

DOMÍNIOS DE LOCALIDADE NA INTERPRETAÇÃO SEMÂNTICA

Vitor Augusto Nóbrega

Conceitos-chave do capítulo:

- Significados polissêmicos e idiomáticos
- Morfemas internos vs. morfemas externos
- Alossemia contextual
- Domínios de fase internamente à estrutura da palavra
- Ciclicidade em contextos de alossemia e de idiomaticidade

Uma das grandes questões envolvendo o estudo da linguagem humana é a da significação, ou seja, o que quer dizer *significar* e como isso se processa internamente à Faculdade da Linguagem. Para fornecer uma explicação a essa questão, a linguística, a filosofia, a lógica, a psicologia, entre outras áreas do conhecimento, costumam recorrer ao Princípio da Composicionalidade, comumente creditado ao lógico e filósofo Gottlob Frege (JANSSEN, 2012). Esse princípio determina que a interpretação das expressões linguísticas emerge da soma das partes que as compõem, considerando seu arranjo sintático, tal como sintetizado em (1):

(1) Princípio da Composicionalidade

O significado de uma expressão complexa é uma função dos significados de suas partes e do modo como elas estão combi-

nadas sintaticamente.^{108/109}

(adaptado de Partee, 1984, p. 281, tradução nossa)

No entanto, diversos tipos de objetos linguísticos, de variadas extensões, incluindo palavras derivadas, palavras compostas, sintagmas e sentenças, carregam significados que desafiam as diretrizes do princípio em (1), ou seja, o significado de suas partes constituintes – sejam elas morfemas ou palavras – não contribui, diretamente, para o significado global do elemento formado. Exemplos estão listados em (2):

- (2) a. (marido) banana
 b. fogão
 c. caixa-preta
 d. entregar o ouro

Em (2a), notamos que o nome *banana* é polissêmico: neste sintagma, ele não corresponde ao fruto da bananeira, mas é interpretado, metaforicamente, como “covarde” ou “sem iniciativa”. Em (2b), o sufixo aumentativo *-ão* não compõe a extensão de *fogão*, visto que *fogão* não é um “fogo/incêndio de grandes proporções”, mas um aparelho doméstico utilizado para cozinhar. Em (2c), o composto nominal *caixa-preta* denota a aparelhagem que grava os dados sobre o funcionamento de uma aeronave; porém, não necessariamente diz respeito a uma caixa, e, curiosamente, sua cor

108 No trecho original: “*The meaning of a compound expression is a function of the meanings of its parts and of the way they are syntactically combined*”.

109 A composicionalidade é uma propriedade fundamental do nosso conhecimento linguístico. É ela que permite, por exemplo, que qualquer indivíduo possa compreender e produzir sentenças que nunca tenha ouvido. É também a composicionalidade que possibilita que uma criança em processo de aquisição faça inferências sobre o significado das expressões linguísticas a que é exposta e levante hipóteses sobre sua organização estrutural, desenvolvendo, assim, uma Gramática interna (IRURTZUN; GALLEG0, 2007, p. 513). De modo amplo, o Princípio da Composicionalidade impede que o significado das expressões geradas pela Faculdade da Linguagem seja arbitrário, fato que assegura uma compreensão mútua entre o que é produzido e interpretado pelos falantes de uma língua (FODOR; LEPORE, 2002, p. 14-15).

não precisa ser preta. Finalmente, o sintagma verbal em (2d) tem seu complemento nominal, *o ouro*, interpretado, idiomáticamente, como “segredo” ou “informação valiosa”.

Para explicar o emprego de dados semanticamente irregulares como aqueles em (2), as abordagens lexicalistas geralmente assumem que objetos linguísticos, simples ou complexos, com significados polissêmicos ou idiomáticos estão armazenados em um componente independente, pré-sintático – o léxico –, o qual é caracterizado como o repositório de idiosincrasias da Gramática (CHOMSKY, 1981; DI SCIULLO; WILLIAMS, 1987; JACKENDOFF, 1997, 2002, entre outros). Para esses autores, tais objetos linguísticos são inseridos no componente sintático já formados e com suas respectivas idiosincrasias semânticas. Consequentemente, a composicionalidade, a regularidade e a produtividade atreladas à sintaxe não são afetadas, pois eles serão interpretados como uma unidade sintática atômica.

Vimos que a Morfologia Distribuída não assume um léxico pré-sintático. Sendo assim, de que maneira esse modelo de Gramática explica a presença de polissemia e idiomática nas expressões linguísticas? Para responder essa questão, avançamos, neste capítulo, a discussão iniciada no capítulo *Lista 3: Enciclopédia*, ao explorar de que modo se dá a articulação entre a sintaxe e a Lista 3. Discutimos, primeiramente, o papel das raízes na emergência de significados polissêmicos e especiais e a proposta de decomposição plena assumida pelo modelo. Em seguida, mostramos que a determinação desses significados leva em consideração o ambiente sintático em que a raiz está inserida, bem como os morfemas estruturalmente próximos a ela. Por fim, manipulamos os mecanismos sintáticos utilizados para delimitar os domínios de polissemia e idiomática, com o intuito de aprender como a Morfologia Distribuída faz a sintaxe fazer sentido.

1. Onde, na Gramática, estão os significados polissêmicos e idiomáticos?

Uma das funções características do léxico, para os modelos que o assumem como componente gerativo, é a de listar as irregularidades semânticas de uma língua, tais como formas polissêmicas e formas complexas cujos significados não são previstos pelas regras da Gramática. Essa função estabelece uma dicotomia entre os papéis do léxico e da sintaxe: o léxico é o espaço das irregularidades, ao passo que a sintaxe é o espaço das regularidades. A Morfologia Distribuída rejeita essa dicotomia ao identificar que não há uma correspondência uniforme entre estrutura sintática e significado regular, uma vez que as irregularidades semânticas não se restringem ao domínio da palavra (MARANTZ, 1996, 1997). Tal posicionamento teórico leva a Morfologia Distribuída a questionar qual seria a fonte dos significados irregulares. Mais precisamente, o que permite que um sistema gerativo composicional, como a sintaxe, dê origem a objetos linguísticos com significados idiossincráticos, a exemplo daqueles elencados em (2).

Devemos, primeiramente, perceber que, embora a Morfologia Distribuída rejeite a hipótese de um armazenamento integral pré-sintático, o modelo não descarta a necessidade de se listarem certos tipos de informações. A multiplicidade de significados das raízes, por exemplo, sinaliza a importância de listas internamente à Faculdade da Linguagem, como vimos no capítulo *Lista 3: Enciclopédia*. Para recuperar esse ponto, tomemos o nome *coração*, derivado da raiz $\sqrt{\text{CORAÇÃO}}$, o qual apresenta inúmeras interpretações: “órgão muscular”, “pessoa a quem se ama”, “sentimentos”, “parte central de um lugar”, “generosidade”, etc. Os múltiplos significados não são previsíveis, pois toda raiz pode, eventualmente, adquirir significados inesperados, tal como é o caso da raiz $\sqrt{\text{BANAN}}$, em

(2a). Assim sendo, é indispensável que a Faculdade da Linguagem contenha uma lista de informações concernentes à polissemia e à idiomaticidade. Porém, essa lista não precisa ser, necessariamente, pré-sintática e composta por objetos complexos.

O que a Morfologia Distribuída propõe, como uma alternativa à hipótese de armazenamento integral, é que todo tipo de irregularidade semântica esteja associado a uma classe de primitivos – as raízes –, e ao modo como seu significado é determinado pela Lista 3 pós-sintaticamente. Nesse sentido, o significado idiossincrático de um objeto linguístico não decorre de seu tipo estrutural, se morfológico ou sintático, mas da interpretação atribuída à(s) sua(s) raiz(es), seja ele um objeto linguístico mais simples, como os nomes *coração* e *banana*, ou um objeto linguístico mais complexo, como o sintagma verbal *entregar o ouro*. As raízes, portanto, são a fonte das idiossincrasias da Gramática. Para entender como elas adquirem significados diversos, uma distinção fundamental do modelo deve estar clara até este ponto do manual: as raízes não estão alocadas na mesma lista em que seus significados estão listados.

As raízes contidas na Lista 1 – dos primitivos que abastecem a sintaxe – correspondem a um núcleo semântico subespecificado, ou seja, sem uma interpretação fixa, pré-determinada (ARAD, 2003, 2005).¹¹⁰ De acordo com essa visão, as raízes devem ser entendidas como uma potencialidade de significados, já que por serem semanticamente subespecificadas, elas podem ser associadas a diferentes interpretações. Essas diferentes interpretações, por outro

110 Alguns autores, como Marantz (1996), Acquaviva (2009) e Harley (2014), assumem uma postura mais radical que a de Arad (2003, 2005) e defendem que as raízes são, na realidade, vazias de significado. Esses autores sugerem que o significado de uma raiz é determinado tardiamente, no ramo de LF, semelhantemente à inserção tardia de fonologia nos nós terminais sintáticos que verificamos no ramo de PF. De qualquer forma, ambas as hipóteses – subespecificação ou vacuidade de significado – chegam a conclusões paralelas acerca do modo como a Gramática atribui significados idiossincráticos às expressões linguísticas. A ideia básica conectando essas duas hipóteses é a de que a polissemia e a idiomaticidade vinculadas a uma raiz somente são possíveis se ela não estiver rigidamente associada a um dado significado. Em outras palavras, uma raiz não pode ser vista como um “copo cheio”, caso contrário, não seria possível atribuir a ela significados adicionais.

lado, estão listadas como entradas enciclopédicas na Lista 3, o componente pós-sintático que faz interface com o nosso conhecimento de mundo. A Lista 3, como vimos no capítulo *Lista 3: Enciclopédia*, inclui tudo aquilo que sabemos sobre as raízes e seus significados nos diferentes ambientes sintáticos em que elas aparecem (MARANTZ, 1997, 2001; ARAD, 2003, 2005; HARLEY, 2014).¹¹¹

Por exemplo, a raiz $\sqrt{\text{CORAÇÃO}}$ tem sua polissemia codificada na forma de uma entrada enciclopédica. Em (3), listamos os diferentes significados contidos na entrada enciclopédica da raiz $\sqrt{\text{CORAÇÃO}}$, quando ela ocorre em um ambiente nominal. Σ corresponde a cada um dos possíveis significados.

- (3) $\sqrt{\text{CORAÇÃO}}$, no ambiente $[_n]$,
- Σ_1 : órgão muscular que recebe o sangue das veias e o envia para as artérias;
 - Σ_2 : pessoa a quem se ama;
 - Σ_3 : parte central de um lugar (ex., *o coração da cidade*);
 - Σ_4 : sentimentos (ex., *conquistou o coração de seus alunos*);
 - Σ_5 : generosidade (ex., *presidente sem coração*);
 - etc.

Essa mesma raiz pode ainda adquirir outros significados em construções idiomáticas, tais como $[_n \text{ coração}] \text{ mole}$, *com o* $[_n \text{ coração}] \text{ nas mãos}$, *cortar o* $[_n \text{ coração}]$, os quais estão, igualmente, listados na Lista 3, juntamente com seu contexto de ocorrência. Podemos dizer então que, para a Morfologia Distribuída, entender a distribuição de significados polissêmicos e especiais é entender o

111 É válido destacar que o núcleo semântico subespecificado sugerido por Arad (2003, 2005) diz respeito a um denominador comum, responsável por conectar semanticamente as palavras derivadas de uma mesma raiz. As entradas enciclopédicas presentes na Lista 3 correspondem, por sua vez, aos significados que podem ser correlacionados a esse denominador comum, muito embora essa correlação não seja direta em muitos casos.

modo como as raízes da Lista 1 são associadas às suas entradas enciclopédicas na Lista 3.

Da mesma forma que a Morfologia Distribuída determina o que deve estar listado, ela também determina o que não deve estar listado, a saber: formas complexas com significados idiomáticos, os chamados “significados especiais”, tais como os associados às formas ilustradas em (2c) e (2d). Isso implica que todo objeto linguístico complexo, ou seja, formado por mais de um morfema, deve ser derivado pela sintaxe toda vez que for empregado. Esse pressuposto, retomado em (4), prevê uma decomposição plena das formas derivadas pela Gramática, impedindo que a Lista 1 armazene objetos complexos pré-formados, isto é, que ela armazene palavras e/ou sintagmas, como discutido no capítulo *Lista 3: Enciclopédia*.

(4) Decomposição plena

Não há objetos complexos armazenados em nossa memória. Todo objeto complexo deve ser derivado pela Gramática toda vez que for empregado.¹¹²

(adaptado de Embick, 2015, p. 21, tradução nossa)

A decomposição plena nos permite identificar quais são os ambientes sintáticos sensíveis a idiosincrasias semânticas, em vez de apenas listar objetos irregulares em um léxico pré-sintático. Como ressaltam Epstein e Seely (2006, p. 7), listar um objeto complexo é deixar de explicá-lo. Ainda, a decomposição plena estabelece que qualquer tipo de idiosincrasia semântica é determinado na interface interpretativa pela Lista 3. Em linhas gerais, todos os objetos linguísticos – palavras e sintagmas – são derivados composicionalmente pela sintaxe. Os significados idiosincráticos

112 Trecho original: “No complex objects are stored in memory; i.e., every complex object must be derived by the grammar”.

que esses objetos podem apresentar dizem respeito a informações enciclopédicas associadas às suas raízes, juntamente com o contexto sintático que os desencadeia. Vale salientar novamente que a aquisição de informações enciclopédicas não se restringe ao período crítico da aquisição da linguagem, pois podemos associar novos significados às raízes a qualquer momento. Para entender como a associação entre raízes e significados acontece, iremos, na próxima seção, explorar as consequências da decomposição plena e o papel do contexto sintático na interpretação semântica.

2. Contextualizando o significado

Tendo em vista o que discutimos até aqui, você deve estar se perguntando: como as raízes – primitivos semanticamente subespecificados – têm seu conteúdo fixado? Para responder essa questão, devemos investigar os limites estruturais das idiosincrasias semânticas, ou seja, até que ponto podemos combinar morfemas a uma raiz e ainda ser capaz de derivar palavras e sintagmas com significados inesperados. Desenvolveremos essa questão, examinando, inicialmente, a variedade de significados presentes em uma família de palavras, ou seja, em um conjunto de palavras derivadas de uma mesma raiz.

Um dos exemplos mais expressivos de família de palavras é fornecido pela morfologia do hebraico. Como descrito no capítulo *Lista 1: traços morfossintáticos e raízes*, o hebraico contém raízes formadas por três consoantes, \sqrt{CCC} , que se tornam verbos ou nomes quando associadas a um determinado padrão de vogais. A combinação de uma raiz a diferentes padrões vocálicos dá origem a uma família de palavras cujos significados podem variar consideravelmente. Alguns exemplos são as palavras formadas a partir das raízes $\sqrt{X\check{S}B}$ e \sqrt{SGR} , em (5) e (6):

(5) Hebraico

	Raiz: $\sqrt{X\check{S}B}$	Padrão	Palavra
a.	CaCaC (v)	<i>xašav</i>	‘pensar’
b.	CiCCeC (v)	<i>xišev</i>	‘calcular’
c.	hiCCiC (v)	<i>hexšiv</i>	‘considerar’
d.	maCCeC (n)	<i>maxšev</i>	‘computador’
e.	maCCaCa (n)	<i>maxšava</i>	‘pensamento’
f.	taCCiC (n)	<i>taxšiv</i>	‘cálculo’
g.	CiCCon (n)	<i>xešbon</i>	‘explicação’

(ARAD, 2003, p. 750)

(6) Hebraico

	Raiz: \sqrt{SGR}	Padrão	Palavra
a.	CaCaC (v)	<i>sagar</i>	‘fechar’
b.	hiCCiC (v)	<i>hisgir</i>	‘extraditar’
c.	hitCaCCeC (v)	<i>histager</i>	‘fechar-se’
d.	CeCeC (n)	<i>seger</i>	‘fechamento’
e.	CoCCayim (n)	<i>sograyim</i>	‘parênteses’
f.	miCCeCet (n)	<i>misgeret</i>	‘quadro’

(ARAD, 2003, p. 746)

Ao observar as palavras em (5) e (6), percebemos que elas incorporam múltiplos significados, muito embora apresentem, em certo nível, um denominador semântico comum. Percebemos também que os diferentes significados somente são determinados quando as raízes $\sqrt{X\check{S}B}$ e \sqrt{SGR} se combinam aos padrões vocálicos, os quais correspondem a diferentes expoentes fonológicos para as categorias lexicais verbais (v) e nominais (n) (ARAD, 2003, p. 741).

Essa correlação entre raízes e padrões vocálicos nos sugere que a categorização, isto é, a atribuição de uma categoria lexical a uma raiz, tem um papel fundamental na determinação do significado das raízes, já que é nesse ambiente sintático que verificamos a atribuição de uma variedade de conteúdos.

Para avaliar a validade de tal afirmação, observemos o que ocorre quando derivamos novas palavras a partir dos itens formados em (5) e (6). Se os significados das novas palavras derivadas forem dissociados do significado das palavras que lhes servem como base, a afirmação acima é falsa, pois isso nos indica que as raízes $\sqrt{X\check{S}B}$ e \sqrt{SGR} não tiveram seu conteúdo fixado quando foram categorizadas pela primeira vez. Consideremos, para tanto, os verbos denominais formados a partir dos nomes *maxšev* ‘computador’ e *misgeret* ‘quadro’, em (7) e (8):

(7) Hebraico

	Raiz: $\sqrt{X\check{S}B}$	Padrão	Palavra
a.	maCCeC (n)	<i>maxšev</i>	‘computador’
b.	CiCCeC (v)	<i>mixšev</i>	‘computar’

(ARAD, 2003, p. 750)

(8) Hebraico

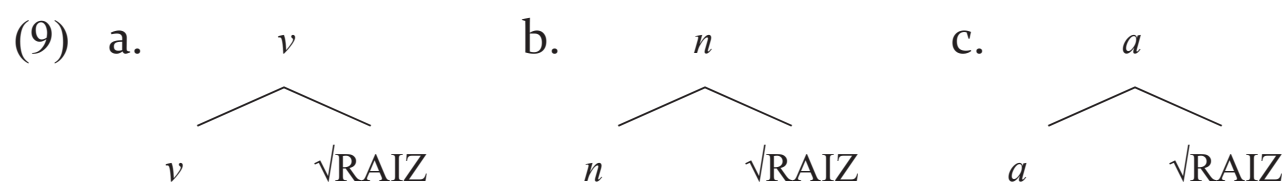
	Raiz: \sqrt{SGR}	Padrão	Palavra
a.	miCCCeCet (n)	<i>misgeret</i>	‘quadro’
b.	CiCCeC (v)	<i>misger</i>	‘enquadrar’

(ARAD, 2003, p. 746)

Podemos afirmar que os verbos *mixšev* ‘computar’, em (7b), e *misger* ‘enquadrar’, em (8b), são derivados dos nomes *maxšev*

‘computador’, em (7a), e *misgeret* ‘quadro’, em (8a) – mas não diretamente das raízes $\sqrt{X\check{S}B}$ e \sqrt{SGR} –, pois ambos mantêm o prefixo *m-* correspondente ao padrão nominal, como visto primeiramente em (5d) e em (6f) (ARAD, 2003, p. 750). Vemos que os significados desses verbos denominais são composicionais, já que eles se somam ao significado da base nominal, ou seja, o verbo *mixšev* ‘computar’ inclui o significado do nome *maxšev* ‘computador’, da mesma forma que o verbo *misger* ‘enquadrar’ inclui o significado do nome *misgeret* ‘quadro’. Esses dados, portanto, endossam a hipótese de que a categoria lexical opera como uma fronteira na delimitação do significado de uma raiz. Isso significa que as raízes $\sqrt{X\check{S}B}$ e \sqrt{SGR} tiveram sua interpretação fixada quando se tornaram um nome, o que impediu que derivações posteriores, tal como a formação dos verbos denominais em (7) e (8), apresentassem conteúdos imprevisíveis e não composicionais.

A hipótese de que o primeiro núcleo categorial atua como uma fronteira na determinação do significado das raízes foi, primeiramente, proposta por Marantz (2001, 2007) e endossada nos trabalhos de Arad (2003, 2005), com os dados que acabamos de discutir. Em termos sintáticos, essa hipótese pode ser reescrita do seguinte modo: as raízes têm sua interpretação negociada no ambiente sintático delimitado pelo primeiro núcleo categorial. Como discutido no capítulo *Lista 1: traços morfossintáticos e raízes*, as raízes são primitivos gramaticais sem uma especificação funcional, sendo associadas a uma categoria durante a derivação sintática, através da combinação a um núcleo categorial, tais como *v*, *n* ou *a*. Para lembrar como isso ocorre, vejamos as estruturas de categorização em (9):



Com base na distribuição semântica observada acima, Marantz (2001, 2007) e Arad (2003, 2005) fornecem uma generalização sobre a determinação dos significados idiossincráticos de uma palavra, descrita em (10).¹¹³

- (10) Generalização sobre a determinação dos significados idiossincráticos de uma palavra
- a. Palavras formadas diretamente das raízes podem apresentar significados variados.
 - b. Palavras derivadas de palavras já formadas – ou seja, de raízes categorizadas – apresentam somente um significado previsível, composto ao significado de sua base.

Há, nesse sentido, uma restrição de localidade na determinação do significado de uma raiz, pois apenas o núcleo categorial estruturalmente próximo a ela, seguindo a configuração sintática em (9), pode condicionar sua interpretação. Essa restrição de localidade nos indica que temos, pelo menos, duas classes de morfemas na Gramática, as quais se distinguem por sua localização na estrutura sintática da palavra, a saber: morfemas que se concatenam localmente às raízes, os quais chamaremos de morfemas internos, e morfemas que se concatenam a outros elementos mais distantes das raízes, os quais chamaremos de morfemas externos (MARANTZ, 2001, 2007). No capítulo *Domínios de localidade na forma fonológica*, pudemos ver como essa distinção entre morfemas opera no que se refere às previsões relacionadas à forma. No que compete à interpretação semântica, as principais características dessas duas classes de morfemas estão resumidas em (11) e esquematizadas em (12), em que x representa um núcleo categorial (v, n, a):

113 Essa generalização basicamente revisa, em termos sintáticos, a distribuição de morfemas em diferentes níveis lexicais, tal como foi assumido por algumas hipóteses lexicalistas. Para entender melhor como o léxico se organiza nessa abordagem em níveis, ver Kiparsky (1982b).

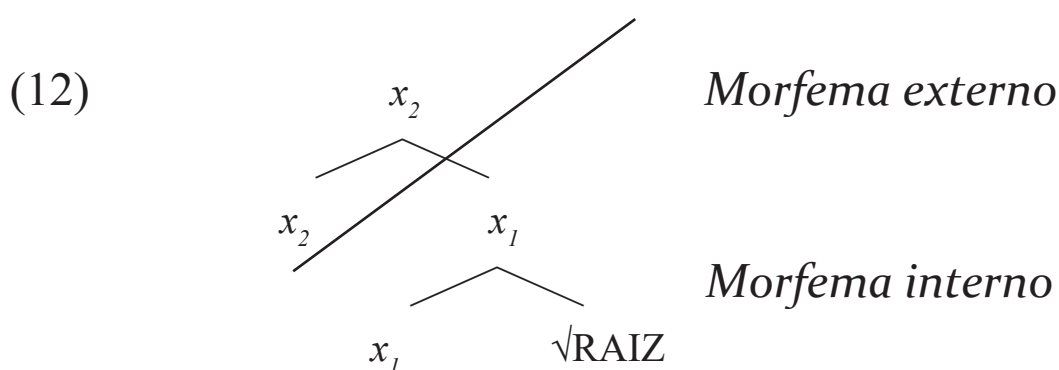
(11) Morfemas internos vs. morfemas externos (MARANTZ, 2001, 2007)

a. **Morfemas internos**

- i. Posição estrutural: correspondem ao primeiro núcleo categorial e aos morfemas que se concatenam internamente ao domínio que tal núcleo delimita;
- ii. Interpretação: podem desencadear significados idiossincráticos.

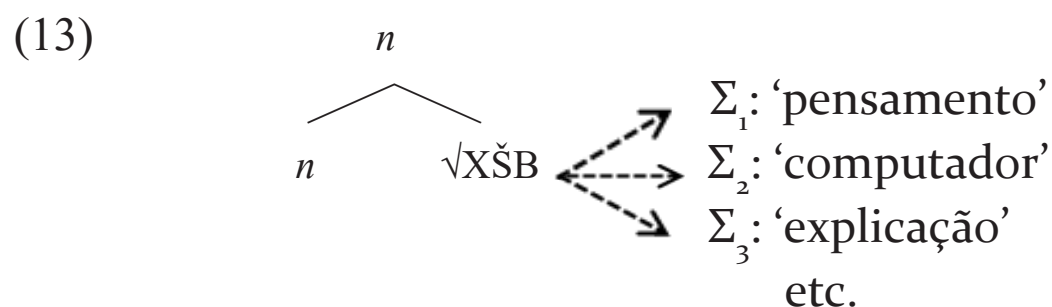
b. **Morfemas externos**

- i. Posição estrutural: correspondem a todo morfema que se concatena externamente ao domínio delimitado pelo primeiro núcleo categorial;
- ii. Interpretação: apresentam somente significados previsíveis.



Para entender essa distinção, recuperemos os dados do hebraico discutidos em (5). De acordo com as definições em (11), podemos dizer que os padrões vocálicos nominais *maCCeC*, *maCCaCa* e *CiCCon* correspondem aos expoentes fonológicos de um núcleo categorial nominal *n*, que se distribui sintaticamente como um morfema interno, ou seja, concatena-se localmente à raiz. Tal afirmação se justifica pelo fato de que esse núcleo categorial *n*, embora realizado por diferentes padrões vocálicos, pode condicionar o significado da raiz $\sqrt{X\check{S}B}$ e extrair dela diferentes interpretações, notadamente: *maxšev* ‘computador’, quando realizado pelo padrão

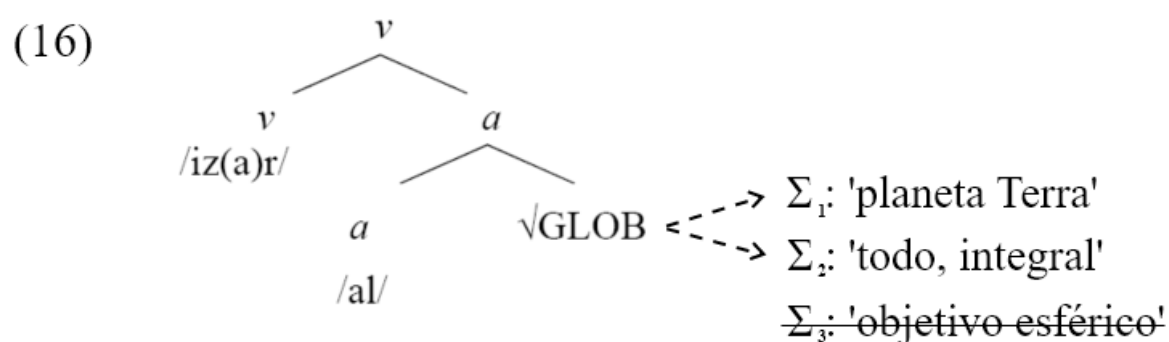
maCCeC, *maxšava* ‘pensamento’, quando realizado pelo padrão *maCCaCa*, e *xešbon* ‘explicação’, quando realizado pelo padrão *CiCCon*.¹¹⁴ A configuração de sua posição estrutural está ilustrada em (13) abaixo; as setas pontilhadas indicam os significados que a raiz $\sqrt{X\check{S}B}$ pode apresentar em um contexto nominal, catalogados na Lista 3 pós-sintaticamente:



Por outro lado, o núcleo categorial verbal v , realizado fonologicamente pelo padrão *CiCCeC*, corresponde a um morfema externo nesse contexto, já que ele se concatena em uma posição distante da posição ocupada pela raiz, acima do núcleo categorial n , representado em (13). Observe, em (14), que não há localidade entre o núcleo v e a raiz $\sqrt{X\check{S}B}$, pois n intervém entre eles. Consequentemente, v não pode condicionar o significado da raiz $\sqrt{X\check{S}B}$ e induzir uma interpretação idiossincrática. Dada a fronteira que o núcleo n estabelece, o significado da palavra formada a partir da verbalização da estrutura em (13) deve se compor ao significado negociado no ambiente nominal. É por essa razão que o verbo denominal *mixšev* ‘computar’ tem sua interpretação composicionalmente relacionada ao significado do nome *maxšev* ‘computador’:

114 Arad (2003, 2005) argumenta, com efeito, que cada padrão vocálico corresponde a um núcleo categorial n distinto. A fim de simplificar a exposição, admitimos, neste capítulo, que esses padrões representam a realização fonológica de um mesmo núcleo categorial.

-iz(a)r, ao adjetivo *global* e formar o verbo *globalizar*. Uma vez que *v* corresponde a um morfema externo nesse contexto, ele não pode acessar a raiz $\sqrt{\text{GLOB}}$ e induzir uma interpretação diferente daquela negociada no ambiente adjetival. Isso se confirma pelo fato de que o verbo *globalizar* não significa “fazer algo se tornar esférico”, mas pode significar ou (i) “fazer algo se tornar mundialmente distribuído/conhecido” ou (ii) “totalizar, integralizar”, tal como ilustrado em (16).



Verificamos, portanto, que a escolha do significado de uma raiz polissêmica é regida por uma restrição de localidade estrutural, a qual delimita a extensão do contexto sintático capaz de influenciar sua interpretação. Note que, opostamente à proposta de armazenamento integral das abordagens lexicalistas, será a distinção morfema interno *vs.* morfema externo que determinará a irregularidade semântica das palavras. Nesse sentido, a Morfologia Distribuída, em vez de recorrer a um léxico pré-sintático, propõe que as irregularidades semânticas sejam definidas por localidade estrutural.

Marantz (2013) caracteriza a influência da estrutura sintática na interpretação semântica dos morfemas que a compõem como *alossema contextual*. Esse condicionamento estrutural assemelha-se ao que observamos nos casos de alomorfia contextual, ou seja,

quando um dado morfema condiciona a realização fonológica de outro em contextos locais, assunto extensamente explorado no capítulo *Domínios de localidade na forma fonológica*. A alossemia contextual pode ser definida, por sua vez, do seguinte modo:

(17) Alossemia contextual

Influência estrutural na escolha do significado de um nó terminal sintático, ou seja, de um alossema contextualmente apropriado.

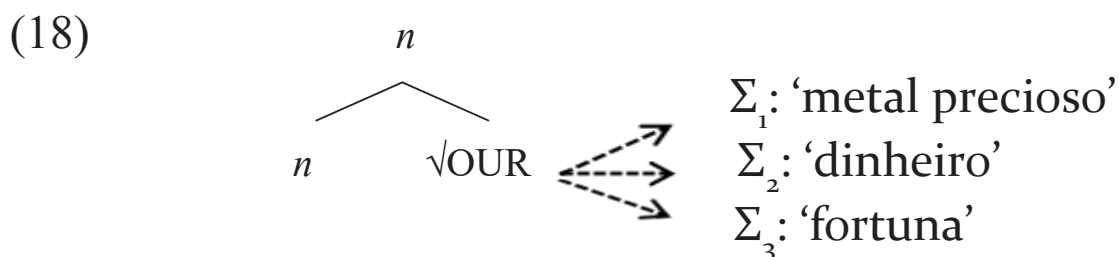
É importante ressaltar que a alossemia contextual circunscreve-se a casos de polissemia, ou seja, a casos de interpretações distintas – porém, semanticamente relacionadas – de uma mesma raiz. Um exemplo disso é a multiplicidade de significados das raízes que discutimos até aqui, nomeadamente, as raízes $\sqrt{\text{BANAN}}$, $\sqrt{\text{CORAÇÃO}}$ e $\sqrt{\text{GLOB}}$, do português, e $\sqrt{\text{XŠB}}$ e $\sqrt{\text{SGR}}$, do hebraico. Por outro lado, raízes que compartilham uma mesma forma fonológica, mas que não apresentam nenhuma correlação semântica, dizem respeito a casos de homofonia acidental, ou seja, raízes cuja identidade fonológica é um mero acaso histórico. Exemplos dessa natureza são as raízes do português $\sqrt{\text{MANG}}_1$, que significa “fruto da mangueira”, e $\sqrt{\text{MANG}}_2$, que corresponde a “uma parte da vestimenta”. Casos de homofonia acidental não caracterizam um exemplo de alossemia contextual, visto que correspondem a raízes distintas e, por consequência, a conjuntos distintos de entradas enciclopédicas.

Em síntese, podemos concluir que os casos de polissemia são resolvidos no domínio estrutural delimitado pelo primeiro núcleo categorial. Esse é, portanto, o contexto sintático que restringe sua

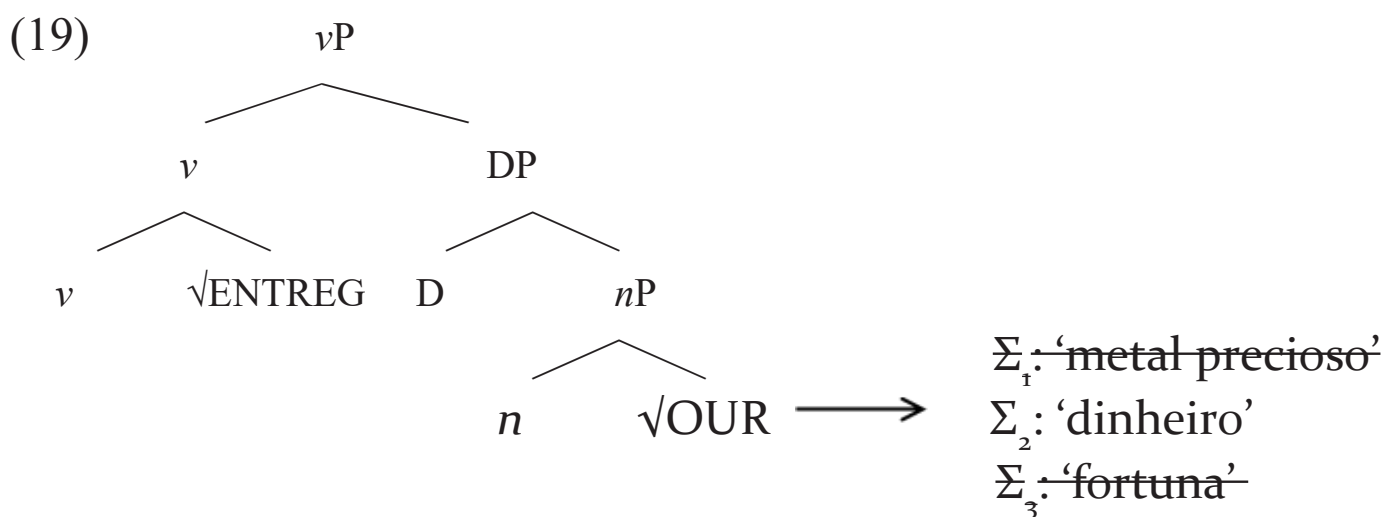
aplicação.¹¹⁵ Porém, tal conclusão nos coloca uma questão: dado que as raízes têm sua polissemia decidida no ambiente do primeiro núcleo categorial, de que forma obtemos significados idiomáticos em estruturas sintáticas mais complexas, tais como a expressão *esconder o ouro*? Mencionamos, na última seção, que as expressões idiomáticas não estão listadas em um léxico pré-sintático – tendo em vista a proposta já mencionada de decomposição plena – e que seu significado idiomático está igualmente listado na Lista 3. Para responder essa questão, devemos identificar o limite estrutural da idiomaticidade, o qual é evidentemente mais extenso que o limite estrutural da polissemia, já que as expressões idiomáticas são constituídas por um conjunto maior de elementos.

Marantz (2013) sugere, como resposta à questão acima, que as interpretações especiais atribuídas às expressões idiomáticas correspondem a um tipo de significado estabelecido sobre o significado fixado de uma raiz. Isso quer dizer que os significados das raízes podem ser sobrescritos caso haja uma entrada enciclopédica sinalizando outra interpretação em um contexto sintático mais elaborado. Para ilustrar, tomemos a expressão idiomática *esconder o ouro*, que, como vimos em (2d), tem seu complemento nominal interpretado como “segredo”. Ao iniciar sua derivação sintática, formamos o sintagma nominal *o ouro*, através da concatenação da raiz $\sqrt{\text{OUR}}$ a um núcleo categorial *n*. Nesse ambiente estrutural, um dos alossemas da raiz $\sqrt{\text{OUR}}$ deverá ser selecionado, tal como representado em (18):

115 É relevante apontar que diversos autores vêm desafiando a assunção de que o primeiro domínio categorial compreende o único ambiente sintático em que significados idiossincráticos podem ser negociados. Há inúmeros casos de morfemas derivacionais que ocorrem acima do primeiro núcleo categorizador e levam a interpretações não composicionais (BORER, 2013; ANAGNOSTOPOULOU; SAMIOTI, 2014; HARLEY, 2014). Fenômenos semelhantes são observados na formação de palavras compostas, as quais também envolvem mais de um núcleo categorizador e compreendem uma interpretação não composicional (NÓBREGA; PANAGIOTIDIS, 2020; NÓBREGA, 2020). Marantz (2013) e Panagiotidis (2014) oferecem análises alternativas que mantêm a assunção acima e explicam os casos de idiossincrasia que extrapolam o primeiro núcleo categorizador.



Observe, entretanto, que o nome *ouro*, empregado fora da expressão idiomática *entregar o ouro*, não apresenta o alossema “segredo”, fato que se verifica pela ausência desse alossema em um contexto nominal isolado. Como, então, a raiz $\sqrt{\text{OUR}}$ acaba adquirindo essa interpretação? À medida em que a derivação sintática continua, a estrutura em (18) é combinada ao verbo *entregar*, dando origem ao sintagma verbal *entregar o ouro*, estruturalmente esquematizado em (19). Suponha, então, que o alossema da raiz $\sqrt{\text{OUR}}$ escolhido no ambiente de n , em (18), tenha sido “dinheiro”. Esse alossema será mantido até o momento em que a estrutura sintática em (19) for acessada integralmente pela Lista 3.

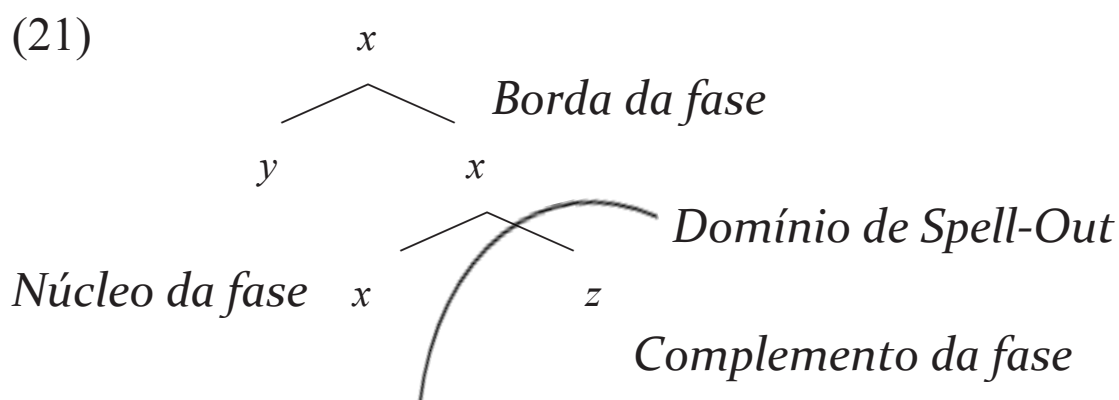


Seguindo Marantz (1997, 2013) e Harley (2014), admitimos que o limite estrutural para a atribuição de significados idiomáticos corresponde ao domínio circunscrito pelo núcleo funcional introdutor de argumentos externos, Voice. Esse núcleo é concatenado logo acima de vP . Como resultado, toda a estrutura sintática incluída em vP será vista como um ambiente passível de ser interpretado

vez que essa porção é removida da sintaxe, ela já não pode mais ser manipulada ou alterada por suas operações. Marantz (2001, 2007) e Arad (2003, 2005) veem na proposta de derivação por fases de Chomsky (2000) uma maneira de explicar os contextos de idiosincrasia identificados na formação de palavras. Os autores, então, estendem essa proposta ao domínio da palavra e buscam, a partir dos recortes estabelecidos pelas fases, derivar os domínios de localidade previstos na generalização em (10).

Na visão de Chomsky, as fases são determinadas por um conjunto fixo de núcleos sintáticos. O complemento desses núcleos é transferido para as interfaces semântica e fonológica logo após se concatenarem ao núcleo. Nesse sentido, o núcleo da fase é a unidade sintática responsável por desencadear o envio de seu complemento estrutural para as interfaces; em outras palavras, o núcleo da fase é responsável por desencadear o *Spell-Out* de seu complemento. Todos os elementos que estão fora do complemento da fase, ou seja, (i) o núcleo, (ii) seus especificadores e qualquer (iii) material adjungido acima do núcleo, constituem a borda da fase e são mantidos ativos na computação sintática.

O mapa estrutural da fase está representado, esquematicamente, na estrutura em (21), em que *x* corresponde ao núcleo da fase, *y* ao especificador do núcleo e *z* ao complemento do núcleo:



Na seção 2, chegamos à conclusão de que as raízes, em contexto

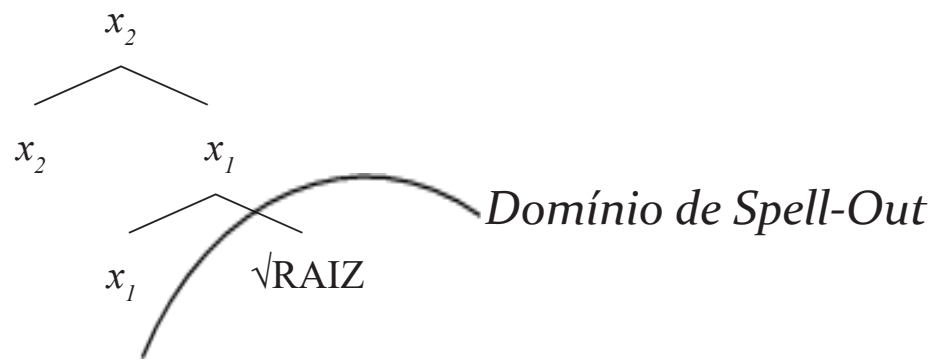
isolado, não apresentam um significado fixo, mas correspondem a uma potencialidade de significados. Sua interpretação somente é determinada quando elas se combinam a um núcleo categorial. Tendo em vista que o núcleo categorial serve como uma fronteira na determinação do significado das raízes, podemos assumir, seguindo a proposta de derivação por fases, que todo núcleo categorial corresponde a um núcleo de fase, e, por esse motivo, desencadeia o envio de seu complemento para as interfaces interpretativas. Nomearemos a associação entre a generalização em (10) e a abordagem de derivação por fases em (21) de abordagem Marantz-Arad, definida em (22):

(22) Abordagem Marantz-Arad

As raízes recebem uma interpretação no contexto do primeiro núcleo de fase – isto é, do primeiro núcleo categorial – a que se combinam, seja ele v , n , ou a . Sua interpretação é, então, preservada por toda a derivação.

Nossa função, a partir de agora, é associar a implementação formal descrita em (22) à discussão empírica que fizemos na última seção. Podemos dizer que, de acordo com a abordagem Marantz-Arad, qualquer caso de alossemia contextual, ou seja, de polissemia de uma raiz, deve ser resolvido no domínio de fase determinado pelo primeiro núcleo categorial. Para ilustrar como esse mecanismo sintático se aplica, tomemos a estrutura genérica em (23), admitindo, para tanto, que x corresponde a um núcleo categorial. Nessa estrutura genérica, temos a concatenação de uma $\sqrt{\text{RAIZ}}$ a dois núcleos categoriais, x_1 e x_2 , ou, em termos mais gerais, a dois morfemas derivacionais:

(23)

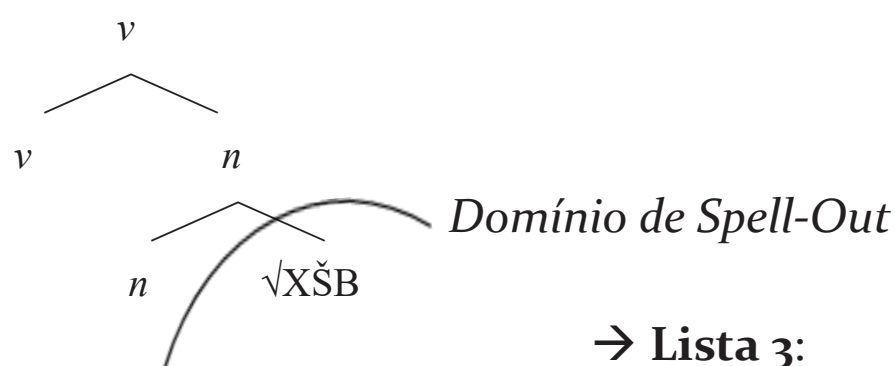


O núcleo x_1 , no momento em que se concatena a $\sqrt{\text{RAIZ}}$, desencadeia o envio dessa $\sqrt{\text{RAIZ}}$ – seu complemento – para as interfaces, já que todo núcleo categorial é um núcleo de fase, de acordo com a abordagem em (22). Entretanto, x_1 ainda permanece acessível para a $\sqrt{\text{RAIZ}}$ nas interfaces, pois ele se concatenou a ela localmente, o que faz dele um morfema interno. Em um segundo momento, x_2 se concatena à combinação [$x_1 + \sqrt{\text{RAIZ}}$]. Nesse passo da derivação, a $\sqrt{\text{RAIZ}}$ já está inacessível para x_2 , pois ela foi enviada para as interfaces no primeiro domínio de fase, estabelecido por x_1 . Portanto, x_2 não tem acesso direto à $\sqrt{\text{RAIZ}}$. Por outro lado, x_1 permanece acessível à x_2 , dada sua localidade estrutural, e, uma vez que x_2 é também um núcleo categorial, x_1 será enviado para as interfaces, pois está localizado em seu complemento.

Tendo em mente esse mecanismo formal, iremos, agora, explicar a atribuição dos significados das palavras *mixšev* ‘computar’, do hebraico, e *globalizar*, do português, seguindo a abordagem em (22).

Iniciemos pelo verbo *mixšev* ‘computar’. Primeiramente, a raiz $\sqrt{\text{XŠB}}$ é concatenada ao núcleo categorial n , que a envia para as interfaces semântica e fonológica. Nesse momento, a Lista 3 passa a ter acesso à raiz $\sqrt{\text{XŠB}}$ e atribui a ela um de seus alossemas no contexto nominal, tal como ilustrado em (24).

(24)



→ Lista 3:

 $\sqrt{X\check{S}B}$, no ambiente [n],

 Σ_1 : ‘pensamento’

 Σ_2 : ‘computador’

 Σ_3 : ‘explicação’

Uma vez que a raiz foi enviada para a interface semântica, ela já não está mais acessível para a sintaxe, fazendo com que seu significado não possa ser alterado por núcleos funcionais mais externos, ou seja, por núcleos que se concatenem acima de n . Em decorrência disso, a concatenação do núcleo categorial v acima de n terá acesso somente ao alossema negociado no primeiro domínio de fase, o que faz com que o verbo *mixšev* ‘computar’ tenha sua interpretação composta ao significado do nome *maxšev* ‘computador’.

O mesmo raciocínio se aplica na formação da palavra *globalização*. A raiz \sqrt{GLOB} concatena-se ao núcleo categorial a , e, uma vez que todo núcleo categorial é um núcleo de fase, a raiz \sqrt{GLOB} é enviada para as interfaces. A Lista 3, então, atribui a essa raiz um dos alossemas que podem ocorrer no ambiente adjetival. Assim que o significado da raiz \sqrt{GLOB} é fixado, ele é preservado. Em seguida, v concatena-se a a , mas não tem acesso direto à raiz, pois ela já foi enviada para as interfaces, como ilustrado em (25). Consequentemente, v não pode condicionar a escolha de seus alossemas. O verbo formado a partir do adjetivo [$a + \sqrt{GLOB}$] tem, como consequência, seu significado dependente do significado que foi determinado no primeiro domínio de fase, resultando em uma

interpretação correspondente a “fazer algo se tornar mundialmente distribuído/conhecido, universal”.

(25)



→ **Lista 3:**

\sqrt{GLOB} , no ambiente [a],

Σ_1 : ‘**planeta Terra**’

Σ_2 : ‘~~todo, integral~~’

Σ_3 : ‘~~objeto esférico~~’

RESUMINDO

Concluindo, aprendemos, neste capítulo, como a Morfologia Distribuída explica a presença de polissemia e idiosincrasias nas expressões linguísticas. Mostramos que as irregularidades semânticas podem ser explicadas em termos de localidade estrutural e que essa localidade estrutural, por sua vez, pode ser formalizada em termos da assunção de fases no domínio da palavra. Essa nova perspectiva sobre a determinação das idiosincrasias semânticas nos permite explorar as irregularidades de uma língua de forma mais abrangente, desvendando quais são os domínios da estrutura sintática sensíveis a significados polissêmicos e especiais, em vez de apenas listar todas as idiosincrasias na forma de átomos pré-sintáticos não analisáveis. Para a Morfologia Distribuída, essa é uma das demonstrações de como fazer a sintaxe fazer sentido.

PARA SABER MAIS

Há, atualmente, diversos trabalhos debatendo a relevância e a melhor maneira de se delimitar os domínios de localidade na interpretação semântica, bem como a associação desses domínios aos domínios de localidade na interpretação fonológica. Além da abordagem Marantz-Arad, o leitor pode consultar a proposta de Heidi Harley, em *On the identity of roots* (2014), que vem ganhando notoriedade nos últimos anos. Há também diversos trabalhos que investigam os domínios de localidade semântica na formação de palavras do português, por exemplo: *Nominalizações em -ada em construções com o verbo leve dar em português brasileiro*, de Ana Paula Scher (2006), sobre as nominalizações em -ada; *Uma abordagem localista para morfologia e estrutura argumental dos verbos complexos (parassintéticos) do português brasileiro*, de Indaiá Bassani (2013), sobre a não composicionalidade em formações verbais parassintéticas; *A relação entre gênero e morfologia avaliativa nos nominais do português brasileiro: uma abordagem sintática da formação de palavras*, de Paula Armelin (2015), sobre as formações não composicionais com sufixos diminutivos e aumentativos; *Sobre aumentativos de verbos*, de Alessandro Boechat de Medeiros (2015), sobre a interpretação dos sufixos aumentativos de verbos; *Headedness and exocentric compounding*, de Vitor Nóbrega e Phoivos Panagiotidis (2020), sobre a interpretação não composicional de palavras compostas, entre outros. É importante destacar também que a Abordagem Marantz-Arad vem sendo desafiada por diferentes autores, entre eles Hagit Borer, em *The syntactic domain of content* (2013), Elena Anagnostopoulou e Yota Samioti, em *Domains within words and their meanings: a case study* (2014), e por Heidi Harley, em *On the identity of roots* (2014), entre outros.

EXERCÍCIOS

Exercício 1. Contraste, em termos gerais, a posição lexicalista e a posição não lexicalista defendida pela Morfologia Distribuída sobre a formação e o uso de formas polissêmicas e idiomáticas na Gramática.

Exercício 2. Considere as formações com sufixos diminutivos, *-inho/a* e *-zinho/a*, e aumentativos, *-ão/ona* e *-zão/ona*, abaixo, retiradas de Armelin (2015):

<i>Caipirinha</i>	<i>Quentão</i>	<i>Mapazinho</i>
<i>Carrinho</i>	<i>Panelazona</i>	<i>Coroinha</i>
<i>Pastelzão</i>	<i>Poetazinho</i>	<i>Cadernão</i>
<i>Pazona</i>	<i>Camisinha</i>	<i>Serpentezinha</i>
<i>Sacolão</i>	<i>Meninozinho</i>	<i>Palavrão</i>

A partir desse conjunto de dados, explique o contraste morfema interno vs. morfema externo, identificando a natureza interna ou externa desses sufixos. Justifique suas conclusões com base na interpretação semântica de cada palavra.

Exercício 3. Scher (2006) assinala que, no português, há verbos denominais cujos significados não acarretam a existência de seus nomes correspondentes, ou seja, a ação que esses verbos descrevem não implica o uso do nome associado, tal como ilustrado na sentença (a). Paralelamente, a autora mostra que há verbos cujos significados acarretam o nome correspondente, como se observa na não aceitação da sentença (b).

(1)

- (a) Ibn-e-Muljin o apunhalou com uma espada envenenada.
(SCHER, 2006, p. 37)
- (b) *Ele parafusou aquele quadro na parede com um prego.
(SCHER, 2006, p. 38)

Com base em nossas discussões sobre localidade estrutural e interpretação semântica, explique a formação dos verbos *apunhalar* e *parafusar* utilizando a abordagem Marantz-Arad. Forneça, em sua explicação, o passo a passo da derivação sintática de cada verbo.

