

## Segmentação de palavras na alfabetização de crianças: um estudo sobre consciência de palavra

*Word segmentation in children's literacy: a study about word awareness*

**Débora Mattos Marques**  
**Aline Lorandi**

Universidade do Vale do Rio dos Sinos – UNISINOS – Porto Alegre – Brasil



**Resumo:** A presente pesquisa teve por objetivo investigar de que maneira a consciência linguística acerca do conceito de palavra pode influenciar as segmentações na escrita de crianças nos primeiros anos no Ensino Fundamental, observando, então, dados de hiper e hiposegmentação e relacionando-os com a consciência de palavra. Para análise dos dados de consciência linguística, utilizou-se o Modelo de Redescrição Representacional, de Karmilloff-Smith (1986, 1992), que postula quatro níveis em que o conhecimento é redescrito na mente humana, tornando-se, ao longo do tempo, acessível para consciência e verbalização. A metodologia de pesquisa consistiu em seis testes, sendo quatro para verificação nos níveis de consciência de palavra e dois para obtenção de dados de escrita. Na análise dos dados percebeu-se que grande parte dos equívocos segmentais na escrita tem relação com o entendimento que esse sujeito possui acerca do que é uma palavra. Como resultado, então, encontraram-se evidências de que as segmentações incomuns não são motivadas apenas por questões prosódicas ou fonológicas, mas também por questões de consciência linguística a respeito do que esse sujeito entende por palavra.

**Palavras-chave:** Consciência de palavra. Hipersegmentação. Hiposegmentação; modelo de Redescrição Representacional.

**Abstract:** The present research aimed to investigate how linguistic awareness regarding the concept of "word" may influence some mistakes on segmenting words in children's writing in the Elementary School. The observed data comprised those of hyper and hyposegmentation which were then related to word awareness. For the analysis of linguistic awareness data, the Representational Redescription, proposed by Karmilloff-Smith (1986-1992), has been used. It postulates four levels where knowledge is redescrbed in the human mind, becoming accessible for awareness and verbalization along with the time. The research methodology consisted of six tests, out of which four were applied in order to verify word awareness, and, the other two tests, to obtain samples of writing data. Thus, it was noticed that a great part of the segmentation mistakes identified in the collected writings are related to the informants' ability to distinguish between different words until the moment they were observed. As a result, the uncommon segmentation mistakes found in the analyzed data evidenced that not only are they motivated by prosodic or phonological matters, but they are also influenced by linguistic awareness issues involving the informants' understanding of word.

**Keywords:** Word Awareness. Hypersegmentation. Hyposegmentation. Representational Redescription Model.

## Introdução

Embora seja um assunto muito recorrente dentro e fora da sala de aula, os casos de hipo e hipersegmentação são pouco estudados em termos de pesquisa no Brasil, representando um campo com muitas dúvidas, mas também muitas possibilidades de estudo.

A presente pesquisa debruça-se sobre essa temática, relacionando-a com a consciência linguística, a fim de perceber como a consciência que as crianças têm acerca do que é uma palavra influencia na segmentação da escrita durante o processo de alfabetização. Grande parte dos estudos sobre segmentações incomuns<sup>1</sup> da escrita recaem sobre questões prosódicas e fonológicas para explicar esse fenômeno, mas acredita-se que a consciência de palavra também pode influenciar na escrita infantil.

Para a análise de dados, a opção teórica e metodológica é o modelo de Redescrição Representacional (doravante MRR), que propõe olhar para a aquisição do conhecimento em uma perspectiva desenvolvimental. Abordar a consciência de palavra e as segmentações incomuns da escrita a partir desse viés permite perceber não apenas como as crianças interagem com a própria língua, mas também como a compreendem e constroem hipóteses acerca do sistema linguístico.

## 2 Consciência Linguística

Um dos conceitos mais conhecidos de consciência linguística vem dos autores Tunmer e Herriman (1984<sup>2</sup> apud Lorandi, 2011a), que afirmam que a consciência é a capacidade de pensar sobre e manipular traços estruturais da língua, tratando a linguagem como objeto de pensamento e não como um simples processo de compreender e produzir

<sup>1</sup> A expressão “segmentações incomuns” é, nesse texto, sinônimo de “segmentações não convencionais”.

<sup>2</sup> TUNMER, W. E.; HERRIMAN, M. L. The development of metalinguistic awareness: a conceptual overview. In: TUNMER, W. E.; PRATT, C.; HERRIMAN, M. L. (editors). **Metalinguistic awareness in children: theory, research and implications**. Berlin: Springer-Verlag, 1984, p. 12-35.

palavras. Para esses autores essa emergência representa a passagem de um conhecimento implícito para um conhecimento explícito.

Já a autora Karmiloff-Smith (1986, 1992) propõe quatro níveis de consciência, sendo um implícito e três explícitos. Em sua teoria, ela trabalha com a ideia de redescrição mental, em que, ao atingir um determinado nível, o conhecimento é redescrito, tornando-se, progressivamente, acessível para consciência e verbalização.

Dessa forma, pensando em consciência linguística a partir do Modelo de Redescrição Representacional, Lorandi (2011b) propõe um novo conceito:

(...) a consciência linguística pode ser definida como o alcance de determinados níveis de representação mental, nos quais o conhecimento está em um formato no qual é possível:

- mostrar sensibilidade aos recursos da língua (E1<sup>3</sup>);
- trabalhar em tarefas off-line (E2);
- pronunciar-se sobre o conhecimento (E3).

Esse é o conceito de consciência linguística adotado neste trabalho, uma vez que abrange os diferentes níveis de redescrição pelo qual o conhecimento passa até a consciência possa emergir.

Quando se pensa especificamente na consciência de palavra, é preciso deixar clara qual a concepção de *palavra* que se adota, já que ela pode ser vista de diferentes maneiras, variando conforme o foco do estudo. Segundo Rosa (2006),

Afora o uso na escrita, podemos entender palavra: (a) como uma unidade fonológica; (b) como o elemento mínimo da estrutura sintática; (c) como um elemento do vocábulo da língua.

Neste trabalho, adotou-se como conceito de palavra como “um elemento do vocábulo da língua, sendo entendida como parte do léxico do português brasileiro e, na escrita, separada pelos espaços gráficos.

<sup>3</sup> Esses níveis de explicitude serão apresentados detalhadamente na seção 2.2.

Esse conceito não é, obviamente, tão fácil para quem está no início da escolarização. Em função disso, destaca-se a importância de se investigar qual é o entendimento das crianças sobre o que é uma palavra. Cardoso e Castro (2012) retomam o conceito de Christensen (1991<sup>4</sup> apud Cardoso e Castro, 2012, p. 136), de que

a consciência de palavra é a capacidade de compreensão de uma palavra como um elemento constitutivo do discurso e que inclui a capacidade de segmentar frases em palavras, de separar as palavras do seu referente, de as substituir e de reconhecer sinônimos e antônimos.

Essas autoras realizaram uma pesquisa com crianças em idade pré-escolar e escolar, falantes do português europeu, e buscaram descrever a consciência de palavra a partir de tarefas de segmentação, levando em conta itens funcionais e lexicais. E, além disso, “identificar quais as dimensões do conhecimento linguístico (fonológico, morfológico e sintático) que estão presentes quando uma criança identifica as fronteiras de palavras num enunciado linguístico apresentado oralmente” (CARDOSO e CASTRO, 2012, p. 137). Em seus resultados, as frases que continham clíticos<sup>5</sup> apresentaram um menor índice de acertos, estando os maiores índices nas frases com nomes e verbos.

De suas conclusões, duas são particularmente interessantes: 1) Que os resultados apontam para o fato de que é a escolaridade que favorece o desenvolvimento da consciência de palavra (uma vez que as crianças em idade escolar apresentaram resultados melhores do que as em idade pré-escolar), e 2) “As crianças apresentam pouca consciência das palavras funcionais não acentuadas, daí realizarem maior número de segmentações não convencionais” (*ibid*, p. 149).

Assim como Cardoso e Castro, grande parte dos estudos sobre a emergência da consciência de palavra afirma que a manifestação metalinguística sobre o que é palavra dá-se com a alfabetização e em torno dos sete anos. Karmiloff-Smith (1996)

<sup>4</sup> Christensen, C. (1991). An Evaluation of the Metalinguistic Awareness Program. The University of Queensland: Recuperado em 7 Outubro, 2010, de [//www.aare.edu.au/91pap/chric91042.txt](http://www.aare.edu.au/91pap/chric91042.txt).

<sup>5</sup> Palavras funcionais não acentuadas. Ex: “mas”, “de”, “em”.

realizou uma pesquisa com crianças menores, em idades de 4;6 a 5 anos e verificou que mais de 50% das crianças de 4 anos e mais de 90% das de 5 anos apresentaram respostas corretas em um teste de identificação da última palavra ouvida em uma história (o mesmo teste que foi aplicado nesta pesquisa). A autora mostra que esses dados provocam uma nova reflexão, especialmente no que diz respeito às possíveis relações existentes entre consciência de palavra e alfabetização, uma vez que antes dos sete anos as crianças já conseguem perceber o que é uma palavra.

Em seu estudo foram utilizadas palavras de classe aberta e classe fechada<sup>6</sup>, e os dados só mostraram diferença significativa no grupo de crianças de 4 anos e meio, sendo as palavras de classe aberta mais reconhecíveis para as crianças do que as palavras de classe fechada. Porém, em contexto geral, não houve diferença significativa, e o que foi possível perceber é que a consciência se desenvolve progressivamente.

Clearly the question of what age children "have" metalinguistic knowledge is ill-formed. Rather, metalinguistic awareness is a matter of degree (Clark, 1978; Chaney, 1989) and develops progressively over the early years of life rather than emerging abruptly at 6 or 7 years of age. But that development starts much earlier than previous work suggested. Tasks which allow young children to represent both open and closed class words during normal linguistic processing and then to focus on those representations during metalinguistic reflection will succeed in pinpointing the early developmental abilities<sup>7</sup> (p. 214).

Nesse estudo, as crianças obtiveram sucesso na tarefa de diferenciar “a última palavra” de “a última

<sup>6</sup> A classe aberta possui um número ilimitado de palavras, estando constantemente sujeita a processos de modificação e surgimento de novas palavras (nomes, verbos, por exemplo). Classe fechada possui um número limitado de palavras, que dificilmente sofrem alteração ou surgimento de novas palavras (como preposições e conjunções, por exemplo).

<sup>7</sup> Tradução: Claramente, a questão sobre em que idade a criança “tem” conhecimento metalinguístico está mal formada. Mais especificamente, consciência metalinguística é uma questão de grau (Clark, 1978; Chaney, 1989) e é desenvolvida progressivamente ao longo dos anos iniciais, em vez de emergir abruptamente aos 6 ou 7 anos de idade. O desenvolvimento inicia muito antes do que alguns estudos prévios sugeriram. Tarefas que permitam às crianças mais novas representar tanto palavras de classe abertas quanto de classe fechada durante o processamento linguístico normal e que então foquem nessas representações durante a reflexão metalinguística vão ser mais bem sucedidas em mostrar as habilidades de desenvolvimento mais precoces.

coisa” que ouviram, mostrando que possuem noções diferentes para as duas tarefas.

Silva e Lorandi (2013) adaptaram e aplicaram testes baseados no estudo acima referido, com crianças de 6 a 11 anos, falantes do português brasileiro, porém não encontraram diferenças significativas entre essas duas tarefas, já que a maioria das crianças pareceu não encontrar distinção entre “palavra” e “coisa”. Todavia, os estudos de Silva e Lorandi incorporaram outros testes e, então, por meio de uma análise qualitativa, foi possível perceber as hipóteses que a criança faz sobre o conceito de palavra, muitas vezes mostrando estar em níveis diferentes para cada teste. O trabalho dessas autoras (2013) utiliza o Modelo de Redescrição Representacional, que será mais amplamente abordado no item a seguir, como análise e mostra que os resultados independem da idade, uma vez que há dados de crianças de 6 e de 10 anos no mesmo nível de representação mental. Isso nos mostra que a consciência de palavra é progressiva, e, claro, desenvolvimental, não estando necessariamente relacionada à idade.

### 3 Modelo de Redescrição Representacional

O Modelo de Redescrição Representacional foi proposto por Karmiloff-Smith (1986,1992) e rompe com a dicotomia implícito/explicito, postulando quatro níveis em que o conhecimento é redescrito na mente humana, tornando-se, paulatinamente, ao longo do desenvolvimento, acessível para acesso consciente e verbalização. A importância dessa teoria está em ser capaz de captar as sutilezas que acontecem no processo de aquisição do conhecimento, abrindo espaço para o entendimento de fenômenos que são obscurecidos quando se restringe uma análise apenas ao que pode ser expresso deliberadamente em tarefas *off-line*<sup>8</sup>.

Dessa forma, o Modelo de Redescrição Representacional postula quatro níveis em que o

<sup>8</sup> De acordo com Lorandi (2011a), tarefas *off-line* são aquelas que exigem reflexão por parte do informante. Ele precisa manter a informação na mente, trabalhar com ela, para só depois dar uma resposta.

conhecimento é redescrito: Implícito (I), Explícito 1 (E1), Explícito 2 (E2) e Explícito 3 (E3), que serão explicados a seguir, tomando-se a linguagem como exemplificação de cada um dos níveis.

No nível **Implícito (I)**, o conhecimento implícito apresenta-se como um bloco fechado, não está definido representacionalmente (KARMILOFF-SMITH, 1986). Nesse nível, a criança imita o **input** que recebe, “devolve” a informação exatamente como a recebeu, ainda não sendo capaz de manipular as estruturas da língua. Segundo Lorandi (2011a), ela é capaz de, por exemplo, produzir verbos irregulares como “fui” ou “trouxe”, ainda que não exista um conhecimento explícito sobre a utilização e o funcionamento dessas formas na língua, constituindo-se de mera repetição do *input* recebido. É um nível procedimental.

A redescrição do nível I para o **Explícito1 (E1)** envolve a sensibilização para os recursos da língua. Nesse momento passa a ocorrer uma série de análises internas ao sistema, mas que ainda não estão disponíveis para acesso consciente. Agora a criança já não está mais tão ligada ao *input* externo e é capaz de extrair informações (como as regularidades) do seu sistema. Podem surgir as *formas morfológicas variantes* (LORANDI 2007, 2011a) como “sabo” (em vez de “sei”) ou “comei” (em vez de “comi”), mostrando que começa a aparecer uma análise interna.

O acesso consciente é possibilitado quando o conhecimento é redescrito para o nível **Explícito 2 (E2)**. Nesse nível, a criança já tem consciência de determinado recurso da língua. Isso quer dizer que ela consegue ir além da sensibilidade aos recursos da língua para a consciência sobre eles, o que permite que ela os manipule intencionalmente. Segundo Lorandi e Karmiloff-Smith (2012), nesse momento a criança consegue entender, por exemplo, que entre “padaria”, “mercearia” e “bilheteria” existem traços em comum, mas não é capaz de explicar exatamente o quê.

No nível **Explícito 3 (E3)** as diferentes representações de um mesmo conhecimento conectam-se de maneira explícita e apresentam-se

disponíveis para acesso consciente e verbalização. Seu conhecimento metalinguístico atinge o ápice, sendo o produto final da redescrição representacional e permitindo que a criança não apenas entenda, mas seja capaz de falar sobre a própria língua e explicar que “padaria”, “mercearia” e “bilheteria” têm a mesma terminação porque representam um lugar (mesmo que explique isso com outras palavras).

Esses quatro níveis de redescrição representacional são um grande ganho para entender os processos que a criança (e também o adulto) passa ao estar diante de um novo conhecimento, uma vez que explica como uma análise dicotômica é falha e incompleta, uma vez que o não pronunciar-se sobre o conhecimento é entendido como sendo conhecimento implícito, e não explícito ou consciente. Além disso, vale destacar que, segundo Karmiloff-Smith, nem tudo está disponível em E3 e, por isso, torna-se fundamental o professor ter em mente outras formas de avaliar o conhecimento do seu aluno.

O MRR possui um olhar diferenciado para o desenvolvimento, à medida que entende que o acesso consciente independe da idade em muitos aspectos. Dessa maneira, uma criança pode estar no nível I para certo conhecimento e no nível E3 para outro. Da mesma forma, um mesmo conhecimento pode estar representado como I para uma criança de 10 anos e E3 para uma de 7.

Ancorado no MRR de Karmiloff-Smith (1986, 1992), este trabalho busca perceber o processo de aquisição do conhecimento de um grupo de informantes quanto à consciência de palavra em sua relação com a escrita. Para isso é preciso olhar, também, para o processo de aquisição desta habilidade.

#### 4 Aquisição da escrita

O processo de aquisição da escrita é repleto de singularidades, e muitos estudos se debruçam sob essa temática na busca de tentar perceber como a criança entende e vive essa aquisição de um novo conhecimento. Cunha e Miranda (2009) retomam as

ideias de Piaget (1972) a respeito do desenvolvimento para dizer que a criança é sujeito na construção do próprio conhecimento, interagindo com o mundo ao seu redor e organizando o pensamento.

Ferreiro e Teberosky (1999) também retomam a teoria piagetiana, ampliando seu olhar para as questões da escrita e assim mostrando que é um engano pensar que a aprendizagem depende exclusivamente do ensino específico. As autoras (*ibid*, p. 296) afirmam que,

Num terreno em que, classicamente, e apesar da variedade de enfoques, pensou-se sempre que não podia haver aprendizagem sem um ensino específico, e em que a contribuição do sujeito se considerava como dependente e subsidiária do método de ensino, pudemos descobrir uma linha evolutiva que passa por conflitos cognitivos semelhantes, até nos detalhes do processo, aos conflitos cognitivos constitutivos de outras noções fundamentais.

Essas considerações são importantes para que se entenda que a criança não é uma tábula rasa em que apenas se depositam conhecimentos, mas sim protagonista e sujeito do próprio aprendizado, constituindo a aquisição da escrita como um processo reflexivo, intrincado de erros e acertos. Os erros, de forma especial, nos dão pistas para entender como esse novo mundo começa a se organizar na mente infantil, mostrando as hipóteses que a criança vai elaborando a respeito da escrita, dos limites entre palavras e da relação fonema/grafema.

Abaurre, Fiad e Mayrink-Sabinson (1997, p.16-17<sup>9</sup> *apud* PEREIRA, 2011) chamam a atenção para este olhar que é preciso ter sob os “erros”:

[...] durante um longo período, os estudos e práticas pedagógicas ignoraram o fato de que os ‘erros’ cometidos pelos aprendizes de escrita/leitura eram, na verdade, 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.

Neste trabalho serão abordados dados de segmentações incomuns na escrita, que revelam as

<sup>9</sup> ABAURRE, M. B. M.; FIAD, R. S. e MAYRINK-SABINSON, M. L. (Orgs.) **Cenas de aquisição da escrita**. Campinas, SP: Associação de Leitura do Brasil (ALB): Mercado das Letras, 1997. 204p.

hipóteses que a criança constrói acerca da própria língua. Essas segmentações são divididas em: hipossegmentação (ausência indevida de espaço entre palavras) e hipersegmentação (presença indevida de espaços entre uma palavra), que serão mais amplamente abordadas na seção a seguir.

## 5 Hipo e Hipersegmentação

Dados como hiper e hipossegmentações mostram como a criança percebe não apenas o *continuum* de fala, mas também a gramática da língua.

Estudos como o de Cunha e Miranda (2009) olham para os dados de segmentação por meio da fonologia prosódica, e várias questões importantes são levantadas. Primeiramente, as autoras puderam perceber que os dados de hipossegmentação são mais comuns, especialmente aqueles em que ocorre a junção de uma palavra gramatical a uma palavra fonológica (como *derepente*).

Esses achados confirmam o que diz Ferreiro e Teberosky (1999) sobre a dificuldade da criança em reconhecer, na fase de aquisição da escrita, sequências de uma ou duas letras como palavra (p. 143-144).

Há ainda dados de hipossegmentações, que formam frases fonológicas ou frases entonacionais (como em: *miaroupa*), dados de junção de palavra fonológica e palavra gramatical (ex: *pegala*) e de duas palavras gramaticais (ex: *oque*). As autoras concluíram que são os constituintes mais altos da hierarquia prosódica que motivam esse processo de junção.

Os dados de hipersegmentação são mais incomuns e caracterizam-se, principalmente, pela separação de uma palavra em duas, em que uma é gramatical e a outra fonológica (ex: *da quela*), ou a separação em duas palavras fonológicas (*chapeu sinho = chapeuzinho*). Percebe-se, também, uma tendência de manter o pé troqueu<sup>10</sup>, padrão do português brasileiro.

<sup>10</sup> “Troqueu: pé métrico que consiste de uma sílaba acentuada seguida de uma sílaba não acentuada que pode ser representada em grade como (\* .) e, em árvore, como (s w)”,

Tenani (2010, p. 260) buscou investigar até que ponto as segmentações não convencionais representam também decisões sobre a grafia das palavras, estabelecendo relações entre práticas sociais orais e práticas sociais letradas. Em seu estudo, a autora encontrou dados de escrita que revelam uma reflexão sobre a língua por parte das crianças, mostrando dados ora ligados à oralidade ora à escrita enquanto norma ortográfica. Foi possível encontrar dados de flutuação na escrita de uma mesma palavra, mostrando as hipóteses que a criança faz, tanto nos espaços quanto em sua grafia (ex: “*então, em tão, en tão, emtão*” (p.260)).

As conclusões às quais a autora chega é de que não é possível dizer que a escrita inicial é apenas a transcrição da fala, uma vez que ela mostra dados claros de tentativas de escrita voltadas às normas ortográficas próprias da língua. Além disso,

(...) a palavra é objeto de reflexão, cujos significados e funções não são estáveis e cuja grafia não é dada exclusivamente por uma motivação morfosintática, mas também por motivação prosódica e informação letrada, que é historicamente constituída (TENANI, 2010, p. 265).

Abaurre (2010) também mostra que a escrita infantil não é apenas uma tentativa de reprodução da fala. Em seu estudo, a autora apresenta exemplos em que as crianças se mostram muito sensíveis aos recursos da língua escrita, especialmente aquelas de classe média alta e que, supostamente, possuem um maior contato com textos escritos. Segundo ela, as hipóteses de segmentação feitas pelas crianças “revelam a representação que fazem das categorias morfológicas e semânticas da língua e das relações que se estabelecem entre os elementos destas categorias” (ABAURRE, 2010, p. 136). O essencial é perceber que a modalidade escrita diferencia-se da modalidade falada, e a criança, desde muito cedo, percebe essa diferença e constrói suas hipóteses acerca da segmentação.

(...) as crianças de um modo geral recorrem à oralidade para fazer várias hipóteses sobre a escrita, *mas usam também a escrita*,

como na palavra “bala” [‘ba.la], em que a primeira sílaba é forte e a segunda é fraca em termos de tonicidade (SILVA, 2011, p. 214).

dinamicamente, para conduzir uma análise da própria fala e elaborar propostas de representação que com o tempo se constituirão em representações canônicas da língua (ABAURRE, *ibid*, p. 140). [Grifo da autora].

A partir desses estudos é possível perceber que muitas podem ser as motivações que levam um aluno a realizações essas segmentações não-convencionais. Neste trabalho, o foco está em buscar entender qual a concepção que cada aluno tem acerca do que é uma palavra e como essa mesma concepção poderá influenciar suas escolhas na escrita.

## 6 Metodologia

Esta pesquisa foi realizada com alunos de escola pública, do 1º ao 5º ano (de 6 a 11 anos, aproximadamente), com 10 de cada série, sendo 5 do sexo masculino e 5 do feminino, todos falantes do português brasileiro e nenhum bilíngue, somando-se 50 participantes.

A coleta consistiu em 6 testes, sendo 4 de consciência de palavra e 2 de escrita. Os instrumentos de avaliação da consciência de palavra utilizados foram os testes elaborados por Silva e Lorandi (2013), que foram ampliados e reformulados para responderem melhor aos objetivos desta pesquisa.

A seguir tem-se a descrição dos testes.

### 5.1. TESTE 1 – HISTÓRIA – “QUAL FOI A ÚLTIMA PALAVRA QUE EU DISSE?”

No primeiro teste, a experimentadora conta uma história para a criança e a interrompe em momentos específicos, com o questionamento: “qual foi a última **palavra** que eu disse? ”, e a criança precisa, então, repetir o que entende como sendo a última palavra. As pausas se dão após algumas palavras de classe aberta e de classe fechada. Esta é uma tarefa *on-line*, ou seja, não exige que a criança mantenha a informação por muito tempo na mente para dar uma resposta.

Esse teste foi reformulado a partir da história criada por Silva e Lorandi (2013), porque optou-se por

incluir na história alguns casos comuns de problemas na segmentação na escrita, como “de repente” e “em cima”, estabelecendo ligação com o teste nº 5, que também possui essas sequências (e será explicado posteriormente).

### 5.2. TESTE 2 - HISTÓRIA – “QUAL FOI A ÚLTIMA COISA QUE EU DISSE?”

Esse teste tem a mesma sequência do anterior, porém, nas pausas, a experimentadora pergunta “Qual foi a última **coisa** que eu disse?” Esse segundo teste é avaliado em conjunto com o primeiro, de modo a perceber se a criança diferencia “coisa” e “palavra”, de modo especial com relação a palavras de classe fechada.

### 5.3. TESTE 3 – CONTAGEM DE PALAVRAS

Esse teste consiste em uma tarefa de contagem de palavras em uma frase. A experimentadora explica para a criança que esta será uma tarefa de contar palavras, em que a criança deverá ouvir uma frase e depois contar quantas palavras ela contém. Nesse teste algumas frases possuem numerais, para verificar se a criança é influenciada pelo numeral para realizar a contagem.

Ao todo foram 15 frases, sendo 4 com numerais, 4 com casos comuns de hipossegmentação, 4 com casos comuns de hipersegmentação, e 3 frases sem dificuldades, funcionando como distratoras.

### 5.4. TESTE 4 – RESPONDA PARA O URSINHO CHICO

O último experimento consiste em uma lista de 15 palavras (mesclando classe aberta e classe fechada) e tem como objetivo perceber o que a criança considera como palavra, a partir de seu julgamento sobre se cada item da lista é ou não uma palavra, além de sua explicação sobre o julgamento realizado. Ela é convidada a explicar para um ursinho de pelúcia (o chico) se “anel” (por exemplo) é uma palavra ou não e o porquê. Esse experimento avalia a consciência em nível E3, quando a criança, além de acessar conscientemente seu conhecimento, também

é capaz de falar sobre esse ele. Trata-se de uma tarefa *off-line*.

Nos testes 1, 3 e 4, os alunos realizaram um pré-teste, que consistiu em uma pequena amostra do que viria a seguir. Esse pré-teste visou a garantir que os informantes entendessem a tarefa à qual seriam submetidos. O teste 2, por ser igual ao teste 1 em sua estrutura, não foi precedido de um pré-teste.

#### 5.5. TESTE 5 – DITADO DE FRASES

Para verificar a escrita das crianças, dois testes foram realizados. O primeiro foi um ditado de frases. A intenção de inserir esse teste foi proporcionar a escrita dessas sequências, para verificar possíveis ocorrências de segmentações incomuns que poderiam não aparecer na escrita espontânea. Foram ao todo 17 frases, sendo incluídas palavras normalmente tidas como alvo de hipossegmentação<sup>11</sup> (7) e hipersegmentação (5), assim como frases distratoras (5).

#### 5.6. TESTE 6 – ESCRITA DE NARRATIVA

O último teste consistiu em uma escrita motivada a partir de uma tirinha da Turma da Mônica (sequência de três imagens) sem diálogos (e sem balões de diálogo), para que a criança ficasse livre para produzir uma história.

A experimentadora explicou a tarefa detalhadamente, para que a criança entendesse que deveria escrever uma história e não, por exemplo, escrever apenas balões na tirinha.

## 7 Resultados e análise dos dados

Os dados desta pesquisa serão apresentados tanto de forma quantitativa quanto de forma qualitativa, na busca por perceber se existe ou não uma relação da consciência de palavra com a escrita.

<sup>11</sup> As palavras e sequências comumente alvo de segmentações incomuns foram retiradas dos estudos de Cunha e Miranda (2007, 2009), Tenani (2012), Cardoso e Castro (2012) e Pereira (2011).

#### 6.1. TESTE 1 – QUAL FOI A ÚLTIMA PALAVRA QUE EU DISSE?

Percebeu-se que há uma progressão nas séries, uma vez que, dos 10 informantes do 5º ano, 9 responderam com a última palavra e apenas um mesclou palavras e frases. Os dois primeiros anos são os mais instáveis, e as crianças apresentaram informações bastante distintas se comparadas aos três anos posteriores. Ainda assim, houve crianças de 1º ano que responderam com a última palavra (ainda que tenham apresentado alguns equívocos, juntando vogal e palavra de conteúdo, por exemplo), e crianças do 4º ano que repetiram predominantemente com frases e não com palavras. Isso reforça um ponto importante do MRR, que afirma que os níveis de consciência independem da idade, dado esse que só é possível observar por meio de uma análise qualitativa (mostrando também aqui a relevância dos trabalhos que optam por esse tipo de análise).

#### 6.2. TESTE 2 – QUAL FOI A ÚLTIMA COISA QUE EU DISSE?

O teste 2 servia de suporte para o teste 1, buscando perceber se os sujeitos da pesquisa estabeleciam alguma diferença entre os dois testes. Apenas 26% conseguiram diferenciar “palavra” e “coisa”, repetindo a última palavra no teste 1, e frases no teste 2. O inverso também ocorreu, mostrando a concepção ainda equivocada sobre o que é palavra.

Outro ponto importante desses dois testes é que alguns informantes respondiam a última palavra de classe aberta. Assim, em uma pausa diante de “Pedro e o”, a resposta foi “Pedro”. Isso fornece indicativos de que muitas crianças não consideram como palavra os vocábulos curtos (ou não acentuados), como vogais, o que posteriormente foi confirmado nos demais testes.

#### 6.3. TESTE 3 – CONTAGEM DE PALAVRAS

No teste de contagem de palavras as crianças apresentaram quatro tipos de respostas: (i) contagem de sílabas; (ii) contagem de palavras; (iii) mescla de palavras e sílabas; e (iv) números aleatórios.



Perceberam-se grandes diferenças entre as séries. Enquanto no 1º ano houve instabilidade, com respostas muito variadas, no 5º ano 90% dos informantes contaram palavras. A maior parte dos informantes realizou a contagem em voz alta, permitindo, assim, que a experimentadora pudesse perceber e registrar as segmentações por eles realizadas. O destaque está em “de repente”, que foi contada como uma só palavra pela grande maioria dos informantes, e também para “demais”, que sofreu diversas hipersegmentações, sendo contabilizada como duas palavras (*de mais*).

Vale salientar que a contagem de sílabas e a mescla entre sílabas e palavras se deu, predominantemente, no 1º e 2º anos. Esse dado chama a atenção porque parece sugerir uma forte influência da alfabetização, uma vez que foi possível perceber, nas observações da rotina de sala de aula dos estudantes, muita ênfase ao trabalho com as sílabas, seja na didática docente ou na linguagem visual das salas de aula.

Nesse teste havia quatro perguntas com numerais, para observar se as crianças seriam influenciadas por eles ao dar sua resposta, porém nenhum informante relacionou o numeral ao número total de palavras da frase.

#### 6.4. TESTE 4 – RESPONDA PARA O CHICO

O quarto teste é o que possibilita a observação do nível Explícito 3, em que acesso consciente ao conhecimento é possibilitado, além do relato verbal sobre esse conhecimento.

Em um primeiro momento era necessário julgar o item ouvido como sendo palavra ou não, para, logo em seguida, explicar o porquê. As respostas podem ser concebidas, de modo geral, em torno de cinco eixos, que estão descritos abaixo acompanhados de um exemplo.

- 1) Julgamento correto – a criança julga corretamente, mas não consegue explicar. Ex: “Quanto” é uma palavra? *É. Por quê? Não sei dizer por quê.* (Inf. 314)
- 2) Explicação semântica (buscando uma aplicação da palavra na vida cotidiana). Ex: “Anel” é

uma palavra? *Não. Por quê? Porque a gente coloca, não é uma palavra, é uma coisa.* (Inf.112)

- 3) Explicação morfológica (buscando pensar utilização de palavras em frases ou considerando da quantidade de letras ou sílabas). Ex: “Com” é uma palavra? *Sim. Por quê? Porque a gente usa “com muito carinho”.* (Inf. 415)

- 4) Semelhança fonológica (relacionando a outras palavras, a partir da aproximação entre sílabas). Ex: “Que” é uma palavra? *Sim! Por quê? Porque parece queijo!* (Inf. 126)

- 5) Respostas que não possuem uma explicação plausível (também se enquadram aqui respostas do tipo *sim, porque todo mundo fala*). Ex: “Até” é uma palavra? *Não. Por quê? Porque faz mal à pessoa.* (Inf. 114)

Esse teste revelou uma grande confusão das crianças em julgar e explicar o que é ou não palavra. Muitas questões muito importantes foram observadas por meio desse teste, como a noção de palavra e a relação entre palavras de classe aberta e de classe fechada, as quais serão desenvolvidas na seção 6.7.

#### 6.5. TESTE 5 – DITADO DE FRASES

Percebeu-se que a maior ocorrência de segmentações incomuns se dá no 1º ano, que é o momento inicial de aquisição da escrita. Foi possível perceber que os informantes estão fazendo hipóteses, pois, embora utilizem poucos espaços, também os colocam em lugares equivocados, causando as hipersegmentações.

Há uma visível progressão na escrita com o avanço da escolaridade, e é interessante perceber como alguns equívocos estão presentes em todos os anos escolares. As hiposegmentações mais recorrentes são em “de repente” (escrito como *derepente* ou *derrepente*) e “em cima” (escrito como *encima* ou *emcima*), e, nas hipersegmentações, “embora” foi o alvo mais frequente (escrito como *em bora*).

#### 6.6 TESTE 6 – ESCRITA DE NARRATIVA

O segundo teste de escrita era de narrativa espontânea e, por isso, a quantidade de dados de

hipo e hipersegmentações foi menor, uma vez que as crianças, de uma maneira geral, escreveram textos bastante curtos, contendo em torno de 3 a 4 linhas. É importante ressaltar que os textos apresentam uma progressão em relação ao seu tamanho ao longo dos anos escolares, excetuando-se alguns poucos casos.

O 2º ano apareceu em destaque nas hipossegmentações, porém, se os textos do 1º ano fossem maiores, o resultado possivelmente seria outro. Destaca-se, portanto, a importância de haver dois testes de escrita, para que se tenham mais dados a serem analisados e com maior precisão.

### 6.7. NOÇÃO DE PALAVRA

A noção de palavra dos informantes circulou por diversas esferas: 1) palavra é aquilo que se pode explicar semanticamente; 2) palavra é tudo aquilo que se fala; 3) palavra é o que tem mais de uma sílaba; 4) palavra é o que tem mais de três letras, 5) palavra são os nomes, objetos, verbos; 6) palavra é aquilo que tem sentido sozinha, sem precisar ser colocada em frase; 7) palavra é aquilo que as pessoas falam ou escrevem com frequência; 8) palavra é o que se pode usar em frases.

Alguns dados merecem destaque, como a relação classe aberta e classe fechada. Foi possível observar que muitos informantes estabeleceram uma dicotomia entre as classes, classificando como palavra as de classe aberta e como não palavra as de classe fechada. As palavras de classe aberta foram predominantemente explicadas pelo viés semântico, como no exemplo dado anteriormente, em que um informante verbaliza que considera palavra aquilo que pode explicar pela existência no mundo real. Essa parece ser a percepção da maior parte dos informantes dessa pesquisa. Já as palavras de classe fechada, em sua maioria, não são consideradas palavras, ou, se são, muitos informantes não conseguem explicar o porquê. Para ilustrar essa ocorrência, dois informantes foram selecionados:

O informante 4120<sup>12</sup> julga classe aberta como palavras, valendo-se de explicações semânticas e também morfológicas, com a utilização da palavra em frases. As de classe fechada são consideradas não palavras, porque precisam ser usadas “em frases”, ou seja, não têm sentido sozinhas. Exemplos:

*Amigo. É palavra. Porque vem de um sentimento e quando a gente diz “amigo” é uma palavra.*

*Até. Não é palavra. Porque é usada em frases, tipo “até mais”, daí não é palavra.*

Já o informante 311 julga classe aberta como palavra e justifica a resposta pelo viés morfológico. Classe fechada julga como não palavra, afirmando que não são palavras inteiras. Exemplos:

*Caneta. É palavra. Porque não é uma frase, é só uma palavra, tipo casa, porta.*

*De. Não, porque eu acho que não é uma palavra inteira.*

Essa diferença entre classe aberta e classe fechada também foi percebida nos testes 1 e 2, quando, diante da solicitação de repetir a última palavra dita pela experimentadora, muitas crianças respondiam com a última palavra de classe aberta.

Esses dados são interessantes porque mostram a dificuldade que as crianças têm em reconhecer clíticos como palavra, o que pode ser uma das causas das segmentações não convencionais, uma vez que a maior incidência desses equívocos está nas hipossegmentações do tipo classe fechada + classe aberta (como em *acasa*).

### 6.8. NÍVEIS DE REDESCRIÇÃO MENTAL E ESCRITA

Os dados de consciência de palavra foram analisados via Modelo de Redescricao Representacional. A seguir tem-se a tabela com os resultados dos dados analisados em cada série. Ressalta-se que o enquadramento no nível E2 (em que há consciência sobre as estruturas da língua) deu-se por uma avaliação do todo, em que os

<sup>12</sup> 4120 – 4º ano, sexo feminino, 20º participantes a realizar os testes.

informantes respondem satisfatoriamente aos três primeiros testes e possuem julgamentos corretos e respostas de E2 ou E3 no teste 4. É importante salientar que muitos informantes parecem estar em transição entre os níveis, em função de apresentarem alguma inconsistência em suas respostas que, ora parecem indicar um nível, ora outro. Foram, então, classificados nas legendas: I -> E1 e E1 -> E2.

Quadro I: Quantidade de informante em cada nível de redescrição mental. (N=10)

	1º ano	2º ano	3º ano	4º ano	5º ano
I	6	4	1	--	--
I -> E1	2	2	1	--	--
E1	2	4	3	3	2
E1 -> E2	1	--	2	3	2
E2/E3	--	--	3	4	6

Fonte: As autoras

Com base nessa análise, um dos principais pontos a se destacar é o de que, embora seja possível perceber uma progressão da consciência de palavra ao longo das séries, isso não é aplicável a todos os indivíduos, uma vez que temos dados de E1 no 1º e no 5º ano, o que reforça a questão de que a idade não é estanque e os processos de consciência são muito particulares e individuais. A informante 311, por exemplo, não consegue julgar corretamente todas as questões do teste 4, mas produz uma resposta de E3 que ilustra bem esse nível de consciência com acesso à verbalização: “Chapeuzinho” é palavra? É! Por quê? *Se fosse “chapeuzinho vermelho foi levar doces para a vovó” seria uma frase, mas só chapeuzinho é uma palavra.* Não foram encontrados informantes que possam ser classificados plenamente em nível E3, considerando-se todas as palavras, mas há respostas desse nível em todas as séries.

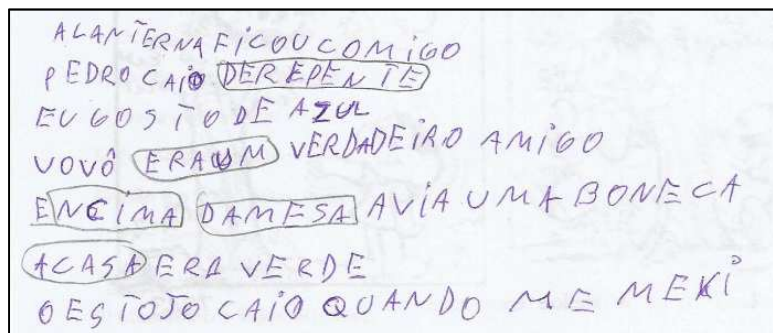
Em relação à metodologia de pesquisa vale a pena destacar o exemplo da informante 515. No primeiro teste responde, satisfatoriamente, repetindo a última palavra; no segundo teste repete a última frase, mostrando estabelecer uma diferença entre

*palavra e coisa*; no teste 3 faz a contagem correta de 15 (de 16 questões), acertando, inclusive, a contagem de *de repente*, o que pouquíssimos informantes acertaram. Seu teste de escrita não possui nenhuma segmentação não convencional. No teste 4, porém, a informante surpreende, respondendo que 13, das 15 questões, não eram palavras. Algumas de suas respostas foram: *beleza não é palavra, é um adjetivo; até não é palavra, é um cumprimento de “tchau”; chapeuzinho não é palavra, é uma personagem dos contos.* Olhando-se para todos os dados apresentados pela informante, percebe-se que há indícios de consciência nos demais testes, o que leva a pensar que a informante pode não ter entendido a tarefa. Esse exemplo reforça a importância da variedade de testes para pesquisar um mesmo fenômeno, especialmente com relação a dados de consciência linguística.

O interessante é perceber que, com a progressão escolar, algumas dessas questões parecem ser sanadas, mas muitas permanecem. Há segmentações não convencionais que acompanham os alunos em todas as séries, como é o caso de “de repente” e “embora”. Isso pode ocorrer pela ausência de uma explicitação em aula sobre o que é uma palavra, pois, como já dito anteriormente, os alunos acabam por construir suas próprias hipóteses, e as concepções equivocadas vão acompanhando-os, muitas vezes, até a vida adulta.

Verificou-se também que os dados de escrita acompanham os dados de consciência de palavra. Conforme há o avanço dos níveis de consciência, há também a redução das segmentações incomuns. De um modo geral, percebeu-se uma clara relação entre os testes de consciência de palavra e escrita. Observe-se o exemplo da informante 213. No teste 4, a informante julga que palavras de classe fechada não são palavras, justificando com “*Não, porque é só ‘s’ e ‘a’, tem que ser palavra grande*” (para “as”), ou “*não, porque não dá para escrever direito*” (para “de”, “que”, “com”). No teste de escrita, ela comete apenas equívocos de hipossegmentação, quase sempre juntando uma palavra de classe fechada com uma palavra de classe aberta;

Figura 1 – Texto informante 213



Fonte: As autoras

Esse exemplo ilustra o resultado dos testes e mostra que existe uma relação entre a consciência que se tem acerca do que é uma palavra e a escrita, confirmando a hipótese deste trabalho. Dos 50 informantes desta pesquisa, apenas um apresentou dados que não parecem estabelecer essa correspondência, uma vez que o informante apresentou um teste de escrita sem segmentações não convencionais, mas não apresentando indícios de consciência nos testes.

Os dados também mostram que os dois primeiros anos de alfabetização são os que apresentam mais variação de níveis e respostas, especialmente o 2º ano. Olhando para os dados de consciência linguística, pode-se perceber que essa é a série de maior diversidade de níveis e muitas hipóteses sobre o sistema linguístico, visíveis também nos dados de escrita. Por isso justifica-se a importância de que conceitos como o de palavra sejam trabalhados de forma explícita em sala de aula, a fim de auxiliar esses indivíduos no processo de construção do conhecimento.

É importante dizer que não foram encontradas diferenças significativas nos níveis de consciência entre os sexos.

## 10 Considerações finais

Este trabalho buscou investigar dados de consciência de palavra e dados de segmentações incomuns na escrita, de modo a estabelecer uma relação entre eles, percebendo se a consciência que

as pessoas têm acerca do que é uma palavra pode influenciar sua escrita. Os dados obtidos mostraram que há uma relação entre consciência de palavra e escrita, e esse resultado fomenta diversas reflexões nos campos teórico, metodológico e educacional.

Em relação à metodologia, destaca-se a importância de ter uma variedade de testes para a análise de fenômenos de consciência linguística, e até mesmo de escrita. Tarefas *on-line* e *off-line* mostram conhecimentos diferentes e, por isso, é também importante que essas duas dimensões sejam contempladas na metodologia de pesquisa. Os testes precisam ser capazes de obter respostas nos diferentes níveis de explicitação e de consciência, para que análise possa verificar o processo de emergência das habilidades previstas em cada nível. Ainda, o pesquisador precisa estar atento à própria prática, percebendo as possíveis falhas que estejam ocorrendo e buscando alternativas de resolução, de maneira que não haja um entendimento equivocado das tarefas por parte dos informantes, comprometendo a pesquisa.

No campo teórico e educacional, a relevância deste trabalho está em mostrar que os casos de segmentações incomuns não são um fenômeno apenas prosódico, mas também é um fenômeno ligado à consciência linguística, mais especificamente à consciência morfossintática. Há a necessidade de expandirem os trabalhos na área da consciência morfossintática para que chegue aos contextos educativos a importância de se trabalhar de forma explícita com conhecimentos de morfossintaxe, nesse caso em especial, com o conceito de palavra, para que o indivíduo seja capaz de pensar sobre a própria língua, e assim construir o próprio conhecimento com base em níveis E2 e E3, e não em nível Implícito, reproduzindo conceitos que lhe foram passados sem refletir sobre eles.

Também se ressalta a importância das análises qualitativas para perceber as singularidades presentes em cada indivíduo, especialmente ancorando-se no MRR, em que é possível perceber dados de consciência ainda que o indivíduo não seja capaz de falar sobre o próprio conhecimento. Espera-

se, também, contribuir para o surgimento de novas pesquisas sobre dados de segmentações não convencionais, fenômeno recorrente nas classes de alfabetização brasileiras, mas ainda pouco estudado na perspectiva aqui abordada.

## Referências

- ABAURRE, Maria Bernadete Marques. O que revelam os textos espontâneos sobre a representação que faz a criança do objeto escrito. In: KATO, Mary A. *A concepção da escrita pela criança*, 4ª ed. Campinas, SP: Pontes, 2010 p. 135-142.
- BISOL, Leda (Ed.). *Introdução a estudos de fonologia do português brasileiro*. 4 ed. ver. E ampl. – Porto Alegre: EDIPUCRS, 2005.
- CARDOSO, Sônia, CASTRO, Ana. Consciência de palavra em crianças de idade pré-escolar e escolar. *Textos Selecionados, XXVII Encontro Nacional da Associação Portuguesa de Linguística*, Lisboa, APL, pp. 136-152, 2012.
- CUNHA, Ana Paula Nobre; MIRANDA, Ana Ruth Moresco. A influência da hierarquia prosódica em hiposegmentações da escrita de crianças de séries iniciais. *Revista Virtual de Estudos da Linguagem – ReVEL*. Edição especial n. 1, 2007. ISSN 1678-8931 [www.revel.inf.br].
- CUNHA, Ana Paula Nobre; MIRANDA, Ana Ruth Moresco. A hipo e a hipersegmentação nos dados de aquisição de escrita: a influência da prosódia. *ALFA: Revista de Linguística*, v. 53, n. 1, 2009.
- KARMILOFF-SMITH, Annette. *Beyond modularity: a developmental perspective on cognitive science*. Cambridge (MA): MIT, 1992.
- KARMILOFF-SMITH, Annette. From meta-processes to conscious access: evidence from children's metalinguistic and repair data. *Cognition*, n. 23, p. 95-147, 1986.
- KARMILOFF-SMITH, Annette. et al. Rethinking metalinguistic awareness: representing and accessing knowledge about what counts as a word. *Cognition*, v. 58, n. 2, p. 197-219, 1996.
- LORANDI, Aline. *From sensitivity to awareness: the morphological knowledge of Brazilian children between 2 and 11 years old and the representational redescription model*. 2011. Tese (Doutorado em Letras) – Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2011a.
- LORANDI, Aline. A consciência linguística e o modelo de Redescrição Representacional: como explicar a discrepância entre os processos de consciência em diferentes microdomínios? In: FERREIRA GONÇALVES, G; BRUM DE PAULA, M. R.; KESKE-SOARES, M. *Estudos em Aquisição Fonológica*. v. 4. Pelotas: Editora e Gráfica Universitária UFPel, 2011b, p. 205 a 217.
- LORANDI, Aline; KARMILOFF-SMITH, Annette. *From sensitivity to awareness: the morphological knowledge and the representational redescription model*. *Letras de Hoje*, v. 47, p. 6-16, 2012.
- PEREIRA, Tânia Maria Augusto. A segmentação no processo de aquisição da linguagem escrita. *Veredas on line*, 2011, p. 273-88.
- PIAGET, Jean. *A epistemologia genética*. Petrópolis: Vozes, 1972.
- ROSA, Maria Carlota. *Introdução à Morfologia*. 5 ed. São Paulo: Editora Contexto, 2006.
- SILVA, Cristiana lack; LORANDI, Aline. O que é uma palavra: a consciência linguística em crianças de 6 a 11 anos falantes do português brasileiro. *Signo*, v. 38, p. 61-83, 2013.
- SILVA, Thaís Cristófar; GUIMARÃES, Daniela Oliveira; CANTONI, Maria Mendes. *Dicionário de fonética e fonologia*. São Paulo: Editora Contexto, 2011.
- TENANI, Luciani. A grafia dos erros de segmentação não convencional de palavras. *Cadernos de educação*, n. 35, 2012.
- SOUZA, Maurício de. Disponível em <<https://dreamsophies.files.wordpress.com/2010/08/tira-turma-da-monica-2441.gif>>– Acesso em: 21/06/2015.

### COMO CITAR ESSE ARTIGO

LORANDI, Aline; MARQUES, Débora Mattos. Segmentação de palavras na alfabetização de crianças: um estudo sobre consciência de palavra. *Signo*, Santa Cruz do Sul, v. 41, n. 71, set. 2016. ISSN 1982-2014. Disponível em: <<https://online.unisc.br/seer/index.php/signo/article/view/7243>>. Acesso em: \_\_\_\_\_. doi: <http://dx.doi.org/10.17058/signo.v41i71.7243>.