forte), quando hipersegmentadas, passaram a corresponder à combinação duas sílabas (“em tão”). Já em relação às palavras trissílabas, que predominantemente correspondiam a um  $\Sigma$  troqueu, quando hipersegmentadas, eram realizadas em sílabas + pé métrico.

Em relação aos dados de hífen, percebemos que foram mobilizadas apenas duas sequências, sendo a mais predominante a  $\omega + \omega$ , seguida de  $\omega + \text{cl.}$  É interessante observar que, dos cinco dados de  $\omega + \omega$ , três envolvem verbos (“ficamos-la”), o que pode constituir uma pista de que tais hifens correspondam a usos não convencionais, uma vez que pode se tratar de uma ênclise verbal de pronomes átonos, estrutura que, provavelmente, gera dúvidas ao aluno.

Por meio da análise, percebemos que, quanto à hipossegmentação, os sujeitos têm dados semelhantes aos típicos do EF II. No entanto, dois sujeitos produziram outras combinações de palavras, os sujeitos 16 e 27, que foram diferentes das típicas. Quanto às hipersegmentações, a maioria dos dados apresenta características típicas do EF II. Percebemos, ainda, que, em relação aos aspectos prosódicos, os clíticos fonológicos favorecem grafias não convencionais, como as hipossegmentações e os híbridos, característica comum aos demais sujeitos do EF II. Dessa forma, podemos observar que, apesar de serem considerados *outliers* devido à alta frequência de hipossegmentações, tais sujeitos produzem, no geral, dados muito próximos aos típicos dos demais sujeitos do EF II.

### 3.4 Um sujeito singular?

Na parte anterior de nossa análise, privilegiamos a discussão das ocorrências que permitiram o apontamento de regularidades linguísticas que contribuíram para que caracterizássemos as segmentações não convencionais de palavras dos sujeitos que foram estatisticamente considerados *outliers*, segundo Fiel (2018). Por meio dessa análise, observamos que a maioria dos sujeitos apresentaram tendências semelhantes às aquelas típicas do EF II, segundo estudos anteriores feitos por Paranhos (2012), Silva (2014), Tenani (2016), Fiel (2018), dentre outros. No entanto, o sujeito 27 se mostrou particular em relação ao grupo analisado, produzindo dados que chamaram a nossa atenção não só pelas estruturas linguísticas mobilizadas, mas também porque algumas dessas características foram observadas em textos do EF I, como descrito por Chacon (2005, 2006, 2011), Capristano (2004, 2007a, b), Torquette (2016).

Nesta seção, o nosso intuito é o de realizar uma análise qualitativa das segmentações de palavra, mostrando não apenas a estrutura linguística mobilizada, mas também apresentando uma hipótese sobre sua motivação linguística. Tal análise se mostra necessária, pois essas grafias reforçam a nossa interpretação de que essas marcas indicam o modo como o “escrevente lida com sentidos que perpassam o texto ao segmentar os enunciados em porções cujos limites não coincidem com aqueles da palavra convencional” (TENANI, 2008, p. 242). Portanto, não é somente o dado (enquanto estrutura) que nos interessa observar, mas, sim, o modo como a grafia se apresenta no funcionamento do enunciado em que ocorre.

Dada essa nossa motivação, analisaremos os quatro textos produzidos pelo sujeito 27 em que ocorreram os dados de híbridos. A escolha se dá pelo fato de, como já anunciado, esse sujeito ter produzido dados quantitativamente acima da média dos demais alunos do banco em dois dos quatro anos do EF II. Ele também apresentou dados de todos os tipos: hipossegmentações, hipersegmentações e híbridos. Outro ponto a se destacar é que das sete grafias de híbridos produzidas por todos os sujeitos de nossa pesquisa, ele foi responsável por cinco delas. Iniciamos a nossa análise trazendo o texto da Figura 9, reapresentada a seguir para o leitor, em que podemos observar a ocorrência de todos os tipos de segmentações não convencionais, além dos dados de híbrido “as davano” (andavam no) e “da eronave” (da aeronave).

Figura 9. Texto 1 do sujeito 27 (texto apresentado na figura 8)

Um país estranho

Eu tinha exquite uma carta de Roberto e a professora falou que ia se sortindo era Costa.

O sortido ia se dia 07/11/2008 ai tarde da noite ia sair de casa para ir mais país com uma noite especial de Diego Sanchez foi sortido para ir ao meu amigo Andre Compamba ai tinha um sorriso de lado ai logo da noite não foi.

Ai não foi na janela porage o país não, rio jupiter e noturno e sol e lua não passou a lua chegaram 7:00 da noite.

Não desceu da oronave o país era, muito estranho os corpos os danados os os comida de la era muito estranho de mais.

Ai não viu um ate ele tinha 4 olhos.

Os marx de ele era uma Bola tinha um monte de luz.

Os esdora deles era quadrados.

Os comida de la ate que era um leuco costoso e ele era um irais astros.

Era o ultimo dia de viagem para voltar a Terra.

No dia 20/11/08 não chegou a terra as 1:00 hora da manhã ate que foi na era i na que se país estranho.

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

Para uma análise qualitativa dos dados, partimos do texto produzido no 6º ano do Ensino Fundamental II. O critério de seleção foi a ocorrência de dois dados híbridos, que são pouco comuns nos textos do EF, isto é, com baixas ocorrências em relação às demais segmentações tanto na amostra de textos do EF II quanto em textos do EF I, como observam Capristano (2021) e Miranda, Cunha (2009), mas, principalmente, por esse tipo de dado se mostrar como bastante complexo, como buscaremos demonstrar a partir de análise de “as davano” e “da eronave”.

A análise toma como ponto de partida a afirmação de Capristano (2021) sobre esse tipo de dado. Para a autora:

[...] não há, ainda, [...] uma efetiva descrição dessas segmentações, ou seja, ainda não tem sido explorado de forma profícua como ocorre a emergência de mesclas/híbridos/mistos na escrita da criança. Ademais, não há uma problematização satisfatória do seu funcionamento, uma vez que os debates ainda não contemplaram, com exclusividade, o que essas segmentações especiais representariam para o funcionamento da escrita da criança. (CAPRISTANO, 2021, p. 792)

Salientamos, também, que a atenção ao dado de híbrido se justifica por suas características específicas em relação aos outros dois tipos de segmentação, uma vez que

As pesquisas sobre escrita da criança consideram mesclas/híbridos/mistos como uma segmentação não possível de ser explicada de forma satisfatória se classificada, em separado, como episódios de hipossegmentação ou de hipersegmentação. São, portanto, “casos especiais”, merecedores de classificações, descrições e explicações diferentes. Pesquisas como as de Cunha (2010) e de Chacon (2004), por exemplo, entendem que, nesses modos não convencionais de emprego dos espaços em branco, podem-se “[...] observar motivações relativamente iguais às que aparecem nas hipo e hipersegmentações” (CUNHA, 2010, p. 347, grifo meu). Esses dados parecem, também, em alguns momentos, não reproduzir - como nos casos de hiper e hipossegmentação - “[...] padrões rítmico-entonacionais da oralidade e/ou que não se explicam com base em algoritmos como aqueles que definem constituintes da hierarquia prosódica, conforme proposta por Nespor e Vogel (1986)”. (CHACON, 2004, p. 225) Ou seja, identificam-se convergências e divergências entre o funcionamento de mesclas/híbridos/mistos quando comparado ao de hipo e hipersegmentações. (CAPRISTANO, 2021, p. 791-792)

A partir das considerações de Capristano (2021), assumimos que esse tipo de registro gráfico de fronteiras de palavras permite tratar, de modo privilegiado, da relação do sujeito com a noção de palavra e com o texto. Assim, por meio das análises dos híbridos, pudemos observar, de modo mais detalhado, de que maneira o sujeito se relaciona com sua escrita. Também defendemos a hipótese de que a ocorrência de híbridos está fortemente ancorada nas relações de sentidos que podem ser depreendidas do texto analisado. Desse modo, não deixamos de abordar questões linguísticas já antes discutidas, uma vez que, mesmo a partir de ocorrências, grande parte das vezes, diferentes das tendências encontradas, é possível, em



alguma medida, estabelecer vínculos com aspectos mais gerais dos dados analisados.

Antes de fazer a análise do dado híbrido, cabe observar a presença de várias grafias diferentes daquelas previstas pela convenção ortográfica no texto da Figura 9 a fim de melhor compreender as ocorrências desse tipo de dado. Iniciamos pela análise de hipossegmentações produzidas pelo aluno 27. Na hipossegmentação “ea” (ea), é importante lembrar que, na fala, pode ocorrer um sândi vocálico por ditongação [ia]. Essa grafia poderia ser interpretada como uma mera “interferência da fala na escrita”, mas nota-se que, embora na fala tal sequência seja produzida como [ia], o aluno ainda registra a vogal com a letra <e>, provavelmente por ele reconhecer que, embora a pronúncia seja a vogal alta, na escrita a grafia é com <e>, por exemplo “menino > minino”. Interpretamos que o aluno não escreve o enunciado se baseando inteiramente na fala, o que evidencia a “[...] inserção do aprendiz em práticas de letramento (dentro e fora do contexto escolar) que privilegiam um conhecimento metalinguístico sobre a língua em seu modo de enunciação escrito” (CHACON, 2005, p.83).

Em “aurna”, é possível que o aluno tenha sido atravessado pela percepção do sândi vocálico, em que “a”, artigo átono, pode ser prosodizado junto ao hospedeiro “urna”, palavra que tem acento. Essa juntura entre palavras gera uma ditongação, que consiste em um processo “de formação de ditongos com a vogal final de um vocábulo e a inicial de outro, desde que uma das vogais da sequência seja alta e átona” (BISOL, 1996, p. 160).

Em “paraze”, há a junção entre a preposição “para” e o verbo “ver”. A partir de “ver”, observamos a troca da consoante: o aluno troca [v], fricativa labiodental vozeada, por [z], uma fricativa alveolar vozeada. Esses fonemas partilham traços de modo e vozeamento, mas são diferentes quanto ao ponto articulatorio. Ainda observamos o apagamento do <r> da forma do infinitivo “ver”, tipo de registro anteriormente comentado. Desse registro, cabe comentar que, embora na variedade linguística do aluno seja comum a redução da preposição de “para” para “pra” ou “pa” (MARCATO, 2013), o que resultaria em “pravê”, o aluno grafa “para”, segundo a convenção ortográfica, um registro fortemente ancorado em práticas letradas, pois “para” ocorreu apenas na fala de informantes com ensino superior completo, segundo investigação de Marcato (2013) em dados de fala de pessoas de diferentes faixas etárias e pertencentes a diferentes faixas de escolaridade, residentes na região de São José do Rio Preto, mesma cidade onde residiu o aluno 27 durante o desenvolvimento das atividades de produção escrita aqui analisadas.

Passamos a tratar das hipersegmentações que ocorrem no texto analisado. Em “esquito”, há a hipersegmentação da palavra “escrito”. Tal separação pode ter sido motivada por



um aspecto fonológico, uma vez que a fronteira gráfica se dá na sílaba em que há uma proeminência prosódica “esCRItO > es QUIto”, aspecto característico das hipersegmentações, como já descrito anteriormente. Chama a atenção também a troca da letra <c> pelo dígrafo <qu>. Essa troca tem relação com o fato de <c> e <qu> representarem o fonema [k] em certas palavras, como em “casa” e “quilo”. No entanto, as convenções ortográficas preveem que essa letra e dígrafo têm valores distintos em contextos complementares a depender da combinação com a vogal que forma a sílaba. Para o emprego da letra ou do dígrafo, o aluno deve lembrar que: <c> diante de vogais <a, o, u> tem valor de [k] (“casa, coral, curral”) e diante de <e, i>, o valor de [s] (“cedo, cidade”); <qu> diante de <a, o> tem valor de [kw] (“quase, sequoia”) e diante de <e, i> tem valor de [k] (“quero, quibe”). Além da troca de letra por um dígrafo onde não é previsto, houve uma simplificação da estrutura silábica, passando de CCV para o padrão CV, que é preferencial do português brasileiro (PB). Vale observar que essa é uma estrutura sujeita a dois processos variáveis na fala, a saber: o apagamento do tepe, como em “peto” (preto) – cf. Miranda (2007); e a metátese, como em “preda” (pedra) – cf. Coelho (2016). Lançamos a hipótese interpretativa de que a troca entre <c> e <qu> possa ter ocasionado a simplificação de <cri> a <qui>, uma vez que “quri” não é uma sequência gráfica da escrita da língua portuguesa. Dessa descrição salta aos olhos que no ponto de segmentação não convencional de palavra há, também, uma grafia não convencional da consoante e a simplificação da sílaba. É esse conjunto de registros não convencionais de segmento consonantal, sílaba e fronteira de palavra, que torna a escrita desse aluno particularmente interessante por tornar possível explicitar a complexidade do processo de registro da palavra ortográfica.

Em “de mais” (demaís), interpretamos que se trata de um advérbio, uma vez que indica excesso, ou seja, o quão demasiadamente estranhas eram aquelas criaturas. Dessa forma, a grafia esperada pela convenção é “demaís”. Além da distinção morfosintática entre “demaís” e “de mais”, Tenani (2016) analisou que a sílaba pretônica isolada pode corresponder a uma preposição átona e, portanto, um clítico fonológico, que é delimitado por branco conforme ortografia da língua portuguesa. Essas características são típicas dessas hipersegmentações no EF II, como a autora argumenta:

Desse modo, as ocorrências de “de mais” atestadas nos textos do EF II são classificadas como casos de segmentação não-convencional do advérbio “demaís”, sendo as segmentações não-convencionais decorrentes 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”. (TENANI, 2016, p. 103)

Em “de ele” é possível notar a atuação de outra informação da escrita convencional, referente “à escolha de letras para representar certos segmentos que estão nos limites ou fronteiras dos pontos de corte do continuum fônico-gráfico” (TENANI, 2009, p. 115). Soma-se essa característica à duplicação do registro da letra <e> em “de ele” uma vez que a palavra deveria ser grafada “dele”, o que demonstra que o aluno reconhece as palavras “de” e “ele”, que existem na língua portuguesa, ao lado de “dele”, que é registrado em outros trechos de seu texto. Assim, há um trânsito do escrevente entre possibilidades de grafia a partir do fônico e as possibilidades de grafia na escrita.

Já em “na que le”, temos uma dupla fronteira por espaço em branco em uma única palavra. Segundo Tenani (2016), esses tipos de hipersegmentações “[...] permitem tratar de modo privilegiado de características linguístico-discursivas imbricadas a aspectos verbo-visuais envolvidos no processo de produção de enunciados escritos.” (TENANI, 2016, p. 88-89). Além disso, a autora também explicita se tratar de dados pouco recorrentes, uma vez que a maioria das hipersegmentações no EF II apresentam apenas uma fronteira não convencional.

Ao cabo da descrição das grafias não convencionais que ocorreram em um único texto, foi possível demonstrar que estão em jogo a representação de segmentos consonantais e vocálicos, de estruturas silábicas e de limite de palavras simultaneamente. Também buscamos, por meio de análise detalhada, argumentar que informações fonéticas, fonológicas e morfossintáticas são mobilizadas na produção do texto. Não por acaso, passamos à análise da grafia dos dados “da eronave” e “as davano” e, em seguida, argumentamos que essas ocorrências são mais bem interpretadas se consideradas a organização textual e a proposta de produção a que o aluno foi solicitado responder. Iniciamos retomando o trecho em que os dados ocorrem.

Trecho 1. Noi deseu da eronave o pais era, muito estranho os carro as davano ar as comida de la era muito estranho de mais.

Interpretação segundo as convenções ortográficas e gramaticais:

*Nós descemos da eronave. O país era muito estranho. Os carros andavam no ar. As comidas de lá eram muito estranhas demais.*

O dado “da eronave” chama a atenção devido à sua complexidade linguística, uma vez que as fronteiras de hipossegmentação e de hipersegmentação não são tão evidentes. Seria possível dizer que ocorre uma hipossegmentação entre “da” e “aeronave”, já que o processo degeminação que ocorre na fala “d[ae]ronave”, descrito como uma fusão de vogais idênticas, é registrado na escrita, pois a letra <a> em “aeronave” é registrada apenas junto a palavra “da”. Cabe salientar, ainda, que o processo de degeminação no PB é descrito como fusão de vogais idênticas ou semelhantes, e, segundo Bisol (2002), há a justaposição dessas vogais, resultando em uma vogal longa, que, por sua vez, será encurtada por degeminação, como em na adolescência > n[a]dolescência. No caso do fenômeno da degeminação, também há a desassociação de um nó silábico, pois a vogal longa se associa à segunda sílaba da sequência para, em seguida, acontecer a fusão ou o encurtamento dessa vogal (Bisol 2002, p. 234).

Já no caso de “as davano”, retomamos a análise feita anteriormente para uma melhor caracterização de sua complexidade. Inicialmente, constatamos que “as davano” corresponde à forma convencional “andavam no”, porém, constatamos que, neste texto, não há concordância de plural entre sujeito e verbo, como em “nói deseui” e “as comida era”, de modo que supomos que o trecho em análise possa ser interpretado como “os carro andava no ar”. A partir dessa interpretação, supomos que ocorre uma hipersegmentação da primeira sílaba pretônica “an”, grafada como “as” (há uma troca de <n> por <s> em coda silábica), e uma hipossegmentação entre “andava” e “no”, o que pode ser capturado como um processo em que “andava no” passa a “as davano”. Observa-se que essa junção à esquerda da preposição átona “no” em relação ao hospedeiro é característica menos recorrente nos dados do EF II, como demonstrou Fiel (2018). Já a troca entre consoantes revela características complexas do próprio processo de escrita, sendo elas: aspecto fonético, pois os sons [s] e [n] são alveolares, sendo um traço em comum entre as consoantes, apesar de modo e vozeamento serem diferentes; aspecto fonológico, já que tanto <s> quanto <n> são consoantes que ocorrem em fim de sílaba, portanto, são letras que representam consoantes que podem ocupar a mesma posição de coda silábica; e aspecto ortográfico, uma vez que não se emprega <n> em coda silábica em fim de palavra. Desse modo, a troca de <n> por <s> em fim de sílaba não é aleatória, mas ancorada em aspectos fonético, fonológico e ortográfico do português.

Outro fato linguístico importante a considerar é que a inserção do branco, que gera a

hipersegmentação, ocorre justamente entre a sílaba pretônica e tônica da forma verbal. Essa consideração encontra respaldo nos estudos sobre hipersegmentação em textos do EF II, pois é uma característica típica haver segmentação da sílaba pretônica da palavra em razão da proeminência acentual da sílaba tônica de palavra prosódica, como apontado por Silva (2014, p. 161):

Na maioria das hipersegmentações, apreendemos informações fonológicas importantes ligadas a esses constituintes, como a presença do acento prosódico, que se constituiu em um lugar privilegiado para a inserção de um limite gráfico não-convencional na fronteira em que esteve presente essa proeminência, como em, por exemplo, “em **bora**” (embora).

Em síntese, a análise de “da eronave” e, mais especificamente, de “as davano” propicia explicitar a complexidade linguística desse tipo de dado que, embora raro quanto à sua frequência, demanda mobilizar conceitos sobre qualidade fonética de consoantes e vogais, estrutura silábica e proeminência acentual, além de relações entre fronteiras de palavra prosódica e ortográfica. Lembramos, no entanto, que defendemos a hipótese de que a ocorrência desse tipo de dado está fortemente ancorada nas relações de sentidos que podem ser apreendidas do texto em análise, como argumentaremos a partir da análise do texto enquanto uma resposta à proposta de produção textual.

Observamos que o sujeito inicia seu texto relatando sobre a data de um suposto sorteio e também que seriam os dois ganhadores para ir “nesse país”. Entretanto, embora utilize o pronome demonstrativo “nesse”, indicando algo que já foi referido e está sendo retomado, não há essa referência no texto, uma vez que o aluno sequer mencionou, até então, qual país seria esse. Ao ler a proposta, podemos inferir sobre o país que o aluno está tratando. Vejamos o enunciado da proposta na figura abaixo.

**Figura 10. Enunciado de proposta de produção textual**

- Imagine que você é um astronauta que foi mandado a algum planeta do Sistema Solar. Quando você chegou lá, encontrou alguns habitantes daquele planeta estranho.
- Da mesma maneira que o extraterrestre descreveu o planeta Terra, escreva uma narrativa, em que você seja o personagem principal, contando como que era o planeta e seus habitantes.
- Seu texto deve conter de 20 a 25 linhas e deve ser escrito à tinta. Seu texto não deve ultrapassar os limites designados para a escritura. Dê um título a seu texto.

Fonte: Tenani (2015)

A proposta solicita uma narrativa a partir da apresentação de um conto em que um extraterrestre descreve a Terra e seus habitantes de maneira pouco comum. Na proposta, o tema é “Viagem para outro planeta”. A partir dessa informação, podemos inferir que o tal país a que se refere o aluno, na verdade, seja um outro planeta indicado na proposta de produção textual. Essa interpretação é respaldada pelo trecho “para ir nesse país em uma nave espacial”, já que não se usa nave espacial para chegar a nenhum país na Terra. Embora a proposta sugira que o sujeito imagine ser um astronauta, ele ainda parece ocupar, na narrativa, o papel de aluno na Terra. Em seguida, o aluno passa a descrever o momento em que ele “desce *da eronave*”, pisa no tal país estranho e começa a relatar o que ele encontra de diferente por lá. Nesse momento, ele passa a responder ao que foi solicitado pela proposta. É interessante observar, ainda, que é nesse momento que ele dá início ao seu passeio ao “país estranho”; é o momento em que ele de fato conhece um outro planeta. É nesse momento da narrativa que nos deparamos com as grafias de híbrido “da eronave” e “as davano”. Essas grafias ocorrem num momento de tensão da narrativa, em que o aluno finalmente chega no tal planeta e se depara com características muito diferentes das quais ele conhece. Esse tipo de registro gráfico se configura como efeito dessa tensão da narrativa, quando o aluno dá resposta à proposta textual de conhecer outro planeta com seres e características tão diferentes.

Um outro aspecto importante a ser apontado sobre a escrita desse aluno é a ocorrência de outro híbrido, “come cheia” para “comecei a”, em outro texto, também num ponto de tensão da sua narrativa. Detectamos que a produção de híbridos em momentos importantes do texto parece uma característica típica desse sujeito. Portanto, tais grafias podem ser interpretadas como efeito de uma organização linguístico-textual do enunciado escrito.

A fim de ampliar a ancoragem dessa hipótese interpretativa, passamos, a seguir, à análise de outro texto do sujeito 27, em que ocorre o dado híbrido “no estados zunidos”.

Figura 11. Texto 3 do sujeito 27

- Oi o meu nome é Diego eu vou comta en meu sendo eu, queria i para a pequim eu que ria ver en jogar olimpico, de la comdo a caba en jogar eu vou imhora de pequim dela eu queria i para a paraguai compra uma baixe um pc e mais.

Dela eu queria ir para o japão que ria com ese, la eu queria ver os carros ferrari os corvette emais carros eu queria com pra uma coisa.

Dela eu queria ir para garuja na prai, eu queria mora la ~~na~~ na, dera no mar com uma casa comtudo carros campo pesinas e mais coisas eu ~~que~~ era um empresario de um, esquirtorio eu tinha mais uma casa no estados, guimidos e jiu.

Fonte: Banco de Dados de Escrita do Ensino Fundamental II (Z08\_5C\_09M\_03)

Neste texto, podemos observar algumas segmentações não convencionais produzidas pelo sujeito 27: as hipossegmentações “dela” (de lá”), “emais”, “comtudo”; as hipersegmentações “que ria” “a caba”, “com ese”, “com pra”; e o híbrido “estados zunidos”. Em relação as grafias não convencionais, podemos observar: “Comta”, “i”, “olímpico”, “comdo”, “imbora”, “i”, “baixe”, “Garuja”, “prai”, “dera no mar”, “pesinas”, “esquirtorio”.\*

Novamente, partimos do híbrido para a seleção desse texto para análise. O dado “estados zunidos” chama a atenção também devido à sua complexidade linguística, uma vez que as fronteira de hipossegmentação e de hipersegmentação não são tão claras nesse dado. Seria possível dizer que ocorre uma hipossegmentação entre “estados” e “zunidos”, já que o sândi externo que ocorre na fala “esta[dozu]nidos” é registrado na escrita, pois a letra <s> em coda final de “estados” é marcada como “z” na palavra “unidos” a que se junta resultando em “zunidos”, havendo, então, uma ressilabificação de <s>: a consoante em fim de sílaba “os” é grafada junto à vogal da palavra que lhe é adjacente “unidos”.

Também nos deparamos com as segmentações “que ria”, “a caba”, “dela” e outras grafias não convencionais: “comta, i, olinpico, comdo, baixe”. Interessa comentar características comuns às hipersegmentações “que ria” e “a caba”: ambas apresentam fronteira na sílaba tônica, a sílaba pretônica que fica isolada corresponde a um clítico fonológico os quais são palavras existentes na língua portuguesa (“que” e “a”). Já a

hipossegmentação “dela” (“de lá”) trata-se de um clítico (cl) mais hospedeiro (w), como também já analisado anteriormente.

Percebemos, ainda, que, neste texto, o aluno descreve o sonho de futuro da vida dele. Aqui, o aluno inicia seu texto descrevendo vários elementos dos quais ele gostaria de ter ou lugares em que gostaria de morar. No fim, ele termina o seu texto dizendo que teria uma casa nos “estados zunidos”. Assim, é possível supor que esse fato esteja relacionado ao imaginário social do aluno, uma vez que parece deixar claro o “sucesso” do escrevente em relação ao seu sonho, já que tal fato está ligado ao *status* social de conseguir adquirir uma casa em um país como os Estados Unidos. Logo, há um destaque para um momento importante do texto, em que o sujeito atinge o máximo do seu sonho, atendendo à proposta de produção textual.

### **3.5 Flutuação entre o convencional e o não convencional**

Nesta seção, o foco será a flutuação entre grafias convencionais e não convencionais de segmentação de palavra. Novamente, selecionamos a escrita do sujeito 27 para tratar dessa flutuação e apresentamos outro texto dele para a análise.







convencional e não convencional tendo como partida o híbrido “come” e “cheia”, registro da junção entre “comecei” e “a”, no enunciado “comecei a jogar lá”. Passaremos a analisar as segmentações que ocorrem no texto selecionado, visando construir nossa argumentação respaldada na visão expressa por Capristano (2021).

Inicialmente, chama bastante atenção o dado “come cheia jogala” (comecei a jogar lá) por se tratar de um híbrido: hipersegmentação de “comecei”, e hipossegmentação da sílaba tônica de “comecei” e a preposição “a”, e, ainda, outra hipossegmentação entre “jogar” e “lá”. A seguir, indicamos as fronteiras de palavra (w) e de pé métrico ( $\Sigma$ ) das palavras grafadas convencionalmente, em (10), e das palavras quando grafadas não convencionalmente, em (11).

10. Comecei a jogar lá  $\Rightarrow$  [comecei]w [a]w [jogar]w [lá]w

11. Come cheia jogala  $\Rightarrow$  [come]  $\Sigma$  [cheia]  $\Sigma$  [jogala]  $\Sigma$

Partindo da hipersegmentação de “comecei”, observamos que o branco ocorre antes da sílaba tônica “cei” e, por isso, as sílabas pretônicas “come” ficam delimitadas por branco. Soma-se o fato de que essa sequência de sílabas corresponde à forma flexionada do verbo “comer”: a forma de terceira pessoa do singular no presente do indicativo é “come”. Além dessa hipersegmentação, há a troca de “ce” (“comeCEi”) por “che” (“come CHEia”). É interessante perceber que a troca pode ter sido baseada em critérios fonéticos, uma vez que as consoantes [ʃ], representada pelo dígrafo “ch”, e [s], representada pela letra “c”, partilham traços em comum, pois ambas são fricativas desvozeadas, mas diferem entre si pelo ponto de articulação: [s] alveolar e [ʃ] pós alveolar.

É possível que a sequência de palavras “comecei a” tenha sido grafada a partir da junção de “a”, uma preposição átona, junto ao hospedeiro “comecei”, palavra que tem acento em “cei”. Tal processo de sândi vocálico gera, mais especificamente, uma ditongação, que consiste em um processo “de formação de ditongos com a vogal final de um vocábulo e a inicial de outro, desde que uma das vogais da sequência seja alta e átona” (BISOL, 1996, p. 160). Desse modo, na fala, é possível prever a junção entre a sílaba final de “comecei” com a preposição “a” a qual pode ter motivado a junção na escrita e teríamos “comeceia”, e, com a hipersegmentação, geraria “come ceia” que também correspondem a palavras do português. Mas constatamos que a inclusão de “h” na sequência “ceia” cria uma outra palavra “cheia”. Portanto, o dado chama a atenção não só por haver a hipersegmentação seguida da

hipossegmentação, mas por haver, ao mesmo tempo, o acréscimo de letra de modo a resultar na palavra “cheia”, cuja cadeia fônica não corresponde exatamente à sequência fônica representada em “comecei a”. A complexidade das alterações ortográficas fica, pois, evidenciada nessa descrição. Portanto, há uma complexidade extra quanto aos processos de representação de palavras que estão envolvidos na sequência de palavras.

Já no dado “jogala”, trata-se do registro das palavras “jogar lá” e ocorre uma junção não prevista entre o verbo “jogar” e o advérbio “lá”, pois é uma hipossegmentação entre palavras acentuadas enquanto a tendência é haver hipossegmentação entre uma palavra não acentuada e outra acentuada. Há o apagamento do “r” final do verbo “jogar”, resultando em uma grafia que corresponderia a uma estrutura verbal enclítica: “joga-la”, ou seja, a grafia seria o verbo “jogar” e o pronome “la”. Nessa possível grafia, a ausência de hífen entre verbo e pronome, como descrito por Fiel (2018), configura uma junção de verbo mais um pronome em posição enclítica e, segundo a autora, “[...] parece menos motivada na configuração prosódica preferencial do [português brasileiro] PB, pois a direção da prosodização de pronomes em relação ao verbo no PB é proclítica.” (acréscimos nossos, p. 86). Entretanto, a autora afirma que a colocação enclítica é ensinada e valorizada no ambiente escolar e é possível que seja essa a motivação por trás dessas grafias – principalmente no sexto ano, uma vez que seria o ano escolar em que esse conteúdo programático relativo à colocação enclítica do verbo é ensinado. Desse modo, também haveria relação com o eixo do código escrito institucionalizado (CORRÊA, 1997) Fiel (2018) também ressalta que a maioria das hipossegmentações de verbo mais pronome é verbos com duas ou três sílabas, como em “pagala” e “estudalos”, o que resulta em palavras paroxítonas com pés ternários. Assim, tal junção resulta em uma configuração métrica de troqueu silábico, que, segundo Bisol (1996), é a configuração métrica mais geral do português em relação à distribuição de acento.

Por meio da análise do dado “come cheia jogala”, descrevemos possíveis motivações linguísticas para sua ocorrência. Passamos a tratar da flutuação entre essa grafia não convencional e grafias da mesma sequência de palavras, ou uma sequência similar, que seguem as convenções ortográficas. Observe os destaques no trecho transcrito abaixo:

### Trecho 1 selecionado do texto 4, sujeito 27

*Ai o meu jolega começou a jogar ne uma quandra tertó da minha casa quartigan come cheia jogala<sup>1</sup> ai eu pare de jogar la com 12 anos o meu amigo que jogava la<sup>2</sup> ele saio dela primeiro do que eu sai na rua ele (rasura) tava na frente da casa dele eu pergunte i onde ele jogava ele falo no jo(V)ec era ne uma sexta-feira ele mechamou para ir la jogar no joVec e(u) fui e rostei de la e comecei a jogar la<sup>3</sup> jano primeiro dia eu fis um gol eu gosti de la fiquei jogando (rasura)\* 10 anos jogando.*

*Eundia foi olieiro do Santos F.C eles mechanaran para fazer um testi eu pasei eu comesei a jogar<sup>4</sup> sub 16 (rasura\*) na sexta-feira tive ir para Satos porque eu ia morar em Santos mi(n)ha mae foi com migo para a lugar um apartamento, ai na quarta-feira eu tinha ja jogar (rasura)\* era o primeiro jogo que eu ia jogar com sub 16 os profesores falava que eu jogava Bei fazia muitos gol um dia o professor doryal junior o (raura)\* tecnico do Santos F.C (e)le e do time profisional.*

Nos trechos sublinhados, há a flutuação entre as grafias “come cheia jogala” (sinalizada por 1) e “comecei a jogar la” (sinalizada por 3). Também há flutuação entre o registro do clítico junto ou separado da forma verbal hospedeira, como nas grafias “jogala” (sinalizada em 1) e “jogava la” (sinalizada 2). Outra flutuação a observar é entre “comesei a jogar” (sinalizada por 4) e “come cheia jogala” (sinalizada em 1). Essas flutuações podem revelar que não se trata apenas de um caso de uma sequência “problemática” para o aluno, ou seja, não seria suficiente e nem adequado entender que as grafias não convencionais foram motivadas, unicamente, em razão de haver dificuldade com a escrita da sequência “comecei a jogar lá”, já que ora é ora não é grafada de acordo com a convenção.

De modo geral, a flutuação entre grafia convencional e não convencional se mostra bastante instigante, uma vez que nos levaria ao seguinte questionamento: por que em determinados momentos a grafia segue a convenção ortográfica e, em outros momentos, não? O que levaria o sujeito a produzir um dado tão diferente e complexo em certos trechos do texto, mas convencional em outros? Outra questão que nos é feita: o aluno, afinal, está submetido, ou não, às convenções ortográficas em relação a certas estruturas? Para responder aos questionamentos, tecemos uma análise geral do texto, buscando indícios em outras grafias não convencionais, bem como na organização do texto e no enunciado da proposta de

produção textual.

Inicialmente, o aluno começa seu texto com o seguinte parágrafo destacado a seguir.

### **Trecho 2 selecionado do texto 4, sujeito 27**

*Eu senpre soiei que, eu queria ser jogador de futeBol eu quando eu tinha 6 anos eu não ligava para futeBol os meus amigos mechamava para jogar Bola na rua eu falava não.*

No trecho acima, o aluno inicia seu texto a fim de desenvolver a proposta “Tornar-se famoso” e conta sobre o seu sonho de ser jogador de futebol. Inicialmente, em relação às grafias não convencionais, nos deparamos com a palavra “senpre”, grafado com “n” ao invés de “m”. A regra ortográfica de emprego do “m” antes de “p” e “b” é ensinada nos primeiros anos do processo de alfabetização e, segundo os documentos oficiais de educação, essa grafia não é esperada em um texto produzido na 5ª série/6º ano do EF. Em seguida, nos deparamos com a grafia de “soiei” (sonhei), que evidencia que o aluno não grafou a nasal palatal em posição de ataque, tendo trocado o dígrafo “nh” pelo grafema “i”. Essa troca tem relação com o fato de que a nasal palatal poder ser realizada como uma vogal alta anterior, pois, conforme Cagliari (1981, p. 85), “alguns falantes usam [i] ou [j] onde outros usam uma nasal palatal em início de sílaba dentro de palavras.”. Por meio dessa descrição, observamos que se trata de grafias mais próximas da oralidade do que da escrita, segundo o trabalho de Cagliari (1981). Essas grafias podem ser interpretadas como uma marca da heterogeneidade da escrita, pois o sujeito busca representar características fonético-fonológicas detectadas em sua variedade linguística falada em enunciados escritos, o que constitui indícios do modo heterogêneo de constituição da escrita, segundo Corrêa (2004). Ainda no trecho em análise, há a ocorrência da hipossegmentação “mechamava”, na qual há a falta de espaço em branco entre “me” e “chamava”. Acrescenta-se que há, ainda, “um entrecruzamento de critérios de natureza prosódica e critérios de natureza gráfica na composição dessas estruturas” (CHACON, 2006, p. 8), já que o sujeito também respeita as fronteiras anterior a “me” e posterior a “chamava”, coincidindo com certos limites estabelecidos pela convenção ortográfica dessas palavras. Há também a falta de concordância entre o verbo e o sujeito: é previsto “chamavam” no trecho escrito “meus amigos mechamava”. Essa ausência de concordância verbal, bem como a segmentação não convencional, parece decorrer de uma pressuposição do aluno de que algumas características da fala pudessem ser transferidas diretamente para a escrita, não só

por uma questão de “reprodução” de características que o sujeito observa em sua variedade linguística, mas também uma representação das fronteiras observadas no fluxo da linguagem oral. Novamente, detectamos nessas grafias a heterogeneidade da escrita apontada por Corrêa (2004), uma vez que mostra o trânsito do sujeito entre práticas orais/faladas e letradas/escritas. Esse trânsito do sujeito é bastante relevante para a hipótese interpretativa da flutuação entre grafias convencionais e não convencionais em análise.

Avançamos na análise, recuperando o trecho a seguir apresentado.

### **Trecho 3 selecionado do texto 4, sujeito 27**

*Ai o meu jolega começou a jogar ne uma quandra tertó da minha casa l quartigan come cheia jogala ai eu pare de jogar la com 12 anos o meu amigo que jogava la ele saio dela primeiro do que eu sai na rua ele (rasura) tava na frente da casa dele eu pergunte í onde ele jogava ele falo no jo(V)ec era ne uma sexta-feira ele mechamou para ir la jogar no joVec e(u) fui e rostei de la e comecei a jogar la jano primeiro dia eu fis um gol eu gostí de la fiquei jogando (rasura)\* 10 anos jogando.*

No trecho, deparamo-nos com o dado “come cheia a jogala”. Além das características descritas anteriormente, notamos que essa grafia tem ancoragem no ritmo da fala, uma vez que é segmentada em pés métricos troqueus (sílabas fortes seguidas por uma átona), a saber: [come]Σ [cheia]Σ [jogala]Σ. Cabe observar que, neste momento, o aluno inicia o relato do sonho de ser jogador de futebol. Nesse trecho, identificamos o cerne da resposta à proposta textual. Vejamos: até então, o sujeito apenas relatou que seus amigos o chamavam para jogar bola e ele recusava o convite, porém, no trecho (5), há uma mudança nesse comportamento, uma vez que o sujeito passa a começar a jogar futebol. A partir desse momento, o seu processo de se tornar famoso dá o primeiro passo. Assim, temos indícios de que os registros não convencionais das fronteiras de palavras, nesse ponto do seu texto, resultam de uma tentativa de destaque para uma parte da narrativa considerada importante e em um recurso que contribui para a construção de sentido, da resposta à proposta de o aluno supor como seria ser famoso (seria o início do marco importante em sua vida da passagem em que ele deixa de recusar os convites para jogar e começa a jogar “lá” e ser famoso). Desse modo, a sequência “come cheia jogala” deixa de ser uma mera grafia de segmentação não convencional e passa a representar uma maneira de destacar o marco inicial da trajetória do sujeito em processo de se tornar famoso.

Detectamos, pois, um momento de conflito tanto no texto quanto no escrevente. Assim, tal grafia pode ser interpretada como efeito de uma organização rítmica do enunciado, mais silabado, para chamar a atenção do leitor para a parte central do texto: o sonho de ser jogador de futebol.

Essa hipótese explicaria o motivo de haver essa grafia bastante diferente neste ponto do texto em contraposição ao que se observa na mesma sequência “comecei a jogar lá”. Dessa forma, argumentamos que as motivações para grafias não convencionais estão não só no domínio da palavra e de sua representação gráfica, mas também nas práticas sociais orais/letradas dos escreventes. Essa linha de interpretação encontra respaldo em Chacon (2021, p 14-15) quando o autor afirma que

[...] as flutuações ortográficas indiciam a ancoragem da escrita das crianças tanto em características que se detectam em enunciados falados quanto em características que se detectam em enunciados escritos. Trata-se, portanto, de uma escrita constitutivamente heterogênea, já que sustentada, simultaneamente, por características que provêm desses dois tipos de práticas de circulação dos atos de fala e de escrita – plasmadas na escrita, assim como ocorre com a fala.

Passamos à análise de outro trecho selecionado do texto 4.

#### **Trecho 4 selecionado do texto 4, sujeito 27**

*Eu comesei A jogar no Santos F. C profisional (rasura)\* eu jo(v)o a olado de Neymar, RoBinho, e PH Ganso.*

*Pela primeira Ve(s)e nor ira jogar comtra o corintians nor jogamos com ele nor gaiou (rasura)\* o Santos ganhou de 13 a 0 foi a maior golia do Santos no Brasileiro de 2010 eu fis os 6 gol do Santos era que eu sempre queria jogar comtra o corintias os corintianos ficagal triste de pois desa goleada os corintiano fora(m) so zuado fim.*

Nesse momento do texto, identificamos as grafias não convencionais “comesei” (comecei), “profisional” (profissional), “comtra” (contra), “gaiou” (ganhou), “fis” (fiz), “ficagal” (ficaram) e “desa” (dessa), além das segmentações não convencionais “a olado” (ao

lado) e “de pois” (depois). Interessa-nos destacar a grafia híbrida “a olado”, em que ocorre uma hipersegmentação entre “ao” e uma hipossegmentação entre “ao” e “lado”. É possível que o aluno tenha grafado a letra “a” separada do “o” por reconhecê-la como uma palavra escrita possível de ser grafada isoladamente, processo semelhante como em “a lugar”. Essa grafia, assim como observamos em “come cheia jogala”, se trata de um novo momento importante do texto, uma vez que o aluno está contando sobre sua excitação de estar jogando ao lado de jogadores importantes como Neymar, Ganso e Robinho. Essas informações têm relação com o imaginário social do escrevente, uma vez que lida com relações de status social, por exemplo, do jogador de futebol. Desse modo, ao trazer um novo destaque ao texto, gera-se um novo momento de conflito, uma vez que “põe em destaque a presença do escrevente na partida de futebol como um jogador tão importante como aqueles que destaca: estar “ao lado” dos seus ídolos na mesma condição deles, e não apenas como um admirador” (SILVA; TENANI, 2014, p. 150). Assim, há um destaque, ainda, para o exato momento em que o sujeito atinge a fama, atendendo à proposta de produção textual.

Valemo-nos, ainda, da interpretação de Capristano (2021) sobre dados dessa natureza identificados em textos do EF I para assumir sua interpretação para dados do EF II. Para a autora:

[...] a análise qualitativa das mesclas/híbridos/mistos, apenas resumidamente apresentada aqui, permite defender que essas segmentações raras e especiais são sinais de que as crianças segmentam seus enunciados escritos não exatamente sob o efeito da busca por recortar palavras, mas, sim, sob o efeito da busca por recortar diferentes dimensões de suas experiências sócio-históricas, ou, ainda, da busca por encontrar sentidos para a (sua) escrita. Seriam, portanto, assim como todas as segmentações convencionais e não convencionais vistas na escrita de crianças, manifestações de como a língua e a escrita se mostram, se colocam ou se impõem ao sujeito e, ao mesmo tempo, marcas do sujeito na busca de alçar-se e/ou encontrar-se com sentidos da/para a (sua) escrita. (CAPRISTANO, 2021, p. 802).

Com Capristano (2021), buscamos entender o funcionamento do que a autora nomeia como mesclas/híbridos/mistos e nossa análise aponta para a complexidade linguística dos dados e para pesquisas futuras que analisem os dados como pistas da complexa relação do sujeito da linguagem com a (sua) escrita. Assim como para dados do EF I, interpretamos que as segmentações não convencionais do EF II ora analisadas podem ser interpretadas como

“[...] pistas das movências de sentidos e da própria língua atuantes na enunciação escrita da criança, movências que [...] são permitidas pelos sempre complexos e inéditos encontros entre o eu e outro, constitutivos desse modo de enunciação.” (CAPRISTANO, 2021, p. 801)

Com base na análise apresentada dos dois textos, compreendemos com Tenani (2008) que a segmentação de palavras ora convencional ora não convencional ganha uma dimensão textual na produção desse escrevente, já que a maneira de colocar as fronteiras de palavras não ocorre de forma aleatória quando consideramos a construção de todo o texto. Embora os dados de híbridos tenham poucas ocorrências (apenas seis) e na produção de um aluno mais em específico, esses dados se diferenciam em relação aos tipos de híbrido encontrados no EF I, conforme é possível observar em Cunha (2004). Essa autora, apesar de também não ter identificado muitos dados desse tipo, percebeu, no entanto, que, quando eles ocorriam, as crianças primeiro hipossegmentavam duas palavras para depois hipersegmentar parte da segunda palavra, como em: “pofa vor” (para “por favor”) (CUNHA, 2004, p. 112). Como vimos nesta análise, nos dados “as davano”, “come cheia” e “a olado”, o escrevente primeiramente hipersegmenta para depois hipossegmentar.



## CONSIDERAÇÕES FINAIS

Nesta dissertação, buscamos compreender o percurso das segmentações não convencionais encontradas nos textos de alunos considerados *outliers* quanto ao número de hipossegmentações no EF II. Em função dessa classificação inicial em relação a esses sujeitos, conforme descrito por Fiel (2018), e por sua escrita não ter sido anteriormente analisada, realizamos, por um lado, uma descrição mais geral das características linguísticas das suas produções e, por outro, uma análise de ocorrências que se distanciavam, ligeiramente, das demais regularidades aqui encontradas.

Inicialmente, trouxemos uma discussão sobre noção de palavra, uma vez que se trata de uma unidade linguística constituída de características fonológicas, morfológicas, sintáticas e semânticas, caracterizando a sua natureza complexa. Também resenhamos alguns trabalhos que tratam dos diferentes tipos de segmentação não convencional (hipo, hiper e híbridos). Tratamos brevemente das diferentes perspectivas e suas respectivas interpretações atribuídas aos casos de segmentação não convencional, como a perspectiva escolar, que enxerga os casos de hipo e hipersegmentação como erros ortográficos decorrentes de indesejada interferência do oral no escrito, resquícios de problemas de alfabetização, e/ou decorrentes de distúrbio específico de escrita; e a perspectiva linguística, que interpreta esses dados como efeito do trânsito dos sujeitos escreventes por práticas sociais orais/faladas e letradas/escritas. Para a nossa pesquisa, assumimos a perspectiva linguística, a qual interpreta os casos de segmentação não convencional de palavras como pistas do trabalho do sujeito com a língua(gem). Em nossa análise, descrevemos as características prosódicas dos dados dos sujeitos *outliers* por meio da análise prosódica dos dados, baseada no arcabouço teórico da Fonologia Prosódica de Nespor e Vogel (1986), que organiza os enunciados falados em constituintes fonológicos arranjados hierarquicamente. Com base nessa abordagem teórica, explicamos de que maneira a organização rítmico-entoacional dos enunciados constitui os textos escritos e motiva os casos apresentados.

Em relação aos resultados, observamos que a junção entre clítico e palavra prosódica foi a estrutura que mais gerou dados de hipossegmentação nos textos dos sujeitos *outliers*. Nesse conjunto de dados de hipossegmentação, observamos, em todos os anos do EF II, um possível não reconhecimento do estatuto autônomo dos clíticos, que seriam interpretados,

pelos escreventes, como sílabas que integram as palavras que os acompanham. Além disso, essas características vão ao encontro das típicas do EF II. Buscamos compreender, também, os critérios linguísticos subjacentes às grafias de hipersegmentação. Por meio da organização dos dados a partir da estrutura prosódica das palavras convencionais, identificamos pistas de informações prosódicas e (orto)gráficas. Sobre o aspecto prosódico, as fronteiras gráficas não convencionais puderam ser relacionadas a fronteiras de constituintes prosódicos, com destaque, para os constituintes sílaba e pé métrico. Na maioria das hipersegmentações, apreendemos informações fonológicas importantes ligadas a esses constituintes, como a presença do acento prosódico, que se constituiu em um lugar privilegiado para a inserção de um limite gráfico não convencional na fronteira em que esteve presente essa proeminência, como constatado na seção 3.4.

Outras grafias, como as derivadas da presença não convencional do hífen, deram pistas de informações letradas/escritas, na medida em que se basearam em formas verbais (por exemplo, *volteilá* [*voltei lá*]) como se fossem sequência verbo-clítico (por exemplo, *“ajude-me”*). Foi possível observarmos que, em todas as situações descritas, foram informações relativas à própria escrita que estiveram subjacentes às hipóteses acerca dos limites de palavras, fundamentadas no que o escrevente “supõe ser – a partir não só do que aprendeu na escola, mas, em grande parte, do que assimilou fora dela – a visão escolarizada do código institucionalmente reconhecido” (CORRÊA, 1997, p. 271), já que demonstram estarem envolvidas pelo “caráter reprodutor de uma prática instituída, [pela] tentativa de alçamento à escrita culta formal e [pelos] discursos estabilizados da instituição escolar” (CORRÊA, 2004, p. 167).

Argumentamos que os dados apresentados parecem referir-se à representação que o sujeito escrevente faz do escrito institucionalizado para a sua escrita, pois consideramos a junção do clítico com o verbo em posição enclítica como tentativa do escrevente de alçar a (sua) escrita ao que ele imagina ser a escrita convencional. Essa tentativa do escrevente é estabelecida por meio dos entrelaçamentos do oral/falado e do letrado/escrito (CORRÊA, 2004).

A respeito dos híbridos, observamos que essas grafias, principalmente as produzidas pelo sujeito 27, podem ser compreendidas como hipóteses particulares sobre a noção de palavra escrita “reveladoras [por sua vez] das singularidades dos sujeitos e da relação por eles estabelecida com a linguagem” (ABAURRE, 1997, p. 24).

Vimos, ainda, por meio do processo interpretativo desses dados, semelhanças com

processos de segmentação em palavras observados em enunciados escritos produzidos no início do processo formal de aquisição da escrita e, ao mesmo tempo, semelhanças com momentos relacionados aos aspectos mais gerais observados em todo o *cópus*. Observamos, por meio dos casos particulares, que o modo não convencional de segmentação das palavras teve como efeito demarcar destaques de noções ou ações centrais em que se desenvolvem narrativas sobre as quais se construíram o dado texto.

Quanto à trajetória longitudinal de cada aluno, identificamos que grande parte deles iniciou e terminou o EF II produzindo segmentações não convencionais. No entanto, as ocorrências se tornaram menos frequentes no decorrer dos anos, quando cotejada à quantidade de dados em relação à extensão dos textos produzidos (dada pelo número de palavras escritas) e às dificuldades em grafar tipos específicos de palavra (como “demais”, por exemplo) ou tipos de palavras que apresentaram sílabas correspondentes (como amigo, apareceu, apaixonado, por exemplo).

As segmentações não convencionais, ao observarmos as suas distribuições, caracterizaram-se por apresentar uma diminuição ao longo dos anos. Notamos, ainda, que os escreventes tiveram dúvida quanto aos limites das mesmas palavras, ou seja, muitos dos dados encontrados aqui eram do mesmo tipo (por exemplo, muitas ocorrências de “oque”). Assim, não houve somente o decréscimo do número de dados, mas, também, os tipos de palavras que geram dúvidas quanto aos seus limites.

Com base nos resultados alcançados, compreendemos que a utilização de recursos gráficos, com o intuito de dividir um enunciado escrito em palavras, continua, em alguma medida, sendo um desafio para os escreventes que já completaram o período da alfabetização. Entendemos, ainda, que reconhecer os limites de palavra implica, ao escrevente da língua, lidar com a complexidade da própria definição do conceito de palavra, a qual não se restringe à forma de registro escrito e perpassa várias esferas da linguagem. Desse modo, para “acertar” onde começa e termina uma palavra escrita, o escrevente deve estar atento a diferentes critérios linguísticos a fim de alçar às exigências da ortografia, valorizadas, sobretudo, socialmente. Chegamos, neste momento final, vislumbrando a relevância de nossos resultados para a compreensão dos processos de segmentação não convencional de palavras que ainda ocorrem no período do EF II, na medida em que, por meio deles, podem ser traçados paralelos que permitem observar semelhanças e diferenças quanto ao modo de delimitação nas palavras em etapas distintas do processo escolar.

Dessa forma, as grafias não convencionais se mostram como indícios do modo como o

aluno lida com a escrita, com a convenção ortográfica, com a noção de palavra e com a linguagem. Notamos, ainda, que algumas ocorrências mais particulares também podem ser pistas de um imaginário social desse sujeito, atrelado às relações de status sociais da sociedade. Essas grafias não convencionais, como evidenciado no decorrer deste texto, também configuram fortes indícios da heterogeneidade constitutiva da escrita, que “constitui marca inequívoca da relação sujeito/linguagem, indício da circulação do escrevente por práticas orais/letradas e reatualização, em termos de funcionamento,” (CAPRISTANO, 2010, p. 177). Dessa forma, analisamos os processos mobilizados pelo sujeito em direção à escrita/palavra ortográfica. E, por meio desses processos, percebemos, ainda, os entrelaçamentos entre o oral/falado e o letrado/escrito, como bem elabora Capristano (2010, p. 175):

[...] ao assumir a escrita como processo – e não como produto –, assume-se, também, que o caráter heterogêneo da escrita não está restrito ao seu material significante, mas propaga-se para outras dimensões que constituem o processo de sua produção. Nesse sentido, na avaliação, no exame e/ou na análise de fatos linguísticos escritos, a atenção volta-se prioritariamente para aquilo que, no produto escrito, aponta para seu processo de constituição.

Defendemos, por tudo o que foi discutido, a necessidade de não só considerar a quantidade de palavras escritas por texto e a distribuição de ocorrências por escrevente como informações importantes, mas também considerar a natureza linguística dos dados, que deixam à mostra as imagens que os escreventes têm sobre a constituição da escrita, a qual, junto a Corrêa (2004), defendemos, neste trabalho, ser heterogênea. Tal abordagem se mostra importante, uma vez que a afirmação de alto número de ocorrência de junção entre palavras não é critério suficiente para validar diagnóstico de “distúrbio específico de escrita”, uma prática bastante comum em ambiente escolar ou clínico.

## REFERÊNCIAS

- ABAURRE, M. B. M. Os estudos linguísticos e a aquisição da escrita. In: CASTRO, M. F. P. (Org.) *O método e o dado no estudo da linguagem*. Campinas: Editora da Unicamp, p. 111-78, 1996.
- BARRETO, D. A. R. J.; MASSINI-CAGLIARI, G. O apagamento das consoantes róticas finais: Um estudo comparativo entre o Português Arcaico e o Português Brasileiro. *Revista do GEL*, v. 16, p. 37-52, 2019.
- BERTI, L; C, CHACON, L; SANTOS, A. P. A escrita de /aN/ por pré-escolares: pistas acústico-auditivas. *Cadernos de Educação – Pelotas (UFPel)*. 2010 JanAbr; 19 (35): 195-219.
- BISOL, L. Constituintes prosódicos. In: \_\_\_\_\_. *Introdução a estudos de fonologia do português brasileiro*. Porto Alegre: EDIPUCRS, p. 247-261, 1996.
- BISOL, L. A degeminação e a elisão no VARSUL. In: BISOL, L.; BRESCANCINI, C. (orgs.) *Fonologia e variação: recortes do português brasileiro*. Porto Alegre: EDIPUCRS, 2002. p.231-250.
- BRASIL. Ministério da Educação. Secretaria de Educação Fundamental. *Parâmetros curriculares nacionais: língua portuguesa*. Brasília, 1998.
- BOOP DA SILVA, Taís. *Formação de palavras compostas em português brasileiro: uma análise de interfaces*. 195f. Porto Alegre: Tese (Doutorado em Letras) Universidade Federal do Rio Grande do Sul, 2010.
- BORSATO, A. *Grafias não convencionais no EF II: Estudo de sujeitos outliers*. Relatório de Iniciação Científica, Universidade Estadual Paulista “Júlio de Mesquita Filho”, campus de São José do Rio Preto, 2019.
- CAGLIARI, L. C. *Elementos de fonética do português brasileiro*. 185 cf. Tese (Livre Docência) – Instituto de Estudos da Linguagem, Universidade de Campinas, Campinas, 1981.
- CALLOU, D. 1987. *Variação e distribuição da vibrante na fala urbana culta do Rio de Janeiro*. Tese de Doutorado, Universidade Federal do Rio de Janeiro.
- CALLOU, D.; Leite, Yonne & Moraes, João. 1994. Variação e diferenciação dialetal: a pronúncia do /r/ no português do Brasil. In Koch, Ingedore Grunfeld Villaça. Gramática do Português Falado, 465-493. Campinas: Editora da Unicamp/FAPESP.
- CALLOU, D.; SERRA, C. Variação do rótico e estrutura prosódica. *Revista do GELNE* 14: 41-58. 2012.
- CÂMARA JR., J. M. *A Estrutura da língua portuguesa*. Petrópolis: Vozes, 2004.

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: <https://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/13918>.

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. Por uma concepção heterogênea da escrita que se produz e que se ensina na escola. *Cadernos de Educação*. Ano 19 n. 35, Ed. UFPel - Pelotas, RS - Janeiro-Abril, 2010.

CAPRISTANO, C. C. Nos (não) limites da palavra: movências do modo escrito de enunciação. *Revista Brasileira de Linguística Aplicada*, v. 21, p. 789-812, 2021.

CARDOSO, Denise Porto. *Fonologia da Língua Portuguesa*. São Cristóvão: Universidade Federal do Sergipe: Cesad, 2009.

CARMO, M. C.; TABORDA, I. R. Apagamento de /R/ em coda silábica na variedade do interior paulista. *Letras Escreve*, Macapá, v. 9, n. 3, p. 39-51, jul./dez. 2019.

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](http://www.marilia.unesp.br/Home/Pesquisa/Gpel/Artigos/2005_-_Chacon_-_Hipersegmentacoes.pdf)>

CHACON, L. Algumas palavras sobre a relação entre oralidade e letramento em hipersegmentações na aquisição da escrita. In: CORRÊA, M. L. G. (org.). *Práticas escritas na escola: letramento e representação*. São Paulo: Convênio CAPES/COFECUB, p. 57-61, 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.

CHACON, L. A relação fala/escrita em dados não convencionais de escrita infantil. *Cadernos de Linguística*, v. 2, n. 1, p. 01- 17, 2021.

COELHO, I. L.; GÖRSKI, E. M.; NUNES de SOUZA, C. M. N e MAY, G. H. Para conhecer sociolinguística. São Paulo: *Contexto*, 2015.

COELHO, B. C. (2016), *Transposições ortográficas e estrutura da sílaba na escrita infantil*. Dissertação de mestrado, Universidade Estadual Paulista, São José do Rio Preto.

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

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

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

CUNHA, A. P. M.; MIRANDA, A. R. M. A influência da hierarquia prosódica em hipossegmentações da escrita de crianças de séries iniciais. *Revista virtual de estudos da linguagem – ReVEL*. Edição especial, n. 1, 2007, p. 1-19.

CUNHA, A. P. N.; MIRANDA, A. R. M. A hipo e a hipersegmentação nos dados de aquisição de escrita: a influência da prosódia. *ALFA: Revista de Linguística*, São Paulo, v. 53, n. 1, 2009. Disponível em: <https://periodicos.fclar.unesp.br/alfa/article/view/1681>

DONADEL, G. Breve reflexão acerca da escrita: a natureza das segmentações não convencionais. *Cadernos do Instituto de Letras*. Porto Alegre, nº 42, junho de 2011. p. 201-219.

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

FIEL, R. P. *Estudo longitudinal de hipossegmentações em textos do Ensino Fundamental II*. 2018. 123 f. Dissertação (Mestrado em Estudos Linguísticos) - IBILCE, UNESP, S.J.R.Preto, 2018.

GAYER, J. E. L. Uma breve história dos constituintes prosódicos. *Diadorim*. Rio de Janeiro, v. 2, n. 17, p. 149-172, dez., 2015.

MARCATO, F. *Análise prosódica de preposições monossilábicas*. 2013. 158 f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual Paulista Júlio de Mesquita Filho, Instituto de Biociências, Letras e Ciências Exatas, 2013.

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

NESPOR, M.; VOGEL, I. *Prosodic Phonology*. Dordrecht: Foris Publications, 1986.

PARANHOS, F. C. *Segmentações não-convencionais de palavra nos quatro últimos anos do ensino fundamental: um estudo longitudinal*. 2014. 170 f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual Paulista. São José do Rio Preto, 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.

PEREIRA, E. C. *Grammatica expositiva*. 7. ed. São Paulo, Rio de Janeiro: Weisflag Irmãos, 1918.

PIACENTE, T. QUEREJETA, M. La separación de palabra en la escritura infantil. *Revista Neuropsicología Latinoamericana*, [s. l.], v. 4, n 1, p. 1-17, 2012.

OLIVEIRA, E. C. *Um outro olhar para os erros de segmentação*. Tese (Doutorado em Linguística – Instituto de Estudos da Linguagem. Campinas: Universidade Estadual de Campinas, 2007.

REIS, M. C.; TENANI, L. E. Registros da heterogeneidade da escrita. São Paulo: *Cultura Acadêmica*, 2011.

RÚBIO, C. F. 2012. *Padrões de concordância verbal e de alternância pronominal no português brasileiro e europeu: estudo sociolinguístico comparativo*, Tese de Doutorado, Universidade Estadual Paulista, São José do Rio Preto.

SELKIRK, E. On derived domains in sentence phonology. *Phonology Yearbook*, nº 3, 1986.

SILVA, L. M. *Um estudo longitudinal sobre as hipersegmentações de palavras escritas nos anos finais do ensino fundamental*. 2014. 171 f. Dissertação (Mestrado em Estudos Linguísticos) - IBILCE, UNESP. S.J.R.Preto, 2014.

SILVA, L. M.; TENANI, L. *Hipersegmentações de palavra no ensino fundamental*. 1. ed. São Paulo: *Cultura Acadêmica*, 2014. (Coleção PROPG Digital- UNESP). ISBN 9788579835834. Disponível em: <http://hdl.handle.net/11449/126224>.

SIMIONI, T. O clítico e seu lugar na estrutura prosódica em português brasileiro. *Alfa: revista de linguística*, São Paulo, n. 52(2), p. 431-446, 2008. Disponível em: <https://periodicos.fclar.unesp.br/alfa/article/view/1526>.

TENANI, L. *Banco de dados de escrita do Ensino Fundamental II*. FAPESP/UNESP, 2005. Disponível em: <http://www.convenios.grupogbd.com/redacoes/Login>.

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

TENANI, L. *Prosódia e escrita: uma análise a partir de (hiper)segmentações de palavra*. Tese (Livre-Docência) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2016.

TORQUETTE, A. O. *Segmentações não-convencionais de palavras em escolas públicas e privadas*. 2016. 139 f. Tese. (Doutorado em Estudos Linguísticos) Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2016.

VASSOLER, A.; CAMACHO, R. Crenças e atitudes linguísticas: a variante retroflexa na variedade rio-pretense. *Revista do GEL*, 13.163-191, 2016. Disponível em: <https://revistas.gel.org.br/rg/article/view/1426>.

ZORZI, J. L. Alterações ortográficas nos transtornos de aprendizagem. In: MALUF, M. I. (Org.). *Aprendizagem: tramas hipersegmentação de palavras do conhecimento, do saber e da subjetividade*. Rio de Janeiro: *Vozes*; São Paulo, 2006. p. 144-162.