

Segmentação de palavras e convenções ortográficas: um olhar para a aquisição da escrita ao longo dos séculos

DOI: <http://dx.doi.org/10.21165/el.v48i1.2354>

Adelaide Maria Nunes Camilo¹

Resumo

O presente trabalho tem como objetivo central a comparação entre as segmentações não-convencionais feitas por crianças em período de aquisição de escrita no século XX e segmentações não-convencionais encontradas em textos produzidos em português entre os séculos XIV e XVIII. Sendo assim, buscamos discutir a relação que o sujeito escrevente estabelece com a linguagem no momento da construção de seu texto, observando como sua inserção em práticas sociais letradas pode influenciar a sua produção escrita. Os dados de segmentação não-convencional encontrados serão analisados qualitativamente de acordo com os pressupostos do *paradigma indiciário*, proposto por Ginzburg (1987).

Palavras-chave: segmentação de palavras; aquisição de linguagem; oralidade; letramento.

¹ Universidade Estadual de Campinas (UNICAMP), Campinas, São Paulo, Brasil; adelaidecamilo@gmail.com; <http://orcid.org/0000-0003-1158-0883>

Word segmentation and spelling conventions: a look at the writing acquisition over the centuries

Abstract

The purpose of this paper is to observe and compare the process of word segmentation in texts produced by Brazilian students which are acquiring the writing language and texts written in Portuguese between the sixteenth and eighteenth centuries. In this sense, our aim is to discuss the relationship that the author establishes with language at the time of the construction of his text, observing how his or her insertion in literate social practices can influence his or her written production. The unconventional segmentation data will be qualitatively analysed according to the assumptions of the evidential paradigm, proposed by Ginzburg (1987).

Keywords: word segmentation; language acquisition; orality; literacy.

Introdução²

Sabemos que muitos são os trabalhos que notam as semelhanças existentes entre a segmentação de palavras em textos de crianças em fase de alfabetização e nos primeiros textos produzidos em Língua Portuguesa; todos eles passam tangencialmente pela questão, mas sem tomá-la como foco principal de seu estudo.

Em busca de compreender e justificar tais semelhanças, a partir da concepção de que a segmentação de palavras é marcada por fatores próprios da língua e que se torna explícita no processo de aquisição de escrita, buscaremos, aqui, discutir as semelhanças e eventuais diferenças existentes nas segmentações de palavras presentes nos dois tipos de textos, investigando como a organização da língua, em termos de estrutura prosódica e morfológica, pode se refletir nas tentativas de segmentação daqueles que ainda não dominam as convenções ortográficas.

Comparando sistematicamente as produções infantis e de adultos de períodos anteriores da língua, apontando suas similaridades e divergências, buscamos comprovar que a imagem que o escrevente faz da escrita é parte de um imaginário socialmente constituído e partilhado. Dessa maneira, acreditamos que, na escrita, materializam-se as práticas sociais das quais ele, direta ou indiretamente, faz parte, assim como proposto por Corrêa (2004) ao tratar de redações de vestibulandos.

² O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Assim, pretendemos discutir quais características da segmentação da escrita podem ser tidas como herança do processo de padronização ortográfica da Língua Portuguesa e quais características são próprias do processo de aquisição da língua escrita. Além disso, pretendemos averiguar quais características estão mais presentes na escrita de crianças e quais aparecem mais frequentemente na escrita de adultos, refletindo ainda sobre se há alguma diferença marcante na escrita dessas duas categorias de sujeitos.

Revisão bibliográfica

Uma vez que a questão da segmentação de palavras se vincula às diferentes convenções ortográficas definidas ao longo do tempo para a escrita de uma língua, investigamos, na história do português, os acordos e as normas que buscavam estabelecer e fixar uma ortografia para nossa língua.

Segundo Sousa (2004, p. 316, grifos da autora), as primeiras tentativas de padronização do português datam do século XVI, com as publicações das gramáticas de Fernão de Oliveira e João de Barros.

É fato consensual que o estabelecimento de uma norma sistematizada por força institucional se dá tardiamente na história da língua portuguesa. Ou seja, apesar da publicação de compêndios que propõem sistematizações para a grafia – Barros (1532), Gandavo (1572), Lião (1576), Alvaro Ferreiro Véra (1631) – não se pode falar em uma norma ortográfica para o português antes das reformas de fins do século 18.

Em função do tardio estabelecimento de normas ortográficas, atestava-se um sentimento de liberdade entre os escritores, já que cada um organizava seus textos de acordo com as normas que acreditava serem as mais adequadas, ou, muitas vezes, as que deixavam o texto mais próximo aos padrões adotados para o latim, abrindo espaço para uma variação individual que, como afirma Cagliari (1994, p. 108), trouxe o “caos para a escrita”.

Naqueles casos, porém, em que havia incerteza ou esquecimento da escrita usual de outrora e, por mingua de conhecimentos etimológicos, ou não acudia ao espírito a imagem do respectivo termo latino, ou não se percebia a relação fonética entre os vocábulos de uma e outra língua, nesses casos vacilava-se na grafia, escrevendo vocábulo ora de um, ora de outro modo, ou então firmava-se a maneira de escrever muitas vezes em pura contradição com a prática do passado. (SAID ALLI, 1921 apud SOUSA, 2004, p. 315).

A respeito desse “caos” instalado na escrita da época, em 1878, Barbosa Leão, célebre escritor de gramáticas e compêndios de regras ortográficas, destaca as consequências para o ensino da língua nas escolas: “A orthografia portugueza é pois um cahos, um verdadeiro Protheu de mil fórmas caprichosas, que se transforma, na escola primária, em cabeça de Medusa para tormento dos alumnos e desespero dos professores.” (LEÃO, 1878, p. 102).

É exatamente diante desse cenário de variação individual que centramos nossa comparação entre os textos de crianças e os textos de períodos anteriores de nossa língua, uma vez que, como já apresentado por Cagliari (1994, p. 107),

É curioso e interessante notar que as crianças em fase de alfabetização costumam cometer erros de grafia de palavras produzindo formas de escrita semelhantes às desses primeiros escreventes do Português. Na verdade, quer os antigos, quer as crianças, ao escrever, encontram-se diante do desconhecido, não sabendo qual seria a forma ortográfica das palavras.

Assim, acreditamos que, ao se deparar com o desconhecido, o escrevente se coloca frente à língua à procura, nesse mesmo sistema, de algo em que possa fundamentar suas indagações, de forma que as âncoras mais utilizadas nos momentos de dúvidas gráficas tendem a ser tanto a fonética, ora ligando cada letra a um som específico ou segmentando a cadeia escrita de modo semelhante à segmentação da cadeia fônica, quanto os conhecimentos prévios que o escrevente traz do que é a escrita e a imagem que faz dela.

Para além das variações individuais no que diz respeito à grafia das palavras, à medida que a língua portuguesa assumiu a função de língua de cultura, houve a necessidade de escrevê-la e, com isso, uma das principais dificuldades enfrentadas foi a segmentação da fala em palavras, como apontam Ilari e Basso (2006, p. 25),

Ainda hoje, a segmentação é, às vezes, motivo para hesitações, e até pessoas familiarizadas com a escrita ‘erram’, escrevendo <derepente> em vez de <de repente>, <porisso> em vez de < por isso>, ou mesmo < esta-mos> em vez de <estamos>. A separação de palavras foi uma das grandes aquisições do período medieval, mas os autores medievais e clássicos utilizaram, às vezes, em seus textos, segmentações diferentes das que são consideradas corretas hoje.

É interessante notar que essa variação na segmentação de palavras pode ser encontrada até mesmo dentro de textos de um único autor, ou mesmo dentro de uma mesma obra, como constatado por Cortesão (1967, p. 90) ao fazer sua transcrição da Carta de Caminha:

Assim em *obatel*, *aboca*, formações mais simples, *o* e *a* “são elementos átonos desprovidos de sentido lexical, que formam corpo com o elemento seguinte, de cujo acento participam e cujo género e número têm por função sinalar”. E o autor acrescenta: “No plural, Pêro Vaz de Caminha escrevia: *os batees*, *as bocas*, etc., separando o artigo do substantivo. Tal separação era de prever, porquanto a presença, também no artigo, do sinal do plural, dava àquele um particular relevo, fazendo-o aparecer como elemento gramatical distinto. Assim, a grafia da *Carta*, unindo o artigo ao substantivo, no singular, e separando-o, no plural, traduzia o estado da língua falada com um vivo sentimento de realidade”.

Desse modo, assumimos o já apontado por Kajita (2009, p. 64-65), segundo o qual

[...] a segmentação não-canônica obviamente não é um fenômeno historicamente recente, antes, trata-se de um fenômeno que independe do tempo, e que tem como variáveis o padrão ortográfico vigente na época e o nível de adequação de um determinado sujeito a esse padrão.

No entanto, apesar das inúmeras mudanças ortográficas relativas à grafia dos fonemas sofridas pela língua portuguesa durante todo seu processo de se tornar uma língua autônoma e oficial, não pudemos encontrar, nos documentos das reformas ortográficas, menções ao estabelecimento do critério para a separação dos elementos clíticos. O que nos leva a pensar que, a princípio, as mudanças se deram de forma individual por meio do uso das regras ortográficas utilizadas pelas tipografias nas publicações de textos de autores consagrados e, vagarosamente, alcançaram a língua como um todo, tornando-se o padrão a ser seguido.

Prova dessa tentativa de normatização ortográfica por meio das diferentes regras utilizadas nas casas de publicação é dada pelo seguinte comentário de Barbosa Leão (1878, p. 102) em seus documentos e estudos a favor da reforma ortográfica:

Nas diferentes typographias nacionaes, a primeira cousa que se pergunta a um escriptor, que não queira ser revisor orthographico das suas próprias obras, é, se quer a orthographia da casa. Um dos mais conspícuos membros d’essa real academia, ha pouco fallecido,

dizia: “Que na impossibilidade de saber qual era a orthographia mais auctorizada e preferível dos clássicos portuguezes, aceitava a orthographia variável das typographias onde mandava imprimir os seus trabalhos literários”.

Nesse sentido, a instabilidade da segmentação de palavras até hoje se apresenta nos textos daqueles em fase de aquisição da escrita, ou daqueles que, em algum momento da vida, foram introduzidos no mundo letrado, mas que, por um motivo ou outro, não puderam avançar em seus estudos, mantendo-se estagnados quanto ao domínio das convenções estabelecidas para o código escrito.

Acreditamos, assim como Marquilhas (2003, p. 285), que:

[...] dada uma cultura escrita, as realizações gráficas que contrastem com convenções coesas são uma metalinguagem dos seus autores; no desvio à convenção encontram-se vestígios de uma conceptualização linguística, facto que permite fundamentar hipóteses sobre a natureza de estruturas da língua.

1.3 A segmentação na escrita da criança

No que diz respeito, especialmente, à escrita infantil em período de alfabetização, como já discutido anteriormente, ao tentar lidar com as convenções ortográficas, parece que as crianças buscam se ancorar em padrões já conhecidos por elas e, muitas vezes, tomam como ponto de partida as semelhanças entre o caractere gráfico e o som da língua.

Sendo assim, a percepção fonética seria o primeiro passo para a segmentação. A partir do contato com textos escritos, há uma expansão dos critérios considerados para segmentar a escrita, de maneira que passam também a utilizar os conhecimentos adquiridos de semântica e morfologia, como já defendido por Abaurre e Cagliari (1985), para definir o local dos espaços em branco. Tal procedimento pode acabar gerando dúvidas, de forma que os escreventes tendem a segmentar mais do que a ortografia prevê ao atribuírem, por exemplo, sentidos a subpartes de palavras.

Se é verdade que, em alguns momentos, por trás das hipóteses de escrita está a fala, revelando-se através de características que a criança (ainda não influenciada pela própria escrita!) demonstra perceber muito bem, é também verdade que ela já incorpora em muitos outros momentos as marcas específicas dessa escrita que está sendo chamada a contemplar. (ABAURRE, 1988, p. 137).

Ainda nesse viés, Capristano (2007) e Chacon (2004), ao analisar as fugas às convenções ortográficas na segmentação de palavras, destacaram que esses desvios poderiam ser motivados por um movimento de subjetivação do escrevente, uma vez que este demonstra uma “preocupação” com a colocação dos espaços em branco previstos pelas convenções. Os mesmos autores defendem, ainda, que esses dados forneceriam sinais da organização prosódica da língua,³ assim como da maneira como os sujeitos lidam com informações letradas.

No que diz respeito ao trabalho do escrevente com as informações letradas, múltiplos aspectos linguístico-discursivos estariam envolvidos no processo de segmentação do enunciado, como aponta Corrêa (2004). De acordo com o autor, tanto o enunciado falado como o enunciado escrito se caracterizam pela convivência, em um mesmo enunciado, de características do oral/falado e do letrado/escrito. Em sua teoria, a heterogeneidade da linguagem se manifestaria em três eixos: (i) o que diz respeito à ideia que o escrevente faz da constituição da escrita; (ii) a representação que o indivíduo faz do código escrito institucionalizado; e (iii) o eixo da dialogia falado/escrito. A partir desses três eixos, é possível considerar as representações da escrita que passam pelo escrevente.

Ao considerar a heterogeneidade constitutiva da linguagem, o enunciado produzido pelas crianças em fase de aquisição de linguagem seria, portanto, uma combinação de

(1) informações lingüísticas que circulam em práticas de letramento, nas quais se dá o confronto da criança com informações [...] difundidas por meio de propriedades gráfico-visuais, e (2) informações lingüísticas que circulam em práticas de oralidade, nas quais se dá o confronto da criança com informações [...] difundidas por meio de propriedades acústico-auditivas (CHACON, 2013, p. 3).

Nesse sentido, as observações de Chacon (2013) se coadunam com as de Corrêa (2004), concluindo que as flutuações ao segmentar as palavras estariam relacionadas aos diferentes modos de planejamento do texto, que também variam conforme as experiências do escrevente no que tange ao modo como se relaciona e convive com práticas orais e práticas escritas:

O conhecimento teórico que permite ao analista reconhecer a flutuação do escrevente está também, em parte, relacionado com os diferentes modos de planejamento do texto oral e do texto escrito, os quais, não explicitamente e apenas em termos práticos,

3 Segundo a Teoria dos Constituintes Prosódicos organizada por Nespor e Vogel (1986), conforme apresentada em Bisol (2005), a hierarquia prosódica seria constituída por sete domínios: sílaba, pé, palavra fonológica, grupo clítico, frase fonológica, frase entoacional e enunciado fonológico.

são dominados também por qualquer falante/escrevente. Tal conhecimento, porém, frequentemente deixa de atuar na produção escrita do escrevente, pondo-o em situação de variar o grau de convivência entre o oral/falado e letrado/escrito em função de seu envolvimento com o assunto abordado ou com o destinatário a quem se dirige. (CORRÊA, 2004, p. 19).

Dessa forma, buscamos aproximar os textos de crianças em fase de aquisição de escrita a textos produzidos por adultos entre os séculos XVI e XIX, apontando que, a depender do grau de intimidade que o escrevente apresenta com o código escrito institucionalizado, as dificuldades apresentadas quanto à segmentação das palavras se mantêm, justamente, em função da instabilidade própria da língua, marcada historicamente pelo processo de evolução e de tardia padronização por ela sofrida.

Procedimentos metodológicos: material e método de análise

Para esse estudo, tomamos como *corpus* de análise textos de crianças de diversas escolas públicas e particulares pertencentes ao Banco de Dados de Aquisição de Escrita, organizado no âmbito do Projeto Integrado de Pesquisa intitulado *A relevância teórica dos dados singulares na aquisição da linguagem escrita*, financiado pelo CNPq e desenvolvido no Instituto de Estudos da Linguagem da UNICAMP entre 1999 e 2005, sob a coordenação da professora Maria Bernadete Marques Abaurre.

Também compõem o material textos produzidos por adultos e pertencentes a períodos anteriores da língua, englobando um período entre os séculos XVI e XIX. Como fonte de textos provenientes dos séculos XVI e XVII, utilizamos Gramáticas da Língua Portuguesa publicadas em Portugal, como a *Gramática da linguagem portuguesa*, publicada em 1536 por Fernão de Oliveira. No que tange aos registros de português brasileiro nos séculos, foram utilizados textos pertencentes ao Corpus Histórico do Português, constituído no âmbito do Projeto *Para a História do Português Brasileiro* (PHPB), sob coordenação do professor Ataliba de Castilho.

Para a análise dos dados, o material foi dividido em dois grandes grupos: (i) textos atuais produzidos por crianças e (ii) textos produzidos por adultos em períodos anteriores. O segundo grupo teve, ainda, uma segunda divisão: autores consagrados, representados por documentos oficiais, e os textos produzidos por mãos inábeis.⁴ Nesse sentido, a

4 Termo cunhado por Marquilhas (2003, p. 278) e usado para tratar de escreventes com pouca familiaridade com o código escrito. Em suas palavras, “falantes em fase incipiente de aquisição da escrita”. Em seus trabalhos, Marquilhas discute a escrita em Portugal no século XVII, caracterizando as práticas sociais de leitura e escrita a partir de arquivos da Inquisição Portuguesa. Ao discutir essas produções, a autora observou, dentre outros aspectos, a ausência de uma ortografia

divisão acima proposta para o *corpus* diacrônico se sustenta na hipótese de que, apesar de haver variação em todos eles, as mãos inábeis apresentariam mais semelhanças com as segmentações não-convencionais produzidas por crianças, enquanto as outras se manteriam mais em variações estruturais convencionizadas, como a colocação ou não do espaço em branco entre elementos clínicos em função da ausência de regras claras.

Cabe ainda ressaltar que os textos foram analisados qualitativamente com base no *paradigma indiciário* de Ginzburg (1989, p. 149), “método interpretativo centrado sobre resíduos, sobre os dados marginais, considerados reveladores”, cuja aplicação aos estudos de aquisição da escrita foi proposta e discutida em Abaurre, Fiad e Mayrink-Sabinson (1997). Essa escolha se justifica uma vez que, nesse trabalho, teve-se como intuito demonstrar que os “erros” cometidos pelo aprendiz tratam-se de “indícios de um processo em curso de aquisição da representação escrita da linguagem, registros de momentos em que a criança torna evidente a manipulação que faz da própria língua” (ABAURRE; FIAD; MAYRINK-SABINSON, 1997, p. 16).

Discussão e análise

Nesta seção, para melhor apresentação dos dados encontrados, a análise será feita, primeiramente, aproximando os *corpora* e discutindo suas semelhanças para, em um segundo momento, distanciá-los ressaltando as particularidades existentes em cada um deles.

Como representante dos textos de adultos letrados do século XVI, apresentamos trechos da Gramática da linguagem portuguesa, de Fernão de Oliveira (1536):

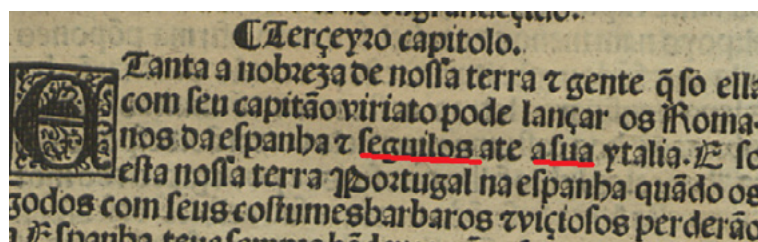


Imagem 1. Gramática da linguagem portuguesa – pronome⁵

homogênea entre escreventes pouco familiarizados com o código escrito e escreventes de documentos oficiais, tanto no que diz respeito ao uso das letras, quanto à segmentação das palavras.

5 Lê-se: “Tanta a nobreza de nossa terra e gente que só ela com seu capitão Viriato pode lançar os Romanos da Espanha e ‘seguilos’ até ‘asua’ Itália”.

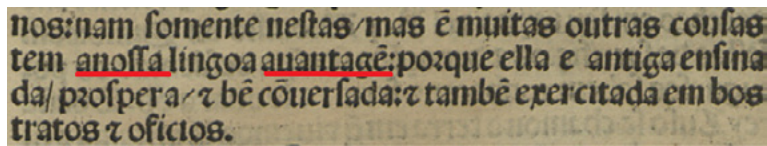


Imagem 2. Gramática da linguagem portuguesa – artigo⁶

É interessante notar que, nas duas imagens acima, encontramos elementos clíticos unidos à palavra de conteúdo que os sucede. Tanto na imagem 1, como na 2, observa-se a hipossegmentação do artigo definido “a” à palavra fonológica a seguir, como em “asua”, “anossa” e “avantaçãem”. Contudo, a hipossegmentação do artigo não parece ser uma regra, visto que, na imagem 1, ele aparece separado do substantivo “nobreza”, em “tanta a nobreza de nossa terra”. Apesar de este tipo de dado ser comumente encontrado em textos de iniciantes na escrita, é importante lembrarmos que, neste caso, a obra analisada se trata, justamente, de uma gramática prescritiva, que tem como principal intuito apontar a riqueza da língua portuguesa e discorrer sobre como ela deve ser usada. Nesse contexto, não seria esperado haver variações no uso da escrita. Todavia, não podemos afirmar com certeza que essa variação é de responsabilidade do autor ou se o responsável por isso seria o revisor ou editor do texto, dado que se trata de um material impresso, diferentemente do *corpus* infantil e de mãos-inábeis aqui analisados.

Ainda no que diz respeito às imagens anteriores, no início do Capítulo 3, vemos o pronome átono “lo” hipossegmentado com relação ao verbo “seguir”, em “seguilos”, construção também muito comum em textos infantis, como poderemos observar abaixo.

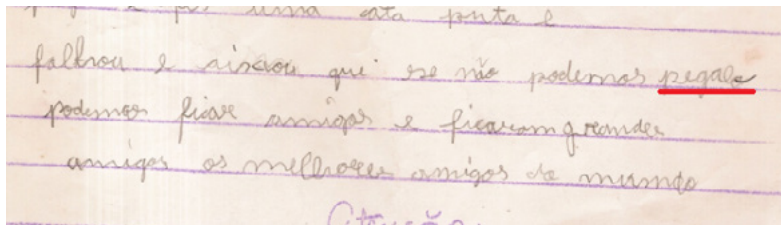


Imagem 3. Texto infantil – pronome⁷

Como já apontado por Abaurre (1991) e Tenani (2011), nos textos de crianças em fase de aquisição da escrita, parece haver uma tendência maior para a hipossegmentação,

6 Lê-se: “Não somente nestas, mas em muitas outras coisas tem ‘anossa’ língua ‘avantaçãem’: porque ella é antiga, ensinada, próspera e bem conservada, e também exercitada em bons tratos e officios”.

7 Lê-se: “Falhou e avisou que se não podemos ‘pegalo’ podemos ficar amigos e ficaram grandes amigos os melhores amigos do mundo”.

uma vez que, assim como Ferreiro e Pontecorvo (1996, p. 45) afirmam, “há coincidência nos dados no que concerne à maior facilidade para distinguir substantivos, verbos e adjetivos e uma grande dificuldade para aceitar que artigos, conjunções, preposições e outros elementos de ligação sejam palavras”. O mesmo pode ser observado nos textos dos adultos: há uma maior quantidade de casos de hipossegmentações de elementos clíticos do que hipersegmentações desses mesmos elementos quando parte integrante de palavras.

No que concerne ao uso dos pronomes átonos ao longo de toda a gramática de Fernão de Oliveira, assim como nas gramáticas de seus contemporâneos João de Barros (1540) e Pêro de Magalhães de Gândavo (1574), nos parece que tendem a aparecer unidos às formas verbais. Uma exceção a esse comportamento se dá na obra de Duarte Nunes de Leão (1576) que, apesar de não abordar o assunto em sua teorização de normas ortográficas, apresenta uma variação no uso dos pronomes, ora unindo-os ao verbo, como em “dobraremse”, ora separando-os, como ocorre em “tomando os”.

Essa instabilidade na colocação dos espaços em branco entre os pronomes átonos e os verbos nos primeiros estudos de ortografia da língua portuguesa se mantém ao longo dos anos. No Compêndio de Orthografia de Frei Luis de Monte Carmelo (1767), por exemplo, podemos, até mesmo, encontrar variação na colocação do hífen como forma de separação entre o pronome e a palavra de conteúdo, ocorrendo tanto em casos de próclise, como de ênclise: se-escrevem, se-explicam e dividi-se.

O clítico “se” é, também, um grande causador de dúvidas na segmentação infantil, de modo que encontramos dados em ele aparece hipossegmentado à palavra fonológica ou, até mesmo, sendo reconhecido, quando parte de uma palavra, como um elemento átono:



Imagem 4. Texto infantil⁸

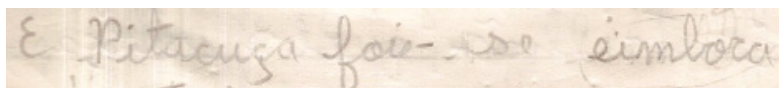


Imagem 5. Texto infantil⁹

8 Lê-se: “Chegando na praça, Pitacuça ‘sentouse’”.

9 Lê-se: “E Pitacuça foi-se eimbora”.

Não obstante, é importante ressaltar que a variação na escrita dos autores de gramáticas finda, no século XX, com a Reforma Ortográfica de 1911, em razão da definição do uso do hífen para a união de pronomes pessoais enclíticos aos verbos. Entretanto, o mesmo não acontece nos textos de adultos mãos-inábeis, posto que, assim como as crianças, sua fuga à convenção não se dá em função de um padrão atestado pelas gramáticas, mas pela dificuldade em reconhecer e trabalhar com os elementos átonos que mobilizam em sua escrita. São exemplos disso as imagens abaixo, pertencentes ao Banco de Dados do PHPB (1906-2000), a primeira coletada de uma carta escrita em 1908 e a segunda escrita em 1955.

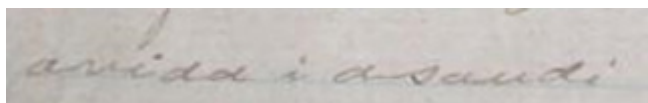


Imagem 6. Texto adulto¹⁰

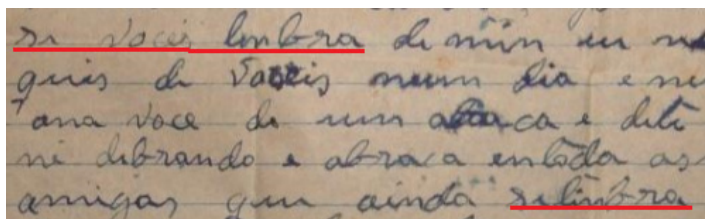


Imagem 7. Texto adulto¹¹

As segmentações não-convencionais, porém, vão muito além do trabalho com elementos clíticos. Em alguns dos textos analisados, encontramos hipossegmentações de estruturas que não podem ser consideradas átonas, como em uma das cartas pertencentes ao banco do PHPB, em que aparece a seguinte sentença: "me aceutouhum Cavallo, da/ sua terra e empenhado de seus Parentes". O mesmo tipo de ocorrência pode ser visto no exemplo abaixo:

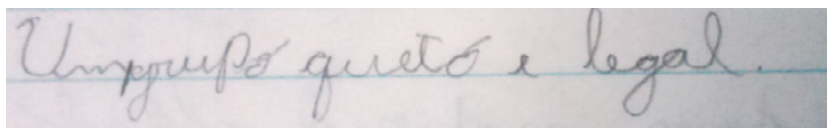


Imagem 8. Texto Infantil

10 Lê-se: "avida i asaudi".

11 Transcrição: "se voces lenbra de min eu nuca mis-lquis de voeis num dia e nu| ana voce de um abraço e dete| ni debrando e abraço entoda as | amigas que ainda selinbra".

Entretanto, é interessante notar que, no dado produzido por adulto, o pronome oblíquo “me” que antecede a construção não foi hipossegmentado ao verbo, apenas o numeral. Todavia, nesses dados, não nos é possível afirmar que a fuga às convenções ocorreu em função da percepção de um elemento mais fraco e que, por conta de sua atonicidade, não teria sido reconhecido como uma palavra de conteúdo pelo escrevente.

Como já afirmou Abaurre (1991, p. 6), “é impossível, naturalmente, determinar com certeza o critério que está por trás das escolhas de segmentação de cada criança, em cada caso particular”, porém, por meio de uma observação atenta dos textos, podemos levantar hipóteses. Por mais que o numeral e artigo indefinido “um” seja tônico, ele é constituído por apenas duas letras (ou três, no caso da escrita adulta) e, como já comprovado por Ferreiro (2013), o escrevente em fase de aquisição da escrita tem dificuldades em reconhecer monossílabos como palavras. Uma das possíveis causas dessa hipossegmentação, portanto, poderia ser o reduzido tamanho da palavra.

No tocante à imagem 8, especialmente, vale notar que a criança retornou à sua escrita e inseriu uma barra separando o artigo indefinido “um” e o substantivo “grupo”. Como já dito anteriormente, não nos é possível identificar o que levou a criança a retornar ao seu texto e reescrevê-lo. Podemos declarar, no entanto, que esse retorno foi espontâneo, uma vez que não há marcas de correção externa. Também podemos notar que esse retorno ocorreu após o término da escritura do texto, já que pelo espaçamento do texto e pelo traçado do lápis, a barra foi inserida após a escritura da palavra “legal”.

Acreditamos assim que, ao reelaborar sua escrita nas imagens 4, 5 e 8, o escrevente deixa transparecer indícios de seu trânsito do escrevente por práticas sociais orais/faladas e letradas/escritas. Nas imagens 4 e 5, respectivamente, a criança hipossegmenta o clítico “se” em “sentouse” e, na imagem seguinte, separa-o por meio do hífen. Em todos os casos mencionados, é possível notar o trabalho do escrevente com marcas específicas da escrita, nomeadamente, o hífen e a barra. Nesse âmbito, as explicações para as segmentações não-convencionais vão mais além da percepção do ritmo da língua e envolvem seu trânsito por práticas sociais orais/faladas e letradas/escritas, assim como o imaginário que faz da escrita, uma vez que, no português brasileiro, a próclise é mais produtiva do que a ênclise; apesar disso, o escrevente opta por utilizar o pronome enclítico, com o intuito de alçar sua escrita ao ideal do código escrito institucionalizado.

Há ainda, no Banco de Dados do PHPB, ocorrências um pouco idiossincráticas, que diferem do que foi encontrado nos demais materiais de textos de adultos e nos textos produzidos por crianças em fase de aquisição de escrita, são elas:

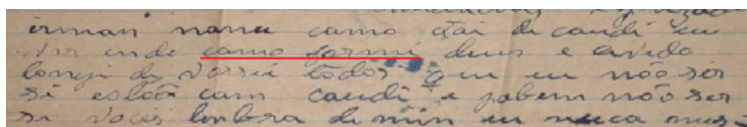


Imagem 9. Carta 50 – como formi¹²

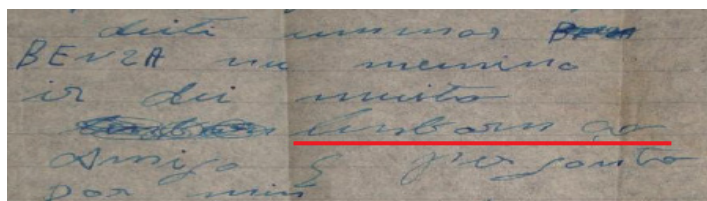


Imagem 10. Carta 21 – lenbarn ça¹³

No dado apresentado em 9, acreditamos que a segmentação da palavra “conforme” tenha sido feita como “como formi” por conta de uma hipercorreção e percepção de semelhança prosódica entre as duas formas. Entretanto, o mesmo critério não pode ser usado para explicar a ocorrência de “lenbarn ça” para “lembrança”, como vemos na imagem 10.

Em nosso trabalho de levantamento de dados, não encontramos ocorrências semelhantes a essa em textos de crianças, o que nos leva a indagar quais os motivos que fazem com que haja diferenças entre as segmentações de crianças em fase de aquisição de escrita e de adultos que não concluíram devidamente o processo de alfabetização.

Um primeiro argumento a respeito das diferenças entre os dados de mãos-inábeis e os dados infantis poderia ser o fato de que, enquanto as crianças ainda estão sistematizando o aprendizado da segmentação e utilizam sua escrita como momento de teste de possibilidades, os adultos estão estagnados nessa fase e, ao invés de testarem hipóteses, tendem a se arriscar menos, acabando, muitas vezes, por hipercorrigir suas segmentações e optar por segmentações que, aos seus olhos, fiquem mais próximas do padrão do português, como alteração da estrutura silábica complexa CCV para CVC, mais produtiva no português brasileiro. Apesar de ser comum observar essa alteração da estrutura silábica em crianças aprendendo a falar, em nossos dados de escrita, ela não foi observada.

12 Transcrição: “Irman nanu como vai de caudi eu| vor indo como formi deus e civido| longi di vosses todos que eu não sir| si estão com caudi e pabem não ser| se voces lenbra de min eu nuca mis-|quis de voceis num dia e nunhoro| ana voce de um abraço”.

13 Transcrição: “Conpadi eu estoul| com muita saudadi| du senhor du menino| da mia con madi| Almerinda conpadi| deiti ummas BESA | BENSU nu menino| ir dei muita| lenbarn lenbarn ça| Amigo que pergontal| por min”.

Outra possibilidade seria o tipo de material analisado; enquanto os textos infantis são produções escolares, passíveis de leitura e correção dos professores ou colegas, os textos dos adultos mãos-inábeis são cartas para familiares ou profissionais em cargos superiores aos deles. Para esses adultos, o código escrito não é visto como lugar para hipóteses ou brincadeiras, pelo contrário, a mensagem deve ser emitida e compreendida pelo leitor, sem espaço para dúvidas e verificações de conteúdo. Além disso, provavelmente, seu período de alfabetização (muitas vezes, não escolar), foi interrompido, de maneira que não houve momento para alçar sua escrita à convenção, ou para discussões a respeito dos critérios utilizados para separação das palavras.

3. Considerações finais

Ao comparar os textos de crianças em fase de aquisição de linguagem com os textos produzidos por adultos em séculos anteriores, saltam-nos aos olhos a importância das relações que o sujeito estabelece entre si e seu texto, o que faz com que deixe marcas de trânsito por práticas sociais orais/faladas e letradas/escritas, assim como pistas do processo de formação de seu enunciado e da língua.

No que tange aos textos de adultos letrados, foi possível identificar que a variação em sua segmentação se deve, basicamente, à ausência de critérios estabelecidos para a segmentação dos elementos clíticos. Desse modo, a partir da aprovação de uma reforma ortográfica e definição de critérios morfológicos de separação entre palavras, a instabilidade na escrita dos clíticos diminui, dando espaço para uma segmentação estável.

Os textos de adultos mãos-inábeis, por sua vez, apresentam uma segmentação mais semelhante aos textos de infantis, mantendo uma tendência para a hipossegmentação de elementos clíticos a palavras fonológicas. Porém, uma vez que o trânsito por práticas sociais orais/faladas e letradas/escritas parece ser diferente entre os dois sujeitos, há discrepâncias nos tipos de segmentações não-convencionais apresentadas: enquanto as crianças demonstram mais liberdade para testar hipóteses e tentar alçar sua escrita ao imaginário que fazem do código escrito, os adultos parecem estar estabilizados em uma fase anterior de alfabetização, refletindo em sua escrita tendências de hipercorreção e manutenção de padrões silábicos mais produtivos no português.

Sendo assim, se torna possível afirmar que, por mais que uma língua evolua e se distancie de sua forma primeira, sua instabilidade continuará sendo representada por aqueles que se mostram mais sensíveis às estruturas e ainda pouco capturados pelas normas rígidas definidas por tratados gramaticais: os aprendizes em alfabetização e aqueles com menor domínio do código escrito.

REFERÊNCIAS

ABAURRE, M. B. M. O que revelam os textos espontâneos sobre a representação que faz a criança do objeto escrito? *In*: KATO, M. A. (org.). *A concepção da escrita pela criança*. v. 1. 2. ed. Campinas: Pontes Editores, 1988. p. 135-142.

ABAURRE, M. B. M.; CAGLIARI, L. C. Textos espontâneos na primeira série: evidências da utilização, pela criança, de sua percepção fonética da fala para representar e segmentar a escrita. *Cadernos CEDES – Antropologia e Educação Interfaces do Ensino e da Pesquisa*, São Paulo, v. 14, p. 25-29, 1985.

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

BARROS, J. de. *Grammatica da língua portuguesa*. Olyssipone apud Ludouicum Rotorigium Typographum, 1540.

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

BISOL, L. Mattoso Câmara Jr. e a Palavra Prosódica. *DELTA*, São Paulo, v. 20, n. especial, p. 59-70, 2005.

CAGLIARI, L. C. Algumas reflexões sobre o início da ortografia da língua portuguesa. *Cadernos de Estudos Lingüísticos*, Campinas, v. 27, p. 103-111, 1994.

CARMELO, F. L. do M. *Compendio de Orthographia, com sufficientes catálogos, e novas regras* [...]. Lisboa: Na officina de António Rodrigues Galhardo, 1767.

CAPRISTANO, C. C. *Aspectos de segmentação na escrita infantil*. São Paulo: Martins Fontes, 2007.

CHACON, L. Constituintes prosódicos e letramento em segmentações não-convencionais. *Revista 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, n. 2, p. 369-383, jan./jun. 2013. DOI: <http://dx.doi.org/10.11606/issn.2176-9419.v15i2p369-383>.

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

CORTESÃO, J. *A carta de Pero Vaz de Caminha: com um estudo de Jaime Cortesão*. Rio de Janeiro: Portugália, 1967.

FERREIRO, E. *O ingresso na escola e nas culturas do escrito: seleção de textos de pesquisa*. São Paulo: Cortez, 2013.

FERREIRO, E.; PONTECORVO, C. Os limites entre as palavras. A segmentação em palavras gráficas. In: FERREIRO, E.; PONTECORVO, C.; MOREIRA, N.; HIDALGO, I. G. *Chapeuzinho Vermelho aprende a escrever*. São Paulo: Ática, 1996. p. 38-66.

GÂNDAVO, P. de M. de. *Regras que ensinam a maneira de escrever e orthographia da língua portuguesa, com hum dialogo que adiante se seggue em defensam da mesma língua*. Lisboa: Na officina de Antonio Gonçalues, 1574.

GINZBURG, C. Sinais: raízes de um paradigma indiciário. In: GINZBURG, C. *Mitos, emblemas, sinais*. Tradução Federico Carotti. São Paulo: Companhia das Letras, 1989.

ILARI, R.; BASSO, R. *O português da gente: a língua que estudamos, a língua que falamos*. 2. ed. São Paulo: Contexto, 2012.

KAJITA, A. S. *A segmentação inábil: um estudo da segmentação ortográfica não-canônica*. 2009. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2009.


LEÃO, D. N. de. *Orthographia da Língua Portuguesa*. Lisboa: João de Barreira, 1576.

LEÃO, J. B. *Coleção de estudos e documentos a favor da refórma da ortografia em sentido sónico*. Lisboa: Imprensa Nacional, 1878.

MARQUILHAS, R. *A faculdade das letras: leitura e escrita em Portugal no século XVII*. Bragança Paulista: EDUSF, 2003.

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

OLIVEIRA, F. *Gramática da Língua Portuguesa*. Lisboa: Imprensa Nacional, 1536.



SOUSA, M. C. P. *Língua Barroca: sintaxe e história do português nos 1600*. 2004. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2004.

TENANI, L. E. Letramento e segmentações não-convencionais de palavras. *In: TFOUNI, L. Letramento, escrita e leitura: Questões Contemporâneas*. Campinas: Mercado de Letras, 2011. p. 229-243.