



UNIVERSIDADE ESTADUAL DE CAMPINAS  
SISTEMA DE BIBLIOTECAS DA UNICAMP  
REPOSITÓRIO DA PRODUÇÃO CIENTÍFICA E INTELLECTUAL DA UNICAMP

**Versão do arquivo anexado / Version of attached file:**

Versão do Editor / Published Version

**Mais informações no site da editora / Further information on publisher's website:**

<https://revistas.gel.org.br/rg/article/view/3134>

**DOI: 10.21165/gel.v18i2.3134**

**Direitos autorais / Publisher's copyright statement:**

©2021 by Grupo de Estudos Linguísticos do Estado de São Paulo. All rights reserved.

DIRETORIA DE TRATAMENTO DA INFORMAÇÃO

Cidade Universitária Zeferino Vaz Barão Geraldo

CEP 13083-970 – Campinas SP

Fone: (19) 3521-6493

<http://www.repositorio.unicamp.br>

# AS PALAVRAS QUE OS BEBÊS NÃO DIZEM: REVISITANDO O PROBLEMA DA NOÇÃO DE “PALAVRA” À LUZ DA AQUISIÇÃO DA LINGUAGEM

Mauricio RESENDE<sup>1</sup>

<http://dx.doi.org/10.21165/gel.v18i2.3134>

**Resumo:** A caracterização da unidade “palavra” tem sido um dos problemas centrais das teorias em morfologia, e a falta de critérios consistentes para defini-la tem levado inclusive certos modelos teóricos (os modelos construcionistas) a abandonarem o estatuto especial dessa unidade e caracterizá-la como fruto de operações sintáticas – as mesmas empregadas para gerar sintagmas/sentenças. No bojo dessas considerações, este artigo tem por objetivo revisitar esse problema à luz de uma outra perspectiva: a da aquisição da linguagem. Mais especificamente, o presente trabalho, em um estudo que analisa *corpora* de fala de crianças adquirindo o PB, revisita os principais dilemas que a noção de “palavra” coloca para as teorias linguísticas bem como para as metodologias em aquisição da linguagem e discute o estatuto das unidades mínimas de produção da aquisição. Ao final, este estudo mostra que as palavras não são unidades linguísticas mínimas em nenhum nível e que, assim como a gramática do adulto, a criança se apoia em unidades menores; portanto, a intuição sobre “palavra” deve ser produto de algum outro fator.

**Palavras-chave:** Palavra. Aquisição da Linguagem. Morfologia. Análise de *corpora*.

---

<sup>1</sup> Universidade de São Paulo (USP), São Paulo, São Paulo, Brasil; [mauri\\_cio\\_resende@hotmail.com](mailto:mauri_cio_resende@hotmail.com); <https://orcid.org/0000-0001-7487-5043>

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

## *BABIES’ NON-SAID WORDS: A REVIEW OF THE “WORDHOOD” ISSUE IN THE LIGHT OF LANGUAGE ACQUISITION*

**Abstract:** The proper characterization of wordhood has been an important topic in morphological theories, and the lack of consistent criteria for defining it has led certain frameworks (the constructionist ones) to give up on the special status of “word” and to characterize it as an output of syntactic operations – the ones employed to generate phrases and sentences. That been said, this paper aims at reviewing this problem in the light of another perspective, namely, language acquisition. Specifically, the present work, built on a corpus analysis of children acquiring Brazilian Portuguese, reviews the main issues emerging from the wordhood targeting both linguistic theory and language acquisition methodology, and also discusses the status of the minimal unities of infant utterances. In short, this case study shows that words are not the minimal linguistic unities at any level at all and as well as target-grammar (i.e. adult-grammar), children lean on smaller unities. Therefore, the intuition behind wordhood must come from somewhere else.

**Keywords:** Wordhood. Language Acquisition. Morphology. Corpus Analysis.

### **Introdução**

Ao serem interpelados com a pergunta “o que se sabe quando se sabe uma língua”, muitos falantes chegariam, de uma forma ou de outra, à observação intuitiva de que “saber uma língua” significa conhecer as palavras dessa língua, mas também conhecer regras de combinação dessas palavras de modo a formar frases. Superficialmente, seria possível dizer que o conhecimento linguístico é constituído de *léxico* (= palavras) e *sintaxe* (regras de combinação). Nessa visão superficial (e, em um certo sentido, senso-comum), o conhecimento fonológico (= pronúncia) e o conhecimento semântico (= significado) seriam redutíveis àquilo que se sabe sobre as palavras (ou seja, sobre o léxico), pois a pronúncia e o significado das sentenças seriam derivados das palavras, quando combinadas nas frases.

Sob a mesma perspectiva, a noção de que as “palavras” são unidades privilegiadas do conhecimento linguístico aparece na observação de que “qualquer falante nativo é capaz de reconhecer uma palavra de sua língua”, isto é, é capaz de, intuitivamente, responder se uma dada sequência de sons poderia ou não ser uma palavra – ancorando-se na mesma visão de que um falante nativo sabe identificar uma sentença como sendo bem ou malformada. Como base nessas considerações, é possível notar que a noção daquilo a que “palavra” deveria corresponder parece muito intuitiva não apenas aos falantes, mas principalmente a muitos linguistas na hora de formular suas perguntas de pesquisa.

Da perspectiva do falante, como observa Petter (2008), há vários ditos e expressões vernáculas que refletem a intuição de que, para o leigo, as palavras podem ser identificadas como uma unidade formal, como pode ser visto em: *para bom entendedor meia palavra basta, palavra de honra, últimas palavras, medir as palavras, pedir a palavra, sem meias palavras, palavra de mestre/escoteiro, palavras ao vento, homem de/sem palavra, em outras/poucas palavras, belas/sábias palavras* e mesmo *palavras-chave* etc.

A respeito dessa “intuição” do falante sobre a unidade “palavra” (qualquer que ela seja), Borges Neto (2020, p. 24, grifo do autor) argumenta que “a questão que está por detrás do uso extensivo da noção de ‘palavra’ se deve à *naturalização das teorias tradicionais*, cuja ontologia é, de forma geral, assumida acriticamente”. Mais especificamente, o autor defende que o reconhecimento da unidade formal básica “palavra” está enraizado na visão greco-latina da tradição gramatical, que é assimilada via processo de escolarização. Para Borges Neto, é a naturalização das noções (e da visão) teóricas da Gramática Tradicional que constroem a ideia de que “palavra” seja uma noção intuitiva.

Seja como for, um dos principais objetivos da Linguística (Formal) é tanto explicitar a intuição que o falante nativo tem quanto definir e formalizar o conhecimento linguístico. Nesses termos, cabe ao linguista se distanciar da sua “intuição” de falante (a qual pode, de fato, ser a naturalização da *teoria* linguística tradicional) e buscar por critérios claros que explicitem e/ou desconstruam o que “intuitivamente” se poderia considerar uma unidade primitiva e/ou mínima do conhecimento linguístico. Assim, como adverte Petter (2008, p. 59), diferentemente do que é para um falante leigo, “para o linguista, não é tão simples articular a entidade que [a palavra] representa”.

No bojo dessas considerações, o problema com a noção de “palavra” não está na identificação dessa unidade por parte do falante nativo, cujo reconhecimento se materializa em um certo número de ditos populares e/ou em respostas “intuitivas” a respeito daquilo de que é feito o conhecimento linguístico, mas sim, na *naturalização* dessa unidade por parte dos linguistas, de tal modo que suas perguntas e sua metodologia de pesquisa assumam “acriticamente” – para empregar o termo de Borges Neto (2020) – a ideia de que a palavra é uma unidade privilegiada e, inclusive, primitiva – cf. Lazzarini-Cyrino e Mattos (2020).

Nesse mesmo cenário, embora a caracterização mais elementar do campo de estudos da “Morfologia” – o estudo da estrutura interna das palavras, suas diferentes formas e seus processos de formação – seja trivialmente recorrente, é bem verdade que não há, entre os morfólogos, um consenso nem sobre aquilo que poderia ser uma palavra em uma dada língua (coerentemente com os vários níveis de análise), nem sobre aquilo

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

que deveria ser uma palavra do ponto de vista translinguístico; uma noção “universal” de palavra – cf. Marantz (1997) e Haspelmath (2011). A respeito disso, Villalva (2012, p. 126) inclusive observa que “não há nada no termo ‘morfologia’ que leve a pensar em palavras”.

Do ponto de vista teórico, a “palavra” recebeu dois papéis que acabaram por instituir diferentes visões de morfologia: modelos de MORFOLOGIA BASEADA EM PALAVRA e modelos de MORFOLOGIA BASEADA EM MORFEMA. Ortogonalmente, outra questão teórica versa sobre o componente da gramática em que se formam as palavras; na GRAMÁTICA GERATIVO-TRANSFORMACIONAL (GGT), há linguistas que defendem que elas são formadas no léxico (LEXICALISTAS), e há aqueles que defendem que estas são criadas na sintaxe por meio dos mesmos mecanismos usados para construir sintagmas e sentenças (CONSTRUCIONISTAS)<sup>2</sup>.

Não é objetivo deste artigo revisitar toda essa discussão nem mesmo reproduzir as análises e descrições empíricas que dão suporte a uma ou a outra visão – para tanto, cf. Bruening (2018). Alternativamente, o presente estudo visa atualizar essa discussão à luz de uma outra perspectiva, a saber, da aquisição da linguagem<sup>3</sup>. Especificamente, este trabalho analisa um *corpus* constituído por gravações/transcrições longitudinais de cinco crianças adquirindo o português brasileiro (PB), em um período entre 12 e 27 meses, e discute se, dessa perspectiva, há evidências de que a “palavra” é uma unidade relevante para a aquisição, ou seja, para o conhecimento linguístico da criança.

Como argumentado a seguir, essa unidade não parece ser privilegiada pelas crianças adquirindo a língua, em nenhum nível, o que dá suporte à ideia de Borges Neto (2020) de que talvez a noção “intuitiva” de palavra venha da naturalização da visão tradicional via escolarização, uma vez que, estando em processo de aquisição da linguagem, os bebês não podem se apoiar nessa “intuição”. Em síntese, o presente artigo mostra que há motivação empírica, teórica e metodológica para defender que “palavra” não é uma unidade relevante para a aquisição e que, das perspectivas semântica, morfológica e fonológica, as crianças adquirem a linguagem se apoiando em morfemas, não em palavras – cf. Resende (2021).

---

2 É preciso deixar claro, todavia, que o panorama teórico apresentado se refere a trabalhos e discussões de orientação gerativista e formalista. À luz de outras teorias, como a funcional-cognitivista e/ou a discursiva, diferentes questões e diferentes respostas emergiriam e, por essa razão, não é possível contemplar ou discutir toda a bibliografia relevante sobre o tema. O mesmo comentário se aplica à Lexicografia, uma área relevante e importante para a discussão proposta, mas que, por motivos de espaço e escopo, não foi abordada.

3 Naturalmente, existem diferentes “perspectivas” (visões e/ou teorias) de/sobre aquisição da linguagem. Esse termo está sendo empregado aqui com o sentido de que a discussão deste artigo se baseia em dados oriundos da fala infantil, em processo de aquisição de língua materna.

Para tanto, este trabalho está dividido da seguinte maneira: na primeira seção, são feitas algumas considerações sobre o estatuto e as definições de “palavra” na literatura e são levantados problemas a respeito da falta de isomorfia entre as noções de “palavra” nos diferentes níveis de análise. Na segunda seção, é discutida a contraparte teórica e mitológica do problema que emerge da assunção de “palavra” na aquisição da linguagem e, em seguida, as subseções apresentam e discutem os dados de fala infantil que dão suporte à ideia de que a palavra não é uma unidade privilegiada ou característica do processo de aquisição. Por fim, na terceira seção, tecem-se considerações acerca do que, na ausência de “palavra” como primitivo, distingue aquisição de morfologia de aquisição de sintaxe.

### **A unidade “palavra” na gramática do adulto**

Embora frequentemente se atribua à Morfologia o estatuto de “campo de estudos no domínio da palavra”, bastam apenas alguns conjuntos de dados para se chegar à conclusão de que tal definição de “Morfologia” não implica o reconhecimento de fronteiras claras para esse domínio, já que o reconhecimento daquilo que é uma “palavra” não é claro. Ainda assim, é comum que se tenha uma ideia de qual seja o domínio da Morfologia, pois existe alguma intuição sobre o que deveria ser uma palavra. Seja como for, à luz de uma análise mais detalhista, pode-se constatar que existe um grande problema empírico que subjaz à noção de “palavra”, que é a falta de critérios precisos para sua identificação.

Esse problema aparece detalhado a seguir. Ao final, a ideia é mostrar que a falta de critérios claros para explicitar a suposta intuição sobre “palavra” é capaz de influenciar – ou enviesar – o tratamento de certos fenômenos linguísticos como, por exemplo, o estatuto de “palavra” na aquisição da linguagem. Em última instância, se não há critérios claros para definir ou restringir, com segurança, a que corresponde a unidade “palavra”, não há como identificar, com segurança, o que diferencia aquisição de sintaxe de aquisição de léxico, se a ideia corrente é a de que a criança dá pistas de que deu início à aquisição de sintaxe quando profere enunciados contendo mais de uma palavra.

O problema empírico versa sobre qual é exatamente o nível de análise em que os falantes – e os linguistas – se apoiam para classificar uma dada sequência de sons como sendo ou não uma palavra. A pergunta formulada em termos de “qual é *exatamente*” já é suficiente para levantar suspeitas de que a intuição sobre “palavra” se apoia em diferentes propriedades, isto é, em características relevantes para mais de um nível de análise. Ainda assim, a resposta (ou a expectativa) mais óbvia seria a de que essa unidade é definida no domínio da morfologia. Contudo, diferentes visões de *morfologia* caracterizam a “palavra” de formas bastante diferentes.

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

Do ponto de vista fonológico, as palavras (fonológicas) têm sido identificadas, em termos estruturalistas, a partir de um GRUPO DE FORÇA, isto é, de um acento primário ou secundário em uma sequência de sons. Assim, “o acento (de grau 2 ou 3, conforme sua posição no grupo de força) é a marca nítida do vocábulo fonológico” (CAMARA JR, 1970, p. 97). Embora essa definição (ainda) seja relativamente bem aceita entre os fonólogos, a falta de isomorfia entre um VOCÁBULO FONOLÓGICO e uma PALAVRA MORFOLÓGICA – um “VOCÁBULO FORMAL”, nos termos de Camara Jr. (1970) – gera problemas para uma teoria de morfologia, mas sobretudo, cria discrepâncias a respeito da intuição sobre “palavra”.

Como o próprio Camara Jr. observa, certas unidades, às quais se costuma chamar de “palavra”, não têm o estatuto de um vocábulo fonológico, como é o caso de formas átonas (isto é, sem acentos 2 ou 3) que se cliticizam à palavra fonológica (ou seja, ao grupo de força), visto em artigos como *o* e *a*; preposições como *de* e *a*, pronomes como *me* e *te* etc. A princípio, nenhum morfólogo (ou sintaticista) estaria disposto a dizer que essas unidades átonas não são palavras, uma vez que existem “classes de palavras”, e algumas delas são prototipicamente caracterizadas por elementos dessa natureza, isto é, átonos. Com isso, não se quer dizer que as classes de palavras (pelo menos, da forma como aparecem na Gramática Tradicional) são critérios sólidos de análise, mas o ponto é que a ideia daquilo que seja uma palavra, na visão da fonologia, não condiz com a “intuição” sobre essa unidade.

Uma maneira de renunciar ao problema da falta de isomorfia entre esses dois níveis seria, por exemplo, prescindir do nível fonológico para depreender essa unidade. Contudo, observações empíricas atestam que essa manobra não está livre de consequências, se não se está disposto a abrir mão da intuição sobre o que é uma palavra. O primeiro problema vem das interjeições, especificamente de formas como *putz*, *eita*, *ixi*, *aff*, *hummm*, *eca*, *ai*, *ui*, *opa* e outras. Apesar de constituírem uma das “classes de palavras”, muitos linguistas defendem que estas não são verdadeiras unidades linguísticas. Sem entrar nos detalhes dessa questão, assumo-se<sup>4</sup> que as interjeições fazem parte do conhecimento linguístico do falante.

Como aparece discutido em Ameka (1992), a suspeita sobre o estatuto vocabular das interjeições vem da observação de que elas são “anomalias fonológicas” ou “vocalizações paralinguísticas” [que expressam estados da alma]. Dito de outro modo, a suspeita sobre o estatuto vocabular das interjeições vem do seu caráter *fonológico* – não apenas por parte

---

4 Isso pode ser assumido, pois as interjeições são signos linguísticos genuínos, da perspectiva saussuriana – ou seja, são uma sequência de sons associada a um significado – e também porque falantes nativos – mas não falantes estrangeiros – têm intuição sobre quando, ou como, usar uma dada interjeição na sua língua.

dos linguistas, mas também por parte dos falantes. Alguns falantes hesitariam ao classificar “*uhum*” como uma palavra do português de uma forma que não hesitariam ao classificar “*sim*”; o mesmo se aplica a *putz* versus *minha nossa* ou *meu Deus*, ainda que ambas as formas sejam semelhantes dos pontos de vista sintático e semântico.

Logo, a “intuição do falante” parece levar em conta, sim, propriedades fonológicas para classificar algo como “palavra”. Raciocínio semelhante pode ser observado na análise do que se chama de “estrangeirismos”. Especificamente, de um ponto de vista estritamente linguístico (ou seja, desconsiderando fatores culturais etc.), os falantes tendem a considerar uma palavra como “estrangeira” quando sua realização fonológica não obedece a padrões da fonologia do português, como *budget* e *déjà vu*, para além da ortografia, como discutido a seguir. Há inclusive trabalhos que analisam as *adaptações fonotáticas* de palavras estrangeiras quando incorporadas ao léxico português – cf., por exemplo, Massini-Cagliari (2011) e referências. Esse movimento só pode ser resultado da conclusão de que a intuição sobre as palavras é sensível a propriedades fonológicas.

Do ponto de vista morfológico, para Camara Jr. (1970, p. 106), “o vocábulo formal é a unidade a que se chega quando não é possível uma nova divisão em uma ou mais formas livres” – sendo “formas livres” aquelas que podem ocorrer isoladamente e funcionam como comunicação independente. Além dos problemas fonológicos que emergem dessa definição (discutidos pelo próprio Camara Jr.), mesmo na morfologia, tal caracterização encontra problemas. O primeiro vem de advérbios em *-mente*: ainda que *rapidamente* seja um vocábulo formal, ele pode ser subdividido em outra forma livre, como visto em *rápida e sorrateiramente*. Vale observar que essa “elipse morfológica” (RESENDE, 2017) é licenciada apenas em um contexto sintático específico e não faz parte da “natureza vocabular/lexical” deste ou de outros advérbios do PB.

Outro problema, notado por Borges Neto (2020), tem a ver com a possibilidade de o sufixo diminutivo *-zinho* (uma forma presa) aparecer “livre” em *Ana perdeu a promoção por causa de uma zinha que trabalha com ela* e também dos prefixos *ex-* ou *pós-* em *João ainda pensa no ex* e *Maria ainda não terminou a pós* – cf. Resende (2017) para discussão – ou de casos como *analisar os prós e os contras* etc.

Finalmente, um argumento empírico derradeiro contra a solidez dessa caracterização vem de itens como *tranco* (em “pegar no tranco” ou “aos trancos e barrancos”), *zás* (em “num zás-traz”), *toa* (em “à toa”), *eira* (em “sem eira nem beira”) ou *lé* e *cré* (em “não juntar lé com cré”), que constituem unidades fonologicamente bem formadas, com acento

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

primário, mas que não funcionam *isoladas* como comunicação independente<sup>5</sup>. Vale notar que esse é o caso mais extremo de uma escala que contém compostos (tais como *guarda-chuva*) e expressões idiomáticas (como *chutar o balde*), em que suas palavras constitutivas não podem ocorrer isoladamente, com a mesma “função comunicativa”.

Do ponto de vista sintático, há, pelo menos, duas maneiras de articular a (aparente) relevância da unidade “palavra” para a sintaxe. A primeira delas é falar de uma PALAVRA MORFOSSINTÁTICA, como discutido por Borges Neto (2020, p. 20), que é “qualquer forma que um item lexical pode assumir para funções gramaticais”. Acerca disso, suponha-se que “...para funções gramaticais” seja um critério claro, com consistência preditiva; além disso, suponha-se que seja claro a que corresponde um “item lexical”.

À primeira vista, dada a divisão de tarefas clássica entre *morfologia derivacional no léxico* e *morfologia flexional na sintaxe* (ANDERSON, 1982), seria possível aventar que uma palavra morfossintática – frequentemente tida como primitivo dos modelos PALAVRA-E-PARADIGMA – seria a unidade que é sensível ao contexto sintático como, na morfologia flexional. Todavia, é igualmente possível aventar que um “item lexical” verbal assuma uma função gramatical de “nome”, por exemplo, (como em uma nominalização) (CHOMSKY, 1970) e também seja, no domínio derivacional, uma “palavra morfossintática”. Seja como for, essa definição não é nem necessária nem suficiente para recuperar a “intuição” sobre “palavra”, tampouco para capturar a relevância sintática de certos processos morfológicos; ela parece reproduzir apenas a separação circular entre palavras diferentes e formas de uma mesma palavra.

Sob outra perspectiva, para Di Sciullo e Williams (1987), as palavras – itens lexicais – são os átomos sintáticos – elementos que ocupam a posição de núcleo ( $X^0$ ) na sintaxe e cuja estrutura interna não pode ser acessada ou referida pelas regras sintáticas – o que ficou conhecido na literatura como “HIPÓTESE DA INTEGRIDADE LEXICAL”. Já há muito debate em torno dessa hipótese, e não é objetivo do presente artigo reproduzir essa discussão – mas cf. Bruening (2016). A ideia é apenas frisar que essa é uma noção (ainda) corrente na visão lexicalista, a qual tem sido veementemente desafiada, com base em observações empíricas translinguísticas advindas de fenômenos como elipse, escopo de advérbio e operadores, estrutura argumental, entre outros.

---

<sup>5</sup> Um parecerista anônimo ressaltou que o item “tranco” pode aparecer em outras construções com verbos leves, tais como, *dar um tranco na mesa* (com a leitura de “esbarrar na mesa”). Nesse sentido, é possível que a interpretação contextual desses itens possa variar a depender do dialeto – agradeço a esse parecerista por sugerir o exemplo “à toa”.

Do ponto de vista semântico, a ideia corrente – ou intuição – é a de que as palavras são unidades atômicas de significado. Essa ideia é trivialmente questionável sob a ótica da morfologia, no sentido de que – com alguma ressalva para vogais temáticas – os *morfemas* é que são as unidades mínimas de (som e) significado. No entanto, contingenciada pela visão de que a sintaxe não tem acesso à estrutura interna das palavras – isto é, aos morfemas – a ideia é que a unidade mínima de significado *relevante para a sintaxe* é o item lexical – a palavra. Ainda assim, mesmo na acepção de que são as palavras – formas livres e formas dependentes – as unidades mínimas de significado, alguns fenômenos empíricos depõem contra a análise de que o significado dessas formas é atômico ou especial.

Como discute Marantz (1997), para defender que os itens lexicais codificam alguma noção especial de significado, seria necessário mostrar empiricamente duas propriedades fundamentais. A primeira delas é a de que deveria haver algo atômico ou especial (do ponto de vista do significado) em *desejar*, *apoiar* e *comprar* que não pode ser recuperado por construções (sintáticas) com verbos leves correlatos como *ter desejo*, *dar apoio* e *fazer uma compra*. O mesmo raciocínio se aplica a *suavizar/tornar suave*, *derrubar/fazer cair* e a *pós-guerra/depois da guerra*. Se os significados desses itens lexicais podem também ser construídos na sintaxe, então eles não são significados atômicos.

A segunda propriedade seria demonstrável na justificativa de que o significado não composicional das palavras, a partir dos morfemas, é especial – e diferente – em relação ao significado não composicional de sintagmas, a partir de itens lexicais. Dito de outro modo, seria necessário mostrar que a junção de morfemas que gera um significado especial – por exemplo, “camisinha” não significa “camisa” + “inha”, mas *preservativo* – é diferente da junção de palavras que gera um significado especial – por exemplo, “bater as botas” não significa “bater” + “as” + “botas”, mas *morrer*. No entanto, a generalização parece ser a de que existem construções de vários tamanhos que não significam o que a soma de suas partes normalmente significa.

Outro problema ainda para a noção semântica de “palavra” é oferecido por alguns compostos como *sofá-cama*, *surdo-mudo*, *bomba-relógio* e *público-alvo*, para citar alguns, que, embora possam ser caracterizados como “átomos sintáticos” – nos termos de Di Sciullo e Williams (1987) – são transparentemente composicionais ou, pelo menos, apresentam uma leitura composicional disponível. Portanto, existem coisas que se costuma chamar de “item lexical”/“palavra”, mas que não são atômicos ou especiais nem do ponto de vista do significado.

Finalmente, existe ainda uma última noção de palavra – a *palavra ortográfica* – que merece alguns comentários. Naturalmente, da perspectiva da aquisição da linguagem,

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

essa noção é trivialmente irrelevante, uma vez que as crianças adquirem a fala antes de aprender a escrever. Ainda assim, a partir da discussão de Camara Jr. (1970) e Borges Neto (2020) e à luz da discussão desta subseção, o reconhecimento de uma “palavra ortográfica” pode iluminar algumas questões pertinentes à presente discussão<sup>6</sup>.

Como argumentado nesta subseção, a intuição do falante a respeito do que seja uma palavra não é empiricamente delimitada em nenhum nível, isto é, não só não há isomorfia entre “palavra fonológica” e “palavra morfológica”, como também não há correspondência entre estas e “palavra semântica” (ou “lexema”) ou entre alguma delas e “palavra (morfo) sintática” (ou “item lexical”). Curiosamente, todavia, o único domínio em que todas essas noções coincidem é na língua escrita.

Do ponto de vista ortográfico, *palavra* é uma unidade identificável por dois espaços em branco – um antes e um depois. Essa classificação é a única que fundamentalmente faz com que uma única *forma* congregue informações gramaticais, fonológicas e semânticas, não atômicas e não correspondentes entre si – indo inclusive na direção contrária à língua falada. Nesse sentido, como observa Borges Neto (2020), é bastante plausível supor que a intuição sobre “palavra” que os falantes – e os linguistas – parecem ter esteja intimamente associada a uma intuição sobre a escrita, cujo critério de identificação é bastante claro.

Evidência adicional para essa hipótese é oferecida na segunda seção, a partir de uma análise de *corpus* de aquisição, com base no qual se pode concluir que “palavra” não é uma unidade privilegiada da aquisição da linguagem. Se tal análise estiver na direção correta, pode se dar suporte adicional à ideia de que a intuição sobre palavra é construída via escolarização e por um processo de naturalização da Gramática Tradicional (BORGES NETO, 2020), um modelo *teórico* PALAVRA-E-PARADIGMA. Vale notar que crianças em início de alfabetização cometem muitos erros de segmentação de palavras, porque levam em conta critérios da língua falada (ou seja, *da sua língua*). Se houvesse alguma intuição “primitiva” de palavra (antes da escolarização), seria esperado que as crianças apresentassem menos dificuldades na segmentação vocabular.

Seja como for, a intuição sobre palavra ortográfica também enfrenta problemas, mas, nesse caso, a hesitação da classificação vem não de uma análise empírica, mas da falta de proficiência no conhecimento das regras ortográficas vigentes, da língua padrão. Exemplos conhecidos são: *de novo*, *de repente* e *de cor* – que se escrevem separados – diferentemente de *devagar* (cf. *vagarosamente*) e demais (cf. *de menos*) que se escrevem junto. Outros casos clássicos são: *em cima versus embaixo*, *afim versus a fim*, *senão versus*

---

6 Cf. Scliar-Cabral (2021) para referências e para uma discussão recente e importante sobre o processamento das unidades de alfabetização.

*se não* etc. Essas segmentações não parecem levar em conta critérios consistentes ou “intuitivos”, que vão além de motivações extralinguísticas (históricas, culturais etc.) para/ sobre a escrita.

Como adiantado, outra evidência para “palavra” que leva em conta a ortografia tem a ver com a grafia de palavras estrangeiras, ou seja, os falantes parecem ter alguma noção sobre o fato de que uma palavra não é do português, porque estas não se conformam normalmente ao padrão de escrita vigente. Cumpre notar que, da perspectiva da aquisição da linguagem, não existem estrangeirismos; tudo é português. Pareceria pouco convincente a ideia de que crianças com 3, 4 ou 5 anos de idade, adquirindo palavras como *notebook*, [nɔ.tʃi.'bu.kɪ], *tablet* ['ta.ble.tʃi] ou *Mc* ['mɛ.kɪ], teriam mais dificuldade com estas do que teriam com *tutifruti* [tu.tʃi.'fru.tɪ], *plástico* ['plas.tʃi.kɔ] ou *leque* ['lɛ.kɪ], diante do fato de que estas são palavras estrangeiras.

Finalmente, um outro problema subjacente à hesitação para classificar uma palavra ortográfica tem a ver com o uso do hífen, sobretudo quando ocorrem mudanças nas normas gramaticais vigentes – um novo ACORDO ORTOGRÁFICO. Dentro desse domínio – a escrita –, os critérios para a delimitação de uma palavra são ainda menos óbvios. Vale notar que “estar entre dois espaços em branco” é um critério circular e artificial, pois a questão jaz na motivação de *por que* uma certa unidade (e não outra) deve aparecer separada das demais por meio desses espaços – ou mesmo de um hífen.

No seu uso mais geral, o hífen serve para rejuntar partes de uma palavra que, por causa da translineação, ficaram separadas. Sob tal critério, em empregos mais específicos, palavras como “surdo-mudo”, “anti-inflamatório” e “falar-se-ia” também reuniriam partes *originalmente* separadas. Esse movimento, todavia, acaba obscurecendo a distinção entre COMPOSIÇÃO e PREFIXAÇÃO, já que *alguns* prefixos são grafados com hífen e outros, não. É preciso observar que, naturalmente, os linguistas têm critérios morfológicos, fonológicos e semânticos para distinguir prefixação e composição (ainda que esta seja uma discussão acalorada), mas os falantes, desconhecendo tais critérios, só podem se apoiar na ortografia, o que constitui evidência de que a intuição sobre palavra é ortográfica; ou seja; a hesitação sobre classificar explicitamente o que é uma palavra é fruto dos critérios inconsistentes do sistema de escrita – cf., por exemplo, “fala-se” *versus* “se fala”.

Naturalmente, o intuito desses questionamentos não é de apontar as incongruências do sistema de escrita e/ou da tradição gramatical; a relevância dessas observações para o problema em questão é a de que a intuição sobre o que é ou não uma palavra pode estar baseada na língua escrita, em vez de na língua falada. Se esse for o caso, a intuição sobre as demarcações dessa unidade não faz parte do conhecimento linguístico inato, mas sim, do conhecimento gramatical construído por meio da escolarização.

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

Seja como for, é preciso ficar claro que não existe fundamentalmente um problema na identificação de palavras por parte do falante, mesmo que este se baseie na escrita. O problema surge unicamente quando linguistas insistem, contrariamente às evidências, em assumir tácita ou acriticamente que tais unidades *intuitivas* são primitivas ou especiais, sob algum ponto de vista relevante ou crucial nas teorização e análise linguísticas.

### **A unidade “palavra” na gramática da criança**

Independentemente de a intuição sobre “palavra” emergir (ou não) da naturalização da Gramática Tradicional e/ou da intuição sobre “palavra ortográfica”, é fato que, da ótica da fonologia, da sintaxe, da semântica e da própria morfologia, a falta de isomorfia entre as diferentes facetas dessa unidade coloca um problema fundamental para a teoria linguística. Assim, uma saída metodológica interessante – muitas vezes adotada, explicitamente ou não – é: se se trata de uma análise fonológica, recorre-se à palavra fonológica; se é semântica, ao lexema; se sintática, recorre-se ao item lexical – o que quer que este signifique. Porém, do ponto de vista da aquisição da linguagem, não é óbvio como essa metodologia pode ser adotada se a aquisição da linguagem não constitui um nível de análise linguística *per se*; ou seja, do ponto de vista empírico, a aquisição é antes uma certa perspectiva para um dado fenômeno ou nível de análise do que um objeto de análise autônomo.

Scarpa (2012, p. 243) esclarece que “a aquisição da linguagem é, pelas suas indagações, uma área híbrida, heterogênea e multidisciplinar. No meio do caminho entre teorias linguísticas e psicológicas, tem sido tributária das indagações advindas tanto da Psicologia quanto da Linguística”. Por essa razão, para a autora, “as questões suscitadas pela aquisição da linguagem, bem como os muitos problemas metodológicos e teóricos colocados pelos próprios dados aquisicionais, têm, não raro, levado tanto a Psicologia quanto a própria Linguística a se repensarem e se renovarem” (SCARPA, 2012, p. 243).

No bojo dessas considerações, é possível vislumbrar que a interface entre a aquisição da linguagem e a metateoria linguística pode seguir dois caminhos, a saber, a apropriação das generalizações empíricas e da descrição dos dados de uma língua (a gramática-alvo) como ponto de partida para a caracterização do desenvolvimento linguístico infantil e/ou a tomada dos dados de aquisição como ferramenta metodológica de análise teórica (YANG; ROEPER, 2011).

Trilhando o segundo caminho, o presente artigo visa iluminar um problema teórico e metodológico que tem a ver como a noção de “palavra”. Parece razoável supor que um linguista que pretende compreender a aquisição de oclusivas em uma língua tenha critérios

claros de como identificar sons oclusivos. O mesmo raciocínio se aplica a um pesquisador que busca compreender a aquisição de orações relativas ou de quantificadores; não parece ser possível fazer previsões/generalizações consistentes com base em definições imprecisas – ainda que elas sejam “intuitivas”. É nesse sentido que a unidade “palavra”, assumida tácita e/ou acriticamente, pode enviesar as respostas de certas perguntas de pesquisa<sup>7</sup>.

O problema empírico (§ 1) por trás da noção de palavra acaba criando um problema metodológico e um problema teórico para os estudos em aquisição. No primeiro caso, se não há critérios claros para definir o que é uma palavra ou um item lexical, não é possível fazer previsões ou generalizações sobre (o que é) aquisição de léxico. Nessa mesma linha, outro problema metodológico é o de que sem saber, com clareza, o que é um item lexical ou seu “tamanho” e as propriedades que o caracterizam, não é possível determinar, com precisão, a diferença entre o período HOLOFRÁSICO (cf. QUADROS, 2008) – enunciados de uma única palavra – e o período de início da aquisição de sintaxe – enunciados de duas ou mais palavras –, já que essa diferença entre os estágios está fundamentalmente construída na visão de que as palavras configuram uma unidade especial, pré-sintática.

Simplificadamente, o percurso de aquisição – nesse caso, de produção – da linguagem é normalmente entendido da seguinte forma: os bebês armazenam palavras em seu léxico mental; então entre 10-12 meses, começam a proferir as primeiras palavras. Em seguida, entre 16-20 meses, ocorre a EXPLOÇÃO DE VOCABULÁRIO e, posteriormente, entre 18-20 meses, os bebês passam a produzir sintaxe, isto é, enunciados de mais de uma palavra. Por fim, entre 20-30 meses, as crianças começam a empreender a análise da estrutura interna das palavras e a formar palavras novas, o que caracterizaria a aquisição de morfologia<sup>8</sup>.

Porém, para que este fosse um percurso metodologicamente coerente, seria esperado que houvesse (i) diferença qualitativa entre a concatenação de morfemas e a concatenação de itens lexicais; (ii) diferença qualitativa entre a aquisição de morfologia derivacional e morfologia flexional; (iii) evidência de que os primeiros enunciados da criança – os enunciados holofrásicos – fossem monoblocos e não estivessem subdivididos em unidades menores, na gramática infantil, naquele estágio de aquisição. Assim, com vistas a analisar empiricamente essas três expectativas, o presente estudo realizou uma pesquisa baseada em *corpora* da maneira descrita a seguir.

---

7 Uma consequência dessa discussão – não abordada no presente artigo – diz respeito ao cálculo da EME (EXTENSÃO MÉDICA DO ENUNCIADO), que leva em conta as unidades mínimas constitutivas das produções infantis; especificamente a razão entre o número de morfemas e o número de palavras – cf., por exemplo, Del Re e Hilário (2013) para discussão e referências.

8 Costa e Santos (2003, cap. 2), Quadros (2008), Scarpa (2012), Grolla e Figueiredo Silva (2014, cap. 2).

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

## Metodologia e hipóteses

Com o intuito de avaliar a hipótese de que as crianças armazenam e produzem “palavras”, o presente estudo realizou uma coleta de dados longitudinais – provenientes de gravações/transcrições – de cinco crianças adquirindo o PB. As faixas etárias analisadas compreendem um intervalo entre 12 meses – período associado à produção dos primeiros enunciados – e 28 meses – período associado ao início da produção de sobregeneralizações – fenômeno conectado à produtividade morfológica<sup>9</sup>. O conjunto de dados de cada criança, dentro desse intervalo, esteve sujeito à disponibilidade nos *corpora*. Assim, a constituição do *corpus* desta pesquisa aparece resumida no Quadro 1<sup>10</sup>.

**Quadro 1.** Constituição do corpus Criança

	Fonte	Período disponível analisado
R	CEDAE	01;01.08 – 02;02.28
L	CEDAE	01;06.14 – 02;03.14
F	CEDAE	01;09.05 – 02;04.21
AC	CEAL	01;08.12 – 02;03.00
G	CEAL	01;10.21 – 02;03.17

**Fonte:** Elaboração própria

A composição do conjunto de dados sob análise foi implementada por meio da leitura das transcrições, seguida da coleta dos enunciados de cada criança, independentemente do seu “tamanho”, de sua complexidade estrutural e também de se o enunciado proferido era “bem-formado”, isto é, se apresentava um pareamento forma-significado que corresponde àquele observado na gramática do adulto. Naturalmente, porém, construções foneticamente malformadas (isto é, diferentes da gramática-alvo) foram metodologicamente consideradas como sendo fruto de um aparelho articulatório em desenvolvimento, e não de uma falha na representação interna dessas estruturas – da LÍNGUA-I.

9 Cf. Costa e Santos (2003, cap. 2), Quadros (2008), Scarpa (2012), Grolla e Figueiredo Silva (2014, cap. 2).

10 A criança é identificada pela primeira letra do seu nome. O *corpus* do CEDAE (Centro de Documentação Cultural *Alexandre Eulálio*) pertence ao Instituto de Estudos da Linguagem (IEL) da Universidade Estadual de Campinas (UNICAMP), e o *corpus* do CEAL (Centro de Estudos de Aquisição da Linguagem) pertence à Pontifícia Universidade Católica do Rio Grande do Sul (PUC-RS). Na descrição do período, os números indicam a idade da criança, especificamente ano; mês. dia.

Entretanto, foram excluídos da amostra (i) enunciados ininteligíveis, aqueles aos quais não foi possível atribuir nenhuma interpretação por causa de sua forma, nem mesmo dentro do contexto; (ii) enunciados que claramente configuravam uma mera repetição da fala do adulto; (iii) enunciados excessivamente recorrentes, tais como, algumas formas verbais e vocativos etc. – nesses casos, apenas a primeira ocorrência foi considerada; (iv) enunciados onomatopaicos ou unicamente interjetivos.

Assim, aplicando esses procedimentos metodológicos, o conjunto de dados analisado constitui-se de 2154 enunciados. Naturalmente, este artigo não pretende apresentar, discutir nem analisar todos eles; a ideia é oferecer uma amostra relevante, tanto quantitativa quanto qualitativamente, de um pequeno conjunto de fenômenos com vistas a determinar quais, de fato, parecem ser as unidades com que o bebê opera nos seus primeiros enunciados. Como discutido anteriormente, a ideia corrente é a de que essas unidades são palavras e que há uma diferença qualitativa entre o período holofrásico (sem sintaxe) e o período do início de combinação dos itens lexicais (com sintaxe). No bojo dessas considerações e a partir da primeira seção, esta análise se vale de três critérios para a identificação das unidades primitivas, a saber, um fonológico, um semântico e um sintático.

Da perspectiva da fonologia, já foi amplamente demonstrado na literatura que as unidades fonológicas e a sensibilidade para certos fenômenos fonéticos, fonológicos e prosódicos desempenham um papel crucial na aquisição da linguagem, sobretudo no nível do desencadeamento (*'bootstrapping'*) – cf. Guasti (2002, cap. 2) e também Gout e Cristophe (2006) para discussão e referências. Seja como for, do ponto de vista da *produção*, basta identificar quais são as unidades mínimas “de forma” que a criança produz, observando a relação de adjacência com outras unidades fonológicas; isto é, por meio da identificação dos contextos (antes, depois, isoladamente) em que a unidade aparece.

Da perspectiva semântica, nem sempre é fácil determinar quanto contexto estrutural é necessário para a interpretação de uma unidade; ainda assim, pode-se supor que unidade *mínima*, quando bem interpretada, pode aparecer em vários contextos diferentes – já que é preciso saber o que essa unidade significa para empregá-la de forma apropriada. Por fim, do ponto de vista sintático, a questão versa principalmente sobre qual é a operação da gramática responsável por formar (compor) as diferentes unidades. Portanto, o objetivo central desse recorte metodológico é mostrar – e não simplesmente assumir – quais são as unidades mínimas de produção na fala infantil. Vale lembrar que o problema da delimitação das unidades de desencadeamento bem como de armazenamento é ortogonal à delimitação das unidades de produção – cf. Resende (2021) para discussão e referências.

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

Em outras palavras, a questão que subjaz a quais são as unidades envolvidas no processo de aquisição da linguagem envolve três subníveis, a saber, (i) as unidades que desencadeiam o processo; (ii) as unidades que a criança efetivamente armazena e (iii) as unidades que a criança produz e junta para produzir enunciados. Nesse cenário, o objetivo desta pesquisa é abordar a questão (iii), embora a delimitação de quais sejam as unidades de produção possa criar condições para uma melhor apreensão das unidades armazenadas no léxico mental da criança que, por seu turno, podem iluminar a discussão sobre as unidades desencadeadoras do processo, isto é, as primeiras pistas linguísticas relevantes para o bebê.

Em termos teóricos, este estudo cria condições para que se discutam os primitivos formais dos enunciados; no caso, se são palavras ou unidades menores – os morfemas. De todo modo, dada a dificuldade de delimitar, com segurança, o que é uma palavra (cf. a primeira seção), os morfemas, diferentemente, podem ser metodologicamente entendidos como unidades mínimas de forma e de sentido – mas cf. terceira seção para discussão. Portanto, a hipótese deste estudo é a de que as unidades mínimas – som e significado – de produção são os morfemas, e o ônus da prova recai sobre aqueles que defendem que as unidades mínimas devem ser maiores do que os morfemas; por exemplo, as palavras.

## Resultados e discussão

Via de regra, em se tratando de unidades discretas na aquisição, as primeiras formas fonológicas produzidas pela criança são as sílabas. Embora já apareçam “sílabas” durante o período do BALBUCIO, alguns estudos têm mostrado que a articulação *fonológica* dos sons da língua-alvo e que as primeiras associações dessas unidades a algum significado – o que fundamentalmente caracterizaria o início da produção *linguística* – ocorrem alguns meses mais tarde (cf. GUSTI, 2002), entre 10 e 12 meses. Portanto, a busca pelos primeiros enunciados infantis deve partir mais ou menos desse período (cf. a subseção “Metodologia e hipóteses”).

Acerca disso, no entanto, é preciso ficar claro que defender que os *morfemas*, e não as palavras, são as unidades básicas/mínimas “de forma” não significa que se espera que os morfemas apareçam isolados em enunciados “holomorfêmicos”, pois essa hipótese seria trivialmente falsa, uma vez que existem restrições de ordem fonológica (como restrições fonotáticas), morfológica (morfemas são, a rigor, formas presas), semântica (a interpretação de uma unidade depende de sua estrutura), entre outras, contra essa possível visão. Nesse sentido, então, basta mostrar que, embora os primeiros enunciados

da criança (ainda que monossilábicos) não sejam monomorfêmicos, é possível analisá-los como sendo unidades mínimas dentro dos enunciados se a criança parece poder ser capaz de “isolá-los” e mobilizá-los em diferentes enunciados, como pode ser visto no Quadro 2.

**Quadro 2.** Evidência das raízes como unidades mínimas de forma

Criança	1º contexto	2º contexto
R	[a.bi] ('abre') 01;03.06	[a.bu] ('abriu') 01;04.24
R	[o.tu] ('outro') 01;05.10	[o.tɐ] ('outra') 01;07.00
R	[tʃi.lu] ('tiro') 01;06.03	[tʃi.lɐ] ('tira') 01;06.22
R	[fa.'so] ('fechou') 01;06.22	[fɛ.sɐ] ('fecha') 01;06.29
R	[e.si] ('esse') 01;06.29	[ɛ.sɐ] ('essa') 01;07.13
L	[bo.bɐ] ('boba') 01;06.14	[bo.bu] ('bobo') 01;06.14
L	[ej.u] ('feio') 01;06.19	[fej.ɐ] ('feia') 01;07.30
L	[mi.'ni.nu] ('menino') 01;07.08	[mi.'ni.nɐ] ('menina') 01;07.08
L	[fa.tu] ('chato') 01;07.30	[fa.tɐ] ('chata') 01;07.30
L	[da] ('dá') 01;10.28	[dew] ('deu') 01;10.28
F	[kɛ.lu] ('quero') 01;09.05	[kɛ] ('quer') 01;09.05
F	[a.'mɛ.lu] ('amarelo') 01;09.05	[ma.lɛ.'li.ɲu] ('amarelinho') 01;09.05
F	[bĩ.'ka] ('brincar') 01;09.19	[bĩ.ku] ('brinco') 01;09.19
F	[a.'fo] ('achou') 01;09.19	[a.'fej] ('achei') 01;09.19
F	[pu.'luɐ] ('perua') 01;09.19	[pi.lu.'i.ɲɐ] ('peruinha') 01;09.19
AC	[a.ma.'lɛ.lu] ('amarelo') 01;08.12	[a.ma.lɛ.'li.ɲu] ('amarelinho') 01;08.12
AC	[bɔj.ɐ] ('bola') 01;08.12	[bɔ.'li.ɲɐ] ('bolinha') 01;10.25
AC	[u.su] ('urso') 01;08.12	[u.'si.ɲu] ('ursinho') 01;10.25
AC	[kẽ.'ʃa.du] ('cansado') 01;08.12	[kẽ.'so] ('cansou') 02;01.26
AC	[se.bo.'li.nɐ] ('cebolinha') 01;10.25	[se.'bo.lɐ] ('cebola') 02;01.26
G	[bo.lɛ.'tʃi.ɲɐ] ('borboletinha') 01;10.21	[bo.bo.'lɛ.tɐ] ('borboleta') 01;10.21
G	[ɛ.sɐ] ('essa') 01;10.21	[e.si] ('esse') 01;10.21
G	[bõw] ('bom') 01;10.21	[bo.ɐ] ('boa') 02;01.27
G	[a.'biw] ('abriu') 01;10.21	[a.'bi] ('abrir') 02;01.27
G	[ko.'mẽ.du] ('comendo') 02;01.27	[ko.'mɛ] ('comer') 02;01.27

Fonte: Elaboração própria

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

O Quadro 2 contém cinco pares de formas produzidos por cada uma das crianças do *corpus*. A ideia é mostrar que os bebês, desde muito cedo – isto é, desde antes da explosão do vocabulário ou desde antes de serem proficientes em formar sentenças – já são capazes de empreender diferentes combinações com as raízes de sua língua que (por hipótese) são as unidades mínimas armazenadas no seu léxico mental – cf. Albuquerque, Bezerra e Ferrari-Neto (2012) para um trabalho que dá suporte a essa visão. Convém observar, no entanto, que muito embora os bebês não dominem, com segurança e estabilidade, quais as informações gramaticais e semânticas associadas a cada um dos morfemas e combinadas com as raízes, é possível concluir que as raízes (e os morfemas) são peças *fonologicamente* decomponíveis e reconhecíveis já nos primeiros enunciados infantis<sup>11</sup>.

Dito de outro modo, embora os bebês não empreguem corretamente certos morfemas desde o início de suas produções – algo já notado por Santos e Scarpa (2003), cf. terceira seção para discussão –, é possível afirmar que, da perspectiva da *forma*, as raízes já são identificadas. Nessa visão, com base no Quadro 2, pode-se observar que as crianças já são capazes de identificar as raízes, combinando-as com diferentes morfemas, tais como, algumas flexões verbais, vogais temáticas, marcas de gênero e/ou diminutivos.

Portanto, esses dados servem como evidência de que as palavras – da perspectiva da *forma* – não são as unidades mínimas, uma vez que as raízes aparecem empregadas em diferentes ambientes morfológicos (de sufixação). Com relação à metodologia de isolar a raiz a partir de um morfema sufixal, Clark (2014) observa que parece haver uma preferência de processamento, no sentido de que o radical é prontamente identificado através do ATAQUE (*onset*), especialmente quando este não é obscurecido por algum material prefixal. Assim, uma vez que a raiz foi identificada, tornam-se mais fáceis as modulações de significado que resultam da adição de algum material sufixal ao radical. Logo, com base nessa metodologia, é possível concluir que as crianças identificam as raízes como unidades formais, pela adição dos diferentes morfemas sufixais.

É verdade que os bebês não parecem dominar, desde o início, *todos* os pareamentos forma-significado no que concerne ao emprego dos morfemas – cf. § 3 para discussão e hipóteses. Ainda assim, mesmo que a criança profira *vai nanar* se referindo a ela mesma – destonado da gramática-alvo, que geraria *vou nanar* –, pode-se afirmar que ela já é capaz de reconhecer *v-* ou *-ai* como unidades mínimas de forma. Bagetti e Corrêa (2011), por

---

<sup>11</sup> Como um dos pareceristas anônimos observou, os dados apresentados no Quadro 2 também poderiam ser vistos como argumento para um modelo palavra-e-paradigma, em um processo chamado de “lematização”. E é natural que os dados empíricos possam, às vezes, servir de evidência para mais de uma teoria. A razão para rejeitar o modelo palavra-e-paradigma, no entanto, é apresentada na primeira seção e reforçada na terceira seção.

exemplo, mostram experimentalmente que crianças entre 9 e 12 meses já são sensíveis a alterações morfofonológicas dos padrões dos afixos verbais.

Pondo de lado esses casos, é possível notar ainda que, para além da *forma*, os bebês já parecem mostrar que identificam partes da palavra – ou ainda, parte dos seus enunciados “holofrásicos” – como decomponíveis do ponto de vista do *significado*. Especificamente, como mostra o Quadro 3, a utilização recorrente de certos morfemas mostra que os bebês já são capazes de mobilizar significados menores que a palavra.

**Quadro 3.** Evidência dos morfemas como unidades mínimas de *significado*

Criança	Morfema	Ocorrências		
R	-ou	[a.'bo] ['bo] ['so] [tʃi.'lo] [go.'do] [pa.'so] [so.'lo] [ko.'do]	('acabou') ('quebrou') ('fechou') ('tirou') ('guardou') ('passou') ('chorou') ('acordou')	01;03.06 01;03.06 01;06.06 01;06.22 01;06.22 01;07.00 01;07.08 01;07.13
L	-a	['fey.ɐ] ['ʃa.tɐ] [me.'ni.nɐ] ['bu.ʃɐ] ['ba.vɐ] ['li.dɐ] ['su.zɐ]	('feia') ('chata') ('menina') ('bruxa') ('brava') ('linda') ('suja')	01;07.30 01;07.30 01;10.08 01;11.00 02;01.13 02;01.16 02;01.16
F	-Vndo	[a.'bĩ.dɔ] [ka.'ĩ.dɔ] [pĩ.'sẽ.dɔ] [su.'bĩ.dɔ] [ẽ.'dẽ.nɔ] [fa.'zẽ.dɔ] [bĩ.'kẽ.dɔ]	('abrindo') ('caindo') ('pensando') ('subindo') ('andando') ('fazendo') ('brincando')	01;09.05 01;09.05 01;09.19 01;09.19 01;09.19 01;09.25 01;09.25
AC	-inhV	[ka.'zi.nɐ] [pa.sa.'i.nɔ] [ko.ko.'zi.nɔ] [a.mi.'gi.nɐ] [bo.nɛ.'ki.nɐ] [pe.ke.'ni.nɔ] [pe.'si.nɐ]	('casinha') ('passarinho') ('cocoquinho') ('amiguinha') ('bonequinha') ('pequeninho') ('pecinha')	01;08.12 01;08.12 01;08.12 01;08.12 01;08.12 01;08.12 01;08.12
G	-o	[ve.'me.jɔ] ['o.tɔ] ['ga.tɔ] ['po.kɔ] ['pa.tɔ] [ka.'da.dɔ] [bo.'ni.tɔ]	('vermelho') ('outro') ('gato') ('porco') ('pato') ('quadrado') ('bonito')	01;10.21 01;10.21 01;10.21 01;10.21 01;10.21 01;10.21 01;10.21

Fonte: Elaboração própria

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

No que tange ao Quadro 3, o objetivo é mostrar que os bebês são capazes de, muito cedo, identificar unidades portadoras de significado dentro das palavras. A rigor, o próprio Quadro 2 já constitui evidência para essa análise, pois se a criança é capaz de empregar os morfemas lexicais (raízes) em diferentes contextos (morfemas sufixais), ela já dá pistas de que associa um certo significado àquela *forma* (raiz). Adicionalmente, o Quadro 3 oferece o mesmo tipo de evidência, mas com base em morfemas funcionais (ou gramaticais).

Mais especificamente, o Quadro 3 apresenta cinco morfemas sufixais – todos eles exemplificados com sete ocorrências para cada uma das cinco crianças. Porém, convém mencionar que todos os morfemas foram registrados na fala de todas as crianças; a ideia é apenas exemplificar o argumento com uma maior variedade de dados. Vale a pena notar que os primeiros registros desses morfemas gramaticais surgem a partir dos 15 meses e passam a se tornar mais recorrentes a partir dos 18 meses<sup>12</sup> – cf. Santos e Scarpa (2003).

A respeito do morfema *-ou*, ainda que não associem, desde o início da aquisição, esse morfema às informações de 3ª pessoa, singular, 1ª conjugação, tempo pretérito e aspecto perfeito, é possível supor que os bebês associam *alguma* informação a [’o] que se mantém constante nas suas ocorrências, uma vez que este não é o único morfema verbal com que a criança opera antes dos 18 meses. Assim, os dados de R no Quadro 3 constituem evidência de que algum significado já é identificado – que é menor do que uma palavra.

Quanto aos morfemas *-a* (de L) e *-o* (de G), eles também servem como evidência de que essas partes da palavra já são significativas para as crianças. Augusto e Corrêa (2005), por exemplo, mostram que crianças de 22 meses já são capazes de reconhecer diferentes marcas de gênero – inclusive do gênero não marcado. Logo, não é necessário ‘esperar’ que as crianças produzam sobregeneralizações – tais como *passarinha* (AC 01;10.25) e *vaco* (F 02;02.00) – para concluir que ela já adquiriu a oposição morfológica de gênero no PB.

Especificamente, a ideia corrente de que as crianças só mostram que são capazes de analisar a estrutura interna das palavras *quando* produzem sobregeneralizações é uma contingência da ideia de que as palavras, no período dito “holofrásico”, são monoblocos. Todavia, é perfeitamente possível concluir com base em *formas gramaticais* que a criança

---

12 Convém observar que os dados disponíveis provenientes de F, AC e G são mais tardios do que aqueles de L e, principalmente, de R. Por essa razão, a ausência de exemplos em intervalos anteriores é resultado de uma questão metodológica – a falta de dados disponíveis – e não deve ser entendida como ausência de produção dessas formas na fala infantil nesses períodos. Por exemplo, ao passo que foram necessários quatro meses de transcrições para reunir sete exemplos da fala de R para o morfema *-ou* e seis meses da fala de L para o morfema *-a*, todas as ocorrências de *-inh-* para AC e de *-o* para G foram coletadas em uma única transcrição – o que mostra que, entre 1;08-1;10, as crianças já fazem usos muito produtivos dos morfemas gramaticais.

já empreende a análise da estrutura interna das palavras e, por isso, raízes e morfemas gramaticais aparecem em diferentes contextos. Com relação aos diminutivos (de AC) e aos gerúndios (de F), também é perfeitamente plausível concluir que essas crianças já têm uma representação clara de qual pode ser a contribuição semântica dessas partes de palavra.

Por fim, com relação à sintaxe das raízes e dos morfemas (em oposição à sintaxe dos sintagmas e sentenças que supostamente é diferente), atenta-se para o tipo de mecanismo que a gramática utiliza para formar palavras e para formar sintagmas e sentenças. Acerca disso, Chomsky (2005, p. 11-12) entende que

Um fato elementar a respeito da faculdade da linguagem é o de que esta é um sistema de infinitude discreta. Qualquer sistema desse tipo é baseado em uma operação primitiva que toma  $n$  objetos já construídos e constrói a partir destes um novo objeto: no caso mais simples, o conjunto desses objetos  $n$ . Chamemos essa operação de “concatenação”. A concatenação ou algum [mecanismo] equivalente é um requerimento mínimo. Dispondo da concatenação, nós temos instantaneamente um sistema ilimitado de expressões hierarquicamente estruturadas.

Nesse sentido, assumindo as versões mais recentes da GGT, a operação (mais básica) disponível na GRAMÁTICA UNIVERSAL (GU) é a CONCATENAÇÃO. Nesse cenário, o primeiro mecanismo de que a criança dispõe para gerar seus enunciados é o de juntar as unidades mínimas que estão no seu léxico mental, que, do ponto de vista da forma (Quadro 2) e do significado (Quadro 3), são raízes e morfemas – não palavras. Portanto, é esta a operação básica empregada para formar palavras (que não são unidades primitivas) e, posteriormente sintagmas e, então, sentenças. Logo, o ônus da prova recai sobre aqueles que defendem que a operação responsável por formar palavras e por formar sintagmas/sentenças é diferente.

Se a criança já dispõe de  $\sqrt{\text{MENIN}}$ , /o/e /a/ no seu léxico mental, é a *mesma* operação de concatenação que vai juntar essas unidades mínimas para gerar enunciados de qualquer tamanho. À medida que o processo de aquisição se desenvolve, essa mesma operação vai ser mobilizada para juntar *menino* com *o*, *o menino* com *alto* etc. Seguindo esse raciocínio, a manipulação da operação de concatenação é muito anterior ao período “holofrásico”, e a identificação de formas (cada vez) maiores do que uma palavra é um processo natural que resulta de seguidas aplicações da operação de concatenação, como descrito por Chomsky.

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

No bojo dessas considerações, o Quadro 4 ilustra a dinamicidade das combinações empreendidas pelas crianças desde muito cedo com base em alguns morfemas de grau (no caso, diminutivos e aumentativos). O objetivo é mostrar que, apesar do repertório limitado das crianças de até 24 meses, há um grande número de combinações diferentes com *-inho* e *-ão*, que mostram o uso produtivo da operação de concatenação com as unidades mínimas já disponíveis.

**Quadro 4.** Dinamicidade e produtividade nas combinações com morfemas de grau

Exemplo	Criança/Idade	Exemplo	Criança/Idade
<b>DIMINUTIVOS</b>		<b>AUMENTATIVOS</b>	
<b>NOMES</b>			
[mi.'gi.nɐ] ('amiguinha')	AC 01;08.12	[pẽ.'no.nɐ] ('penona')	R 01;08.15
[pɛ.'si.nɐ] ('pecinha')	AC 01;08.12	[ma.'lo.nɐ] ('malona')	R 01;10.23
[bɔ.'li.jɐ] ('bolinha')	R 01;08.15	[pɔj.'kẽw] ('porcão')	R 01;11.25
[mẽ.ni.'njɐ] ('menininha')	R 01;08.15	[kæ.le.'lẽw] ('carneirão')	R 01;11.25
[ka.'fɪ.nɐ] ('caixinha')	R 01;08.20		
[fi.'li.nɐ] ('filhinha')	R 01;08.20		
[pa.'tʃi.nɔ] ('patinho')	R 01;08.20		
[ba.'li.nɐ] ('balinha')	R 01;08.20		
[ga.'tʃi.jɔ] ('gatinho')	R 01;09.09		
[ka.ʃo.'hi.nɔ] ('cachorrinho')	F 01;09.19		
[pi.lu.'i.nɐ] ('peruinha')	F 01;09.19		
[pĩ.'tʃi.nɔ] ('pintinho')	R 01;09.20		
[bi.'fɪ.nɔ] ('bichinho')	F 01;09.25		
[u.'si.nɔ] ('ursinho')	R 01;10.00		
[va.'li.nɐ] ('varinha')	R 01;10.00		
[po.'ki.nɔ] ('porquinho')	R 01;10.00		
[ka.'zi.nɐ] ('casinha')	R 01;10.20		
[va.'ki.nɐ] ('vaquinha')	R 01;10.21		
[bõ.'bi.nɐ] ('bombinha')	F 01;10.21		
[sa.'i.nɐ] ('sainha')	R 01;10.23		
[mẽ.tʃi.'li.nɐ] ('mentirinha')	R 01;10.23		
[sa.'pi.nɔ] ('sapinho')	F 01;10.28		
[ka.va.'li.nɔ] ('cavalinho')	F 01;11.06		

[tẽ.'pi.nɐ] ('tampinha')	F 01;11.14		
[kwe.'li.nɔ] ('coelhinho')	F 01;11.21		
[fu.mi.'gi.nɐ] ('formiguinha')	F 01;11.21		
[me.ni.'ni.nɔ] ('menininho')	R 01;11;25		
[ka.'hi.nɔ] ('carrinho')	R 01;11;25		
[bu.la.'ki.nɔ] ('buraquinho')	R 01;11;28		
[pej.'ʃi.nɔ] ('peixinho')	R 01;11;28		
<b>ADJETIVOS</b>			
[ma.le.'li.nɔ] ('amarelinho')	F 01;09.05	[gẽ.'do.nɐ] ('grandona')	F 01;10.14
[bẽ.'ki.nɔ] ('branquinho')	R 01;10.00	[gẽ.'dẽw] ('grandão')	R 01;10.23
[bõ.'zi.nɔ] ('bonzinho')	F 01;10.03	[mi.'zo.nɐ] ('mijona')	R 01;11.03
[pe.ke.ni.'ni.nɔ] ('pequenininho')	F 01;10.03		
[pi.ki.ni.'ni.nɐ] ('pequeninha')	R 01;11.03		
<b>ADVÉRBIOS</b>			
[po.'ki.nɔ] ('pouquinho')	AC 01;08.12		
[de.va.ga.'li.nɔ] ('devagarinho')	R 01;11.28		
[dʒi.ej.'ʃi.nɔ] ('direitinho')	L 02;01.16		

Fonte: Elaboração própria

O Quadro 4 não contém uma lista exaustiva de todas as ocorrências encontradas; apenas um conjunto representativo de casos não repetidos entre as crianças que aparecem até os 25 meses. Após esse período, a ocorrência desses sufixos é ainda mais acentuada, podendo-se encontrar inclusive exemplos como [bi.ʃẽw.'zẽw] ('bichãozinho') de L 02;03.14, em que o sufixo de aumentativo aparece empregado recursivamente. A ideia aqui é apenas mostrar que as crianças parecem ser bastante proficientes em juntar morfemas antes que se mostrem proficientes em juntar palavras para formar sentenças (em torno de 2 anos); no entanto, essa não é uma diferença qualitativa, mas quantitativa no sentido de que tem a ver com o "tamanho" das peças que são juntadas: primeiro menores (= morfemas) e depois, maiores (= palavras), como discutido em § 3.

O fato de diminutivos e aumentativos poderem aparecer em itens pertencentes a mais de uma categoria serve para mostrar que os bebês já estão atentos às várias possibilidades de combinação dos itens que estão em seu léxico mental. Seguindo o padrão já observado em outras línguas, o sufixo de diminutivo é o primeiro morfema derivacional efetivamente recorrente e produtivo na gramática da criança (CLARK, 2014), sendo muito superior ao do aumentativo – o que já reflete um padrão da gramática do adulto no PB.

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

Vale notar ainda que as crianças já implementam corretamente as vogais átonas finais /a/ e /o/ depois dos sufixos de grau, o que mostra que as combinações são mais sofisticadas do que pareceria à primeira vista, já que os bebês produzem corretamente a vogal adjacente ao afixo diminutivo, *mesmo* quando a vogal temática é /e/, o que aparece exemplificado por nomes masculinos como [gɔ.'li.ɲɔ] ('golinho') R 01;10.10, [pej.'fi.ɲɔ] ('peixinho') R 01;11.28, [pa.'ki.ɲɔ] ('parquinho') F 02;01.01, [to.pe.'tʃi.ɲɔ] ('topetinho') F 02;02.26, [ʒa.ka.'li.ɲɔ]<sup>13</sup> ('jacarezinho') F 02;02.12, [lej.'tʃi.ɲɔ] ('leitinho') L 02;02.21 e também pelo nome feminino [lɛ.'bi.ɲɛ] ('lebrinha') L 02;03.07.

### **Sem palavras: distinguindo aquisição de morfologia de aquisição de sintaxe**

Para além do problema das unidades mínimas de produção linguística da criança, a noção de “palavra” envolve uma outra questão teórico-metodológica que tem a ver com a separação tradicional entre morfologia flexional e morfologia derivacional. Essa distinção tem sido frequentemente usada para distinguir palavras diferentes (derivação) de diferentes formas da mesma palavra (flexão). Dada a discussão na primeira seção e levando em conta que nem a própria tradição gramatical oferece critérios claros para definir o que é uma “palavra”, essa oposição parece ainda menos motivada e, em última instância, menos necessária.

Adicionalmente, essa suposta distinção também tem sido empregada pelos linguistas para associar certos processos morfológicos a diferentes componentes da arquitetura da gramática: morfologia derivacional no léxico *versus* morfologia flexional na sintaxe – cf. Anderson (1982) para uma visão clássica. Discutir todos os problemas metodológicos e teóricos por trás dessa distinção vai muito além do propósito deste artigo – mas cf. Borges Neto (2012). Ainda assim, em se tratando de aquisição de morfologia, alguns comentários devem ser feitos.

Se as peças mínimas dos enunciados linguísticos produzidos pelas crianças (e pelos adultos) são morfemas (lexicais e gramaticais), não há diferença entre morfologia flexional e derivacional no que tange à aquisição de morfologia ou a alguma divisão de tarefas entre componentes distintos de formação de palavras, e as evidências empíricas disponíveis dão suporte a essa conclusão. Primeiramente, na esteira de Marantz (2001), as

---

<sup>13</sup> Na gramática alvo, o nome *jacaré* é normalmente considerado atemático, pois se encerra na vogal tônica [ɛ] e não em uma vogal átona, como [ə]. Para esses nomes ditos “atemáticos”, o diminutivo é *-zinho* e não *-inho*, cf. *sofazinho*, *pezinho*, *jabutizinho*, *cocozinho*, *tatuzinho* etc. Como é possível que a criança ainda não tenha adquirido plenamente essa regra morfofonológica do PB, é plausível supor que a aplicação de *-inho* se valha da análise (pela criança) de que o [ɛ] é também uma vogal temática.

“palavras” produzidas pelas crianças são, no mínimo, bimorfêmicas, ou seja, compostas pela raiz e algum sufixo – com exceção de alguns itens funcionais, que só aparecem mais tardiamente na fala da criança (cf. CORRÊA; AUGUSTO, 2017) e ainda de algumas poucas formas verbais como *quer*, que apresentam a mesma forma da raiz √QUER e que, mesmo nesse caso, são tratadas como flexionadas.

De acordo com Clark (2007, p. 374), “as crianças começam a mobilizar os processos de formação de palavras por volta do mesmo período de suas primeiras flexões”, afirmação esta que é corroborada pelos dados encontrados, pois os primeiros registros atestam tanto formas verbais e marcas de gênero quanto vogais temáticas nominais e as então chamadas “formas nominais do verbo” (infinitivos, gerúndios e participios) – o que já foi notado por Santos e Scarpa (2003). Nesse sentido, a novidade, então, não parece ser *empírica*, mas *metodológica*.

A rigor, as vogais temáticas nominais /a/, /e/ e /o/ são solenemente desconsideradas como pistas de aquisição de morfologia ou de formação de palavras, se se assume que as palavras já saem formadas do léxico. Entretanto, uma análise como aquela que subjaz ao Quadro 2, em que a combinação com diferentes vogais temáticas dá pistas de que a criança está atenta à raiz, mostra que as vogais finais não constituem a mesma unidade do ponto de vista da *morfologia*. O ruído metodológico vem, porém, de uma análise que não trata as vogais temáticas nominais como unidades relevantes para a aquisição, mas trata as vogais temáticas verbais como tal, ao considerar as formas *ama*, *come* e *abre* como flexionadas para a 3ª pessoa do singular, apesar de elas corresponderem ao mesmo templatão “√RAIZ + vogal temática” – crucialmente, o que caracteriza essas formas dentro do paradigma verbal do PB é a *ausência* de morfologia flexional.

Além disso, do ponto de vista da aquisição, no domínio nominal, as vogais são /a/, /e/ e /o/ e no verbal são /a/, /e/ e /i/ – e, em alguns contextos há neutralização das duas últimas em [ə] (CAMARA JR., 1970); isto é, não há evidência para uma separação metodológica entre essas informações, tampouco há evidência empírica. Mais uma vez, o ônus da prova recai sobre aqueles que querem defender que deve haver uma separação qualitativa entre *morfologia* derivacional e flexional; em particular, da perspectiva da aquisição.

Em todo caso, se esse raciocínio estiver na direção correta, uma questão permanece em aberto e versa sobre o comportamento de um certo tipo de morfologia, que desencadeia concordância sintática. Conforme Costa e Lopes (2016, p. 563), “adquirir a sintaxe de uma língua implica atingir o desempenho do adulto na produção e compreensão, pelo menos, dos três domínios seguintes: estrutura oracional, movimento e dependências referenciais”.

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

Nesse sentido, existe um tipo de morfologia que serve para sinalizar certas “dependências estruturais”, manifesta pelos morfemas de concordância.

Vale notar que, nesse quesito, a diferença *qualitativa* não está no tipo de morfologia, mas sim no tipo de sintaxe; ou seja, os morfemas em si podem ser (e são) adquiridos muito antes de a criança dominar as regras de concordância da sua língua; o ponto é que, depois de um certo período, a criança se dá conta de que alguns morfemas (verbais e nominais) devem estar “alinhados” na sentença, mas isso tem muito mais a ver com a aquisição de sintaxe, do que com a aquisição dos morfemas propriamente. O mesmo raciocínio pode ser aplicado à regência: desde muito cedo na aquisição, a criança junta unidades menores para formar unidades maiores; é apenas mais tardiamente que ela aprende que certas “junções” devem ser mediadas por itens funcionais, como preposições etc.

Fundamentalmente, os morfemas flexionais não são (puramente) funcionais em todas as palavras em que ocorrem; por exemplo, no verbo, [’o] (-ou) e [’ej] (-ei) são diferentes morfemas (cumulativos) e são adquiridos bastante cedo, e ainda que não esteja claro para a criança que -ou significa 3ª pessoa do singular e primeira conjugação, é razoável supor (com base no Quadro 3) que alguma informação aspectual/temporal já esteja em jogo. Assim, uma previsão seria a de que as primeiras associações que a criança empreende têm a ver com informações semânticas em vez de puramente gramaticais – o que formalmente, na GGT, pode ser implementado em termos de traços interpretáveis e não interpretáveis – cf. Corrêa e Augusto (2017) para uma discussão acerca do nível do sintagma nominal.

Logo, morfemas flexionais – sendo cumulativos – têm uma contraparte “gramatical” /“funcional”, mas também têm uma contraparte “interpretável” (isto é, semântica), e, na verdade, a propriedade caracterizadora da morfologia flexional faz parte da aquisição da sintaxe, e não tem a ver crucialmente com a produção dos morfemas em si, mas com a consciência da relação sintática que eles sinalizam. Logo, a diferença qualitativa que se esperaria na aquisição dos diferentes tipos de morfemas se perde, e o que se mantém é a observação de que, no processo de aquisição de sintaxe, algumas dessas formas precisam ser “repetidas” para que se possam estabelecer certas relações entre os constituintes.

Portanto, à luz dessas considerações, a aquisição de sintaxe não precisa ser definida em termos de um período de produção de uma para duas palavras – a menos que se possa definir, de forma consistente, o que é uma “palavra” –, basta apenas supor que a aquisição de sintaxe começa quando relações entre os constituintes são formalmente estabelecidas, e “formalmente” quer dizer: por meio de morfemas e partículas *gramaticais*. Para Chomsky (2005), como apontado, a operação que junta unidades menores para formar unidades maiores é dada pela GU, e não precisa ser “adquirida”.

Por hipótese, a criança já nasce equipada para juntar as unidades mínimas do seu léxico mental para compor unidades cada vez maiores, e o “tamanho” de tais unidades resulta do número de operações de concatenação que o bebê empreende em cada “estágio” da aquisição. Depois que a criança for capaz de aplicar um certo número de concatenações, ela começa a estabelecer relações “de longa distância” entre esses elementos – que é a aquisição de sintaxe. Nesse sentido, a separação entre o que é “adquirir morfologia” e o que é “adquirir sintaxe” não é uma questão de “quando formar palavras” *versus* “quando formar frases”, pois a operação de juntar (= formar) objetos linguísticos é dada pela GU.

Se esse raciocínio estiver na direção correta, e “isolar morfemas” e “formar palavras” não caracteriza a “aquisição de morfologia”, a pergunta se torna: o que significa adquirir morfologia? Como aparece discutido em Resende (2021), “adquirir morfologia” é adquirir o pareamento entre os feixes de traços morfossintáticos e semânticos da sintaxe à sua forma fonológica; ou seja, é identificar a quais informações semânticas e gramaticais se associam as peças fonológicas armazenadas no léxico mental. Dito de outro modo, a aquisição de morfologia é o estabelecimento da relação entre unidades fonológicas depreendidas do ambiente e o subconjunto de traços universais dadas pela GU (CHOMSKY, 1995; MARANTZ, 2001), e é justamente durante esse período que aparecem as sobregeneralizações.

Naturalmente, as maiores dificuldades para adquirir esses pareamentos vêm de traços gramaticais e não de informações semânticas, e a manipulação dos traços formais surge com mais afincos depois da aquisição da sintaxe – do modo como formulado por Costa e Lopes (2016). No bojo dessas considerações, as previsões feitas são de que os pareamentos com os traços interpretáveis devem ser adquiridos antes dos não interpretáveis e de que a aquisição destes é concomitante ou posterior à aquisição da sintaxe. De todo modo, a *formação* de palavras é anterior à *formação* de sintagmas/sentenças em virtude do número de aplicações da operação de concatenação. Ademais, todo o processo de desenvolvimento da aquisição da linguagem pode ser acomodado sem assumir que a palavra é especial, quer do ponto de vista formal/gramatical quer do ponto de vista das produções infantis.

## Palavras finais

Este artigo teve como objetivo revisitar o problema clássico da noção de “palavra” à luz de uma perspectiva diferente: a da aquisição da linguagem. Primeiramente, mostrou-se que há um problema empírico e teórico em relação às fronteiras dessa unidade que

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

pode criar um problema metodológico para algumas pesquisas. Em seguida, analisaram-se dados de aquisição da linguagem com vistas a mostrar que o início da produção dos enunciados infantis não deve ser caracterizado como “ausência de sintaxe” (em um sentido que concebe a concatenação como uma operação sintática), tampouco as produções ditas “holofráscas” devem ser analisadas como monoblocos ou como “falta de análise da estrutura interna das palavras”. Em última instância, a palavra não é uma unidade especial da aquisição.

Nesses termos, a intuição do falante sobre o que pode ser uma palavra de sua língua aliada à intuição sobre o que pode ser uma sentença bem formada é redutível a “os falantes são capazes de avaliar, como sendo bem ou malformadas, as estruturas de sua língua, independentemente do tamanho” – contra Villalva (2012).

Portanto, se essa reflexão estiver na direção correta e se as crianças, de fato, não se apoiam em palavras para adquirir a linguagem – ou fundamentalmente não têm a intuição sobre “palavra” que os adultos têm – então, é possível que essa intuição seja produto da escolarização e da naturalização da Gramática Tradicional (BORGES NETO, 2020) e que a intuição sobre palavra é, em última instância, intuição sobre palavra ortográfica.

## Agradecimentos

Gostaria de agradecer à Ana Paula Scher pela leitura prévia deste trabalho, por todas as suas observações, sugestões e também pelo constante diálogo durante o desenvolvimento desta pesquisa. Um agradecimento especial também ao Pablo Faria, pela cessão dos dados provenientes do *corpus* do CEDAE. Agradeço ainda aos pareceristas anônimos por suas observações e apontamentos que ajudaram a levar este artigo à sua forma final. Todos os erros são meus.

## Referências

ALBUQUERQUE, H.; BEZERRA, G.; FERRARI-NETO, J. Percepção infantil da morfologia derivacional: um estudo experimental sobre segmentação de morfemas em português brasileiro. **Signo y Seña**, Buenos Aires, n. 22, p. 119-138, 2012.

AMEKA, F. Interjections: the universal yet neglected part of speech. **Journal of Pragmatics**, North-Holland, v. 18, p. 101-118, 1992.

ANDERSON, S. R. Where’s morphology? **Linguistic Inquiry**, Cambridge, v. 13, n. 4, p. 571-612, 1982.

AUGUSTO, M.; CORRÊA, L. S. Marcação de gênero, opcionalidade e generacidade: processamento de concordância de gênero no DP aos dois anos de idade. **Linguística**, Rio de Janeiro, v. 1, n. 2, p. 207-234, 2005.

BAGETTI, T.; CORRÊA, L. M. S. The early recognition of verb affixes: evidence from Portuguese. *In*: ANNUAL BOSTON UNIVERSITY CONFERENCE ON LANGUAGE DEVELOPMENT, 35, 2011, Boston. **Proceedings...** Boston, ABUCLD, 2011. p. 450-462.

BORGES NETO, J. Flexão e derivação: será que os tratamentos dados a esta distinção não adotam uma perspectiva 'enviesada'? **Cadernos de Estudos Linguísticos**, Campinas, v. 54, n. 2, p. 307-317, 2012.

BORGES NETO, J. Sobre a noção de palavra. **Muitas Vozes**, Ponta Grossa, v. 9, n. 1, p. 10-26, 2020.

BRUENING, B. The lexicalista hypothesis: both wrong and superfluous. **Language**, Washington, v. 94, n. 1, p. 1-42, 2018.

CAMARA JR, J. M. **Estrutura da língua portuguesa**: edição crítica. Petrópolis: Vozes, 2019 [1970].

CHOMSKY, N. Remarks on nominalization. *In*: JACOBS, R.; ROSENBAUM, P. (ed.). **Readings in English transformational grammar**. Waltham: Ginn and Company, 1970. p 184-221.

CHOMSKY, N. **The Minimalist Program**. Cambridge: MIT, 1995.

CHOMSKY, N. Three factors in language design. **Linguistic Inquiry**, Cambridge, v. 36, n. 1, p. 1-25, 2005.

CLARK, E. Morphology in language acquisition. *In*: SPENCER, A.; ZWICKY, A. (ed.). **The handbook of morphology**. Cambridge: Blackwell, 2007. p. 374-379.

CLARK, E. Acquisition of derivational morphology. *In*: LIEBER, R.; ŠTEKAUER, P. (ed.). **The Oxford handbook of derivational morphology**. Oxford: Oxford, 2014. p. 424-439.

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

CORRÊA, L. M. S.; AUGUSTO, M. R. A. Primeiros passos na aquisição da sintaxe: o sintagma nominal. *In*: FREITAS, M. J.; SANTOS, A. L. (ed.). **Aquisição de língua materna e não materna**: questões gerais e dados do português. Berlin: Language Science, 2017. p. 121-154.

COSTA, J.; SANTOS, A. L. **A falar como os bebês**. Lisboa: Caminho, 2003.

COSTA, J.; LOPES, R. E. V. Acquisition of Portuguese syntax. *In*: WETZELS, L.; COSTA, J.; MENUZZI, S. (ed.). **The handbook of Portuguese Linguistics**. Blackwell: Malden, 2016. p. 562-577.

DEL RÉ, A.; HILÁRIO, R. N. Limites e contribuições do uso da EME para pesquisas de cunho qualitativo na aquisição do PB. **Protolíngua**, João Pessoa, v. 8, n. 2, p. 121-144, 2013.

DI SCIULLO, M. A.; WILLIAMS, E. **On the definition of word**. Cambridge: MIT, 1987.

GOUT, A.; CRISTOPHE, A. O papel do *bootstrapping* prosódico na aquisição da sintaxe e do léxico. *In*: CORRÊA, L. S. (org.). **Aquisição da linguagem e problemas do desenvolvimento linguístico**. São Paulo: Loyola, 2006. p. 103-128.

GROLLA, E.; FIGUEIREDO SILVA, M. C. **Para conhecer aquisição da linguagem**. São Paulo: Contexto, 2014.

GUASTI, M. T. **Language acquisition**: the growth of grammar. Cambridge: MIT, 2002.

HASPELMATH, M. The indeterminacy of word segmentation and the nature of morphology and syntax. **Folia Linguistica**, Berlin, v. 45, n. 1, p. 31-80, 2011.

LAZZARINI-CYRINO, J. P.; MATTOS, E. B. Do morfema para a palavra: padrões de associação entre morfemas a partir de métodos estatísticos. *In*: CONGRESSO NACIONAL EM ESTUDOS INTERDISCIPLINARES DA LINGUAGEM, 1, e-book, 2020.

MARANTZ, A. Sem escapatória da sintaxe: não tente fazer análise morfológica na privacidade do seu próprio léxico. **ReVEL**, Porto Alegre, v. 13, n. 24, p. 8-33, 2015 [1997].

MARANTZ, A. **Words** (manuscrito), 2001.

MASSINI-CAGLIARI, G. Adaptação fonológica de nomes próprios de origem estrangeira: comparação entre português arcaico e português brasileiro. **Estudos Linguísticos**, São Paulo, v. 40, n. 2, p. 795-807, 2011.

PETTER, M. Morfologia. *In*: FIORIN, J. L. (org.). **Introdução à Linguística: princípios de análise**. São Paulo: Contexto, 2008. p. 59-79.

QUADROS, R. M. O paradigma gerativista e a aquisição da linguagem. *In*: FINGER, I.; QUADROS, R. M. (org.). **Teorias de aquisição da linguagem**. Florianópolis: UFSC, 2008. p. 45-82.

RESENDE, M. S. Elipse morfológica: uma abordagem sintática para a formação de palavras. **Estudos Linguísticos e Literários**, Salvador, n. 57, p. 341-365, 2017.

RESENDE, M. S. Bebês são não lexicalistas: considerações sobre o estatuto de 'palavra' na aquisição da linguagem. **Diadorim**, Rio de Janeiro, 2021.

SANTOS, A. L.; LOPES, R. E. de V. Primeiros passos na aquisição de sintaxe: direcionalidade, movimento do verbo e flexão. *In*: FREITAS, M. J. (ed.). **Aquisição de língua materna e não materna: questões gerais e dados do português**. Berlin: Language Science, 2017. p. 155-176.

SANTOS, R. S.; SCARPA, E. M. A aquisição da morfologia verbal e sua relação com o acento primário. **Letras de Hoje**, Porto Alegre. v. 38, n. 4, p. 249-260, 2003.

SCARPA, E. M. Aquisição da linguagem. *In*: MUSSALIM, F.; BENTES, A. C. (org.). **Introdução à Linguística: domínios e fronteiras**. 8. ed. São Paulo: Cortez, 2012. p. 241-271.

SCLIAR-CABRAL, L. New perspectives on the reading and early literacy processing. **Cadernos de Linguística**, Campinas, v. 2, n. 1, p. 1-14, 2021.

VILLALVA, A. Palavras, que as há. **Revista de Estudos da Linguagem**, Belo Horizonte, v. 20, n. 2, p. 125-139, 2012.

YANG, C.; ROEPER, T. Minimalism and language acquisition. *In*: BOECKX, C. (ed.). **The Oxford handbook of linguistic Minimalism**. Oxford: Oxford, 2011.

- | As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem

---

COMO CITAR ESTE ARTIGO: RESENDE, Mauricio. As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem. **Revista do GEL**, v. 18, n. 2, p. 160-159, 2021. Disponível em: <https://revistadogel.gel.org.br/>

Submetido em: 11/05/2021 | Aceito em: 29/06/2021.

---