

# LISTA 3: ENCICLOPÉDIA

Paula Roberta Gabbai Armelin

Vitor Augusto Nóbrega

## Conceitos-chave do capítulo:

- Conteúdo conceitual
- Interpretação contextual de raízes
- Polissemia e idiomaticidade
- Enciclopédia e a arquitetura da Gramática

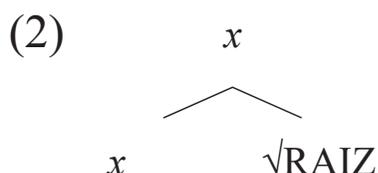
Este capítulo é dedicado à apresentação da Lista 3, também conhecida como Enciclopédia, um dos componentes assumidos pelo modelo de Gramática proposto pela Morfologia Distribuída. Em linhas gerais, a Lista 3 armazena os significados de natureza conceitual, sejam eles “literais” ou idiossincráticos, que serão atribuídos às raízes a partir do ambiente sintático em que elas estão inseridas. Em comparação aos modelos lexicalistas, é possível dizer que a Lista 3 é responsável por armazenar o tipo de informação semântica que, no lexicalismo, estava atrelada ao léxico.

Na Morfologia Distribuída, um componente como a Lista 3 é necessário na medida em que o modelo assume como premissa a decomposição integral de todo objeto linguístico, em especial, a decomposição de objetos complexos com conteúdo não composicional, tais como palavras derivadas, palavras compostas e expressões idiomáticas. Como veremos com mais detalhes no capítulo *Domínios de localidade na interpretação semântica*, essa assunção está codificada no princípio de *decomposição plena*, definido mais diretamente em Embick (2015, p. 21), apresentado em (1):

## (1) Decomposição plena

Não há objetos complexos armazenados em nossa memória. Todo objeto complexo deve ser derivado pela Gramática toda vez que for empregado<sup>27</sup> (adaptado de Embick, 2015, p. 21, tradução nossa).

A assunção desse princípio descarta o armazenamento de unidades lexicais para além dos primitivos assumidos pelo modelo, notadamente, raízes e traços morfossintáticos. Assim sendo, qualquer objeto linguístico formado por mais de um morfema será necessariamente derivado pela sintaxe, ainda que esse objeto envolva um significado idiossincrático. Para o modelo, isso implica que mesmo palavras aparentemente simples, como *flor* e *café*, devem ser derivadas toda vez em que forem empregadas, já que elas são resultado da concatenação entre uma raiz e um núcleo categorizador, como ilustrado na estrutura em (2), em que *x* está para os categorizadores *n*, *v* ou *a*:



Ainda que derivada de maneira composicional, a estrutura em (2) está sujeita a receber uma interpretação idiossincrática. É válido destacar que foi a interpretação idiossincrática, em especial, de alguns tipos de nominalizações, uma das principais motivações para se postular que sua formação deveria ocorrer em um componente independente, o léxico – caracterizado como o espaço das idiossincrasias –, e não na sintaxe propriamente dita, que, desse ponto de vista, deveria se ocupar somente do que

<sup>27</sup> No trecho original: “No complex objects are stored in memory; i.e., every complex object must be derived by the grammar”.

é considerado previsível e sistemático (CHOSMKY, 1970, 1981; HALLE, 1973; DI SCIULLO; WILLIAMS, 1987; SCALISE, 1984). Dada essa cisão de papéis atribuídos ao léxico e à sintaxe, toda teoria que tente reintegrar a formação de palavras à computação sintática, tal como faz a Morfologia Distribuída, precisa explicar de que forma um sistema computacional composicional, como a sintaxe, é capaz de gerar objetos linguísticos com conteúdos inesperados, não composicionais.

Ainda, é importante salientar que assumir uma teoria derivacional não lexicalista para a formação de palavras não significa renunciar à ideia de que significados arbitrários precisam ser, de alguma maneira, listados. Logo, uma vez que a Morfologia Distribuída não descarta a ideia de listagem e, ao mesmo tempo, rejeita a assunção de um léxico como repositório de idiossincrasias, esse modelo deve esclarecer: (i) como significados inesperados são atribuídos às estruturas derivadas composicionalmente pela sintaxe e (ii) em que lugar da Gramática tais significados idiossincráticos estão armazenados. Como veremos neste capítulo, para a Morfologia Distribuída, as irregularidades semânticas observadas em palavras – e também em sintagmas – decorrem da articulação entre uma classe de primitivos, as raízes, e uma lista de significados que cada raiz pode apresentar, considerando os ambientes sintáticos em que estão inseridas. Dessa forma, o repertório de significados associados a cada raiz de uma língua, bem como dos contextos estruturais em que esses significados aparecem, é o que chamamos de Lista 3.

Para apresentar o papel da Lista 3, exploramos, neste capítulo, dois fenômenos que, geralmente, estão atrelados a irregularidades semânticas, a saber: (i) a polissemia das raízes que, em diferentes ambientes estruturais, podem veicular significados muito distintos, e (ii) a idiomaticidade de expressões complexas, como as expressões idiomáticas, cujo significado global não é recuperável a partir da composição do significado de suas partes constituintes, como na

expressão *bater as botas*, que corresponde a “morrer”. Este capítulo também coloca em discussão o local de acesso da Lista 3 na arquitetura de Gramática assumida pelo modelo. Tendo em vista que a Lista 3 organiza conteúdos relevantes para a interpretação dos objetos gerados pela sintaxe, assume-se que ela deve estar necessariamente alocada em seu ramo interpretativo, a Forma Lógica, comumente referida como LF. Veremos, no entanto, que, ao longo dos anos, há variações no local de acesso à Lista 3, em diferentes propostas, que resultam diretamente das diferentes visões a respeito da natureza das raízes, ainda um tema de intenso debate na literatura.

## **1. O conteúdo conceitual e a interpretação contextual das raízes**

Para compreendermos a natureza dos conteúdos associados à Lista 3, é necessário estabelecermos uma distinção entre dois tipos de significados que são relevantes para a interpretação das expressões linguísticas: um propriamente linguístico, derivado a partir da organização de traços morfossintáticos em arranjos estruturais dispostos hierarquicamente, e outro extralinguístico, que não apresenta propriedades gramaticais ou formais, e corresponde, em grande parte, ao conhecimento de mundo dos falantes de uma língua, ou seja, compreende um conhecimento de caráter enciclopédico.

A distinção entre esses dois tipos de significados não é uma novidade da Morfologia Distribuída. Ela é frequentemente recuperada na literatura morfológica e sintática de diversas maneiras e com nomenclaturas variadas. Grimshaw (1993), por exemplo, faz uma distinção entre “estrutura semântica” e “conteúdo semântico”, sendo a primeira noção relacionada a aspectos estruturais, enquanto a segunda diz respeito aos significados lexicais, os quais podem ser idiossincráticos. Rappaport Hovav

e Levin (1998), na mesma direção, estabelecem uma distinção entre o que chamam de “componente estrutural” e “componente idiossincrático” do significado, em que o primeiro interage com a sintaxe, ao passo que o segundo é irrelevante para a computação sintática. Distinções semelhantes podem também ser verificadas em Lieber (2004).

Em geral, nas visões gerativistas de Gramática, podemos dizer que esses dois tipos de significados são processados a partir de estratégias distintas. Mais especificamente, o primeiro está relacionado a aplicações funcionais operadas por LF, que atuam sobre as informações codificadas pelos traços morfossintáticos e sobre os aspectos estruturais que compõem a interpretação de uma expressão linguística. A sintaxe deriva estruturas cuja interpretação é processada composicionalmente a partir do conteúdo codificado em seus núcleos funcionais.<sup>28</sup> Nesse sentido, o significado operado por LF segue o princípio da composicionalidade – creditado ao lógico e filósofo Gottlob Frege –, que constrói a interpretação das expressões linguísticas a partir da maneira como seus traços formais foram organizados pelo componente sintático.

O segundo tipo de significado é arbitrário e convencionalizado, não desempenhando um papel direto no funcionamento do sistema combinatorial, ou seja, ele não influencia as operações sintáticas na formação de um objeto linguístico. A inoperância sintática desse significado é um dos motivos pelos quais a Morfologia Distribuída postula que ele não precisa estar especificado pré-sintaticamente, diferentemente do que propõem as abordagens lexicalistas. O significado enciclopédico, que não é derivado a partir dos traços morfossintáticos contidos nos núcleos funcionais, está armazenado na forma de entradas enciclopédicas alocadas na Lista 3. Para termos uma noção mais clara desse significado, retomamos as

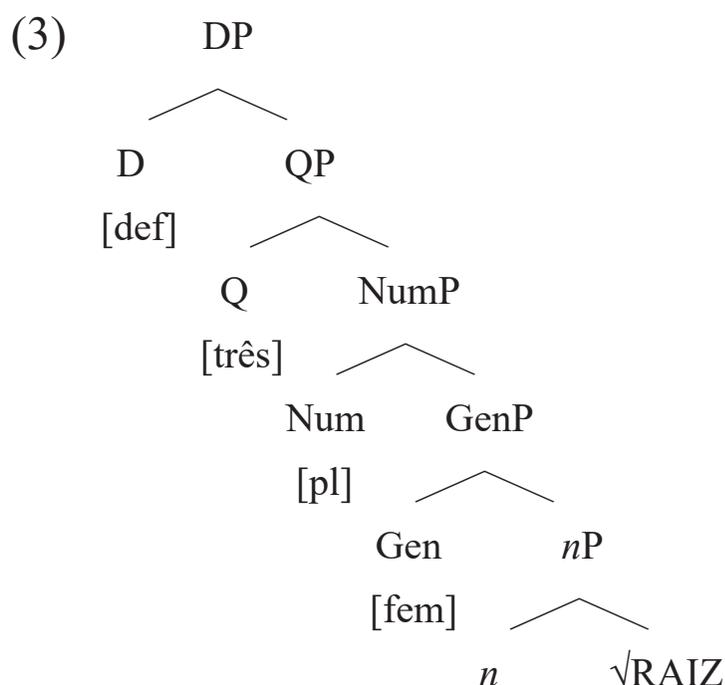
---

28 Como exemplo do modo como aplicações funcionais operam na composição do significado das sentenças, ver Heim e Kratzer (1998).

observações de Marantz (1996, p. 3), que busca explicitar a natureza do significado conceitual e arbitrário articulado às raízes e destacar sua irrelevância no funcionamento sintático:

[...] a diferença semântica entre “cachorro” e “gato” não desencadeia nenhuma operação ou restrição sintática ou nenhum princípio semântico composicional; então, qualquer que seja o traço que distingue “cão” e “gato”, ele não estaria relacionado aos traços dos itens lexicais<sup>29</sup> (MARANTZ, 1996, p. 3, tradução nossa).

O ponto levantado por Marantz, no trecho acima, pode ser ilustrado com a estrutura em (3).



O conteúdo dos traços morfossintáticos que compõem a estrutura em (3) promove uma interpretação invariável do arranjo derivado, que, em português, corresponderia a algo como “As três [n ]-(e)s”. Ainda, esses traços devem estar disponíveis para as operações sintáticas, a fim de permitir, por exemplo, que a informação de gênero e número, presentes nos núcleos Gen e

29 No trecho original: “[...] the semantic difference between “dog” and “cat” drives no syntactic or compositional semantic principle, rule, or constraint; thus, whatever feature of “dog” that distinguishes “dog” from “cat” would not be a feature of Lexical items”.

Num, respectivamente, sejam copiadas no núcleo determinante, D, influenciando sua expoência fonológica. Por outro lado, a realização da raiz, em (3), como  $\sqrt{\text{CACHORR}}$  ou  $\sqrt{\text{GAT}}$ , em nada afeta a composição estrutural e semântica desses traços morfossintáticos. Logo, é razoável admitir que a lista de significados enciclopédicos atribuídos a uma estrutura sintática seja somente acessada no momento em que a estrutura já esteja formada.<sup>30</sup>

Isso não significa dizer, entretanto, que a estrutura sintática é insignificante na determinação de qual será o conteúdo enciclopédico atribuído a uma raiz, pelo contrário. Um núcleo categorizador concatenado diretamente a uma raiz, por exemplo, tem a capacidade de influenciar seu significado, delimitando um domínio de interpretação idiossincrática. Isto é, a depender da categoria sintática associada a uma raiz, o conteúdo atribuído poderá variar – o que, em certo sentido, reforça a hipótese de que as raízes são fundamentalmente acategoriais. A esse respeito, Marantz (2001, p. 12) destaca que

A interação entre a semântica de uma raiz e a semântica dos núcleos categorizadores que formam nomes, verbos e adjetivos determinará quão boa será uma combinação entre uma raiz e um núcleo [categorizador]<sup>31</sup> (MARANTZ, 2001, p. 12, tradução nossa).

Com isso em mente, tomemos a raiz  $\sqrt{\text{GAT}}$  do português. É possível notar que essa raiz é capaz de participar de formas igualmente gramaticais em contextos nominais e adjetivais, e até mesmo verbais, cada qual veiculando diferentes interpretações, como exemplificado em (4).

- (4) a. O  $[[\sqrt{\text{GAT}}]_{\text{N}}]$ -o miou a noite inteira.  
 b. O amigo da Maria é muito  $[[\sqrt{\text{GAT}}]_{\text{A}}]$ -o.

30 A porção da estrutura relevante para o acesso à Enciclopédia será mais bem discutida no capítulo *Domínios de localidade na interpretação semântica*.

31 No trecho original: “*The interaction of root semantics and the semantics of the heads that create nouns, verbs and adjectives determines how good a combination of a root and such a head will be*”.

- c. O João [[√GAT]-*ific*-<sub>v</sub>]-*ou* a casa toda.  
 d. O João [[√GAT]-*e*-<sub>v</sub>]-*ou* pela casa por horas.

Enquanto em (4a) temos a interpretação de *gato* como uma entidade, correspondente a um felino, mamífero, que é adotado como animal de estimação; em (4b), a interpretação é adjetival e pode ser descrita como *bonito*, ou seja, trata-se da atribuição de uma propriedade a um indivíduo, no caso, *o amigo da Maria*. Essa variação interpretativa nos aponta que, tanto o conteúdo mais convencional – aqui atrelado ao ambiente nominal –, quanto o conteúdo mais idiomático de uma raiz devem estar especificados na Lista 3.

Já em (4c), temos a formação de um verbo que pode não ter um significado óbvio e mesmo um uso tão comum, mas que é passível de uma interpretação contextual, como, por exemplo, a denotação de um “processo de adequação de um ambiente para gatos, incluindo a instalação de tocas e prateleiras”. É interessante notar também que (4c) pode ser bastante usual para um grupo particular de pessoas; por exemplo, aquelas que têm gatos como animais de estimação. Isso nos aponta que o conhecimento enciclopédico reflete uma certa variação individual, determinada pelas experiências particulares que temos no mundo. Além disso, esse conhecimento não está circunscrito ao período crítico da aquisição da linguagem, como é o caso do conteúdo veiculado pelos traços morfossintáticos, uma vez que podemos, a partir de novas experiências, adquirir e manipular novos conteúdos enciclopédicos.

Em alguns casos, a ocorrência de uma raiz em um ambiente categorial ainda pouco comum pode levar a uma interpretação pragmaticamente anômala, embora a estrutura em si seja bem formada do ponto de vista sintático, como ocorre em (4d). Marantz (2001, p.12), ao recorrer a um exemplo paralelo a (4d), aponta que seria possível imaginar um contexto em que João estaria imitando

o comportamento de um gato, por exemplo, miando e arranhando os móveis, para chamar a atenção de alguém. Dessa forma, uma raiz pode ser realizada em um contexto estrutural que leve a uma interpretação incompatível com o nosso conhecimento de mundo. Aliás, o clássico exemplo de Chomsky (1957, p. 15), em (5a), traduzido em (5b), já apontava para a autonomia da sintaxe em relação ao significado enciclopédico das palavras.

- (5) a. *Colorless green ideas sleep furiously.*  
 b. Ideias verdes incolores dormem furiosamente.

Em (5), a boa formação sintática independe do fato de a expressão formada ser compatível ou não com uma interpretação plausível, dado nosso conhecimento de mundo.<sup>32</sup> Um exemplo semelhante é oferecido em Siddiqi (2009), com a sentença abaixo:

- (6) a. #O sanduíche de presunto comeu o cachorro.

A estranheza de (6), graficamente indicada por #, está na incompatibilidade entre as interpretações geralmente atribuídas às raízes e os contextos estruturais em que elas foram inseridas, notadamente, a relação de predicação inusitada ali estabelecida. Para Siddiqi (2009, p. 93):

A Enciclopédia marca essa sentença como malformada [...]. No entanto, podemos ajustar nosso conhecimento do mundo real para que ele se adeque à sentença ao interpretar o *sanduíche de presunto* de forma que um sanduíche de presunto seja, de algum modo, capaz de comer um cachorro. Também podemos ajustar o significado de *comer* em vez do significado de seus argumentos<sup>33</sup> (SIDDIQI, 2009,

32 Para uma discussão aprofundada a respeito do conceito de aceitabilidade, em particular, em oposição à ideia de agramaticalidade, ver Guimarães (2017, Capítulo 2).

33 No trecho original: “*The Encyclopedia marks this sentence as ill-formed because ham sandwiches do not make good eaters. However, we can adjust our real world knowledge to fit the sentence by interpreting ham sandwich in such a way that a ham sandwich is somehow capable of eating a dog. We can also adjust the meaning of eat rather than the meaning of its arguments*”.

p. 93, tradução nossa).

Essa articulação entre o conteúdo estrutural e o conteúdo enciclopédico das expressões linguísticas é perfeitamente sintetizada por Borer (2005, p. 108), ao assinalar que o conteúdo enciclopédico “não apresenta nenhuma propriedade formal, e é equivalente, [...] a um ‘material’ que é despejado no molde estrutural a ser atribuído a propriedades gramaticais”, bem como vimos na descrição da estrutura em (3), sendo o molde estrutural, nesse caso, as raízes.

Nesse sentido, a postulação da Lista 3 auxilia na distinção entre morfemas funcionais e raízes assumida pelo modelo da Morfologia Distribuída. A esse respeito, especificamente, Harley (2019, p. 162) aponta que, enquanto as raízes apresentam uma entrada enciclopédica, os morfemas funcionais não o fazem:

As raízes podem ser especificadas para interpretação idiomática, às vezes restritas por contextos sintáticos e semânticos bastante elaborados, enquanto os morfemas funcionais não podem. Morfemas funcionais devem contribuir com suas denotações padrão com qualquer estrutura em que eles são encontrados<sup>34</sup> (HARLEY, 2019, p. 162, tradução nossa).

Além disso, outra distinção de conteúdo entre morfemas funcionais – chamados neste manual de traços morfossintáticos (veja o *Capítulo Lista 1: Traços morfossintáticos e raízes* para lembrar) – e raízes pode ser feita: enquanto os primeiros não estão sujeitos à polissemia, as raízes representam uma potencialidade de significados, podendo, em tese, ser vinculadas a novos significados de modo ilimitado. Sua natureza inerentemente polissêmica é o que demanda a assunção de uma lista dedicada a esses conteúdos. Podemos dizer, em resumo, que nossa Faculdade da Linguagem, em um sentido mais amplo, é capaz de associar estruturas linguísticas a

<sup>34</sup> No trecho original: “*Roots can be specified for idiomatic interpretation, sometimes restricted to very elaborate interdependent syntactic and semantic contexts, while functional morphemes cannot. Functional morphemes must contribute their standard denotation to any structure which they find themselves in*”.

conteúdos enciclopédicos, vinculando o significado composicional que emerge do arranjo estrutural de traços morfossintáticos ao significado extralinguístico veiculado pelas raízes, armazenado através de instruções na Lista 3.

Para exemplificar a assimetria entre estrutura e conteúdo enciclopédico, exploramos, na próxima seção, a interpretação de expressões idiomáticas.

## 2. Lista 3 e idiomaticidade

As expressões idiomáticas podem ser definidas como um constituinte – ou uma série de constituintes – para os quais a interpretação semântica não é uma função composicional de suas partes (MATEU; ESPINAL, 2007). Exemplos de expressões idiomáticas no português são, além de *bater as botas*, *esconder o ouro*, *abrir o jogo*, *acabar em pizza*, *chutar o pau da barraca*, *pisar na bola*, *tirar água do joelho*, *quebrar um galho*, entre inúmeras outras.

A natureza não composicional dessas construções serviu como o principal argumento para abordagens lexicalistas sugerirem que elas deveriam estar listadas de modo integral em um léxico pré-sintático, sendo inseridas na sintaxe como um bloco, semelhantemente a um item lexical, isto é, como um elemento de nível mínimo. Isso impedia, conseqüentemente, que a estrutura interna dessas formações fosse sintaticamente derivada, tendo em vista que a sintaxe atua na formação de construções regulares, não devendo se ocupar, portanto, da formação de objetos linguísticos com conteúdos não composicionais (CHOMSKY, 1980, 1981; DI SCIULLO; WILLIAMS, 1987; JACKENDOFF, 1997, 2002, dentre outros). Logo, expressões idiomáticas operariam de modo equivalente a palavras complexas que, ao serem inseridas no componente sintático, se comportariam como um item atômico, sintática e semanticamente.

Isso é o que se observa na análise sugerida por Chomsky (1981, p. 146). O autor busca manter a estrutura interna das expressões idiomáticas, tal como *kick the bucket* (lit. chutar o balde, ‘morrer’), assumindo que um nó V(erb) está sobre uma estrutura complexa listada no léxico, a qual será inserida em uma posição V na sintaxe, tal como representado em (7).

(7) [<sub>V</sub>[<sub>VP</sub>[<sub>V</sub> *kick*] [<sub>NP</sub>[<sub>Det</sub> *the*] [<sub>N</sub> *bucket*]]]]

Nunberg, Sag e Wasow (1994) mostram, entretanto, que nem todas as expressões idiomáticas são atômicas, tal como *kick the bucket*, em (7), pois alguns tipos podem carregar certa composicionalidade entre seus constituintes. Para os autores, há, pelo menos, dois tipos de expressões idiomáticas a serem consideradas, a saber: (i) expressões combinadas de modo idiomático (ECIs) e (ii) sintagmas idiomáticos (SIs). As ECIs, embora associadas a conteúdos não convencionais, têm seu significado distribuído entre seus constituintes, como ocorre com a expressão *engolir sapos*, em que *engolir* está associado, sugestivamente, a “enfrentar”, e *sapos* a “problemas”. As SIs, por sua vez, não têm seu significado distribuído por seus constituintes, mas um significado atômico, como é o caso de *kick the bucket* e de *bater as botas*, em que a expressão idiomática, como um todo, expressa o significado “morrer”.

A fim de tornar a diferença entre esses dois tipos mais evidente, os autores recorrem a cinco testes sintáticos para identificar uma ECI: modificação, quantificação, topicalização, elipse e anáfora. Enquanto ECIs admitem intervenções sintáticas em sua estrutura interna, sem perder seu conteúdo idiossincrático, SIs tornam-se agramaticais nesses contextos, como pode ser visto, de modo ilustrativo, nos contrastes de (8) a (10):

- (8) Modificação
- a. Eu engoli um baita sapo.
  - b. \*Eu bati uma baita bota.
- (9) Quantificação
- a. Eu tive que engolir dois sapos.
  - b. \*Eu tive que bater duas botas.
- (10) Topicalização
- a. Sapos, eu já engoli hoje.
  - b. \*Botas, ele bateu hoje.

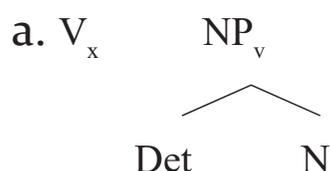
As ECIs nos indicam que uma solução como a de Chomsky (1981), no entanto, não é capaz de explicar sua flexibilidade sintática, como visto nos testes de (8) a (10).

Jackendoff (1997, 2002), por sua vez, procura abarcar diretamente as diferenças estruturais presentes nesses dois tipos de expressões idiomáticas. Opostamente ao tratamento fornecido pela visão gerativista padrão, ilustrada em (7), o autor endossa a argumentação de Nunberg, Sag e Wasow (1994), ao afirmar que as expressões idiomáticas não são periféricas, mas parte da linguagem humana, já que contêm uma estrutura fonológica, sintática e conceitual. Esse posicionamento exige uma explicação das expressões idiomáticas como objetos da Gramática, e não como itens armazenados na memória, tal como conhecimento pragmático, fatos da história e habilidades práticas (JACKENDOFF, 1997, p. 156). Nesse sentido, Jackendoff propõe que as expressões idiomáticas — sejam elas ECIs ou SIs — estão listadas sintaticamente no léxico como sintagmas verbais comuns (VPs), através de “estruturas conceituais lexicais” (do inglês, *lexical conceptual structure*, LCS).

Em sua proposta, ECIs carregam uma estrutura interna

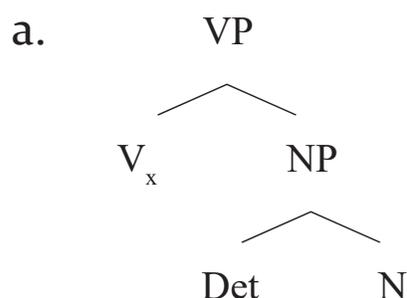
flexível, já que suas partes são identificadas separadamente, como se verifica no VP listado no léxico, descrito em (11), em que a expressão idiomática do inglês *bury the hatchet* (lit. ‘fazer as pazes’) será interpretada composicionalmente, notadamente: *bury* ‘enterrar’, interpretado como “reconciliar”, e *hatchet* ‘machado’, como “desentendimento”. Por outro lado, SIs, tal como *kick the bucket* ‘morrer’, embora sejam igualmente uma unidade de nível mínimo, conterão uma estrutura VP fixa, cujo NP objeto não é unificado como um argumento em sua LCS, tal como representado em (12b):

(11) *bury the hatchet* lit. enterrar o machado ‘fazer as pazes’



b. LCS: [RECONCILIAR ([ ]<sub>A</sub>, DESENTENDIMENTO)]<sub>y</sub>]<sub>x</sub>

(12) *kick the bucket* lit. chutar o balde ‘morrer’



b. LCS: [MORRER ([ ]<sub>A</sub>)]<sub>x</sub>

Jackendoff (1997, 2002) sugere que uma estrutura sintática se torna idiomática pelo fato de que nem todos os seus constituintes sintáticos, em (11a) e (12a), correspondem aos seus constituintes

conceituais, em (11b) e (12b). Em sua análise, o que notamos é uma exigência de se listar estruturas quase-sentenciais, sintaticamente endocêntricas, em um léxico pré-sintático, em especial a estrutura das SIs, como em (12). Tal posicionamento explicita uma constante nas análises lexicalistas, que é o de suprimir a estrutura interna de objetos linguísticos a fim de abarcar seu significado não composicional pré-sintaticamente. Um problema adicional para essas abordagens reside no modo como itens lexicais são inseridos na estrutura sintática, dado que as expressões idiomáticas supramencionadas são tratadas como itens lexicais sintagmáticos, ou seja, itens lexicais maiores que um único nó sintático. Além disso, um terceiro problema está em dar conta de sua endocentricidade, tendo em vista que são construídas como sintagmas ou orações convencionais da língua, e, paralelamente, devem estar alocadas no componente lexical.

No entanto, esse tipo de abordagem para as expressões idiomáticas não seria razoável em um modelo como a Morfologia Distribuída, especialmente se considerarmos a assunção da decomposição plena, em (1), a qual impede que elementos complexos sirvam como átomos sintáticos. Na verdade, a própria existência de objetos complexos com significados arbitrários, como as expressões idiomáticas, parece apontar para a fragilidade da hipótese de que exista uma separação categórica entre a fronteira das palavras e dos sintagmas. Isso porque, como vimos na seção anterior, significados especiais também emergem no nível da palavra como resultado da associação entre significados enciclopédicos e contextos sintáticos específicos.

Como alternativa, a Morfologia Distribuída busca explicar o significado arbitrário das expressões idiomáticas através do mesmo mecanismo que licencia o conteúdo conceitual no âmbito da palavra, notadamente: como resultado da associação entre o ambiente sintático que abriga um conjunto de raízes e uma entrada

enciclopédica que relaciona um significado particular a esse contexto específico. Aliás, significados não composicionais podem surgir em variados níveis, como nas palavras, nos sintagmas e até mesmo em sentenças inteiras, como podemos ver nos exemplos a seguir:

- (13) a. Palavra: *barbeiro* significado: “motorista ruim”  
 b. Sintagma: *bater as botas* significado: “morrer”  
 c. Sentença: *A vaca foi para o brejo* significado: “dar errado”

Além disso, é importante destacar que os significados ditos “literais” das raízes que compõem os dados em (13) estão, da mesma forma, disponíveis para serem atribuídos a essas formações. Por exemplo, em um contexto apropriado, é possível interpretar (13a) como “profissional que faz barbas”; (13b), como um “evento de sacudir as botas” e, (13c), como a descrição de um evento em que uma dada vaca tenha, de fato, se dirigido até um brejo. Se expressões idiomáticas são estocadas no léxico, como propõem as abordagens lexicalistas, então, seria preciso assumir tratamentos muito diversos para uma e outra interpretação.

Na Morfologia Distribuída, por outro lado, a estrutura interna derivada pela sintaxe para essas formações seria a mesma, independentemente do significado atribuído às suas raízes. As diferentes interpretações – idiomáticas ou mais “literais” – ficam, por outro lado, a cargo das diferentes porções da estrutura sintática a que o conteúdo enciclopédico é atribuído. Nessa perspectiva, as abordagens sintáticas de formação de palavras têm focado em delimitar um domínio sintático bem definido de estrutura dentro do qual uma interpretação não composicional pode ser licenciada, como veremos em detalhes no capítulo *Domínios de localidade na interpretação semântica*.

Aliás, a própria presença de certa composicionalidade entre os constituintes das ECIs, como mostraram Nunberg, Sag e Wason (1994), sugere uma inadequação das abordagens lexicalistas. Avançando essa perspectiva, McGinnis (2002) argumenta, por exemplo, que as propriedades aspectuais das expressões idiomáticas funcionam de modo bastante sistemático, sem que haja diferenças em relação à semântica aspectual das formações composicionais. Considerando que a informação aspectual das expressões linguísticas, de modo geral, compõe a sua organização sintática, então, também seria esperado que as propriedades aspectuais das expressões idiomáticas fossem composicionalmente derivadas. Dessa forma, o significado das expressões idiomáticas não deve ser visto como algo totalmente arbitrário. O componente estrutural do significado, pelo menos as propriedades aspectuais exploradas em McGinnis (2002), é sistemático e composicional. Esse tipo de evidência aponta para a relevância de propostas que derivam a estrutura sintática das expressões idiomáticas na sintaxe, tal como a Morfologia Distribuída. Vejamos os exemplos a seguir:

- (14) a. O João já vinha morrendo há semanas com aquela doença terrível.  
 b. ?O João já vinha batendo as botas há semanas com aquela doença terrível.

Se, como propõe a Morfologia Distribuída, há uma estrutura sintaticamente derivada associada às expressões idiomáticas, então essa estrutura deve igualmente estar sujeita a consequências semânticas, como a composicionalidade aspectual detectada por McGinnis (2002). O contraste entre os exemplos em (14) pode, então, ser explicado pelo fato de que *bater as botas* não se comporta sintaticamente como *morrer*, ainda que sua interpretação global idiomática seja essa. Isso porque a expressão idiomática é formada

por um verbo mais seu objeto direto e essa estrutura sintática apresenta consequências importantes (MARANTZ, 1997, p. 212) na comparação com *morrer*, por exemplo. Dessa forma, (14b) não pode ser interpretado idiomáticamente, restando apenas a interpretação literal, o que não é compatível com o restante da sentença. Tal perspectiva prevê, então, que há um componente composicional até mesmo para sintagmas verbais que tenham interpretação idiomática.

Em resumo, vimos que soluções lexicalistas que envolvem a listagem de construções complexas no léxico sinalizam, na verdade, que as fronteiras entre a estrutura da palavra e da sentença são tênues. Opostamente ao tratamento periférico fornecido pela visão lexicalista, a Morfologia Distribuída requer uma explicação das expressões idiomáticas como objetos gerados pela Gramática, não como itens armazenados na memória, e que domínios estruturais sujeitos a irregularidades semânticas são previstos pela própria Gramática, tal como veremos no capítulo *Domínios de localidade na interpretação semântica*, ao explicar a derivação de expressões idiomáticas no modelo. Assim, a Morfologia Distribuída deriva o significado das expressões idiomáticas considerando que há uma atribuição de conteúdo conceitual à porção da estrutura sintática envolvida na interpretação não composicional, sem necessariamente se ocupar da listagem de itens internamente complexos.

### **3. O lugar da Lista 3 no modelo da Morfologia Distribuída**

Como vimos nas seções anteriores, a Lista 3 armazena o conjunto de instruções que mapeiam os objetos gerados pelo sistema computacional em conteúdos enciclopédicos. Essa atribuição de conteúdo à estrutura sintática está necessariamente atrelada a dois aspectos cruciais para discutirmos o local de acesso da Lista 3 no

modelo da Morfologia Distribuída.

O primeiro aspecto é o de que a Lista 3 é necessariamente pós-sintática, uma vez que o significado conceitual não faz parte dos primitivos que compõem a Lista 1, sendo irrelevante para as operações do componente sintático, como discutimos anteriormente. O segundo aspecto relevante é que, considerada a bifurcação da derivação sintática para as interfaces de PF e LF, a Lista 3 precisa estabelecer uma conexão com o ramo de LF, uma vez que esse é o componente relacionado ao Sistema Conceitual-Intencional, o componente interpretativo da cognição humana com o qual a Faculdade da Linguagem faz interface. Em todas as propostas de representação da Lista 3 na arquitetura do modelo, esses dois aspectos são contemplados, ou seja, a Lista 3 é pós-sintaticamente posicionada, estando intimamente vinculada ao ramo de LF.

Por outro lado, o lugar da Lista 3 em relação ao ramo de PF não é consensual no modelo, estando sujeito a dois tipos de abordagens na literatura. Uma dessas abordagens assume que a Lista 3 tem acesso unicamente a LF, não havendo, portanto, qualquer relação entre a Lista 3 e PF. A segunda abordagem, por sua vez, aponta a necessidade de que a Lista 3 acesse, na verdade, tanto LF como PF. Crucialmente, a filiação teórica a uma ou outra posição depende, em linhas gerais, do que se assume em relação à natureza das raízes<sup>35</sup> ou, mais especificamente, em relação ao momento em que as raízes são propriamente individualizadas na formação de uma expressão linguística. Em outras palavras, o que está em questão é o ponto da derivação em que é possível identificar, de alguma forma, qual é exatamente a raiz que será empregada em uma expressão linguística.

Nas versões iniciais do modelo da Morfologia Distribuída

---

35 Uma discussão mais aprofundada a respeito da natureza da raiz como primitivo teórico é feita no capítulo *Lista 1: traços morfossintáticos e raízes*.

(HALLE; MARANTZ, 1993; MARANTZ, 1996), as raízes não eram individualizadas por nenhuma informação na Lista 1. Consequentemente, não é possível diferenciar uma raiz de outra nem na Lista 1, nem durante a computação no componente sintático. Dessa forma, as diferenças fonológicas ou de conteúdo conceitual, por exemplo, entre  $\sqrt{\text{GAT}}$  e  $\sqrt{\text{CACHORR}}$ , não estão presentes seja na Lista 1, seja na sintaxe.

Nessa perspectiva, as propriedades fonológicas das raízes são definidas apenas pós-sintaticamente através da operação de *Inserção de Vocabulário*<sup>36</sup>, que atribui conteúdo fonológico aos nós terminais da sintaxe. Portanto, raízes, assim como traços morfossintáticos, estão sujeitos à inserção tardia de fonologia. Especificamente em relação às raízes, sendo elementos destituídos de traços morfossintáticos, não é possível estabelecer uma competição entre elas em termos do *Princípio do Subconjunto* a fim de definir qual raiz deverá ser inserida. Por isso, Marantz (1996) propõe que o falante teria uma escolha livre, ou seja, uma opção não forçada pela derivação, a respeito de qual raiz seria inserida no nó sintático relevante.

Nesse sistema, consequentemente, a representação fonológica da raiz, definida através de *Inserção de Vocabulário*, também precisa ser considerada como informação relevante para a atribuição de conteúdo enciclopédico, como apontado por Marantz (1996, p. 17):

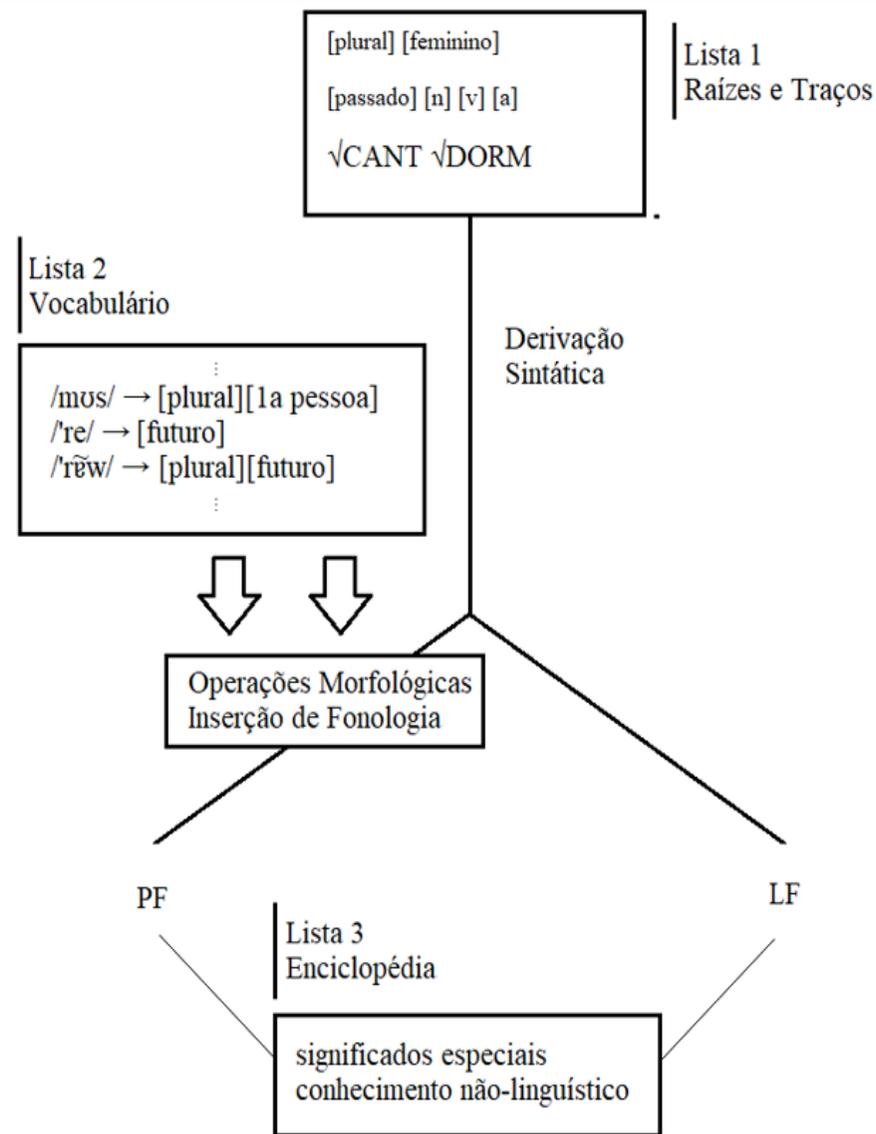
A inserção tardia envolve fazer uma afirmação específica sobre a conexão entre LF e a interpretação semântica. LF não pode, por si só, ser a entrada para a interpretação semântica. Se “gato” é inserido na fonologia em um nó no qual “cachorro” poderia muito bem ter sido inserido - e se, como assumimos, a diferença entre “gato” e “cachorro” faz diferença na interpretação semântica - então a representação fonológica, especificamente a escolha dos itens de Vocabulário, também deve ser uma entrada para a interpretação

<sup>36</sup> Para uma apresentação mais aprofundada desse mecanismo, indicamos a leitura do capítulo *Lista 2: Vocabulário*.

semântica<sup>37</sup> (MARANTZ, 1996, p. 17, tradução nossa).

A interpretação é, então, construída a partir do resultado de toda a derivação de uma expressão linguística, incluindo também os objetos decorrentes dos ramos de PF e LF, como ilustrado na arquitetura de Gramática a seguir:

Figura 1 – Arquitetura da Morfologia Distribuída.



Fonte: elaboração própria.

Nessa perspectiva de funcionamento do modelo, como bem

37 No trecho original: “Late insertion involves making a specific claim about the connection between LF and semantic interpretation. LF can’t by itself be the input to semantic interpretation. If “cat” is inserted in the phonology at a node at which “dog” could just as well have been inserted – and if, as we assume, the difference between “cat” and “dog” makes a difference in semantic interpretation – then the phonological representation, specifically the choice of Vocabulary items, must also be input to semantic interpretation”.

salienta Harley (2014), o acesso simultâneo às duas interfaces é necessário para evitar a possibilidade de que no ramo de PF ocorra a inserção de um item como  $\sqrt{\text{GAT}}$ , enquanto o conteúdo conceitual associado a outro elemento, por exemplo, CACHORRO, seja acessado no ramo de LF. Em outras palavras, o acesso simultâneo da Lista 3, tanto a PF, como a LF, garante que o item inserido no ramo de PF seja exatamente aquele considerado para fins de interpretação enciclopédica.

Uma proposta diferente a respeito do ponto de acesso da Lista 3 surge na literatura a partir da ideia de que é necessário individualizar as raízes já na Lista 1. Nessa abordagem, diferentes implementações foram desenvolvidas para delinear a identificação pré-sintática das raízes. Para Embick (2000), e trabalhos subsequentes, por exemplo, as raízes são diferenciadas uma das outras na Lista 1 pela sua representação fonológica subjacente, como apontado em Embick (2015, p. 23)<sup>38</sup>:

As raízes possuem uma representação fonológica como parte de sua constituição primitiva. Nesse sentido, então, a raiz  $\sqrt{\text{CAT}}$  em inglês é especificada com a matriz fonológica /kæt / desde o início<sup>39</sup> (EMBICK, 2015, p. 23, tradução nossa).

Essa representação fonológica subjacente é, em geral, suficiente para identificar a raiz. No entanto, um potencial desafio para essa individualização fonológica da raiz são os casos de homofonia, em que duas raízes distintas apresentam acidentalmente a mesma fonologia, como em *manga*, no português, cuja representação fonológica pode estar conceitualmente relacionada à interpretação de “parte de uma peça de roupa” ou “tipo de fruta”. Para solucionar essa questão, Embick (2015) propõe que as raízes homófonas,

38 Recomendamos o trabalho de Armelin, Gonçalves e Melo (2020), que discute em detalhes o trabalho de Embick (2000).

39 No trecho original: “*Roots possess a phonological representation as part of their primitive make-up. In this sense, then, the Root  $\sqrt{\text{CAT}}$  in English is specified with the phonological matrix /kæt/ from the beginning*”.

além de possuírem uma matriz fonológica como todas as outras raízes, também apresentam um índice, recurso que atua como um identificador não fonológico para individualizar essas raízes. Utilizando instruções alfanuméricas para representar essa identificação via índice, poderíamos ilustrar as raízes homófonas da seguinte forma na Lista 1:

- (15) a.  $\sqrt{\text{MANG}_{157}}$   
 b.  $\sqrt{\text{MANG}_{945}}$

Com essa identificação prévia das raízes na Lista 1, a atribuição de significado enciclopédico não precisa mais “checar” o resultado de PF, uma vez que as raízes, nessa perspectiva, não estão sujeitas à inserção tardia. Consequentemente, o acesso da Lista 3 à LF é suficiente para a atribuição do conteúdo conceitual à estrutura sintática.

O conteúdo fonológico, no entanto, não é a única maneira proposta na literatura para a individualização das raízes na Lista 1. Alguns autores, como Pfau (2000, 2009) Acquaviva (2009) e Harley (2014), propõem que as raízes são individualizadas exclusivamente por meio de um índice alfanumérico, sem qualquer representação fonológica. Para fins de uma ilustração um pouco mais concreta, nós podemos pensar nesse índice como o “endereço” das raízes, assim como os números das casas, por exemplo, identificam e diferenciam uma casa da outra em uma mesma rua.

Considerados os contextos estruturais em que as raízes estão inseridas, esse índice atua como uma instrução para a atribuição de conteúdo fonológico às raízes no ramo de PF e para a atribuição de conteúdo conceitual a tais elementos no ramo de LF, como pontua Harley (2014, p. 242-243):

Os elementos da Lista 1 que ocuparão o nó terminal raiz podem, portanto, ser notados como  $\sqrt{_{279}}$ ,  $\sqrt{_{322}}$ ,  $\sqrt{_{2588}}$ , etc. A Lista 2 consiste

em instruções para definir cada uma dessas entidades em um determinado contexto morfossintático. A Lista 3 consiste em instruções para interpretação dessas entidades em um determinado contexto morfossintático (HARLEY, 2014, p. 242-243, tradução nossa)<sup>40</sup>

Esse sistema torna os ramos de PF e LF bastante paralelos no modelo de Gramática, uma vez que a competição de Itens de Vocabulário para a inserção das raízes será definida em relação ao índice da mesma forma que a identificação da interpretação correta de uma dada raiz naquele mesmo contexto também tomará o índice como instrução.

Mais especificamente, Harley (2014) argumenta, através de diferentes fatos empíricos, que não é razoável admitir que as raízes apresentem uma forma fonológica prévia, tendo em vista que elas estão sujeitas a alomorfias de diversas ordens, inclusive as supletivas, em que a fonologia de um mesmo elemento da Gramática é realizada por fonologias distintas. Nesse sistema, as alomorfias supletivas são, na verdade, resultado da associação de um mesmo índice a diferentes expoentes fonológicos, considerados os contextos estruturais relevantes. A título de exemplificação apenas, se considerarmos o par *boi – vaca* do português como um pareamento supletivo, cuja fonologia é definida pelo traço de gênero presente na estrutura, podemos representar tal supleção a partir de um mesmo Item de Vocabulário, tal como representado abaixo:

(16) Supleção: um único índice para várias instruções na Lista 2

$$\begin{aligned} \sqrt[471] & \leftrightarrow /boj/ \ / \ [masc] \\ & \leftrightarrow /vak\epsilon/ \ [fem] \end{aligned}$$

Harley (2014) aponta ainda que individualizar as raízes por

<sup>40</sup> No trecho original: “The root terminal node elements occurring in List 1 can thus be notated as  $\sqrt[279]$ ,  $\sqrt[322]$ ,  $\sqrt[2588]$  etc. List 2 consists of instructions for spelling out each of these entities in a given morphosyntactic context. List 3 consists of instructions for interpreting these entities in a given morphosyntactic context”.

seu conteúdo conceitual também não é uma opção interessante dado o caráter polissêmico desses elementos que terão seus significados estabelecidos através das relações construídas no componente computacional. Essa polissemia, no entanto, seria análoga à supleção no ramo de PF, ou seja, como representado em (16) para a Lista 2, o mesmo índice pode estar atrelado a diferentes interpretações também na Lista 3. Para ilustrar, vamos assumir que  $\sqrt{_{987}}$  corresponda à raiz que participa da formação *gato*, por exemplo. Fixando a estrutura em um contexto nominal para simplificação, podemos atentar para a multiplicidade de conteúdo enciclopédico que pode estar atrelada ao índice relevante na Lista 3 no sistema de Harley (2014):

(17) Polissemia: um único índice para várias instruções na Lista 2

a. Instruções na Lista 2

$\sqrt{_{445}} \leftrightarrow /gat/$

b. Instruções na Lista 3

$\sqrt{_{445}} \leftrightarrow$  “animal de estimação que mia e dorme muito”  
 $/[n \ \_\_]\sqrt{}$   
 $\leftrightarrow$  “pessoa provida de beleza” /  $[n \ \_\_]\sqrt{}$   
 $\leftrightarrow$  “instalação irregular e ilegal de serviços como internet” /  $[n \ \_\_]\sqrt{}$   
 $\leftrightarrow$  “atleta que altera a sua idade real para parecer mais jovem” /  $[n \ \_\_]\sqrt{}$   
 $\leftrightarrow$  “indivíduo que realiza pequenos furtos”  
 $/[n \ \_\_]\sqrt{}$

Também a interpretação das raízes em expressões idiomáticas pode estar contida nas instruções agrupadas em torno de um único índice que mapeia para os conteúdos enciclopédicos relevantes em contextos estruturais específicos, tal como ilustrado abaixo:

## (18) Idiomaticidade: um único índice para várias instruções na Lista 2

a. Instruções na Lista 2

 $\sqrt{_{591}} \leftrightarrow /bawd/$ 

b. Instruções na Lista 3

 $\sqrt{_{591}} \leftrightarrow$  “recipiente com alça usado para carregar água” /  $[n \ [ \_ ] \sqrt{ }]$   
 $\leftrightarrow$  “trocar de uma linha de metrô para outra em um mesmo trajeto” /  $[v \ [ \_ ] \sqrt{ }]$   
 $\leftrightarrow$  “desistir” /  $[_{YP}[_V \ chutar] \ [_{DP}[_{Det} \ o] \ [N \ [n \ [ \_ ] \sqrt{ }]]]]]$ 

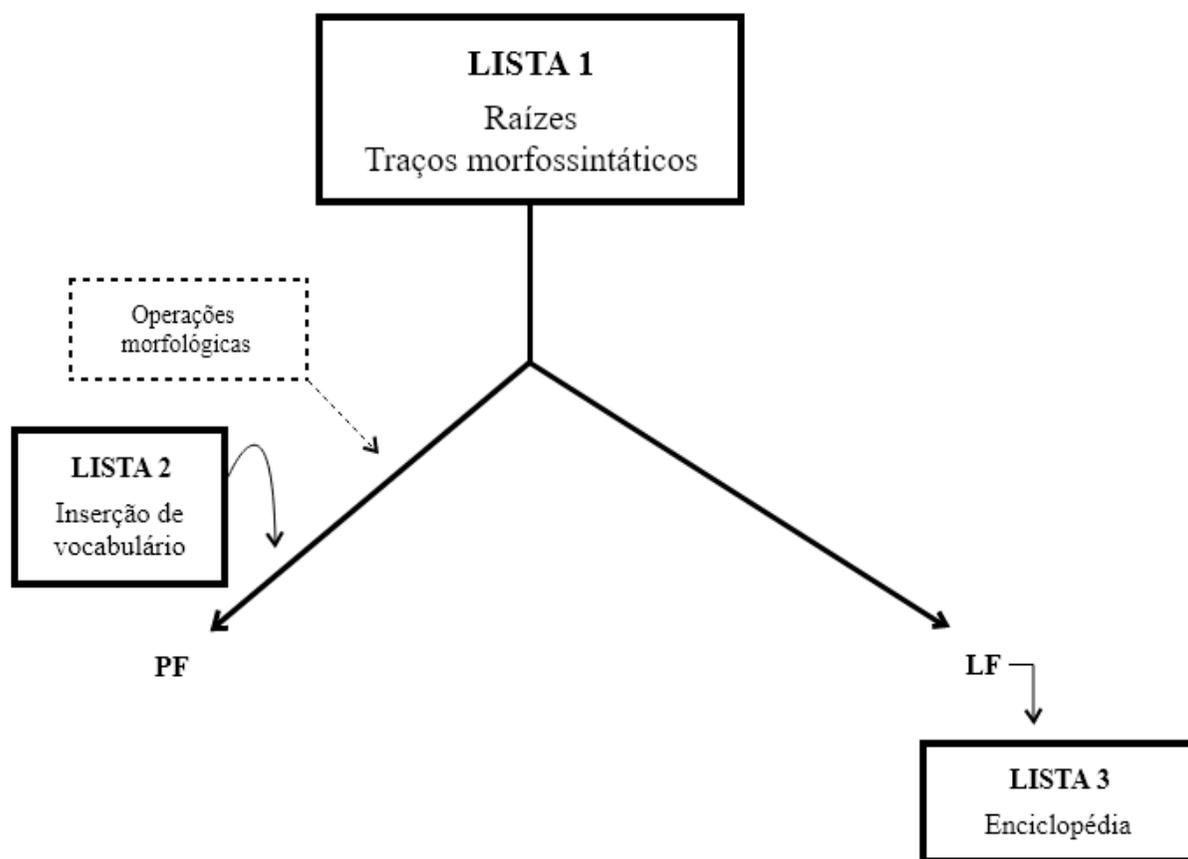
Na coluna (b), de (18), temos diferentes instruções relevantes para interpretação da raiz, cada uma delas atribuída a um contexto sintático distinto. Mais especificamente, na primeira delas encontramos a raiz  $\sqrt{_{591}}$  categorizada por uma estrutura nominal, o que forma, por exemplo, o nome *balde* na língua, com o seu conteúdo enciclopédico relevante, qualquer que ele seja. Por sua vez, a segunda interpretação enciclopédica apresenta a mesma raiz e, portanto, o mesmo índice, dessa vez em um contexto verbal, o que formaria, por exemplo, o verbo *baldear* na língua. Esse verbo é empregado para indicar, por exemplo, a mudança de uma linha de metrô para outra, e que vem a formar a nominalização derivada, bastante mais frequente, *baldeação*. Por fim, na terceira das interpretações ilustradas, a raiz relevante aparece como complemento do verbo *chutar* na expressão idiomática *chutar o balde*, que será interpretada como “desistir”.

A fatoração dessas informações associadas às raízes pela arquitetura da Gramática faz com que cada ramo do modelo opere, exclusivamente, com as informações que lhe cabem, evitando assim que a sintaxe tenha acesso a traços fonológicos e a conteúdos enciclopédicos, os quais não são empregados por suas operações.

De qualquer forma, em abordagens nas quais as raízes apresentam uma identidade previamente determinada, seja por

uma representação fonológica subjacente, seja por um índice alfanumérico, não haverá a necessidade de que a Lista 3 consulte o resultado de PF. Assim, a individualização prévia da raiz é suficiente para que a entrada enciclopédica relevante possa ser acessada, tal como ilustrado na arquitetura de Gramática a seguir:

Figura 2 – Localização do componente enciclopédico: acesso a LF.



Fonte: Bassani e Minussi (2015, p. 140).

Por fim, uma noção importante no que diz respeito ao acesso às instruções da Lista 3 é a ideia de localidade. De maneira geral, as abordagens sintáticas de formação de palavras têm proposto que a interpretação não composicional listada na Enciclopédia só pode ser acessada através de domínios sintáticos bem definidos, como apontamos algumas vezes. Assim, uma questão bastante saliente para a teoria é investigar em que ponto da derivação sintática esse

significado pode ser acessado. Essa questão, no entanto, extrapola os limites deste capítulo e será retomada no capítulo *Domínios de localidade na interpretação semântica*.

## RESUMINDO

Neste capítulo, aprendemos que alguns aspectos da interpretação das expressões linguísticas não podem ser derivados da constituição de traços formais da estrutura sintática, como a distinção entre dois substantivos animados como *gato* e *cachorro*, por exemplo. Tais informações arbitrárias e irrelevantes para o sistema computacional são localizadas, no modelo da Morfologia Distribuída, na Lista 3 ou Enciclopédia. Em linhas gerais, as entradas enciclopédicas estão associadas a raízes específicas inseridas em contextos hierárquicos sintaticamente construídos, de modo que a Lista 3 atribui significado conceitual a uma porção de estrutura sintática.

Além disso, vimos que, dados os pressupostos de decomposição plena, a Morfologia Distribuída precisa buscar uma explicação para as expressões idiomáticas como objetos da Gramática, não como itens armazenados na memória. Assim, o modelo deriva o significado das expressões idiomáticas considerando que há uma atribuição de conteúdo conceitual à porção da estrutura sintática envolvida na interpretação não composicional.

Por fim, exploramos os possíveis lugares de representação da Lista 3 no modelo da Morfologia Distribuída, discutindo a que tipo de informação esse componente precisa ter acesso para que a atribuição de conteúdo conceitual possa acontecer. A esse respeito, vimos que alguns autores propõem a necessidade de que a Lista 3 seja alimentada por informação de ambas as interfaces, PF e LF. Tal

visão está associada à necessidade de uma checagem da informação fonológica inserida na raiz no ramo de PF. Por outro lado, para autores que individualizam a raiz na Lista 1, seja via representação fonológica, seja via índice, o acesso à PF deixa de ser necessário e a Lista 3 passa, então, a ter acesso unicamente à interface de LF.

## **PARA SABER MAIS**

Para o leitor interessado em saber um pouco mais a respeito do conteúdo enciclopédico e da maneira como ele é acessado, recomendamos primeiramente o manuscrito de 1996 de Alec Marantz, *Cat as a phrasal idiom*. Nesse trabalho, estão muitas das ideias seminais no modelo a respeito das informações associadas à Enciclopédia. Ressaltamos a argumentação de que todas as palavras são estruturalmente complexas e, portanto, sua interpretação não pode estar listada e precisa ser associada a informações na Lista 3. Além disso, o capítulo *The Lexicon-Encyclopedia Interface* de Heidi Harley e Rolf Noyer, publicado em 2000, é também uma referência importante, pois nele os autores revisitam as nominalizações investigadas em Chomsky (1970) e reanalisam alguns desses dados como anomalias pragmáticas causadas exatamente na consulta enciclopédica. Além disso, remetemos o leitor ao artigo *The syntactic domain of content* publicado em 2013 por Hagit Borer, que explora, também a partir de uma abordagem sintática da formação de palavras – o modelo Exoesqueletal – a atribuição de conteúdo enciclopédico a variados tamanhos de estrutura sintática. Por fim, o artigo *A interface sintaxe-pragmática na formação de palavras: avaliando os pontos de acesso da Enciclopédia na arquitetura da Gramática*, de Rafael Minussi e Vitor Nóbrega, explora o funcionamento da Lista 3 através de dados do português brasileiro, propondo reformulações na atuação desse componente no modelo.

## EXERCÍCIOS

**Exercício 1.** Busque diferenciar o significado estrutural linguístico do significado conceitual idiossincrático através de 3 exemplos de expressões idiomáticas. Como modelo, você pode tomar o exemplo abaixo:

(i) *Chutar o pau da barraca*

Significado linguístico estrutural: existe um evento, denotado pelo verbo *chutar*, que toma como argumento interno o sintagma nominal *o pau da barraca*. Essa relação entre predicado-argumento está envolvida na atribuição da interpretação de *tema* desse sintagma nominal, que fará dele o elemento que foi chutado. Além disso, o verbo *chutar* requer um argumento externo, que receberá o papel temático de *agente* e será saturado, em uma sentença, através de outro sintagma nominal, como *o escoteiro*, por exemplo, em *o escoteiro chutou o pau da barraca*.

Significado conceitual idiossincrático: idiomáticamente a expressão é conceitualmente equivalente a desistir.

A partir das distinções de significado que você trouxe com os seus exemplos, reflita se é possível detectar algum sinal de composicionalidade nos significados idiossincráticos das expressões idiomáticas.

**Exercício 2.** Escolha uma raiz qualquer do português brasileiro e verifique se ela pode ser inserida em diversos contextos sintáticos, como diversas categorias e/ou empregados em expressões idiomáticas. A partir daí, forneça a entrada enciclopédica relevante nos moldes de Harley (2014). Como modelo, você pode usar o

exemplo (15) da seção 3.

**Exercício 3.** Discuta as diferentes abordagens apresentadas no capítulo em relação ao local de acesso à Lista 3 na arquitetura de Gramática da Morfologia Distribuída, apontando como essas propostas se relacionam a diferentes visões a respeito da natureza das raízes.

