

# Flutuação de grafia de palavras: aspectos morfossintáticos

---

DOI: <http://dx.doi.org/10.21165/el.v49i1.2701>

**Roberta Pereira Fiel<sup>1</sup>**

## Resumo

Neste artigo, foram analisadas situações em que a sequência “o que” apresentava-se grafada de modo convencional (“o que”) e de modo não convencional (“oque”). Essas flutuações de grafia foram extraídas de textos de alunos que cursaram os quatro últimos anos do Ensino Fundamental (EF) em uma escola pública do interior paulista. Quanto aos resultados, verificou-se que: (i) o conjunto de sujeitos selecionados terminam o EF II com índice maior de “o que” grafado convencionalmente, todavia o índice de “oque” grafado não convencionalmente também aumenta; (ii) na flutuação gráfica “o que” e “oque”, são mobilizados dois *type* baseados na informação morfossintática, um *type* com funcionamento pronominal e um *type* com funcionamento interrogativo; e (iii) essa flutuação pode estar relacionada a aspectos prosódicos advindos da relação entre fala e escrita.

**Palavras-chave:** segmentação de palavras escritas; fala e escrita; flutuação de grafia de palavras.

---

<sup>1</sup> Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), São José do Rio Preto, São Paulo, Brasil; [roberta.fiel@unesp.br](mailto:roberta.fiel@unesp.br); <https://orcid.org/0000-0001-7456-7905>.

## Word spelling fluctuation: morphosyntactic aspects

### Abstract

In this paper, we analyzed sequences in which “o que” was spelled conventionally (“o que”) and unconventionally (“oque”). These fluctuations in spelling were extracted from texts of students who attended the last four years of Elementary School in a public school of the countryside of São Paulo, Brazil. Regarding the results, it was found that: (i) the subjects finish the Elementary School with a higher number of “o que” conventional form, however, the unconventional form “oque” also increases; (ii) in the graphical fluctuation between “o que” and “oque” two *types* based on morphosyntactic information are mobilized, one *type* with pronominal functioning and one *type* with interrogative functioning; and (iii) this fluctuation may be related to prosodic aspects from the relationship between speech and writing.

**Keywords:** segmentation of written words; speech and writing; word spelling fluctuation.

### Introdução

Segmentar sequências de letras é, possivelmente, um dos grandes desafios enfrentados por escreventes em processo de aquisição de escrita, uma vez que para essa tarefa é necessário mobilizar critérios fonológicos, morfológicos, sintáticos e semânticos que colocam em evidência a complexidade da noção de palavra.

Esses desafios em relação à segmentação de palavras têm sido explorados em trabalhos que evidenciam a relevância do estudo da escrita tanto dos anos iniciais (CAPRISTANO, 2007a, 2007b; CHACON, 2004, 2013), quanto dos anos finais (FIEL, 2018; TENANI, 2016; SILVA, 2014) do Ensino Fundamental (de agora em diante, EF).

Assim como as pesquisas citadas, neste trabalho também nos interessa a delimitação de fronteiras de palavras na escrita, mais especificamente, os momentos de flutuação em relação a fronteiras de palavras escritas, isto é, flutuação entre grafia convencional e não convencional de palavras. Sobre esses momentos flutuantes, Abaurre (1991) argumenta que uma mesma criança pode mobilizar, simultaneamente, critérios de segmentação aparentemente conflitantes em seu texto escrito:

Não é de se estranhar, portanto, que possam ser encontradas soluções diferentes para o mesmo problema de segmentação em um mesmo texto (da mesma forma como tão frequentemente encontramos soluções diferentes para uma mesma palavra no mesmo texto, em termos das letras que a compõem), já que a elaboração de um texto escrito delimita, para a criança, um espaço particular

de solução de problemas onde ela será chamada a (re)construir a linguagem através de sua transposição para a forma de representação escrita. Enquanto estiver trabalhando em tal construção, a criança, naturalmente elaborará hipótese conflitantes, porque estará testando deferentes critérios possíveis para resolver os mesmos problemas. (ABAURRE, 1991, p. 4).

Dentre os diversos tipos de flutuação de palavras, elegemos para investigação a flutuação entre as grafias “o que” e “oque”. Essa escolha baseia-se em Fiel (2018), trabalho no qual analisamos dados de segmentação não convencional de palavras em que há a ausência de fronteira entre palavras, como “denovo”, “concerteza”, e encontramos grande quantidade da grafia “oque” não convencional ao lado da grafia “o que” convencional em textos de mesmo aluno e entre diferentes alunos ao longo do EF II. Nesse trabalho, não era nosso objetivo analisar grafias convencionais em relação a grafias não convencionais, por essa razão, avançamos em relação a esse trabalho ao propor uma análise específica da flutuação “o que” e “oque”. É importante observar que a sequência “o que” pode ser segmentada de diversas maneiras na escrita, entretanto, nos textos do EF II a que temos acesso, encontramos somente a flutuação entre a grafia “o que” e “oque”.

Assim, neste artigo propomos: (i) descrever quantitativamente ocorrências de “o que” quando segmentado convencionalmente e não convencionalmente ao longo do EF II; (ii) descrever o funcionamento morfossintático de “o que” e “oque” em textos escritos por alunos do EF II; (iii) investigar se, nessa flutuação, existe uma correlação entre função morfossintática e segmentação gráfica de palavra; (iv) analisar a possível atuação de componentes fonológicos na segmentação convencional e não convencional de “o que”.

No que diz respeito à organização deste artigo, na seção seguinte, apresentamos os pressupostos teóricos que adotamos para trabalhar com os dados de flutuação gráfica. Posteriormente, descrevemos algumas características do material de análise e, em seguida, na seção de análise dos dados, realizamos uma descrição longitudinal do conjunto de dados no decorrer do EF II. Além disso, descrevemos as características morfossintáticas das grafias “o que” e “oque” e propomos uma interpretação baseada em pressupostos da fonologia prosódica para um texto com essa flutuação gráfica. Para concluir, nas considerações finais, retomamos os principais pontos trabalhados e evidenciamos possíveis desdobramentos da análise realizada.

## **1. Alguns pressupostos teóricos para o estudo da escrita**

Considerada a natureza do dado, isto é, a flutuação de “o que” em textos do EF II, faz-se necessário explicitar o posicionamento teórico assumido no que se refere aos conceitos de palavra, de segmentações não convencionais de palavras escritas e de escrita.

Definir o que é palavra não é uma tarefa simples, uma vez que seus limites podem variar a depender do critério considerado, ou seja, uma palavra pode ser delimitada por critérios gráficos, fonológicos, sintáticos, lexicais, entre outros (ROSA, 2000). Dentre esses diversos critérios de definição de palavra, interessa-nos aqueles concebidos no âmbito da fonologia e da morfossintaxe.

Na abordagem fonológica, a palavra é definida a partir da presença de um acento principal, como a palavra “casa”, que possui acento principal na primeira sílaba (“ca”) e, por isso, é considerada uma palavra fonológica. Todavia, no Português brasileiro (doravante, PB), também existem formas átonas, isto é, não acentuadas, como preposições (de, em, com), pronomes (o, a, me, nos) e conjunções (e, se). Na fonologia, essas formas não são chamadas de palavras fonológicas e sim de clíticos fonológicos, pois funcionam sempre ancoradas às palavras acentuadas que lhe são adjacentes (BISOL, 2000, 2005)<sup>2</sup>. Assim, na sequência “a casa amarela” temos duas palavras fonológicas (“casa” e “amarela” e um clítico fonológico – o artigo “a”). Entretanto, se o critério for morfossintático, nessa mesma sequência, identificam-se três palavras, pois, nessa abordagem, elementos átonos, como preposições, conjunções e artigos, são considerados palavras gramaticais, porque atuam gramaticalmente na sentença; já palavras de conteúdo lexical, como “casa” e “amarela”, são consideradas palavras lexicais, visto que fazem parte do léxico (ROSA, 2000).

Silva (2018), Fiel (2018) e Tenani (2016) mostraram que é justamente nessa não coincidência entre os limites de palavra fonológica e palavra morfossintática que se encontram as maiores dificuldades dos escreventes do EF II em relação à segmentação de palavras, pois não conseguem identificar como palavras algumas preposições, artigos e conjunções, devido a sua atonicidade, e, muitas vezes, hipossegmentam-nas com outros elementos. Nesse sentido, a flutuação entre “o que” e “oque”, evidencia essa dificuldade presente no EF II, já que os escreventes parecem não conseguir identificar se a sequência “o que” se trata de duas palavras separadas ou de uma palavra só (“oque”).

No que diz respeito às segmentações não convencionais de palavras escritas, a grafia não convencional com ausência do espaço em branco, como “oque”, é chamada de hipossegmentação na literatura que trata de dados de segmentação de palavras. As hipossegmentações e os outros tipos de segmentações não convencionais de palavras escritas<sup>3</sup> são, muitas vezes, interpretadas como problemas de alfabetização (ZORZI,

---

2 Para outros trabalhos que lidam com a noção fonológica de palavra cf. Toneli (2009, 2014), Fernandes (2007b), Schwindt (2013, 2007).

3 A literatura sobre segmentação não convencional de palavras define três tipos de ocorrências desse fenômeno: (i) hipossegmentação: quando há a ausência não-convencional de fronteira entre as palavras, como em “concerteza”, em que, convencionalmente espera-se “com certeza”; (ii) hipersegmentação: decorre da presença não-convencional de uma fronteira gráfica no interior da palavra, como na ocorrência “co migo”, para a qual a convenção prevê “comigo”; e (iii) Mescla ou híbrido: ocorrência simultânea de hipossegmentação e hipersegmentação na mesma sequência,

2006) ou como indícios de patologia (FERNÁNDEZ *et al.*, 2010), todavia entendemos que compreendê-las desse modo é desconsiderar a relação heterogênea que todos nós temos com a língua. Assim, neste trabalho, fazemos nossa afirmação de Abaurre, Fiad e Mayrink-Sabinson (1997, p. 16-17) de que:

[...] “erros” cometidos pelos aprendizes de escrita/leitura são, de fato, [...] preciosos indícios de um processo em curso de aquisição da representação escrita da linguagem, registros dos momentos em que a criança torna evidente a manipulação que faz da própria linguagem, história da relação que com ela (re) constrói ao começar a escrever/ler.

Desse modo, assumimos uma abordagem linguística que interpreta as segmentações não convencionais de palavras escritas como dados epilinguísticos de trabalho do sujeito da linguagem, seja em seu modo de enunciação falado, seja em seu modo de enunciação escrito. Essa abordagem está relacionada à concepção de escrita que adotamos para interpretar dados de escrita, a qual defende que a presença de marcas do oral no escrito não se dá por um processo de interferência da fala na escrita, pois a escrita não é homogênea, é, na verdade, heterogeneamente constituída pelos modos de enunciação falado e escrito (CORRÊA, 2004). Assim, nessa proposta, fala e escrita são concebidas como modos de enunciação intrinsecamente ligados às práticas sociais orais/faladas e letradas/escritas, isto é, os fatos da língua (falado/escrito) só se realizam por serem modos de enunciação concretizados por meio de práticas de linguagem.

Para argumentar como se estabelecem essas relações entre práticas orais/faladas e letradas/escritas, trabalhos que, assim como este, assumem uma concepção da escrita constituída de modo heterogêneo (CHACON, 2013, 2004, 1998; TENANI, 2016, 2013, 2004; CAPRISTANO, 2007a, b, dentre outros) se valem de aspectos da fonologia prosódica, mais especificamente, do arcabouço teórico da fonologia prosódica proposto por Nespor e Vogel (1986, 2007), pois a partir dessa teoria é possível trilhar caminhos para trabalhar com a complexidade da dimensão fônica da língua(gem).

Nespor e Vogel (1986, 2007) propuseram a existência de sete constituintes prosódicos que estruturam os enunciados das línguas do mundo. Segundo as autoras, esses constituintes são organizados hierarquicamente, assim, do menor para o maior tem-se: sílaba ( $\sigma$ ), pé métrico ( $\Sigma$ ), palavra prosódica ( $\omega$ ); grupo clítico (CG)<sup>4</sup>, frase fonológica

---

como, por exemplo, no dado “pora qui”, no qual há, em um primeiro momento, a hipossegmentação da palavra “por” com a primeira sílaba da palavra “aqui” e, posteriormente, a hipersegmentação da palavra “aqui”.

4 A definição do *status* do grupo clítico tem causado bastante repercussão e discussões. Entretanto, faz-se importante explicitar que, neste trabalho, nos limitamos a reconhecer que existe uma estrutura prosódica para o clítico e seu hospedeiro. Essa estrutura assumimos ser o chamado grupo clítico, defendido por Nespor e Vogel (1986, 2007) e Bisol (2000, 2005).

(Φ), frase entoacional (I) e enunciado fonológico (U). Neste trabalho, caracterizamos somente os domínios que são mais relevantes para a análise dos dados com os quais trabalhamos, que são esses: (i) a palavra prosódica ou fonológica<sup>5</sup> (ω) caracterizada, como já mencionado, por ser portadora de um acento lexical ou acento primário, como, por exemplo, a palavra morfológica “guarda-roupa” que se constitui de duas palavras prosódicas [guarda]ω [roupa]ω, por haver dois acentos: um na primeira sílaba de “guarda” e outro na primeira sílaba de “roupa”; e (ii) a frase fonológica (Φ) caracterizada a partir da configuração de sintagmas definidos pela sintaxe, por exemplo, no enunciado “[As meninas] Φ [comeram] Φ [o bolo] Φ [de pêssego] Φ”, encontram-se quatro sintagmas, três nominais e um verbal e, desse modo, quatro frases fonológicas.

Na seção seguinte, descrevemos o material do qual os textos e os dados trabalhados neste artigo foram extraídos.

## 2. Material e métodos

Os textos dos quais selecionamos os dados de flutuação gráfica pertencem a uma amostra longitudinal de textos que integra o “Banco de Dados de Escrita do Ensino Fundamental II” (TENANI, 2015)<sup>6</sup>. A elaboração das amostras que constituem esse banco deu-se por meio do projeto de extensão “Desenvolvimento de oficinas pedagógicas de leitura, interpretação e produção textual no Ensino Fundamental”, coordenado pelas professoras doutoras Luciani Ester Tenani e Sanderléia Longhin-Thomazi (IBILCE/UNESP), e desenvolvido ao longo de quatro anos (2008 a 2011) em uma escola pública estadual, situada em São José do Rio Preto, noroeste paulista, com a colaboração de alunos de graduação e de pós-graduação do IBILCE/UNESP.

A produção dos textos era precedida de discussão coletiva de variados temas com base na leitura de textos de diversos gêneros em sala de aula. Tenani e Longhin-Thomazi (2014)<sup>7</sup> mostram que todas as discussões e atividades desenvolvidas na escola foram pautadas em uma concepção de escrita enquanto modo de enunciação inserida em um processo de letramento vinculado a práticas de oralidade (CORRÊA, 2004).

---

5 Esse constituinte, em inglês *phonological word*, foi traduzido por Bisol (1996) como palavra fonológica e, posteriormente, em trabalhos sobre o português europeu (VIGÁRIO, 2007) e sobre o PB (FERNANDES, 2007), foi traduzido como palavra prosódica. Neste trabalho, optamos por usar a segunda opção, já que é a mais difundida em trabalhos sobre esse tema atualmente (cf. TENANI, 2016; GUZZO, 2015; TONELI, 2014; FERNANDES-SVARTMAN, 2012).

6 O “Banco de Dados de Escrita do Ensino Fundamental II” está disponível *online* em <http://www.convenios.grupogbd.com/redacoes/Login> e pode ser acessado, por meio de um cadastro, após autorização das coordenadoras.

7 Tenani e Longhin-Thomazi (2014) trazem uma descrição da metodologia e das atividades desenvolvidas no projeto de extensão.

Com base nessa abordagem teórica sobre escrita, as propostas de produção de texto foram realizadas visando proporcionar aos alunos experiências de leitura e de escrita vinculadas a práticas sociais orais e letradas (TENANI; LONGHIN-THOMAZI, 2014). Essas propostas foram elaboradas, em geral, com base em uma coletânea de textos ou em uma tira de humor. A partir desse material, os alunos eram convidados a discutir o tema em sala de aula e, posteriormente, a escrever seus textos seguindo as características de diferentes gêneros e tipos textuais, conforme previsto pelo programa curricular de cada ano letivo<sup>8</sup>.

Quanto à metodologia, selecionamos, do material descrito, 435 textos que possuíam as duas grafias ou ao menos uma delas, isto é, a hipossegmentação “oque” e/ou a grafia convencional “o que”. A fim de não perder de vista o aspecto longitudinal deste trabalho, os textos selecionados foram escritos por 20 alunos, 10 meninas e 10 meninos, que produziram textos em todos os anos do EF II.

Para descrição dos dados de um ponto de vista longitudinal, consideramos, baseadas em Fiel (2018) e Tenani (2016), a quantidade de palavras escritas em cada texto para gerar um coeficiente de ocorrência que leva em conta a razão entre o total de ocorrências e o total de palavras escritas ao longo dos quatro anos do EF II<sup>9</sup>.

No que diz respeito à análise dos dados, valemo-nos das abordagens quantitativa e qualitativa utilizadas por Bybee (2003). De acordo com a autora, existem dois métodos relevantes para organizar a frequência de dados em estudos linguísticos: a frequência *token*, que se refere à frequência textual de ocorrência de um item independentemente de seu significado, e a frequência *type*, que é contabilizada a partir da função que determinado item possui dentro de seu contexto. Desse modo, neste artigo, a frequência *token* diz respeito à contagem geral das ocorrências de “o que” e “oque” nos 435 textos e a frequência *type* refere-se à análise e à quantificação das diferentes funções morfossintáticas que essas grafias apresentam nos textos.

### 3. Resultados e discussão

Na primeira parte desta seção, trazemos uma análise quantitativa, na qual buscamos descrever a quantidade de dados de “o que” e “oque” no EF II a partir de um ponto de vista longitudinal. Na segunda parte, descrevemos as características morfossintáticas da flutuação da grafia “o que” e, na terceira parte, propomos análise, baseada em aspectos da fonologia prosódica, de um texto com flutuação gráfica de “o que”.

---

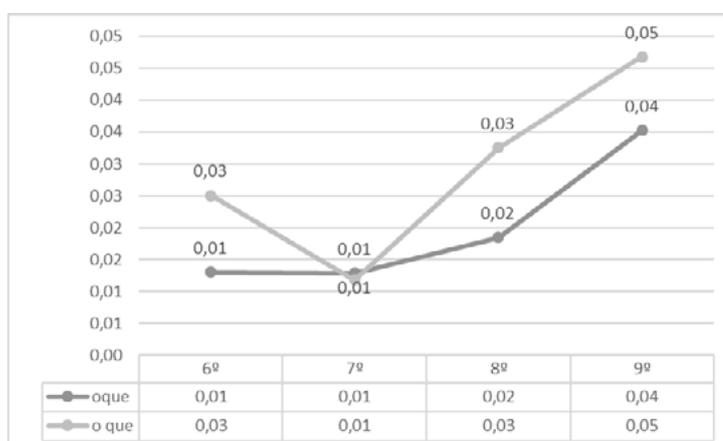
8 A descrição dos gêneros e de tipos textuais que constituem o Banco está disponível, por meio de cadastro, em: <http://www.convenios.grupogbd.com/redacoes/Login>.

9 Exemplo do cálculo realizado: Se o sujeito possui seis ocorrências da hipossegmentação “oque” em um total de 983 palavras escritas ao longo do EF II, o cálculo será o seguinte:  $6/983 = 0,00610$ . Na elaboração dos gráficos, consideramos o valor arredondado, no caso do exemplo 0,01.

### 3.1. “Oque” e “o que”: características longitudinais

A partir dos critérios descritos anteriormente, levantamos 106 ocorrências da sequência “o que” grafadas convencionalmente e 74 ocorrências de “oque” grafadas não convencionalmente.

Ao levarmos em conta o índice gerado a partir da média entre dados e quantidade de palavras escritas nos textos, observamos que tanto os dados convencionais, quanto os não convencionais apresentam um aumento ao longo do EF II, como se observa no gráfico 1:



**Gráfico 1.** Índice de dados de “oque” e “o que” ao longo do EF II  
**Fonte:** Elaboração própria

No que diz respeito aos dados de “o que”, no sexto ano, houve índice de 0,03 ocorrências; no sétimo ano, 0,01 ocorrências; no oitavo ano, 0,03 ocorrências e, no nono ano, 0,05 ocorrências, isto é, o maior índice em todos os anos. Desse modo, o sexto ano, em relação ao nono ano, apresenta um aumento de 0,02 de ocorrências.

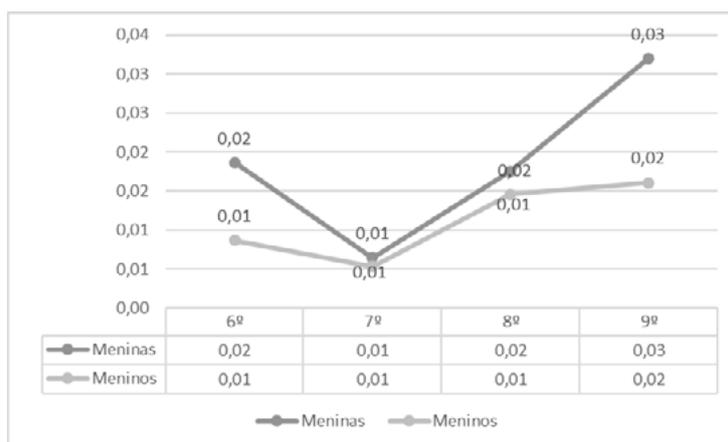
Os dados hipossegmentados de “oque” também evidenciam um aumento. No sexto e sétimo ano, houve índice de 0,01 ocorrências em cada ano, no oitavo ano, 0,02 ocorrências e, no nono ano, 0,04 ocorrências. Assim, se considerarmos o sexto ano em relação ao nono ano, há um aumento de 0,03 ocorrências.

Desse modo, nessa primeira análise, podemos constatar que para esse conjunto de sujeitos o tempo de escolarização parece atuar de modo significativo no aprendizado de convenções ortográficas, visto que os sujeitos terminam o EF com um maior índice de dados “o que” segmentados convencionalmente. Todavia, não há diminuição nos dados de “oque” não convencionais ao longo do EF II, evidenciando que pode haver outros

fatores, além da escolarização, atuando na manutenção desses dados não convencionais ao longo do EF II.

A fim de investigar outros fatores que podem colaborar com essa diminuição, propomos a análise do efeito do sexo/gênero<sup>10</sup> dos sujeitos, visto que nosso material é composto por 10 textos escritos por meninas e 10 textos escritos por meninos. Essa decisão baseia-se em Tenani (2016) que investigou dados de segmentação não convencional de palavras em textos escritos por alunos do EF II e observou que há diferença quantitativa entre ocorrências de meninas e meninos, sendo que as meninas constituem 51,71% da amostra de sujeitos e alcançam a taxa de 58,76% dentre os sujeitos com os menores índices de dados. Desse modo, nosso intuito é observar se essa tendência ocorre nos dados da flutuação “o que” e “oque”.

Assim, no Gráfico 2, ilustramos a distribuição do índice de dados de “o que” ao longo do EF II para meninas e meninos.



**Gráfico 2.** Índice de dados de “o que” ao longo do EF II para meninas e meninos

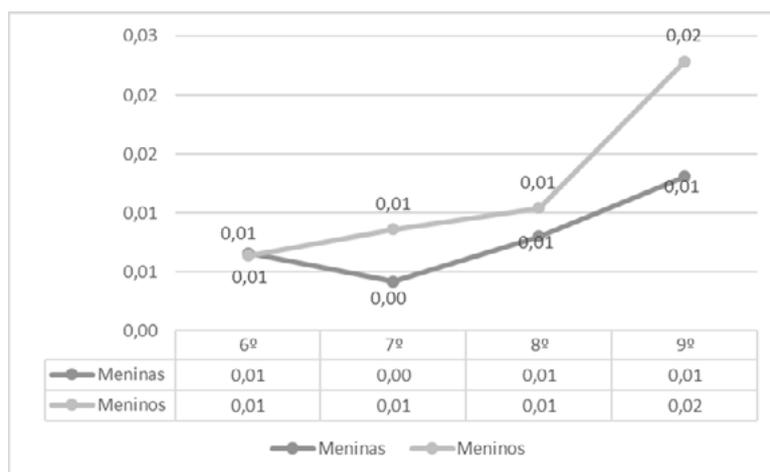
**Fonte:** Elaboração própria

A partir do Gráfico 2, é possível notar que, no que diz respeito à grafia convencional, as meninas mostram um maior índice, evidenciando um aumento de 0,01 ocorrências do sexto ano em relação ao nono ano. No sexto ano, elas apresentam um índice de 0,02 ocorrências; no sétimo ano, há uma diminuição para 0,01 ocorrências; no oitavo ano, volta a subir para 0,02 e terminam o EF com 0,03 ocorrências.

<sup>10</sup> Assim como Tenani (2016), emprestamos o termo “sexo/gênero” da Sociolinguística para evitar uma possível confusão com a noção de “gênero” textual/discursivo. Além disso, na coleta de dados do Banco de Escrita do EF II, não foi considerada informação quanto à identidade de gênero dos alunos, assim, só temos informação quanto ao sexo dos sujeitos.

Os meninos também mostram um aumento de 0,01 ocorrências do sexto ano em relação ao nono ano, todavia, eles mantêm uma frequência de 0,01 ocorrências durante os três primeiros anos do EF II, aumentando para 0,02 ocorrências no nono ano.

As informações sobre a grafia hipossegmentada “oque” para meninas e meninos seguem no Gráfico 3.:



**Gráfico 3.** Índice de dados de “oque” ao longo do EF II para meninas e meninos  
**Fonte:** Elaboração própria

No que se refere à grafia “oque”, observamos que, de modo geral, as meninas hipossegmentam menos do que os meninos. No sexto ano, elas apresentam índice de 0,01 ocorrências, no sétimo ano, o índice é zero, já no oitavo e nono ano, o índice é de 0,01 para cada ano. Desse modo, do sexto ano em relação ao sétimo ano não há mudança significativa no índice de ocorrências das meninas. Os resultados dos meninos mostram frequência de 0,01 ocorrências nos três primeiros anos do EF II e um aumento para 0,02 ocorrências no nono ano, ou seja, do sexto ano em relação ao nono ano há uma diferença de 0,01 ocorrências.

A partir desses resultados, observamos que os meninos, em relação às meninas, hipossegmentam mais e grafam menos convencionalmente a sequência “o que”, assim podemos argumentar que eles parecem ser responsáveis pelo aumento dos dados “oque” ao final do EF. Esse resultado vai ao encontro dos resultados estatísticos mostrados por Tenani (2016) de que, no EF II, meninas hipossegmentam menos do que meninos, visto que apresentam um maior índice da grafia convencional e um menor índice de hipossegmentação de “o que” e são as responsáveis pelo aumento da grafia convencional ao final do EF.

Assim, parece ser relevante considerar o efeito sexo/gênero dos sujeitos para descrever ocorrências de flutuação da sequência “o que”, no entanto, para um resultado mais preciso, é necessário submeter os dados a uma análise estatística, a fim de confirmar ou refutar a relevância desse efeito para esse conjunto de dados. Essa tarefa ficará para uma pesquisa futura, visto que não é nosso intuito realizar esse tipo de análise neste trabalho.

Na seção seguinte, descrevemos os aspectos morfossintáticos que a flutuação entre “o que” grafado convencionalmente e “oque” hipossegmentado desempenharam no conjunto de textos selecionados.

### 3.2. “O que” e “oque” no EF II: aspectos morfossintáticos

No conjunto de textos selecionados, extraímos 106 *token* de “o que” e 74 *token* de “oque”. Em ambos os *token* distinguimos dois *type* com base na informação morfossintática.

No *type* 1, “o que” e “oque” evidenciam funcionamento pronominal, isto é, função de substituir um nome, como em (1), (2) e (3).

(1) “Fica aí uma opinião minha, é **o que** eu penso e não to nem aí para **o que** vocês acham” (Z11\_8B\_05F\_05)

(2) “Cada um faz **o que** quer, o corpo não é meu, se a pessoa quiser estragar o corpo fumando o problema é dela [...]” (Z11\_8B\_05F\_06)

(3) “Também quero que todos consigam **oque** querem, pois nada é impossível [...]” (Z08\_5C\_26F\_03)

Dentro do conjunto de dados de “o que”, o *type* 1 teve frequência total de 85,8% (91/106), a maior frequência nesse conjunto de dados. No sexto ano, teve frequência de 9,4% (10/106), no sétimo ano, de 6,6% (7/106), no oitavo ano, de 16% (17/106) e no nono ano, de 53,8% (57/106).

No conjunto de dados “oque”, o *type* 1 apresentou frequência de 85,1% (63/74). No sexto ano, ocorreu em 7% (5/74) dos dados, no sétimo ano, em 11,3% (8/74) dos dados, no oitavo ano, em 15,5% (11/74) dos dados e no nono ano, em 54,9% (39/74) dos dados.

A maioria dos dados de “o que” e “oque” com funcionamento pronominal ocorreu nos gêneros textuais “relato de experiência” (29,2% - 45/154) e “artigo de opinião” (29,2% - 45/154).

No *type 2*, “o que” e “oque” mostraram funcionamento de pronome interrogativo, como em (4), (5) e (6).

(4) “**O que** você veio fazer na minha casa?” (Z11\_8B\_05F\_04)

(5) “**(O) que** me encomoda?” (Z09\_6C\_21F\_02)

(6) “**Oque** eu ia fazer lá na Disneylandia??” (Z08\_5D\_04M\_06)

O *type 2* mostrou frequência baixa se comparada ao *type 1*. No conjunto de dados de “o que”, ocorreu somente em 14,2% (15/106) dos dados. O sexto ano apresentou frequência de 5,7% (6/106), o sétimo, oitavo e nono ano apresentaram a mesma frequência de 2,8% (3/106) em cada ano.

Nas grafias “oque”, o *type 2* ocorreu em 14,9% (11/74). No sexto e sétimo ano, houve frequência de 4,2% (3/74) em cada ano, no oitavo ano, houve frequência de 1,4% (1/74) e no nono ano, 5,6% (4/74).

A maioria dos dados de “o que” e “oque” com funcionamento interrogativo ocorreu em gêneros textuais que, muitas vezes, possuem características de diálogos, como “carta pessoal” (30,8% - 8/26) e “narrativa de ficção” (15,4% - 4/26).

Ainda nesse conjunto de dados, somente a grafia convencional é grafada na periferia direita da estrutura sintática, isto é, como elemento QU- *in situ* (MIOTO; KATO, 2005); e isso ocorre em somente duas ocorrências (13,3%), as quais seguem transcritas em (7) e (8). No restante das ocorrências, o pronome interrogativo é grafado na periferia esquerda da estrutura sintática, ou seja, fora de seu lugar canônico (MIOTO; KATO, 2005), como em (4), (5) e (6). Essa configuração já era esperada, pois, segundo Miotto e Kato (2005), no PB, sentenças interrogativas com QU- deslocado apresentam frequência de 65,75% nas produções de sentenças interrogativas, contra apenas 8,84% de frequência de realizações de sentenças interrogativas com QU- *in situ*.

(7) “Vai fazer **o que**?” (Z09\_6C\_22M\_04)

(8) “E ela dise **o que**?” (Z08\_5D\_19F\_03)

Sobre as características morfossintáticas desses dados, podemos pontuar que as grafias “o que” e “oque” podem ocorrer tanto no *type 1* quanto no *type 2*, o que nos dá evidências de que, para esse conjunto de dados, a flutuação gráfica independe da estrutura morfossintática.

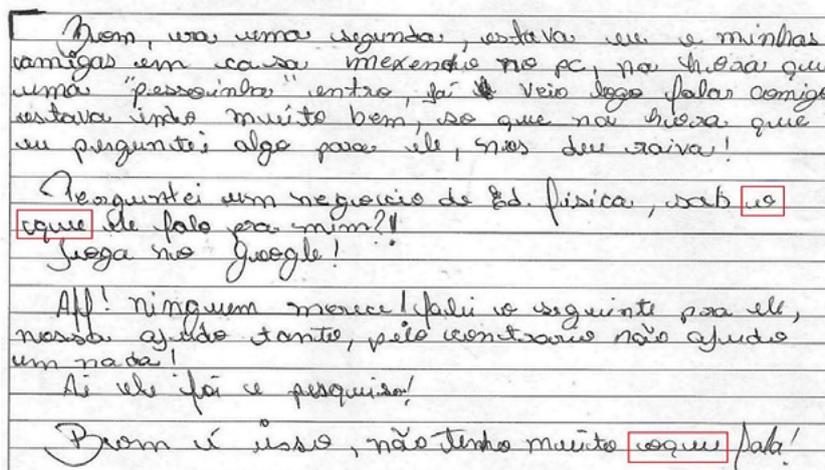
No que se refere à diferença quantitativa do *type* 1 em relação ao *type* 2, podemos argumentar que ela se dá devido à configuração do Banco de Dados de Escrita do EF II, pois, das 26 propostas de produção textual presentes nesse banco, 57,7% (15/26) são do gênero “relato de experiência” e “artigo de opinião”, enquanto somente 15,4% (4/26) são dos gêneros “carta pessoal” e “narrativa de ficção”.

Assim, uma hipótese inicial seria que essa configuração pode ter possibilitado que os escreventes produzissem mais textos com as características de relatos e artigos de opinião, viabilizando maior ocorrência de dados com funcionamento pronominal em relação ao funcionamento interrogativo. Corroborando com essa hipótese, Ticianel (2016), em trabalho que analisa dados de escrita infantil, mostra que dados com funcionamento interrogativo são mais comuns em gêneros textuais em que ocorre diálogo face a face. Segundo essa autora, dados de “oque” com funcionamento interrogativo acontecem, no EF I, em contextos em que há fala de personagens. Essa hipótese deve ser mais explorada em trabalho em que um maior número de ocorrências seja mobilizado, a fim de confirmá-la ou refutá-la.

### **3.3. “O que” versus “oque” em um mesmo texto: relação entre fala e escrita**

Com base nas análises feitas neste trabalho, observamos que, em uma primeira análise, a flutuação entre as grafias “o que” e “oque” parecem acontecer independentemente da estrutura morfossintática, desse modo, as informações morfossintáticas não nos deram pistas sobre o funcionamento dessa flutuação gráfica. Todavia, se voltarmos nosso olhar para a relação entre fala e escrita, mais especificamente, entre prosódia e escrita, podemos encontrar pistas que nos ajudam a elaborar possíveis hipóteses explicativas para a permanência dessa flutuação nos anos finais do EF.

A fim de ilustrar essa relação, trazemos, a seguir, um texto escrito no oitavo ano do EF, no qual há a grafia convencional “o que” e a hipossegmentação “oque”:



**Figura 1.** Texto com flutuação da grafia “o que”  
**Fonte:** Z10\_7A\_04F\_03<sup>11</sup>

No texto da Figura 1, o escrevente narra uma situação de sua vida na qual ocorreu um conflito. Em (11), temos o trecho em que a grafia “o que” é segmentada convencionalmente. Notamos que, nesse trecho, parece haver uma possível ênfase em “o que” que é caracterizada na escrita pelo espaço entre “o” e “que”. Mesmo esse espaço sendo esperado dentro da convenção ortográfica, esse recurso pode ter sido utilizado na busca por enfatizar que o conflito entre as personagens será evidenciado naquele momento.

(11) “Perguntei um negócio de ED. Física, sab **o que** ele falo pra mim? (Z10\_7A\_04F\_03)

Sob um ponto de vista que leva em conta a relação entre prosódia e escrita, Fernandes (2007) argumenta que a ênfase prosódica pode ser projetada na escrita por meio de espaços entre as palavras. Além disso, Perini (2010) e Chutumiá (2013) explicam que interrogativas abertas do tipo “Q-” constituem sozinhas sintagmas interrogativos, essa informação permite-nos interpretar “o que” em (11) como correspondente a limites de frase fonológica, já que segundo Rosignoli (2017), fronteiras de frase fonológica podem marcar elementos enfáticos. Além disso, Toneli (2009), em trabalho que analisa o estatuto prosódico de palavras funcionais em ocorrências de fala experimental, argumenta que palavras funcionais em contexto de focalização podem receber acento tonal, independentemente da posição em que ocupam dentro da frase entoacional<sup>12</sup>.

11 Sugestão de leitura: "Bom, era uma segunda, estava eu e minhas amigas em casa mexendo no pc, na hora que uma "pessoinha" entro, já (rasura\*) veio logo falar comigo, estava indo muito bem, so que na hora que eu perguntei algo para ele, nos deu raiva! Perguntei um negocio de Ed.fisica, sab o que ele falo pra mim?! Joga no Google! Aff! Ninguem merece! falei o seguinte pra ele, nossa ajuda tanto, peio contrario não ajuda em nada! Ai ele foi e pesquisou! Bom é isso, não tenho muito oque fala!"

12 Para discussões importantes sobre o estatuto prosódico de clíticos no PB cf. Silva (2018) e Bisol (2000).

No texto da Figura 1, ainda temos a grafia hipossegmentada, a qual se encontra no trecho transcrito em (12).

(12) “Bom é isso, não tenho muito **oque** fala” (Z10\_7A\_04F\_03)

No que diz respeito a essa grafia, ainda sob um ponto de vista prosódico, podemos interpretá-la como uma palavra prosódica, pois o “que” não apresenta acento lexical, entretanto, ao se unir ao “o”, recebe um acento prosódico e passa a se comportar como uma possível palavra prosódica (TENANI, 2011), isto é, um só elemento funcionando no sistema linguístico. Nesse trecho, diferente do anterior, não parece haver ênfase, visto que se trata da conclusão do texto.

Assim, defendemos que a flutuação entre as grafias “o que” e “oque” no texto da Figura 1 parece mobilizar fatores de natureza fonológica e gráfica. Esses fatores nos auxiliam na construção de hipóteses explicativas para a permanência e a grande ocorrência desse tipo de dado no final do EF II. Corrêa (2004), ao lidar com dados desse tipo em textos de vestibulandos, pontua que eles parecem dar indícios de tentativas de atribuição de fidelidade à representação dos enunciados falados nas produções textuais.

Desse modo, com base nos argumentos citados, concordamos com as afirmações de Capristano e Ticianel (2014) de que a dificuldade dos escreventes em segmentar “o que” decorre da relação entre fatores fonológicos e gráficos, por isso a sua permanência em séries mais avançadas do EF e até mesmo em textos de vestibulandos, como mostrou Corrêa (2004).

#### **4. Considerações finais**

Neste trabalho, descrevemos dados de flutuação gráfica entre “o que” e “oque” retirados de textos escritos por alunos no EF II e observamos que, longitudinalmente, há um uso maior de “o que” grafado convencionalmente. Todavia, a hipossegmentação “oque” não findou ao final do EF II, evidenciando, ao contrário, um aumento.

Esse resultado, como mostramos, pode estar relacionado ao sexo/gênero dos escreventes, visto que o grupo de meninos detém a maior quantidade de dados ao longo do EF II. Essa hipótese deve ser investigada mais profundamente em trabalho com um maior número de escreventes e que se proponha a analisar a questão sexo/gênero em textos escritos com base em análise estatística.

Sobre o funcionamento morfossintático, observamos que, no conjunto de textos analisados, tanto a grafia “o que” quanto a grafia “oque” ocorrem em maior número com funcionamento pronominal e em gêneros textuais em que há descrição de fatos e argumentação em busca da defesa de um ponto de vista, como em “relatos de experiência” e “artigos de opinião”. Esse resultado diferencia-se de resultados de escrita do EF I, pois Ticianel (2016) mostrou que a maioria dos dados de “oque” no EF I evidenciam funcionamento interrogativo e acontecem, majoritariamente, em contexto de discurso direto, isto é, nos momentos em que há fala de personagens.

Os poucos casos em que observamos funcionamento interrogativo, tanto “o que” quanto “oque” foram grafados fora da posição canônica na estrutura sintática. Esse resultado vai ao encontro da tendência da língua, pois Mito e Kato (2005) mostram que, no PB, sentenças interrogativas com QU- deslocado apresentam frequência de 65,75% nas produções de sentenças interrogativas, contra apenas 8,84% de frequência de realizações de sentenças interrogativas com QU- *in situ*.

Além dessas características morfossintáticas, observamos que, no conjunto de dados analisados, as grafias “o que” e “oque” podem ocorrer tanto no *type 1* quanto no *type 2*, evidenciando que para esse conjunto de dados a flutuação gráfica independe da estrutura morfossintática. Entretanto, se considerarmos relações entre fala e escrita, veremos que a flutuação entre “oque” e “o que” dão pistas de configuração prosódica de enunciados relacionada às maneiras pelas quais saliências métricas e ênfase entoacional podem ser associadas a porções dos enunciados (TENANI; LONGHIN-THOMAZI, 2015). Desse modo, defendemos que a permanência da flutuações “o que” e “oque” ao longo do EF II evidencia marcas da heterogeneidade constitutiva da escrita (CORRÊA, 2006, 2004) que caracterizam a escrita do EF II e que as considerar como problemas de alfabetização é desconsiderar a relação heterogênea que todos nós temos com a língua(gem).

Concluimos apontando para a necessidade de se ampliar a quantidade de textos e, por conseguinte, de dados, a fim de investigar se os resultados e as hipóteses presentes neste trabalho encontram respaldo em um conjunto maior de ocorrências.

## REFERÊNCIAS

ABAURRE, M. B. M. A relevância dos critérios prosódicos e semânticos na elaboração de hipóteses sobre segmentação na escrita inicial. *Boletim da Abralín*, Campinas, v. 11, p. 203-217, 1991.

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

BISOL, L. O clítico e seu hospedeiro. *Letras de hoje*, Porto Alegre, v. 40, n. 3, p. 163-184, 2005.

BISOL, L. O clítico e seu status prosódico. *Revista de estudos da Linguagem*, Belo Horizonte, v. 9, n. 1, p. 5-30, 2000.

BISOL, L. *Introdução a estudos de fonologia do Português Brasileiro*. Porto Alegre: EDIPUCRS, 1996.

BYBEE, J. Mechanisms of change in grammaticalization: the role of frequency. In: JOSEPH, B.; JANDA, R. (ed.). *The handbook of historical linguistics*. Oxford: Blackwell, 2003.

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

CAPRISTANO, C. C. *Segmentação na escrita infantil*. v. 1. São Paulo: Martins Fontes, 2007b.

CAPRISTANO, C. C.; TICIANEL, G. F. Discurso direto e hipossegmentações na escrita infantil. *Filologia e Linguística Portuguesa*, São Paulo, v. 16, n. 1, p. 233-259, jan./jun. 2014.

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

CHACON, L. Flutuação na segmentação de palavras: relações entre constituintes prosódicos e convenções ortográficas na escrita infantil. *Filologia e Linguística Portuguesa*, São Paulo, v. 15, p. 369-383, 2013.

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

CHUTUMIÁ, D. I. *As interrogativas -q do português de Moçambique: contribuição para uma análise comparativa com o português europeu e o português brasileiro*. 2013. Dissertação (Mestrado) – Faculdade de Letras, Universidade do Porto, Porto, 2013.

CORRÊA, M. L. G. Heterogeneidade da escrita: a novidade da adequação e a experiência do acontecimento. *Filologia e Linguística Portuguesa*, São Paulo, n. 2, p. 269-286, 2006.

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

FERNÁNDEZ, A. Y. *et al.* Avaliação e intervenção da disortografia baseada na semiologia dos erros: revisão da literatura. *Revista CEFAC*, São Paulo, v. 12, n. 3, p. 499-504, 2010.

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

FERNANDES, F. R. Tonal association in neutral and subject-narrow-focus sentences of Brazilian Portuguese: a comparison with European Portuguese. *Journal of Portuguese Linguistics*, v. 6, n. 1, p. 91-115, 2007b.

FIEL, R. P. *Estudo longitudinal de hipossegmentações em textos do ensino fundamental II*. 2018. Dissertação (Mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto, 2018.

GUZZO, N. *A prosodização de clíticos e compostos em Português Brasileiro*. 2015. Tese (Doutorado em Letras) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2015.

MIOTO, C.; KATO, M. As interrogativas Q do português europeu e do português brasileiro atuais. *Revista da ABRALIN*, v. 4, n. 1 e 2, p. 171-196, dez. 2005.

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

NESPOR, M.; VOGEL, I. *Prosodic phonology: with a new foreword*. Berlin: New York: Mouton de Gruyter. 2007.

PERINI, M. A. Interrogativas abertas (interrogativas –Q). *In*: PERINI, M. A. *Gramática do português brasileiro*. São Paulo: Ática, 2010. p. 124-125.

ROSA, M. C. *Introdução à morfologia*. São Paulo: Contexto, 2000.

ROSIGNOLI, C. C. O padrão entoacional das sentenças interrogativas parciais do português brasileiro em fala manipulada. *Estudos linguísticos*, São Paulo, v. 46, n. 1, p. 41-54, 2017.

SCHWINDT, L. C. S. *et al.* A influência da variável escolaridade em fenômenos fonológicos variáveis: efeitos retroalimentadores da escrita. *Revista Virtual de Estudos da Linguagem – ReVEL*, v. 5, n. 9, p. 1-12, ago. 2007.

SCHWINDT, L. C. S. *et al.* Palavra fonológica e derivação em português brasileiro: considerações para a arquitetura da gramática. In: BISOL, L.; COLLISCHONN, G. (org.). *Fonologia: teorias e perspectivas*. Porto Alegre: EDIPUCRS, 2013. p. 15-28.

SILVA, L. M. da. *Grafias não convencionais de preposições e sílabas pretônicas: pistas de prosodização de clíticos preposicionais*. 2018. Tese (Doutorado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto, 2018.

SILVA, L. M. da. *Um estudo longitudinal sobre as hipersegmentações de palavras escritas nos anos finais do ensino fundamental*. 2014. Dissertação (Mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto, 2014.

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

TENANI, L. E. *Banco de dados de escrita do ensino fundamental II*. São José do Rio Preto: Instituto de Biociências, Letras e Ciências Exatas, UNESP, 2015. Disponível em: <http://www.convenios.grupogbd.com/redacoes/Login>. Acesso em: 10 nov. 2015

TENANI, L. E. Hipersegmentação de palavras: análise de aspectos prosódicos e discursivos. *Linguagem & Ensino*, Pelotas, v. 16, n. 2, p. 305-326, 2013.

TENANI, L. E. A segmentação não-convencional de palavras em textos do ciclo II do Ensino Fundamental. *Revista da ABRALIN*, Brasília, v. 10, n. 2, p. 91-119, 2011.

TENANI, L. E. Segmentações não-convencionais e teorias fonológicas. *Letras de Hoje*, Porto Alegre, v. 39, n. 3, p. 233-244, 2004.

TENANI, L. E.; LONGHIN-THOMAZI, S. R. Oficinas de leitura, interpretação e produção textual no Ensino fundamental. *Em Extensão*, Uberlândia, v. 13, n. 1, p. 20-34, 2014.

TICIANEL, G. F. *Discurso direto e hipossegmentações na aquisição da escrita*. 2016. Dissertação (Mestrado em letras) – Centro de Ciências Humanas, Letras e Artes, Universidade Estadual de Maringá, Maringá, 2016.

TONELI, P. M. *A palavra prosódica em Português Brasileiro: o estatuto prosódico das palavras funcionais*. 2007. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2009.

TONELI, P. M. *A palavra prosódica no português brasileiro*. 2014. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2014.

VIGÁRIO, M. O lugar do grupo clítico e da palavra prosódica composta na hierarquia prosódica: uma nova proposta. *In: LOBO, M.; COUTINHO, M. A. (org.). Actas do XII Encontro Nacional da Associação Portuguesa de Linguística – Textos seleccionados*. Lisboa: Colibri Artes Gráficas, 2007, p. 673-688.

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.