

A SEGMENTAÇÃO DE PALAVRAS E O PROCESSO DE AQUISIÇÃO DE LINGUAGEM

Adelaide Maria Nunes Camilo¹

RESUMO

O presente trabalho tem como tema a segmentação na fala da criança. Seu objetivo é esboçar a tendência do sujeito em fase de aquisição de linguagem em segmentar as palavras de acordo com padrões mais comuns na língua, como, por exemplo, o padrão silábico, assim como reconhecimento por parte do sujeito falante de palavras de uso frequente e cotidiano na língua. Para tanto, serão utilizados como fundamentação teórica os trabalhos de Peters (1983,1985) a respeito do processo de segmentação durante a aquisição, Figueira (1995), que trata das inovações lexicais na fala da criança, assim como seu trabalho de 1996, acerca do erro na investigação linguística. Com o intuito de melhor demonstrar os desvios de segmentação, pretende-se aproximar os dados de fala aos da escrita, pois se acredita que ao se debruçar sobre o código escrito, a criança deixa transparecer dúvidas a respeito da linguagem que até então poderiam passar despercebidas. Para essa análise, tomar-se-á por base os textos de Abaurre (1988) sobre a representação feita pela criança do objeto escrito, Ferreiro e Pontecorvo (1996) e Camilo (2011) sobre segmentações não-convencionais. Por fim, conclui-se que independente dos critérios usados para a segmentação, a palavra é o principal instrumento da criança em busca de alçar a um padrão de funcionamento do sistema linguístico.

Palavras chave: Palavra. Segmentação de palavras. Aquisição de linguagem.

1 CONSIDERAÇÕES INICIAIS

Além de uma grande preocupação dos estudos linguísticos, a aquisição da língua oral vem sendo amplamente discutida no âmbito das ciências humana. A

impressionante forma com que as crianças alcançam a fluência em sua língua materna, ainda nos primeiros anos de vida, tem impressionado e chamado atenção para a capacidade linguística do ser humano, acreditando-se que, assim como defendeu Chomsky, essa seria uma capacidade inata e universal de nossa espécie.

Ainda muito cedo, as crianças obrigatoriamente trabalham com as palavras a fim de saciar suas necessidades físicas e emocionais. É por meio delas que nomeiam as coisas, pedem alimento e se relacionam com o mundo ao seu redor, contudo, é justamente esse mesmo elemento que causa tanta dúvida e dificuldade na hora de (re)produzir enunciados. Para a criança não há clareza quanto à definição do que é *palavra*, deixando evidente ser este um ponto de instabilidade da língua. Dessa forma, neste trabalho, daremos ênfase a um processo muito comum e que explicita o trabalho do sujeito com a língua: a segmentação de palavras.

2 DEFINIÇÃO DE PALAVRA

O enunciado, tanto falado quanto escrito, pode ser segmentável em inúmeras unidades: pés rítmicos, sílabas, palavras, morfemas, sintagmas, etc. Dentre todas essas unidades a que mais se destaca é a palavra, tanto, que é a partir dela que segmentamos nossos enunciados escritos.

Dizemos que ao escrever separamos as palavras. Seria mais adequado dizer que a escrita define a unidade 'palavra', já que a escrita nos oferece a melhor definição *prática* (não teórica) de 'palavra': conjunto de letras separadas por espaços em branco. (FERREIRO; PONTECORVO, 1996, p. 40).

De acordo com a fonologia, a palavra prosódica seria um dos constituintes mais baixos da hierarquia prosódica e seria definida pela presença de apenas um elemento proeminente, não podendo, portanto, ter mais de um acento primário. Já para a morfologia, a palavra seria o elemento em que se concretizam as categorias morfológicas como gênero, número e flexão verbal, ou seja, verbos, adjetivos, substantivos, palavras funcionais e palavras derivadas.

Segundo Bisol (1996), não há "isomorfismo entre palavra fonológica e palavra morfológica, como exemplificam os verdadeiros compostos do português, que formam uma só palavra morfológica, mas duas fonológicas: $[[gwarda]ω[rowpa]ω]φ$." (BISOL, 1996, p. 233).

Tentando solucionar o problema de definição, Blanche-Benveniste (1993 apud FERREIRO; PONTECORVO, 1996, p. 39) propõe uma diferenciação entre palavra no enunciado e palavra no sistema, sendo a primeira a palavra com limites imprecisos e variáveis, de acordo com sua posição na enunciação, já a segunda corresponderia à entrada no dicionário, concebida como previamente recortada.

Pensando nessa instabilidade de conceitos de palavra, demonstraremos como ela se apresenta na fala e na escrita da criança.

3 SEGMENTAÇÃO ORAL

Segundo Peters (1983), a segmentação de palavras na fala da criança ocorre a partir da identificação de padrões sintáticos ou morfológicos na cadeia sonora provenientes da fala de um adulto “modelo”². Para ela, a criança seria capaz de identificar “chunks” (blocos) nos enunciados e, a partir desses blocos, o sujeito procuraria por pontos em que fosse possível estabelecer cortes. Tais pontos seriam localizados a partir de padrões fonológicos, como acento, ritmo e entonação que seriam reforçados pelo adulto ao se comunicar com a criança, ou então por saliências morfológicas, como fim ou começo de palavras, além de repetições de palavras inteiras ou subpartes delas.

Ainda, segundo essa autora, haveria línguas que facilitariam a identificação de fronteiras, como, por exemplo, o Hebraico, em que há a repetição do morfema de número ao final de cada palavra. Esse tipo de repetição ajudaria o sujeito a localizar os limites de palavra na medida em que ele se repetiria ao final de todas as palavras da sentença³, marcando claramente o fim dos sintagmas nominais e verbais.

A segmentação também pode ocorrer, segundo Peters (1983), a partir de similaridades fonológicas entre palavras, de maneira que seria possível o sujeito estabelecer comparações entre palavras da língua. Essas comparações poderiam ocorrer de duas formas: *i-* quando uma parte de uma palavra pode ser reconhecida como um segmento existente na língua; *ii-* quando duas palavras compartilham de semelhanças fonológicas, fazendo com que as partes semelhantes sejam identificadas como segmentos da língua. Para o português, um exemplo do primeiro tipo de comparação seriam as palavras *embora* e *embaixo*, que teriam em sua estrutura a preposição *em*.

A autora cita ainda que esse tipo de comparação pode ocorrer instantaneamente, que seria quando a criança produziria a segmentação imediatamente à escuta do enunciado adulto, ou a partir de um processo de reflexão sobre a língua, que seria quando a segmentação fosse proveniente de um enunciado cotidiano, de alta frequência. Nos casos de memorização de unidades longas, em que suas subpartes façam parte do vocabulário frequente da criança, há uma maior tendência para que ocorram erros de segmentação.

"I want you to be good" and "I want you to behave" are frequently heard, such a segmentation principle could be the basis of productions such as Rachel Scollon's "I'm going to be very very /heyv/," or Norman Gibson's "Daddy, Laura's not being /heyv/. (PETERS, 1983, p. 38).

Ainda, segundo a proposta de Peters (1983), a informação estrutural não seria imediatamente avaliada pelo falante após a segmentação de uma longa unidade em duas unidades menores pela aplicação das combinações. A princípio, o falante trabalharia com a justaposição de unidades dentro de *frames*, que seriam uma espécie de quadros de estruturas em que apareceriam as formas X + Y (ex. be + good, be + nice). Assim, primeiro o sujeito identificaria que X precisa de uma forma Y que complete seu significado, mas, sem compreender ao certo a que classe de palavras o Y faria parte. Com base no trabalho com várias unidades semelhantes, o falante identificaria os traços de classe em comum e trabalharia, portanto, com uma abstração.

Valendo-se do processo de abstração, percebemos que a criança passaria de uma identificação e extração de unidades dos blocos de enunciados produzidos pelo adulto, para um trabalho com as subunidades pertencentes a esses blocos.

Cabe aqui ressaltar que, apesar de apresentar com detalhes a proposta de Peters (1983) de como se daria o processo de segmentação na fala da criança, nos distanciamos de sua visão, por acreditar que esta pressupõe certo grau de consciência por parte do sujeito em relação ao seu trabalho epilinguístico.

4 SEGMENTAÇÃO DIVERGENTE⁴

Sendo o erro na linguística aquilo que “destoa ou é diferente de uma realização tomada como modelo ou padrão” (FIGUEIRA, 1996, p.55), ele ocupa um olhar privilegiado nos estudos linguísticos, já que é por meio dele que podemos

perceber os indícios do trabalho da criança com a língua, “um sinal eloquente de que a criança está construindo a língua como sistema de regras” (FIGUEIRA, 1996, p.69). Nesta seção, explicitaremos os desvios de segmentação apresentados pelas crianças durante seu processo de aquisição da língua.

De acordo com Peters (1985), podemos encontrar três tipos de má segmentação: *i. a undersegmentation* (ou hipossegmentação), que seria quando a criança une mais as unidades do que o previsto pelas convenções; *ii. Mis-segmentation* (ou má segmentação/segmentação confusa), que ocorre quando a criança não compreende a sequência de palavras ou unidade e faz uma segmentação não esperada, trocando elementos de lugar (ex. wizard of oz é entendido como wizarda voz); *iii. Oversegmentation* (ou hipersegmentação), acontece quando a criança tende a separar as palavras ou unidades mais do que o previsto pelas convenções.

A mesma autora dá um exemplo muito interessante de hipersegmentação na fala:

C: Steven I am /heyv/.
 S: What? You hate? What do you hate?
 C: /heyv/. I am /heyv/.
 S: You hate? You hate me? The music? What?
 C: No, I am /heyv/. /heyv/.
 S: I don't know what you are talking about.
 Silence. A bit later:
 C: I /heyv/.
 S: You hate me?
 C: (shakes her head no)
 S: Who do you hate?
 Silence. A bit later:
 C: I am behaving.
 (PETERS, 1985, p. 1049).

Neste exemplo, podemos notar que, como exposto no item acima, a criança tem problemas ao identificar a palavra *behave*, comparando-a as estruturas de *be+x*, que aparecem com frequência na fala dos pais. Segundo Peters (1985), ao não ser compreendida, a fala da criança passa por um processo de reanálise, em que ela faz uma pausa e reestrutura seu enunciado até conseguir ser bem sucedida em sua comunicação.

O mesmo tipo de segmentação divergente é encontrado em um dado do português publicado por Figueira (1995):

(diálogo entre um menino e sua tia, no qual o primeiro pede notícias sobre Teobaldo, o gato com que brincara nas férias, na casa dos avós)

M. Tia, e o Baldo?

A. O quê? Quem?

M. O Baldo.

A. O quê?

M. O Baldo. O meu Baldo, o teu Baldo.

(dado de uma criança de três anos e meio, recolhido por M. Francisca Lier de Vitto)

(FIGUEIRA, 1995, p. 70).

Esta segunda ocorrência pode ser explicada tomando-se por base o uso das combinações propostas por Peters (1983), uma vez que a criança reconhece no nome próprio do gato (Teobaldo) uma unidade existente na língua (teu), o que fica explícito em sua fala, quando, para se fazer entender, ela explica: *O Baldo. O meu Baldo, o teu Baldo.*

Para exemplos de hiposegmentação na fala, também recorremos ao trabalho de Figueira (1995, p. 71), de onde retiramos as ocorrências de “*quantoisso* (enquanto isso), *amoto* (a moto), *o isque* (o uisque)”. Nestes três casos, a criança não reconhece na cadeia sonora que cada uma dessas ocorrências é formada por duas palavras, provavelmente, isso acontece em função da ausência de acentos nos elementos clíticos⁵ em *amoto* e *oisque*, em relação ao *quantoisso*, podemos dizer que a criança reconheceu a fronteira de frase fonológica e não de palavras⁶.

Em relação à segmentação confusa, Peters (1983, p. 53) traz os seguintes exemplos:

guinea pig	beginning pig
shoo fly	shoe fly
Don Quixote	donkey goat
lead us not into temptation	lead us not into Penn Station
the dawn's early light	the donzerly light
life is but a dream life	is butter and cream
it costs an arm and a leg	it costs a nominal egg
to all intents and purposes	to all intensive purposes
qu'est-ce qui se passe?	qu'est-ce qui space

5 SEGMENTAÇÃO NA ESCRITA

Considerando que a segmentação de palavras se apresenta como um ponto instável na língua, apresentaremos como os desvios de segmentação aparecem na escrita, com o intuito de mostrar que, apesar de aparecerem cedo na fala das crianças, eles não são resolvidos antes de seu ingresso na ambiente escolar e

podem, muitas vezes, permanecer por muitos anos, sendo encontrados, até mesmo, em séries avançadas.

Segundo Abaurre (1988), a hipótese de segmentação feita pela criança revela as representações que faz das categorias morfológicas e semânticas da língua e das relações que se estabelecem entre os elementos dessas categorias. Dessa maneira, desde cedo é possível encontrar a incorporação de aspectos convencionais em sua escrita, o que mostra seu contato com práticas letradas.

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).

Em relação à oralidade, a prosódia estaria presente na produção infantil quando, por exemplo, partes separadas pelos espaços em branco parecem corresponder aos constituintes prosódicos propostos por Nespor e Vogel (1986), ou seja, a criança separa a palavra em unidades prosódicas reconhecendo possíveis proeminências da palavra, como é o caso de “sobre mesa” (para “sobremesa”), em que o aluno atribui à palavra dois acentos primários, a partir de sua percepção do acento secundário em “so”. Quanto às marcas específicas de escritas, podemos considerar a colocação de pronomes em ênclise, a manutenção de normas ortográficas como colocação de *ss* e *rr*, além da identificação de partes das palavras como sendo artigos, preposições, etc.

As crianças de um modo geral recorrem à oralidade para fazer várias hipóteses sobre a escrita, mas usam também a escrita, dinamicamente, para conduzir uma análise da própria fala e elaborar propostas de representação que com o tempo se constituirão em representações canônicas. (ABAURRE, 1988, p. 140)

Segundo Ferreiro e Pontecorvo (1996), a segmentação lexical é difícil em idades pré-escolares e se torna sistemática aos 7/8 anos, quando a criança começa a ser alfabetizada. Além disso, as autoras constataam uma maior facilidade nos alunos para distinguir adjetivos, substantivos e verbos e uma grande dificuldade para aceitar que artigos, conjunções, preposições e outros elementos de ligação sejam palavras. Talvez seja esse um dos motivos pelos quais há um grande predomínio de hiposegmentação de clíticos como *a*, *o*, *la*, *lo*, *de*:

a tendência à hipossegmentação parece dominar sobre a tendência à hipersegmentação, qualquer que seja a língua, a tradição escolar e o tipo de script [...] de certa maneira, essa tendência expressa o 'mal-estar' que as crianças sentem diante da segmentação convencional, já que esta apresenta mais segmentações do que as que parecem 'normais' a estes escritores principiantes. (FERREIRO; PONTECORVO, 1996, p. 49-51).

Esse resultado pode ser justificado a partir do já observado por autores como Cunha e Miranda (2007), em que explicam que a criança em fase inicial de alfabetização tende a apresentar um maior número de hipossegmentações, pois ainda percebe a escrita como uma transposição da fala, modalidade em que não há uma distinção notadamente marcada entre os elementos que as constituem, de maneira que a compreendem como uma “cadeia contínua” de constituintes prosódicos. A isso se soma a dificuldade por eles apresentada no reconhecimento de sequências de uma ou duas letras como palavras – o que podemos entender como o caso dos clíticos –, de forma que apresentam uma tendência a aglutinar esses elementos à sequência por eles reconhecida como palavra que, em geral, se trata de uma palavra fonológica.

Ao observar as segmentações nos textos produzidos por crianças é comum nos depararmos com as mesmas sequências, ora escrita segundo as convenções, ora fugindo delas. A respeito dessa instabilidade as autoras afirmam:

A grande instabilidade segmentar que apresentam estes textos pode ser vista em sentido negativo (ausência de conhecimento das segmentações convencionais), mas também em sentido positivo (a produção escrita como espaço de prova de segmentações possíveis, segundo uma multiplicidade de critérios ainda não hierarquizados). (FERREIRO; PONTECORVO, 1996, p. 53).

Como exemplo dessa “instabilidade segmentar”, trazemos os dados apresentados por Camilo (2011):

Figura 1 – Produção textual

Amor sem Limites

Em uma fazenda do Rio de Janeiro morava Mayara uma moça muito bonita dos olhos claros e ela tinha um namorado chamado Maycom ele era um traficante ali da fazenda, ele lutava muito nela mas mesmo assim ela era apaixonada por ele e gostava muito dele não pensava em outros homens todos amigos e parentes dela diziam para ela largar ele mas ela não escutava.

Um dia ela decidiu passear pela praia e viu um rapaz se afogando e ela pulou na água e decidiu ~~salvá-lo~~ ~~salv-lo~~ ajudar pulou na água já que ela sabia nadar muito bem e salvou o rapaz e já no primeiro instante que o viu se apaixonou por ele e ele também por ela, os dois começaram a namorar escondidos e ela a traiu muito bem o contrário de Maycom. Foi quando Maycom descobriu que ela estava a traindo quando saiu correndo atrás dela com uma faca ela ligou para o seu recente namorado que conheceu na praia e disse para ajudar ela que ela estava escondida em uma caverna em um penhasco ~~de~~ o namorado dela chegou e Maycom saiu quem era e correu atrás dele quando chegou com no fim de penhasco ela pegou uma dele e atropelou Maycom e ele morreu, e então ela pode viver em paz com seu namorado, os dois caíram e tiveram dois filhos.

Fonte: CAMILO, A. M. N. Rasuras em fronteiras de palavras: análise de dados do segundo ciclo do ensino fundamental. Relatório científico final apresentado à FAPESP (2011).

Ao longo do texto, o aluno praticamente não apresenta nenhuma rasura ou segmentação não-convencional, tendo, inclusive, poucos desvios ortográficos. Porém, no ápice da narrativa, no nó narrativo do enredo, o escrevente usa a construção “salvá-lo”. Ao escrever pela primeira vez, produz uma forma hipossegmentada e, logo em seguida, rasura, para posteriormente escrever novamente, porém de acordo com as convenções ortográficas. Destaca-se, entretanto, que provavelmente em virtude de sua incerteza a respeito da escrita institucionalizada que almeja alcançar, o escreve rasura novamente e opta por outra palavra para completar o sentido da sentença.

Essa ocorrência chama a atenção em virtude de explicitar a rasura em segmentação não-convencional como um conflito mostrado, ou seja, é uma marca deixada no texto a respeito das incertezas dos alunos a respeito da representação da palavra escrita. Dessa forma, a rasura não é interpretada apenas como um “erro”, mas como marca essencial da subjetividade do escrevente, sendo, portanto, um elemento “revelador das relações do sujeito com a linguagem” (CALIL; FELIPETO; LOPES; 2006, p. 7).

6 CONSIDERAÇÕES FINAIS

Ao longo deste trabalho, foram apresentadas algumas características das segmentações, tanto na fala quanto na escrita, de crianças em fase de aquisição de linguagem. Com isso, pretendemos mostrar que por mais que nos aprofundemos nos estudos, poderemos apresentar apenas algumas faces desse tão complexo processo de trabalho com a língua.

Com base nos estudos aqui elencados, podemos concluir que independente dos critérios usados para a segmentação, a criança, em busca de alcançar a um padrão de funcionamento do sistema linguístico, tem como seu principal instrumento a palavra, “a mais evasiva de todas as unidades que se encontram na segmentação dos enunciados” (MALKIEL, 1970 apud FERREIRO; PONTECORVO, 1996, p.39) e é por intermédio desse instrumento que ela testa e se coloca na língua, não como um sujeito passivo, mas em posição ativa frente a tudo que a língua, como um sistema complexo e de caráter sociocultural, tem a lhe oferecer.

NOTAS

¹ Mestranda em Aquisição de Linguagem do Programa de Linguística do Instituto de Estudos da Linguagem, UNICAMP.

² Cabe aqui chamar a atenção para a importância da fala do adulto como input para a fala da criança, como já foi destacado por DE LEMOS (1982), a respeito da 1ª posição, em que a criança utiliza a fala do adulto como modelo de especulação e criação de hipóteses para o funcionamento da língua.

³ ha-yelad-im ha-gdol-im me-dabr-im
the boy m.pl. the big m.pl. speak m.pl.
The big boys speak
(Berman, 1981, p. 26 apud PETERS, 1983, p. 37)

⁴ Termo cunhado por Figueira (1995) a respeito das inovações lexicais na fala de crianças.

⁵ Segundo, Bisol (2005), são elementos caracterizados por não terem acento primário e por poderem pertencer a diferentes classes gramaticais e, geralmente, por se unirem a uma palavra fonológica, formando, assim, um grupo clítico.

⁶ Apesar de os exemplos serem extraídos de Figueira (1995), nossa análise diverge em alguns pontos.

WORD SEGMENTATION AND THE PROCESS OF LANGUAGE ACQUISITION

ABSTRACT

The present paper is about the segmentation in the child's speech and written. It aims to outline the trend of the subject, in process of acquiring language, segment words according to the most common patterns in the language, such as the syllabic pattern, as well as his recognition of words in everyday language. For this, we used as theoretical background texts of Peters (1983,1985) concerning the segmentation in the process of language acquisition, Figueira (1995), which deals with lexical innovations in the child's speech, Figueira (1996), about error in linguistic research. In order to better demonstrate the segmentation's misapplication is intended to approximate the speech data to the writing, because it is believed that approaching the writing, the child reveals doubts about language which hitherto could go unnoticed. For this analysis, we use texts of Abaurre (1988) about the representation of the written, and also Ferreiro and Pontecorvo (1996) and Camilo (2011) about unconventional segmentations. Finally, we conclude that regardless of the criteria used for segmentation, the word is the main instrument of the child in pursuit of elevating to the language.

Keywords: Word. Word segmentation. Language acquisition.

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*. 2. ed. Campinas: Pontes Editores, 1988. v. 1. p. 135-142.

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

_____. Mattoso Câmara Jr. e a palavra prosódica. *DELTA*, São Paulo, v. 20, n. especial, p. 59-70, 2005.

CALIL; FELIPETO, C.; LOPES, A. de A. O sujeito inexistente: reflexões sobre o caráter da consciência fonológica a partir do “Relatório final de grupo de trabalho alfabetização infantil – os novos caminhos”. *Educação e Pesquisa*, São Paulo, v. 32, n. 1, p. 137-155, 2006.

CAMILO, A. M. N. *Rasuras em fronteiras de palavras: análise de dados do segundo ciclo do ensino fundamental*. Relatório científico final apresentado à FAPESP, 2011.

CUNHA & MIRANDA, A. R. M. A influência da hierarquia prosódica em hipossegmentações da escrita de crianças de séries iniciais. *Revista Virtual de Estudos da Linguagem – ReVEL*. Edição especial, n. 1, 2007.

DE LEMOS, C. G. Sobre a aquisição da linguagem e seu dilema (pecado) original. *Boletim da Abralín*, n. 3, p. 97-126, 1982.

FERREIRO, E.; PONTECORVO, C. Os limites entre as palavras: a segmentação em palavras gráficas. In: FERREIRO, E. et al. *Chapeuzinho Vermelho aprende a escrever*. São Paulo: Ática, 1996. p. 38-66.

FIGUEIRA, R. A. A palavra divergente: previsibilidade e imprevisibilidade nas inovações lexicais da fala de duas crianças. *Trabalhos de Linguística Aplicada*, Campinas, v. 26, p. 49-80, 1995.

_____. O erro como dado de eleição nos estudos da linguagem. In: CASTRO, M. F. P. *O método e o dado no estudo da linguagem*. Campinas: Editora da UNICAMP, 1996.

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

PETERS, A. *The units of language acquisition*. New York: Cambridge University Press, 1983.

_____. A Language segmentation: operating principles for the perception and analysis of language. In: SLOBIN. *The crosslinguistics study of language acquisition*. London: L.F.A. Publishers, 1985. v. 2.