UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LINGUÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM SEMIÓTICA E LINGUÍSTICA GERAL

VITOR AUGUSTO NÓBREGA

Tópicos em Composição: Estrutura, Formação e Acento

Dissertação apresentada ao Programa de Pós-Graduação em Semiótica e Linguística Geral do Departamento de Linguística da Universidade de São Paulo, para obtenção do título de Mestre em Letras.

Área de Concentração: Semiótica e Linguística Geral

Orientadora: Prof.^a Dr.^a Ana Paula Scher

Versão Corrigida
SÃO PAULO
2014

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Catálogo da Publicação

Serviço de Documentação

Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo

Nóbrega, Vitor Augusto.

Tópicos em Composição: Estrutura, Formação e Acento/ Vitor Augusto Nóbrega;
orientadora: Ana Paula Scher. – São Paulo, 2014.

243 f. : fig.

Dissertação (Mestrado – Programa de Pós-Graduação em Semiótica e Linguística Geral. Área de Concentração: Semiótica e Linguística Geral) – Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo.

1. Composição. 2. Classificação dos Compostos. 3. Acento. 4. Teoria do *Merge*. 5. Fonologia Prosódica. 6. Morfologia Distribuída. I. Scher, Ana Paula. II. Título.

Folha de aprovação

VITOR AUGUSTO NÓBREGA

TÓPICOS EM COMPOSIÇÃO: ESTRUTURA, FORMAÇÃO E ACENTO

Dissertação apresentada ao Programa de Pós-Graduação em Semiótica e Linguística Geral do Departamento de Linguística da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo como requisito para obtenção do título de Mestre em Letras.

Área de Concentração: Semiótica e Linguística Geral

Aprovado em:

BANCA EXAMINADORA

Prof.^a Dr.^a Ana Paula Scher
Orientadora
Departamento de Linguística

Prof.^a Dr.^a Maria Filomena Spatti Sandalo
Instituto de Estudos da Linguagem
Universidade Estadual de Campinas

Prof. Dr. Marcello Modesto dos Santos
Departamento de Linguística
Universidade de São Paulo

Prof. Dr. Alessandro Boechat de Medeiros (Suplente)
Departamento de Linguística e Filologia
Universidade Federal do Rio de Janeiro

Prof.^a Dr.^a Esmeralda Vailati Negrão (Suplente)
Departamento de Linguística
Universidade de São Paulo

Aos meus pais, Carlos e Ana Flávia, porque
continuam sempre ao meu lado.

Agradecimentos

Agradeço, primeiramente, à minha querida orientadora, Ana Paula Scher, pelo incentivo desde meus primeiros passos na pesquisa, por acreditar em mim e em minha capacidade. Sou infinitamente grato pelo apoio, pela generosidade, pela compreensão, e por me ensinar a cada instante a ser um profissional sério e dedicado. Sua presença, como mestre e amiga, foi e é essencial para mim.

Agradeço aos meus pais, pela paciência, pelo amor incondicional, pelo amparo nos momentos difíceis, pela preocupação, pelo estímulo, e por me dar forças durante a elaboração do trabalho. Eles que, embora sem entender por que alguém deseja estudar linguística, não mediram esforços em permitir que eu continuasse meus estudos onde quer que fosse e no que quer que fosse. A vocês, meu eterno amor e gratidão.

Agradeço ao Grupo de Estudos em Morfologia Distribuída da USP (GREMD – USP) por contribuir imensamente em minha formação intelectual ao longo desses anos. Através das discussões, pude aprender muito do que sei, não apenas teoricamente, mas profissionalmente. Obrigado a todos os membros, meus amigos de grande estima e respeito, Ana Paula Scher, Rafael Minussi, Indaiá Bassani, Júlio Barbosa, João Paulo Lazzarini-Cyrino, Paula Armelin, Janayna Carvalho e Aline Garcia Roderio-Takahira.

Agradeço ao CNPq pelo apoio financeiro concedido a mim (processo 134069/2012-9), sem o qual eu não poderia ter levado a cabo o projeto desta pesquisa. Agradeço também ao Departamento de Linguística da USP e ao Programa de Pós-Graduação em Semiótica e Linguística Geral por ter me acolhido durante esses dois anos, fornecendo-me material e financiamento para o desenvolvimento e divulgação do meu trabalho.

Aos professores Esmeralda Vailati Negrão e Marcello Modesto pela leitura atenta e pelos comentários feitos em meu Exame de Qualificação. Suas sugestões e apontamentos foram essenciais para os desenvolvimentos posteriores do trabalho.

Aos professores do Departamento de Linguística, Esmeralda Vailati Negrão, Luciana Storto, Marcello Modesto, Raquel Santana Santos, Paulo Chagas de Souza, Marcos Lopes e

Evani Viotti, que me acompanharam durante esses dois anos, contribuindo decisivamente em meu trabalho.

Ao Rafael Minussi, pelo carinho fraterno, pelo exemplo de caráter, pela tolerância e pela paciência. A você, que sempre esteve ao meu lado me motivando e dando forças para continuar o trabalho, na certeza de que vale a pena insistir.

Ao Marcus Lunguinho, pelos conselhos, pelo auxílio nos momentos de dúvida, pela paciência em dedicar seu tempo a ouvir minhas ideias e análises, e também pelas boas conversas e risadas.

À Ceci Farias, pela cumplicidade, pelos momentos de filosofia e confabulação, pelas reuniões tão agradáveis em sua casa, e por ser esse ombro amigo que sempre me acolhe nos momentos de ansiedade.

Aos meus amigos mais que queridos, com os quais dividi momentos muito felizes, e que foram pacientes, compreendendo meus momentos de ausência durante esses dois anos, André Castro, Carol Simões, Gabriel Frey e Henrique Primon.

Aos colegas de departamento, pelas discussões e pelas trocas de experiências, Lara Frutos, Bruna Seixas, Carolina Tomasi, Mariana Resenes, Paula Bauab Jorge, Livia Oushiro, Alexandre Bueno, Lidia Lima, Jéssica Costa, Fernanda Canever, Luciana Sanchez Mendes, Graziela Bohn, Ivan Rocha, Carolina Alves, Karin Vivanco e Fernanda Rosa.

Aos meus irmãos Guilherme e Gustavo, e ao Tufão, meu cachorro-irmão, pelo apoio indireto, e por representarem uma parte tão querida de mim.

À Erica Flávia, auxílio imprescindível em todas as etapas decisivas deste trabalho, a você meu muitíssimo obrigado pela colaboração, pela solicitude e pela amizade.

À professora Antonietta Bisetto, por ter me recebido na Alma Mater Studiorum – Università di Bologna, para discutir partes deste trabalho. Agradeço-lhe também por me fornecer uma vasta quantidade de material, sem os quais, o desenvolvimento desta dissertação não teria sido possível.

Agradeço, por fim, aos professores que se dispuseram a discutir e comentar meu trabalho ao longo desses dois anos, contribuindo em vários aspectos durante sua elaboração, Susi Wurmbrand, Hedde Zeijlstra, Artemis Alexiadou, Marcel den Dikken, Andrew Nevins, Susana Béjar, Sam Mchombo, Philip Miller, Edit Doron, Sergio Scalise, Luiz Carlos Schwindt, Rosa Estopà Bagot e Geert Booij, a vocês meus sinceros agradecimentos.

“It is important to learn to be surprised by simple things – for example, by the fact that bodies fall down, not up, and that they fall at a certain rate; that if pushed, they move on a flat surface in a straight line, not a circle; and so on. The beginning of science is the recognition that the simplest phenomena of life raise quite serious problems: Why are they as they are, instead of some different way?”

Noam Chomsky

Resumo

NÓBREGA, Vitor Augusto. **Tópicos em Composição: Estrutura, Formação e Acento**. 2014. 243 f. Dissertação (Mestrado) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2014.

O objetivo geral desta dissertação é fornecer um quadro amplo da formação dos compostos nas línguas naturais, deflagrando o que deve ser visto como universal em sua formação e o que deve ser visto como particular às línguas. Partindo de uma língua específica, o português brasileiro, demonstramos como a composição se estabelece em um sistema particular, e, no contraste entre as propriedades de seus dados e as propriedades atestadas nesse processo em trabalhos taxonômicos e tipológicos, delineamos as fronteiras entre o geral e o específico. Admitimos que sejam duas as propriedades universais da composição que devem ser abarcadas em um sistema gerativo: (i) o estabelecimento de uma relação gramatical entre os membros de um composto, a saber, uma relação de subordinação, atribuição ou coordenação – fato bem comprovado translinguisticamente nos trabalhos de Bisetto e Scalise (2005), Guevara e Scalise (2009), e Scalise e Bisetto (2009) –, e (ii) a criação de um domínio categorial acima de dois núcleos complexos conectados em determinada relação gramatical, o qual garante que uma estrutura complexa quase-sentencial seja interpretada como uma única unidade sintática. O particular nesse processo de formação de palavras reside no modo como as línguas naturais emolduram morfologicamente seus compostos, ou se a partir de uma combinação entre radicais, ou se a partir de uma combinação entre palavras, bem como os processos fonológicos que determinam a composição em um sistema linguístico específico. Essa assunção é evidenciada pela variedade de estruturas morfológicas encontrada nos compostos translinguisticamente, e pela assistematicidade de processos fonológicos que se aplicam nos compostos a fim de demarcá-los em uma língua. Para explicar esses fatos, assumimos um modelo não-lexicalista de gramática, a Morfologia Distribuída (cf. HALLE; MARANTZ, 1993; EMBICK; NOYER, 2007), e, com base nesse aparato teórico, demonstramos que as propriedades universais da composição são diretamente abarcadas no componente sintático. As relações gramaticais, primeiramente, são formalizadas através dos tipos de aplicação da operação *Merge*, tal como definidos em Chomsky (2000, 2004), nomeadamente, *set-Merge* e *pair-Merge*, enquanto que o domínio categorial é formado a

partir da concatenação de um núcleo definidor de categoria acima de duas ou mais raízes categorizadas. As estruturas morfológicas, por sua vez, serão emolduradas pós-sintaticamente, nos componentes morfológico e fonológico da gramática, onde a variação translinguística se processa. Para tanto, argumentamos que a informação de classe carrega um papel importante na variação estrutural dos compostos translinguisticamente, pois será a presença ou ausência dessa informação que determinará quando uma raiz sintática será um radical ou uma palavra, distribuindo, desse modo, as estruturas sintáticas dos compostos em combinações morfológicas variadas.

Palavras-chave: composição; classificação dos compostos; acento; teoria do *Merge*; Fonologia Prosódica; Morfologia Distribuída.

Abstract

NÓBREGA, Vitor Augusto. **Topics in Compounding: Structure, Formation and Stress**. 2014. 243 f. Master's Thesis – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2014.

The main goal of this thesis is to provide a broad picture of compounding in natural languages, triggering what should be considered universal and what should be considered language-specific in this word formation process. Starting from a particular language, Brazilian Portuguese, we demonstrate how compounding is established in a linguistic system, and by the contrast of the properties present in its compounds with the attested properties provided in taxonomic and typological works on the subject, we outline the boundaries between what is general and what is specific to this language. We argue that there are two universal properties of compounding that should be accounted for a generative system: (i) the establishment of a grammatical relation between the constituents of a compound, namely subordination, attribution and coordination – a crosslinguistically well-proven fact in the works of Bisetto and Scalise (2005), Guevara and Scalise (2009) and Scalise and Bisetto (2009) – and (ii) the formation of a categorial domain above two constituents connected in a specific grammatical relation, which ensures that a quasi-sentential structure will be interpreted as a single syntactic unit. What is language-specific in this word formation process lies in how languages frame their compounds in morphological structures, whether in a combination of stems or in a combination of words, as well as the phonological processes that specify compounding in some linguistic systems. This assumption finds evidence in the variety of morphological structures found in compounds through the languages of the world, and in the unsystematicity of phonological processes that apply to compounds to demarcate it in a particular language. In order to explain all the above-mentioned facts, we assume a non-lexicalist approach to grammar, the Distributed Morphology framework (see HALLE; MARANTZ, 1993; EMBICK; NOYER, 2007), and based on its tenets we demonstrate that the universal properties of compounding are straightforwardly accounted for the syntactic component of the grammar. Primarily, the grammatical relations are formalized through the nature of the operation Merge, as defined in Chomsky (2000, 2004), namely, set-Merge and pair-Merge, while the categorial domain is created by the concatenation of a category-defining head above

two or more categorized roots. The morphological structures of the compounds will be framed post-syntactically at the morphological and phonological components, where crosslinguistic variation takes place. To this end, we argue that class marker information plays an important role in the variation of compounds' morphological structure crosslinguistically, since it will be the absence or the presence of this feature that will define whether a syntactic root will be externalized as a stem or as a word, distributing this way the syntactic structure of a compound in various morphological combinations.

Key-words: compounding; classification of compounds; stress; Merge; Prosodic Phonology; Distributed Morphology.

Lista de Figuras

Figura 1 – A Arquitetura da Gramática para a Morfologia Distribuída.....	32
Figura 2 – O ramo de PF da gramática adaptado de Embick e Noyer (2001).....	34
Figura 3 – A Arquitetura da Gramática e os pontos de acesso das Listas.....	38
Figura 4 – Modelo da Fonologia Lexical no PB para Lee (1995).....	52
Figura 5 – Arquitetura da Gramática para a Fonologia Lexical.....	174

Lista de Tabelas

Tabela 1 – Distribuição percentual das combinações categoriais dos compostos em diferentes famílias linguísticas (GUEVARA; SCALISE, 2009, p. 120).....	81
Tabela 2 – Distribuição percentual das relações gramaticais translinguisticamente (GUEVARA; SCALISE, 2009, p. 119).....	83

Lista de Quadros

Quadro 1 – Combinações da estrutura morfológica dos compostos em tzotzil (ŠTEKAUER; VALERA; KÖRTVÉLYESSY, 2012, p. 42 <i>apud</i> COWAN, 1969, p. 89).....	76
Quadro 2 – Variedade de combinações categoriais dos compostos.....	80
Quadro 3 – Distribuição categorial, relacional e estrutural dos compostos do PB.....	84
Quadro 4 – Exemplos de compostos na classificação de Bisetto e Scalise (2005).....	95
Quadro 5 – Constituição morfológica e fonológica dos compostos (NESPOR, 1999).....	184
Quadro 6 – Constituição morfológica e fonológica dos compostos do PB.....	198

Lista de Abreviaturas e Siglas

3	Terceira pessoa
A/ADJ	Adjetivo
<i>a</i>	Categorizador adjetival
ACC	Caso acusativo
Adv/ADV	Advérbio
AR	<i>Agree</i> Reverso
AP	Sintagma adjetival
Asp	Aspecto
AspP	Sintagma aspectual
BPS	<i>Bare Phrase Structure</i>
CL	Composto lexical
CP	Sintagma complementizador
CPL	Composto pós-lexical
D	Determinante
D _M	Derivação morfológica
DP	Sintagma determinante
DS	Estrutura profunda
D _S	Derivação sintática
EL	Elemento de ligação
FEM	Feminino
FL	Fonologia Lexical
<i>i</i> F	Traço interpretável
INFL	Flexão
IV	Item de Vocabulário
LF	Forma Lógica
MASC	Masculino
MC	Marcador de classe
MCF	Formação de constituinte morfológico
MD	Morfologia Distribuída
N	Nome
NOM	Caso nominativo

Num	Numeral
NP	Sintagma nominal
<i>n</i>	Categorizador nominal
PCF	Formação de constituinte prosódico
PF	Forma Fonética
PL	Plural
PM	Programa Minimalista
PP	Sintagma preposicionado
Prep	Preposição
SG	Singular
SLH	<i>Strict Layer Hypothesis</i>
SPE	<i>Sound Patterns of English</i>
SUJ	Sujeito
T	Tempo
TA	Teoria da Assimetria
R	Traço de borda da raiz
<i>u</i> F	Traço não-interpretável
V	Verbo
<i>v</i>	Categorizador verbal
<i>v</i>	Núcleo funcional introdutor de argumento externo
VF	Vogal final
VP	Sintagma verbal

Lista de Símbolos

$\sqrt{}$	Raiz sintática
$\Re$	Relações gramaticais internas aos compostos
φ	Traços de número, gênero e pessoa
σ	Sílaba
Σ	Pé
ω	Palavra fonológica
C	Grupo clítico
Φ	Frase fonológica
I	Frase entoacional
U	Enunciado

Sumário

Capítulo 1

Introdução	21
-------------------	----

Capítulo 2

Morfologia é Sintaxe	30
-----------------------------	----

2.1. A concepção de gramática	30
2.2 Os primitivos da gramática	36
2.2.1 Raízes	38
2.2.2 A constituição dos morfemas	40
2.2.3 Itens de Vocabulário	41
2.3 Derivação sintática: a Condição do <i>Merge</i> na formação de palavras	42
2.4 Conclusões preliminares	43

Capítulo 3

Estruturas Morfológicas da Composição	44
--	----

3.1 Caracterização do capítulo	44
3.2 As abordagens tradicionais de classificação dos compostos do PB	45
3.3 As estruturas morfológicas da composição	57
3.3.1 Composto as estruturas morfológicas: raízes, radicais e palavras	57
3.3.2 As estruturas morfológicas dos compostos do PB	60
3.3.2.1 Estruturas [radical + radical] e [radical + palavra]	60
3.3.2.2 Estrutura [palavra + palavra]	70
3.3.2.3 A variedade de estruturas morfológicas translinguisticamente	74
3.4 Mapeamento da composição: do geral ao específico	79
3.5 Conclusões preliminares	85

Capítulo 4

Unidades Mínimas da Composição	86
---------------------------------------	----

4.1 Caracterização do capítulo	86
4.2 A universalidade e os universais na composição: o problema da definição	88
4.2.1 A primeira unidade mínima: as relações gramaticais internas à composição	93
4.3 Abordagens lexicalistas sobre a composição	99
4.3.1 Teoria da assimetria e as projeções funcionais tipificadas para a composição	100
4.4 A composição na Morfologia Distribuída	106
4.4.1 A segunda unidade mínima: o domínio categorial	109
4.4.1.1 Problemas de projeção	111
4.4.1.2 Flexão independente nos compostos do grego moderno	113
4.4.1.3 Compostos parassintéticos	114
4.4.1.4 Exocentricidade morfológica	117
4.4.1.5 Estruturas de incorporação nominal	119
4.5 Conclusões preliminares	121

Capítulo 5	
Derivando Compostos	122
5.1 Caracterização do capítulo	122
5.2 A influência da operação <i>Merge</i> nos processos de formação de palavras: a composição	123
5.2.1 As formas da operação <i>Merge</i> e a composição na Morfologia Distribuída	128
5.2.2 O problema do primeiro <i>Merge</i> : a categorização das raízes	131
5.3 Condição do <i>Merge</i> : a formação de palavras a partir da valoração de traços	133
5.3.1 O que constitui as raízes?	139
5.3.2 O que constitui os núcleos categorizadores?	142
5.3.3 Derivando compostos	147
5.3.3.1 Derivação dos compostos subordinados	148
5.3.3.2 Derivação dos compostos atributivos	156
5.4 Conclusões preliminares	166
 Capítulo 6	
O Caminho para PF	167
6.1 Caracterização do capítulo	167
6.2 Ausência de isomorfia entre a sintaxe e a fonologia: a Fonologia Prosódica	169
6.2.1 Os domínios prosódicos	169
6.2.2 A palavra fonológica e a composição	175
6.3 Os padrões acentuais na composição	179
6.3.1 Estruturas morfológicas governando a aplicação de acento	179
6.3.2 Níveis lexicais governando a aplicação de acento	186
6.3.3 Os algoritmos acentuais no PB: entre Bisol (1992) e Lee (1995)	188
6.3.4 Quantas palavras fonológicas são necessárias para um composto no PB?	191
6.3.4.1 Compostos formados por radicais	192
6.3.4.2 Compostos formados por radical e palavra	195
6.3.4.3 Compostos formados por palavras	197
6.4 Morfologia Distribuída: em busca de evidências para um mapeamento indireto dos constituintes sintáticos	199
6.4.1 A noção de fase na palavra	202
6.5 Mapeamento dos constituintes sintáticos em constituintes prosódicos com a aplicação do acento	209
6.6 Conclusões preliminares	214
 Considerações Finais	215
 Referências Bibliográficas	218

CAPÍTULO 1

INTRODUÇÃO

A composição é considerada um processo de formação de palavras capaz de criar unidades complexas sintaticamente inseparáveis e que carregam, quase sempre, novos significados. Embora encontremos uma ampla variedade superficial nos dados gerados por esse processo de formação de palavras nas línguas naturais, defendemos, nesta dissertação, que sua formação é translinguisticamente uniforme e se processa inteiramente no componente sintático da gramática, a partir das mesmas operações responsáveis por gerar sentenças.

Nosso objetivo amplo, nesse sentido, é fornecer uma explicação sobre quais são as informações mínimas necessárias para a formação de um composto nas línguas naturais e quais informações presentes nesse processo são específicas de uma língua particular. Uma vez respondida essa questão, iremos demonstrar como essas informações gerais e específicas à composição estão codificadas na gramática e de que maneira obtemos as variedades superficiais encontradas nos compostos.

Para desenvolver este trabalho, tomamos como exemplar de língua particular o português brasileiro (doravante, PB), e contrastamos as propriedades morfológicas, morfossintáticas e fonológicas de seus dados com as mesmas propriedades atestadas em outros compostos translinguisticamente, as quais foram reunidas a partir de trabalhos descritivos, taxonômicos e tipológicos acerca do processo de composição. É dessa comparação que filtramos suas unidades mínimas, e distribuímos os tópicos a serem discutidos ao longo dos próximos capítulos.

Para levar a cabo tal objetivo amplo, dividimos a dissertação em quatro eixos temáticos principais, nomeadamente, (i) a caracterização do processo de composição em uma língua particular, no caso, o PB; (ii) a reunião das unidades mínimas da composição a serem abarcadas em um sistema gerativo universal; (iii) a explicação da formação sintática dos

compostos translinguisticamente e (iv) a explicação fonológica dos padrões acentuais dos compostos do PB, como um exemplar de externalização desses dados.

Assumimos como hipótese central que as línguas são uniformes de um ponto de vista gerativo e que sua variação superficial se deve aos componentes morfológico e fonológico da gramática, em consonância com o Princípio da Uniformidade, delineado em Chomsky (2001, p. 2), apresentado em (1):

(1) *Princípio da Uniformidade (CHOMSKY, 2001, p. 2, tradução nossa)*

Na ausência de evidências convincentes do contrário, assumo que as línguas são uniformes, com a variação restrita a propriedades facilmente detectáveis dos enunciados¹.

Além disso, defendemos que a formação de objetos linguísticos, sejam eles palavras ou sentenças, ocorre em um mesmo componente gerativo, a sintaxe, com base no arcabouço teórico da Morfologia Distribuída, modelo realizacional não-lexicalista (doravante, MD; cf. HALLE; MARANTZ, 1993; EMBICK; NOYER, 2007).

Assim sendo, partimos da descrição de nossa língua específica, o PB, com o intuito de demonstrar como a composição se distribui internamente em uma língua. Para tanto, percorremos as classificações feitas para os seus compostos, a fim de contrastar quais dessas classificações são parte intrínseca desse processo de formação de palavras, e quais delas podem ser vistas como particulares aos dados do PB. Objetivamos, com isso, estabelecer qual a distribuição mais ampla encontrada nos compostos de uma língua particular, e o que dessa distribuição caracteriza o processo de composição em um sistema linguístico específico.

Das classificações sugeridas para os compostos do PB, reunimos as cinco principais, a saber: (a) classificação dos compostos por combinações de categorias lexicais, (b) classificação dos compostos a partir das relações gramaticais entre seus membros, (c) classificação dos compostos pela forma morfológica de seus membros, (d) classificação dos compostos a partir da junção fonológica entre seus membros e (e) classificação dos compostos de acordo com o componente da gramática em que são formados.

Concentrando-nos nas classificações (a), (b) e (c), notamos que a classificação (c), das formas morfológicas, é uma propriedade da composição específica às línguas, pois reflete o modo como uma língua particular emoldura morfolologicamente seus compostos, por exemplo,

¹ Trecho original (CHOMSKY, 2001, p. 2): “In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances.”

se por uma combinação de radicais ou por uma combinação de palavras. Essa hipótese pode ser evidenciada pelo fato de que as combinações categoriais e as relações gramaticais entre os membros de um composto são uniformes e independem de sua estrutura morfológica para se estabelecer, caracterizando-se, portanto, como uma informação presumivelmente universal desse processo.

Tomemos a relação gramatical de modificação, por exemplo, ocorrendo em compostos de combinação categorial N-N, em línguas tipologicamente distintas, como em (2). O que percebemos é que, embora a relação gramatical e a combinação categorial sejam as mesmas, o modo como os compostos são realizados morfológicamente varia de língua para língua²:

(2) [raiz presa + raiz presa]

a. <i>diàn-zǐ</i>	eletricidade-partícula	‘elétron’	(Chinês)
b. <i>tou-hoku</i>	leste-norte	‘nordeste’	(Japonês)

[radical + radical]

c. <i>gwiazd-o-zbiór</i>	estrela-EL-coleção	‘constelação’	(Polonês)
d. <i>jazyk-o-věda</i>	língua-EL-ciência	‘linguagem científica’	(Theco)

[radical + palavra]

e. <i>dent-o-bucal</i>	dent-EL-bucal		(PB)
f. <i>ayri-o-ánthropos</i>	selva-EL-homem	‘homem selvagem’	(Grego)

[palavra + palavra]

g. <i>sword fish</i>	espada-peixe	‘peixe espada’	(Inglês)
h. <i>célul-a madr-e</i>	célula-mãe	‘célula mãe’	(Espanhol)

Por outro lado, se considerarmos as estruturas morfológicas dentro de uma mesma língua, percebemos que os compostos não são formados com base em uma única combinação estrutural, mas podem ser realizados por duas ou mais combinações. O PB, por exemplo, conta com três tipos estruturais para a realização de seus compostos, a saber: [radical + radical] (e.g., *palpitometria*), [radical + palavra] (e.g., *neurociência*) e [palavra + palavra]

² Os dados presentes no exemplo (2) foram retirados dos seguintes autores: do chinês, Pirani (2008); do japonês, Nóbrega (2011); do polonês, Mukai (2013); do theco, Štichauer (2009); do grego, Nespor e Ralli (1996) e, do espanhol, Kornfeld (2009).

(e.g., *algodão doce*), embora privilegie um tipo estrutural em particular, a combinação [palavra + palavra]. É importante ressaltar, no entanto, que os dois primeiros tipos estruturais supramencionados, [radical + radical] e [radical + palavra], foram muitas vezes negligenciados nas descrições dos compostos do PB, por serem considerados periféricos ou “menos comuns”, já que são formados, em grande medida, por radicais presos de origem grega ou latina (e.g., *antropomorfia*, *hidroginástica*, *neuroestimulação*), comumente rotulados de compostos neoclássicos ou compostos eruditos.

Entretanto, essas estruturas negligenciadas são muito mais abrangentes do que a literatura vinha assumindo até então, pois suas formações não se restringem a radicais neoclássicos exclusivamente, uma vez que tais estruturas também são constituídas pelos demais radicais da língua (e.g., *cervejochato*, *dentobucal*, *visuogestual* e *dilmofóbico*). Desse modo, os compostos neoclássicos deixam de ser vistos como um conjunto de dados periféricos e passam a ser diluídos na visão geral dos compostos do PB, já que essa língua é detentora, segundo nossa visão, de três possíveis formas de externalização para os compostos.

O que argumentamos, dentro dessa ideia, é que as estruturas [radical + radical] e [radical + palavra] estão em um mesmo nível que a estrutura [palavra + palavra], já que, de acordo com nossa hipótese, a estruturação morfológica de um composto é um reflexo de como sua estrutura sintática é realizada fonologicamente – assunção essa que se estende de modo translinguístico. Essa proposta sintetiza o Princípio da Uniformidade na formação de palavras, pois admite que sua geração independa do modo como sua estrutura abstrata será externalizada.

Tendo em vista o fato de que as estruturas morfológicas são específicas às línguas, sua emolduração deve ocorrer nos componentes da gramática onde a variação translinguística se processa, a saber, nos componentes morfológico e fonológico. Por outro lado, para dar conta da formação sintática desses dados, devemos reunir as unidades mínimas a que o componente sintático deve ter acesso a fim de garantir a formação de um composto em qualquer língua natural – de acordo com o Princípio da Uniformidade – e, para obter tais unidades mínimas, devemos olhar para o fenômeno da composição de modo amplo.

Para tanto, buscamos nas generalizações feitas em trabalhos taxonômicos de caráter universalista (cf. BISETTO; SCALISE, 2005; GUEVARA; SCALISE, 2009; SCALISE; BISETTO, 2009) quais informações podem ser consideradas universais nesse processo de formação de palavras. Levando em consideração que os trabalhos supracitados foram elaborados a partir de um grande volume de dados, filtramos, a partir de seus levantamentos, nossa primeira unidade mínima: as relações gramaticais entre os membros de um composto.

Essas relações gramaticais internas à composição são três: (i) subordinação (i.e., relação predicação-argumento), (ii) atribuição (i.e., relação núcleo-modificador) e (iii) coordenação (i.e., relação conjuntiva ou disjuntiva), as quais são descritas como sendo as mesmas relações estabelecidas entre constituintes sentenciais (cf. GUEVARA; SCALISE, 2009, p. 104). Sendo assim, o componente sintático consegue abarcar o estabelecimento dessas relações gramaticais de modo direto, a partir da natureza da operação *Merge*³ que venha a se estabelecer entre dois elementos (cf. CHOMSKY, 2000, 2004). Essa uniformidade, a nosso ver, é o que evidencia a razão pela qual as relações gramaticais constituem parte intrínseca da combinação sintática dos membros de um composto.

A segunda unidade mínima da composição é um domínio categorial criado acima da concatenação entre dois ou mais constituintes em determinada relação gramatical. A assunção de que tal domínio é formado encontra evidências em um conjunto variado de dados, e é capaz de explicar uma série de problemas relacionados à composição, tais como: (i) a exocentricidade categorial (i.e., a rotulação de um composto por uma categoria diferente daquela de seus membros; e.g., $N+N \rightarrow A$); (ii) a liberdade de projeção (i.e., fato que ocorre em compostos V-N, em que tanto V quanto N podem rotular a estrutura), (iii) o duplo licenciamento de traços de seleção (i.e., compostos verbais que exigem argumentos adicionais além daqueles previstos na grade argumental de um de seus membros) e (iv) os efeitos de integridade lexical (i.e., a impossibilidade de um dos constituintes interno ao composto servir como um elemento anafórico em relações de ligação).

A soma dessas duas unidades mínimas – relação gramatical entre os membros de um composto e o domínio categorial acima de sua estrutura – leva-nos a definir a formação sintática dos compostos nos seguintes termos:

(3) *Composto*

Um composto é formado quando dois ou mais núcleos complexos, em determinada relação sintática, são recategorizados por um núcleo definidor de categoria – *n*, *v* ou *a*.

Nessa definição, os núcleos complexos devem ser interpretados como uma raiz concatenada a um ou mais núcleos definidores de categoria, que serão, posteriormente, concatenados em uma relação sintática, seja ela de subordinação, atribuição ou coordenação, e, então, essa estrutura sintática é categorizada por um outro núcleo definidor de categoria,

³ Embora alguns autores tenham traduzido o termo para o português (e.g., “*Compor*”, na tradução do Programa Minimalista feita por Eduardo Raposo, e “*Mergir*”, cf. LEMLE, 2013) faremos uso do termo em inglês.

responsável por fornecer a informação categorial, as informações flexionais e a utilização do composto como uma única unidade sentencial.

No desenvolvimento da dissertação, iremos nos concentrar, mais detidamente, nos compostos formados por uma relação de subordinação e de atribuição, com vias a fornecer uma alternativa de formalização para esses tipos de compostos, desde a sua geração no componente sintático até à sua externalização em PF. Tomamos como objeto principal os compostos do PB, apontando, paralelamente, as extensões que as análises sugeridas para seus dados podem ter para os demais tipos de dados utilizados ao longo dos capítulos.

Por último, concentramo-nos no ramo de PF, a fim de demonstrar como a estrutura sintática de um composto é mapeada em uma estrutura fonológica, ao demonstrar o modo como se dá a variação morfológica supramencionada e como as regras de acento são aplicadas em sua estrutura linearizada. Assumimos, para tanto, a visão da fonologia prosódica, de que a estrutura sintática não é mapeada diretamente em uma estrutura fonológica, (cf. NESPOR; VOGEL, 1986), e que a formação de domínios prosódicos faz parte do ramo de PF, sugerindo, nesse sentido, que a formação da palavra fonológica pode ser regulada de acordo com a teoria de fase proposta em Embick (2010).

Argumentamos, ainda, que a presença ou ausência da informação de classe nos núcleos categorizadores de uma raiz carrega um papel importante na variação estrutural dos compostos translinguisticamente, pois será essa informação que determinará quando uma raiz sintática será um radical ou uma palavra, e quando a regra de acento deverá se aplicar à sequência linearizada. Desse modo, distribuímos as estruturas sintáticas dos compostos em combinações morfológicas variadas, e garantimos, ao mesmo tempo, a aplicação correta das regras de acento.

A seguir, ilustramos o modo como entendemos a formação de um composto, desde sua formação no componente sintático até à sua externalização, nos componentes morfológico e fonológico da gramática, a partir dos compostos V-N, em (4):

(4) *Distribuição dos compostos V-N translinguisticamente*⁴

[radical + palavra]

a. *lam-i-strajk*

(Polônês)

quebra-EL-greve

‘fura-greve’

⁴ EL = elemento de ligação; MC = marcador de classe; VF = vogal final.

- b. *scheid-s-rechter* (Holandês)
 divide-EL-juiz
 ‘árbitro’

(BISETTO, 2013, p. 2)

[palavra + palavra]

- c. *m-tol-a-nkhâni* (Chichewa)
 MC-pega-VF-notícias
 ‘repórter’

(MCHOMBO, 2004, p. 117)

- d. *sī-jī* (Chinês)
 opera-máquina
 ‘motorista’

(BASCIANO; KULA; MELLONI, 2011, p. 233)

O que temos, de acordo com nossa proposta, é que embora os compostos em (4) sejam realizados a partir de estruturas morfológicas distintas, [radical + palavra] e [palavra + palavra], sua formação sintática ocorre de modo uniforme translinguisticamente, ao combinar um verbo mais um nome complemento através da satisfação de regras sintáticas gerais, as quais não têm influência direta sobre a maneira como tal combinação sintática será externalizada pela língua em questão.

É por essa razão que a realização de um composto como “formado por radicais” ou “formado por palavras” é uma consequência do modo como uma estrutura sintática básica é externalizada, uma vez que noções como “radical” e “palavra” dizem respeito às formas como uma raiz sintática será realizada fonologicamente no ramo de PF, a depender de como os traços gramaticais presentes em seu núcleo categorizador serão lidos pelo componente morfológico pós-sintaticamente, e, então, preenchidos com conteúdo fonológico.

Defendemos, portanto, que as *estruturas* morfológicas da composição são um reflexo das possibilidades que uma língua particular tem para realizar fonologicamente uma estrutura sintática, ao passo que a *formação* da estrutura sintática de um composto ocorre de modo uniforme translinguisticamente. A aplicação das regras de *acento*, por sua vez, tal como todo processo fonológico responsável por demarcar a composição em uma língua particular, deve ser tratada como uma informação específica, a ser resolvida nos componentes da gramática onde a variação translinguística ocorre.

Tendo em vista o que se apresentou até o momento, a dissertação segue estruturada do seguinte modo: no Capítulo 2, introduzimos o arcabouço teórico assumido nesta dissertação, a MD. Nele, expomos a visão clássica do modelo, assinalando, quando necessário, os pontos em que a análise a ser proposta no decorrer dos capítulos irá alterar ou endossar seus pressupostos teóricos.

No Capítulo 3, apresentamos a composição em nossa língua específica, o PB, discutindo como esse processo foi abordado e classificado pelos principais autores que trataram do tema. Com base nas classificações propostas para esses dados, apontamos quais delas dizem respeito a propriedades intrínsecas da composição, nomeadamente, as classificações por categorias lexicais e pelas relações gramaticais, e qual delas descreve uma propriedade particular à língua, a saber, as estruturas morfológicas, tal como argumentado acima.

Em seguida, descrevemos as estruturas morfológicas dos compostos do PB, a fim de fornecer um panorama geral da composição nessa língua, o qual embasará, empiricamente, os demais capítulos desta dissertação. Por fim, começamos a demarcar as fronteiras entre o geral e o específico a esse processo de formação de palavras, através da comparação das propriedades atestadas nos compostos do PB com as demais propriedades atestadas translinguisticamente, com o intuito de endossar o caráter particular associado às estruturas morfológicas.

No Capítulo 4, reunimos um conjunto de informações relacionais e estruturais mínimas para a formação sintática dos compostos, com vias a elencar o que é puramente sintático na geração desses dados. Para tanto, discutimos a universalidade da composição nas línguas do mundo, e o que desse processo pode ser considerado translinguisticamente bem atestado em seus dados. A partir desse levantamento, filtramos nossa primeira unidade mínima: as relações gramaticais internas à composição.

Logo após, apresentamos uma breve revisão das abordagens lexicalistas para esse processo, dando maior destaque à Teoria da Assimetria, proposta por Di Sciullo (2005b, 2009), uma vez que seus propósitos se assemelham àqueles levados a cabo nesta dissertação. Serão descritas as projeções funcionais internas aos compostos propostas pela autora, e será demonstrado que os argumentos em favor dessas projeções são amplamente questionáveis. Além disso, argumentamos pela presença de uma segunda unidade mínima para a composição que é a formação de um domínio categorial acima da relação estabelecida entre dois núcleos complexos, o que garante um conjunto de propriedades relativas a esse processo de formação de palavras.

No Capítulo 5, formalizamos as unidades mínimas da composição elencadas no Capítulo 4, a fim de proceder à derivação dos compostos. Para tanto, contextualizamos a operação *Merge* tal como elaborada em Chomsky (2000, 2004) e o modo como suas versões, *set-Merge* e *pair-Merge*, são realizadas no espaço computacional, com o propósito de traduzir as relações gramaticais nos termos dessa operação. Em seguida, desenvolvemos uma abordagem derivacional *bottom-up* para a MD, assumindo que a operação *Merge* é governada por um sistema de valoração de traços, de acordo com o modelo de seleção proposto por Wurmbrand (no prelo).

Posteriormente, serão revistos alguns posicionamentos teóricos, principalmente, a adequação da operação *Merge* para a formação de palavras à operação *Merge* assumida pela sintaxe minimalista. Assim, com base no sistema operacional supramencionado, demonstramos como ocorre a derivação de compostos subordinados e atributivos na gramática. Estabelecemos, nesse capítulo, uma uniformidade entre as operações (viz., *Merge*, *Agree*) e o modo como a derivação de uma sentença ocorre, ao mostrar que a formação de palavras e a formação de sentenças se encadeiam em um mesmo espaço computacional.

No Capítulo 6, após termos visto as estruturas morfológicas dos compostos do PB, no Capítulo 3, e o modo como os compostos são derivados no componente sintático, no Capítulo 5, discutimos outra informação específica às línguas, a saber: os processos fonológicos na composição. Retomando os dados do PB, nossa língua específica, concentramo-nos em um aspecto fonológico frequentemente reiterado acerca desses dados, a saber, a aplicação das regras de acento, mostrando como ela ocorre nas três diferentes estruturas morfológicas atestadas nessa língua. Por fim, trazemos as considerações finais sobre a dissertação.

Tópicos em Composição: Estrutura, Formação e Acento, portanto, visa abordar o fenômeno da composição de modo amplo e explicar como a formação dos compostos se processa como parte da linguagem humana. O trabalho reúne as informações mínimas para a *formação* sintática dos compostos das línguas naturais, e as informações específicas – *estrutura* e *acento* – baseando-se nos compostos do PB para sua elaboração. Como resultado, trazemos evidências para uma uniformidade na formação desses dados, argumentando que ela pode ser facilmente abarcada em uma visão sintaticocêntrica da formação de palavras.

CAPÍTULO 2

MORFOLOGIA É SINTAXE

Neste capítulo, apresentamos o arcabouço teórico sob o qual se assenta a dissertação. Empregamos uma visão não-lexicalista de gramática, tal como a delineada no modelo teórico da MD, na qual um único componente gerativo é responsável pela formação de objetos complexos, sejam eles palavras ou sentenças, a saber: a sintaxe. Expomos a visão clássica do modelo, apontando, quando necessário, os pontos em que a análise a ser proposta no decorrer da dissertação irá alterar ou endossar seus pressupostos teóricos.

2.1. A concepção de gramática

Adotamos a concepção de gramática delineada pela MD, principalmente, as versões de Halle e Marantz (1993, 1994), Marantz (1997) e Embick e Noyer (2001, 2007). Dentro dessa visão, dois pressupostos globais norteiam o modo como a arquitetura da gramática está estruturada e a divisão de trabalhos dentro da gramática para a formação de objetos complexos, a saber:

(5) *Pressupostos globais (adaptado de BOBALJIK, no prelo)*

- a. ***Syntax-all-the-way-down***: o modo básico de combinação na gramática tanto acima quanto abaixo do nível da palavra é sintático. A sintaxe opera nas unidades internas à palavra e, portanto, a formação de palavras é sintática.
- b. ***Inserção Tardia***: os itens manipulados pela sintaxe (i.e., os morfemas funcionais) são abstratos, ou seja, destituídos de conteúdo fonológico. O pareamento de traços fonológicos com os terminais sintáticos ocorre pós-sintaticamente, no mapeamento da sintaxe para a forma fonológica (PF).

Dentro da Gramática Gerativa, pelo menos nos últimos 40 anos, a resposta dominante para o modo como as palavras (i.e., unidades constituídas de propriedades gramaticais, fonológicas e semânticas) são formadas, era fornecida dentro de uma abordagem lexicalista. Nessa visão lexicalista, os pressupostos em (5) estão fora de questão, e a gramática é constituída de dois componentes gerativos, o léxico e a sintaxe, onde palavras são vistas como unidades privilegiadas e fora da jurisdição da sintaxe no que diz respeito à sua formação.

A assunção de (5a), como um pressuposto da gramática, estabelece um rompimento com uma visão de gramática composta por dois módulos gerativos, e relega somente à sintaxe a formação de objetos complexos, sejam eles palavras ou sentenças. Nesse sentido, a interface entre a sintaxe e a morfologia inexistente, tal como pontuam Embick e Noyer (2007, p. 290), pois a estrutura morfológica é estrutura sintática, e qualquer interface entre níveis que leve em consideração a formação das palavras é uma interface com a sintaxe.

Como consequência, o léxico, seja ele entendido como (i) unidades mínimas de significado para a construção de palavras, (ii) pareamento mínimo de som e significado ou (iii) local de armazenamento de idiossincrasias, é retirado da arquitetura da gramática, e em seu lugar fica disponível apenas uma lista de átomos sintáticos, os quais serão manipulados e, portanto, acessados pelo componente sintático para a construção de nós terminais complexos.

Essa proposta do funcionamento da gramática pode ser endossada com a formação dos compostos das línguas naturais, os quais representam uma clara demonstração da presença de estrutura sintática interna à palavra. Os trabalhos de Bisetto e Scalise (2005), Guevara e Scalise (2009) e Scalise e Bisetto (2009), embora dentro de uma abordagem lexicalista, fornecem evidências que endossam o pressuposto básico da MD, ao apontarem que as relações gramaticais presentes entre os constituintes dos compostos são basicamente as mesmas relações presentes entre constituintes sintagmáticos.

É importante notar que a gramática delineada pela MD critica, mais especificamente, a presença de dois sistemas gerativos, configurando-se como um modelo contrário à tese central do lexicalismo sobre a modularidade e a afirmação de que a “palavra” é um objeto especial quando o assunto é a gramática (cf. EMBICK; NOYER, 2007, p. 209). Portanto, a mudança de perspectiva proposta pela MD e, de modo geral, pelas abordagens não-lexicalistas⁵, não se

⁵ E aqui não nos referimos somente à MD, mas também às abordagens sintáticas para a morfologia, tais como Baker (1988), Pesetsky (1995) e Borer (2003, 2005, 2012).

circunscreve inteiramente na defesa de uma sintaxe interna às palavras, mas traz uma ruptura arquitetônica ao defender um único componente gerativo⁶.

Com isso, a arquitetura da gramática sofre algumas reformulações, embora mantenha o modelo em Y, e pode ser representada como na Figura 1, abaixo, conforme as representações trazidas por Embick e Noyer (2007) e Bobaljik (no prelo):

Figura 1 - A Arquitetura da Gramática para a Morfologia Distribuída

A arquitetura representada na Figura 1 mostra um único sistema gerativo, a sintaxe – indicada no esquema por “derivação sintática” – a qual faz uso das mesmas operações assumidas pela sintaxe minimalista, nomeadamente, *Merge*, *Move* e *Agree*. Os princípios que regem a formação de palavras são, por assim dizer, sintáticos, e a morfologia é vista como parte do mapeamento dos elementos formados pela sintaxe até à sua representação fonológica em PF.

Embick e Noyer (2007, p. 292) argumentam, ainda, que a abordagem proposta pela MD encontra apoio no Programa Minimalista (CHOMSKY, 1995; e trabalhos subsequentes), pois não há razão, a priori, para se assumir que uma teoria é mais econômica ao postular dois sistemas gerativos distintos em sua constituição do que apenas um. Os avanços desenvolvidos pelos trabalhos que endossam a presença de um único sistema gerativo responsável pela formação de objetos complexos – palavras e sentenças – contribuem para uma visão mais econômica e elegante da gramática.

Retomando o pressuposto global em (5b) e o interpretando dentro da arquitetura representada na Figura 1, percebemos que os átomos que entram no sistema computacional

⁶ Borer (2012) mostra que a presença de uma sintaxe interna à palavra já está presente em diversos trabalhos lexicalistas, o que pode ser visto nas propostas de Selkirk (1982), com a assunção de regras de reescrita e estrutura sintagmática; de Williams (1981), com a proposta de núcleo nas palavras, e na de Lieber (1980), com a proposta de quadros de subcategorização e estrutura de constituintes para os afixos.

não são signos – no sentido saussureano – tal como propunham as visões lexicalistas, pois, para a MD, a informação fonológica não é intrínseca aos feixes de traços gramaticais, e são inseridas apenas pós-sintaticamente. Dessa forma, a sintaxe gera objetos complexos constituídos de feixes de traços gramaticais dispostos hierarquicamente, os quais serão preenchidos com conteúdo fonológico após o *Spell-Out*.

Assim, o ramo de PF abriga um conjunto de operações pós-sintáticas que, além de serem responsáveis pela inserção de conteúdo fonológico, podem modificar e reelaborar a estrutura gerada pela sintaxe de maneira limitada. Por exemplo, requerimentos específicos de uma língua, tal como a realização de dois traços gramaticais por um único expoente fonológico, pode forçar a fusão dos nós terminais em que esses traços estão localizados a fim de viabilizar a sua realização por esse único expoente fonológico disponível (e.g., *-ot*, do hebraico, que é a realização fonológica dos traços de gênero e número, tal como *mor-ot* lit. professor-FEM.PL ‘professoras’).

A morfologia – enquanto um componente da gramática – é o que designa esse conjunto de processos pós-sintáticos que podem alterar a estrutura gerada pela sintaxe, a fim de adequá-la aos itens de vocabulário disponíveis na língua, ou seja, às regras que fornecem o conteúdo fonológico aos terminais sintáticos. Desse modo, o ramo de PF prepara a estrutura sintática para a o recebimento de informação fonológica, visto que a relação entre o feixe de traços gramaticais e o conteúdo fonológico pelos quais eles serão realizados não é sempre de um para um.

O ramo de PF, portanto, é entendido como uma derivação sequencial que resulta em uma representação fonológica, porém não deve ser visto como um sistema gerativo separado. Isso porque os processos que ocorrem em PF correspondem simplesmente a uma adequação estrutural ocasionada pela reorganização dos traços nos nós terminais sintáticos a fim de possibilitar a inserção do vocabulário de uma língua particular, e também organizar os expoentes fonológicos linearmente após sua inserção.

À guisa de ilustração, trazemos a Figura 2, abaixo, que exemplifica o sequenciamento das operações presentes no ramo de PF, as quais são responsáveis por adequar a estrutura gerada pela sintaxe ao vocabulário de uma língua particular. Os processos apontados na Figura 2 serão detalhados no Capítulo 6, quando discutirmos a externalização dos compostos do PB:

Figura 2: O ramo de PF da gramática adaptado de Embick e Noyer (2001)

Tendo em vista os pressupostos globais em (5) e a arquitetura da gramática representada na Figura 1, acreditamos que a concepção de gramática descrita neste capítulo é a que formaliza mais diretamente o Princípio da Uniformidade (CHOMSKY, 2001, p. 2), descrito na Introdução deste trabalho, pois permite que a sintaxe opere de maneira uniforme com os átomos sintáticos disponíveis, e que a morfologia, pós-sintaticamente, corresponda ao componente gramatical responsável pela variação superficial encontrada nas línguas naturais.

Ao mesmo tempo, esse modelo de gramática expressa claramente a concepção de que uma língua não é som com significado, como propunha Aristóteles, mas significado com som, tal como sugerido por Chomsky (2013)⁷:

“(T) Ordem e outros arranjos são uma parte periférica da linguagem, relacionados exclusivamente à externalização na interface SM [sensório-motora], na qual, claramente, são necessários. Se (T) configura-se como um princípio da GU, então a máxima de Aristóteles deveria ser modificada: a linguagem não é som com significado, mas sim significado com som (ou algum outro tipo de externalização), um conceito muito diferente, que reflete uma ideia tradicional diferente: a de que língua é essencialmente um instrumento do pensamento – “pensamento audível”, “a instrumentalidade falada do pensamento”, tal como William Dwight Whitney expressou a concepção tradicional.” (CHOMSKY, 2013, p. 36, tradução nossa)⁸

⁷ Vale mencionar que esse posicionamento já estava presente em Chomsky (2005, p. 3-4).

⁸ Trecho original (CHOMSKY, 2013, p. 36): “(T) order and other arrangements are a peripheral part of language, related solely to externalization at the SM [sensory-motor] interface, where of course they are necessary. If (T)

“O melhor resultado seria o de que a ordem linear nunca está disponível para a computação em língua-I, a não ser para a externalização em SM. Essa é a tese de maior alcance (T), que foi mencionada anteriormente em conexão com a teoria da ligação: a de que a ordem e outros arranjos são uma parte periférica da linguagem, reflexo das propriedades da interface SM [...] Assumindo-se que essa tese pode ser mantida, seja universalmente (o melhor resultado) ou com algumas qualificações, retornamos a uma concepção tradicional de linguagem como “um instrumento do pensamento”, tal como observado anteriormente. Disso segue-se que os usos particulares da linguagem que dependem da externalização, entre eles a comunicação, são aspectos ainda mais periféricos da linguagem, ao contrário da crença tão difundida como um dogma virtual, sem uma verificação séria.” (CHOMSKY, 2013, p. 39-40, tradução nossa)⁹

Admitimos, portanto, que a diferença superficial entre as línguas naturais esteja a cargo dos componentes morfológico e fonológico, localizado no ramo de PF. A sintaxe, aqui entendida, gera estruturas que independem das particularidades de uma língua, enquanto que o componente morfológico manipula os nós terminais organizados pela sintaxe a fim de adequá-los à distribuição superficial de uma língua específica. O componente fonológico, por sua vez, é o responsável pelos processos fonológicos que se aplicam na sequência linearizada de uma língua particular.

Na próxima seção, apresentamos os primitivos operados pela gramática para a formação de objetos complexos e como eles se distribuem pela sua arquitetura. Em seguida, introduzimos as reformulações propostas para esses primitivos de acordo com o que será defendido nos próximos capítulos.

holds generally as a principle of UG, then Aristotle’s dictum should be modified: language is not sound with meaning, but rather meaning with sound (or some other externalization), a very different concept, reflecting a different traditional idea: that language is primarily an instrument of thought – “audible thinking”, “the spoken instrumentality of thought”, as William Dwight Whitney expressed the traditional conception.”

⁹ Trecho original (CHOMSKY, 2013, p. 39-40): “The best result would be that linear order is never available for computation in I-language, apart from externalization to SM. That is more far-reaching thesis (T) mentioned earlier in connection with binding theory: the thesis that order and other arrangements are a peripheral part of language, reflexes of SM interface properties [...] Assuming the thesis can be upheld, either universally (the best outcome) or with some qualifications, we return to a traditional conception of language as “an instrument of thought”, as noted earlier. It would follow that particular uses of language that depend on externalization, among them communication, are even more peripheral aspects of language, contrary to belief so widespread as to be a virtual dogma, without serious support.”

2.2. Os primitivos da gramática

A noção de palavra, tal como concebida no lexicalismo, é desconstruída na MD com a retirada do léxico. Enquanto no lexicalismo, a palavra é inserida no componente sintático já com informações gramaticais, fonológicas e semânticas; na MD, essas informações estão dispostas pela gramática na forma de listas, cada uma com um ponto de acesso distinto na derivação – o que motiva o nome dado ao modelo.

A *Lista 1* contém os primitivos que entrarão na computação sintática para a formação de objetos complexos. Podemos pensar a *Lista 1* em analogia à constituição do léxico segundo Chomsky (2000). Para o autor, um léxico L é concebido como um subconjunto de traços $\{F_1\}$ selecionado a partir de um conjunto universal de traços $\{F\}$ disponibilizado pela Gramática Universal. Nesse sentido, podemos admitir que a *Lista 1* é um subconjunto de traços $\{F_1\}$ – os quais são estritamente gramaticais – selecionados por uma língua particular Y para serem utilizados como átomos sintáticos. Traços de ordem fonológica, por sua vez, embora sejam selecionados de um conjunto universal de traços $\{F\}$, ficarão alocados em outra lista, e serão acessados pós-sintaticamente.

Além desse subconjunto de traços gramaticais $\{F_1\}$, o inventário de primitivos que compõe a *Lista 1* é formado também por raízes, as quais são primitivos destituídos de traços gramaticais, na visão clássica do modelo. Em resumo, podemos definir os primitivos sintáticos contidos na *Lista 1* como em (6):

(6) *Primitivos sintáticos*

- a. **Morfemas abstratos:** são compostos exclusivamente de feixes de traços gramaticais destituídos de informação fonológica;
- b. **Raízes:** são unidades detentoras de conteúdo lexical e destituídas de conteúdo gramatical.

Essa divisão entre morfemas abstratos e raízes faz com que essas últimas entrem na derivação sintática destituídas de informação gramatical – e, consequentemente, categorial –, adquirindo-a através da concatenação com um morfema abstrato definidor de categoria (cf. MARANTZ, 1996, 1997). Isso permite que uma mesma raiz possa entrar em diferentes contextos sintáticos, a depender do núcleo categorizador com o qual se combina, a saber: n , v ou a . Essa proposta da acategorialidade das raízes está resumida em (7), e representada em (8), abaixo:

(7) *Hipótese das raízes acategorias (EMBICK; NOYER, 2007, p. 296, tradução nossa)*

Raízes não podem aparecer sem estar categorizadas. Raízes são categorizadas a partir da combinação com um núcleo funcional definidor de categoria.¹⁰

Dessa forma, após a sintaxe combinar os traços em estruturas hierárquicas, essas estruturas sofrerão *Spell-Out* e serão enviadas para as interfaces fonológica (PF) e semântica (LF) da gramática. No caminho para a interface fonológica, os nós terminais receberão informação fonológica através do processo de Inserção de Vocabulário. Nesse momento, será acessada a segunda lista, a *Lista 2* ou *Vocabulário*, onde estão os expoentes fonológicos, aqui chamados de itens de vocabulário (doravante, IVs).

Podemos entender o *Vocabulário* como a lista dos expoentes fonológicos correspondentes aos terminais sintáticos, em que cada um desses expoentes contém uma condição para inserção em um determinado nó¹¹. Dessa forma, cada pareamento de expoente fonológico com determinada informação sintática é chamado de IV. Como pontua Marantz (1996, p. 13), os IVs não são signos linguísticos, muito menos os átomos da computação sintática, mas sim os correspondentes fonológicos dos feixes de traços arranjados pelo componente sintático e, em alguns casos, rearranjados pelo componente morfológico.

Por fim, há o acesso à *Lista 3* ou *Enciclopédia*, a qual abriga as entradas enciclopédicas que caracterizam o conhecimento de mundo do falante, e que são responsáveis por relacionar os IVs a significados. Na *Lista 3*, também estão localizadas as informações extralinguísticas relacionadas à interpretação semântica dos objetos formados pela sintaxe, tais como significados idiomáticos e interpretações particulares das raízes em um contexto sintático-semântico específico.

¹⁰ Trecho original (EMBICK; NOYER, 2007, p. 296): “Categorization assumption: Roots cannot appear without being categorized; Roots are categorized by combining with a category-defining functional head.”

¹¹ Não entraremos na discussão sobre o caráter abstrato das raízes, ou seja, se elas também estão sujeitas à inserção de vocabulário tal como os morfemas abstratos, pois essa questão não influi diretamente nos problemas levantados neste trabalho. Para o leitor interessado no assunto, sugerimos os seguintes autores: Embick e Noyer (2007), Embick (2010), para uma visão de raízes com conteúdo fonológico; e Siddiqi (2009), Pfau (2009), Acquaviva (2009), Minussi (2012), Harley (no prelo) para uma visão abstrata das raízes.

Em resumo, a arquitetura da gramática, juntamente com o acesso das listas supramencionadas, fica esquematizada como na Figura 3, abaixo, adaptada de Embick e Noyer (2007, p. 301):

Figura 3 - A Arquitetura da Gramática e os pontos de acesso das Listas

Tendo em mente a organização da gramática e sua composição, apresentamos, brevemente, nas próximas subseções, algumas alterações ao modelo propostas nesta dissertação. Essas informações dadas de antemão serão relevantes para se entender o desenvolvimento e o percurso argumentativo dos próximos capítulos.

2.2.1. Raízes

A raiz ($\sqrt{}$) sofre algumas modificações na proposta a ser desenvolvida nesta dissertação. Na visão padrão da MD, as raízes são consensualmente destituídas de traços gramaticais e, a depender do autor, podem conter ou não traços fonológicos e semânticos desde a sua inserção no componente sintático. Entretanto, notamos um problema na realização do primeiro *Merge*, a saber, o *Merge* entre uma raiz e um núcleo categorizador, o qual envolve uma revisão da noção ontológica da raiz nesse modelo.

Como mencionado na última seção, a MD assume que as mesmas operações presentes na sintaxe minimalista são aquelas responsáveis pela formação das palavras. Desse modo, se assumirmos a definição de *Merge* tal como em Chomsky (2000, 2004), temos que o *Merge* ocorre, em linhas gerais, conjuntamente com a operação *Agree*¹², na qual um traço seletor

¹² A operação *Agree*, de acordo com Chomsky (2000, 2001), é definida da seguinte maneira: (i) um traço não-valorado (uma *sonda*) em um núcleo H sonda seu domínio de c-comando para encontrar um outro exemplar do

presente em um elemento sintático deverá ser valorado por outro elemento sintático a fim de possibilitar a concatenação.

A questão que se coloca, portanto, é a seguinte: se o núcleo definidor de categoria seleciona a raiz, e essa seleção ocorre como uma relação de *Agree*, analogamente a uma relação de sondagem (cf. CHOMSKY, 2000; COLLINS, 2002), qual é a forma do *Merge* responsável pelo primeiro *Merge* na MD? E se a relação entre esses elementos for, de fato, um caso de seleção, quais são os traços que estão em jogo nessa relação de *Agree*, uma vez que a raiz não contém traços gramaticais?

Com o intuito de adequar a teoria do *Merge* à formação de palavras e manter uma uniformidade operacional na sintaxe, assumimos que a raiz contém um traço de borda R não-interpretável não-valorado¹³, o qual possibilita a sua concatenação a um núcleo categorizador, seja ele *n*, *v* ou *a*, tal como sintetizado na Hipótese do Primeiro *Merge*, em (9):

(9) *Hipótese do Primeiro Merge*

Toda raiz contém um traço de borda R não-interpretável não-valorado [*u*R: ____] que poderá ser valorado apenas por um traço de borda R interpretável e valorado como *n*, *v* ou *a*.

A hipótese do Primeiro *Merge* intenciona abarcar de modo direto a hipótese das raízes acategoriais, apresentada em (7), e garante a liberdade de concatenação da raiz com qualquer núcleo definidor de categoria, ao mesmo tempo em que a torna visível no espaço computacional. Além disso, a assunção da hipótese do Primeiro *Merge* faz com que expliquemos a formação de palavras sem recorrer à postulação de um *Merge ad hoc*, algo indesejado para um modelo que pretende uniformizar o sistema gerativo.

A hipótese em (9) será desenvolvida no Capítulo 5, a partir de uma avaliação das possíveis explicações de como o primeiro *Merge* pode ocorrer dentro do modelo.

traço F (um *alvo*), com o qual concordará; (ii) se o alvo tem um valor, esse valor será atribuído como o valor da sonda. Essa operação será mais bem discutida no Capítulo 5.

¹³ A noção de traço de borda (do inglês, *edge feature*) é retirada de Chomsky (2005), e tenta captar a ideia de que os itens lexicais – no nosso caso, as raízes – contêm uma propriedade que os permite serem combináveis, a saber, um traço que sinaliza “eu sou concatenável”. Para nós, além dessa informação, esse traço de borda formaliza a necessidade que uma raiz tem de adquirir traços gramaticais e de se tornar visível no espaço computacional, algo que ocorre apenas quando elas são concatenadas a um núcleo categorizador.

2.2.2. A constituição dos morfemas

Na visão clássica da MD, os núcleos categorizadores fornecem apenas informação categorial às raízes quando concatenados. Na abordagem desenvolvida nesta dissertação, assumimos que os núcleos categorizadores carregam feixes de traços gramaticais selecionados do conjunto $\{F_1\}$ de traços disponíveis na língua, e não apenas o traço categorial. Esses núcleos categorizadores somam um conjunto de traços gramaticais que modelam o funcionamento de determinada raiz no ambiente computacional.

Dessa forma, hipotetizamos que há um conjunto de traços gramaticais passíveis de constituir um determinado núcleo categorial, e a opcionalidade de cada um desses traços fazer parte ou não do feixe de traços desse núcleo é o que ocasiona certos comportamentos computacionais e condiciona certos reflexos no componente morfológico da gramática, relacionados a uma raiz particular.

Por exemplo, para formar um nome, propomos que a *Lista 1* tem disponíveis os seguintes traços candidatos a compor o feixe de traços do núcleo categorizador nominal – os quais, por sua vez, podem dar origem a mais de um tipo de categorizador nominal a depender dos traços selecionados para sua constituição:

(10) *Conjunto de traços possíveis dos categorizadores nominais (n)*

- a. Traço de borda R interpretável e valorado como n [$iR: n$]
- b. Traço de tempo T não-interpretável valorado [$uT: val$]
- c. Traços ϕ interpretáveis não-valorados/valorados [$i\phi: __$]/[$i\phi: val$]
- d. Traço de classe não-interpretável valorado [$uC: 1$]

Um exemplo de comportamento computacional relacionado à seleção dos traços é a presença do traço de borda R, responsável por fornecer informação categorial a uma raiz, satisfazendo a hipótese em (9). O traço de tempo T, por sua vez, possibilita a concatenação de um núcleo categorizador a um sintagma verbal, ao satisfazer as exigências temporais do verbo, transformando esse sintagma em um composto¹⁴. Já um comportamento fonológico decorrente dessa seleção é dado pela presença ou ausência do traço de classe C, que indicará

¹⁴ Associamos a esse traço T um valor genérico de tempo responsável por interromper a formação de uma estrutura sentencial acima de um sintagma verbal (e.g., para a formação de um composto V-N, como *salva-vidas*, *limpa-vidros*). A implementação de como esse e os demais traços são operados pela sintaxe será vista, em detalhes, no Capítulo 5.

ao componente morfológico a possibilidade de inserção ou não de uma vogal temática, diferenciando, assim, um radical de uma palavra.

Vale mencionar que a teoria de traços assumida nesta dissertação enquadra-se em um sistema de traços multi-valentes, de acordo com a terminologia de Adger (2010). Nesse sistema, não há a noção de oposição binária, mas sim um conjunto maior de valores, no qual as semelhanças sintáticas são expressas no próprio sistema de traços. É por esse motivo que temos na estrutura de alguns traços a noção gramatical representada diretamente, por exemplo, a informação de número representada com um valor ‘singular’. Além disso, o sistema ainda permite traços com um valor positivo ou negativo, tal como ocorre com o traço de tempo T no feixe de traços do categorizador nominal, marcado positivamente como ‘*val*’, em (10).

Sinteticamente, a estrutura dos traços assumida nesta dissertação é dada como se segue:

(11) *Estrutura dos traços gramaticais (ADGER, 2010, p. 193, tradução nossa)*

- a. um traço é um par ordenado <Att, Val> em que
- b. Att é retirado de um conjunto de atributos, {A, B, C, D, E, ...}
- c. e Val é retirado de um conjunto de valores, {+, -, a, b, c, ...}.^{15/16}

No capítulo 5, fornecemos o conjunto de traços que podem integrar os núcleos categorizadores, bem como suas motivações, ao detalharmos como ocorre a derivação passo-a-passo de um composto na nossa proposta.

2.2.3. Itens de Vocabulário

Como vimos, os IVs representam regras que associam contextos sintáticos a um material fonológico. No entanto, algumas teorias de fase encontram problemas na distribuição acentual das estruturas sintáticas, pois ao sofrer *Spell-Out*, o acento é aplicado e fica congelado no segmento a que foi atribuído, impossibilitando que ele se mova para a borda direita da palavra fonológica quando mais núcleos funcionais forem preenchidos fonologicamente (cf. LOWENSTAMM, 2010).

¹⁵ Trecho original (ADGER, 2010, p. 193): “a. a feature is an ordered pair <Att, Val> where; b. Att is drawn from the set of attributes, {A, B, C, D, E, ...}; c. and Val is drawn from the set of values, {+, -, a, b, c, ...}.”

¹⁶ Um atributo pode ser, por exemplo, T, D, Caso, *n*, definitude, número, enquanto os valores correspondem à marcação desse atributo, como {D: val; definitude: val; Caso: nom; número: pl}.

Nossa proposta, em contrapartida, adiciona a alguns IVs a possibilidade de desencadear uma reaplicação da regra de acento, de acordo com a Hipótese da Ciclidade Pós-Sintática, delineada abaixo:

(12) *Hipótese da ciclidade pós-sintática*

Um IV marcado desencadeia a aplicação da regra de acento deslocando o acento para a borda direita da palavra fonológica.

Isso se deve pela presença de sufixos portadores de acento (e.g., *-ia*, em *historiografia*) ou sufixos que, uma vez concatenados a uma sequência fonológica, adicionam um número considerável de sílabas fazendo com que o acento seja deslocado para a borda direita da palavra fonológica (e.g., *-ico*, em *historiográfico*). Essa hipótese será apresentada, mais detalhadamente, no Capítulo 6, ao tratarmos da externalização dos compostos do PB.

2.3. Derivação sintática: a Condição do *Merge* na formação de palavras

Assumimos, ainda, um sistema operacional governado pela Condição do *Merge*, proposta em Wurmbrand (no prelo)¹⁷. Essa proposta traz uma preocupação relevante para a construção da estrutura sintática, pois tenta explicitar como as relações de seleção entre os núcleos presentes na estrutura sentencial estão codificadas na gramática, ou seja, o que permite que determinado núcleo se concatene a outro durante a derivação.

Munidos dessa questão, expandimos a proposta de Wurmbrand para a formação de palavras, admitindo, com a autora, uma restrição à operação *Merge*, ao exigir que sua realização ocorra apenas quando há a valoração de um traço, tal como definido em (13):

(13) *Condição do Merge (WURMBRAND, no prelo, tradução nossa)*

Concatene α e β se α puder valorar um traço de β .¹⁸

Essa condição exige um refinamento dos traços gramaticais presentes nos nós terminais sintáticos, algo desejado para um modelo de gramática que opera apenas com traços desde a construção de unidades abaixo do nível da palavra, quanto entre sintagmas para a formação de sentenças. Desse modo, mostramos, a partir dessa condição, como a derivação de

¹⁷ Agradeço ao Marcello Modesto pela indicação e disponibilização do texto.

¹⁸ Trecho original (WURMBRAND, no prelo): “Merge Condition: Merge α and β if α can value a feature of β ”.

um composto procede desde a concatenação entre uma raiz com um núcleo categorizador, até a sua utilização como uma única unidade sintática em uma sentença.

2.4. Conclusões Preliminares

Neste capítulo, elencamos os pressupostos teóricos que sustentam a argumentação e as propostas desenvolvidas nesta dissertação, bem como as alterações sugeridas ao modelo teórico da MD a serem implementadas na análise da formação dos compostos. Acreditamos que a visão de gramática delineada acima é a que explica, de modo mais econômico, o processo de composição das línguas naturais, ao fazer uso de um conjunto restrito de operações para a geração de qualquer objeto linguístico, seja ele uma palavra ou uma sentença. Tendo em mente tais pressupostos, podemos prosseguir ao estudo da composição nas diferentes facetas propostas nessa dissertação.

CAPÍTULO 3

ESTRUTURAS MORFOLÓGICAS DA COMPOSIÇÃO

Para dar início às discussões, apresentamos uma propriedade da composição específica às línguas, a saber, suas estruturas morfológicas, discutindo, em especial, a distribuição estrutural dos compostos do PB. Ao destacar essa característica, demonstramos como uma língua particular emoldura combinações categoriais gerais em estruturas morfológicas. Em paralelo, neste capítulo, percorremos a literatura que descreveu a composição no PB, a fim de fornecer um panorama geral desse processo de formação de palavras nessa língua, o qual embasará, empiricamente, os capítulos seguintes da dissertação.

3.1. Caracterização do capítulo

Tendo em vista que o objetivo principal desta dissertação é determinar quais informações são específicas a uma língua e quais informações são decorrentes de um sistema gerativo universal para a formação dos compostos, colocamos em discussão a seguinte questão: qual a distribuição mais ampla encontrada nos compostos de uma língua particular? Desenvolvemos essa questão a partir da descrição dos compostos de uma língua específica – o PB – discutindo, detalhadamente, as distribuições categorial, relacional e estrutural¹⁹ de seus compostos.

Ao explorar os dados de composição do PB, queremos tornar explícitas as informações que devem ser vistas como específicas para essa língua, com vistas a elencar, em seguida, as informações que são intrínsecas ao processo de composição de um ponto de vista universal. Com isso, demonstramos que uma das características específicas de cada língua é o modo como os compostos são estruturados morfológicamente, ou seja, quais combinações morfológicas, tais como [radical + radical] e [palavra + palavra], estão disponíveis para a realização fonológica dos compostos em uma língua particular, tendo em vista que assumimos

¹⁹ Referimo-nos aqui às estruturas morfológicas dos compostos, nomeadamente, a variedade de combinações feitas a partir de seus estatutos morfológicos principais, como raiz, radical e palavra.

que as palavras compostas advêm de uma combinação sintática translinguisticamente uniforme²⁰.

O capítulo, portanto, segue estruturado da seguinte maneira: na seção 3.2, percorremos a literatura sobre os compostos do PB, focalizando o modo como esse processo foi abordado e classificado pelos principais autores, comentando, em paralelo, os problemas que circundam tais classificações. Na seção 3.3, iniciamos a discussão sobre as estruturas morfológicas propriamente ditas, ao definir suas unidades constituintes, e apresentar as propriedades morfológicas, morfossintáticas e fonológicas advindas dessas combinações. Em seguida, estabelecemos um breve panorama das estruturas morfológicas translinguisticamente, com o intuito de evidenciar seu caráter particular às línguas no processo de composição.

Na seção 3.4, retomamos os levantamentos feitos ao longo das duas últimas seções, a fim de apresentar nossa proposta para o modo como a formação dos compostos deve ser interpretada. Essa proposta, por sua vez, além de unificar as sugestões de classificação dos compostos do PB, delineia o que há de mais geral e o que há de mais específico nesse processo de formação de palavras. Por fim, na seção 3.5, fazemos um resumo dos apontamentos discutidos neste capítulo.

3.2. As abordagens tradicionais de classificação dos compostos do PB

Nas gramáticas, manuais e teses que buscaram apresentar a composição no PB, encontramos uma variedade de sugestões para a classificação dos compostos, as quais podem ser distribuídas nos seguintes tipos: (i) classificação dos compostos por combinações de categorias lexicais, (ii) classificação dos compostos a partir das relações gramaticais entre seus membros, (iii) classificação dos compostos pela forma morfológica de seus membros, (iv) classificação dos compostos a partir da junção fonológica entre seus membros e (v) classificação dos compostos de acordo com o componente da gramática em que são formados.

Embora alguns autores privilegiem um tipo de classificação, muitos empregam, paralelamente, os demais tipos supramencionados. Por exemplo, Bechara (2006) estabelece uma divisão dos compostos a partir das combinações categoriais atestadas na língua, e subdivide uma dessas combinações – nomeadamente, a combinação N-N – em dois tipos, com base nas relações gramaticais que podem se estabelecer entre seus membros, a saber: subordinação ou coordenação.

²⁰ Em consonância com o Princípio da Uniformidade, delineado em Chomsky (2001, p.2); cf. Capítulo 1.

Por outro lado, Sandmann (1991) faz uma divisão inversa a de Bechara (2006), pois divide os compostos de acordo com as relações gramaticais de subordinação ou coordenação, e, então, aponta como as combinações categoriais estão distribuídas em cada uma dessas relações. Alves (2007), por sua vez, coloca em um mesmo nível as relações gramaticais e a forma morfológica, dividindo esse processo de formação de palavras em: compostos subordinados, compostos coordenados, compostos satíricos²¹, compostos entre bases não-autônomas, compostos sintagmáticos e compostos por siglas ou acronímia.

De todo modo, a classificação dos compostos por categorias lexicais é a mais recorrente. Said Ali (1971), Lima (1972), Sandmann (1991), Lee (1995, 1997), Kehdi (2002), Bechara (2006) e Cunha e Cintra (2007), atestam as seguintes possibilidades de combinação categorial nos compostos do PB:

(14) *Combinações categoriais dos compostos do PB*

- a. **N-N**: peixe-espada, trem-bala, ator-cantor.
- b. **N-A**: mesa-redonda, pé-frio, obra-prima.
- c. **A-N**: puro-sangue, belas-artes.
- d. **V-N**: lança-perfume, porta-voz, beija-flor.
- e. **N-V**: sangue-suga.
- f. **A-A**: surdo-mudo, luso-brasileiro, auriverde.
- g. **V-V**: vai-e-vém, corre-corre, pega-pega.
- h. **V-Adv**: pisa-mansinho, ganha-pouco.
- i. **Adv-N/V/A**: malcriação, benquerer, sempre-viva.
- j. **N-prep-N**: pé-de-moleque, baba-de-moça.
- k. **Pron-N**: seu-vizinho (dedo anular), nosso senhor, vossa excelência.
- l. **N-Pron**: pai-nosso, João-ninguém.
- m. **Num-N**: segunda-feira, zero-quilômetro (carro novo), trigêmeos.

Embora possamos verificar uma variedade razoável de combinações categoriais, nem todas são produtivas na língua. Alves (2007), com base em informações de neologismos produzidos na imprensa escrita, aponta, em seu manual, uma maior quantidade de dados formados a partir das combinações N-N, V-N, A-N e Num-N. A autora comenta, ainda, que

²¹ Para Alves (2007, p. 46-48), compostos satíricos são aqueles cunhados a fim de chamar a atenção do receptor através da quantidade de elementos concatenados (e.g., candidato-deputado-cantor), ou através de analogias (e.g., automóvel → papamóvel).

compostos de combinação categorial N-prep-N, embora sejam encontrados, são pouco frequentes no PB (2007, p. 43).

A segunda classificação, que divide os compostos a partir das relações gramaticais entre seus membros, é a que suscita mais debate. Tal como já mencionamos, as relações gramaticais geralmente aludidas são: subordinação e coordenação. Essas relações, no entanto, recebem uma definição e uma interpretação variada de autor para autor. Iniciemos por Bechara (2006); para o autor, apenas os compostos de combinação categorial N-N se subdividem em coordenados e subordinados, de acordo com as seguintes definições:

(15) *Relações gramaticais em Bechara (2006, p. 355-356)*

- a. *Coordenação* – quando há sequência de coordenação de elemento: 1 – o determinante precede: *mãe-pátria, papel-moeda*; 2 – o determinante vem depois: *peixe-espada, carro-dormitório, couve-flor*.
- b. *Subordinação* – quando há subordinação de um elemento, isto é, de um determinante a outro determinado: *arco-íris, estrada de ferro, pão de ló*.

O recorte que o autor estabelece tem como base a presença ou não de uma preposição implícita, ou seja, de um núcleo subordinador. Percebamos que, em sua visão, a coordenação apenas exige uma sequência de elementos contíguos, em que “um substantivo funciona como aposto do outro, em geral, o segundo: *peixe-espada* (peixe que se parece com uma espada)” (2006, p. 355), mas não carrega, necessariamente, a concepção geral de coordenação, ou seja, a de que os membros de um composto são unidos por uma relação conjuntiva (e.g., $x \text{ E } y$) ou disjuntiva (e.g., $x \text{ OU } y$).

A relação de subordinação, por sua vez, diz que os membros de um composto são unidos por uma relação de complementação, a qual pode ser intermediada por uma preposição. Entretanto, o autor aponta que a omissão da preposição ‘de’ é muito natural no português, como acontece em *arco-íris* (por *arco da Íris*, *Íris* é um nome mitológico), *porco-espinho* (= porco de espinho), *beira-mar* (= beira do mar), dentre outros. Logo, se a combinação N-N permite uma paráfrase a partir de uma preposição, o composto poderá ser classificado como subordinado.

O que podemos concluir da distinção feita por Bechara (2006) é que tanto a coordenação quanto a subordinação podem estabelecer uma mesma relação entre os membros de um composto – a de modificação –, algo que é evidenciado pela rotulação dos membros como determinantes e determinados nos dois tipos de compostos. Ora, se a relação de

coordenação para Bechara (2006) classificasse, de fato, os elementos unidos em uma relação conjuntiva, não seria esperado que um dos membros determinasse o outro, uma vez que, na coordenação, ambos estariam em uma configuração simétrica.

Azeredo (2008), por outro lado, faz uma distinção mais refinada dessas relações. O autor sugere que os compostos podem ser classificados em dois tipos, de acordo com as definições em (16):

(16) *Classificação dos compostos a partir das relações gramaticais em Azeredo (2008, p. 445)*

- a. *Tipo 1* – o que a palavra composta denota é entendido como a simples união – ou *coordenação* – dos significados particulares dos lexemas que a integram: *infanto-juvenil, sócio-proprietário, diretor-gerente, rádio-gravador, editor-chefe*.
- b. *Tipo 2* – o significado total do composto é distribuído de forma desigual entre seus componentes, de sorte que um deles figura como base (B) do significado e o outro exprime uma especificação qualquer (E). Os lexemas componentes podem ocorrer na ordem B + E (base + especificação; e.g., *matéria-prima, salário-mínimo, peça-chave*), mais comum e única no uso popular, ou na ordem E + B (especificação + base; e.g., *hidromassagem, lipoaspiração, motosserra*), própria dos usos técnicos e científicos da língua.

O que vemos nas definições de Azeredo (*op. cit.*), ao contrário de Bechara (*op. cit.*), é que as relações internas se distinguem claramente, pois apenas os compostos do Tipo 2 carregam uma relação assimétrica de modificação, em que um membro especifica o significado do outro.

Sandmann (1991) faz uma distinção semelhante e mais detalhada que a de Azeredo (*op. cit.*), pois nota que essas relações não se restringem aos compostos de combinação N-N, mas estão presentes nas demais combinações categoriais atestadas na língua. Para o autor, os compostos podem ser divididos da seguinte forma:

(17) *Classificação dos compostos a partir das relações gramaticais em Sandmann (1991, p. 62-69)*

- a. *Compostos copulativos ou coordenados*:
 - i) *Substantivo composto copulativo*: rotula algum ser que rotula em si, em igualdade de condições, rótulos menores ou mais específicos. Assim,

alguém pode ser simultaneamente cantor e compositor, sendo então um “cantor-compositor” (1991, p. 62).

- ii) *Adjetivo composto copulativo*: no adjetivo composto copulativo encontramos mais de uma predicação, em geral duas, todas reportando-se diretamente e em igualdade de condições ao núcleo, o substantivo ou expressão em função substantiva; e.g., “acordo brasileiro-paraguaio” (1991, p. 64).

b. *Compostos determinativos ou subordinativos*:

- i) *Compostos N-N*²² *vernáculos e não-vernáculos*: nos compostos N-N vernáculos, o segundo substantivo qualifica ou determina o primeiro, estabelecendo uma relação de subordinação do segundo substantivo pelo primeiro, ou de determinação do primeiro pelo segundo (e.g., *estado-mecenas*). Nos compostos N-N não-vernáculos, a direção se inverte (e.g., *chutometria*).
- ii) *Compostos V-N*: o determinante é o segundo elemento, o substantivo, em função de objeto direto (e.g., *pega-rapaz*).
- iii) *Compostos N-A e A-N*: nesses compostos, o substantivo é sempre o núcleo e o adjetivo, sempre o adjunto (e.g., *peso-pesado*, *pronto-socorro*).
- iv) *Compostos N-Num e Num-N*: o substantivo é o núcleo e o numeral o adjunto (e.g., *onze-letras* ‘alcoviteiro’, *camisa-dez* ‘centroavante’).
- v) *Compostos N-de-N*: o primeiro substantivo é o núcleo e a preposição com o segundo substantivo é o adjunto (e.g., *pé-de-meia* ‘economia’).

Em Sandmann (1991), notamos uma distribuição das combinações categoriais em que as relações de coordenação (ou copulativas) e de subordinação (ou determinativas) podem ser estabelecidas. Porém, se nos detivermos nos compostos subordinados elencados pelo autor, o que parece ocorrer é a presença de mais de uma relação entre os membros desses compostos. Por exemplo, em que sentido um nome na posição de complemento, tal como nos compostos V-N, serve como determinante do verbo em questão? Essa relação parece ser bem distinta da relação entre os nomes de um composto N-N, em que um nome não-núcleo modifica, ou determina, o nome núcleo.

²² O autor emprega S (de substantivo) em vez de N. Substituímos a abreviatura para manter uma uniformidade ao longo do texto.

Assim, tendo em vista essas sugestões de classificação pautadas nas relações internas entre os membros de um composto, o que podemos notar é que elas não são sistemáticas, e, em alguns casos, ou negligenciam a presença da mesma relação em outras combinações categoriais (e.g., a divisão de Bechara), ou agrupam em uma mesma relação, demais relações que podem estar presentes (e.g., a divisão de Sandmann)²³.

O que iremos sugerir em nossa abordagem, é que a subordinação, geralmente entendida como uma relação de modificação nas descrições do PB, deve ser cindida em (a) uma relação subordinativa propriamente dita, em que há uma exigência predicado-complemento, como nos compostos V-N, e em (b) uma relação atributiva, ou seja, uma relação de modificação, presente nos compostos N-N e N-A, seguindo a proposta de classificação dos compostos de Bisetto e Scalise (2005)²⁴.

A terceira classificação dos compostos, feita a partir da forma morfológica de seus membros – ou seja, se os compostos são formados por formas livres (e.g., *navio escola*) ou por formas presas (e.g., *hidrologia*) – é apontada de modo paralelo na quase totalidade dos autores verificados (cf. SANDMANN, 1991; MORENO, 1997, 2002; BECHARA, 2006; CUNHA; CINTRA, 2007; ALVES, 2007; BASILIO, 2007; AZEREDO, 2008; BOPP DA SILVA, 2010). Em alguns trabalhos, por outro lado, compostos formados por formas presas não são sequer mencionados diretamente (e.g., LEE, 1995, 1997; KEHDI, 2002), fato que pode estar relacionado ao caráter periférico ou “menos comum” atrelado a essas formações no PB.

Entre os compostos formados por formas presas, os autores fazem referência àqueles constituídos por, no mínimo, um radical preso de origem grega ou latina (e.g., *antropologia*, *hidromassagem*), comumente chamados de neoclássicos ou eruditos. Essa distinção é feita, muitas vezes, com o intuito de apontar uma diferença na posição do núcleo, pois, como mencionado em (16) e (17), os compostos com formas presas têm núcleo semântico à direita (e.g., *tomaticultura*), contrariamente aos compostos com formas livres, com núcleo semântico à esquerda (e.g., *peixe-espada*). Bopp da Silva (2010), em particular, é a única que busca fornecer uma explicação mais detalhada para essas formações, levantando, em seu trabalho,

²³ Poderíamos citar ainda a divisão de Alves (2007) que, como mencionamos, coloca em um mesmo nível as relações gramaticais, de subordinação e coordenação, e a variedade de formas que constituem um composto (e.g., compostos entre bases não-autônomas e compostos sintagmáticos), sugerindo que tais relações não estão presentes nessas formas. Como veremos na seção 3.4, o estabelecimento dessas relações gramaticais independe da estrutura morfológica do composto (i.e., se são formados por radicais ou por palavras).

²⁴ Cf. seção 3.4.

um conjunto de informações sobre os aspectos fonológicos e morfológicos desses compostos²⁵.

A quarta classificação para os compostos do PB parte de critérios fonológicos, separando-os, principalmente, em compostos por justaposição ou compostos por aglutinação (cf. LAROCA, 1994, KEHDI, 2002). A diferença entre esses compostos reside nas seguintes especificações:

(18) *Classificação dos compostos por critérios fonológicos (cf. KEHDI, 2002, p. 36-37)*

- a. *Compostos por justaposição* – quando os termos associados conservam a sua individualidade: *passatempo, sempre-viva*.
- b. *Compostos por aglutinação* – quando os vocábulos ligados se fundem num todo fonético, com um único acento, e o primeiro perde alguns elementos fonéticos (acento tônico, vogais ou consoantes): *boquiaberto, pernalta*.

Villalva (2000), em sua crítica às visões restritas sobre esse processo de formação de palavras, comenta que a divisão dos compostos entre justapostos e aglutinados²⁶ é uma consequência do entendimento da composição como um processo de lexicalização, uma abordagem frequentemente tomada pelas gramáticas, a qual, por sua vez, “não dá conta de dois diferentes modos de realização do processo, mas de dois estágios, ou graus, ou tipos de lexicalização” (2000, p. 347).

A autora argumenta que os compostos por justaposição sofrem apenas uma lexicalização semântica, ao passo que, nos compostos por aglutinação, a lexicalização não é apenas semântica, mas também formal, ou seja, parte dos segmentos é apagada. Em contrapartida, Villalva (*op. cit.*) sugere que compostos por aglutinação e compostos por justaposição devem ser substituídos pelos rótulos de compostos não-lexicalizados e compostos lexicalizados, respectivamente. Nesse sentido, uma classificação dos compostos baseada, exclusivamente, em uma proposta como essa, contempla apenas os aspectos

²⁵ Vale mencionar que citamos apenas Bopp da Silva (2010), nesse contexto, pois nos restringimos a autores que analisaram o processo de composição em sua totalidade. Há trabalhos, por outro lado, que se ativeram exclusivamente aos compostos neoclássicos ou formados por formas presas (cf. GONÇALVES, 2011a, 2011b; SCHER; NÓBREGA, 2014; NÓBREGA, no prelo).

²⁶ Fazemos uma distinção entre compostos por aglutinação e *blends*, ou cruzamentos lexicais (e.g., *apartamento, boilarina, boacumba*), pois enquanto a perda de segmentos nos compostos por aglutinação decorre de uma regra fonológica entre os radicais (e.g., a degeminação em *água + ardente > aguardente*, cf. BOPP DA SILVA, 2010, p. 82), nos *blends*, a extração de segmentos nem sempre é motivada por uma regra fonológica, além do fato de que a extração nesses dados visa efeitos estilísticos, diferentemente dos compostos por aglutinação (cf. GONÇALVES, 2003a, 2003b; MINUSSI; NÓBREGA, 2014).

diacrônicos dessas formações, e deixa de lado uma série de informações sincrônicas acerca do processo, não sendo capaz, portanto, de demonstrar como ele opera na língua em questão.

A quinta classificação dos compostos – que os divide de acordo com o componente da gramática em que são formados – pressupõe uma visão particular de como a gramática se estrutura. Nessa proposta, apresentada nos trabalhos de Lee (1995, 1997), compostos são divididos em lexicais e pós-lexicais, com o intuito de diferenciar: (a) um conjunto de dados que não admite acesso à sua estrutura interna, sendo vistos como objetos morfológicos atômicos, de (b) um conjunto de dados que permite tal acesso, e que deve ser interpretado como construções sintáticas reanalisadas em uma palavra.

Lee (*op. cit.*), assumindo uma visão lexicalista de gramática, que toma por base o modelo teórico da Fonologia Lexical Prosódica (cf. INKELAS, 1989, 1993), sugere que a arquitetura da gramática para o PB contém a presença de dois níveis para a aplicação de regras fonológicas e morfológicas, a saber: um nível derivacional (α) e um nível flexional (β), enquanto que a composição se distribui pela gramática em dois espaços separados, como pode ser visto na Figura 4, abaixo:

Figura 4 - Modelo da Fonologia Lexical no PB para Lee (1995)

A divisão entre Composição I e Composição II, apresentada na Figura 4, remete à separação mencionada logo acima. Compostos do tipo I ou *compostos lexicais*, como rotula o autor, são considerados os verdadeiros compostos do PB e sua formação se dá no léxico, o que os torna sintaticamente opacos; exemplos desses dados estão listados em (19). Além deles, há os compostos do tipo II, chamados de *compostos pós-lexicais* (ou pseudo-compostos), os quais são gerados no componente sintático, e são sintaticamente ativos; exemplos estão elencados em (20):

(19) *Compostos lexicais* (doravante, CL)

- N-N (e.g., autopeça, cineclube, ferrovia, tomaticultura, rádio-táxi, espaçonave)
- A-A (e.g., ítalo-brasileiro, judeu-americana, socioeconômico, sócio-cultural)
- V-N (e.g., guarda-chuva, toca-discos, porta-aviões)

(20) *Compostos pós-lexicais* (doravante, CPL)

- N-N (e.g., sofá-cama, homem-rã, trem-bala, garota propaganda)
- A-A (e.g., surdo-mudo)
- N-A (e.g., boia-fria, carro-forte, mesa-redonda, pão-duro, dedo-duro)
- A-N (e.g., curto-circuito, primeiro ministro)
- N-prep-N (e.g., fim de semana, pé-de-moleque)

A distinção entre esses compostos é feita, basicamente, a partir de cinco testes: (a) formação do plural, (b) formação do diminutivo, (c) derivação, (d) concordância e (e) ordem dos elementos. A aplicação desses testes segue, grosso modo, o seguinte critério: se houver a possibilidade de marcação flexional ou derivacional internamente ao composto, isso significa que ele foi formado na sintaxe, pois sua estrutura interna é sintaticamente transparente. Por outro lado, se a marcação flexional ou derivacional ocorrer apenas em sua borda direita, isso significa que o composto foi formado no léxico, e se comporta como um objeto atômico na sintaxe.

Para os testes (a) e (b), o autor mostra que, enquanto CLs recebem morfemas de plural e diminutivo apenas em sua borda direita, como em (21a) e (22a), nos CPLs a marcação de plural e diminutivo ocorre no primeiro membro do composto, como em (21b) e (22b); e, no caso do plural, a marcação pode ocorrer mais de uma vez:

(21) a. CL: rádio-táxis, ferrovias, guarda-chuvas.

b. CPL: fins-de-semana, trens-bala, surdos-mudos, presidentes-ministros.

(22) a. CL: guarda-roupinha/*guardinha-roupa

b. CPL: horinha-extra/ *hora-extrinha

Com o teste (c), o autor mostra que CLs podem formar novas palavras através de afixação, enquanto CPLs só podem se envolver com derivações específicas, tal como em (23a), uma vez que apenas prefixos podem ser acrescentados a eles, como em (23b):

- (23) a. CL: [[fotograf]_{Nar}]_V, [[puxa-saco]_A ismo_N]
 b. CPL: ex-homem-rã, super-primeiro-ministro

No teste (d), CLs admitem marcação de concordância apenas em sua borda direita, (24a), enquanto nos CPLs a concordância de gênero e número se aplica em todos os constituintes, (24b).

- (24) a. CL: judeu-americano/ *judia-americana/ judeu-americana(s)
 b. CPL: surdos-mudos/ *surdo-mudos/ *surdos-mudo/ *surdas-mudos

E, por fim, no teste (e), CLs apresentam a ordem determinante + determinado, enquanto compostos CPLs apresentam tanto a ordem determinante + determinado, quanto a ordem determinado + determinante:

- (25) a. CL: espaçonave, autopeça
 b. CPL: curto-circuito, primeiro-ministro, trem-bala, funcionário-fantasma

De imediato, o que podemos verificar na classificação proposta por Lee (1995, 1997) é que a estrutura morfológica do composto, ou seja, se o composto é formado por radicais ou por palavras, não é uma informação levada em consideração, tampouco as relações gramaticais que se estabelecem entre seus membros. Há que se notar, ainda, que combinações categoriais idênticas são encontradas nos dois tipos de compostos sugeridos, tal como é o caso das combinações N-N e A-A.

Moreno (1997, 2002) critica amplamente a divisão feita por Lee (*op. cit.*), argumentando que seus testes não evidenciam a presença de CLs, principalmente, os testes relativos à marcação de plural e de concordância, chegando à conclusão de que a formação dos compostos é, essencialmente, sintática. O autor mostra que a distribuição de marcas flexionais sofre a interferência de fatores sintáticos e semânticos, e não apenas das categorias lexicais dos membros de um composto, como sugere Lee, o que torna sua classificação em CLs e CPLs pouco sustentável.

Um exemplo para esse argumento de Moreno pode ser dado com compostos que carregam uma interpretação ambígua, tal como aqueles de combinação categorial N-N. Para Lee, esses compostos são classificados tanto como CLs quanto como CPLs, e em seus dados,

há apenas uma relação de modificação entre seus membros, em que um nome núcleo é modificado por um nome não-núcleo (e.g., CL: *rádio-táxi*; CPL: *trem-bala*). Contrário a essa visão, Moreno argumenta que além de existir mais de uma relação gramatical entre esses nomes, a saber, a relação de coordenação²⁷, o que está em jogo na marcação do plural é a identificação de qual membro é o núcleo do composto.

Tomando como pressuposto que a tendência natural do falante é interpretar os compostos segundo as relações mais usuais no sintagma nominal, e que a estrutura N-A é uma relação sistemática de modificação no PB, Moreno irá argumentar que, em compostos N-N, o nome não-núcleo é interpretado como um adjetivo, e que quando esse vem à direita do núcleo, há a possibilidade de concordância, como em (26), visto que essa concordância também é desencadeada nos sintagmas nominais em que o adjetivo pospõe o nome (e.g., *jovens bonitos*, *mulheres ricas*).

(26) a. escola modelo	>	escola <u>s</u> modelo	ou escola <u>s</u> model <u>os</u>
b. palavra chave	>	palavra <u>s</u> chave	ou palavra <u>s</u> chave <u>s</u>
c. operário padrão	>	operário <u>s</u> padrão	ou operário <u>s</u> padrõe <u>s</u>
d. café concerto	>	café <u>s</u> concerto	ou café <u>s</u> concert <u>os</u>

(MORENO, 2002, p. 110)

Por outro lado, se o nome não-núcleo vier à esquerda, a concordância com o nome núcleo não ocorrerá, como pode ser visto em (27), uma vez que, quando o adjetivo precede o nome em sintagmas nominais, ele pode deixar de flexionar (e.g., *livre pensadores*, *livre arbitrios*, cf. MORENO, 2002, p. 111).

- (27) auto peças cine clubes ferro vias

(MORENO, 2002, p. 111)

Dessa forma, o que está em jogo na formação desses compostos não é o componente da gramática em que são gerados, mas sim o modo como seus membros são organizados através de relações gramaticais no componente sintático. Outra evidência para que a distribuição flexional leve em consideração apenas aspectos relacionais e estruturais – mas não o componente da gramática – vem da relação de coordenação presente nos compostos N-

²⁷ Algo notado também notado por outros autores (e.g., SANDMANN, 1991; BECHARA, 2006; ALVES, 2007).

N, em que ambos os nomes são vistos simetricamente, e não em uma relação assimétrica de modificação.

Por exemplo, tendo em vista que a marcação do plural ocorre, preferencialmente, no núcleo semântico dos compostos (e.g., *fins-de-semana*, *trens-bala*), desencadeando concordância interna em determinados contextos, é de se esperar que os compostos coordenados exibam também uma variabilidade na marcação (e.g., *democrata-cristãos*; *democratas-cristãos*), pois não fica claro nesses dados qual dos nomes figura como núcleo do composto. Esse é um dos motivos pelos quais compostos N-N e A-A são classificados tanto como CLs e como CPLs, pois a relação de coordenação carrega um problema inerente, oriundo de seu caráter simétrico, que é a indeterminação de qual membro figura como núcleo da relação²⁸.

O que podemos depreender, portanto, é que os problemas de marcação flexional e de concordância apontados por Lee não decorrem do componente da gramática em que os compostos são gerados, mas advêm das relações gramaticais presentes entre seus membros, interpretadas de acordo com a presença ou ausência de núcleo (se considerarmos que os compostos N-N coordenados não têm um núcleo). Por esse motivo, deixamos de lado a divisão proposta por Lee, pois nosso intuito, nesse momento, é fornecer uma classificação dos compostos ampla e geral, capaz de evidenciar as particularidades desse processo em uma língua específica com relação às demais línguas naturais.

Para os nossos fins, julgamos ser importantes as classificações (i) das categorias gramaticais, (ii) das relações gramaticais e (iii) das formas morfológicas. As classificações (iv) da junção fonológica e (v) do componente da gramática são restritivas e não se adequam aos propósitos deste capítulo, pois não visamos tratar dos aspectos diacrônicos da composição e não concordamos com a divisão dos compostos sugerida pela última classificação, visto que além de conter problemas, essa proposta faz uso de uma teoria da gramática incompatível com a que assumimos, além de se circunscrever a propriedades específicas de uma língua, não podendo ser ampliada de modo satisfatório translinguisticamente.

Na seção 3.4, iremos retomar as classificações (i), (ii) e (iii), a fim de apontar dentre elas quais decorrem de propriedades universais e intrínsecas a esse processo de formação de

²⁸ Sandmann (1991, p. 63) sugere dois testes para identificar se o composto é coordenado ou subordinado (i.e., contém uma relação de modificação): a paráfrase e o ordenamento de seus membros. Segundo o autor, se um composto como *meia-calça* fosse subordinado, a ordem seria determinante-determinado, *meia* seria o núcleo e *calça* o adjunto. Uma paráfrase, nesse sentido, poderia ser “meia em formato de calça”, o que, para Sandmann (*op. cit.*) não corresponderia aos fatos, pois *meia-calça* é ao mesmo tempo *meia* e *calça*, e uma inversão na ordem dos nomes levaria ao mesmo resultado, *calça-meia*. O mesmo não pode ser dito de um composto subordinado como *trem-bala*, em que a ordem não pode ser alterada, **bala-trem*.

palavras e quais devem ser vistas como específicas à língua, com base em um conjunto de dados translinguísticos acerca do processo. Antes disso, iremos nos ater às estruturas morfológicas dos compostos do PB, com o intuito de explicitar as formas que os compostos dessa língua podem tomar.

3.3. As estruturas morfológicas da composição

Nesta seção, detalhamos as estruturas morfológicas dos compostos do PB. Para tanto, na subseção 3.3.1, fornecemos instruções sobre quais são os constituintes que compõem as estruturas morfológicas dos compostos; na subseção 3.3.2, exploramos mais detalhadamente as estruturas morfológicas dos compostos do PB, descrevendo suas propriedades principais, e, por fim, na subseção 3.3.3, contrastamos essas estruturas com outras formas de estruturação presente nas línguas naturais, a fim de evidenciar o caráter particular das estruturas morfológicas translinguisticamente.

3.3.1. Composto as estruturas morfológicas: raízes, radicais e palavras

Como visto no Capítulo 2, o que temos como primitivo da gramática para a MD é uma unidade intitulada “raiz”. Em princípio, não podemos dizer se essa unidade será fonologicamente bem estruturada, pois a sintaxe apenas garante sua combinação no espaço computacional, ao passo que questões relativas à sua boa formação fonológica são avaliadas pós-sintaticamente, no caminho para PF, e não em um componente lexical pré-sintático, tal como sugerem as abordagens lexicalistas.

É notável, por outro lado, uma distribuição diversificada das possibilidades de realização fonológica das raízes sintáticas, as quais são geralmente descritas pela morfologia tradicional em três grandes grupos: raízes, radicais e palavras. Essas noções circunscrevem-se, a nosso ver, à fonologia, e, a depender da língua, à complexidade morfológica associada a uma raiz sintática. Nesse sentido, a noção de raiz na MD contém, grosso modo, apenas conteúdo semântico, o que a aproxima da noção de *lexema*²⁹ das abordagens lexicalistas.

Embora as três noções supramencionadas sejam encontradas sob um mesmo rótulo na literatura, elas não apresentam uma uniformidade no modo como são definidas, variando, em

²⁹ De acordo com Haspelmath e Sims (2010, p. 333), um *lexema* é um conceito abstrato representando um significado nuclear partilhado por um conjunto de formas de palavras intimamente relacionadas que formam um paradigma (e.g., *viver*, *vive*, *viveu*).

certa medida, de autor para autor³⁰. Em nosso trabalho, uma distinção clara dessas noções se faz necessária, pois assumimos que uma propriedade específica das línguas encontra-se no modo como seus compostos são emoldurados em estruturas morfológicas, embora os compostos provenham de uma combinação sintática translinguisticamente uniforme.

A primeira noção – raiz – é definida por Câmara Jr. (2004 [1977]), como a “parte lexical de um vocábulo”, ou seja, a unidade mínima portadora de conteúdo lexical, que não pode sofrer divisões adicionais, sendo constituída, portanto, de um único morfema. Por exemplo, a raiz da palavra *cristal* é *crist-*, e qualquer desmembramento realizado nessa raiz ocasionará a perda de sua identificação. Em algumas línguas, palavras morfológicamente simples coincidem com suas raízes, uma vez que essas últimas não podem ser desmembradas em partes menores, tal como é o caso do inglês, em que *cat* ‘gato’ é uma palavra simples e ao mesmo tempo a própria raiz.

A segunda noção – radical – é comumente interpretada como a base capaz de receber informação flexional e derivacional, ou seja, ponto a partir do qual demais morfemas podem ser concatenados. O radical, por sua vez, pode coincidir com a raiz, tal como é o caso da palavra *cristal*, em que seu radical, *crist-*, coincide com a forma de sua raiz; por outro lado, o radical de *cristalizar* é *cristaliz-*, embora sua raiz continue sendo *crist-*. A raiz sintática, nesse sentido, deve corresponder à noção de raiz presente na morfologia tradicional, já que o radical é uma unidade passível de ser expandida, pois pode conter informações que não se circunscrevem ao núcleo semântico da palavra, e que são dadas por núcleos funcionais concatenados acima da raiz sintática na MD.

No PB, tal como nas demais línguas românicas, a noção de radical carrega um papel importante nos processos de formação de palavras, pois existe, nessas línguas, a dependência de uma vogal temática ou de morfemas derivacionais para transformar os radicais em unidades fonologicamente autônomas, por exemplo, o nome *gat-*, o verbo *beb-* e o adjetivo *chat-*. Consequentemente, o radical é uma unidade presa, ou seja, não pode ocorrer livremente em uma sentença. Essa característica deflagra um aspecto particular da morfologia dessas línguas, a saber, a impossibilidade da grande maioria de suas raízes constituírem palavras morfológicamente simples³¹.

Os radicais podem ser subdivididos ainda em outras quatro categorias, agrupadas em dois conjuntos: (i) radicais temáticos vs. radicais atemáticos e (ii) radicais simples vs. radicais

³⁰ Por exemplo, se compararmos as definições de Rocha (1998) com as definições de Haspelmath e Sims (2010).

³¹ Exceções são as raízes que não dependem de uma vogal temática para seu licenciamento fonológico, tais como *mar*, *dor*, *papel*, *jardim*, encontradas em menor número.

complexos. Os primeiros pretendem diferenciar os radicais que podem ser licenciados sintaticamente a partir da concatenação a uma vogal temática, ou seja, os radicais ditos vernaculares (e.g., *gat-*, *beb-*, *chat-*), separando-os dos radicais que não permitem a associação a uma vogal temática, sendo licenciados sintaticamente em contextos específicos; a esses radicais atemáticos são associados os chamados radicais neoclássicos (cf. SCHER; NÓBREGA, 2014; e.g., *hidr-*, *agr-*, *antrop-*).

As duas últimas categorias, radicais simples vs. radicais complexos, diferenciam os radicais de acordo com a quantidade de morfemas derivacionais concatenados a eles: logo, se um radical figura como uma única unidade, ou seja, sem morfemas derivacionais adicionados, tal como *crist-* em *cristal*, temos um radical simples, ao passo que nas formas em que esses radicais se compõem de morfemas derivacionais, como em *cristaliz-* de *cristalizar*, em que há um morfema derivacional verbal *-iz* concatenado, temos um radical complexo.

A terceira noção – palavra – define as unidades que podem ocorrer livremente em uma sentença, ou seja, unidades fonologicamente autônomas. A palavra, tal como o radical, pode ser simples ou complexa, e abriga as demais noções supramencionadas. Palavras simples são as raízes do inglês licenciadas sem a intervenção de uma vogal temática (e.g., *cat* ‘gato’), e palavras complexas são os radicais temáticos quando concatenados a uma vogal temática (e.g., *gat-o*, *beb-e-r*, *chat-o*). Além disso, são também palavras os chamados radicais atemáticos quando concatenados a um afixo derivacional (e.g., *hidr-ante*, *agr-ário*, *antrop-ia*).

A partir de agora, iremos nos referir aos radicais como as unidades fonológicas básicas das palavras do PB, as quais não ocorrem por si só, e dependem de uma vogal temática para funcionar como formas independentes, em consonância com a visão de Harris (1991, 1999). As vogais temáticas, sejam elas nominais ou verbais, serão vistas como expoentes fonológicos de marcadores de classe, ou seja, marcas caracterizadoras de grupos nominais ou verbais que declinam ou flexionam da mesma maneira, e que tornam um radical em uma palavra³². Nesse sentido, será a ausência ou a presença da informação de classe que ocasionará a realização de uma raiz sintática como radical ou palavra no componente morfológico da gramática.

No que diz respeito à composição no PB, as noções de radical e palavra são as mais importantes. Como veremos na próxima seção, o PB faz uso da combinação dessas duas possibilidades de realização de uma raiz sintática para a formação de seus compostos, e uma

³² Assumimos, como Alcântara (2003, 2010), que o PB contém quatro classes nominais, nomeadamente: Classe I, cujo morfema é /o/, como em *carr-o*, Classe II /a/ (e.g., *cas-a*), Classe III /e/ (e.g., *pent-e*) e Classe IV /Ø/ (e.g., *mar-Ø*); e três classes verbais: Classe verbal I /a/ (e.g., *am-a-r*, 1ª conjugação), Classe verbal II /e/ (e.g., *com-e-r*, 2ª conjugação) e Classe verbal III /i/ (e.g., *part-i-r*, 3ª conjugação). Retornaremos a essas questões quando demonstrarmos como a informação de classe está codificada na gramática, na seção 5.3.2 do Capítulo 5.

das vantagens de torná-las explícitas é que conseguimos notar um conjunto de padrões morfológicos, morfossintáticos e fonológicos desencadeados pelo modo como radicais e palavras se combinam para a formação de uma palavra composta.

3.3.2. As estruturas morfológicas dos compostos do PB

Nesta subseção, descrevemos, brevemente, as propriedades morfológicas, morfossintáticas e fonológicas contidas nas três formas de realização estrutural dos compostos do PB: [radical + radical], [radical + palavra] e [palavra + palavra]. Ao descrevê-las, argumentamos contra o caráter periférico associado às duas primeiras combinações no processo de composição do PB, alegando que as estruturas [radical + radical] e [radical + palavra] carregam o mesmo valor que a estrutura [palavra + palavra] nessa língua.

3.3.2.1. Estruturas [radical + radical] e [radical + palavra]

Quando apresentamos a sugestão de classificação dos compostos pela forma morfológica de seus membros, na seção 3.2, mencionamos, rapidamente, como a literatura tratou a constituição dos compostos do PB, ao fazer uma distinção entre compostos formados por formas livres e compostos formados por formas presas. Nesta seção, iremos nos concentrar nos compostos do segundo tipo, associando as formas presas à noção de radical, com o intuito de descrever como esses elementos se distribuem nos compostos dessa língua e quais as propriedades morfológicas, morfossintáticas e fonológicas que emergem das estruturas que eles compõem.

De modo geral, os morfólogos associam as estruturas formadas por radicais aos chamados compostos eruditos ou compostos neoclássicos, a saber: compostos formados por pelo menos um radical de origem grega ou latina, tais como *antropologia*, *piscicultura*, *hidroginástica* e *neurociência*. Além de evidenciar a possibilidade restrita de licenciamento sintático presente nesses radicais, os autores também procuraram ressaltar a diferença na posição do núcleo em seus compostos, pois quando comparada aos compostos ditos vernaculares, ou seja, de combinação [palavra + palavra], os primeiros têm núcleo semântico à direita (e.g., *piscicultura*, *hidroginástica*), enquanto os últimos têm núcleo semântico à esquerda (e.g., *peixe-espada*, *casca grossa*).

O que faremos, a partir de agora, é expandir a descrição desses compostos para além da discussão sobre qual a posição do núcleo em suas estruturas. Em consonância com a

proposta de Villalva (2000), que descreveu as estruturas morfológicas do português europeu, assumimos que as estruturas [radical + radical] e [radical + palavra] integram o processo de formação de compostos do PB. Partilhamos com a autora, também, da visão de que o que distingue os compostos formados por radicais dos compostos formados por palavras não são fatores extralinguísticos (como fatores históricos ou etimológicos), mas sim suas propriedades formais (cf. VILLALVA, 2000, p. 352)³³.

É válido mencionar ainda que a composição neoclássica é encontrada em uma variedade de línguas indo-europeias e germânicas (e.g., islandês, espanhol, catalão, francês, romeno, italiano, dentre outras; cf. NESPOR, 1999, p. 130), o que a faz, muitas vezes, ser considerada como um tipo de composição uniforme translinguisticamente. Porém, como ressalta Petropoulou (2009, p. 40), há que se investigar se esse tipo de composição foi realmente incorporado a esses sistemas linguísticos – principalmente, no que concerne à sua estrutura –, e quais desses radicais neoclássicos ainda operam como tais, ou seja, quais não foram gramaticalizados em formas afixais³⁴.

A fim de endossar a incorporação desse tipo de composição no PB, alegando que ele se caracteriza como possibilidades estruturais de seus compostos, exploramos, a partir de agora, suas características gerais, elencadas por Amiot e Dal (2007, p. 323), e as comentamos com base nos dados dessa língua:

(28) *Características gerais dos compostos neoclássicos (adaptadas de AMIOT; DAL, 2007, p. 323)*

a. *Ausência de realização sintática nas línguas-tomadoras*

São formas livres nas línguas de origem, grego antigo e latim, e funcionam, nas línguas tomadoras, como constituintes presos de palavras, não recebendo morfemas gramaticais (e.g., (a₁) *Eu tomei um copo de *hidro*, (a₂) *Eu vi dois *antropos* com chapéu hoje);

b. *Tipo de vocabulário que formam*

São unidades formalmente aprendidas, formadoras de termos técnicos ou científicos e;

³³ É importante mencionar que, ao contrário de Villalva (2000), não fazemos uma distinção entre o que a autora chama de compostos morfológicos (i.e., compostos formados por radicais) e compostos sintáticos (i.e., compostos formados por palavras), pois assumimos que a formação de palavras ocorre no componente sintático da gramática, e tais distinções superficiais percebidas nos compostos são reflexos da presença ou ausência de determinados traços gramaticais em sua estrutura sintática.

³⁴ Essa questão já foi amplamente discutida no trabalho de Gonçalves (2011a).

c. *Presença de uma vogal de ligação entre os componentes*

Há a presença de uma vogal de ligação (o ou i) entre os constituintes do composto no contexto fonológico /...C_fC_i.../, em que C_f e C_i são consoantes na posição final do primeiro radical e na posição inicial do segundo radical, respectivamente (e.g., *antrop-o-logia*, *pisc-i-cultura*).

Começamos pela característica (28a). O caráter preso apontado pelos autores não especifica, exclusivamente, os radicais neoclássicos, pois essa característica pode ser atrelada a qualquer radical das línguas românicas, uma vez que a extensa maioria de seus radicais carece de vogais temáticas para adquirir seu licenciamento fonológico, como comentado na seção 3.3; logo, todo radical é, a priori, preso³⁵. Contudo, os radicais neoclássicos devem ser considerados sempre atemáticos, como sugerem Scher e Nóbrega (2014, p. 78), já que eles não são licenciados fonologicamente quando combinados a um morfema de classe³⁶.

Tomemos, por exemplo, os radicais *cardi-*, *agr-* e *antrop-*, e concatenemos a eles uma vogal temática qualquer – ou seja, um marcador de classe qualquer. O que percebemos é que, diferentemente dos demais radicais do PB, eles não são, de fato, licenciados sintaticamente, algo ilustrado em (29), abaixo:

- (29) a. *cardi-o, *cardi-a, *cardi-e, *cardi-Ø³⁷
 b. *agr-o, *agr-a, *agr-e, *agr-Ø
 c. *antrop-o, *antrop-a, *antrop-e, *antrop-Ø

No entanto, há radicais que, embora tenham sua origem no grego antigo, acabaram adquirindo uma marcação de classe no PB, com significados específicos em contextos particulares, por exemplo, *psíqu-e*, *neur-a*, *morf-e*, *hidr-a*, *fon-e*, *bíbli-a*, *metr-o*, dentre outros. Esses radicais, por sua vez, não podem ser considerados atemáticos, pois carregam

³⁵ É desse caráter preso que nasce uma série de discussões acerca do estatuto morfológico dos radicais neoclássicos na literatura, a saber: se são afixos ou radicais. Como argumentado por Scalise (1984, p. 75-76), eles devem ser tratados como radicais, uma vez que (i) podem transitar internamente nos compostos (e.g., *morf-* que pode estar na primeira posição, *morfologia*, e na segunda posição, *antropomorfia*), diferentemente dos afixos que são “fixos”, (ii) podem receber a adição de afixos, tornando-se formas livres (e.g., *hidr-ico*, *neur-ônio*, *morf-ema*, etc.), algo que não acontece com os demais afixos da língua (e.g., **des-ismo*, **in-mento*) e (iii) carregam conteúdo lexical, diferentemente dos afixos.

³⁶ Para os autores, a ausência de informação de classe não os inclui na Classe IV, a qual é realizada por um morfema zero, Ø, pois mesmo os radicais pertencentes a essa classe são licenciados fonologicamente (e.g., *mar*, *cor*, *amor*).

³⁷ Vale mencionar que *cardio* existe como forma truncada de *cardiologia*, mas não como uma palavra independente que contenha algum significado remetendo a “coração”.

informação de classe nominal e são licenciados fonologicamente, contra a hipótese de Scher e Nóbrega (2014).

Tendo em vista essa informação, podemos, então, assumir que é a ausência da informação de classe que irá determinar o caráter preso de um radical no PB, seja ele de qual origem for. A mesma explicação pode ser expandida para as demais línguas românicas, e para o grego moderno, pois todas essas línguas dependem de uma marca de classe realizada fonologicamente para licenciar seus radicais (cf. HARRIS, 1991, 1999; RALLI, 2000, 2002; ALEXIADOU, 2004).

Como veremos no Capítulo 5, nossa explicação para esse fato é a de que a realização de uma raiz sintática como um radical implica a ausência de marcação de classe em seu categorizador. Nesse sentido, a gramática não faz distinção entre quais raízes sintáticas podem ou não receber um marcador de classe, pois sua presença é variável nos dados, tal como é possível verificar na (i) presença de radicais vernaculares em compostos de estrutura [radical + radical] (e.g., *cervejochato*, *musicólogo*, etc.) – o que não seria esperado, visto que eles deveriam conter suas marcas de classe, e no (ii) licenciamento de radicais neoclássicos via marcadores de classe (e.g., *psíque*, *hidra*, etc.).

O que a gramática interpreta, portanto, é a presença ou a ausência de um traço de classe no categorizador de uma raiz, o qual será responsável por desencadear a inserção de um marcador de classe no componente morfológico, externalizando, assim, uma raiz sintática em uma palavra ou em um radical. Com isso, a distinção entre radicais vernaculares e radicais neoclássicos se perde, pois ambos os tipos de radicais podem ser uniformizados de acordo com a possibilidade ou não de se tornarem autônomos fonologicamente através da marcação de classe. Assim sendo, todas as raízes sintáticas são potencialmente atemáticas, visto que será a ausência de informação de classe em seus núcleos categorizadores que condicionará sua realização como um radical pós-sintaticamente³⁸.

Sobre a característica (28b), apontada por Amiot e Dal (2007), acreditamos que ela é restritiva, pois esse tipo de composição não se circunscreve a termos técnicos ou científicos unicamente, mas pode ser encontrado em uma variedade de outros contextos, como exemplificado nos dados em (30). Além disso, ao abordar tais radicais, e, consequentemente,

³⁸ Percebamos que essa hipótese se circunscreve, a priori, às línguas com marcadores de classe realizados fonologicamente, o que serve como parâmetro para distinguir radicais que podem ou não ser licenciados. Já para o inglês, e para as demais línguas em que não há marcadores de classe realizados fonologicamente, supomos que existam marcadores de classe abstratos capazes de diferenciar quais são seus radicais presos, uma vez que essas línguas também proíbem o licenciamento de formas como *hydr-*, *neur-*, *agr-*, dentre outras. O chinês, por exemplo, que se assemelha ao inglês, também contém um conjunto de raízes presas na formação de compostos (cf. PIRANI, 2008), fato que pode ser explicado da mesma forma caso a hipótese de marcadores de classe abstratos estiver correta.

suas palavras compostas desse modo, os autores sugerem que os falantes não operam intuitivamente com elas, como se figurassem de modo paralelo em seu conhecimento linguístico, embora, sejam utilizadas de modo recorrente na língua.

(30) psicopoeta, neurofunk, agrobey, eletroacupuntura, hidrojump.

A terceira e última característica levantada pelos autores aponta outro aspecto formal desses compostos, que é a exigência de um elemento de ligação conectando os radicais, nomeadamente, as vogais *-o-* e *-i-*. A distribuição dessas vogais é resíduo de uma informação etimológica, a saber: a vogal *-o-* é inserida quando o radical na segunda posição tem origem no grego antigo, e a vogal *-i-* é inserida quando o radical na segunda posição é de origem latina. Essa distribuição está resumida nas regras em (31), e exemplificada em (32):

- (31) a. Se o radical na segunda posição de um composto for de origem grega, insira a vogal *-o-* (cf. 32a)
 b. Se o radical na segunda posição de um composto for de origem latina, insira a vogal *-i-* (cf. 32b)
- (32) a. hidr-o-logia, lesb-o-fobia, cristal-o-grafia, ric-ô-metro, corrupt-o-cracia
 b. pisc-i-cultura, herb-i-cida, vin-í-cola, frut-í-voro

Embora, à primeira vista, essa distribuição sugira que o falante detém conhecimento sobre a etimologia do radical, argumentamos que seu conhecimento linguístico não toma por base essas informações. No PB, pudemos verificar que a vogal *-o-* é a menos marcada para ocupar a posição de ligação, assunção que pode ser evidenciada nas (i) estruturas [radical + palavra], em que a palavra na segunda posição não é oriunda nem do grego antigo nem do latim, como em (33), e nos (ii) compostos em que, em vez de encontramos o marcador de classe *-a-*, particular aos radicais na primeira posição, o que temos é a vogal de ligação *-o-*, como em (34):

(33) hidr-o-pilates, heter-o-machista, neur-o-computador, psic-o-poeta

- (34) a. promess-ó-metro, denunci-o-cracia, music-o-logia, epistol-o-grafia, cam-o-terapia
 b. *promess-a-metro, *denunci-a-cracia, *music-a-logia, *epistol-a-grafia, *cam-a-terapia

Vale mencionar que os dados em (34) excluem a possibilidade de se analisar a vogal de ligação -o- como a realização do marcador de classe I do PB, seguindo a divisão de classes nominais sugerida por Alcântara (2003, 2010), pois se essa posição de junção pudesse ser ocupada por um marcador de classe, as construções em (34b) não seriam barradas pela gramática dessa língua.

Somando-se aos dados elencados acima, os exemplos reunidos em (35) endossam o caráter menos marcado da vogal -o- em contextos de ligação, e também deflagram um aspecto relevante já mencionado ao longo desta seção, a saber: que não são apenas radicais de origem neoclássica que compõem as estruturas [radical + radical] e [radical + palavra] no PB, mas que essas estruturas estão disponíveis a qualquer radical da língua para a formação de um composto:

- (35) a. “Existem também variações digitais dos *cervej-o-chatos*. Eles passam o dia em redes sociais, comentando os passos de pessoas do meio *cervejeiro* (...)”³⁹
 b. “Cara, é impressionante como o IG se tornou *Kassab-o-serrista* depois das eleições.”⁴⁰
 c. “Para você mulher que sempre sonhou em não precisar usar aquela privada suja do banheiro de festa sentada, surgiu a solução: *xerec-o-pênis*.”⁴¹
 d. “Escovas *muc-o-dental* e *dent-o-bucal* específicas para bebês e crianças de baixa idade.”⁴²
 e. “A Casa de Cultura *visu-o-gestual* visa, sobretudo, o aprendizado da língua de sinais brasileira em diferentes níveis de competência comunicativa (...)”⁴³

³⁹ Extraído do site: <http://comidasebebidas.uol.com.br/colunas/conversa-de-bar/2012/12/14/o-surgimento-dos-servejochatos.htm>. Acesso em: 14 dez. 2012.

⁴⁰ Extraído de um comentário presente no site: <http://jornalggn.com.br/blog/luisnassif/a-incrivel-persistencia-de-serra>. Acesso em: 15 jan. 2013.

⁴¹ Extraído do site: <http://www.tvnarua.com.br/index.php?s=noticia&&cat=entretenimento&&id=22332>. Acesso em: 10 mar. 2013.

⁴² Extraído do site: <http://www.patentesonline.com.br/escovas-muco-dental-e-dento-bucal-espec-ficas-para-beb-s-e-crian-as-de-baixa-idade-133502.html>. Acesso em: 15 dez. 2012.

⁴³ Extraído do site: <http://ccev-miralles.blogspot.com.br/>. Acesso em: 15 dez. 2012.

Além dos dados em (35), podemos citar também compostos como *sintatic-o-pragmático*, *politic-o-estrutural*, *franc-o-polonês*, *tóxic-o-dependente* dentre vários outros. Com isso, a distribuição das vogais de ligação pode ser reformulada sem se recorrer à etimologia, sendo necessário, apenas, marcar na gramática quais radicais exigem a presença da vogal de ligação *-i-* – os quais são poucos no PB (e.g., *-cid*, *-vor*, *-col*, *-cultur*) –, pois, nos demais contextos, a vogal *-o-* é o segmento *default* responsável por preencher a posição de ligação quando o primeiro radical termina em uma consoante e o segundo inicia-se em uma consoante, tal como descrito por Amiot e Dal (2007), em (28c). Essa solução, a nosso ver, é suficiente para descrever o fenômeno⁴⁴.

Em contrapartida, Bopp da Silva (2010, p. 94-104), partindo de uma noção semântica do conceito de palavra, admite que os radicais “vernaculares”, embora presos, devem ser considerados palavras, pois carregam conteúdo lexical. Em sua distinção, serão radicais apenas os chamados radicais neoclássicos, o que gera uma quarta divisão estrutural no PB, nomeadamente, a dos compostos [palavra + radical], responsável por abrigar os dados em (36):

(36) a. sexologia	[sexo] _P [logia] _R
b. arquivonomia	[arquivo] _P [nomia] _R
c. historiografia	[histori] _P (o) [grafia] _R
d. cocaínômano	[cocain] _P (o) [mano] _R

(BOPP DA SILVA, 2010, p. 98)

A divisão proposta pela autora não considera, integralmente, a posição de ligação presente nas estruturas [radical + radical] e [radical + palavra], pois analisa a vogal *-o-* de *sex-o* e *arquiv-o*, em (36a) e (36b), como um marcador de classe e não como uma vogal de ligação. Essa última, por sua vez, seria inserida apenas quando o radical de primeira posição contivesse um marcador de classe diferente da vogal *-o-*, como ocorre em (36c) e (36d), com as palavras *históri-a* e *cocaín-a*.

Ora, como já argumentado acima, se a vogal *-o-* em (36a) e (36b) fosse, de fato, um marcador de classe, **histori-a-grafia* e **cocaín-â-mano* deveriam ser perfeitamente gramaticais em PB, e não são. Sendo assim, argumentamos que a combinação estrutural [palavra + radical] não faz parte das possibilidades de realização estrutural dos compostos do

⁴⁴ Scher e Nóbrega (2014) fornecem uma sugestão de como essa distribuição das vogais de ligação nos compostos formados por radicais do PB pode ser implementada na gramática.

PB, e os compostos em (36) são constituídos de uma estrutura [radical + radical], com a vogal -o- sendo o elemento de ligação responsável por concatenar os radicais.

Tendo discutido as características gerais apontadas por Amiot e Dal (2007), passemos, agora, para uma caracterização das duas estruturas morfológicas sob análise. Embora elas estejam incluídas em uma mesma subseção, há algumas diferenças morfossintáticas e fonológicas presentes em seus dados. Sobre as diferenças morfossintáticas, referimo-nos às relações de concordância estabelecidas com o composto. O que notamos é que, em compostos [radical + palavra], parece ser o gênero da palavra na segunda posição que controla a concordância nesses dados, fato ilustrado com os exemplos em (37):

(37)	<i>Palavra</i>	<i>Composto</i>	<i>Concordância</i>
a.	terapia _(FEM)	psic-o-terapia	a _(FEM) psicoterapia _(FEM)
b.	esporte _(MASC)	hidr-o-esporte	o _(MASC) hidroesporte _(MASC)
c.	avaliação _(FEM)	heter-o-avaliação	a _(FEM) heteroavaliação _(FEM)

Já nos compostos [radical + radical], quem controla a concordância é o gênero marcado no sufixo localizado em sua borda direita, como em (38a), ou o gênero natural relativo ao sexo do referente, como em (38b) e (38c):

(38)	<i>Composto</i>	<i>Concordância</i>
a.	-ia _(FEM)	palpit-o-metr-ia a _(FEM) palpímetria _(FEM)
b.	_(MASC)	music-o-graf-o o _(MASC) musicógrafo _(MASC)
c.	_(FEM)	sex-o-log-a a _(FEM) sexóloga _(FEM)

Além disso, é válido mencionar que ambas as estruturas permitem afixos flexionais apenas em sua borda direita, deixando o primeiro radical impossibilitado de receber qualquer morfema flexional:

(39)	<i>[radical + radical]</i>	<i>[radical + palavra]</i>
a.	sex-o-loga-s/*sex-o-s-loga	psic-o-terapia-s/*psic-o-s-terapia
b.	sex-o-logu-inha/*sex-inha-loga	psicoterapia-zinha/*psiquinho-terapia

Ralli e Karasimos (2009) argumentam que essa é uma propriedade inerente aos compostos formados por radicais, e tentam capturá-la impondo uma restrição a suas

formações, a qual eles rotulam de Restrição do Radical Nu (do inglês, *Bare-Stem Constraint*), definida em (40). O que sugerimos, por outro lado, é que a impossibilidade de marcação flexional no primeiro radical se deve à ausência um traço de classe nos radicais dessa posição, como argumentado acima, uma vez que sem essa informação não é possível concatenar qualquer morfema adicional a um radical⁴⁵.

(40) *Restrição do Radical Nu (RALLI; KARASIMOS, 2009, p. 33, tradução nossa)*

A coesão de um composto é mais bem garantida se o primeiro radical estiver o mais nu possível.⁴⁶

No que diz respeito à categoria lexical desses radicais, compartilhamos com Baeskow (2004) a assunção de que os radicais “neoclássicos” carregam categoria nominal, embora seus categorizadores não sejam realizados fonologicamente. Um argumento apontado pela autora é o de que há compostos neoclássicos sem qualquer sufixo nominalizador, e que mesmo assim eles carregam categoria nominal, por exemplo, *telephone* ‘telefone’, formado por *tele* e *phone*, e *microscope* ‘microscópio’, de *micro* e *scope*. Casos semelhantes são encontrados no PB, como os compostos *psicólog-o* e *termômetr-o*, os quais são nominais embora não contenham nenhum afixo nominalizador em seus membros.

Outro argumento para o caráter nominal desses radicais é o fato de que, quando licenciados fonologicamente, eles recebem marcadores de classe nominais (e.g., *psíqu-e*, *neur-a*, *hidr-a*, *morf-e*, *fon-e*, *metr-o*) e não vogais temáticas verbais, por exemplo. Namer e Villoing (2006), por outro lado, expandem suas categoriais lexicais, ao admitir que esses radicais não se circunscrevem apenas à categoria nominal, mas podem também ser verbais, como [*anthrop*_N*phage*_V]_A ‘antropofagia’ e [*gén*_N*cide*_V]_{A/N} ‘genocida’, em que *-fage* e *-cide* parecem exigir *anthrop-* e *gén-* como seus complementos.

Como veremos no Capítulo 5, a categorização é imprescindível para que uma raiz sintática seja manipulada no espaço computacional, seguindo a hipótese das raízes acategoriais apresentada no Capítulo 2, uma vez que uma raiz sintática apenas obterá informações gramaticais ao se concatenar a um núcleo definidor de categoria. Para nós, portanto, não é a presença ou ausência de um núcleo categorizador que tornará raízes

⁴⁵ Cf. Scher e Nóbrega (2014) para outra alternativa de análise, baseada na ausência de informação categorial.

⁴⁶ Trecho original (RALLI; KARASIMOS, 2009, p. 33): “Bare-Stem Constraint: the cohesion of a compound is better guaranteed if the first stem is as bare as possible”.

sintáticas em “radicais presos” no componente morfológico, mas sim a ausência do traço de classe, como vimos argumentando ao longo dessas duas últimas seções.

Por último, queremos apresentar, brevemente, algumas diferenças fonológicas presentes nessas estruturas, algo que iremos explorar, mais demoradamente, no Capítulo 6. Os trabalhos de Nespor e Ralli (1996) e Nespor (1999) apontam que o número de palavras fonológicas de um composto está associado à sua estrutura morfológica. Nespor (1999), particularmente, mostra que as línguas românicas apresentam um padrão semelhante com relação ao número de palavras fonológicas em compostos [radical + radical], algo que pode ser evidenciado pela presença de um único acento primário nesses dados, fato igualmente atestado no PB:

(41) a. *cervej-ó-logo*⁴⁷

b. *palpit-ó-metro*

c. *psic-o-logía*

d. *histori-o-gráfico*

Compostos de estrutura [radical + palavra], por outro lado, constituem duas palavras fonológicas. No PB, por exemplo, percebemos que o acento da palavra na segunda posição é preservado, visto que ele se encontra no mesmo local em que é atribuído à palavra em contexto isolado. Além desse acento, um segundo acento é marcado no radical de primeira posição, como pode ser visto em (42):

(42) a. *hìdr-o-ginástica*

b. *nèur-o-ciência*

c. *psìc-o-biología*

d. *cervèj-o-cháto*

Em resumo, as informações elencadas até aqui visam demonstrar que as estruturas [radical + radical] e [radical + palavra] estão integradas ao sistema linguístico do PB, e não são periféricas como a literatura sugere, uma vez que qualquer radical da língua – ou raiz sintática, se pensarmos em sua formação – podem compor essas palavras. Além disso, todos

⁴⁷ Marcamos o acento primário com um acento agudo (“’”) e o acento secundário com um acento grave (“`”).

esses dados serão explicados da mesma forma de um ponto de vista gerativo, sem a necessidade de se recorrer a qualquer informação etimológica ou nomenclatura especial.

Para finalizar, resumimos as características dos compostos de estrutura [radical + radical] e [radical + palavra], em (43) e (44), abaixo:

(43) *Características morfológicas, morfossintáticas e fonológicas dos compostos de estrutura [radical + radical]*

- a. Presença de um elemento de ligação conectando os radicais;
- b. Informação de gênero determinada pelo sufixo na borda direita ou pelo referente;
- c. Adição de marcas flexionais apenas na borda direita do composto;
- d. Constituição de uma única palavra fonológica.

(44) *Características morfológicas, morfossintáticas e fonológicas dos compostos de estrutura [radical + palavra]*

- a. Presença de um elemento de ligação conectando os radicais;
- b. Informação de gênero determinada pelo gênero da palavra na segunda posição;
- c. Adição de marcas flexionais apenas na borda direita do composto;
- d. Constituição de duas palavras fonológicas.

Na próxima subseção, iremos nos concentrar nos compostos de estrutura [palavra + palavra], avaliando as mesmas características morfológicas, morfossintáticas e fonológicas presentes nesses dados.

3.3.2.2. Estrutura [palavra + palavra]

A estrutura [palavra + palavra] é constituída por unidades fonologicamente autônomas, e se caracteriza como a forma preferencial para a realização fonológica dos compostos do PB. Diferentemente das estruturas tratadas na subseção anterior, elas foram exploradas pela grande maioria dos morfólogos que se propuseram a falar de composição nessa língua, algo que pode ser verificado na atenção fornecida a esses dados em suas classificações.

Esse tipo estrutural abriga uma ampla variedade de combinações categoriais, como visto na seção 3.2, sendo as mais produtivas, as combinações N-N, V-N, A-N e Num-N (cf. ALVES, 2007). Essa variedade de combinações categoriais desencadeia diferentes padrões de marcação do plural e de concordância, as quais podem variar a depender das relações

gramaticais presentes entre os membros de um composto (viz., subordinação ou coordenação), como argumentado por Moreno (1997, 2002).

Kehdi (2002, p. 45-48) e Bechara (2006, p. 129), por exemplo, descrevem a distribuição da marcação de plural nos compostos [palavra + palavra] mais produtivos, da seguinte maneira:

- (45) *Distribuição da marcação de plural em compostos de estrutura [palavra + palavra]*
- a. *Compostos N-A e A-N*: o adjetivo concorda com o nome (e.g., *amor-perfeito* – *amores-perfeitos*);
 - b. *Compostos N-N e N-de-N*: o nome que serve como modificador não recebe marcação de número, pois nomes não estabelecem concordância com outros nomes (e.g., *navio-escola* – *navios-escola*; *pão-de-ló* – *pães-de-ló*);
 - c. *Compostos N-N e A-A em uma relação de coordenação*: como não há um núcleo, ambos podem flexionar (e.g., *surdo-mudo* – *surdos-mudos*; *aluno-mestre* – *alunos-mestres*).
 - d. *Compostos V-N*: apenas o nome permite concordância em número (e.g., *saca-rolha* – *saca-rolhas*);

O que podemos notar, primeiramente, é que, enquanto compostos formados por radicais permitem marcação de plural, exclusivamente, em sua borda direta, os compostos formados por palavras permitem marcação de plural tanto no primeiro quanto no segundo membro de seus compostos, a depender da categoria lexical de cada um (e.g., compostos N-N e N-A). Normalmente, a determinação de qual constituinte receberá ou desencadeará marcação flexional (seja ela marcação de número, gênero⁴⁸ ou diminutivo) é dada pelo núcleo semântico do composto, tal como ocorre nos compostos N-N subordinados (e.g., os *peixes-espada*, os *peixinhos-espada*) e nos compostos N-A (e.g., as *escadas-rolantes*, as *escadinhas-rolantes*).

Moreno (2002) aponta, por outro lado, que a presença de flexão interna aos compostos [palavra + palavra] é um assunto controverso, pois parece não existir, à primeira vista, uma sistematicidade em sua aplicação. Isso se percebe pela hesitação que os falantes apresentam ao tentar determinar quais marcações de plural são aceitáveis nos compostos do PB. Uma das razões para esse fato, segundo o autor, é a ocorrência de reanálises de uma geração para outra,

⁴⁸ Com exceção do gênero natural, ou seja, aquele que depende do referente (e.g., *o/a casca grossa*).

em que enquanto uma geração percebe um composto X+Y, outra o interpreta como uma única unidade, por exemplo: *guarda-marinha* (*guardas-marinha*, *guardas-marinhas* ou *guarda-marinhas*), *navio-escola* (*navios-escola*, *navios-escolas* ou *navio-escolas*) e *decreto-lei* (*decretos-lei* e *decretos-leis*).

Para os compostos em (45a), (45b) e (45c), Moreno (2002), contrariamente às visões de Kehdi (2002) e Bechara (2006), explica a variação na marcação de plural com base nas relações flexionais presentes em um sintagma nominal, visto que, para o autor, os falantes tendem a interpretar os compostos como sintagmas nominais. Além disso, o autor comenta os contrastes flexionais presentes nas relações de modificação e coordenação, algo já apresentado e discutido na seção 3.2 deste capítulo.

Os compostos V-N, por outro lado, apresentam um comportamento mais uniforme, pois apenas o nome na segunda posição recebe marcação de plural (e.g., *porta-malas*, *salva-vidas*, etc.). Entretanto, a não marcação do plural nos nomes internos a esses compostos também se verifica no PB, pois dados como *porta-mala* e *salva-vida* são gramaticais, e ocorrem mesmo quando os nomes na segunda posição carregam um plural irregular, como *lustra-móvel* (de *lustra-móveis*).

Nunes (2007, p. 30) explica que a ausência de marcação do plural nos compostos V-N é consequência de uma reanálise envolvendo indefinidos plurais, tendo em vista que, no PB, em decorrência da marcação de plural ocorrer, geralmente, no determinante (e.g., *os menino*), é possível encontrar sentenças em que o objeto é um sintagma singular nu (e.g., *o João nunca compra livro*) em vez de um sintagma plural nu (e.g., *o João nunca compra livros*), fato que estaria se espraiando para compostos dessa natureza.

Já com relação à concordância de gênero, os compostos V-N apresentam uma distribuição particular, que leva em consideração a interpretação do composto, ou seja, se o composto é agentivo ou instrumental. Dessa forma, enquanto compostos V-N instrumentos desencadeiam concordância de gênero masculino – independentemente do gênero de seu nome interno –, os compostos V-N agentivos desencadeiam concordância de gênero natural, ou seja, as marcas de gênero nos determinantes e adjetivos dependem do gênero do referente. Tais padrões podem ser verificados nos dados trazidos em (46):

(46) a. *Compostos V-N instrumentos*

(o) conta-gota(s); (o) porta-mala(s); (o) para-raio; (o) saca-rolha(s); (o) mata-burro; (o) porta-avião(ões); (o) limpa-vidro(s); (o) lustra-móvel(is); (o) mata-barata; (o) tira-mofo; (o) marca-passo;

b. *Compostos V-N agentivos*

(o/a) desmancha-prazer; (o/a) fura-olho; (o/a) quebra-galho; (o/a) sabe-tudo; (o/a) caga-regra(s)

Retornaremos a essas questões relativas à marcação flexional e concordância nos compostos N-N e V-N quando explicarmos a derivação dos compostos no Capítulo 5. O que devemos reter, no momento, é que compostos [palavra + palavra] permitem uma amplitude flexional maior, se comparada a dos compostos formados por radicais. A nosso ver, essa variedade está intimamente relacionada à estrutura morfológica desses compostos, ou seja, uma vez que ambos os radicais recebem informação de classe, o licenciamento de outros morfemas flexionais é possível, embora sua distribuição seja contextual.

Por fim, de um ponto de vista fonológico, é importante ressaltar que os compostos [palavra + palavra] formam duas palavras fonológicas, o que os assemelha a uma frase fonológica. Tal fato pode ser evidenciado pela presença de dois acentos – um primário e um secundário – distribuídos em cada um dos membros do composto:

- (47) a. vista-gróssa
 b. banho-maría
 c. lua-de-mél
 d. mata-baráta

Acreditamos que essa dupla marcação acentual se deve ao traço de classe, o qual irá regular a identificação de uma palavra no componente fonológico da gramática. A aplicação de acento, por sua vez, tanto de compostos formados por radicais quanto de compostos formados por palavras será discutida em detalhes no Capítulo 6, quando tratarmos, especificamente, da externalização dos compostos de uma língua – no caso, da nossa língua particular sob análise.

Em resumo, podemos dizer que os compostos de estrutura [palavra + palavra] apresentam as seguintes características:

- (48) *Características morfológicas, morfossintáticas e fonológicas dos compostos de estrutura [palavra + palavra]*
 a. Não desencadeia a presença de um elemento de ligação;

- b. Informação de gênero variável a depender da interpretação do composto ou do gênero de seu referente;
- c. Adição de marcas flexionais variável, podendo ocorrer interna ou externamente ao composto;
- d. Constituição de duas palavras fonológicas.

Antes de concluir esta seção, exploramos as estruturas morfológicas dos compostos em outras línguas, a fim de evidenciar seu caráter particular na formação dos compostos das línguas naturais.

3.3.2.3. A variedade de estruturas morfológicas translinguisticamente

O motivo de apresentar um breve panorama translinguístico das estruturas morfológicas é o de demonstrar sua distribuição particular às línguas naturais, uma vez que a variabilidade no processo de composição reside no modo como os compostos são externalizados, algo que se processa no ramo de PF da gramática. Essa variação encontrada nas estruturas morfológicas é o que consideramos como evidência para seu caráter particular translinguisticamente, ao passo que a combinação de seus membros decorre de um sistema universal.

Ralli (2013, p. 10), ao discutir as estruturas morfológicas dos compostos do grego moderno, hipotetiza que a preferência que algumas línguas têm por formar compostos a partir de radicais decorre da presença de um sistema flexional rico, como é o caso do próprio grego moderno. Nessa língua, embora sejam encontradas as mesmas estruturas morfológicas atestadas nos compostos do PB, como é possível verificar em (49), há uma preferência em se realizar os compostos a partir de estruturas formadas por radicais:

(49) [radical + radical]

(Grego Moderno)

a.	<i>xris-ó-skini</i>	<	<i>xris-</i>	<i>skon-</i>
	‘pó dourado’		‘dourado’	‘pó’
b.	<i>karav-ó-pano</i>	<	<i>karav-</i>	<i>pan-</i>
	‘vela’		‘barco’	‘roupa’

[radical + palavra]

c.	<i>elaf-o-kiniyós</i>	<	<i>elaf-</i>	<i>kiniyós</i>
	‘caçador de cervos’		‘cervo’	‘caçador’
d.	<i>xrisavjí</i>	<	<i>xris-</i>	<i>avjí</i>
	‘amanhecer dourado’		‘dourado’	‘amanhecer’

[palavra + palavra]

e.	<i>omáða eryasías</i>	<	<i>omáða</i>	<i>eryasía</i>
	‘grupo de trabalho’		‘grupo’	‘trabalho’
f.	<i>pedí θávma</i>	<	<i>pedí</i>	<i>θávma</i>
	‘menino prodígio’		‘menino’	‘prodígio’

(NESPOR; RALLI, 1996, p. 359-360)

A mesma preferência é encontrada nas línguas eslavas, como no russo, no esloveno, no tcheco, entre outras línguas dessa família (cf. PERKLES, 2008; LIEBER; ŠTEKAUER, 2009b; ŠTICHAUER, 2009). Ralli (2013) comenta também que línguas com um sistema flexional menos robusto, como é o caso do inglês, terão compostos formados por radicais em menor quantidade, além do fato desses radicais se restringirem a um conjunto específico, a saber, o dos radicais “neoclássicos” (e.g., *sociology*, *anthropology*)⁴⁹.

Se expandimos nossa visão e explorarmos os compostos das línguas não-indo-europeias, iremos perceber que a variedade de externalização se amplia. Um exemplo dessa variedade é dado pelo tzotzil, língua maia, que não apenas traz uma variedade de realizações, como também nos permite constatar a complexidade de formas que podem ser unidas formando novos compostos. Segundo Cowan (1969, p. 89 *apud* ŠTEKAUER; VARELA; KÖRTVÉLYESSY, 2012, p. 4), as unidades morfológicas que servem como membros de um composto em tzotzil são:

(50) *Unidades morfológicas que servem como membros de compostos no tzotzil*

- a. raiz (um único morfema);
- b. radical (i.e., uma raiz com um ou mais afixos derivacionais);
- c. um composto, e

⁴⁹ Essa análise pode variar a depender do modo como os radicais neoclássicos são analisados no inglês, língua que permite o licenciamento de suas raízes independentemente de qualquer marcação de classe. Bauer (1998), por exemplo, não considera que essas formas sejam tratadas como radicais, e as analisa como formas combinatórias, tratando a vogal de ligação como parte do primeiro membro do composto.

- d. um radical complexo derivado de um composto (i.e., um composto com um ou mais afixos derivacionais)

A variedade de combinações entre as unidades morfológicas elencadas em (50) dá origem à seguinte distribuição estrutural nos compostos dessa língua:

Quadro 1 - Algumas combinações da estrutura morfológica dos compostos em tzotzil (ŠTEKAUER; VALERA; KÖRTVÉLYESSY, 2012, p. 42 *apud* COWAN, 1969, p. 89)

Combinações	Exemplos
[raiz + raiz]	<i>pĩš-xól</i> embrulhar-cabeça ‘chapéu’
[raiz + radical]	<i>méʔ-[kʰín-ob-al]</i> mãe-névoa ‘arco-íris’
[radical + raiz]	<i>[tʰsúb-il]-lúm</i> pó-solo ‘poeira’
[radical + radical]	<i>[sĩk-il]-[tšám-el]</i> frio-qualidade-doença ‘malaria’
[raiz + composto]	<i>tĩʔ-[mák-téʔ]</i> abertura-[fechamento-madeira] ‘portão de cerca’
[radical + composto]	<i>[kʰán-al]-[ʔóʔn-ton]</i> [amarelo-qualidade]-[coração-pedra] ‘pássaro de peito amarelo’
[composto + radical]	<i>[tóx-tóx]-[tšám-el]</i> [direto-direto]-[ato de morrer] ‘morrer rapidamente’

Diferentemente do tzotzil, há línguas em que a variedade de estruturas para a realização dos compostos é reduzida, tal como é o caso das línguas isolantes, como o chinês e o dâw, língua amazônica do alto Rio Negro. Nessas línguas, a presença de afixos é quase

inexistente, e, geralmente, raízes não flexionadas coincidem com palavras simples, caracterizando-se como as formas básicas para a realização dos compostos. O chinês, por sua vez, além de formar compostos a partir da combinação de raízes, também forma compostos a partir da combinação de um conjunto restrito de raízes presas (cf. PACKARD, 2000; PIRANI, 2008). Exemplos de seus compostos estão apresentados em (51):

(51) [palavra (raiz simples) + palavra (raiz simples)] (Chinês)

a.	<i>bīngshan</i>	<	<i>bīng</i>	<i>shan</i>
	iceberg		gelo	montanha

[raiz presa + raiz presa]

b.	<i>mùban</i>	<	<i>mù-</i>	<i>ban</i>
	quadro		mandeira	quadro

(PACKARD, 2000, p. 90-91)

No dâw, compostos são formados apenas pela concatenação de raízes, que, como mencionamos, caracterizam-se como palavras simples. Tais palavras simples tendem a ser monossilábicas nessa língua, fazendo com que os compostos sejam identificados a partir do número de sílabas, ou seja, como palavras dissilábicas (cf. MARTINS, 2004, p. 126). Exemplos de compostos em dâw estão listados em (52):

(52) [palavra (raiz simples) + palavra (raiz simples)] (Dâw)

a.	<i>jɔn-tòʃ</i>	<	<i>jɔn</i>	<i>tòʃ</i>
	espingarda		tamanduá	nariz
b.	<i>xu-fɛ</i>	<	<i>xu</i>	<i>fɛ</i>
	cachoeira		descer	escorrer

(MARTINS, 2004, p. 143-144)

Para testar, então, a hipótese das estruturas morfológicas como uma propriedade específica às línguas, tomemos a relação de modificação, mais especificamente a modificação presente em compostos N-N, e a consideremos como uma propriedade universal da composição, uma vez que a relação de modificação em si não varia entre as línguas naturais. Ao reunir um conjunto de compostos com essas exigências, percebemos que o único aspecto

variável é o modo como a língua em questão externaliza seus compostos, mas não a combinação entre os dois nomes, nem o estabelecimento da relação de modificação:

(53) [raiz presa + raiz presa]

a.	<i>diàn-zǐ</i>	eletricidade-partícula	‘elétron’	(Chinês)
b.	<i>tou-hoku</i>	leste-norte	‘nordeste’	(Japonês)

[radical + radical]

c.	<i>gwiazd-o-zbiór</i>	estrela-EL-coleção	‘constelação’	(Polonês)
d.	<i>jazyk-o-věda</i>	língua-EL-ciência	‘linguagem científica’	(Tcheco)

[radical + palavra]

e.	dent-o-bucal	dent-EL-bucal		(Português)
f.	<i>ayri-o-ánthropos</i>	selva-EL-homem	‘homem selvagem’	(Grego)

[palavra + palavra]

g.	<i>sword fish</i>	espada-peixe	‘peixe espada’	(Inglês)
h.	<i>célul-a madr-e</i>	célula-mãe	‘célula-mãe’	(Espanhol)

O que podemos inferir a partir dos dados em (53) é que a gramática deriva um composto concatenando duas raízes sintáticas em uma determinada relação gramatical, no caso, duas raízes sintáticas categorizadas como nomes em uma relação de modificação. O modo como essa concatenação se processa na gramática não deve variar entre as línguas, pois, como mencionamos, ela ocorre em um componente sintático de operações universais. Entretanto, o que varia é o modo como essa estrutura de modificação tem seus membros realizados fonologicamente, se como formas dependentes – radicais – ou como formas autônomas – palavras.

Além disso, é importante mencionar que, embora a presença de compostos formados por radicais em uma língua tenha um correlato com sua complexidade flexional, tal como propõe Ralli (2013), assumimos que a emergência de compostos formados radicais está codificada na gramática através da distribuição do traço de classe nos categorizadores de uma raiz sintática, o qual é responsável por determinar quais raízes sintáticas poderão ser realizadas como formas dependentes ou como formas autônomas.

É essa organização operacional que sugerimos para a formação dos compostos nas línguas naturais. Na próxima seção, demonstramos como o caráter universal aludido sobre as combinações categoriais e as relações gramaticais, no último parágrafo, encontra evidência em uma série de dados presentes em trabalhos tipológicos e taxonômicos. Por fim, ilustramos como nossa língua específica – o PB – distribui suas combinações categoriais em relações gramaticais, e como tais combinações são emolduradas morfologicamente.

3.4. Mapeamento da composição: do geral ao específico

Recuperando as classificações dos compostos sugeridas para o PB, apresentadas na seção 3.2 deste capítulo, e anexando a elas a distribuição estrutural vista na última seção, separamos, a partir de agora, mais diretamente, o geral e o específico na composição. Como reiterado diversas vezes, as estruturas morfológicas são propriedades específicas às línguas, e tal afirmação foi demonstrada no final da última seção. Resta-nos, portanto, deflagrar o que das classificações para os compostos, vistas na seção 3.2, pode ser assumido como uma propriedade intrínseca e geral a esse processo de formação de palavras.

Resumidamente, o que podemos perceber até aqui é que o PB – tal como toda língua natural – contém um conjunto de categoriais lexicais disponíveis para a formação de seus compostos, e tais categoriais lexicais podem ser unidas, a priori, em dois tipos diferentes de relações gramaticais, nomeadamente, subordinação e coordenação. Vimos também que essas combinações categoriais conectadas em determinada relação gramatical podem ser realizadas fonologicamente de três modos nessa língua, a saber, através de uma combinação [radical + radical], [radical + palavra] ou [palavra + palavra].

As duas primeiras características supramencionadas estão relacionadas a dois dos tipos de classificação dos compostos do PB explorados na seção 3.2, a saber: (i) classificação por combinações categoriais e (ii) classificação a partir das relações gramaticais entre seus membros. Ambas, como veremos, são atestadas de modo uniforme em qualquer língua natural que apresente a composição como um processo de formação de palavras.

As combinações categoriais, por exemplo, são facilmente identificadas nos trabalhos descritivos feitos sobre a composição. O caráter universal advogado para essas combinações reside na presença de categorias lexicais idênticas nas línguas naturais (viz., nome, verbo e adjetivo, etc.) e na presença de combinações idênticas entre categorias lexicais translinguisticamente, embora uma ou outra combinação esteja ausente em alguma língua por razões de ordem de palavras ou preferência. No Quadro 2, abaixo, trazemos um conjunto de

línguas pertencentes a famílias linguísticas diferentes, a fim de ilustrar a presença das mesmas combinações categoriais em seus compostos⁵⁰:

Quadro 2 – Variedade de combinações categoriais dos compostos

	Chinês	Russo	Finlandês	Francês
[N + N]	<i>dú fàn</i> droga-vendedor ‘traficante’	<i>gorod-geroj</i> cidade-herói ‘herói da cidade’	<i>tee kuppi</i> chá-xícara ‘xícara de chá’	<i>homme-grenouille</i> homem-rã ‘homem-rã’
[N + A]	<i>dǎn xiǎo</i> coragem-pequena ‘covarde’	<i>žizn-e-radostnyj</i> vida-feliz ‘alegre’	<i>tähti kirkas</i> estrela-brilhante ‘estrelado’	<i>poids-lourd</i> peso-pesado ‘caminhão’
[A + N]	<i>zhǔ-yè</i> principal-página ‘página principal’	<i>such-o-frukty</i> seca-EL-frutas ‘frutas secas’	<i>kylmä varasto</i> frio-armazenamento ‘câmara frigorífera’	<i>vert-pomme</i> verde-maça ‘maça verde’
[A + A]	<i>gāo ǎi</i> alto-baixo ‘altura’	<i>těmn-o-sinij</i> escuro-EL-azul ‘azul escuro’	<i>hapan imelä</i> azedo-doce ‘doce e azedo’	<i>aigre-doux</i> agri-doce ‘agridoce’
[V + N]	<i>bèng chuáng</i> pula-cama ‘trampolim’	<i>liz-o-bljud</i> lambe-prato ‘bajulador’	---	<i>ouvre-boîte</i> abre-caixa ‘abridor de latas’
[N + V]	<i>dì zhèn</i> terra-tremer ‘terremoto’	<i>knig-o-ljub</i> livro-ama ‘bibliófilo’	<i>kanta esittää</i> raiz-executar ‘estrear’	---
[V + V]	<i>shì’ài</i> mostrar-amar ‘demonstrar amor’	<i>vert-o-lēt</i> girar-EL-voar ‘helicóptero’	---	<i>cache-cache</i> esconde-esconde ‘esconde-esconde’

Guevara e Scalise (2009), ao desenvolver uma busca por universais na composição, atestam o fato de que as combinações categoriais são uma característica universal da composição, comprovando tal alegação ao fazer um levantamento da distribuição dessas combinações nas línguas naturais. Partindo de uma base de dados constituída por compostos

⁵⁰ Os dados do Quadro 2 foram retirados dos seguintes trabalhos: do chinês, Ceccagno e Basciano (2007, 2009); do russo, Benigni e Masini (2009); do finlandês, Niemi (2009); do francês, Villoing (2012).

das seis macro-áreas linguísticas apontadas por Dryer (1992)⁵¹, os autores elencam as dez combinações categoriais mais produtivas translinguisticamente, as quais podem ser vistas na Tabela 1, abaixo:

Tabela 1 – Distribuição percentual das combinações categoriais dos compostos em diferentes famílias linguísticas (GUEVARA; SCALISE, 2009, p. 120)

Combinações	Média %	Românicas %	Germânicas %	Asiáticas %	Eslavas %
[N+N]	18.08	16.91	14.29	25.42	15.71
[A+N]	7.52	8.09	6.80	10.17	5.00
[N+A]	7.39	11.76	8.16	6.78	2.86
[A+A]	6.22	5.88	7.48	5.08	6.43
[V+N]	7.07	4.41	4.08	16.95	2.86
[N+V]	4.40	1.47	4.08	8.47	3.57
[V+V]	4.38	2.94	2.72	11.86	
[Adv+A]	2.31	3.68	2.72		2.86
[Adv+N]	2.27	1.47	4.76		2.86
[A+V]	2.07	0.74	2.72	3.39	1.43

Podemos notar com esses dados que algumas combinações categoriais são altamente recorrentes translinguisticamente, como é o caso da combinação [N+N], que representa 18.08% da distribuição total das combinações. É interessante notar também que enquanto combinações contendo nomes são preferenciais entre as línguas, as combinações que contêm adjetivos vêm em segundo lugar em produtividade, e as combinações com membros verbais são as que apresentam menor quantidade, sendo produtivas apenas nas línguas asiáticas (e.g., [V+N] com 16.95% e [V+V] com 11.86%).

Esse caráter universal alegado às combinações categoriais também pode ser defendido para as relações gramaticais presente nos compostos. Bisetto e Scalise (2005) e Scalise e Bisetto (2009), ao fornecer uma proposta de classificação universal para a composição, mostram que os membros de um composto são unidos por relações gramaticais de três tipos: subordinação, atribuição e coordenação, as quais estão definidas em (54). Essas relações não são exclusivas à composição, como bem pontuam Guevara e Scalise (2009, p. 104), e estão

⁵¹ Viz., África, Eurasia, sudoeste asiático e Oceania, Austrália e Nova Guiné, América do Norte e América do Sul.

presentes também nas construções sentenciais, contribuindo para a assunção de que elas devem ser consideradas uma característica presumivelmente universal.

(54) *Relações gramaticais internas à composição (BISETTO; SCALISE, 2005)*

- a. **Subordinação:** compostos em que há uma relação de complementação entre seus membros constituintes;
- b. **Atribuição:** compostos em que há uma relação de modificação entre seus membros constituintes;
- c. **Coordenação:** compostos em que se estabelece uma relação conjuntiva ou disjuntiva (conectados por uma conjunção “e” ou “ou”, respectivamente) entre seus membros constituintes.

Percebamos que a relação gramatical de subordinação apresentada nas gramáticas e manuais do PB⁵² é aqui definida como uma relação de atribuição exclusivamente, ao passo que a subordinação caracteriza compostos em que seus membros são concatenados em uma relação de complementação, na qual um membro seleciona o outro. Com isso, a distinção entre as relações presentes em compostos V-N e compostos N-N fornecida nos trabalhos do PB, principalmente no trabalho de Sandmann (1991), fica mais bem esclarecida.

Por exemplo, a relação interna entre os membros de um composto V-N é de subordinação, uma vez que seu verbo exige a concatenação de um nome complemento (e.g., *salva-vidas*, *estraga-prazer*, *limpa-vidros*). Já a relação interna entre os membros de um composto N-N pode ser de três tipos: (i) de subordinação: quando um nome seleciona o outro a partir de uma “relação de” (e.g., *aracnofobia*); (ii) de atribuição, quando um nome não-núcleo modifica o nome núcleo (e.g., *peixe-espada*); ou de (iii) de coordenação, quando os nomes estão em uma relação conjuntiva (e.g., *ator-diretor*)⁵³.

Guevara e Scalise (2009, p. 118) também mostram que, no que concerne às relações gramaticais, as línguas têm preferência em formar compostos a partir de uma relação de subordinação, como pode ser visto na Tabela 2, abaixo, (Subordinação: 40.06%). No entanto, considerando as famílias linguísticas isoladamente, é possível notar, a partir dos dados dessa mesma tabela, que as línguas românicas e eslavas apresentam uma preferência por compostos

⁵² Cf. seção 3.2.

⁵³ É válido mencionar que Rio-Torto e Ribeiro (2012) fizeram uma classificação dos compostos do português europeu nesses termos. Acreditamos, como as autoras, que a assunção dessa tripla divisão das relações gramaticais descreve melhor o modo com os membros de um composto são unidos.

atributivos, enquanto que as línguas do leste asiático têm uma preferência em formar compostos coordenados.

Tabela 2 – Distribuição percentual das relações gramaticais translinguisticamente (GUEVARA; SCALISE, 2009, p. 119)

Classificação	Média %	Românicas %	Germânicas %	Asiáticas %	Eslavas %
Subordinação	40.06	44.85	38.10	37.29	40.00
Atribuição	32.59	34.56	31.29	28.81	35.71
Coordenação	19.62	16.18	12.24	32.20	17.86
Outras	7.73	4.41	18.37	1.69	6.43

O que podemos filtrar, portanto, é que as propriedades consideradas, presumivelmente, universais – combinações categoriais e relações gramaticais – demonstram uma uniformidade capaz de ser abarcada, explicativamente, pelo componente sintático da gramática, local em que operações universais combinam unidades mínimas em objetos complexos. Assim sendo, relegamos o estabelecimento dessas duas propriedades ao componente sintático, o qual é o verdadeiro responsável pela formação dos compostos.

É nesse sentido que a formação de um composto é uniforme entre as línguas, pois a concatenação de categorias lexicais em determinada relação gramatical não varia translinguisticamente na geração de um composto. Já as propriedades específicas às línguas – suas estruturas morfológicas – serão tratadas pós-sintaticamente, no ramo de PF, local da gramática onde a variação superficial e fonológica entre as línguas é determinada.

Nesse sentido, a composição no PB pode ser descrita de um ponto de vista distribucional da seguinte maneira: há um conjunto de categorias lexicais disponível que é concatenado em determinadas relações gramaticais, a partir de operações universais presentes no componente sintático. Depois de concatenadas, essas combinações são emolduradas no componente morfológico da gramática, onde ocorre a inserção de conteúdo fonológico, ou seja, inserção dos IVs. Assim sendo, a distribuição geral atestada nos compostos do PB pode ser representada de modo esquemático no Quadro 3, abaixo:

Quadro 3 – Distribuição categorial, relacional e estrutural dos compostos do PB

Combinações Categoriais		
N-N, N-A, A-N, V-N, N-V, A-A, V-V, V-Adv, Adv-N, N-prep-N, N-Pron, Pron-N, Num-N		
Relações Gramaticais		
Subordinação	Atribuição	Coordenação
N-N, V-N, N-V	N-N, N-A, A-N, V-Adv, Adv-N, N-prep-N, N-Pron, Pron-N, Num-N	N-N, A-A, V-V
Estruturas Morfológicas		
[radical + radical]	[radical + palavra]	[palavra + palavra]
N-N	N-N	N-N
‘aracnofobia’	‘hidroginástica’	‘peixe-espada’
A-N	N-A	N-A
‘loirólatra’	‘cervejochato’	‘mesa-redonda’
N-V	A-N	A-N
‘inseticida’	‘ecoagência’	‘puro-sangue’
	N-V	V-N
	‘teletransportar’	‘lança-perfume’
	A-A	N-V
	‘lusó-brasileiro’	‘sangue-suga’
		A-A
		‘surdo-mudo’
		V-V
		‘vai-vém’
		V-Adv
		‘pisa-mansinho’
		Adv-N
		‘malcriação’
		N-prep-N
		‘pé-de-moleque’
		Pron-N
		‘seu-vizinho’
		N-Pron
		‘pai-nosso’
		Num-N
		‘trigêmeos’

Quando descrevermos o modo como os compostos são derivados na gramática, iremos nos concentrar na análise dos compostos V-N e N-N a fim de contemplar as relações gramaticais de subordinação e atribuição, principalmente. A seleção dos compostos N-N, por sua vez, permite-nos explorar a diversidade estrutural atestada nessa língua, já que essa única combinação categorial pode ser externalizada pelos três tipos estruturais do PB, tal como representado no Quadro 3.

3.5. Conclusões preliminares

Em resumo, apresentamos, neste capítulo, como o processo de composição se distribui em uma língua particular, o PB. Ao revisitar a literatura sobre a composição nessa língua e contrastá-la com as propriedades descritas sobre esse processo de formação de palavras translinguisticamente, pudemos notar que a única propriedade particular a esse processo reside no modo como os compostos são emoldurados em estruturas morfológicas, ao passo que a combinação entre seus membros constituintes é uniforme, e deve ser abarcada por um componente gerativo universal.

Com base na separação entre o geral e o específico desenvolvida neste capítulo, iremos, no capítulo seguinte, concentrar-nos em seus aspectos universais, a fim de elencar quais são as unidades mínimas a que esse componente gerativo, nomeadamente, a sintaxe, deve ter acesso a fim de gerar qualquer composto nas línguas naturais.

CAPÍTULO 4

UNIDADES MÍNIMAS DA COMPOSIÇÃO

Neste capítulo, discutimos a composição de modo amplo, com o intuito de fornecer as unidades mínimas para uma análise sintática da formação dos compostos das línguas naturais. Em sua elaboração, olhamos para o processo de composição de um ponto de vista translinguístico, com vias a filtrar as informações mínimas a que o componente sintático deve ter acesso a fim de gerar um composto.

4.1. Caracterização do capítulo

Da variedade de propostas presentes na literatura para explicar a composição, e dentre elas, as mais relevantes para a nossa discussão: (i) minimalistas; (ii) assimétricas, como a desenvolvida por Di Sciullo (2005b, 2009) e (iii) não-lexicalistas, mais especificamente as desenvolvidas dentro da MD (cf. FÁBREGAS, 2005; ZHANG, 2007; HARLEY, 2009; MINUSSI, 2011; NTHELITEOS, 2013⁵⁴)⁵⁵, não encontramos uma explicação abrangente, preocupada com a composição de modo amplo o suficiente para entender seus mecanismos básicos, salvo, em certa medida, a proposta em (ii).

Um exemplo desse comportamento voltado para detalhes específicos a uma única língua pode ser dado pelas análises propostas para compostos subordinados do tipo V-N (e.g. *lança-mísseis*, *lustra-móveis*, *estraga-prazer*), em que a discussão principal é o ponto da estrutura sintática a ser nominalizado, seja essa nominalização acima de VP, como em Di

⁵⁴ Agradeço à Artemis Alexiadou pela indicação e disponibilização do texto.

⁵⁵ Em modelos não-lexicalistas, como a MD, é comum encontrarmos um hibridismo entre os pressupostos da MD e as estruturas assumidas nas propostas assimétricas (e.g., BARBOSA; MINUSSI, 2009; MINUSSI, 2011; SCHER; NÓBREGA, 2014; NÓBREGA, no prelo). Por outro lado, uma posição dita minimalista assume as operações da sintaxe minimalista (i.e., *Merge*, *Move* e *Agree*) sem os princípios da MD, tais como raízes acategoriais e inserção tardia para a fonologia (e.g., BOK-BENNEMA; KAMPERS-MANHE, 2006; GRAČANIN-YUKSEK, 2006; DELFITTO; MELLONI, 2009; DELFITTO; FÁBREGAS; MELLONI, 2011). A proposta central desta dissertação, por sua vez, assume as operações da sintaxe minimalista juntamente com os princípios da MD, mostrando como as mesmas operações sintáticas para a geração de sentenças são capazes de gerar palavras.

Sciullo e Williams (1987) para os dados do italiano; de *vP*, como em Bok-Bennema e Kampers-Manhe (2006) para as línguas românicas; em *AspP*, como em Schroten (2010) para o inglês e o espanhol, ou em *CP*, como em Franco (2012), para o italiano. Análises como essas não se ampliam para os compostos de mesma combinação categorial no chinês (cf. ZHANG, 2007)⁵⁶, tampouco para os compostos neoclássicos formados por uma relação de subordinação, e.g. *cervejologia*, ‘estudo de cerveja’.

A proposta deste capítulo, por outro lado, visa reunir um conjunto de informações relacionais e estruturais mínimas, sem se fechar a um tipo específico de composição, garantindo, assim, a demarcação entre o que é puramente sintático na formação de um composto e o que deve ser resolvido pós-sintaticamente, ou seja, específico a uma língua. Com isso, buscamos fornecer um refinamento sobre as propriedades mínimas presentes no processo em questão. Vale salientar que esse posicionamento metodológico é derivado do modelo teórico assumido, a MD, pois admitimos que um sistema gerativo com operações universais, a saber, a sintaxe, é responsável pela formação de palavras.

O capítulo segue estruturado da seguinte maneira: na seção 4.2, discutimos a presença da composição nas línguas do mundo e os problemas de identificação e definição dos compostos. Em seguida, iniciamos nossa busca pelas unidades mínimas da composição, ao apresentar as propostas taxonômicas de Bisetto e Scalise (2005) e Scalise e Bisetto (2009). Em consonância com o Capítulo 3, endossamos, nessa seção, que as relações gramaticais internas à composição (i.e., subordinação, atribuição e coordenação) são universais, e devem ser consideradas a primeira unidade mínima devido à sua amplitude translinguística.

Na seção 4.3, apresentamos uma breve revisão das propostas lexicalistas para a composição, dando maior destaque à Teoria da Assimetria (cf. DI SCIULLO, 2005b, 2009), uma vez que seus propósitos se assemelham àqueles levados a cabo nesta dissertação. Serão descritas suas projeções funcionais internas à composição, as chamadas projeções-F (do inglês *f-tree*), e se apontará que os argumentos em favor dessas projeções são amplamente questionáveis.

Na seção 4.4, elencamos as diretrizes metodológicas empregadas na elaboração de uma análise em termos não-lexicalistas, e procedemos à busca da segunda unidade mínima da composição, a saber: a presença de um domínio categorial acima da concatenação de dois ou mais núcleos complexos. Será apresentado um conjunto de fenômenos morfosintáticos que

⁵⁶ Os compostos V-N do chinês podem ter como produto verbos, e.g. *zhi-dao* lit. saber-estrada ‘conhecer’ (cf. ZHANG, 2007, p. 173), algo não previsto nas análises mencionadas.

evidenciam a formação de um domínio nos compostos, em diferentes línguas, tal como a relevância de sua presença em uma visão sintática de formação de palavras.

Na seção 4.5, sintetizamos as unidades mínimas desse processo, para, então, no Capítulo 5, demonstrar como elas podem ser formalizadas na gramática.

4.2. A universalidade e os universais na composição: o problema da definição

A composição vem sendo descrita como um processo universal de formação de palavras desde Greenberg (1963), uma assunção que foi endossada por trabalhos posteriores sobre o assunto (e.g., FROMKIN et. al., 1996; LIBBEN, 2006; DRESSLER, 2006)⁵⁷:

“Não existem, provavelmente, línguas sem composição, afixação, ou sem ambas. Em outras palavras, não existe, provavelmente, nenhuma língua puramente isolante. Há um considerável número de línguas sem flexão, mas talvez nenhuma sem composição e derivação.” (GREENBERG, 1963, p. 92, tradução nossa)⁵⁸

Zamponi (2009, p. 593), por outro lado, ressalta a necessidade de cautela em se assumir que a composição é um processo universal, pois embora seja abundante em algumas línguas (e.g., no kayardild, no nahuatl clássico, no mandarim e em muitas outras línguas isolantes), há línguas em que palavras compostas não são frequentes (e.g., no yuroc, no turkana, no menya), e outras em que são particularmente raras (e.g., no evenki, e nas línguas arauaque maipure-yavitero)⁵⁹.

Bauer (2009, p. 344), em contrapartida, argumenta que a raridade de um fenômeno em uma língua não deve contrariar a afirmação de sua universalidade, mas o que pode contrariá-la é somente a ausência desse fenômeno. Um dos problemas que a questão da universalidade da composição enfrenta – e não apenas essa questão isolada, mas a teoria morfológica como

⁵⁷ De um ponto de vista evolutivo, a composição é apresentada como um fenômeno protogramatical, ou seja, relíquias de estágios anteriores da capacidade linguística (cf. JACKENDOFF, 2009), hipótese que, segundo Jackendoff (2009), encontra evidências na produtividade de compostos em *pidgins* (cf. PLAG, 2006) e em variedades básicas de língua, como, por exemplo, nos estágios de aquisição de segunda língua por alunos estrangeiros (cf. KLEIN; PERDUE, 1997).

⁵⁸ Trecho original (GREENBERG, 1963, p. 92): “There are probably no languages without either compounding, affixing, or both. In other words, there are probably no purely isolating languages. There are a considerable number of languages without inflection, perhaps none without compounding and derivation.”

⁵⁹ Štekauer, Valera e Körtvélyessy (2008 *apud* SCALISE; VOGEL, 2010b, p. 1) também questionam a validade da composição como um processo universal de formação de palavras. A partir da análise de um corpus contendo 55 línguas, os autores mostram que apenas 50 delas apresentam compostos. Dessas línguas, as que não apresentam esse processo de formação de palavras são: ânglia oriental (do inglês, *East Danglia*), karao, groelandês ocidental, diola fogny e kwak’wala.

um todo – é a definição do que é um composto, já que há muitas descrições de línguas em que não fica claro se a língua em questão contém ou não compostos por razões de imprecisão na definição desse processo.

Bauer (2009) comenta o caso do ainu, língua isolante do Japão, sobre a qual os principais trabalhos descritivos não atestam a composição como um processo de formação de palavras (viz., REFSING, 1986; SHIBATANI, 1990 *apud* BAUER, 2009). No entanto, há uma grande variedade de dados nesses trabalhos que sugere a presença de compostos:

(55) a. <i>atuy asam</i>	lit. fundo mar	‘fundo do mar’	(Ainu)
b. <i>mosem apa</i>	lit. pólo entrada	‘entrada do pólo’	
c. <i>kamuy napuri</i>	lit. montanha deus	‘montanha sagrada’	
d. <i>supuya kur</i>	lit. vestígio fumaça	‘vestígios de fumaça’	

(REFSING, 1986, p. 160 *apud* BAUER, 2009, p. 344)

A partir desses dados, devemos concluir que o ainu tem ou não compostos? A resposta para essa questão, como pontua Bauer (*op. cit.*), depende em parte da definição do que é um composto, e em parte da análise proposta para os dados. Nesse sentido, se buscarmos auxílio nas principais definições sugeridas para identificar o processo, perceberemos que não há uma definição satisfatória, e tampouco uniforme, algo que pode ser verificado nas definições elencadas por Scalise e Vogel (2010b, p. 5), em (56):

- (56) a. Uma unidade lexical construída a partir de dois ou mais **elementos**, cada um dos quais funcionando como um **lexema** independente do(s) outro(s) em outros contextos, e que apresenta algum isolamento gramatical e/ou fonológico do seu uso sintático normal. (BAUER, 2001, p. 695)
- b. [...] uma palavra composta contém pelo menos duas bases que são **palavras**, ou pelo menos, **raízes**. (KATAMBA, 1993, p. 54)
- c. Um lexema complexo que pode ser pensado como consistindo de dois ou mais **lexemas**. (HASPELMATH; SIMS, 2002, p. 85)
- d. Sua propriedade definidora é que ele consiste da combinação de dois **lexemas** em palavras maiores. Nos casos simples, a composição consiste da combinação de **duas palavras**, na qual uma palavra modifica o significado da outra, o núcleo. (BOOIJ, 2005, p. 75)
- e. Composição [...] denota a combinação de **formas livres** ou **radicais** para formar um

novo complexo denominado composto. (OLSEN, 2000, p. 280)

f. [...] compostos formados por raízes consistem de dois **radicais** combinados em um, com o composto como um todo carregando a categoria e os traços morfossintáticos do radical à direita. (LIEBER, 2004, p. 47)

g. Quando duas ou mais **palavras** são combinadas em uma unidade morfológica, estamos falando de um composto. (MARCHAND, 1960, p. 11)⁶⁰

Nessas definições, não encontramos um consenso sobre qual é a unidade básica para a formação de um composto, ou seja, qual o estatuto morfológico basilar de seus elementos constituintes – se raízes, radicais, palavras ou lexemas. Essa incongruência nas definições é o que ocasiona a negligência do processo em algumas línguas, visto que, por exemplo, há estatutos morfológicos que variam translinguisticamente (e.g., no inglês, raízes são tipicamente formas livres, ao passo que, nas línguas eslavas, raízes são formas presas)⁶¹.

Lieber e Štekauer (2009b) discutem amplamente esse problema, mostrando que o que temos na literatura são somente visões particulares e restritas para definir a composição. Os autores revelam como muitas dessas definições desconsideram uma gama de dados, pois elas foram propostas com base em línguas em que o tipo de composição prototípico é [palavra + palavra]. Isso faz com que compostos formados por radicais sejam, muitas vezes, desconsiderados; por exemplo, os compostos do grego moderno e das línguas eslavas, exemplificados em (57), abaixo:

(57) a.	<i>gwiazd-o-zbiór</i>	estrela-EL-coleção	‘constelação’	(Polonês)
b.	<i>rychl-o-vlak</i>	rápido-EL-trem	‘trem expresso’	(Esloveno)
c.	<i>ayri-o-ánthropos</i>	selva-EL-homem	‘homem selvagem’	(Grego)

⁶⁰ Trecho original (SCALISE; VOGEL, 2010b, p. 5): “a. A lexical unit made up of two or more **elements**, each of which can function as a **lexeme** independent of the other(s) in other context, and which shows some phonological and/or grammatical isolation from normal usage (Bauer 2001:695); b. [...] a compound word contains at least two bases which are both **words**, or at any rate, **root morphemes** (Katamba 1993:54); c. A complex lexeme that can be thought of as consisting of two or more **lexemes** (Haspelmath 2002:85); d. Its defining property is that it consists of the combination of **lexemes** into larger words. In simple cases, compounding consists of the combination of two **words**, in which one word modifies the meaning of the other, the head (Booij 2005:75); e. Composition [...] denotes the combining of two **free forms** and **stems** to form a new complex word referred to as compound (Olsen 2000:280); f. [...] root compounds consist of two **stems** combined as one, with the compound as a whole bearing the category and morphosyntactic features of the right-hand stem (Lieber 2004:47); g. When two or more **words** are combined into a morphological unit, we speak of a compound (Marchand 1960:11).

⁶¹ Tal como para as línguas românicas; cf. Capítulo 3, para uma discussão sobre as definições de raiz, radical e palavra.

Nesse debate sobre o que vem a ser um composto nas línguas naturais, encontramos trabalhos que procuram sugerir critérios em vez de propor uma definição direta. Donalies (2004, p.76 *apud* LIEBER; ŠTEKAUER, 2009b, p. 6-7), por exemplo, elenca uma lista de critérios para identificar um composto nas línguas germânicas, românicas, eslavas, fino-úgricas⁶² e no grego moderno, os quais estão elencados em (58):

(58) *Critérios para identificação de compostos (DONALIES, 2004, tradução nossa)*

- a. são complexos;
- b. são formados sem afixação;
- c. são pronunciados juntos;
- d. têm um padrão acentual específico;
- e. contêm elementos de ligação;
- f. têm núcleo à direita;
- g. são flexionados como um todo;
- h. são sintaticamente inseparáveis;
- i. são ilhas sintático-semânticas;
- j. são unidades conceituais.⁶³

Obviamente, muitos desses critérios não se aplicam em todas as línguas em questão. Por exemplo, o critério da posição do núcleo é questionável, já que a posição do núcleo parece ser uma informação específica às línguas, e, apesar disso, dentro de uma única língua, ela pode não ser uniforme (e.g., no PB, o composto *sangue frio* tem núcleo semântico à esquerda, enquanto que o composto *vaso constrição* tem núcleo semântico à direita, cf. MORENO, 1997). Além do mais, em se tratando de nuclearidade na composição, é importante explicitar a que núcleo o autor se refere, pois como sugerem Scalise, Fábregas e Forza (2009), o núcleo do composto pode ser de três tipos: categorial (i.e., o constituinte que fornece a categoria do composto), morfológico (i.e., o constituinte que desencadeia concordância) e semântico (i.e., o constituinte que é modificado).

Outro ponto é a presença de elementos de ligação na composição, que, apesar de ser comum em compostos formados por radicais, como vimos no Capítulo 3, não é estritamente

⁶² Sub-família das línguas úgricas, por exemplo, o húngaro, o finlandês e o estoniano.

⁶³ Trecho original (DONALIES, 2004, p. 76 *apud* LIEBER; ŠTEKAUER, 2009, p. 6-7): “Compounds: are complex; are formed without word-formation affixes; are spelled together; have a specific stress pattern; include linking elements; are right-headed; are inflected as a whole; are syntactically inseparable; are syntactico-semantic islands; are conceptual units.”

necessária para garantir a gramaticalidade do composto em algumas línguas, já que compostos formados por palavras, e mesmo compostos formados por radicais, podem ser construídos sem a presença de um elemento de ligação (cf. ŠTEKAUER; VALERA, 2007 *apud* LIEBER; ŠTEKAUER, 2009b, p. 14).

Lieber e Štekauer (2009b, p. 7) comentam, ainda, que o critério (58a), o qual define um composto como complexo, não é capaz de diferenciá-lo de palavras derivadas, as quais também são complexas. Portanto, o que podemos concluir desses critérios é que eles podem ser úteis para a identificação de compostos em uma língua ou grupo de línguas, mas não podem ser assumidos translinguisticamente.

Isso fica mais claro quando olhamos para os critérios fonológicos empregados por Donalies (2004). O critério do padrão acentual não pode ser considerado uma propriedade uniforme, pois uma mesma língua pode conter diferentes distribuições acentuais, a depender do tipo de estrutura morfológica do composto, como vimos com os compostos do PB, no Capítulo 3. Nespor e Ralli (1996) e Nespor (1999), por exemplo, mostram que compostos formados por radicais apresentam um único acento primário, ao passo que compostos formados por palavras apresentam, geralmente, dois acentos primários⁶⁴.

Além disso, o número de acentos primários e o padrão acentual não formam um critério universal. Bauer (2009, p. 345) mostra que, enquanto em línguas como o inglês, o acento pode ser um critério indicativo de composição, em línguas como o chukchi é a harmonia vocálica que desempenha esse papel. Em dêw, compostos são identificados pelo número de sílabas e pela presença de um único tom, já que, nessa língua, enquanto palavras simples são monossilábicas, compostos são dissilábicos (cf. MARTINS, 2004, p. 127). No japonês, o processo fonológico *rendaku*, ou seja, vozeamento de consoantes obstruintes no segundo membro de compostos nativos, é o que marca a composição (e.g., *ta* ‘campo’ → *ina-da* lit. arroz-campo ‘arrozal’).

Uma variedade de trabalhos descritivos mostram outros processos fonológicos responsáveis por distinguir compostos de sintagmas nas línguas naturais. Além do acento, do padrão tonal e do vozeamento, outros efeitos segmentais como vozeamento de fricativas no eslavo (língua atabasca, cf. RICE, 2009), redução de movimentos internos ou repetidos na língua de sinais norte-americana (cf. SANDLER; LILLO-MARTIN, 2006), apagamento de vogais no hebraico (cf. BORER, 2009a) e redução de vogais nas línguas arauaque (cf.

⁶⁴ Nas línguas românicas, por exemplo, compostos formados por palavras têm o mesmo padrão acentual que sintagmas (e.g., compostos V-N, compostos N-A, compostos N-prep-N, dentre outros), o que dificulta uma distinção entre compostos e sintagmas a partir de critérios fonológicos nessas línguas.

ZAMPONI, 2009), levam-nos à conclusão de que a composição não pode ser definida translinguisticamente via processos fonológicos.

Portanto, se buscamos fornecer os ingredientes básicos para a geração de um composto, é possível dizer até aqui que a informação fonológica não constitui um elemento mínimo do ponto de vista universal, caso levemos em consideração os levantamentos feitos por trabalhos tipológicos. Dessa forma, a presença ou não de processos fonológicos uniformes para a identificação de um composto deve ser resolvida por regras fonológicas que se aplicam a uma língua específica, mas, esses processos por si só não contribuem como uma unidade mínima para o entendimento universal da formação dos compostos.

Tendo em vista essa variedade de padrões fonológicos, reiteramos a proposta de que os compostos são derivados uniformemente na gramática, e a variedade superficial supramencionada é resultado de processos pós-sintáticos relativos à externalização. Uma vez que nos baseamos em uma proposta de inserção tardia para a fonologia, noções como raiz, radical e palavra estão relacionadas à boa formação fonológica de raízes sintáticas, como vimos no Capítulo 3. Nesse sentido, compostos formados por radicais, como aqueles em (57), serão gerados da mesma maneira que compostos formados por palavras, e suas diferenças superficiais – radicais ou palavras – não estão no modo como seus membros são concatenados, mas sim no modo como são externalizados.

Por fim, retomando a questão levantada no início da seção, sobre a universalidade da composição, acreditamos que uma resposta só será alcançada quando obtivermos uma noção mais clara de quais são as unidades mínimas para efetuar esse processo de formação de palavras, deixando claro, por um lado, o que é parte intrínseca da formação de um composto e, por outro, o que é específico a uma língua para sua externalização (e.g., estruturas morfológicas, propriedades fonológicas, posição superficial do núcleo, etc.).

Passaremos, agora, para a discussão das propostas de classificação dos compostos, filtrando o que delas pode contribuir para o levantamento dessas unidades mínimas, com o intuito de fornecer, ao longo do capítulo, uma lista de informações necessárias à gramática para a derivação sintática de um composto.

4.2.1. A primeira unidade mínima: as relações gramaticais internas à composição

Na busca pelas características que podem ser vistas como universais – ou mínimas – para a composição, analisamos as propostas taxonômicas de Bisetto e Scalise (2005) e Scalise e Bisetto (2009), as quais foram elaboradas a partir da análise de um grande volume de dados

translinguísticos⁶⁵. Fixamos nossa atenção nos critérios utilizados pelos autores, com o intuito de filtrar quais deles são passíveis de uma aplicação universal.

Numerosas tentativas foram feitas para se desenvolver um esquema classificatório para os compostos (cf. BLOOMFIELD, 1933; MARCHAND, 1969; SPENCER, 1991; FABB, 2001; OLSEN, 2001; BOOIJ, 2005), porém esses esquemas mostram-se restritos quando são aplicados translinguisticamente. Na tentativa de fornecer uma classificação com ampla cobertura empírica, Bisetto e Scalise (2005) e Scalise e Bisetto (2009) identificam os problemas que circundam essas propostas e que impedem uma visão consistente e universal acerca do fenômeno.

Um dos principais problemas é o terminológico, que diz respeito à influência de trabalhos descritivos de línguas anglo-saxãs, principalmente do inglês, na descrição dos compostos das demais línguas. A influência desses trabalhos fez com que diversos autores acabassem forçando os compostos de suas línguas a entrar em dois tipos comuns de compostos anglo-saxões, a saber, compostos formados por raízes (do inglês, *root compounds*)⁶⁶ e compostos sintéticos (do inglês, *synthetic compounds*)⁶⁷.

Além disso, os autores notaram que havia uma sobreposição de critérios nas classificações propostas, dado que alguns teóricos equiparavam, em um mesmo nível classificatório, as relações gramaticais (i.e., subordinação, atribuição e coordenação) e a nuclearidade semântica, a qual é dividida em endocêntrica (i.e., quando o núcleo semântico se encontra interno ao composto; e.g., *peixe espada*) e exocêntrica (i.e., quando o núcleo semântico é externo ao composto; e.g., *(indivíduo) estraga-prazer*).

Por exemplo, nas classificações de Spencer (1991), Haspelmath (2002), Bauer (2001) e Booij (2005), os compostos são divididos nos seguintes conjuntos: (i) compostos endocêntricos, (ii) compostos exocêntricos e (iii) compostos coordenados. De acordo com essa divisão, a composição é vista, basicamente, pela presença ou ausência de núcleo, e a nuclearidade detém a mesma natureza que as relações gramaticais.

Contrariamente a essa divisão dos critérios, Bisetto e Scalise (2005) argumentam que as classificações estão equivocadas, pois todas as relações gramaticais podem apresentar

⁶⁵ Em seus trabalhos, os autores fizeram uso de uma base de dados constituída, exclusivamente, por compostos, pertencente ao Projeto Morbo/Comp (*Morphology at Bologna Research Group/Compounding*, <http://componet.sslmit.unibo.it/index.php?section=home>).

⁶⁶ *Root compounds* são formados pela junção de duas raízes, por exemplo, *sea food* lit. mar-comida ‘frutos do mar’. Esse tipo de composição não pode ser, diretamente, estendido para as línguas românicas, por exemplo, pois nessas línguas os nomes, verbos e adjetivos carregam necessariamente uma vogal temática que não pertence à raiz (cf. BISETTO, 2003 *apud* SCALISE; BISETTO, 2009).

⁶⁷ Compostos sintéticos apresentam uma relação argumental verbo-complemento, acrescida de um elemento sufixal (e.g., do inglês, *truck-driv-er* lit. caminhão-dirigir-NOM ‘motorista de caminhão’).

compostos tanto endocêntricos quanto exocêntricos, ou seja, a nuclearidade deve estar contida em cada uma das relações gramaticais. Com isso, a classificação reformulada pelos autores leva em consideração, primeiro, as relações gramaticais entre os constituintes de um composto, definidas em (59) e, segundo, as relações de nuclearidade presentes em cada uma delas, as quais estão definidas em (60). A classificação final fica esquematizada como na representação em (61), exemplificada com os dados no Quadro 4, abaixo.

(59) *Relações gramaticais internas aos compostos*

- a. **Subordinação:** relação núcleo-complemento;
- b. **Atribuição:** relação núcleo-modificador;
- c. **Coordenação:** relação conjuntiva ou disjuntiva (conectados por uma conjunção “e” ou “ou”, respectivamente).

(60) *Nuclearidade semântica*

- a. **Compostos endocêntricos:** o núcleo semântico, ou seja, o elemento modificado, é interno ao composto (e.g., peixe espada)
- b. **Compostos exocêntricos:** o núcleo semântico é externo ao composto (e.g., [indivíduo] casca grossa).

(61) *Classificação dos compostos (BISETTO; SCALISE 2005)*

Quadro 4 - Exemplos de compostos na classificação de Bisetto e Scalise (2005)

SUBORDINADOS		ATRIBUTIVOS		COORDENADOS	
Endocêntricos	Exocêntricos	Endocêntricos	Exocêntricos	Endocêntricos	Exocêntricos
<i>love story</i>	<i>killjoy</i>	<i>blackboard</i>	<i>paleface</i>	<i>bitter sweet</i>	<i>north-east</i>
amor-história	mata-alegria	preto-quadro	pálido-rostro	amargo-doce	norte-leste
‘romance’	‘estraga-prazer’	‘quadro negro’	‘cara pálida’	‘agridoce’	‘nordeste’

Em (2009), os autores inserem um nível intermediário na classificação de (2005), com o intuito de especificar as relações semânticas e interpretativas encontradas em cada uma das relações gramaticais. Nesse nível intermediário, há a intenção de tornar explícitas quais são as combinações categoriais presentes no estabelecimento da relação gramatical e a natureza categorial do constituinte que determina essa relação.

Uma consequência da inserção desse nível intermediário foi a subdivisão da relação de subordinação entre *verbal-nexus* e *ground*, já que ela pode ocorrer tanto em compostos que contêm um núcleo verbal (e.g., compostos V-N, como *limpa-vidros*, e compostos sintéticos, como *truck-driver*) quanto em compostos sem um núcleo verbal (e.g., compostos nominais com marcação genitiva aberta, como os do islandês, e.g., *laekni-s tösk-ur* lit. médico-GEN.SG mala-NOM.PL ‘bolsas do médicos’⁶⁸, e compostos que implicam uma leitura de “relação de”, e.g. *apron string* lit. avental-fio ‘corda do avental’⁶⁹), e essa distinção passa, então, a ser parte da classificação.

Da mesma forma, os compostos atributivos foram subdivididos em compostos apositivos e atributivos. Os compostos apositivos são aqueles em que o elemento não-núcleo é um nome que expressa uma propriedade do núcleo (e.g., *sword fish* lit. espada_N-peixe_N ‘peixe espada’), e compostos atributivos são aqueles em que a relação de modificação é dada por um adjetivo ou um verbo que age como atributo (e.g., *high school* lit. alta_A-escola_N ‘ensino médio’).

A inserção desse nível intermediário, portanto, deixa a classificação reformulada como em (62):

(62) *Classificação dos compostos (SCALISE; BISETTO, 2009, p. 50)*

⁶⁸ Retirado de Harðarson (2013).

⁶⁹ Essa “relação de” advém da Variável-R proposta por Allen (1978), que distingue a combinação entre nomes nos compostos em duas variedades: (a) Relação É UM (*IS A Relation*) e (b) Relação-R. Na primeira, o produto da composição pode ser interpretado como um tipo de um dos membros do composto (e.g., *sword fish* lit. espada-peixe ‘peixe espada’ É UM tipo de peixe), a qual é definida por Scalise e Bisetto (2009) como uma relação de atribuição. Na segunda relação entre nomes, há uma subordinação implícita, a qual pode ser parafraseada inserindo uma preposição “de” (e.g., *wind mill* lit. vento-moinho ‘moinho de vento’), classificada por Scalise e Bisetto (*op. cit.*) como uma relação de subordinação, sob o rótulo de *ground*.

Para os nossos propósitos, o que parece ser, indiscutivelmente, bem aplicável translinguisticamente são as relações gramaticais, pois elas independem da estrutura morfológica do composto para se estabelecer, tal como vimos no Capítulo 3, além do fato de que elas estão superordenadas com relação aos demais níveis. Acreditamos, com isso, que uma explicação sobre como essas relações gramaticais se estabelecem na gramática, inclua, por conseguinte, uma explicação sobre as especificações feitas pelo nível intermediário – das combinações categoriais –, bem como uma explicação sobre as relações de nuclearidade semântica presente nos compostos.

A assunção de que as relações gramaticais são translinguisticamente atestadas pode ser endossada não apenas pelos dados tratados por Bisetto e Scalise (2005) e Scalise e Bisetto (2009), mas também pelos trabalhos que fizeram uso de sua classificação para a descrição dos compostos, por exemplo: Ceccagno e Basciano (2007), para o chinês; Rosenberg (2007), no francês; Benigni e Masini (2009), no russo; Göskel (2009), no turco; Štichauer (2009), no tcheco; Rio-Torto e Ribeiro (2012), no português europeu; Vercellotti e Mortensen (2012), na língua de sinais norte-americana (ASL), dentre outros.

O nível intermediário, por sua vez, também carrega um caráter universal, já que as combinações categoriais nas quais as relações gramaticais são expressas não variam translinguisticamente (e.g., em um composto V-N, em que N é um objeto direto, a relação entre seus membros deverá ser sempre de subordinação). O que pode variar, nesse sentido, são quais combinações categoriais estão presentes em uma língua particular, e quais relações gramaticais uma mesma combinação categorial pode abrigar (e.g., a combinação N-N que pode formar compostos subordinados, atributivos e coordenados).

Particularmente, ao sugerirmos um tratamento formal para essas generalizações, a diferença entre um composto *ground* (i.e., composto N-N formado por uma relação de subordinação) e um composto apositivo (i.e., composto N-N formado por uma relação de atribuição) será reduzida ao tipo de operação *Merge* que venha a ocorrer entre seus membros. Dessa forma, o que irá distinguir o conjunto dos compostos subordinados do conjunto dos compostos atributivos será a maneira como o sistema computacional irá concatenar seus membros constituintes.

No terceiro nível da classificação, percebemos que as relações de nuclearidade parecem se comportar bem translinguisticamente nos dados de compostos subordinados e atributivos, já que a relação interna entre seus constituintes implica uma hierarquia. No entanto, em compostos coordenados, não é trivial a identificação do núcleo, tal como vimos

com os compostos N-N, no Capítulo 3, pois ambos os constituintes parecem estar em uma configuração simétrica.

A questão da identificação do núcleo em compostos coordenados sugere a existência de um problema mais profundo, que diz respeito à coordenação de modo amplo, a qual não apresenta um consenso na literatura quando o assunto é a nuclearidade⁷⁰. Como mencionado no Capítulo 1, não iremos nos ater aos compostos coordenados, porém, discutiremos as questões referentes à nuclearidade quando apresentarmos nossa sugestão de análise dentro da MD, no Capítulo 5 desta dissertação.

Por fim, o que podemos reter dessa seção é que as relações gramaticais têm se mostrado como uma unidade mínima na composição de um ponto de vista translinguístico. Com base nessa informação, Scalise e Bisetto (2008) e, posteriormente, Guevara e Scalise (2009) sintetizam essa generalização sugerindo que há um protótipo de palavras compostas nas línguas do mundo, cuja essência pode ser capturada pelo esquema apresentado em (63), abaixo:

(63) [X $\Re$ Y]Z

“em que X,Y e Z representam as principais categorias lexicais, e $\Re$ representa uma relação implícita entre os constituintes (uma relação que não é realizada por um item lexical).”⁷¹

Essa generalização, no entanto, embora sintetize nossa primeira unidade mínima, representada como $\Re$, em (63), não é suficiente para a identificação de um composto, já que tais relações gramaticais são as mesmas presentes entre constituintes sintagmáticos (cf. GUEVARA; SCALISE, 2009, p. 107). Quando passarmos a discutir a composição na MD, mais adiante, mostraremos que a composição exige a formação de um domínio categorial, e esse domínio será a segunda unidade mínima necessária para formação de um composto⁷².

⁷⁰ A nuclearidade na coordenação gira em torno de análises que propõem a presença de dois núcleos ou a ausência de núcleos. Para o leitor interessado na questão da nuclearidade em compostos coordenados, sugerimos os trabalhos de Villalva (1995), Olsen (2000, 2001), Ceccagno e Basciano (2007), Guevara e Scalise (2009) e Bisetto (2010).

⁷¹ Trecho original (GUEVARA; SCALISE, 2009, p. 107): “where X, Y and Z represent the major lexical categories, and $\Re$ represents an implicit relationship between the constituents (a relationship not spelled out by any lexical item).”

⁷² Vale mencionar que Guevara e Scalise (2009) colocam como um problema a maneira como $\Re$ se estabelece na gramática, algo para o qual iremos propor uma sugestão de formalização na seção 5.2, do Capítulo 5.

4.3. Abordagens lexicalistas sobre a composição

A composição não figurou como um processo de formação de palavras com visibilidade na maioria dos trabalhos lexicalistas, embora tenha recebido considerável atenção durante o período pré-lexicalista⁷³, consequência da proposta transformacional de Lees (1960)⁷⁴. Os principais trabalhos que se sucederam (e.g., LEVI, 1978, ALLEN, 1978, dentre outros) reviam a proposta de Lees (*op. cit.*) e tentavam torná-la aceitável, apesar de pouco plausível empiricamente, como aponta ten Hacken (2009).

Nos anos 70, os primeiros trabalhos sobre formação de palavras dentro de uma abordagem lexicalista, nomeadamente, Halle (1973) e Aronoff (1976), retiraram a composição do conjunto de dados a ser tratado. É somente nos anos 80 e 90 que propostas sobre esse processo de formação de palavras começaram a ser elaboradas dentro do empreendimento gerativista. Dentre as questões mais pungentes, a principal era a de mapear em qual componente da gramática os compostos eram gerados, se no componente lexical ou no componente sintático.

Por exemplo, Selkirk (1982) fazia uma divisão entre compostos produtivos e compostos pouco produtivos ou idiossincráticos, em que os primeiros eram gerados pela sintaxe, e os últimos, listados no léxico. Dentro da Fonologia Lexical, Shibatani e Kageyama (1988) dividem os compostos em lexicais e pós-lexicais, de acordo com os processos morfofonológicos e morfossintáticos internos à composição, tal como faz Lee (1995, 1997), para os compostos do PB.

Dentre os dados mais explorados, encontramos uma preferência por compostos que possibilitavam um diálogo com os trabalhos desenvolvidos na sintaxe (e.g., compostos sintagmáticos e compostos sintéticos)⁷⁵, e, ainda que muitos autores fizessem uma sintaxe interna à palavra dentro do componente lexical (cf. SELKIRK, 1982; LIEBER, 1983; SCALISE, 1984; DI SCIULLO; WILLIAMS, 1987; LIEBER, 1992, dentre outros), empregando noções de nuclearidade, endocentricidade, estruturas hierárquicas, etc., um

⁷³ O uso do termo pré-lexicalista diz respeito aos trabalhos dentro da abordagem transformacional, anteriores a Chomsky (1970) e às propostas de um componente lexical na arquitetura da gramática.

⁷⁴ Lees (1960), dentro de uma abordagem transformacional, propunha que os compostos e as nominalizações eram resultados de transformações aplicadas sobre sentenças, e, consequentemente, eram interpretados como versões nominais dessas sentenças.

⁷⁵ Essa é uma das principais críticas feitas por Bisetto e Scalise (2005) e Scalise e Bisetto (2009), pois muitas combinações categoriais na composição foram negligenciadas nesses trabalhos.

conjunto de compostos sintagmáticos (e.g., V-N, N-prep-N, N-A) era descrito como gerado pelo componente sintático e enviado ao léxico para ser reutilizado como um nó terminal⁷⁶.

Por outro lado, o domínio empírico descrito foi pouco expandido nesse período, e o que era conhecido sobre a composição restringia-se, de modo considerável, aos trabalhos em línguas germânicas e românicas, as quais ainda ocupam maior destaque nas discussões teóricas sobre o assunto (cf. SCALISE, 1992). Recentemente, a composição vem ganhando espaço e diversos trabalhos descritivos em diferentes línguas estão possibilitando uma compreensão mais ampla sobre o assunto (e.g., BISETTO, 2009; GAETA; GROSSMANN, 2009; LIEBER; ŠTEKAUER, 2009a; SCALISE; VOGEL, 2010a; SCALISE; MASINI, 2012). Esse avanço descritivo permite que as teorias gramaticais consigam visualizar o que há de mais geral nesse processo de formação de palavras, algo que pode ser verificado nos trabalhos elencados e discutidos até este ponto da dissertação.

Dentre as propostas lexicalistas atuais que se propõem a fornecer uma explicação sobre a composição como parte da linguagem humana, tal como é o objetivo geral desta dissertação, temos a Teoria da Assimetria, proposta por Di Sciullo (2005a, 2005b, 2009), a qual será discutida em detalhes na próxima subseção.

4.3.1. Teoria da assimetria e as projeções funcionais tipificadas para a composição

Di Sciullo (2005a) desenvolve uma teoria da gramática em que as relações morfológicas, diferentemente das relações sintáticas, são estritamente assimétricas⁷⁷. A autora, assumindo os pressupostos do Programa Minimalista⁷⁸ (cf. CHOMSKY, 1995, doravante PM), elabora um modelo de gramática, a Teoria da Assimetria (doravante, TA), em que as derivações morfológicas (D_M) ocorrem paralelamente às derivações sintáticas (D_S) no espaço computacional.

⁷⁶ Di Sciullo e Williams (1987) formalizam o envio do composto gerado na sintaxe para o léxico através da seguinte regra:

(i) *Regra de criação de palavras não-morfológicas (Nonmorphological Word Creating Rule)*
 $Y \rightarrow XP$, em que Y representa a categoria lexical e XP as categorias lexicais máximas (VP, NP, PP, AP).

Essa regra é amplamente assumida nos trabalhos em que há compostos sendo gerados pela sintaxe.

⁷⁷ A noção de assimetria aludida pela autora diz respeito à irreversibilidade entre dois elementos de um conjunto. Uma relação, ou seja, um conjunto de pares ordenados em um dado domínio, é assimétrica se o conjunto não contém pares ordenados que possam ser invertidos, por exemplo, $\langle x, y \rangle$ e $\langle y, x \rangle$, contrário a uma relação simétrica que permite tal ordenamento (cf. DI SCIULLO, 2005a, p. 16-20).

⁷⁸ Segundo a autora, o modelo desenvolvido é uma extensão do modelo de derivação-por-fases de Chomsky (2001, 2004).

Em sua proposta, o espaço computacional está dividido em planos paralelos para as computações, como pode ser visto na arquitetura da gramática abaixo:

(64) *Arquitetura da gramática na Teoria da Assimetria (DI SCIULLO, 2005a, p. 22)*

Nessa visão de gramática, compostos são domínios da computação e sua estrutura interna contém uma projeção funcional, a qual deve ser legível na interface semântica (LF) e pode ou não ser legível na interface fonológica (PF). Essa projeção funcional, a que chamaremos projeção-F (do inglês, *f-tree*), garante a assimetria interna entre os membros do composto, uma vez que os compostos são vistos como um objeto morfológico. Nas palavras da autora, a hipótese da projeção-F é descrita como se segue:

(65) *Hipótese da projeção-F (DI SCIULLO, 2005b, 2009, tradução nossa)*

Todos os compostos incluem uma projeção-F⁸⁰.

(66) *Projeção-F*

O primeiro argumento utilizado para se assumir a projeção-F interna a um composto é a presença de uma relação de modificação entre os constituintes de compostos formados por raízes (i.e., *root compounds*), como aqueles listados em (67). Segundo a autora, a relação de modificação é comumente mapeada em uma relação funcional. Dessa forma, o primeiro constituinte do composto, seja ele um nome ou um adjetivo, ocupa a posição de especificador

⁷⁹ Φ e Σ estão para as derivações fonológicas e semânticas, respectivamente.

⁸⁰ Trecho original (DI SCIULLO, 2009, p. 151): “F-Tree Hypothesis: All compounds include an F-tree”.

da projeção-F, enquanto o segundo constituinte ocupa a posição de complemento da projeção-F, como em (68).

(67) a. <i>floppy disk</i>	lit.	flexível-disco	‘disquete’
b. <i>kitchen towel</i>		cozinha-toalha	‘toalha de cozinha’
c. <i>lily white</i>		lírio-branco	‘branco lírio’

No entanto, a exigência da presença de uma projeção funcional para estabelecer uma relação de modificação não é estritamente necessária em termos estruturais. Se assumirmos *Bare Phrase Structure* (doravante, BPS; cf. CHOMSKY, 1995), por exemplo, a relação de modificação ocorre a partir de um tipo de operação *Merge*, a saber, *pair-Merge*, que se aplica a dois constituintes, gerando uma relação assimétrica sem a assunção de uma projeção funcional entre dois elementos⁸¹.

O segundo argumento deriva das condições de interpretabilidade dos compostos nas interfaces. Para tanto, a autora propõe que o núcleo da projeção-F pode apresentar sabores relacionais distintos, motivados pela presença de conectivos como conjunções, disjunções e preposições internas ao composto. Esses sabores relacionais estão exemplificados nos dados em (69), com a conjunção *and* ‘e’ e a disjunção *or* ‘ou’, e nos dados em (70a), com a preposição implícita *with* ‘com’, em (70b), com a preposição implícita *to* ‘para’, e em (70c) com a preposição implícita *in* ‘em’. Dessa forma, o conectivo pode ou não ser legível na interface fonológica, mas precisa estar presente na estrutura a fim de garantir a legibilidade semântica do composto⁸².

⁸¹ Cf. Capítulo 5, seção 5.2.

⁸² No trabalho de (2005), Di Sciullo assumiu três núcleos funcionais distintos: SORT ‘tipo de’, AND ‘e’ e OR ‘ou’. Em (2009), SORT parece ter sido incluído na relação de modificação, como na estrutura em (68), e os núcleos funcionais tipificados passaram a abrigar apenas as relações de coordenação (i.e., AND e OR) e as relações *location* e *locatum* (i.e., IN ‘em’ e WITH ‘com’, respectivamente), estendendo a terminologia de Hale e Keyser (2002).

- (69) a. *bed-and-breakfast*, *hit-and-run*, *truth-or-dare*
 ‘cama e café-da-manhã’ ‘bater e correr’ ‘verdade ou desafio’
 b. **bed-breakfast*, **hit-run*, **truth-dare*

- (70) a. *martini soda*, *vodka soda*
 ‘soda com martini’, ‘soda com vodka’
 b. *Montreal-Boston train*, *New York-Dubai flight*
 ‘trem Montreal-Boston’, ‘voo Nova Iorque-Dubai’
 c. *Paris, Texas*; *Tucson, Arizona*

Os núcleos das projeções (71a) e (71b) abrigam as relações de coordenação, as quais são comumente tratadas em termos de projeções funcionais (e.g., *Boolean phrases*, cf. MUNN, 1992). Nesse ponto, a presença de uma projeção funcional parece ser bem motivada, mas, ainda assim, sua assunção não é consensual entre os autores.

Chomsky (2013, p. 46), por exemplo, assume que uma relação de coordenação deve conter a seguinte estrutura subjacente, [_α CONJ [_β Z W]], a fim de capturar a simetria semântica comum a essas formações. Nela, uma conjunção toma dois objetos idênticos, Z e W, concatenados simetricamente, e um deles se move para uma posição acima da conjunção a fim de garantir a rotulação da estrutura, sem recorrer a projeções funcionais como faz Di Sciullo⁸³.

⁸³ A fim de garantir a relação assimétrica entre os membros de compostos coordenados, uma vez que eles não participam de uma relação de modificação, como em (68), a autora argumenta que o primeiro constituinte desses compostos não pode ocupar a posição de especificador da projeção-F, o que faz com que esse primeiro

As preposições em (71c), (71d) e (71e), por sua vez, não precisam ser partes de uma projeção funcional, mas podem ser vistas como unidades que contêm um quadro de subcategorização próprio, como sugerido por Hale e Keyser (2002), e, assim, selecionam os membros do composto. Dessa forma, a configuração estrutural é alcançada não pela projeção, mas como um resultado da saturação argumental dessas preposições. É válido discutir ainda se os dados trazidos em (70b) e (70c) são, realmente, compostos ou apenas itens enumerados que permitem uma paráfrase empregando tais preposições.

O terceiro argumento utilizado por Di Sciullo para defender a projeção-F é a presença de elementos de ligação nos compostos. A autora mostra que elementos de ligação estão presentes em alguns compostos do inglês e das línguas românicas, e são amplamente encontrados nas línguas eslavas, germânicas e no grego moderno. Como exemplo, trazemos os dados do grego moderno apresentados pela autora:

(72) a. *pagvuno*

(Grego Moderno)

pag-VL-vun- -o
montanha de gelo-NEU NOM-SG
'iceberg'

b. *kapnokalierjia*

kapn-VL-kaliERG- -i- -a
cultivo de tabaco -ivo-FEM NOM-SG
'cultivo de tabaco'

(DI SCIULLO, 2009, p. 153)

constituente tome uma projeção-F como seu complemento, e, então, o segundo constituinte do composto é concatenado na posição de complemento dessa projeção-F, tal como em (ii), retirado de Di Sciullo (2005, p. 28):

Ralli (1992), por outro lado, mostra que a presença da vogal de ligação *-o-* nos compostos do grego moderno é resultado de um requerimento fonológico responsável por eliminar uma sequência de consoantes na fronteira entre radicais, viabilizando, dessa forma, sua realização. Uma evidência para o caráter fonológico desse elemento de ligação é a sua ausência nos compostos em que o primeiro radical termina em vogal e o segundo inicia-se em vogal, tal como pode ser visto nos exemplos em (73):

- (73) a. *ayriaántropos* < *ayri-* *ántropos*
 ‘homem selvagem’ ‘selva’ ‘homem’
 b. *aksiayápitos* < *aksi-* *ayapitós*
 ‘amável’ ‘digno’ ‘amado’

(RALLI, 1992, p. 154)

Portanto, já que a inserção da vogal de ligação decorre de um requerimento fonológico, não há evidências para assumi-la como núcleo de uma projeção funcional na composição, mas sim de relegá-la ao componente fonológico. A mesma exigência fonológica é encontrada nos compostos formados por radicais do russo (cf. PERKLES, 2008) e das línguas românicas – considerando-se os compostos neoclássicos –, o que enfraquece a assunção de uma projeção funcional na composição, como desejado pela autora⁸⁴.

Outro problema na proposta de Di Sciullo, somando-se ao da projeção-F, é o de que a derivação dos compostos não é uniforme entre as línguas. A autora divide a formação de compostos do inglês e do francês em dois domínios computacionais, sendo os primeiros derivados em D_M , enquanto os últimos são derivados em D_S . A principal motivação para essa divisão operacional está relacionada à ordem linear dos constituintes do composto e sua semelhança ou não com constituintes sintagmáticos de mesma combinação categorial.

Uma consequência dessa divisão estabelecida pela autora é que, enquanto os compostos do inglês são gerados pelas operações de D_M ⁸⁵, os compostos do francês são gerados pelas operações sintáticas, *Merge* e *Move*, em D_S . Ora, a análise proposta tem como principal objetivo unificar os compostos a partir de projeções funcionais, as quais, como vimos, são pouco motivadas, e acaba, no fim, deixando sua formação arbitrária na gramática.

⁸⁴ Cf. Scher e Nóbrega (2014), para uma discussão mais aprofundada sobre a ausência de informação morfossemântica atrelada às vogais de ligação.

⁸⁵ As operações de D_M são extensões das operações sintáticas em D_S . Essas operações podem ser vistas como releituras das operações assumidas pela sintaxe minimalista, porém relegadas ao domínio morfológico, e se aplicam, exclusivamente, a estruturas assimétricas. São elas: *M-Shift*, *M-Link*, *Agree* e *M-Flip*.

Ao mesmo tempo, a autora não diz por que e como uma língua seleciona um domínio para a formação de compostos, seja ele D_M ou D_S , e não outro.

Há que se questionar, ainda, como essas projeções funcionais figuram ontologicamente na gramática, já que elas parecem ser adquiridas como tais e não derivadas computacionalmente. Nesse sentido, se as relações gramaticais estabelecidas nos compostos são as mesmas encontradas entre sintagmas, a proposta de Di Sciullo deixa a gramática altamente redundante, já que essas mesmas relações gramaticais deverão ser realizadas por mecanismos distintos, ora via projeções funcionais na formação de compostos, ora pela tipologia do *Merge* na formação de sentenças.

Nesse ponto, podemos recuperar as relações gramaticais discutidas nas classificações de Bisetto e Scalise (2005) e Scalise e Bisetto (2009) e avaliar sua explicação na TA. Em linhas gerais, pudemos perceber que, apesar de objetivar uma explicação universal, baseando-se na questão de por que os compostos são parte da linguagem humana, Di Sciullo explica as relações gramaticais internas à composição via projeções funcionais tipificadas. Tais projeções, além de serem pouco motivadas, não apresentam uma aplicação uniforme, pois embora garantam a assimetria em relações de modificação, acabam deixando a estrutura de alguns compostos nada triviais⁸⁶ a fim de manter sua generalização.

A nossa proposta, como veremos no Capítulo 5, busca explicar a formação dos compostos a partir da natureza da operação *Merge*. Nesse ponto, visamos garantir uma aplicação unificada das ferramentas adotadas, e, ao mesmo tempo, defender que as mesmas operações responsáveis por gerar sentenças, também geram palavras e suas relações internas. Desse modo, a formação dos compostos dentro do modelo a ser exposto não faz uma distinção paramétrica dos locais onde sua derivação pode ocorrer, mas fornece a geração de qualquer composto dentro de um único espaço computacional: a sintaxe.

4.4. A composição na Morfologia Distribuída

A MD, como vimos no Capítulo 2, assume uma teoria sintática de formação de palavras em que as operações responsáveis pela sua formação e o espaço em que são geradas são os mesmos que o das estruturas sentenciais. Dessa forma, se assumirmos que a composição apresenta um caráter de morfologia-como-sintaxe, tal como descreve Harley (2009), o que temos em mãos é um conjunto de dados que parece trazer evidências robustas

⁸⁶ Cf. nota 83, sobre os compostos coordenados.

para as suposições básicas do modelo; por exemplo, a de que compostos sintagmáticos são gerados no mesmo espaço computacional em que as sentenças.

Entretanto, se nos detivermos nos dados de compostos sintagmáticos e continuarmos endossando sua geração no componente sintático, não estaremos fazendo nada além do que as abordagens lexicalistas já fizeram, pois desde Di Sciullo e Williams (1987), para não voltar tanto no tempo, compostos sintagmáticos já eram vistos como sendo gerados pela sintaxe e enviados ao léxico através da *Nonmorphological Word Creating Rule* (DI SCIULLO; WILLIAMS, 1987)⁸⁷.

Dessa forma, se queremos endossar a plausibilidade de uma teoria não-lexicalista para a formação de palavras, é necessário que consigamos explicar, a partir dos mesmos mecanismos, a formação de compostos ditos morfológicos⁸⁸ no componente sintático, ou seja, compostos formados por radicais ou por raízes presas, os quais nada têm a ver, a princípio, com estruturas sentenciais.

Além dessa distinção na seleção de dados relevantes para endossar a abrangência do modelo, temos que levar em consideração as condições sob as quais as palavras são geradas e as implicações que a formação de palavras em um componente sintático traz para o trabalho morfológico. Uma vantagem dessa visão sintaticocêntrica, por exemplo, é a necessidade de se buscar os elementos mínimos nos processos de formação de palavras, visto que o espaço computacional deve garantir sua formação de modo universal, restando, para o componente morfológico, os ajustes estruturais referentes à distribuição superficial e fonológica das unidades morfossintáticas (i.e., nós terminais), tal como apontamos no Capítulo 2.

Nesse sentido, a MD acaba redefinindo muitas das concepções morfológicas tradicionais, desde a noção de morfema, a qual passa a ser vista como um feixe de traços gramaticais, até a ausência de distinção entre morfemas flexionais e derivacionais, uma vez que ambos são vistos como um feixe de traços. O trabalho aqui desenvolvido segue na mesma linha, ao revisar a composição e colocar em questão quais são as unidades mínimas a que o componente sintático deve ter acesso a fim de gerar um composto.

O que queremos enfatizar é que, ao se desenvolver um trabalho morfológico nesses termos, é importante questionar em que medida não se está subjugando ao componente

⁸⁷ A tradição transformacionalista, que antecede as propostas lexicalistas, já tratava compostos sintagmáticos como estruturas sintáticas (e.g., LEES, 1960; MARCHAND, 1965; LEVI, 1978; ROEPER; SIEGEL, 1978). Por outro lado, Benveniste (2006 [1974]), dentro de uma abordagem estruturalista, assume a mesma posição, argumentando que “o impulso que produziu os compostos não veio da morfologia, onde nenhuma necessidade os solicitaria; ele provém das construções sintáticas com suas variedades de predicação” (2006 [1974], p. 163).

⁸⁸ Fazemos alusão à distinção entre compostos sintagmáticos e morfológicos, como feito por Villalva (2000), embora nosso intuito seja o de excluir tal distinção no que diz respeito à geração dessas palavras.

sintático informações que são específicas a uma língua ou conjunto de línguas, e deixando de lado um refinamento das propriedades mínimas presentes no processo em questão. Além disso, do ponto de vista teórico, é necessário ter em mente a abrangência das ferramentas que temos disponíveis para analisar nossos dados, e se elas, de fato, mostram-se adequadas para o trabalho.

Nesse sentido, Borer (1998, p. 152) coloca, de maneira muito simples, as questões básicas que devem nortear o trabalho do morfólogo lexicalista e do morfólogo não-lexicalista na busca por evidências que corroborem suas posições teóricas. Enquanto os primeiros devem mostrar que um componente independente para a formação de palavras contém operações e restrições que não podem ser reduzidas às operações e restrições sintáticas, os últimos devem explicar a riqueza da formação de palavras sem apelar a qualquer operação sintática que encontre motivação apenas na formação de palavras. Dessa forma, se para explicar a formação de palavras na sintaxe, um autor tem que postular mecanismos que não são encontrados nas sentenças, ou vice-versa, ele estará argumentando contrariamente aos seus pressupostos teóricos.

Esse conjunto de diretrizes elencado acima é o que sustenta nossa argumentação com relação à formação dos compostos na MD. Ao discutir a teoria do *Merge* e sua aplicação na formação de palavras, veremos que pouco foi dito sobre o assunto, e que o modo como alguns autores vêm fazendo uso dessa operação acaba mostrando que a formação de palavras não se enquadra nas operações sintáticas (e.g., a proposta de um *Merge* subespecificado para os compostos, cf. ZHANG, 2007; a ser discutido no Capítulo 5). Com o intuito de abarcar as relações gramaticais internas à composição e unificar as operações sintáticas disponíveis para a geração de objetos linguísticos, desenvolveremos, no próximo capítulo, uma proposta que endossa o uso da tipologia do *Merge* na formação de palavras.

Outra questão diz respeito à interpretação da composição dentro do modelo. Harley (2009) define os compostos como uma unidade morfologicamente complexa, a qual é identificada como uma palavra devido ao seu comportamento sintático e fonológico. Em resumo, sua definição fica como se segue:

(74) *Definição de composto (HARLEY, 2009, p. 130, tradução nossa)*

Composto: uma unidade do tamanho de uma palavra que contém duas ou mais raízes.⁸⁹

⁸⁹ Trecho original (HARLEY, 2009, p. 130): “**Compound:** A word-sized unit containing two or more Roots.”

Se essa for a definição de um composto para a MD, encontraremos alguns problemas. Primeiro, qual o tamanho de uma palavra e o que a define como palavra no espaço computacional? Sabemos que os compostos devem conter mais de uma raiz, e que essas duas raízes são interpretadas como uma única unidade sintática – o que a autora chama de palavra. No entanto, sua definição não sintetiza o modo como essas duas raízes se transformam em uma única unidade. Por esse motivo, é preciso refinar a definição de composto dentro do modelo, tornando explícito o que entendemos por palavra, sintaticamente, e qual é a sua constituição interna.

Assumimos que a formação de um domínio categorial acima de uma estrutura, seja ela sentencial ou apenas uma concatenação de duas raízes categorizadas, é o que permite sua interpretação como composto. Para o sistema computacional, o importante é que essa estrutura possa ser utilizada como uma única unidade sintática, independentemente de sua complexidade interna, e nós podemos garantir isso ao assumir que todo composto é resultado de uma categorização estrutural.

A formação de um domínio categorial, resultado da categorização de uma estrutura sintática, não é necessária apenas por razões computacionais, mas também encontra evidências em um conjunto variado de dados que motivam sua presença, os quais serão detalhados na próxima subseção, juntamente com a redefinição dos compostos na MD.

4.4.1. A segunda unidade mínima: o domínio categorial

As relações gramaticais presentes entre os membros constituintes de um composto não são suficientes para definir a composição, pois essas relações são as mesmas entre constituintes sintagmáticos, como vimos na seção 4.2. Assim, propomos que, além das relações gramaticais como unidade mínima, todo composto deve formar um domínio categorial, ou seja, um núcleo definidor de categoria – seja ele *n*, *v* ou *a* – deve ser concatenado acima de sua estrutura a fim de demarcar a composição⁹⁰.

Nesse sentido, assumimos que a característica unificadora da composição é a criação desse domínio categorial acima de uma estrutura sintática construída a partir de determinada relação gramatical, contrariamente à hipótese de que a característica unificadora dos

⁹⁰ Borer (2012) especula a possibilidade de preposições e advérbios também serem núcleos categorizadores, algo que não discutiremos, especificamente, nesta dissertação.

compostos é a presença de uma projeção-F, como sugere Di Sciullo (2005b, 2009). Nossa proposta, portanto, resume-se na seguinte assunção:

(75) *Composto*

Um composto é formado quando dois ou mais núcleos complexos, em determinada relação sintática, são recategorizados por um núcleo definidor de categoria – *n*, *v* ou *a*.

Os núcleos complexos, em (75), devem ser interpretados como uma raiz concatenada a um ou mais núcleos definidores de categoria, que serão, posteriormente, concatenados em determinada relação sintática $\mathfrak{R}$ – seja ela de subordinação, atribuição, ou coordenação –, e, então, essa estrutura sintática formada é categorizada por outro núcleo definidor de categoria, o qual é responsável por fornecer a informação categorial, as informações flexionais e a sua utilização como uma única unidade sentencial.

Essa proposta pode ser representada, de modo sintético, na estrutura genérica abaixo:

(76)

Em que α , β , γ são núcleos categorizadores, e $\mathfrak{R}$ as relações sintáticas de coordenação, subordinação e atribuição.

O nível de complexidade desses núcleos complexos, ou seja, a quantidade de informação gramatical presente em seus núcleos categorizadores, pode variar, pois irá depender de quais traços gramaticais presentes na Lista 1 serão selecionados para formar seu feixe de traços.

Para endossar a presença de um domínio categorial, apresentamos as evidências trazidas pelos (i) problemas de projeção, pela (ii) flexão independente nos compostos do grego moderno, pelos (iii) compostos parassintéticos, pela (iv) incompatibilidade de traços nas relações de concordância em compostos exocêntricos e pelas (v) alterações na grade argumental de estruturas de incorporação nominal. Todos esses fenômenos serão detalhados nas próximas subseções.

Além da questão levantada pela exocentricidade categorial, outros requerimentos de rotulação parecem ser violados nos dados de compostos, por exemplo, compostos com um núcleo predicator que não projeta seu rótulo categorial à estrutura formada. Nos compostos V-N, por exemplo, o rótulo categorial não é o do núcleo predicator, mas o de seu complemento, ou seja, N, ou o de outra categoria gramatical, analogamente ao que ocorre nos compostos coordenados acima⁹². Ilustramos essa questão com os dados de compostos V-N do PB, em (78):

- (78) a. limpa-vidro(s); porta-mala(s) V+N → N
 b. [antena] corta-fita; [porta] corta-fogo V+N → A

O terceiro caso a ser exemplificado diz respeito à incompatibilidade categorial nos compostos que apresentam uma relação interna de modificação. Admitindo que essas relações de modificação sejam decorrentes de uma adjunção, seria esperado que o núcleo da estrutura determinasse o rótulo do complexo formado, da mesma forma que ocorre nos casos de adjunção em sintagmas. Por exemplo, em uma adjunção como [_{NP}[_{NP}fotografia do João]], o rótulo da estrutura é um NP, e não um PP. No entanto, algumas línguas apresentam compostos V-N em que o N não é um complemento de V, mas um adjunto, normalmente interpretado como um locativo, tal como pode ser visto nos dados abaixo:

- (79) a. *traîne-buisson* lit. pendura-arbusto ‘tipo de pássaro’ (Francês)
 b. *reveille-matin* lit. acorda-manhã ‘despertador’
(DESMETS; VILLOING, 2009)
 c. *m-lowammalo* lit. entra no lugar ‘substituto/pronome’ (Chichewa)
 d. *chi-gonambáwa* lit. dorme no bar ‘alcoólatra’
(MCHOMBO, 2004)

⁹² O que estamos chamando de questão de projeção é o que Scalise, Fábregas e Forza (2009, p. 58) consideram a exocentricidade categorial, definida do seguinte modo: “um composto é categorialmente exocêntrico se o constituinte na posição de núcleo não impõe seus traços categoriais à construção como um todo” (do inglês: “a compound is categorially exocentric if the constituent in the head position does not impose its categorial features on the whole construction.”). Nos compostos coordenados em (77), não temos uma resposta precisa de qual dos constituintes é o núcleo do composto, por isso isolamos esses dados como um caso de exocentricidade categorial por excelência, a fim de distingui-los dos casos em que há um núcleo predicator, em termos sintáticos, mas que não rotula a estrutura. Nesse sentido, a assunção de um domínio categorial na composição permite a projeção do núcleo predicator, fazendo com que as questões de rotulação deixem de ser um problema.

O que podemos perceber nos dados em (79) é que, embora a estrutura tivesse que ser rotulada como V, já que o núcleo modificado é o verbo, temos como rótulo um N ou um A. Esse caso, tal como a ausência de endocentricidade e as questões de rotulação em contextos de seleção, não devem ser vistos como argumentos para uma abordagem lexicalista, mas sim como parte de uma característica geral da composição que é a exigência da formação de um domínio categorial⁹³. O que queremos enfatizar é que, apesar da derivação de um composto seguir os passos de uma derivação sintagmática, será a categorização de sua estrutura que fará com que a sintaxe, no nível sentencial, interprete o complexo formado como um único nó, garantindo a endocentricidade de projeção de seus respectivos núcleos internos.

Nos compostos do chinês, por exemplo, a estrutura resultante das relações de coordenação é categorizada por um núcleo definidor de categoria, o qual é responsável pelo valor categorial do composto, independentemente do de seus membros constituintes. Nos compostos V-N, por outro lado, a derivação deve ocorrer tal como a derivação de um sintagma verbal qualquer, porém, em determinado momento, um núcleo definidor de categoria concatena-se à estrutura, transformando esse complexo sintático em um único objeto, sem impedir a projeção do núcleo verbal. O mesmo pode ser dito para os compostos V-N formados a partir de uma relação de modificação, pois tal como nos demais casos, a categoria do composto será fornecida pelo núcleo categorizador responsável por criar o domínio.

4.4.1.2. Flexão independente nos compostos do grego moderno

Outra evidência para a formação de um domínio vem dos casos de flexão independente nos compostos formados por radicais do grego moderno. Nespor e Ralli (1996), Ralli (2008a, 2009) e Ralli e Karasimos (2009) mostram que a informação flexional em compostos [radical + radical] é concatenada ao composto como um todo, uma vez que a marca flexional do composto não é a mesma marca flexional do segundo radical quando utilizado como uma palavra autônoma⁹⁴:

⁹³ Padrosa-Trias (2007), ao tratar dos compostos N-V do catalão, aponta que o estabelecimento de uma relação de modificação feita por N a um núcleo V é problemático para as teorias sintáticas de formação de palavras. Contudo, se nossa proposta estiver correta, essa relação de modificação é decorrente do tipo de *Merge* aplicado sobre esses elementos, ao passo que a categoria do composto é fornecida independentemente, e está a cargo do núcleo categorial responsável por criar o domínio.

⁹⁴ Alguns poderão questionar se o elemento de ligação *-o-* não seria uma marca flexional. Ralli (1992, 2008a, 2008b, 2009) traz uma série de argumentos para defender que essa vogal é um item de conexão, e não um morfema com conteúdo gramatical. Se olharmos para o primeiro radical quando utilizado como uma palavra

(80) a.	<i>nixt-o-lúluδ-o</i>	<	<i>lulúδ-i</i>	(Grego Moderno)
	noite-EL-flor-INFL		flor-INFL	
	‘flor da noite’		‘flor’	
b.	<i>omorf-ó-peδ-o</i>	<	<i>peδ-i</i>	
	bonito-EL-criança-INFL		criança-INFL	
	‘criança bonita’		‘criança’	
c.	<i>sklir-ó-karδ-o</i>	<	<i>karδj-á</i>	
	duro-EL-coração-INFL		coração-INFL	
	‘coração duro’		‘coração’	
d.	<i>mer-ó-nixt-o</i>	<	<i>níxt-a</i>	
	dia-EL-noite-INFL		noite-INFL	
	‘dia e noite’		‘noite’	
e.	<i>riz-ó-γal-o</i>	<	<i>γal-a</i>	
	arroz-EL-leite-INFL		leite-INFL	
	‘pudim de arroz’		‘leite’	

Esses dados nos mostram que um domínio deve ser formado, transformando os objetos concatenados em uma única unidade, a fim de possibilitar a inserção dos traços flexionais verificados nos compostos. Evidências para essa assunção vêm do fato de que os traços flexionais do composto são específicos, e não são os traços flexionais herdados de um dos radicais. Mais adiante, veremos que, ao se concatenar um núcleo definidor de categoria ao complexo formado pelos radicais, será possível permitir que as informações flexionais do composto independam das informações flexionais de seus constituintes, explicando, com isso, os casos de exocentricidade morfológica.

4.4.1.3. Compostos parassintéticos

Os exemplos de parassíntese reiteram a presença de um domínio na formação dos compostos. Esse tipo de composição toma dois radicais, que por si só não constituem um composto, e os concatenam a um morfema sufixal, o qual é responsável por transformar o núcleo complexo em um composto, e por fornecer sua informação categorial.

autônoma, verificaremos que suas marcas flexionais são distintas daquelas que caracterizam o composto, por exemplo, *nixt(a)* ‘noite’ (80a); *mer(a)* ‘dia’ (80c); *riz(i)* ‘arroz’ (80e).

Exemplos desse tipo de composição podem ser encontrados amplamente nas línguas eslavas, em que compostos são, preferencialmente, formados a partir de radicais concatenados a um único sufixo categorizador, tal como pode ser visto nos dados do russo, abaixo, de acordo com a estrutura templática em (81)⁹⁵:

(81) [radical 1 + EL + radical 2 + sufixo derivacional - flexão]

- a. *krasn-o-reč-ivyj* (Russo)
 bonito-EL-fala-ADJ
 ‘eloquente’
- b. *železn-o-dorož-n-yj*
 ferro-EL-estrada-ADJ-MASC.SG.NOM
 ‘relacionado à ferrovia’

(BISETTO; MELLONI, 2008, p. 239)

Os compostos parassintéticos não devem ser confundidos com casos de derivação que tomam como base um composto, tal como ocorre nos dados do PB, em *puxa-saquismo* e *rádio-taxista*. Nesses exemplos, tanto *puxa-saco* quanto *rádio-táxi* são compostos da língua, e a derivação se aplica como um processo independente, ou seja, a concatenação do sufixo não é uma exigência para garantir a gramaticalidade do composto. Contrariamente, os compostos parassintéticos em (81) só se tornam um composto quando o sufixo é adicionado, o que pode ser verificado na agramaticalidade das combinações em (82), tanto nos compostos do russo, quanto nos compostos parassintéticos de outras línguas eslavas e indo-europeias:

- (82) a. *krasn-o-reč-ivyj* (Russo)
 **krasn-o- reč + ivyj/ krasn-o + *rečivyj*
 ‘eloquente’
- b. *blauwogig* (Holandês)
 **blauwoog + ig/ blauw + *ogig*
 ‘de olhos azuis’⁹⁶

⁹⁵ Benigni e Masini (2009, p. 175-176) mostram que esse tipo de composição recebe um estatuto isolado no russo, pela relevância da presença do sufixo garantindo a gramaticalidade do composto.

⁹⁶ Glosa em inglês, ‘blue-eyed’.

- c. *red-blooded* (Inglês)
**red-blood + ed/ red + *blooded*
 ‘avermelhado como sangue’
- d. *kokinomalis* (Grego Moderno)
**kokin-o-mal + is/ kokin-o- + *malis*
 ‘de cabelos vermelhos’
- e. *pescivendolo* (Italiano)
**pescivend(ere) + olol/ pesci + *vendolo*
 ‘vendedor de peixe’
- f. *veriverbium* (Latim)
**ver(i)verb + ium/ veri + *verbium*
 ‘contador de verdades’
- g. *obcokrajowiec* (Polonês)
**obc(o)kraj + owec/ obc(o) + *krajowec*
 ‘estrangeiro’

(BISETTO; MELLONI, 2008, p. 235)

Apesar desse tipo de composição sugerir uma estrutura ternária, Bisetto e Melloni (2008) mostram que as relações de escopo entre seus constituintes evidenciam uma estrutura binária em que o sufixo deve ter escopo sobre os dois radicais, mesmo nos dados em que a combinação entre o radical de segunda posição e o sufixo são palavras atestadas na língua. Concentrando-se nos dados das línguas eslavas, onde esse tipo de composição é realmente produtivo, as autoras chegam a uma generalização estrutural sobre os compostos parassintéticos, a qual está representada na estrutura em (83), ilustrada com um composto parassintético do russo, em (84):

(83) *Padrão configuracional dos compostos parassintéticos (BISETTO; MELLONI, 2008, p. 256)*

Em que α é o radical 1, β é o radical 2 e γ é um sufixo nominal ou adjetival.

(84) a. *strel-o-braz-n-yn*

(Russo)

arco-EL-forma-ADJ- MASC.SG.NOM

‘em forma de arco’

O que podemos perceber a partir desses dados é que um domínio deve ser formado a fim de garantir a gramaticalidade do composto, e a formação desse domínio ocorre após a concatenação dos dois radicais. Portanto, a criação de um domínio categorial na composição, como sugerido nesta dissertação, não deriva apenas de argumentos que possibilitam a rotulação da estrutura independentemente da categoria de seus constituintes, nem somente do conjunto de traços flexionais agregado ao composto, mas decorre também da própria gramaticalidade do composto em algumas línguas, como parte crucial na sua formação.

4.4.1.4. Exocentricidade morfológica

Compostos exocêntricos contribuem para uma visão de domínio na composição, não apenas nos casos de exocentricidade categorial, como visto na subseção 4.4.1.1, mas também nos casos de exocentricidade morfológica, verificado nas relações de concordância estabelecidas com o composto⁹⁷. Seguindo a definição de Scalise, Fábregas e Forza (2009), a exocentricidade morfológica pode ser entendida como se segue:

⁹⁷ Como já mencionado na seção 4.2, a exocentricidade não diz respeito apenas a definições semânticas, ou seja, quando o composto é um hipônimo de um de seus constituintes (e.g., *peixe-espada* é um hipônimo de *peixe*). Reiteramos, aqui, essa afirmação, pois apesar da exocentricidade ser distinguida entre categorial, morfológica e semântica, elas não precisam estar presentes integralmente em um composto para classificá-lo como exocêntrico.

(85) *Exocentricidade Morfológica* (SCALISE; FÁBREGAS; FORZA, 2009, p. 59, tradução nossa)

Um composto é morfologicamente exocêntrico se seus traços morfológicos não forem idênticos aos traços morfológicos de qualquer um de seus membros internos⁹⁸.

Exemplos de exocentricidade morfológica são dados pelos compostos V-N das línguas românicas e bantas⁹⁹. Esses compostos são morfologicamente exocêntricos, pois os traços morfológicos do nome interno ao composto são distintos dos traços presentes nos elementos com os quais o composto estabelece concordância. Nas línguas românicas, a incompatibilidade na concordância pode ser verificada pela diferença de valores nos traços de número e gênero presentes no nome interno ao composto e aqueles presentes no determinante, tal como em (86)¹⁰⁰:

- (86) a. O_{MASC.SG} [para-quedas_{FEM.PL}]_{MASC.SG}
- b. O_{MASC.SG} [porta-malas_{FEM.PL}]_{MASC.SG}
- c. O_{MASC.SG} [salva-vidas_{FEM.PL}]_{MASC.SG}
- d. O_{MASC.SG} [saca-rolhas_{FEM.PL}]_{MASC.SG}

Para garantir que o composto formado contenha informações flexionais independentes daquelas do nome interno ao composto, um núcleo definidor de categoria que conta com traços de número e gênero não especificados – ou seja, não-valorados em termos sintáticos – permite que os traços flexionais do elemento com o qual entra em uma relação de concordância, por exemplo, um determinante ou um adjetivo, sejam transferidos ao composto como um todo. Por sua vez, esse núcleo definidor de categoria é o que forma o domínio da composição. Veremos, detalhadamente, a derivação desses compostos na subseção 5.3.3.1 do Capítulo 5, quando tratarmos da derivação dos compostos subordinados.

⁹⁸ Trecho original (SCALISE; FÁBREGAS; FORZA, 2009, p. 59): “A compound is morphologically exocentric if the morphological features of the compound are not identical to the morphological features of any of its internal constituents”.

⁹⁹ Outro conjunto de dados que poderia ser utilizado como evidência nos casos de exocentricidade morfológica, são os compostos N-A do PB, como *casca-grossa*, *mão-aberta*, *mão-boba*, nos quais a concordância vai ser governada pelos traços de gênero do referente modificado pelo composto (e.g., *o/a casca-grossa*). Quando explorarmos os traços gramaticais presentes nos categorizadores, no Capítulo 5, retomaremos esses dados e indicaremos por que eles também são relevantes para a argumentação de um domínio.

¹⁰⁰ Nas línguas bantas, a diferença no marcador de classe nominal é o que deflagra essa incompatibilidade, pois o marcador de classe nominal do nome interno ao composto é distinto do marcador nominal que caracteriza o composto (cf. BASCIANO; KULA; MELLONI, 2011, p. 220), gerando uma estrutura templática como a seguinte: [(MC_Y) + [[√V + vf] + [(MC_X) + √N]]].

4.4.1.5. Estruturas de incorporação nominal

Os casos de incorporação nominal, ou seja, casos em que um nome é fundido a um verbo para formar outro verbo podem ser vistos como uma evidência adicional ao domínio da composição. A incorporação nominal, encontrada em diversas línguas, exige a explicação de como uma combinação entre um verbo, geralmente transitivo, mais seu objeto direto, consegue tomar outros argumentos além daqueles exigidos pelo verbo em questão.

Um exemplo pode ser dado com o composto verbal *camatrencar* ‘quebrar a(s) perna(s)’, do catalão, em (87), formado por incorporação nominal, em que o verbo de alternância *trencar* ‘quebrar’, ao ter incorporado o objeto direto *cama* ‘perna’, torna-se um verbo transitivo, tal como pode ser visto na sentença em (88):

- (87) *cama-trencar* (Catalão)
 perna-quebrar
 ‘quebrar a(s) perna(s)’

- (88) *El caçador va camatrencar l’ocell.* (Catalão)
 O caçador foi perna-quebrar o-pássaro.
 ‘O caçador quebrou a(s) perna(s) do pássaro’

(PADROSA-TRIAS, 2007, p. 57)

A inserção de novas exigências argumentais pode ser explicada, diretamente, pelo domínio da composição, pois o núcleo verbal categorizador, responsável por criar o domínio, tem em sua constituição traços responsáveis por garantir a seleção de outros argumentos, tal como um núcleo verbal qualquer. É por esse motivo que, embora a incorporação nominal seja a verbalização de uma estrutura quase sentencial, ela ainda é capaz de conter as mesmas propriedades que um verbo simples isolado.

Para ilustrar a amplitude e a variedade do fenômeno, mais dados de incorporação nominal são trazidos em (89), com exemplos de uma variedade de línguas. Todos esses dados são, idealmente, explicados como o do catalão em (87):

- (89) a. *afis-o-kolo* (Grego Moderno)
 cartazes-fixar
 ‘fixar cartazes (na parede)’
- b. *lug-rdzi* (Tibetano)
 ovelhas-cuidar
 ‘pastorear’
- c. *laka-lás-a* (Totonac)
 cara-estapear-IPF
 ‘estapear (alguém) na cara’
- d. *t!šút!š-ul púšil-na* (Tzotzil)
 pequeno.peças-dobrar.repetidamente
 ‘dobrar (alguma coisa) cada vez mais menor’
- e. *yenlhipeya* (Wichí)
o-yen-lhip-ey-a
 3SUJ-fazer-peça-PL-ACC
 fazer-peças
 ‘picar em fatias/ele pica em fatias’
- (ŠTEKAUER; VALERA; KÖRTVÉLYESSY, 2012, p. 72)

É válido mencionar que todo composto verbal exigirá uma explicação sobre como ele, enquanto um núcleo complexo, passa a exigir outros argumentos. Nesse sentido, a mesma explicação sugerida para os casos de incorporação nominal poderá ser estendida para os compostos verbais V-V, exemplificados com dados do japonês em (90), em que a combinação de dois núcleos verbais dá origem a outro verbo:

- (90) a. *de-kakeru* (Japonês)
 sair-pendurar
 ‘começar’
- b. *kaki-naosi*
 escrever-consertar
 ‘reescrever’

c. *yomi-hajimeru*

ler-começar

‘começar a ler’

(SHIBATANI; KAGEYAMA, 1988)

Em conclusão, percebemos que há evidências para se assumir que a formação de um domínio categorial é uma unidade mínima da composição, não apenas para garantir que a sintaxe interprete esse complexo estrutural como um único objeto, mas também para permitir que um conjunto independente de traços categoriais, flexionais e argumentais possam ser inseridos em um composto. Essa relevância vai se tornar mais clara no capítulo seguinte, quando apresentarmos o conjunto de traços presentes em um núcleo categorizador e o modo como a operação *Agree*, tal como concebida em Wurmbrand (no prelo), procede na sua concatenação.

4.5. Conclusões preliminares

Tendo em vista as informações levantadas neste capítulo, podemos concluir que as unidades mínimas para a formação sintática de um composto nas línguas naturais são: (i) as relações gramaticais, seja ela de subordinação, atribuição ou coordenação, entre dois núcleos complexos (i.e., entre duas ou mais raízes sintáticas categorizadas), e (ii) o domínio categorial acima dessa combinação.

No próximo capítulo, formalizamos essas duas unidades mínimas a fim de demonstrar como se processa a derivação dos compostos subordinados e atributivos na gramática, tomando como objeto de análise os compostos V-N e N-N, principalmente, e, trazendo para a discussão, sempre que possível, os demais compostos debatidos neste capítulo, a fim de demonstrar a possibilidade de expansão da análise.

CAPÍTULO 5

DERIVANDO COMPOSTOS

Neste capítulo, formalizamos as unidades mínimas da composição, a fim de demonstrar como o componente sintático procede na formação de um composto. Focalizando os compostos subordinados e atributivos, demonstramos como as mesmas operações sintáticas responsáveis por formar sentenças também são responsáveis por formar palavras e suas relações internas, fato que endossa o pressuposto básico de que a gramática contém um único componente gerativo: a sintaxe.

5.1. Caracterização do capítulo

Concentramo-nos, neste capítulo, na formalização das informações levantadas até este ponto da dissertação, principalmente, no modo como as unidades mínimas apontadas para a composição são manipuladas pela gramática. Assentamos nossa proposta na natureza da operação *Merge*, tal como delineada em Chomsky (2000, 2004), sugerindo como as relações gramaticais internas à composição podem ser estabelecidas a partir de sua aplicação. Ao mesmo tempo, desenvolvemos um refinamento dos feixes de traços presentes nos núcleos categorizadores, a fim de viabilizar uma visão derivacional *bottom-up* para a MD.

Apesar de nossa proposta parecer decorrente do modelo teórico empregado, já que a MD assume as operações da sintaxe minimalista como responsáveis pela formação de palavras (cf. MARANTZ, 1996, 1997, 2001; HARLEY; NOYER, 1999; ARAD, 2003, 2005; EMBICK; NOYER, 2007; HARLEY, 2008, 2009; ARREGI; NEVINS, 2012, dentre outros), essa assunção gera um conjunto de questões acerca da natureza da primeira concatenação no modelo, a saber: a concatenação entre uma raiz e um núcleo categorizador, e sua adequação às concepções atuais da operação *Merge*.

Desse modo, ao mesmo tempo em que procuramos trazer uma contribuição descritiva sobre a formação dos compostos, abordamos também questões teóricas relacionadas à

natureza do primeiro *Merge* na MD (i.e., o *Merge* entre a raiz e o núcleo categorizador). Apontamos que a MD, até então, vem assumindo a operação *Merge* sem refletir sobre suas consequências para um modelo sintaticocêntrico, e, para tanto, buscamos fornecer uma alternativa de unificação entre a proposta de *Merge* dentro da sintaxe minimalista e sua aplicação na formação de palavras desde a primeira concatenação.

O capítulo, portanto, segue dividido da seguinte maneira: na seção 5.2, revisitamos a tipologia da operação *Merge*, tal como elaborada em Chomsky (2000, 2004), e, a partir de suas variedades, traduzimos a primeira unidade mínima da composição nos termos dessa operação, nomeadamente, as relações gramaticais. Posteriormente, comentamos algumas propostas problemáticas para a composição dentro da MD, no que concerne ao modo como empregaram a operação *Merge*, e, em seguida, descrevemos o problema do primeiro *Merge* e uma alternativa de solução.

Na seção 5.3, apresentamos o sistema operacional que assumimos para a sintaxe, a qual se estrutura em um sistema de valoração regido por uma operação de *Agree Reverso*, com base em Wurmbrand (no prelo), e detalhamos a constituição em traços dos primitivos que entram nesse sistema. Por fim, desenvolvemos uma abordagem derivacional *bottom-up* para a MD, derivando compostos subordinados e atributivos de nossa língua específica, o PB, e apontando, sempre que possível, as extensões que as análises sugeridas a esses dados podem ter para os demais tipos de dados tratados ao longo da dissertação.

Como resultado, visamos estabelecer uma uniformidade nas operações e no modo como a derivação de uma sentença ocorre, ao demonstrar que a formação de palavras e a formação de sentenças se encadeiam em um mesmo espaço computacional.

5.2. A influência da operação *Merge* nos processos de formação de palavras: a composição

Chomsky (1993, 1994, 1995, 2000, 2004, e trabalhos subsequentes) afirma que a sintaxe contém uma operação que é dada livremente, a qual é necessária para garantir a recursividade do sistema, a operação *Merge*. Em linhas gerais, essa operação toma dois objetos α e β , *já construídos*, e cria novos elementos a partir da concatenação entre esses dois primeiros – a qual pode ser representada como um conjunto $\{\alpha, \beta\}$.

O *Merge* pode ser visto como a operação básica do sistema gerativo¹⁰¹, pois gera um conjunto infinito de expressões construídas hierarquicamente, e é responsável não apenas pelo estabelecimento da estrutura argumental de núcleos predadores, mas também por concatenar informações morfológicas à estrutura (e.g., *v*, *T*, *Asp*) e por possibilitar o movimento de constituintes já concatenados a posições mais altas.

Essa operação possibilita a formação de unidades maiores a partir de unidades menores, dando origem a novos objetos sintáticos. Um novo objeto formado a partir do *Merge*, diga-se *K*, deve ser constituído dos itens α e β e ter a forma $\{\gamma, \{\alpha, \beta\}\}$, em que γ identifica qual o tipo de *K*, ou seja, sua categoria, e também indica quais são suas propriedades relevantes (cf. CHOMSKY, 1995, p. 243). γ , portanto, é o rótulo de *K*, e essa combinação mínima pode ser resumida como se segue:

(91) $K = \{\gamma, \{\alpha, \beta\}\}$, em que α, β são objetos sintáticos e γ é o rótulo de *K*.

A rotulação de *K* já era, desde o início do Programa Minimalista (cf. CHOMSKY, 1995, p. 244) determinada derivacionalmente, e o rótulo da combinação era construído a partir das informações presentes nos dois constituintes internos ao conjunto, ou seja, α e β . Das possibilidades para a rotulação da estrutura¹⁰², Chomsky (1995) assume que ou α ou β projeta, e, conseqüentemente, o constituinte que projeta deve ser considerado o núcleo de *K*. Dessa forma, se α projeta, então $K = \{\alpha, \{\alpha, \beta\}\}$ e α é o núcleo de *K*. Essa relação está representada de maneira ilustrativa na estrutura arbórea em (92):

Porém, a definição de qual constituinte deveria ser o núcleo de *K* era dada arbitrariamente, pois não havia uma informação anterior que determinasse que α deveria ser o núcleo e não β , já que $\{\alpha, \beta\} = \{\beta, \alpha\}$. Esse problema da arbitrariedade na identificação do núcleo, juntamente com a necessidade de se explicitar a natureza do *Merge* (i.e., diferenciar

¹⁰¹ A operação *Merge* vem ganhando destaque devido ao empreendimento biolinguístico, o qual questiona como o ser humano adquiriu a capacidade de adquirir a linguagem. Como apontam Berwick e Chomsky (2011, p. 30), a evolução da linguagem será reduzida à operação *Merge*, devido à recursividade que essa operação introduz ao componente gerativo, e avaliá-la na formação de palavras é uma tarefa imprescindível.

¹⁰² As possibilidades de rotulação apontadas por Chomsky (1995, p. 244) são: (i) o rótulo da estrutura seria a intersecção de α e β ; (ii) a união de α e β ou (iii) um ou outro, ou seja, ou α ou β . A opção tomada pelo autor é a terceira, pois tanto (i) quanto (ii) gerariam elementos contraditórios.

um caso de seleção/substituição de uma adjunção) levou Chomsky a redefinir a operação *Merge*, propondo uma variedade de operações. Com isso, a operação *Merge* deixa de ter uma única forma, e passa a conter diferentes modos de concatenação, diferenciados de acordo com a presença ou não de um núcleo seletor e pela relação de simetria ou assimetria que estabelece.

Dessa tipificação da operação *Merge*, dois modos de concatenação são distinguidos: *set-Merge* e *pair-Merge*, em que o primeiro forma um conjunto $\{\alpha, \beta\}$ e o segundo forma um par ordenado $\langle \alpha, \beta \rangle$. O *set-Merge* fica, então, responsável pelo estabelecimento das relações argumentais e pelos movimentos na estrutura sintática (de acordo com a distinção entre *Merge* externo e *Merge* interno), e o *pair-Merge* fica responsável pela concatenação de adjuntos.

Set-Merge é a forma mais simples dessas operações, e sua definição pode ser vista como uma versão revisada da noção de concatenação apresentada em (91). Sua reformulação é dada em conjunto com a operação *Agree*, pois o núcleo da concatenação passa a ser identificado a partir de um traço seletor que deverá ser valorado, e sua valoração se dá com a concatenação a um elemento capaz de valorar esse traço. Com isso, o rótulo da estrutura passa a ser previsível, pois o núcleo que tem seus traços valorados é quem rotula a estrutura, tal como definido abaixo:

“*Set-Merge* tem, tipicamente, uma assimetria inerente. Quando α, β se concatenam, ele satisfaz requerimentos (selecionais) de um (o *selector*), mas não de ambos. Por outro lado, de modo bastante geral, o seletor é unicamente determinado pelo par (α, β) , tal como pode ser visto nos casos de revisão. *Set-Merge* de (α, β) tem algumas propriedades de *Agree*: um traço F de um dos elementos concatenados (diga-se α) deve ser satisfeito para que a operação ocorra. Além disso, F é o rótulo de α , e, portanto, detectável de modo direto [...] O traço seletor F para o *Merge* é análogo à sonda em uma relação de *Agree*. Ademais, F é o único elemento de α que entra na operação, logo, o único disponível sem complicações futuras para se determinar o rótulo dos elementos concatenados. Nesse caso também, o rótulo é previsível e não precisa ser indicado: o rótulo do seletor projeta.” (CHOMSKY, 2000, p. 133-134, tradução nossa)¹⁰³.

¹⁰³ Trecho original (CHOMSKY, 2000, p. 133-134): “Set-Merge typically has an inherent asymmetry. When α, β merge, it is to satisfy (selectional) requirements of one (the *selector*) but not both. Fairly general, furthermore, the selector is uniquely determined for a pair (α, β) , as can be seen by reviewing cases. Set-Merge of (α, β) has some of the properties of Agree: a feature F of one of the merged elements (say, α) must be satisfied for the operation to take place. Furthermore, F is in the label of α , hence detectable in an optimal way [...] The selector F for Merge is analogous to the probe for Agree. Furthermore, F is the only element of α that enters into the operation, hence the only one available without further complication to determine the label of the merged elements. In this case too, then, the label is predictable and need not be indicated: the label of the selector projects.”

Vemos, portanto, que *set-Merge* envolve uma valoração de traços para viabilizar sua aplicação e garantir qual é o núcleo da estrutura formada, e, consequentemente, qual o constituinte que a rotula. *Pair-Merge*, por outro lado, não apresenta um traço seletor, e contrário à relação simétrica criada pelo *set-Merge*, *pair-Merge* cria apenas relações assimétricas, a fim de garantir a assimetria inerente às relações de adjunção.

Nos casos em que *pair-Merge* ocorre, α se adjunge a β formando $\langle \alpha, \beta \rangle$, e o elemento adjungido α não altera a categoria do núcleo β ; além disso, é β quem rotula a estrutura. Como define Chomsky (2000, p. 133), a adjunção de α a β forma $K = \{\gamma, \langle \alpha, \beta \rangle\}$, em que γ é o rótulo de β . Nas palavras do autor, esse tipo de operação *Merge* deve ser interpretado do seguinte modo:

“É um fato empírico que há também uma assimetria na operação de adjunção, a qual toma dois objetos β e α e forma um par ordenado $\langle \alpha, \beta \rangle$, α adjungido a β . *Set-Merge* e *pair-Merge* são descendentes de substituição e adjunção em teorias anteriores. Dadas as propriedades básicas da adjunção, podemos pensar intuitivamente α concatenado em um plano separado, com β conservando todas as suas propriedades no “plano principal”, a estrutura simples.” (CHOMSKY, 2004, p. 18, tradução nossa)¹⁰⁴.

Como comentado acima, Chomsky ainda faz uma divisão interna na operação *Merge* a partir de suas funções na sintaxe, que são: o *Merge* externo e *Merge* interno. O *Merge* externo é responsável por concatenar elementos da numeração à estrutura sintática que está sendo derivada, ou seja, concatenam-se dois objetos α e β que estão separados, tal como representado em (93), retirado da representação feita em Langendoen (2003). O *Merge* interno, por sua vez, é responsável por mover elementos que já são parte da estrutura, por exemplo, se β é parte de α e deve ser movido, uma cópia de β é criada e β é concatenado, internamente, mais acima na estrutura, como pode ser visto em (94):

¹⁰⁴ Trecho original (CHOMSKY, 2004, p. 18): “It is an empirical fact that there is also an asymmetric operation of adjunction, which takes two objects β and α and forms the ordered pair $\langle \alpha, \beta \rangle$, α adjoined to β . Set-merge and pair-merge are descendants of substitution and adjunction in earlier theories. Given the basic properties of adjunction, we might intuitively think of α as attached to β on a separate plane, with β retaining all its properties on the “primary plane”, the simple structure.”

(93) *Merge externo*

(94) *Merge interno*

(LANGENDOEN, 2003, p. 308)

Tendo em vista a distinção entre *set-Merge* e *pair-Merge*, podemos redefinir, agora, as relações gramaticais presentes na composição em termos sintáticos. Admitindo que enquanto a subordinação está para uma relação predicado-argumento, a operação *set-Merge* é a que deve se aplicar a fim de garantir a saturação do traço de seleção presente no núcleo desses compostos. Já nos casos de modificação, um adjunto será concatenado a um núcleo; logo, a operação responsável por estabelecer essa relação assimétrica é *pair-Merge*.

Embora não tratemos da coordenação detalhadamente, assumimos que um composto coordenado pode seguir a mesma descrição feita em Chomsky (2013, p. 46), mencionada no Capítulo 4, quando discutimos as estruturas de coordenação na TA de Di Sciullo (2005b, 2009). Nesse caso, uma conjunção toma dois objetos categorialmente idênticos, Z e W, concatenados simetricamente, e um deles se move para uma posição acima da conjunção a fim de garantir a rotulação da estrutura, como representado em (95). No nosso caso, em particular, a relação procede da mesma maneira, porém deverá ser recategorizada para gerar um domínio.

- (95) a. [_α Conj [_β Z W]]
 b. [_γ Z[_α Conj [_β Z W]]]

Em resumo, podemos redefinir, em termos sintáticos, como as relações gramaticais internas à composição devem ser entendidas a partir de agora:

(96) *Relações gramaticais internas à composição em termos sintáticos*a. **Subordinação**

Relação predicado-argumento em que um constituinte α tem seu traço seletor valorado pelo seu complemento β , gerando uma estrutura simétrica.

b. **Atribuição**

Relação de adjunção em que o constituinte não-núcleo α é concatenado ao núcleo β , independentemente da valoração de um traço, gerando uma estrutura assimétrica.

c. **Coordenação**

Relação conjuntiva em que uma conjunção toma dois constituintes categorialmente idênticos, α e β , em uma relação simétrica, desfazendo a simetria via movimento, a fim de permitir a rotulação da estrutura.

Uma vez que as relações em (96) são estabelecidas entre dois núcleos complexos, a formação de um composto se completa quando outro núcleo categorizador se concatena à estrutura criando um domínio, de acordo com a definição dos compostos sugerida no Capítulo 4. Outras propostas não-lexicalistas, no entanto, recorrem a uma relação subespecificada entre os membros de um composto, forçando a gramática a conter uma terceira forma de *Merge*, contextual à formação de palavras. Apresentamos e contra-argumentamos essa proposta na subseção seguinte.

5.2.1. As formas da operação *Merge* e a composição na Morfologia Distribuída

Zhang (2007), fazendo uso da MD, exclui as formas da operação *Merge* vistas na última subseção da formação dos compostos. Ao tratar dos compostos do chinês, a autora elenca um conjunto de problemas para as abordagens sintáticas, mostrando que os dados dessa língua suscitam outras questões que não foram levantadas pelos compostos de línguas românicas e germânicas nas principais análises.

Para a autora, (i) a exocentricidade categorial (i.e., a rotulação de um composto por uma categoria diferente daquelas de seus membros, e.g., $N+N \rightarrow A$, tal como ocorre nos compostos coordenados vistos no Capítulo 4)¹⁰⁵, (ii) a liberdade de projeção (i.e., fato semelhante ao que ocorre nos compostos V-N do PB, em que tanto V quanto N podem rotular a estrutura)¹⁰⁶, (iii) o duplo licenciamento de traços de seleção (i.e., compostos semelhantes às

¹⁰⁵ Cf. Capítulo 4, subseção 4.4.1.1.

¹⁰⁶ Cf. Capítulo 4, subseção 4.4.1.1.

estruturas de incorporação, em que o composto formando seleciona argumentos além daqueles selecionados por seu núcleo verbal interno)¹⁰⁷ e (iv) o efeito da integridade lexical (i.e., a impossibilidade de um dos constituintes internos ao composto servir como um elemento anafórico nas relações de ligação)¹⁰⁸ exigem uma revisão na explicação de como os compostos são formados.

Na tentativa de propor uma solução para os problemas levantados, Zhang (2007) toma uma posição contrária a um sistema sintático operacionalmente uniforme, pois assume um nível subespecificado para a concatenação entre as raízes. Em sua proposta, duas raízes nuas são concatenadas diretamente uma a outra e, apenas posteriormente, um núcleo definidor de categoria atribui a informação categorial ao composto, tal como representado abaixo¹⁰⁹:

O primeiro problema dessa análise diz respeito à ausência de uniformidade nas operações utilizadas para a formação de objetos sintáticos, sejam eles palavras ou sentenças, pois o nível subespecificado para a concatenação não é verificado nos demais contextos sintáticos. Do ponto de vista teórico, não seria ideal a postulação de operações ou tipos de combinações exclusivas à formação de palavras, uma vez que o pressuposto da teoria é o de que as mesmas operações assumidas para a formação de sentenças devem ser capazes também de formar palavras.

Além disso, uma vez que a raiz é definida como uma unidade destituída de traços sintáticos, ela só se torna visível para o componente computacional se for concatenada a um núcleo definidor de categoria, pois será esse núcleo quem fornecerá seus traços gramaticais. Portanto, a combinação de duas raízes nuas é indesejável do ponto de vista computacional,

¹⁰⁷ Cf. Capítulo 4, subseção 4.4.1.5.

¹⁰⁸ Um exemplo dado pela autora é a seguinte sentença (2007, p. 177):

(iii) **Ta xian na-le yi ba cha-hu ranhou ba ta_i dao-ru beizi-li.*
 ele primeiro tomar-PRF um CL chá-pote então BA o verter-dentro xícara-dentro.
 “Ele primeiro pegou um pote de chá_i, e então o_i (o chá) verteu em uma xícara.”

¹⁰⁹ f= morfema funcional.

pois elas por si só estão invisíveis para a sintaxe, o que impede sua manipulação no espaço computacional¹¹⁰.

Um terceiro problema é a aplicação translinguística dessa análise. Se avaliarmos sua capacidade explicativa a partir dos compostos em que há, claramente, uma relação argumental entre seus constituintes, por exemplo, os compostos V-N em que o N é interpretado como argumento interno de V (e.g., *salva-vidas*, *limpa-vidros*, *estraga-prazer*), percebemos que as estruturas em (97) não são capazes de explicar como V seleciona o N complemento, tal como não explicam a diferença entre um composto N-N atributivo (e.g., *peixe-espada*) e um composto N-N coordenado (e.g., *bar-restaurante*).

O quarto e último problema a ser apontado também diz respeito aos compostos V-N, mais precisamente àqueles das línguas românicas, em que tanto V quanto N carregam marcadores de classe específicos (e.g., *bat-e pap-o*, *tir-a mof-o*), e sua distribuição correta não é prevista pela estrutura acima, ou seja, a inserção de um marcador de classe nominal poderia acontecer em qualquer uma das raízes, o que não é verificado na empiria¹¹¹.

Contrários a essa proposta de um nível subespecificado para o *Merge*, acreditamos que o domínio da composição, proposto no Capítulo 4, consegue abarcar os problemas levantados pela autora, já que permite que um novo rótulo categorial seja fornecido à estrutura formada, e que possíveis traços flexionais independentes, tal como um novo quadro argumental, sejam anexados a esse composto, sem ter que postular que duas raízes nuas sejam concatenadas livremente e que essa concatenação aconteça somente com esse conjunto de dados.

Antes de passarmos para a apresentação da derivação dos compostos, temos que dar um passo para trás e explicar a primeira concatenação em uma abordagem não-lexicalista, a saber, a natureza da operação *Merge* responsável por concatenar uma raiz a um núcleo definidor de categoria. Em nossa visão, é imprescindível a manutenção de uma uniformidade nas operações disponíveis na gramática, e, para isso, é importante que explicitemos qual o *Merge* responsável por essa primeira concatenação. Será esse o assunto que discutiremos na próxima subseção.

¹¹⁰ Essa ideia está em consonância com as propostas que retiram da raiz a propriedade de seleção de argumentos (cf. MINUSSI, 2012; BASSANI, 2013), tal como havia sido sugerido por Marantz (1997), Embick (2004), entre outros, já que a raiz só passa a conter propriedades gramaticais e seletivas depois de concatenada a um núcleo categorizador.

¹¹¹ Zwitserlood (2008) propõe uma solução semelhante para os compostos da língua de sinais holandesa. A proposta da autora carrega os mesmos problemas presentes no trabalho de Zhang (2007), ao assumir a concatenação livre entre duas raízes para a formação de um composto. Além disso, essa autora assume que traços de ordem fonológica (e.g., boca, movimento repetido, mão) são manipulados pela sintaxe na formação de objetos complexos, o que não é plausível num modelo como a MD, pois a sintaxe não opera com esse tipo de traços. Tais traços, por sua vez, devem ser vistos como externalizações de traços morfossintáticos reais (e.g., número, pessoa, categoria, etc.).

5.2.2. O problema do primeiro *Merge*: a categorização das raízes

A MD, como repetimos reiteradas vezes, assume que as mesmas operações responsáveis por gerar sentenças, também são responsáveis por gerar palavras. Seus principais autores afirmam, explicitamente, que essas operações são as mesmas assumidas na sintaxe minimalista (cf. MARANTZ, 1996, 1997, 2001; HARLEY; NOYER, 1999; ARAD, 2003, 2005; EMBICK; NOYER, 2007; HARLEY, 2008, 2009; ARREGI; NEVINS, 2011, dentre outros). Porém, não há uma discussão sobre como a tipologia da operação *Merge* estabelecida no nível sentencial influencia e se aplica nos processos de formação de palavras.

Enquanto, para Chomsky (2000, p. 132), os objetos a serem concatenados já estão formados quando a derivação se inicia; na MD, esses objetos precisam, antes de tudo, ser derivados, uma vez que essa teoria assume que as raízes não possuem informação categorial, mas as adquirem a partir da concatenação com um núcleo definidor de categoria (e.g., [$\sqrt{+n}$], [$\sqrt{+v}$]). Portanto, uma consequência dessa visão acategorial das raízes (cf. MARANTZ, 1996, 1997)¹¹² é a necessidade prévia da formação de um item lexical antes de sua concatenação a outro item, o qual também deverá ser formado. A formação desses itens, por sua vez, levanta uma importante questão: qual a natureza da operação *Merge* que ocorre entre uma raiz e um núcleo definidor de categoria?

Em uma abordagem representacional para a MD, a questão do primeiro *Merge* talvez não seja tão relevante, porém, se buscamos desenvolver uma abordagem derivacional para o modelo, é importante esclarecer como ocorre o primeiro *Merge*, qual a sua natureza e como esse objeto formado se concatena a outro posteriormente. Por isso, antes de sugerirmos uma explicação para essa questão, devemos fazer um balanço da constituição dos primitivos do modelo e de sua distribuição no espaço computacional.

O primeiro primitivo que deve ser avaliado é a raiz. Para a grande maioria dos autores, a raiz é uma unidade destituída de traços sintáticos, e pode ser vista ora como um *placeholder* (cf. MARANTZ, 1996), ora como um átomo conceitual (cf. ARAD, 2003, 2005), ora como um conjunto vazio (cf. DE BELDER; VAN CRAENENBROECK, 2011, 2013) e ora como um índice (cf. HARLEY, no prelo). Independentemente da definição que se assuma, há que se demonstrar como o núcleo definidor de categoria seleciona um objeto sintático destituído de traços gramaticais e garante a sua visibilidade no espaço computacional. Por

¹¹² A hipótese das raízes acategoriais também é assumida em outras abordagens não-lexicalistas, como a Teoria Exoesquelética de Borer (2003, 2005, 2012).

outro lado, núcleos definidores de categoria não fornecem informação categorial exclusivamente a raízes, ou seja, esses núcleos podem categorizar sintagmas a fim de formar outras palavras, por exemplo, os compostos V-N, N-A, N-prep-N, A-N, dentre vários outros.

A questão que se coloca, portanto, é a seguinte: se o núcleo definidor de categoria seleciona a raiz, e essa seleção ocorre via a operação *Agree* – tal como vimos na discussão sobre as formas do *Merge* – analogamente a uma relação de sondagem (cf. CHOMSKY, 2000; COLLINS, 2002), qual é a forma do *Merge* responsável pelo primeiro *Merge* na MD? E se a relação entre esses elementos for, de fato, um caso de seleção, quais são os traços que estão em jogo nessa relação de *Agree*, uma vez que a raiz não contém traços gramaticais?

Elencamos três possíveis soluções para essa questão:

(98) *Possíveis soluções para a questão do primeiro Merge na MD*

- a. A concatenação de um núcleo definidor de categoria e uma raiz ocorre a partir de um *Merge* cego, que apenas concatena os dois elementos sem se importar com sua constituição interna;
- b. A concatenação de um núcleo definidor de categoria e uma raiz é uma relação de modificação estabelecida por um *pair-Merge*, gerando uma estrutura assimétrica entre esses dois elementos;
- c. A concatenação entre um núcleo definidor de categoria e uma raiz é uma relação de seleção estabelecida por *set-Merge*, e a sintaxe conta com diferentes sabores de núcleos categorizadores a fim de garantir a seleção de raízes e de estruturas sintagmáticas (e.g., categorização de sintagmas para a formação de compostos);

Começamos pela solução (98a). Essa solução é, evidentemente, a menos adequada dentre as três, já que diminui metodologicamente o poder explicativo da teoria ao postular um *Merge* que ocorre apenas no nível da formação das palavras, com vias a garantir que sua formação continue sendo gerada pela sintaxe. Como vimos na divisão de trabalhos elaborada por Borer (1998), um modelo não-lexicalista deve evidenciar que as mesmas operações sintáticas são capazes de formar palavras e não postular operações adicionais que ocorram apenas em casos isolados – tal como fez Zhang (2007). Se a solução (98a) for realmente plausível, é necessário que apontemos construções sentenciais em que um *Merge* cego também se aplique.

A solução (98b) também nos parece inadequada, pois se há uma relação de modificação entre o núcleo definidor de categoria e a raiz, é de se esperar que esse núcleo

modifique a raiz, uma vez que é ele quem traz os traços gramaticais que a modelam. No entanto, se assim for, o rótulo da estrutura formada nunca será o do categorizador, mas sim o da raiz, pois nos casos de *pair-Merge* – ou em qualquer outra relação de modificação – é o elemento modificado que rotula a estrutura. Se assumirmos o inverso, a solução parece menos óbvia, pois qual seria a modificação realizada entre uma unidade destituída de traços sintáticos a um nó categorizador?

Além disso, o categorizador deve estar ativo para as relações de *Agree*, pois ele serve como um interventor em diversos contextos, principalmente nos casos de exocentricidade morfológica¹¹³, em que os traços morfossintáticos do composto devem se diferenciar dos traços morfossintáticos de seus membros (e.g., $O_{\text{MASC.SG}}$ [*salva-vidas*_{FEM.PL}]_{MASC.SG}). Nesse sentido, se o categorizador for concatenado sempre em um “plano separado”, em decorrência da aplicação de *pair-Merge*, ele nunca estará ativo para servir como um núcleo interventor nesses contextos. Dessa forma, excluimos a solução em (98b) para a explicação dessa relação

¹¹⁴.

Nossa opção de solução, portanto, é a (98c), já que nos parece mais adequado interpretar a relação do primeiro *Merge* como um caso de seleção, em que a raiz atrai o núcleo categorizador. Elaboraremos melhor como essa relação se processa na gramática, após apresentarmos a visão derivacional que assumimos nesta dissertação.

5.3. Condição do *Merge*: explicando a formação de palavras a partir da valoração de traços

A visão derivacional que assumimos para a MD está baseada, essencialmente, na abordagem sintática para seleção proposta em Wurmbrand (no prelo). Apesar de discutir mais profundamente o que definimos como *set-Merge*, essa proposta traz uma preocupação relevante para a construção da estrutura sintática, pois tenta explicitar como as relações de

¹¹³ Cf. subseção 5.3.3.1, onde formalizaremos essa relação de intervenção.

¹¹⁴ De Belder e Van Craenenbroeck (2011) sugerem que a concatenação de um categorizador a uma raiz ocorre via *pair-Merge*. Contudo, a teoria de *Merge* empregada pelos autores assume que todo *Merge* é assimétrico, ou seja, todo *Merge* é um *pair-Merge*, de acordo com as propostas de Zwart (2010). A solução sugerida incorre nos mesmos problemas listados acima, além de conter uma série de redundâncias operacionais relativas à formação de núcleos complexos (e.g., a necessidade que um núcleo complexo tem de retornar à numeração a fim de esperar o momento exato para ser concatenado à estrutura; cf. subseção 5.3.3.1, para uma discussão sobre esse problema).

seleção entre as projeções presentes na estrutura sentencial estão codificadas na gramática, ou seja, o que permite que determinada projeção se concatene a outra durante a derivação¹¹⁵.

A autora desenvolve um mecanismo estritamente local e determinístico para as relações de seleção argumental e morfológica, que induz a um refinamento dos traços gramaticais presentes nos nós terminais, ao impor condições à operação *Merge*. Esse refinamento dos traços gramaticais presentes nos nós terminais é algo desejado para um modelo de gramática que opera apenas com traços, tornando a construção da estrutura sintática direcionada¹¹⁶. Nesta seção, buscamos desenvolver uma extensão do que Wurmbrand fez para a sentença, para a formação de palavras – mais especificamente, para a derivação sintática dos compostos.

Na proposta de Wurmbrand, a operação *Merge* está sujeita à *Last Resort* (99), ou seja, ela ocorrerá somente quando um traço for satisfeito:

(99) *Last Resort* (ABELS, 2003 *apud* WURMBRAND, no prelo, tradução nossa)

Um constituinte somente poderá ser concatenado, ou seja, concatenado na base ou reconcatenado, se isso levar a uma imediata satisfação de traços previamente não satisfeitos¹¹⁷.

Chomsky (2000), em sua definição básica de *Merge*, apresentada no início da seção 5.2, propõe algo semelhante, ao postular que essa operação é desencadeada por um traço seletor que deve ser satisfeito via *Agree*¹¹⁸. Em sua proposta, o traço seletor é análogo à sonda na relação de *Agree*, e desencadeia uma busca em seu domínio de c-comando a fim de encontrar um alvo com o qual possa ser satisfeito e se concatenar.

Wurmbrand segue na mesma linha de Chomsky (*op. cit.*), no que diz respeito à valoração de um traço para o estabelecimento da concatenação. A autora propõe que a operação *Agree* deve ser uma condição sob a operação *Merge*, e seu licenciamento está

¹¹⁵ Embora haja uma discussão na literatura sobre como argumentos são concatenados à estrutura sintática e como alguns constituintes são movidos via *Merge* (e.g., movimento do sujeito sentencial para T a fim de satisfazer o traço EPP), pouco foi dito sobre como as relações de seleção entre as projeções da espinha sentencial estão codificadas na gramática. A proposta de Wurmbrand, por sua vez, procura explicitar o que está em jogo quando qualquer projeção se concatena a outra na estrutura.

¹¹⁶ Levando em consideração que essa direcionalidade depende dos traços presentes nos núcleos envolvidos na derivação.

¹¹⁷ Trecho original (ABELS, 2003 *apud* WURMBRAND, no prelo): “A constituent may only be merged, i.e. base-merged or re-merged, if that leads to the immediate satisfaction of a previously unsatisfiable feature”.

¹¹⁸ Estamos citando apenas Chomsky (2000) com relação à proposta de valoração de traços para a aplicação do *Merge*, porém outros autores também poderiam ser mencionados, e.g. Watanabe (1996), Collins (2002), Abels (2003), dentre outros.

relacionado à satisfação de um traço, o qual deverá ser valorado de acordo com a condição especificada abaixo¹¹⁹:

(100) *Condição do Merge (WURMBRAND, no prelo, tradução nossa)*

Concatene α e β se α puder valorar um traço de β ¹²⁰.

A valoração desses traços, no entanto, não segue a visão padrão de *Agree*, em que o núcleo seletor busca em seu domínio de c-comando um alvo com o qual possa ser valorado, procedendo a uma valoração de baixo para cima, tal como proposto, primeiramente, por Chomsky (2000, 2001), em (101). Wurmbrand vale-se de uma operação direcionalmente inversa de *Agree*, o *Agree Reverso* (do inglês, *Reverse Agree*), a qual é construída a partir de um sistema de valoração que ocorre de cima para baixo na estrutura sintática¹²¹.

O *Agree Reverso* (doravante, AR) é uma operação governada pela valoração de traços, e a interpretabilidade desses traços está dissociada da noção de valoração, ou seja, tanto traços interpretáveis quanto traços não-interpretáveis podem vir valorados (*i/uF: val*) ou não-valorados (*i/uF: __*) da numeração. A definição dessa operação está dada em (102), e ilustrada em (103):

(101) *Agree Padrão – valoração de baixo para cima (cf. CHOMSKY, 2000, 2001)*

- i. Um traço não-valorado F (uma *sonda*) em um núcleo H sonda seu domínio de c-comando para encontrar um outro exemplar do traço F (um *alvo*), com o qual concordará.
- ii. Se o alvo tem um valor, seu valor será atribuído como o valor da sonda¹²².

¹¹⁹ Uma proposta semelhante foi elaborada por Pesetsky e Torrego (2006), na qual os autores argumentam que toda realização da operação *Merge* – seja *Merge* externo ou interno – deve envolver uma relação sonda-alvo, sintetizada no seguinte requerimento: *Vehicle Requirement on Merge* (VRM): “Se α e β se concatenam, algum traço F de α deve sondar um traço F em β ” (do inglês: “If α and β merge, some feature F of α must probe F on β ”). A diferença dessa proposta com relação à proposta de Wurmbrand está relacionada ao modo como os traços são operados, pois enquanto para Pesetsky e Torrego (*op. cit.*) o que ocorre é uma relação de compatibilidade de traços (do inglês, *feature matching*), para Wurmbrand é uma relação de valoração de traços.

¹²⁰ Trecho original (WURMBRAND, no prelo): “Merge α and β if α can value a feature of β ”.

¹²¹ Uma proposta semelhante é defendida por Zeijlstra (no prelo), rotulada de *Upward Agree*. O autor elenca uma série de argumentos para motivar a necessidade de uma operação *Agree* com valoração de cima para baixo, contrariamente à visão padrão delineada em Chomsky (2000, 2001) e em Pesetsky e Torrego (2004, 2007), por exemplo.

¹²² Trecho original (definição seguindo CHOMSKY, 2000, 2001, sintetizada por PESETSKY; TORREGO, 2007, p. 263): “(i) an unvalued feature F (a *probe*) on a head H scans its c-command domain for another instance of F (a *goal*) with which to agree; (ii) If the goal has a value, its value is assigned as the value of the probe”.

(102) *Agree Reverso* – valoração de cima para baixo (WURMBRAND, no prelo, tradução nossa)

Um traço F: ___ em α é valorado por um traço F: val em β , sse

- i. β c-comanda α E
- ii. α está acessível para β [acessível: não sofreu *Spell-Out*]
- iii. α não valora {um traço de β }/ {um traço F de β }^{123/124}.

No *Agree* padrão, o elemento mais alto (i.e., a sonda) deve ser deficiente para motivar uma busca de traços disponíveis em seu domínio. Nessa visão, podemos considerar que *Last Resort* é definido com relação à sonda. Diferentemente, no AR, o elemento mais alto em uma relação de *Agree* não é deficiente, pois deve vir valorado, o que faz com que *Last Resort* seja realizado pela Condição do *Merge* e não a partir de uma sondagem. Com isso, a visão de que *Last Resort* é uma condição para o estabelecimento do *Merge* fica mais evidente.

Dentro de uma abordagem de AR para a sintaxe, podemos explicar a seleção morfológica e a seleção de argumentos via valoração de traços. Nesse caso, a seleção morfológica não é determinada lexicalmente, mas reduz-se à *Agree* e, do mesmo modo, a estrutura argumental do verbo é vista como um conjunto de traços não-interpretáveis não-valorados que precisam ser valorados via *Agree* para que o *Merge* se realize.

Podemos tomar a morfologia verbal como exemplo de seleção morfológica nesse sistema. Os verbos são inseridos com um traço de tempo não-interpretável não-valorado e os verbos não-lexicais carregam um traço de tempo não-interpretável, porém valorado, correspondente ao seu valor semântico (e.g., *iT*: progressivo, *iT*: perfectivo). Desse modo, como todo verbo tem traços não-valorados, eles deverão se concatenar com um elemento não-verbal a fim de satisfazer a Condição do *Merge*. A valoração, por sua vez, associará o traço não-interpretável de um verbo mais baixo com o valor semântico do verbo de cima, derivando

¹²³ Segundo a autora, a terceira condição é necessária para prevenir que dois nós irmãos valorem um ao outro (e.g., quando T se concatena com um AuxP, T valora o *uT*: ___ de AuxP, mas AuxP não pode valorar o *uT*: ___ de T).

¹²⁴ Trecho original (WURMBRAND, no prelo): “A feature F: ___ on α is valued by a feature F:val on β , iff: (i) β c-commands α AND, (ii) α is accessible to β [accessible: not spelled out], (iii) α does not value {a feature of β }/ {a feature F of β }.”

a seleção morfológica. Tal relação está ilustrada no exemplo abaixo em que *v* se concatena a T com o intuito de valorar seus traços de T:

A seleção argumental, por sua vez, processa-se da mesma forma. Os predicados têm um traço não-interpretável que deverá ser valorado (e.g., o verbo *comer* carrega um traço [$u\phi$: $_$] que deverá ser valorado por um DP; o verbo *demandar* carrega um traço de tempo [uT : $_$] que deverá ser valorado por um T subjuntivo), e, após valorados, esses traços são traduzidos semanticamente como se segue: traços ϕ são traduzidos como um “argumento de” tanto para V quanto para *v*, ou seja, o XP que valorar os traços ϕ de V será interpretado como um argumento de V (o argumento interno), ao passo que o XP que valorar os traços ϕ de *v* será interpretado como um argumento de *v* (o argumento externo)¹²⁵.

V ainda contém um traço *v* que carrega os valores de projeções *vP/VP-shell* possibilitando a introdução de outros argumentos na estrutura (e.g., agente, experienciador), o qual é traduzido na identificação do evento após sua valoração¹²⁶. Além disso, um traço-Q interrogativo pode estar presente em V para garantir a inserção de complementos interrogativos em alguns verbos. As relações de valoração dos traços de V e *v* estão sintetizadas em (105):

¹²⁵ Percebamos que a estrutura argumental não é dada via um traço seletor como em Chomsky (2000) ou a partir da checagem de traços, como em Adger (2003), em que um núcleo verbal pode ser subespecificado com um traço [uD : $_$] ou [uC : $_$] para garantir a concatenação de um argumento. Em Wurmbrand, embora a ideia seja a mesma, ou seja, a seleção é realizada via a valoração de traços, a autora estabelece um sistema em que a valoração de determinado traço por certo núcleo lexical faz com que esse núcleo seja interpretado como um argumento (e.g., um núcleo XP que valora os traços [$u\phi$: $_$] de V será interpretado como seu argumento interno pelo sistema).

¹²⁶ Os casos de alternância causativa-incoativa estão codificados nas especificações do traço de *v* em V, pois é a sua presença que permitirá a inserção da projeção de *vP* e, conseqüentemente, de outro argumento. Dessa forma, enquanto verbos como *pular* e *quebrar* diferem no que concerne à opcionalidade do traço de *v* em V, um *vP* pode ser omitido em *quebrar*, mas não em *pular*. O conjunto de traços de V presente nesses verbos pode ser resumido da seguinte maneira (os parênteses indicam opcionalidade):

(iv)	<i>pular</i>	[uT : $_$, uv : $_$, ($u\phi$: $_$)]	João pulou o muro; *O muro pulou.
(v)	<i>quebrar</i>	[uT : $_$, (uv : $_$), $u\phi$: $_$]	João quebrou o vaso; O vaso quebrou.

- (105) a. v/V : $[u\phi: __]$ XP ϕ :val é um argumento de v/V
 b. V : $[uQ: __]$ XP iQ : val é interrogativo
 c. V : $[uv: __]$ Identificação do evento

Em resumo, trazemos a derivação da sentença “O João comeu a torta”, adaptada do trabalho de Wurmbrand, a fim de ilustrar a seleção morfológica e argumental através de um sistema de construção da estrutura sintática baseado na valoração de traços. Em (106a), estão os itens lexicais com seus respectivos traços; e de (106b) a (106h) estão representadas as valorações de cada passo da operação *Merge*, indicada por $< >$, após o AR ter se aplicado (os traços adquiridos via valoração estão sublinhados):

- (106) a.
-
- ```

graph TD
 TP --> DP1[DP
[iφ:3SG.MASC]]
 TP --> T_prime[T']
 T_prime --> T[T
[uφ:__, iT: past]]
 T_prime --> vP[vP]
 vP --> DP2[DP
[iφ: 3SG.MASC]]
 vP --> v_prime[v']
 v_prime --> v[v
[iv:AGENT, uφ:__]]
 v_prime --> VP[VP]
 VP --> V[V
[uT:__, uv:__, uφ:__]]
 VP --> DP3[DP
[iφ: 3SG.FEM]]

```
- b.  $<V + DP>$   $[uT: \_\_, uv: \_\_, u\phi: \underline{3SG.FEM}]$  (VP)  
 c.  $<v + VP>$   $[iv:AGENT, u\phi: \_\_ / uT: \_\_, uv: \underline{AGENT}, u\phi: \underline{3SG.FEM}]$  ( $v'$ )  
 d.  $<DP + v'>$   $[iv:AGENT, u\phi: \underline{3SG.MASC} / uT: \_\_, uv: \underline{AGENT}, u\phi: \underline{3SG.FEM}]$  ( $vP$ )  
 e.  $<T + vP>$   $[iv:AGENT, u\phi: \underline{3SG.MASC} / uT: \underline{past}, uv: \underline{AGENT}, u\phi: \underline{3SG.FEM}]$   
 f. *Transfer; Spell-Out* de VP  $[uT: \underline{past}, uv: \underline{AGENT}, u\phi: \underline{3SG.FEM}]$   
 g.  $[u\phi: \_\_, iT: past]$  ( $T'$ )  
 h.  $<DP + T'>$   $[u\phi: \underline{3SG.MASC}, iT: past]$   
 i. *Transfer; Spell-Out* da estrutura remanescente.

As relações de modificação, por outro lado, não são exploradas nesse trabalho. A autora sugere que os modificadores sejam concatenados tardiamente à estrutura sintática, como uma operação de última hora, que ocorre no momento da transferência da estrutura para as interfaces. Essa proposta deriva da ausência de seleção entre os elementos sintáticos que

entram em uma relação de modificação, pois sua concatenação independe de valoração de traços (i.e., não há uma relação “argumento de” entre o modificador e o modificado).

O *pair-Merge*, então, é uma operação que combina dois elementos independentes em um único domínio de *Spell-Out*, e ocorre em *Transfer*, produzindo a configuração correta a ser lida tanto por LF quanto por PF. Essa visão de inserção de última hora<sup>127</sup> para os modificadores está em consonância com a proposta de Hornstein e Nunes (2008) em que modificadores estão sujeitos apenas à composicionalidade semântica.

Wurmbrand, portanto, oferece um sistema derivacional estritamente local para a construção da estrutura sintática, a partir do mecanismo de AR, juntamente com uma explicitação da distribuição dos traços nos núcleos sintáticos. É essa visão derivacional *bottom-up* que queremos implementar na formação de palavras dentro da MD, pois isso nos leva a um refinamento dos traços presentes em cada nó sintático. Nas próximas subseções, desenvolvemos a tarefa de apresentar a constituição dos nós sintáticos em termos de traços e mostrar como sua combinação se processa no espaço computacional.

### 5.3.1. O que constitui as raízes?

Como vimos no final da subseção 5.2.2, o *Merge* entre a raiz e um núcleo definidor de categoria deve ocorrer a partir de uma relação de seleção entre o núcleo e a raiz. Porém, para que a seleção ocorra, seja ela em termos chomskianos ou dentro da proposta de Wurmbrand, é necessário que se estabeleça uma relação de *Agree* entre esses dois elementos. Para que isso aconteça, a raiz precisa conter um traço que possibilite sua participação em uma relação de *Agree* com o núcleo categorizador, caso contrário o *Merge* não ocorrerá<sup>128</sup>.

Por essa razão, postulamos que a raiz contém um traço de borda R não-interpretável não-valorado [*uR*: \_\_\_\_] que será satisfeito apenas quando um núcleo definidor de categoria, contendo um traço de borda R interpretável valorado, entrar em uma relação de *Agree* com a raiz<sup>129</sup>. Essa relação de valoração, a que chamamos de hipótese do Primeiro *Merge*, pode ser resumida como se segue:

<sup>127</sup> Estamos nos referindo a essa concatenação como “inserção de última hora” para não confundir com a noção de inserção tardia para a fonologia, na MD.

<sup>128</sup> Vale mencionar que, se as relações de modificação se processam tardiamente no espaço computacional, como uma operação de última hora (cf. WURMBRAND, no prelo), a visão de que o *Merge* entre a raiz e um núcleo definidor de categoria é um *pair-Merge* perde ainda mais plausibilidade, pois a informação categorial deve estar presente no início da derivação a fim de permitir a concatenação de outras projeções à estrutura sintática.

<sup>129</sup> Agradeço a Susi Wurmbrand, comunicação pessoal, pelas sugestões e discussões sobre como a relação entre a raiz e o categorizador poderiam ser estabelecidas no modelo.

(107) *Hipótese do Primeiro Merge*

Toda raiz contém um traço de borda R não-interpretável não-valorado [ $uR$ : \_\_\_\_] que poderá ser valorado apenas por um traço de borda R interpretável e valorado como  $n$ ,  $v$  ou  $a$ .

O traço de borda, nesse sentido, é um atributo – seguindo a terminologia de Adger (2010), e pode conter os seguintes valores:  $n$ ,  $v$  ou  $a$ , tal como ilustrado em (108), equiparando, assim, sua estrutura a dos demais traços<sup>130/131</sup>. Portanto, o sabor categorial da raiz dependerá do valor do traço de borda com o qual ela entra em uma relação de *Agree*, como representado em (109).

(108) *Traço de borda R* (109) *Categorização de uma raiz*

Como apenas os núcleos definidores de categoria carregam o traço [ $iR$ : val] serão somente eles que poderão se concatenar à raiz. Apesar de parecer trivial, essa proposta visa adequar a visão de *Merge* utilizada pela sintaxe minimalista às questões primárias do *Merge* na MD, sem recorrer à postulação de um *Merge ad hoc* para a formação de palavras. Além disso, a hipótese do Primeiro *Merge* abarca de modo direto a hipótese das raízes acategoriais e garante a liberdade da raiz em receber qualquer informação categorial, ao mesmo tempo em que a torna visível no espaço computacional.

Uma vez que a concatenação está estabelecida, é necessário rotular a estrutura formada. Deve ficar claro, no entanto, que a proposta de Wurmbrand, contrária ao sistema de Chomsky (2000, 2004) não envolve um traço seletor, pois a Condição do *Merge* exige que o sistema

<sup>130</sup> Agradeço ao Hedde Zeijlstra, comunicação pessoal, por me sugerir a mesclagem do traço de borda R com as informações categoriais, algo que simplificou de modo considerável a proposta sugerida para o primeiro *Merge* nas versões anteriores deste trabalho.

<sup>131</sup> Por exemplo, comparando-se à estrutura de um traço de número, esse contém um atributo – número – e os seguintes valores: singular, plural, dual, etc.

computacional valora um traço a fim de que um elemento qualquer, seja ele um argumento ou um morfema de tempo, seja concatenado à estrutura. Portanto, a rotulação não é definida em termos de um traço seletor nesse sistema.

Embora as questões de rotulação não sejam discutidas explicitamente em seu trabalho, Wurmbrand, comunicação pessoal, comenta que, ao invés de rótulos, o que temos são os feixes de traços do núcleo. Porém, para saber qual elemento é o núcleo, é necessário que estabeleçamos a seguinte condição abaixo:

(110) *Condição para identificação do núcleo em uma estrutura formada via Agree Reverso*

O núcleo  $\beta$  responsável por valorar um traço deficiente  $F$  de  $\alpha$  – em que  $\alpha$  é o primeiro núcleo abaixo de  $\beta$  – a fim de satisfazer a Condição do *Merge*, será o núcleo da estrutura formada e o responsável pela projeção dos traços após a concatenação.

Assim, como o mecanismo de AR funciona inversamente ao mecanismo do *Agree* Padrão, seria esperado que o elemento que valorasse os traços de um núcleo deficiente fosse o responsável por rotular a estrutura. Portanto, de acordo com a definição acima, como é o núcleo definidor de categoria com o traço de borda  $R$  valorado quem valora o traço [ $uR$ : \_\_\_] da raiz, será ele o núcleo da estrutura, e, ao mesmo tempo, o elemento responsável por projetar seus traços ao conjunto formado após a concatenação<sup>132</sup>.

Outra questão com relação às questões de projeção deve ser discutida. Na concatenação de uma raiz com um núcleo definidor de categoria, a projeção decorrente dessa concatenação é ou não uma projeção máxima? Em BPS, uma projeção é máxima se ela já não mais projeta, logo, se formamos um nome após a concatenação de um núcleo definidor de categoria a uma raiz, a projeção deve ser máxima, como em (111), caso nenhum outro núcleo definidor de categoria seja concatenado à estrutura. Porém, se temos em mãos um caso de derivação com dois ou mais núcleos definidores de categoria (e.g., [ $N[v[N[\sqrt{cris}t]al]iz-a]ção$ ]), então a projeção não é máxima e somente o será quando não houver mais núcleos definidores de categoria a serem concatenados, como em (112):

<sup>132</sup> Susi Wurmbrand, comunicação pessoal, comenta também que, outra alternativa para a rotulação pode ser dada a partir da natureza do traço de borda da raiz, uma vez que traços de borda são tratados de modo diferente na sintaxe, por exemplo, o fato deles exigirem a ocorrência de um *Merge*. Isso, supõe Wurmbrand, seria a razão pela qual o traço de borda da raiz não projeta, fazendo com que, consequentemente, apenas os traços do núcleo categorizador sejam transmitidos acima na estrutura.

Essa questão da projeção é relevante para a formação dos compostos, uma vez que deve estar claro se o que estamos concatenando são dois *nPs* ou dois *ns* na formação de um composto N-N por exemplo, já que não é explícito como essa distinção se dá no modelo<sup>133</sup>.

Em resumo, propomos que as raízes não são definidas pela ausência de traços, mas pela presença de um traço de borda R não-interpretável não-valorado [*uR*:   ]<sup>134</sup>. Por outro lado, não assumimos que a raiz seja definida pelo traço de borda por si só, ou seja, uma espécie de *place-holder* demarcado pela presença desse traço, mas sim como uma unidade detentora de conteúdo semântico desde o início da derivação e que carrega apenas um traço sintático capaz de torná-la visível no espaço computacional após sua valoração.

### 5.3.2. O que constitui os núcleos categorizadores?

Em nossa visão, os núcleos categorizadores são feixes de traços que modelam a raiz com a qual são concatenados. O conjunto de traços que cada núcleo categorizador traz tem influências não apenas na sintaxe, no que diz respeito às questões de valoração, mas também na fonologia, no que concerne às regras de aplicação de acento<sup>135</sup>.

Assim, torna-se imprescindível explicitar quais são os traços e suas possíveis distribuições nos núcleos categorizadores, tal como as consequências de cada combinação de traços para a gramática. Antes de explicitá-los, temos que ter em mente que o conjunto de traços dos categorizadores deve garantir que eles sejam concatenados tanto a raízes quanto a estruturas mais complexas a fim de gerar um composto.

<sup>133</sup> Vale salientar que em casos de não satisfação de algum traço, a estrutura também não é rotulada como máxima, por exemplo, o traço EPP em T promove a formação de um nível intermediário a fim de garantir o movimento do sujeito sentencial. Na formação de palavras, acreditamos que tais casos não são encontrados.

<sup>134</sup> Boeckx (2010, p. 29) também propõe que o que ele postula como endereços conceituais (do inglês, *conceptual addresses*) – os quais são utilizados num sentido próximo ao de raiz para a MD – contêm um traço de borda por razões distintas daquelas apresentadas aqui. Para o autor, a sintaxe opera apenas concatenando endereços conceituais uns aos outros e é somente após o *Spell-Out* que cada um desses endereços será interpretado categorialmente.

<sup>135</sup> Tal como discutido sobre o traço de classe, no Capítulo 3, a ser mais bem ilustrado no Capítulo 6.

Começemos pelo domínio nominal. De acordo com Lieber (1992), os nomes contêm os seguintes traços: categoria, gênero e classe. Ritter (1993), ao tratar do domínio nominal, adiciona a esse conjunto de traços a informação de número, tratando-a como uma projeção funcional independente. Tal como esses autores, assumimos que todos os traços supramencionados integram o feixe de traços dos núcleos definidores de categoria nominal, estando o traço de categoria formalizado pelo traço de borda R e os demais sintetizados sob o rótulo de *traços  $\varphi$*  (exceto o traço de classe).

Embora esses últimos traços estejam sintetizados sob o rótulo de traços  $\varphi$ , eles devem ser entendidos como independentes, pois nem todos os traços  $\varphi$  são acionados nos casos de concordância, tal como mostram Béjar (2003), com os sistemas de concordância com múltiplos argumentos, e Baker (2013), com as diferenças de concordância entre verbos e adjetivos. Mantê-los autônomos, portanto, garante o acesso individual de cada um desses traços nas possíveis relações de dependência que possam ser estabelecidas.

O traço de borda, além de fornecer informação categorial, é responsável pela concatenação desse feixe de traços presentes no núcleo categorizador à raiz, satisfazendo a Condição do *Merge*. Já para os contextos sentenciais, sugerimos que a concatenação do feixe de traços de um núcleo categorizador se dá a partir de um traço de tempo T não-interpretável valorado [ $\mu$ T: val], o qual permite que um sintagma verbal seja categorizado em um composto (e.g., compostos V-N) e também que sentenças sejam categorizadas (e.g., do inglês, *she had a [devil-may-care approach]* ‘ela tinha uma abordagem com a qual nem o diabo pode’; *this is a [God-is-dead theology]* ‘essa é uma teologia de Deus está morto’). Esse traço satisfaz as exigências temporais do constituinte verbal, impedindo que um núcleo de tempo seja concatenado à estrutura.

Dessa forma, admitimos que um categorizador nominal pode conter os seguintes traços:

(113) *Conjunto de traços possíveis dos categorizadores nominais (n)*

- a. Traço de borda R interpretável e valorado como  $n$  [ $i$ R:  $n$ ]
- b. Traço de tempo T não-interpretável valorado [ $\mu$ T: val]
- c. Traços  $\varphi$  interpretáveis não-valorados/valorados [ $i\varphi$ : \_\_\_\_]/[ $i\varphi$ : val]
- d. Traço de classe não-interpretável valorado [ $\mu$ C: 1]<sup>136</sup>

<sup>136</sup> O traço de classe  $\mu$ C pode ter os seguintes valores: 1, 2, 3 e 4, e cada um desses valores corresponde a inserção de uma vogal temática específica, a saber /o/, /a/, /e/ e  $\emptyset$ , respectivamente, em consonância com o trabalho de Alcântara (2003, 2010).

Embora alguns traços sejam imprescindíveis no categorizador, tal como o traço de borda, nem todos precisam estar presentes em seu feixe, como o traço de tempo e o traço de classe. A ausência do traço de classe, por exemplo, é o que permite a formação de compostos com radicais, já que o componente morfológico não terá informação gramatical para a inserção de um marcador de classe, o que trará consequências em seu padrão acentual<sup>137</sup>.

A estrutura formada pelo categorizador nominal será concatenada a um núcleo determinante (D) a partir da valoração dos traços  $\phi$  de  $n$  por D<sup>138</sup>. Contudo, não é comum a presença de definitude nos nomes internos aos compostos. Se tomarmos como exemplo os compostos V-N, a presença de um determinante torna o composto agramatical, já que esse nome não pode fazer referência a uma entidade ou objeto no mundo, tal como vemos nos dados em (114):

- (114) a. \*Esse objeto é um marca-as-páginas.  
       b. \**Cet objet est un grille-le-pain.* (Francês)  
           este objeto é um torra-o-pão  
           “Esse objeto é uma torradeira”  
       c. \**Quest’oggetto è um asciuga-i-capelli.* (Italiano)  
           este objeto é um enxuga-os-cabelos  
           “Este objeto é um secador de cabelos.”

Então, como explicamos a presença de traços  $\phi$  no nome interno ao composto se parece não haver a presença de um D correspondente à sua valoração? A primeira alternativa é dizer que o núcleo determinante pode não conter a informação de definitude, mas está presente para garantir a valoração dos traços  $\phi$  de  $n$ , os quais, como vimos, podem ser distintos dos traços  $\phi$  do composto. Essa alternativa, no entanto, exige que expliquemos como os traços  $\phi$  desse D, destituído de definitude, não são preenchidos no componente fonológico.

Outra alternativa é dizer que os traços  $\phi$  dos nomes internos à composição já vêm valorados, e, por isso, um núcleo D não se concatena a eles. Nessa solução, garantimos que os

<sup>137</sup> Cf. Capítulo 6.

<sup>138</sup> Assumimos como Cyrino e Espinal (no prelo) que o traço de número é morfossintaticamente valorado em D, e, então, seu valor é atribuído ao categorizador nominal via *Agree*. Com relação ao traço de gênero, não assumimos que ele esteja na raiz, como alguns autores dentro da MD sugerem (e.g., ALCÂNTARA, 2003. 2010), mas que, tal como número, ele está valorado em D, e é atribuído ao nome via a mesma operação. Essa última assunção encontra evidências no caráter contextual que gênero pode ter (cf. LAZZARINI-CYRINO; ARMELIN; MINUSSI, 2013), o que pode ser capturado ao relegá-lo ao domínio discursivo atrelado ao núcleo D.



traços  $\phi$  do nome interno ao composto possam, em alguns contextos, ser distintos dos traços  $\phi$  do composto, ao mesmo tempo em que explicamos por que os nomes usados na composição são diferentes dos nomes utilizados isoladamente<sup>139</sup>. Acreditamos que essa segunda alternativa é a mais plausível, pois parece captar de maneira direta as informações dadas pela empiria.

Em resumo, assumimos que os traços  $\phi$  dos categorizadores nominais de nomes internos a compostos podem estar valorados, enquanto os traços  $\phi$  dos categorizadores nominais responsáveis por criar o domínio da composição não são valorados, já que o composto é visto como um nome utilizado isoladamente.

Nos núcleos determinantes D, portanto, encontramos o seguinte conjunto de traços:

(115) *Conjunto de traços dos núcleos determinantes (D)*

- a. Traços  $\phi$  interpretáveis valorados [ $u\phi$ : val]
- b. Traço de  $v$  não-interpretável não-valorado [ $uv$ : \_\_\_]

O traço de  $v$ <sup>140</sup> acima está relacionado com a marcação de Caso do DP. Tal como Wurmbbrand (no prelo), seguimos a proposta de Pesetsky e Torrego (2007) para a marcação de Caso dos DPs, em que o Caso corresponde a um traço não-interpretável de  $v$  – no caso, V para os autores – em um DP, o qual será lido como nominativo se for valorado por T, ou acusativo, se for valorado por  $v$ , tal como ilustrado abaixo:

<sup>139</sup> Legate (2010, 2012 *apud* WURMBRAND, no prelo) recorre a uma solução semelhante para diferenciar o  $v$  transitivo e o  $v$  de construções passivas. A autora propõe que o  $v$  de construções passivas é inserido com seus traços  $\phi$  valorados, o que elimina a necessidade de inserção de um argumento externo e, ao mesmo tempo, fornece a semântica de que há um sujeito implícito.

<sup>140</sup> O  $v$  está para V, e se refere ao núcleo definidor de categoria verbal, já que, em nossa abordagem, o núcleo verbal deve ser formado sintaticamente. Dessa forma, fazemos uma distinção entre  $v$  (sem itálico), núcleo definidor de categoria, e o  $v$  (italicizado), introdutor de argumento externo.

Embora já tenhamos discutido, brevemente, os traços de *v* e de *v*, ao apresentarmos como ocorre a seleção argumental na abordagem de Wurmbrand (no prelo), no início da seção 5.3, retomamos os traços já discutidos e somamos a eles os traços requeridos para a formação de palavras. Em resumo, os traços possíveis dos categorizadores verbais e de *v* ficam como segue:

(117) *Conjunto de traços possíveis dos categorizadores verbais (v)*

- a. Traço de borda R interpretável e valorado como *v* [*i*R: *v*]
- b. Traço de tempo T não-interpretável não-valorado [*u*T: \_\_\_\_]
- c. Traço de tempo T não-interpretável valorado [*u*T: val]
- d. Traços  $\phi$  não-interpretáveis não-valorado [*u* $\phi$ : \_\_\_\_]
- e. Traço de *v* não-interpretável não-valorado [*uv*: \_\_\_\_]
- f. Traço de Q não-interpretável não-valorado [*u*Q: \_\_\_\_]
- g. Traço de classe verbal não-interpretável valorado [*uCv*: 1]<sup>141</sup>

(118) *Conjunto de traços possíveis de (v)*

- a. Traço de *v* interpretável valorado [*iv*: val]
- b. Traços  $\phi$  não-interpretáveis não-valorados/valorados [*u* $\phi$ : \_\_\_\_]/[*u* $\phi$ : val]

Nesses núcleos, os traços  $\phi$  e o traço Q estão relacionados exclusivamente à estrutura argumental, enquanto os demais traços possibilitam a seleção morfológica desses núcleos. Cada um dos traços de tempo T em *v* corresponde a funções específicas: (i) se valorado, ele servirá como um categorizador verbal responsável por categorizar uma estrutura sentencial, e (ii) se não-valorado, ele corresponde ao traço de T em *v* que permite a concatenação de um núcleo de tempo T à estrutura sintática após sua valoração, tal como ocorre em (106a). Além disso, um traço T não-valorado em *v* permite a seleção de argumentos de núcleos verbais que exigem complementos com marcas temporais, por exemplo, o verbo *decide*, do inglês, que exige um [*u*T: subjuntivo] (cf. WURMBRAND, no prelo).

Por fim, elencamos os traços presentes nos categorizadores adjetivais, em (119):

<sup>141</sup> As classes verbais no português, por exemplo, são três: *-a-* (1ª conjugação), *-e-* (2ª conjugação) e *-i-* (3ª conjugação).

(119) *Conjunto de traços possíveis dos categorizadores adjetivais (a)*

- a. Traço de borda R interpretável e valorado como *a* [*i*R: *a*]
- b. Traços  $\phi$  não-interpretáveis não-valorados [*u* $\phi$ : \_\_\_\_]
- c. Traço de tempo T não-interpretável valorado [*u*T: val]

Em síntese, assumimos que os núcleos definidores de categoria são feixes de traços responsáveis por modelar a raiz ou a estrutura com a qual se combinam, não apenas adicionando informação categorial, mas um conjunto de traços que definirão essa unidade sintática, morfológica, e até mesmo fonologicamente, como veremos no Capítulo 6.

Os rótulos “núcleo definidor de categoria”, “núcleo categorizador” ou “categorizador” serão mantidos para evitar confusões terminológicas e possíveis interpretações que sugiram um rompimento com a MD. Porém, é importante que esteja claro que, para nós, esses núcleos fornecem mais informação do que apenas a categoria lexical, e o modo como seus traços constituintes estão descritos – se valorados ou não, se interpretáveis ou não – têm consequências para a construção e interpretação da estrutura sintática, tal como para a inserção de conteúdo fonológico.

### 5.3.3. Derivando compostos

Desenvolvemos, a partir de agora, o passo-a-passo da derivação de um composto subordinado, o qual será exemplificado pelos compostos V-N, e, em seguida, a derivação de um composto atributivo, representada por alguns tipos de compostos N-N. Com esses exemplos, demonstramos como o sistema de valoração de traços apresentado acima se processa em uma visão não-lexicalista de gramática na formação de palavras.

Em paralelo às derivações, retomamos os dados apresentados na subseção 4.4.1, do Capítulo 4, utilizados como argumentos para a criação de um domínio categorial, a fim de elucidar sua relevância na formação dos compostos. Esse domínio categorial, como veremos, é responsável por adicionar à estrutura um conjunto de traços gramaticais que permite explicar (i) a aparente ausência de endocentricidade sintática, (ii) a incompatibilidade entre os traços  $\phi$  contidos nos constituintes internos ao composto e aqueles pertencentes ao composto e (iii) a inserção de uma grade argumental quando o composto formado é um núcleo verbal.

### 5.3.3.1. Derivação dos compostos subordinados

Iniciamos a demonstração das derivações com os compostos subordinados V-N nominais. Esses compostos são caracterizados por apresentar um constituinte verbal e um constituinte nominal concatenados em uma estrutura sintaticamente endocêntrica, na qual se estabelece uma relação de subordinação predicado-argumento. Suas propriedades gerais estão elencadas em (120):

(120) *Propriedades gerais dos compostos V-N*

- i. São fundamentalmente nominais<sup>142</sup>;
- ii. São formados por verbos transitivos (ou transitivizados);
- iii. O constituinte nominal é interpretado como um argumento interno do verbo;
- iv. O constituinte nominal deve ser nu;
- v. São exocêntricos semanticamente;
- vi. Formam nomes com interpretação de agente ou instrumento.

Essas propriedades são atestadas em todas as línguas em que esse tipo de composição é produtivo, por exemplo, nas línguas românicas (cf. DI SCIULLO, 1992; BISETTO, 1999, para o italiano; RAINER; VARELA, para o espanhol; DESMETS; VILLOING, 2009, para o francês), e nas línguas bantas (cf. MCHOMBO, 2004, para o chichewa; BASCIANO; KULA; MELLONI, 2011, para o bemba, o swahili e o gĩkũyũ, embora nessas três últimas esses compostos sejam encontrados em menor quantidade).

Para uma explicação satisfatória da estrutura sintática dos compostos V-N, uma análise precisa contemplar os seguintes fatos empíricos, a saber: (a) como são saturados os requerimentos argumentais de seus núcleos verbais, (b) como se dá a interpretação agentiva (e.g., no PB, *desmancha-prazer*; no chichewa, *tola nkhâni* lit. pega-notícias ‘repórter’, cf. MCHOMBO, 2004, p. 117) ou instrumental desses compostos (e.g., no PB, *lustra-móveis*; no swahili, *ki-choma mguu* lit. fura-pé ‘erva com sementes farpadas’, cf. BASCIANO; KULA;

<sup>142</sup> A propriedade (120i) pode ser vista como redundante ao explicitar que a categoria do composto é nominal, já que nas línguas românicas e germânicas essa parece ser a categoria comum das palavras compostas em geral. No entanto, isso não é verificável translinguisticamente. Se tomarmos como exemplo alguns dados de compostos V-N do chinês, a categoria do composto pode ser nome ([*zhi<sub>V</sub>-ji<sub>N</sub>*]<sub>N</sub> lit. conhecer-REFL ‘amigo íntimo’), adjetivo ([*kai<sub>V</sub>-xin<sub>N</sub>*]<sub>A</sub> lit. abrir-coração ‘feliz’) ou verbo ([*zhi<sub>V</sub>-dao<sub>N</sub>*]<sub>V</sub> lit. conhecer-estrada ‘conhecer’), cf. Zhang (2007).

MELLONI, 2011 *apud* CONTINI MORAVA, 2007, p. 1131) e (c) como se dá a incompatibilidade dos traços  $\phi$  do nome interno ao composto com os traços  $\phi$  do composto<sup>143</sup>.

Com relação à (a), sobre a saturação dos argumentos exigidos pelo núcleo verbal, temos que responder à seguinte questão: visto que o verbo interno aos compostos V-N é transitivo (i.e., prevê dois argumentos), e que, superficialmente, apenas o argumento interno é saturado, como explicar a saturação – ou não – do argumento externo previsto?

Grimshaw (1990) argumenta, para esses casos, que o argumento externo não pode ser saturado internamente ao composto, caso contrário, a formação não seria um predicado bem formado. Contreras (1985 *apud* NUÑES CEDEÑO, 1991) e Di Sciullo (1992) tomam uma posição oposta a de Grimshaw (1990) e assumem que o argumento externo é saturado internamente ao composto, ou como um sujeito nulo, ou como um *pro*, respectivamente, tal como pode ser visto nas representações em (121) e (122):

(CONTRERAS, 1985)

(DI SCIULLO, 1992)

Di Sciullo (1992) argumenta que a presença de um *pro*, nos compostos do italiano, está em oposição à presença de sufixos agentivos nos compostos sintéticos do inglês, tal como o sufixo *-er* (e.g., *truckdriver* ‘motorista de caminhão’). Esses sufixos são realizações do argumento externo do núcleo verbal, e a assunção de um *pro* interno à estrutura desses compostos decorre da possibilidade de omissão do sujeito no italiano (cf. RIZZI, 1986 *apud* DI SCIULLO, 1992, p. 66-67).

Em favor de sua análise, a autora mostra que os compostos V-N do italiano não utilizam o sufixo agentivo *-tore*, semelhante ao sufixo *-er* do inglês, caso contrário o argumento externo seria preenchido mais de uma vez. Esse fato pode ser atestado na agramaticalidade dos dados em (123b) e (123c), nas combinações feitas a partir do composto *taglia-carte* lit. corta-carta ‘abridor de cartas’:

<sup>143</sup> Nas línguas bantas, a incompatibilidade  $\phi$  mencionada diz respeito às classes nominais do nome interno ao composto que não são as mesmas do composto: [(MC<sub>V</sub>) + [√V + vf] + [(MC<sub>X</sub>) + √N]] (cf. BASCIANO; KULA; MELLONI, 2011).

- (123) a. *taglia-tore* (Italiano)  
 cortar-dor  
 ‘cortador’  
 b. *\*tagliatore-carte*  
 cortador-carta  
 c. *\*carte-tagliatore*  
 carta-cortador

(DI SCIULLO, 1992, p. 72)

Além das questões argumentais, a assunção de um *pro* é responsável por garantir a interpretação agentiva ou instrumental, mencionada em (c), acima. Isso implica que a posição de sujeito sentencial deve ser preenchida por um DP vazio que carrega o papel temático de agente ou instrumento, antes que o VP seja nominalizado para a formação do composto.

As propostas de Contreras (1985) e Di Sciullo (1992) foram atualizadas no trabalho de Bok-Bennema e Kampers-Manhe (2006), pois o argumento externo deixou de ser inserido pelo núcleo verbal e passou a ser uma exigência de um núcleo funcional *v* acima de VP – nesta dissertação, *vP*. Com isso, as autoras propõem que os compostos V-N são o resultado da nominalização de um *vP*, e não mais de um VP (*vP*). Argumentos para essa análise são encontrados na transitividade verbal (i.e., trata-se de verbos transitivos) e na checagem de Caso do objeto, já que é o núcleo de *vP* que desencadeia a sonda para o Caso acusativo do complemento de *v*<sup>144</sup> – tal como propõem Pesetsky e Torrego (2007) – descrito na última seção.

Parece, portanto, ser plausível a presença de um *pro* nos compostos V-N, visto que ele garante a interpretação agentiva e instrumental do composto. Da mesma forma, a nominalização de *vP* é a mais adequada, em termos estruturais, já que explica como se dá a saturação verbal sem recorrer a bloqueios na saturação do argumento externo, garantindo, ao mesmo tempo, a marcação de Caso acusativo do complemento nominal.

Com relação à (c), ou seja, à incompatibilidade dos traços  $\phi$  do nome interno ao composto com os traços  $\phi$  do composto, apresentada na subseção 3.3.2.2, do Capítulo 3, e na subseção 4.4.1.4, do Capítulo 4, e retomada em (124), assumimos que é o domínio categorial o responsável por garantir essa diferença nos valores dos traços. Dessa forma, a presença de

<sup>144</sup> Para uma discussão mais detalhada sobre as propostas de estrutura argumental dos compostos V-N, cf. Nóbrega (no prelo).

um traço [ $i\phi$ : \_\_\_] no categorizador nominal responsável por categorizar a estrutura verbal é o que permite (i) a concatenação de um núcleo D acima do composto, via a valoração dos traços  $\phi$  do categorizador nominal, e (ii) a diferença nos valores dos traços  $\phi$  do composto, quando for o caso.

- (124) a.  $O_{\text{MASC.SG}} [\text{para-quedas}_{\text{FEM.PL}}]_{\text{MASC.SG}}$   
 b.  $O_{\text{MASC.SG}} [\text{porta-malas}_{\text{FEM.PL}}]_{\text{MASC.SG}}$   
 c.  $O_{\text{MASC.SG}} [\text{salva-vidas}_{\text{FEM.PL}}]_{\text{MASC.SG}}$   
 d.  $O_{\text{MASC.SG}} [\text{saca-rolhas}_{\text{FEM.PL}}]_{\text{MASC.SG}}$

A mesma solução é capaz de explicar a incompatibilidade das classes nominais presente em compostos V-N das línguas bantas. Essas classes nominais são associadas a gênero e número (cf. CORBETT, 1991; CARSTENS, 2008 *apud* BASCIANO; KULA; MELLONI, 2011, p. 212), e diferem internamente no composto V-N, uma vez que a classe nominal do nome interno ao composto é diferente da classe nominal do composto:

- (125) a. *m-pala-nkhâni* < *nkhâni* (Chichewa)  
 CL1-pegar-notícias < CL9-notícias  
 ‘repórter’  
 b. *chi-pha-dz’uwa* < *dz’uwa*  
 CL7-matar-sol < CL5-sol  
 ‘mulher bonita’

(MCHOMBO, comunicação pessoal)

Dessa forma, o nome interno ao composto, como vimos na subseção 5.3.2, vem com os traços  $\phi$  do categorizador já valorados, e sua concatenação ao núcleo verbal se dá pela valoração dos traços  $\phi$  não-interpretáveis não-valorados de vP, devido à Condição do *Merge*<sup>145</sup>. Esse categorizador nominal deve conter, ainda, em seu feixe de traços, um traço [ $uv$ : \_\_\_], que garante a atribuição de Caso acusativo ao nome em questão. Por outro lado, o categorizador nominal responsável por recategorizar a estrutura sintática vem com seus traços  $\phi$  não-valorados, já que o composto é visto como um nome utilizado isoladamente.

<sup>145</sup> O mesmo deve ocorrer para as línguas bantas, pois a ausência do aumento no constituinte nominal, o qual é frequentemente associado ao determinante nas línguas românicas (cf. de DREU, 2008 *apud* BASCIANO; KULA; MELLONI, 2011), exige que o nome interno venha com seus traços  $\phi$  já valorados.

Contudo, antes do núcleo verbal ser concatenado ao núcleo nominal, ambos devem ser formados no espaço computacional. Assumimos, para tanto, que a formação do núcleo verbal e a formação do núcleo nominal, ou seja, a categorização de duas raízes com determinado feixe de traços, ocorre paralelamente no plano computacional antes de serem concatenadas entre si.

Na sintaxe minimalista, Nunes (2012) mostra que o sistema computacional consegue operar com mais de uma raiz de estrutura (do inglês, *root syntactic tree*) ao mesmo tempo, devido à retirada de DS (do inglês, *deep structure*) como um nível de representação e com a assunção da Condição de Extensão (cf. CHOMSKY, 1995)<sup>146</sup>, o que possibilita a construção de objetos sintáticos paralelamente no espaço computacional. A formação de dois núcleos lexicais, nesse sentido, ocorre de modo similar à formação e concatenação de especificadores complexos a uma estrutura verbal (e.g., a concatenação de [*o garoto*] a [*comprou maçãs*]).

Admitindo que um conjunto de raízes e núcleos categorizadores esteja disponível em uma numeração para a entrada no sistema computacional, a formação e a concatenação de um núcleo verbal a um núcleo nominal ocorre da seguinte maneira:

$$(126) \text{ a. } N = \{\sqrt{A_1}; \sqrt{B_1}; n_1; v_1\}$$

$$\text{b. } N' = \{\sqrt{A_0}; \sqrt{B_1}; n_1; v_1\}$$

$$K = \sqrt{A}$$

$$\text{c. } N'' = \{\sqrt{A_0}; \sqrt{B_1}; n_1; v_0\}$$

$$K = \sqrt{A}$$

$$L = v$$

$$\text{d. } N''' = \{\sqrt{A_0}; \sqrt{B_1}; n_1; v_0\}$$

$$M = [\sqrt{A} \ v]$$

$$\text{e. } N'''' = \{\sqrt{A_0}; \sqrt{B_0}; n_1; v_0\}$$

$$M = [\sqrt{A} \ v]$$

$$O = \sqrt{B}$$

<sup>146</sup> A Condição de Extensão exige que as operações de projeção operem apenas em nós raízes.



$$f. N''' = \{\sqrt{A_0}; \sqrt{B_0}; n_0; v_0\}$$

$$M = [\sqrt{A} \ v]$$

$$O = \sqrt{B}$$

$$P = n$$

$$g. N''' = \{\sqrt{A_0}; \sqrt{B_0}; n_0; v_0\}$$

$$M = [\sqrt{A} \ v]$$

$$Q = [\sqrt{B} \ n]$$

$$h. N''' = \{\sqrt{A_0}; \sqrt{B_0}; n_0; v_0\}$$

$$R = [[\sqrt{A} \ v] [\sqrt{B} \ n]]$$

A visão sugerida acima contraria a proposta de De Belder e Van Craenenbroeck (2011) de que os núcleos lexicais, depois de formados, voltariam à numeração para serem tomados por um núcleo predicator. Em sua proposta, o sistema categorizaria uma raiz como um nome, e essa construção voltaria à numeração para esperar a formação do núcleo verbal. Acreditamos que essa alternativa é custosa para o sistema, além de ocorrer em um contexto muito específico, o que a torna pouco motivada.

No que concerne ao sistema de valoração de traços, propomos que a estrutura de um composto V-N deve ser como em (127a), na qual está representada a derivação de um composto com incompatibilidade nas marcas flexionais, tais como aqueles presentes no exemplo (124), por exemplo,  $O_{\text{MASC-SG}} \text{salva-vidas}_{\text{FEM-PL}}$ , expansível igualmente aos compostos V-N em (125). Além disso, cada morfema abstrato que compõe a estrutura sintática está apontado com seus respectivos feixes de traços na estrutura. De (127b) a (127h), estão representadas as valorações de cada passo da operação *Merge*, indicados por  $\langle \rangle$ , após o AR ter se aplicado (os traços adquiridos via valoração estão sublinhados):

- b.  $\langle n_1 + \sqrt{\phantom{x}} \rangle$  [iR: *n*, iφ: PL.FEM, uv: \_\_, uC: 2/ uR: *n*] ( $n_1P$ )
- c.  $\langle v + \sqrt{\phantom{x}} \rangle$  [iR: *v*, uφ: \_\_, uv: \_\_, uT: \_\_, uCv: 1/ uR: *v*] (*v*)
- d.  $\langle v + n_1 \rangle$  [iR: *v*, uφ: PL.FEM, uv: \_\_, uT: \_\_, uCv: 1/ iR: *n*, iφ: PL.FEM, uv: \_\_, uC: 2] ( $vP$ )
- e.  $\langle v + vP \rangle$  [iv: agente, uφ: \_\_/ iR: *v*, uφ: PL.FEM, uv: agente, uT: \_\_, uCv: 1] ( $v'$ )
- f.  $\langle pro + v' \rangle$  [iφ: 3.SG, uv: \_\_/ iv: agente, uφ: 3.SG, uT: \_\_] ( $vP$ )
- g.  $\langle n_2 + vP \rangle$  [iR: *n*, uT: val, iφ: \_\_/ iv: agente, uφ: 3.SG, uT: val] ( $n_2P$ )
- h.  $\langle D + n_2P \rangle$  [uφ: SG.MASC, uv: \_\_/ iR: *n*, uT: val, iφ: SG.MASC] (DP)

A atribuição de Caso, embora não marcada no passo-a-passo da concatenação, estabelece-se da seguinte maneira: enquanto o traço [iv: agente] em *v* é responsável pela atribuição de Caso acusativo ao complemento de *vP*, o traço de tempo [uT: val] em *n<sub>2</sub>* é o responsável por atribuir Caso nominativo a *pro*, além de valorar o traço [uT: \_\_] de *v*, garantindo a nominalização do sintagma verbal. Já o Caso do núcleo D dependerá dos traços φ que esse núcleo valorar quando for concatenado a uma estrutura sentencial, ou seja, será acusativo, se D valorar os traços φ de um *v*, ou nominativo, se D valorar os traços φ de *v*.

A incompatibilidade de traços φ encontrada em compostos N-A, tais como *casca-grossa*, *mão-aberta*, *mão-boba*, nos quais a concordância de gênero depende do referente (e.g., *o/a casca-grossa*), será explicada do mesmo modo que os compostos V-N em (127), a partir



no modo com a operação *Merge* concatena um verbo a um nome, a saber, via *pair-Merge*, em uma relação de adjunção.

Na próxima subseção, tratamos da derivação dos compostos atributivos, explorando como os modificadores são concatenados a um núcleo via *pair-Merge* para a formação de um composto.

### 5.3.3.2. Derivação dos compostos atributivos

Em contraste com os compostos subordinados, os compostos atributivos têm entre seus constituintes uma relação de modificação, a qual, em termos derivacionais, processa-se como uma adjunção. O que podemos perceber com isso é que a discussão entre argumentos *vs.* adjuntos está presente na explicação dos compostos, a partir do contraste na formação dos compostos subordinados e dos compostos atributivos.

Ao contrário dos argumentos, os adjuntos podem ser descritos como objetos sintáticos opcionais, pois não são requeridos por um núcleo lexical. Ao mesmo tempo, os adjuntos não têm participação efetiva na computação sintática, como salientam Iruntzun e Gallego (2007), pois esses objetos não formam configurações temáticas, não entram em mecanismos de atribuição de Caso, não são exigências lexicais imprescindíveis para a boa formação da sentença, entre outras propriedades. Em resumo, os adjuntos simplesmente estão “lá”.

Para explicar esse tipo de compostos, temos que levar em consideração quais são os mecanismos que licenciam adjuntos na composição. Com o intuito de fornecer uma explicação para a ocorrência de adjunção em compostos na gramática, tomamos os compostos atributivos mais representativos do PB para análise, nomeadamente, a combinação N-N.

Tendo em vista o caráter opcional dos adjuntos e a possibilidade de se concatenar um número ilimitado desses elementos a um núcleo lexical, uma primeira propriedade pode ser notada acerca desses dados: a recursividade. É apenas nos compostos atributivos que encontramos uma recursividade fornecida pela concatenação de adjuntos, semelhante àquela encontrada na adjunção de modificadores nas sentenças, tal como ilustrado no dado de Bresnan (1982 *apud* IRURTZUN; GALLEGO, 2007, p. 191), trazido em (130).

Essa recursividade é produtiva em compostos nominais atributivos, embora varie de língua para língua no modo como é externalizada. Nas línguas germânicas, por exemplo, Bisetto (2013)<sup>147</sup> mostra que os compostos nominais podem ser expandidos com ou sem a

<sup>147</sup> Agradeço à Antonietta Bisetto a disponibilização e envio do artigo antes de sua publicação.

presença de um elemento de ligação, tal como nos dados apresentados em (131), enquanto que, nas línguas românicas, essa recursividade é encontrada, em grande medida, nos compostos formados por radicais, como em (132), ou em casos paralelos, como ocorre nos compostos V-N, em que o nome interno ao composto carrega núcleos modificadores, tal como em (133).

(130) *Fred deftly handed the toy to the baby by reaching behing his back over lunch at noon in a restaurant last Sunday in Back Bay without interrupting the discussion.*

“Fred habilmente entregou o brinquedo para o bebê, passando-o por trás dele (/pelas suas costas) na hora do almoço, ao meio-dia, em um restaurante, no domingo passado, em Back Bay sem interromper a conversa.”

(131) a. *university teaching award committe member* (Inglês)

universidade.ensino.prêmio.comitê.membro

‘membro do comitê do prêmio de ensino superior’

b. *Donau. dampf. schiff.fahrt -s. gesellschaft -s. kapitän -s. mütze* (Alemão)

Danúbio.vapor.barco.viagem EL.oficial-SUF.EL. capitão EL. Cap

‘cap do capitão da companhia de barcos a vapor do Danúbio’

(NEEF, 2009)

c. *weer -s. voorspelling -s. deskundige -n. congress* (Holandês)

tempo EL. previsão EL. especialistas EL. conferência

‘conferência de especialistas em previsão do tempo’

(DON, 2009)

(132) a. aero-foto-geo-graf-ia

b. auto-foto-bio-graf-ia

c. oftalm(o)-otorrin(o)-laring(o)-log-ista

(133) *L’attentatore ha collocato l’ordigno nel piccolo vano [...] dove si trova il [porta [rotolo delle strisce di carta che si usano per coprire la tavolleta del wc]].*

“O terrorista colocou a bomba no pequeno espaço [...] onde se encontra o [porta[rolo de tiras de papel que se usa para cobrir a borda do vaso sanitário]].”

(GAETA; RICCA, 2009, p. 45)

Essa peculiaridade reitera um fato já atestado empiricamente que é, enquanto os adjuntos são, idealmente, ilimitados, os argumentos podem ser de no máximo três, tal como pontuam Hornstein e Nunes (2008, p. 60). Por esse motivo, um composto subordinado do tipo V-N não terá mais do que três constituintes, a não ser que um de seus argumentos nominais contenha um ou mais modificadores, tal como ocorre no exemplo (133).

Hornstein e Nunes (*op. cit.*) mencionam que, além da recursividade, uma segunda propriedade das estruturas de adjunção é a manutenção da informação categorial do núcleo, tal como a preservação da nuclearidade. Isso pode ser visto nos dados em (134), em que a informação categorial é mantida intacta, independentemente do número de adjuntos adicionados.

- (134) a. [VP[VP[VP comer um bolo] rapidamente] no trabalho]  
 b. [NP[NP estudante de linguística] da USP]

Contudo, nos dados de composição, a manutenção da informação categorial não é diretamente preservada em alguns dados, algo já discutido na seção 4.4.1.1, do Capítulo 4, quando apresentamos os problemas de projeção. O domínio da composição proposto nesta dissertação conserva a possibilidade de manutenção da nuclearidade nas estruturas de adjunção internas aos compostos, ao permitir que um núcleo categorial de qualquer natureza seja concatenado a uma estrutura sintática para a formação do composto, tal como ilustrado em (135):

- (135) a. [<sub>NP</sub>[<sub>NP</sub> casca] grossa]]  
 b. [<sub>NP</sub>[<sub>NP</sub> pau] mandado]]

A assunção desse domínio categorial não vem como uma alternativa local a esse problema, mas é um fato bem motivado empiricamente, como visto no Capítulo 4, além de ser responsável por garantir o caráter atômico dos compostos, ou seja, a incapacidade de apenas um dos membros ser retomado anaforicamente, ou de apenas um dos membros ser movido – tal como ocorre com a topicalização de adjuntos nas sentenças em (136), ao contrário do composto em (137):

- (136) a. ele [comeu um bolo] [no trabalho] [rapidamente].  
 b. ... e [um bolo] [no trabalho] ele comeu [rapidamente].  
 c. ... e [um bolo] [no trabalho] [rapidamente] ele comeu.

- (137) a. ele é um [casca grossa].  
 b. \*e [grossa] ele é um [casca].  
 c. e [casca grossa] ele é.

É por esse motivo que alguns dados de compostos N-prep-N ainda carecem de uma análise mais detalhada, pois embora sejam considerados uma palavra composta em algumas análises, muitos desses dados se comportam mais como sintagmas do que como compostos propriamente ditos. Assim sendo, um teste para identificação de um sintagma N-prep-N é a possibilidade de movimento de parte de seus constituintes, como ocorre nas sentenças de fala espontânea em (138), algo que não é permitido quando temos uma palavra composta, tal como em (137), acima:

- (138) a. O bolo<sub>t</sub> com leite condensado diet que eu faço é ótimo [~~bolo~~<sub>t</sub> de chocolate].  
 b. Esse patê<sub>t</sub> aqui é uma delícia [~~patê~~<sub>t</sub> de soja].  
 c. Todos os exames<sub>t</sub> que eu fiz [~~exame~~<sub>t</sub> de sangue] saíram ótimos.  
 d. Eu quero um doce<sub>t</sub> daquele [~~doce~~<sub>t</sub> de abóbora].

Para abrigar todas as propriedades supramencionadas, a formalização clássica da operação de adjunção propunha que, quando aplicada, retornava um sintagma do mesmo tipo do sintagma tido como alvo. O resultado dessa operação está representado em (139), em que uma mesma categoria é cindida em (dois) segmentos. Seu objetivo é tornar o adjunto (i.e., o YP, em (139)) estruturalmente fora da visão de operações que levem em consideração a dependência estrutural para se aplicar (e.g., dominância, c-comando, irmandade, etc.):

Na sintaxe minimalista, como vimos na seção 5.2, o sistema computacional conta com apenas uma operação responsável pela construção de estruturas sintáticas: a operação *Merge*,

a qual vem em dois sabores, *set-Merge* e *pair-Merge*. Vimos também que para os casos de adjunção, *pair-Merge* é a versão responsável pela concatenação de adjuntos à estrutura.

Diferentemente de *set-Merge*, *pair-Merge* é uma forma de concatenação que ocorre sem a valoração de um traço e que, como resultado, nenhum rótulo é criado. A assimetria inerente à adjunção é capturada pela formação de um par-ordenado, em vez de um conjunto, fazendo com que os adjuntos sejam concatenados em um “plano separado”, a fim de deixá-los fora das relações de dependência. A diferença representacional dessas operações está ilustrada em (140), abaixo:

(140) a. *set-Merge*

input:  $\alpha, \beta$

output:  $\{\alpha, \{\alpha, \beta\}\}$

b. *pair-Merge*

input:  $\alpha, \beta$

output:  $\{\alpha, \langle \alpha, \beta \rangle\}$

(IRURTZUN; GALLEG0, 2007, p. 184)

Nesse sentido, um composto atributivo, seja ele N-N ou N-A, por exemplo, deve ter, em tese, as seguintes sequências da operação *Merge* na sua formação, seguindo a representação em (141). Nela,  $\{x\}$  representa um *set-Merge*, e  $\langle x \rangle$  um *pair-Merge*:

(141)  $\{z, \langle \{x, \sqrt{\}}, \{y, \sqrt{\}} \rangle\}$

No sistema operacional empregado nesta dissertação, Wurmbrand (no prelo) sugere que os modificadores são concatenados tardiamente à estrutura sintática, como uma operação de última hora, a qual ocorre no momento da transferência da estrutura para as interfaces. No entanto, assumimos, contrariamente à suposição da autora<sup>148</sup>, que caso um núcleo complexo – depois de formado – não contenha um traço capaz de satisfazer a Condição do *Merge*, ele irá se concatenar à estrutura via *pair-Merge*, na forma de um adjunto.

<sup>148</sup> Essa proposta é defendida, propriamente, nos trabalhos de Lebeaux (1991) e Stepanov (2001).



Entretanto, antes de avaliar o momento em que o adjunto é concatenado à estrutura, temos outra questão a ser esclarecida, a saber: uma vez que a concatenação de um núcleo categorizador ocorre apenas quando há valoração de traços, tal como visto nas últimas seções, qual traço deve ser valorado a fim de permitir a concatenação de um núcleo categorizador a estruturas de adjunção para formação de um domínio?

Para respondê-la é necessário ter em mente que o adjunto é um constituinte isolado na composição, já que será concatenado em um “plano separado”, o que faz com que o categorizador responsável pelo domínio seja concatenado à estrutura a partir da valoração de algum traço do categorizador do núcleo da adjunção (i.e., do constituinte modificado), seguindo o AR. Nesse sentido, para a formação de um composto N-N como *hidromassagem*, temos, no espaço computacional, dois nomes formados paralelamente, contendo os seguintes traços em seus núcleos categorizadores:

A concatenação do núcleo categorizador do domínio ao núcleo da adjunção, para a formação de *hidromassagem*, ocorre através da valoração dos traços  $\phi$  do categorizador  $n_2$  pelos traços  $\phi$  do núcleo  $n$  do domínio. Por sua vez, os traços  $\phi$  presentes no núcleo categorizador  $n$  do domínio já terão sido valorados por D, pois como vimos na subseção 5.3.2, é o núcleo de D que carrega os valores desses traços<sup>149</sup>.

Nesse sentido, propomos que quando um núcleo D se concatena a um núcleo categorizador isoladamente, através da valoração de seus traços  $\phi$ , e, posteriormente a uma estrutura complexa formada por um nome mais um adjunto, temos um composto, tal como ocorre na derivação sequencial em (143). Por outro lado, se o núcleo de D se concatena diretamente a uma estrutura complexa formada por um nome mais um adjunto, através da valoração direta dos traços  $\phi$  do categorizador desse nome, temos um sintagma.

<sup>149</sup> Desse modo, as descrições das relações de concordância apresentadas nas subseções 3.3.2.1 e 3.3.2.2, do Capítulo 3, devem ser interpretadas pela assunção geral de que os valores dos traços  $\phi$  vêm de D, e não estão previamente valorados nos categorizadores do composto, tal como era sugerido ao dizer que sufixos, como *-ia*, bem como o sexo do referente, controlam as relações de concordância dos compostos.

$$(143) \text{ a. } N = \{n_{A1}, n_{B1}, n_{C1}, \sqrt{A_1}, \sqrt{B_1}, D_1\}$$

$$\text{b. } N' = \{n_{A1}, n_{B1}, n_{C1}, \sqrt{A_0}, \sqrt{B_1}, D_1\}$$

$$K = \sqrt{A}$$

$$\text{c. } N'' = \{n_{A0}, n_{B1}, n_{C1}, \sqrt{A_0}, \sqrt{B_1}, D_1\}$$

$$K = \sqrt{A}$$

$$L = n_A$$

$$\text{d. } N'' = \{n_{A0}, n_{B1}, n_{C1}, \sqrt{A_0}, \sqrt{B_1}, D_1\}$$

$$M = [\sqrt{A} \ n_A]$$

$$\text{e. } N''' = \{n_{A0}, n_{B1}, n_{C1}, \sqrt{A_0}, \sqrt{B_0}, D_1\}$$

$$M = [\sqrt{A} \ n_A]$$

$$O = \sqrt{B}$$

$$\text{f. } N''' = \{n_{A0}, n_{B1}, n_{C0}, \sqrt{A_0}, \sqrt{B_0}, D_1\}$$

$$M = [\sqrt{A} \ n_A]$$

$$O = \sqrt{B}$$

$$P = n_C$$

$$\text{g. } N''' = \{n_{A0}, n_{B1}, n_{C0}, \sqrt{A_0}, \sqrt{B_0}, D_1\}$$

$$M = [\sqrt{A} \ n_A]$$

$$Q = [\sqrt{B} \ n_C]$$

$$\text{h. } N'''' = \{n_{A0}, n_{B0}, n_{C0}, \sqrt{A_0}, \sqrt{B_0}, D_1\}$$

$$M = [\sqrt{A} \ n_A]$$

$$Q = [\sqrt{B} \ n_C]$$

$$R = n_B$$

$$i. N'''' = \{n_{A0}, n_{B0}, n_{C0}, \sqrt{A_0}, \sqrt{B_0}, D_0\}$$

$$M = [\sqrt{A} n_A]$$

$$Q = [\sqrt{B} n_C]$$

$$R = n_B$$

$$S = D$$

$$j. N'''' = \{n_{A0}, n_{B0}, n_{C0}, \sqrt{A_0}, \sqrt{B_0}, D_0\}$$

$$M = [\sqrt{A} n_A]$$

$$Q = [\sqrt{B} n_C]$$

$$T = [D n_B]$$

$$k. N'''' = \{n_{A0}, n_{B0}, n_{C0}, \sqrt{A_0}, \sqrt{B_0}, D_0\}$$

$$U = \langle \sqrt{B} n_C [\sqrt{A} n_A] \rangle$$

$$T = [D n_B]$$

$$l. N'''' = \{n_{A0}, n_{B0}, n_{C0}, \sqrt{A_0}, \sqrt{B_0}, D_0\}$$

$$V = [[D n_B \langle \sqrt{B} n_C [\sqrt{A} n_A] \rangle]$$

Sendo assim, a derivação dos compostos N-N atributivos *antropomorfe*, *hidroginástica* e *peixe-espada* ocorre de acordo com as estruturas em (144), (145) e (146). Nelas, é a valoração dos traços  $\phi$  do núcleo do composto que permitirá a concatenação do categorizador nominal responsável por criar o domínio. Além disso, será a presença ou ausência de um traço de classe no feixe de um categorizador nominal que desencadeará a variedade superficial notada nesses compostos, a partir da interpretação feita do feixe de traços de seus núcleos categorizadores pelo componente morfológico pós-sintaticamente.

(144) *antropomorfe*

(145) *hidroginástica* (146) *peixe-espada*

As estruturas em (144) e (145) abarcam uma das primeiras características apontadas aos compostos formados por radicais, no Capítulo 3, que é a presença de traços flexionais realizados apenas em sua borda direita (e.g., *antromorfes*, *\*antroposmorfe*), através da distribuição dos traços de classe pela estrutura. De acordo com a hipótese levantada no Capítulo 3, a impossibilidade de marcação flexional no primeiro radical desses compostos se deve à ausência de um traço de classe nos categorizadores de suas raízes, o que impede a realização fonológica dos demais traços presentes em seu feixe, na morfologia.

Dessa forma, caso o traço de classe esteja presente apenas no domínio categorial, a estrutura externalizada será [radical + radical], com traços flexionais realizados apenas em sua borda direita, tal como representado na estrutura em (144). Por outro lado, caso o traço de classe esteja marcado, exclusivamente, no categorizador de algum dos núcleos categorizadores internos ao composto (e.g., no  $n_1$ , da estrutura em (145)), a estrutura

externalizada será [radical + palavra], e será somente na borda do núcleo complexo formado por esse categorizador que os traços flexionais serão realizados fonologicamente.

Assim sendo, será apenas quando o feixe de traços de um núcleo categorizador contiver um traço de classe valorado que os traços  $\phi$  presentes em seu feixe serão realizados fonologicamente, no componente morfológico. Nesse sentido, na estrutura em (144), embora a concatenação do categorizador  $n_3$  responsável por formar o domínio tenha que valorar os traços  $\phi$  do núcleo  $n_1$  para satisfazer a Condição do *Merge*, os traços  $\phi$  do núcleo  $n_1$  não serão realizados fonologicamente após o *Spell-Out*, pois seu feixe de traços não contém um traço de classe.

A estrutura em (144) abarca, diretamente, os compostos de estrutura [radical + radical] do grego moderno, vistos na subseção 4.4.1.2, do Capítulo 4, e os compostos parassintéticos vistos na subseção 4.4.1.3, do Capítulo 4. Nesses dados, a informação flexional do composto é concatenada apenas em sua borda direita, tal como pode ser verificado com os dados em (80), repetidos parcialmente em (147), o que, de acordo com nossa hipótese, ocorre devido à presença de um traço de classe apenas no categorizador do domínio<sup>150</sup>.

- |                                                                                                                                                                                                                                                    |                        |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| <p>(147) a. <i>nixt-o-lúluδ-o</i><br/>                   noite-EL-flor-INFL<br/>                   ‘flor da noite’<br/>               b. <i>riz-ó-γal-o</i><br/>                   arroz-EL-leite-INFL<br/>                   ‘pudim de arroz’</p> | <p>(Grego Moderno)</p> |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|

Por sua vez, se ambos os categorizadores dos membros de um composto contiverem um traço de classe, o que temos é um composto como *peixe-espada*, em (146). É importante mencionar que o nome adjungido nesses compostos, no caso, *espada*, normalmente se comporta como os nomes internos a compostos V-N, em que seus traços  $\phi$  já vêm valorados da numeração, o que explica o fato deles não concordaram com o núcleo (e.g., *peixes-espada*, \**peixes-espadas*), tal como discutido no Capítulo 3. Nos casos em que a concordância com o núcleo é possível (e.g., *navios-escola*, *navios-escolas*), admitimos que, uma vez valorados os

<sup>150</sup> Vale ressaltar que, ao contrário do PB, o grego moderno tem sufixos flexionais *portmanteau*, em que um único expoente fonológico denota, ao mesmo tempo, as informações de Caso, número e classe (cf. ALEXIADOU, 2004).

traços  $\varphi$  do núcleo do composto pelo categorizador do domínio, esses valores serão percolados ao núcleo categorizador do nome adjungido.

Demonstramos, dessa forma, como ocorre a formação dos compostos atributivos na sintaxe, e como a configuração dos feixes de traços de seus categorizadores pode influenciar no modo como sua estrutura sintática será externalizada, bem como onde seus traços  $\varphi$  serão realizados fonologicamente. Acreditamos que as estruturas descritas acima podem ser estendidas, de modo uniforme, aos compostos de realização [radical + radical], [radical + palavra] e [palavra + palavra] nas línguas naturais, pois será a partir dessa configuração de traços que obteremos tal variedade superficial nos compostos.

#### 5.4. Conclusões preliminares

Apresentamos, neste capítulo, como ocorre a formação dos compostos subordinados e atributivos no componente sintático, a partir de um sistema de valoração de traços baseado na Condição do *Merge*, de acordo com a proposta de Wurmbrand (no prelo). Nele, argumentamos que (i) as relações gramaticais são diretamente explicadas a partir da natureza da operação *Merge* que venha a ocorrer entre dois núcleos complexos, e que (ii) a criação de um domínio na composição se dá com a concatenação de um núcleo categorizador acima de uma estrutura sintática.

Explicitamos, além disso, quais são os traços que podem compor os núcleos categorizadores, bem como o que permite sua concatenação a uma raiz no espaço computacional para formação de palavras compostas. Além disso, com o intuito de manter uma uniformidade operacional na sintaxe, argumentamos que a raiz contém um traço de borda capaz de viabilizar a sua concatenação a um feixe de traços gramaticais, permitindo, dessa forma, que ela seja manipulada no espaço computacional para a formação de palavras.

## CAPÍTULO 6

### NO CAMINHO PARA PF

Depois de apresentadas as estruturas morfológicas da composição no PB e o modo como os compostos são formados – levando-se em consideração informações translinguísticas para uma análise sintática do fenômeno – retornamos aos assuntos específicos às línguas, a saber, os aspectos fonológicos da composição, ao mostrar como ocorre a aplicação das regras de acento nos três tipos estruturais de compostos do PB. Em paralelo, discutimos os descompassos entre o modo como a fonologia prosódica (cf. NESPOR; VOGEL, 1986) descreve a formação do domínio da palavra fonológica e sua adequação em um modelo não-lexicalista de gramática com inserção tardia para a fonologia, tal como assumido pela MD.

#### 6.1. Caracterização do capítulo

Após termos visto as estruturas morfológicas dos compostos, ao descrever os padrões estruturais do PB, e o modo como são derivados no componente sintático, passamos, agora, a discutir outra informação específica às línguas, a saber: os aspectos fonológicos da composição. Retomando os compostos do PB, nossa língua específica, iremos nos concentrar em um aspecto fonológico frequentemente reiterado acerca desses dados, nomeadamente, a aplicação das regras de acento, mostrando como ela se distribui em suas três estruturas morfológicas.

Com isso, visamos explorar a interface sintaxe-fonologia e o modo como a estrutura sintática na formação de palavras é mapeada em uma estrutura fonológica. Baseando-nos, principalmente, nos trabalhos de Nespor e Vogel (1986), Nespor e Ralli (1996), Lee (1995, 1997), Moreno (1997), Nespor (1999) e Vigário (2003), colocamos em discussão o descompasso entre a fonologia prosódica e a assunção de estruturas sintáticas como estruturas morfológicas, tal como assumido no modelo não-lexicalista da MD.

O descompasso mencionado acima diz respeito às assunções feitas pela fonologia prosódica sobre o modo como a palavra morfológica é mapeada em uma palavra fonológica, e

também às informações morfológicas relevantes para o mapeamento do padrão acentual dos compostos. Na maioria dessas abordagens, a palavra fonológica é construída a partir do estatuto morfológico dos constituintes de um composto (i.e., se radicais ou palavras) para, então, receber a aplicação da regra de acento. Essa visão entra em descompasso com os primitivos teóricos assumidos pela MD, pois essa última não se vale das mesmas noções empregadas pelas demais teorias morfológicas, e o que pode ser, grosso modo, chamado de palavra é construído em termos sintáticos.

Nesse contexto, buscamos fornecer evidências para um modelo não-lexicalista de gramática ao assumir que a atribuição de acento nos compostos não precisa, necessariamente, levar em consideração o estatuto morfológico das noções da morfologia tradicional (i.e., raiz, radical, palavra), como proposto por Nespor e Ralli (1996), Pepperkamp (1997), Nespor (1999) e Vigário (2003), mas que tais noções são epifenomenais, uma vez que elas correspondem ao modo como o feixe de traços morfossintáticos do categorizador de uma raiz sintática é preenchido fonologicamente.

Para governar a mediação entre a estrutura sintática e a estrutura fonológica, assumimos que a noção de fase no nível da palavra (cf. MARANTZ, 2001, entre outros) é responsável por demarcar o conteúdo da derivação enviado para a forma fonológica (PF) e, assim, permitir a aplicação das regras de acento nas estruturas geradas. Essa proposta vem na esteira de outros trabalhos que fizeram uso da teoria de fase no mapeamento prosódico, a saber, Marvin (2003) e Lowenstamm (2010), para a palavra; e Kahnemuyipour (2004, 2009), Kratzer e Selkirk (2007) e Irwin (2011), na sentença.

O capítulo segue dividido da seguinte maneira: na seção 6.2, apresentamos as noções que fundamentam a fonologia prosódica, pontuando as assunções teóricas que conflitam com um posicionamento não-lexicalista, a fim de sugerir, no decorrer da discussão, como a fonologia prosódica pode ser integrada a essa visão de gramática. Em seguida, focalizamos a construção da palavra fonológica, com o intuito de assinalar suas relações com a composição e seus pontos de conflito com os pressupostos da MD.

Na seção 6.3, expomos duas grandes diferenças descritivas elaboradas para a divisão dos compostos e seus reflexos na aplicação do acento, mostrando que, nos debates sobre a distribuição acentual dos compostos do PB (e.g., LEE, 1995, 1997; MORENO, 1997), os compostos de estrutura [radical + radical] (e.g., *cervejólogo*, *historiografia*) e [radical + palavra] (e.g., *neurocirurgia*, *psicoafetivo*) foram pouco explorados. Por esse motivo, desenvolvemos a aplicação de testes fonológicos aos seus dados, com o intuito de demonstrar como esses dois tipos de compostos são mapeados prosodicamente.



Na seção 6.4, recuperamos os processos que compõem o ramo de PF para a MD – apresentados, primeiramente, no Capítulo 2 – e sinalizamos as consequências que a assunção da fase no nível da palavra traz para o modelo. Nessa discussão, as questões de um único ou de múltiplos *Spell-Outs*, tal como a defesa de um mapeamento indireto das estruturas sintáticas, entram em pauta, a partir de um balanço das propostas de Arad (2003, 2005), Marvin (2003), Lowenstamm (2010) e Embick (2010), principalmente.

Em seguida, na seção 6.5, desenvolvemos nossa análise, mostrando que uma tipologia de núcleos categoriais baseada em traços, tal como a desenvolvida no Capítulo 5, é capaz de diferenciar a variedade superficial dos compostos do PB. Juntamente com essa discussão, demonstramos como as regras de acento são aplicadas nos compostos dessa língua, assumindo, para tanto, a proposta acentual de Lee (1995), e a traduzindo na proposta de fase fornecida em Embick (2010). Por fim, na seção 6.6, reunimos as conclusões preliminares fornecidas no capítulo.

## **6.2. Ausência de isomorfia entre a sintaxe e a fonologia: a Fonologia Prosódica**

Iniciamos a discussão, apresentando, na subseção 6.2.1, os princípios que governam a fonologia prosódica, tomando por base, principalmente, as propostas de Nespor e Vogel (2007 [1986]). Na subseção 6.2.2, concentramo-nos na criação do domínio da palavra fonológica e sua relevância no mapeamento prosódico de palavras compostas, com vias a elaborar um conjunto de questões que devem ser explicadas por um trabalho que ousa assumir uma visão de domínios prosódicos em um modelo não-lexicalista.

### **6.2.1. Os domínios prosódicos**

A fonologia prosódica, tal como delineada em Nespor e Vogel (2007 [1986]), opera com a ideia principal de que a fala é construída hierarquicamente em constituintes prosódicos, vistos como domínios para a aplicação de regras fonológicas específicas. Esses constituintes prosódicos são motivados pela ausência de isomorfia entre os elementos formados nos componentes gerativos da gramática<sup>151</sup> e aquilo que o componente fonológico interpreta como domínio para a aplicação de suas regras.

---

<sup>151</sup> Tendo em vista que Nespor e Vogel (2007 [1986]) assumem um componente lexical gerativo.

Essa visão de domínios específicos para a fonologia vem contra modelos que assumem um acesso direto às informações morfossintáticas ou sintáticas pelo componente fonológico. Um exemplo típico das teorias de acesso direto é a apresentada em *Sound Patterns of English* (doravante, SPE; CHOMSKY; HALLE, 1968), que propunha que a aplicação de regras fonológicas tomasse como *input* a estrutura sintática. No exemplo em (148), tirado de Chomsky e Halle (1968, p. 13), a estrutura sintática é diretamente acessada pelo componente fonológico, que, com base na própria estrutura sintática, estabelece os domínios para aplicação de regras fonológicas.

(148)

(CHOMSKY; HALLE, 1968, p. 13)

Uma teoria de acesso direto, como a assumida em SPE, não consegue abarcar os casos em que a aplicação de regras fonológicas desrespeita os domínios sintáticos. Um exemplo é a aplicação da regra fonológica *linking-r*, presente em dialetos não-róticos do inglês britânico, a qual foi apresentada por Nespor e Vogel (2007 [1986]) como um claro exemplo da irrelevância da estruturação sintática para sua aplicação, pois ocorre não somente entre palavras de uma mesma sentença (149a), como também entre palavras de sentenças diferentes (149b):

(149) a. *I saw him fa[r] away.*

(BROADBENT, 1991, p. 282)

b. *There's my mothe[r], I've got to go.*

(NESPOR; VOGEL, 2007 [1986], p. 4)

Nos casos em que a aplicação da regra fonológica desconsidera as fronteiras estabelecidas pela estrutura sintática, as propostas de mapeamento direto, como a de SPE, perdem a plausibilidade. Na regra apresentada acima, não há um constituinte sintático que agrupe as duas sentenças a fim de desencadear a aplicação da regra em (149b). Fatos como o de *linking-r* sugerem que o sistema não é homogêneo e, por conseguinte, há a necessidade do estabelecimento de outros domínios para a aplicação das regras fonológicas.

Nesse contexto, Nespore e Vogel (2007 [1986]) definem um conjunto de domínios para compor a aplicação das regras fonológicas, disposto de modo hierárquico, como pode visto na representação retirada de Bisol (1999), abaixo:

Essa hierarquia ficou conhecida como a hierarquia prosódica, e foi, desde sua elaboração, assumida como uma escala universal. Dessa forma, é esperado que, em todas as línguas naturais, cada um desses constituintes fonológicos – e somente esses – tenham influência na fonologia<sup>152</sup>. Nespore e Vogel (2007 [1986], p. 11-12), por outro lado, salientam

<sup>152</sup> Há, no entanto, divergências entre alguns autores sobre a motivação real para alguns desses domínios. Por exemplo, o Grupo Clítico (C) é um constituinte que gera intenso debate sobre sua presença na hierarquia prosódica, recebendo diversos contra-argumentos por parte de alguns autores (cf. SELKIRK, 1984; VIGÁRIO, 2003). Outro ponto de discussão é qual constituinte é o mais baixo na hierarquia prosódica, se é a sílaba (cf. SELKIRK, 1984; NESPOR; VOGEL, 2007 [1986]; VIGÁRIO, 2003), ou a palavra fonológica (cf. HAYES, 1989 *apud* VIGÁRIO, 2003).

que a ausência de regras que fazem referência a um domínio prosódico específico em uma língua particular não significa, necessariamente, que o domínio em questão não exista, pois ou (i) uma regra que faz referência a esse domínio ainda não foi encontrada, ou (ii) aquele domínio não é relevante para seus padrões fonológicos (cf. SELKIRK, 1980 *apud* NESPOR; VOGEL, 2007 [1986], p. 11).

Ainda na esteira da universalidade, a hierarquia prosódica é governada por quatro princípios gerais, listados em (151), e a estrutura interna de cada constituinte da hierarquia é caracterizada por uma mesma configuração geométrica alcançada por meio de regras de formação para cada domínio prosódico, as quais seguem a regra geral dada em (152), abaixo, conforme sugerem Nespor e Vogel (2007 [1986], p. 7):

(151) *Princípios que regem a hierarquia prosódica (NESPOR; VOGEL, 2007[1986], p. 7, tradução nossa)*

*Princípio 1:* uma dada unidade não-terminal da hierarquia prosódica,  $X^P$ , é composta de uma ou mais unidades da categoria imediatamente abaixo,  $X^{P-1}$ .

*Princípio 2:* uma unidade de um dado nível da hierarquia está exaustivamente contida na unidade super-ordenada da qual é parte.

*Princípio 3:* as estruturas hierárquicas da fonologia prosódica são ramificações n-árias.

*Princípio 4:* a relação de proeminência relativa definida por nós irmãos é tal que para determinado nó é atribuído um valor forte e para todos os demais um valor fraco<sup>153</sup>.

(152) *Construção do constituinte prosódico (NESPOR; VOGEL, 2007[1986], p. 7, tradução nossa)*

Una em uma ramificação n-ária  $X^P$  todos os  $X^{P-1}$  incluídos na sequência delimitada na definição do domínio de  $X^P$ <sup>154</sup>.

<sup>153</sup> Trecho original (NESPOR; VOGEL, 2007 [1986], p. 7): “*Principle 1.* A given nonterminal unit of the prosodic hierarchy,  $X^P$ , is composed of one or more units of the intermediate lower category,  $X^{P-1}$ ; *Principle 2.* A unit of a given level of hierarchy is exhaustively contained in the superordinate unit of which it is part; *Principle 3.* The hierarchical structures of prosodic phonology are n-ary branching; *Principle 4.* The relative prominence relation defined for sister nodes is such that one node is assigned the value strong (s) and all the other nodes are assigned the value weak (w).”

<sup>154</sup> Trecho original (NESPOR; VOGEL, 2007 [1986], p. 7): “Prosodic Constituent Construction: Join into an n-ary branching  $X^P$  all  $X^{P-1}$  included in a string delimited by the definition of the domain of  $X^P$ .”

Os princípios apresentados acima asseguram as condições de boa-formação das estruturas prosódicas, restringindo um número ilimitado de possíveis estruturas. Tais princípios formulados por Nespor e Vogel (2007[1986], p. 7) recuperam as condições de boa-formação das representações prosódicas propostas, inicialmente, por Selkirk (1984), as quais são conhecidas pelo rótulo de *Strict Layer Hypothesis* (SLH). As estruturas abaixo, retiradas de Ladd (1996 *apud* VIGÁRIO, 2003, p. 20), ilustram algumas das configurações barradas pela SLH:

(153) a. Não há dominação múltipla<sup>155</sup>

b. Não há nós irmãos distintos<sup>156</sup>

b. Não é possível pular níveis<sup>157</sup>

d. Não há recursividade<sup>158</sup>

A construção desses domínios prosódicos foi vista, em um primeiro momento, ocorrendo pós-lexicalmente. No entanto, Nespor e Vogel (*op. cit.*) já sinalizavam que a construção da palavra fonológica teria problemas em seu mapeamento caso ocorresse pós-lexicalmente, pois a criação desse domínio, em específico, deveria fazer referência a certas informações morfológicas que já não estariam presentes nesse ponto da derivação (assumindo-se o modelo em Y da gramática). Uma solução, portanto, foi assumir que a construção de domínios mais baixos na hierarquia prosódica, tais como a sílaba ( $\sigma$ ), o pé ( $\Sigma$ ) e a palavra fonológica ( $\omega$ ), deveria ocorrer no componente lexical.

Nesse sentido, a fonologia prosódica foi adaptada à concepção de gramática proposta pela Fonologia Lexical (doravante, FL; cf. KIPARSKY, 1982; MOHANAN, 1982), na qual a gramática é composta de dois componentes básicos: o nível lexical e o nível pós-lexical. Essa organização estabelecia duas grandes classes de processos fonológicos – os lexicais e os pós-lexicais – que se aplicavam a depender do local da arquitetura da gramática em que o domínio prosódico era criado.

<sup>155</sup> E.g., dois pés ( $\Sigma$ ) formados por uma sílaba ( $\sigma$ ) em comum.

<sup>156</sup> E.g., uma palavra fonológica ( $\omega$ ) formada por um pé ( $\Sigma$ ) e uma sílaba ( $\sigma$ ) que estejam em relação de irmandade.

<sup>157</sup> E.g., uma palavra fonológica ( $\omega$ ) construída a partir de duas sílabas ( $\sigma$ ).

<sup>158</sup> E.g., uma palavra fonológica ( $\omega$ ) formada por outras duas palavras fonológicas.

No modelo da FL, tanto a morfologia quanto uma parte da fonologia operam no nível lexical, e suas operações devem estar interligadas, já que o resultado das operações morfológicas serve como *input* para a aplicação das regras fonológicas, tal como pode ser visto na representação da arquitetura da gramática ilustrada abaixo, retirada de Booij e Rubach (1987, p. 3 *apud* VIGÁRIO, 2003, p. 12):

**Figura 5** - Arquitetura da gramática para a Fonologia Lexical

No entanto, contrariamente à FL, o modelo não-lexicalista de gramática assumido nesta dissertação, propõe que as palavras são geradas pelo componente sintático e a inserção de conteúdo fonológico ocorre pós-sintaticamente, como descrito no Capítulo 2. Consequentemente, o local da gramática em que se processa a construção dos domínios prosódicos deve ser estritamente pós-sintático nesse modelo. A partir disso, uma questão se coloca: a criação de domínios prosódicos ocorre simultaneamente à inserção de conteúdo fonológico (i.e., simultaneamente à inserção dos IVs) ou somente quando a estrutura sintática já foi linearizada?

Como veremos na seção 6.4, a estrutura sintática que compõe a palavra é enviada parcialmente para o componente fonológico, devido à aplicação de fases durante a sua derivação. Em razão disso, como o componente fonológico desenvolve a criação de domínios prosódicos? É possível assumir que há regras de aplicação de acento que ocorrem de modo cíclico pós-sintaticamente? Ou, contrário às assunções da fonologia prosódica, a MD é um modelo que exige um mapeamento direto da estrutura sintática – tanto de palavras quanto de sentenças – pelo componente fonológico?

Antes de desenvolver essas questões, iremos discutir a formação de um domínio prosódico comumente relacionado aos dados de composição, a saber, a palavra fonológica ( $\omega$ ), comentando, em paralelo, como a distribuição e o número de acentos primários são relevantes para sua delimitação. Perceberemos, mais adiante, que essas definições são essenciais para a compreensão dos problemas aqui levantados, pois elas desencadeiam uma série de questões para uma proposta de domínios prosódicos em um modelo de inserção tardia para a fonologia.

### 6.2.2. A palavra fonológica e a composição

Quando o assunto é a prosódia em palavras compostas, um dos primeiros julgamentos que vem à mente para a identificação do número de palavras fonológicas formadas é o número de acentos principais. Esse acento principal, comumente conhecido como acento primário, é um dos diagnósticos mais intuitivos para a identificação da palavra fonológica. Como aponta Vigário (2003, p. 23), é geralmente aceito que a palavra fonológica deve carregar um e apenas um acento primário. Com isso, podemos estabelecer uma generalização acerca da identificação desse domínio:

(154) *Primeiro critério para a identificação de uma palavra fonológica ( $\omega$ )*

Uma palavra fonológica ( $\omega$ ) deve carregar um e apenas um acento primário.

Dado que a palavra fonológica é sempre o domínio prosódico focalizado quando o assunto é a composição (cf. SELKIRK, 1984; NESPOR; VOGEL, 1986; NESPOR; RALLI, 1996; NESPOR, 1999, dentre outros), iremos comentar as principais características de seu mapeamento e como ocorre a sua criação. De imediato, podemos depreender, a partir do que já foi exposto até aqui, que a palavra fonológica não espelha diretamente a palavra morfológica.

Câmara Jr. (2010 [1971], p. 47), por exemplo, já assinalava dessemelhanças na relação entre a palavra morfológica e a palavra fonológica, apresentando como exemplos os clíticos, exemplificados em (155), em que duas unidades formais criam uma única unidade fonológica, e as palavras compostas, em (156), em que duas palavras fonológicas correspondem a uma única unidade morfológica<sup>159</sup>:

<sup>159</sup> Há que se mencionar que, no estruturalismo, os morfológos já haviam notado que o domínio morfológico e o domínio fonológico de uma língua podem ser discrepantes, como aponta Moreno (1997, p. 17). Por sua vez, o

(155) a. [feríu-se]<sub>ω</sub><sup>160</sup>

b. [o bélo]<sub>ω</sub>

(156) a. [guárda]<sub>ω</sub>[chúva]<sub>ω</sub>

b. [cóna]<sub>ω</sub>[gótas]<sub>ω</sub>

Nespor e Vogel (2007 [1986]), dentro de uma abordagem gerativista, argumentam que o domínio da palavra fonológica pode ser menor ou igual a um nó terminal sintático e essa distribuição varia de língua para língua<sup>161</sup>. Por exemplo, enquanto no grego moderno e no latim, o domínio da palavra fonológica coincide com o nó sintático, no turco e no sânscrito o domínio fonológico é menor que um nó sintático.

Desse modo, partindo de regras fonológicas que se aplicam, necessariamente, dentro do domínio da palavra fonológica nas línguas supramencionadas<sup>162</sup>, e visando explicitar os diferentes modos de criação da palavra fonológica translinguisticamente, as autoras propõem as seguintes definições para a criação de seu domínio, as quais subdividem as línguas em dois tipos:

(157) *Domínio da palavra fonológica (ω)*

*Tipo Linguístico 1*

O domínio da palavra fonológica é igual a Q (Q = nó terminal sintático).

ou

*Tipo Linguístico 2*

I. O domínio da palavra fonológica consiste de:

- a. um radical;
- b. qualquer elemento identificado por critérios fonológicos e/ou morfológicos;
- c. qualquer elemento marcado com os diacríticos [+W].

---

trabalho de Nespor e Vogel (2007 [1986]), que vinha na esteira de outras discussões sobre o assunto, traz uma formalização dessas atestações empíricas dentro de um modelo gerativo de gramática.

<sup>160</sup> O acento primário está sendo marcado por um acento agudo na vogal/sílaba relevante.

<sup>161</sup> Alguns autores afirmam que a palavra fonológica pode ser considerada maior se levarmos em consideração os clíticos, principalmente nas propostas em que o domínio do Grupo Clítico (C) está ausente.

<sup>162</sup> As regras fonológicas específicas à palavra fonológica nas línguas analisadas são: (i) assimilação da nasal no grego moderno, (ii) regra de acento principal no latim, (iii) vozeamento final no sânscrito e (iv) harmonia vocálica no turco (cf. NESPOR; VOGEL, 2007 [1986], p. 110-122).



II. Quaisquer elementos soltos dentro de Q fazem parte da palavra fonológica adjacente mais próxima ao radical. Se não há uma palavra fonológica próxima, eles formam uma palavra fonológica por si só<sup>163</sup>.

Nespor e Ralli (1996) e Nespor (1999) mostram, posteriormente, que há compostos no grego moderno que não são correspondem ao nó sintático, diferentemente do que previa Nespor e Vogel (2007 [1986]) ao classificar essa língua como pertencente ao Tipo Linguístico I, em (157). De acordo com Nespor e Ralli (1996) e Nespor (1999), há uma flutuação no número de palavras fonológicas formadas a partir de um único nó sintático, já que no grego moderno, tal como no italiano, é possível formar tanto uma única palavra fonológica a partir de seus compostos, como em (158), quanto duas palavras fonológicas, como em (159), e essa flutuação no número de palavras fonológicas decorre da estrutura morfológica do composto:

- |          |                   |   |                 |              |
|----------|-------------------|---|-----------------|--------------|
| (158) a. | <i>kuklóspito</i> | < | <i>kúkla</i>    | <i>spíti</i> |
|          | ‘casa de boneca’  |   | ‘boneca’        | ‘casa’       |
| b.       | <i>sociólogo</i>  | < | <i>soci-(o)</i> | <i>log-</i>  |
|          | ‘sociólogo’       |   | ‘soci-VL’       | ‘logo’       |
- 
- |          |                      |   |              |                          |
|----------|----------------------|---|--------------|--------------------------|
| (159) a. | <i>zóni asfalías</i> | < | <i>zóni</i>  | <i>asfalías</i> (GEN.SG) |
|          | ‘cinto de segurança’ |   | ‘cinto’      | ‘segurança’              |
| b.       | <i>pórtagióie</i>    | < | <i>pórta</i> | <i>gióie</i>             |
|          | ‘porta-jóias’        |   | ‘porta’      | ‘jóias’                  |

Apesar de não termos detalhado a argumentação de Nespor e Vogel (2007 [1986]) nos desmembramentos propostos para a formação da palavra fonológica nas línguas do Tipo Linguístico 2, o que queremos enfatizar de sua proposta é que as regras de mapeamento que definem a palavra fonológica representam a interação entre a morfologia e a fonologia. Ou seja, elas reagrupam elementos terminais da estrutura morfológica de modo a que as

<sup>163</sup> Trecho original (NESPOR; VOGEL, 2007 [1986], p. 141): “ $\omega$  domain. A. The domain of  $\omega$  is Q. Or. B. I. The domain of  $\omega$  consists of: a stem; b. any element identified by specific phonological and/or morphological criteria; c. any element marked with the diacritic [+W]. II. Any unattached elements within Q form part of the adjacent  $\omega$  closest to the stem; if no such  $\omega$  exists, they form a  $\omega$  on their own”.

estruturas resultantes não tenham que corresponder, necessariamente, aos constituintes morfológicos.

Nisso, encontramos um descompasso entre a maneira como o mapeamento da palavra fonológica foi proposto e a arquitetura não-lexicalista da gramática assumida nesta dissertação. Enquanto a palavra fonológica representa a interação entre o componente morfológico e componente fonológico (cf. NESPOR; VOGEL, 2007 [1986], p. 109), a frase fonológica é o domínio prosódico que representa a relação entre o componente sintático e o componente fonológico.

No entanto, se, na MD, as palavras são construídas sintaticamente, de acordo com o primeiro pressuposto global que governa esse modelo – *Syntax all the way down*<sup>164</sup> – como deve ocorrer o mapeamento da estrutura sintática da palavra em uma estrutura fonológica? Além disso, como as informações supra-segmentais, mais especificamente, a atribuição do acento primário, ocorre em um modelo de inserção tardia para a fonologia? E, por fim, em que ponto do caminho para PF a construção dos domínios prosódicos se processa?

Embora a definição da palavra fonológica ainda seja um assunto em aberto, como aponta Vigário (2003), as perguntas elencadas acima são de extrema importância caso seja assumida a formação de domínios prosódicos na MD. Idealmente, um modelo que assume a formação de palavras ocorrendo a partir das mesmas operações sintáticas e com o mesmo modo de estruturação, é necessário que haja um algoritmo capaz de fazer referência à estrutura sintática da palavra e à criação da palavra fonológica correspondente, tal como ocorre no nível sentencial para a criação da frase fonológica.

Antes de apresentar a distribuição de acento nos compostos, queremos salientar que não há um consenso na literatura sobre qual domínio prosódico é responsável pelas regras fonológicas relacionadas à composição. Vogel (1990) assume que a prosodização dos compostos varia de acordo com o modo como as línguas constroem o domínio do Grupo Clítico (C). Nespor e Ralli (1996) e Nespor (1999), por outro lado, assumem que os domínios prosódicos relevantes para a prosodização dos compostos são a palavra fonológica e a frase fonológica, enquanto que Vigário (2003) argumenta pela presença de um domínio adicional – a *palavra prosódica máxima* ( $\omega_{max}$ ) –, o qual está superordenado à combinação de duas palavras fonológicas, e é responsável por abrigar as palavras compostas.

Em nossa discussão, exploramos a visão de Nespor e Ralli (1996) e Nespor (1999), por considerar que sua proposta conta com maior cobertura empírica, e por querer endossar

---

<sup>164</sup> Cf. Capítulo 2, exemplo (5).

a manutenção da não-recursividade na estrutura prosódica, ou seja, não permitir que determinado elemento domine outros de mesma categoria (e.g., duas palavras fonológicas sendo dominadas por outra palavra fonológica, cf. exemplo (153d)), tal como ocorre na proposta de Vigário (2003).

Como veremos, na próxima seção, a palavra fonológica é o domínio prosódico mais reclamado quando o assunto é a atribuição de acento primário; porém, a aplicação de acento nos compostos pode ser dividida em duas abordagens: (i) a abordagem estrutural, a qual toma por base a estrutura morfológica do composto e (ii) a abordagem em níveis lexicais, em que a aplicação de acento leva em consideração o nível lexical em que o composto é formado.

### 6.3. Os padrões acentuais na composição

Nesta seção, expomos as diferentes propostas para os padrões acentuais nos compostos. O que veremos é que, enquanto alguns autores utilizam a estrutura morfológica do composto para o mapeamento do número de palavras fonológicas formadas, tal como Nespor e Ralli (1996), Nespor (1999) e Vigário (2003), outros, como Lee (1995, 1997), fazem uso de critérios baseados nos processos fonológicos e morfológicos que ocorrem nos compostos, com o intuito de mapear o local de sua formação na gramática (i.e., se são formados no léxico ou na sintaxe) e, então, distribuir sua aplicação de acentos.

#### 6.3.1. Estruturas morfológicas governando a aplicação de acento

As propostas que levam em consideração a estrutura morfológica do composto para a distribuição do acento, e, por conseguinte, para a criação da palavra fonológica, iniciam-se com o trabalho de Nespor e Ralli (1996) sobre os compostos do grego moderno. As autoras descrevem três tipos estruturais na língua e cada tipo apresenta uma distribuição acentual distinta, a saber:

(160) *Distribuição acentual nos compostos do grego moderno (NESPOR; RALLI, 1996)*

##### *Tipo 1*

- a. Compostos [radical + radical]: um único acento é atribuído na antepenúltima sílaba do composto;

*Tipo 2*

- b. Compostos [radical + palavra]: um único acento é atribuído ao composto e está localizado no mesmo ponto em que é atribuído à palavra – segundo membro do composto – em contexto isolado;

*Tipo 3*

- c. Compostos [palavra + palavra]: há a presença de dois acentos principais, e ambos caem no mesmo ponto em que são atribuídos às palavras quando pronunciadas isoladamente.

## (161) Tipo 1 [radical + radical]

|      |                    |   |               |             |
|------|--------------------|---|---------------|-------------|
| i.   | <i>anθókipos</i>   | < | <i>anθ-</i>   | <i>kip-</i> |
|      | ‘jardim de flores’ |   | ‘flor’        | ‘jardim’    |
| ii.  | <i>tirópita</i>    | < | <i>tir-</i>   | <i>pit-</i> |
|      | ‘torta de queijo’  |   | ‘queijo’      | ‘torta’     |
| iii. | <i>karavópano</i>  | < | <i>karav-</i> | <i>pan-</i> |
|      | ‘lona’             |   | ‘barco’       | ‘pano’      |

Tipo 2 [radical + palavra]<sup>165</sup>

|     |                     |   |               |                |
|-----|---------------------|---|---------------|----------------|
| iv. | <i>kreatayorá</i>   | < | <i>kreat-</i> | <i>ayorá</i>   |
|     | ‘mercado de carne’  |   | ‘carne’       | ‘mercado’      |
| v.  | <i>tamosaláta</i>   | < | <i>taram-</i> | <i>saláta</i>  |
|     | ‘salada de caviar’  |   | ‘caviar’      | ‘salada’       |
| vi. | <i>elafokiniyós</i> | < | <i>elaf-</i>  | <i>kiniyós</i> |
|     | ‘caçador de cervos’ |   | ‘cervo’       | ‘caçador’      |

## Tipo 3 [palavra + palavra]

|      |                      |   |                |              |
|------|----------------------|---|----------------|--------------|
| vii. | <i>atomikí vómva</i> | < | <i>atomikí</i> | <i>vómva</i> |
|      | ‘bomba atômica’      |   | ‘atômica’      | ‘bomba’      |

<sup>165</sup> A principal motivação para distinguir compostos [radical+radical] de compostos [radical+palavra] no grego moderno vem da flexão. As marcas finais flexionais dos itens pertencentes aos compostos do primeiro tipo são, em geral, diferentes dos finais flexionais dos membros na segunda posição do composto, quando usados como uma palavra autônoma (e.g., *nixtolúludo* ‘flor da noite’ vs. *lulúdi* ‘flor’; *omorfôpeðo* ‘criança bonita’ vs. *peði* ‘criança’).

|       |                       |   |              |                |
|-------|-----------------------|---|--------------|----------------|
| viii. | <i>omáda eryasías</i> | < | <i>omáda</i> | <i>eryasía</i> |
|       | ‘grupo de trabalho’   |   | ‘grupo’      | ‘trabalho’     |
| ix.   | <i>pedí thávma</i>    | < | <i>pedí</i>  | <i>thávma</i>  |
|       | ‘menino prodígio’     |   | ‘menino’     | ‘prodígio’     |

Dada a distribuição acentual, e seguindo a generalização de que a presença de um acento primário demarca uma palavra fonológica, como explicitado em (154), as autoras concluíram que: compostos do Tipo 1 e do Tipo 2 formam uma única palavra fonológica, enquanto compostos do Tipo 3 formam duas palavras fonológicas, constituindo, posteriormente, uma frase fonológica. A explicação formal apresentada naquele momento era a de que havia uma divisão de trabalhos na formação dos compostos entre o léxico – para Tipos 1 e 2 – e a sintaxe – para o Tipo 3.

Mais especificamente, os compostos do Tipo 1, formados pela combinação de dois radicais, transformam-se em uma palavra morfológica quando afixados por morfemas flexionais. Dessa forma, através de uma regra de mapeamento morfologia-fonologia, o conjunto todo (i.e., os dois radicais e os demais sufixos) é, então, mapeado em uma palavra fonológica, a qual recebe a atribuição de acento primário na antepenúltima sílaba.

Nos compostos do Tipo 2, o acento do segundo membro do composto, ou seja, da palavra, é preservado. Esse membro já apresenta o estatuto de palavra fonológica no nível em que a composição ocorre, carregando um acento primário. Uma explicação para esse tipo de composto, segundo as autoras, exige a assunção de que há um princípio de preservação do acento, de acordo com o qual o acento é preservado durante o processo de formação de palavras quando envolve constituintes já acentuados (cf. BURZIO, 1994 *apud* NESPOR; RALLI, 1996).

Por fim, nos compostos do tipo 3, apesar de haver dois acentos primários oriundos das palavras que formam o composto, o acento do segundo membro é o mais proeminente dos dois, o que levou as autoras a assumirem que tais compostos constituem uma frase fonológica, pois se comportam fonologicamente como um sintagma qualquer. Ambos os acentos são primários no nível da palavra, enquanto que a proeminência do segundo é determinada no nível da sentença. Dessa forma, não é a fonologia desses compostos que os fazem ser

interpretados como tais, mas é o seu comportamento morfossintático e, frequentemente, seu significado específico<sup>166</sup>.

Nespor (1999) retoma as três combinações estruturais supramencionadas e mostra que elas são atestadas em uma variedade de línguas, sendo que cada língua pode fazer uso de mais de uma dessas combinações estruturais para formação de seus compostos, embora apresente preferência por uma estrutura ou outra (cf. NESPOR, 1999), tal como mencionamos no Capítulo 3<sup>167</sup>, ao contrastar as estruturas morfológicas dos compostos do PB com a de línguas tipologicamente distintas.

No que diz respeito aos compostos do Tipo 1, Nespor (1999) nota que esse tipo de composição está presente em um grande número de línguas (e.g., islandês, espanhol, catalão, francês, romeno, entre outras) as quais fazem uso de radicais neoclássicos para formação de seus compostos (e.g., *hidr-*, *neur-*, *log-*). A autora salienta que esses radicais devem ser considerados pseudoradicais, uma vez que não são flexionados e não podem ser utilizados livremente em uma sentença<sup>168</sup>. Exemplos desse tipo de composição em italiano estão elencados no exemplo (162), abaixo:

- |                           |   |                     |            |
|---------------------------|---|---------------------|------------|
| (162) a. <i>barómetro</i> | < | <i>baro metro</i>   | (Italiano) |
| ‘barômetro’               |   |                     |            |
| b. <i>onnívoro</i>        | < | <i>onni voro</i>    |            |
| ‘onívoro’                 |   |                     |            |
| c. <i>filántropo</i>      | < | <i>filo antropo</i> |            |
| ‘filantropo’              |   |                     |            |
| d. <i>téo[z]ofo</i>       | < | <i>teo [s]ofo</i>   |            |
| ‘teósofo’                 |   |                     |            |

O padrão acentual desses compostos pode variar de duas maneiras, segundo a autora: ou (i) o acento é atribuído por uma regra específica, ou (ii) a aplicação do acento segue as regras gerais de aplicação de acento das línguas em que aparecem. Os dados do italiano em

<sup>166</sup> Esse fato pode ser abarcado pela Enciclopédia (i.e., Lista 3) na MD, pois embora *pão duro* (i.e., um pão duro) e *pão-duro* (i.e., um indivíduo mesquinho, sovina) formem ambos uma frase fonológica, a Enciclopédia garante a interpretação não-composicional do segundo exemplo a partir do contexto sintático em que ele está inserido.

<sup>167</sup> Por exemplo, enquanto as línguas românicas apresentam uma preferência por compostos [palavra + palavra] (cf. NESPOR, 1999), o grego moderno prefere formar composto a partir de uma estrutura [radical + radical] (cf. RALLI, 2009).

<sup>168</sup> Esse último argumento deve ser analisado com muita cautela, pois, como vimos no Capítulo 3, os radicais neoclássicos não se diferenciam amplamente dos chamados radicais vernaculares, principalmente no que diz respeito ao seu caráter preso.

(162), por exemplo, tal como os do espanhol, têm a aplicação de acento seguindo regras específicas, uma vez que o acento recai na antepenúltima sílaba da palavra, uma posição marcada para essas línguas.

Casos em que a aplicação do acento segue a regra geral da língua para os compostos do Tipo 1 são encontrados no holandês e no francês. Por exemplo, no holandês, o acento primário é atribuído na última ou penúltima sílaba, de acordo com o peso da última rima, tal como pode ser visto em (163), e, no francês, ele é atribuído sempre à última sílaba, tal como em (164):

|       |                          |                  |            |
|-------|--------------------------|------------------|------------|
| (163) | a. <i>psycholoog</i>     | ‘psicólogo’      | (Holândes) |
|       | b. <i>filantroop</i>     | ‘filantropo’     |            |
|       | c. <i>psychologe</i>     | ‘psicóloga’      |            |
| (164) | a. <i>aérodrome</i>      | ‘campo aéreo’    | (Francês)  |
|       | b. <i>fratricide</i>     | ‘fatricida’      |            |
|       | c. <i>anthropomorphe</i> | ‘antropomórfico’ |            |

Nos compostos do Tipo 2, Nespor (1999) nota que há também uma variação translinguística no seu mapeamento em estruturas fonológicas, pois eles poderão ser mapeados em uma única palavra fonológica, como vimos nos dados do grego moderno em (161), ou em duas palavras fonológicas, como é o caso das línguas românicas e germânicas, por exemplo. Nessas últimas, o radical na primeira posição é, na maioria das vezes, neoclássico, tal como aqueles utilizados nos compostos do Tipo 1, em (163) e (164).

Para ilustrar os compostos do Tipo 2 que são mapeados em duas palavras fonológicas, trazemos os dados do holandês apresentados em Nespor (1999, p. 134). Esses compostos, além de carregar dois acentos primários, não permitem que ocorra uma regra comum à palavra fonológica nessa língua, a saber, o apagamento da vogal no ponto de junção entre dois elementos, o que evidencia o seu mapeamento em duas palavras fonológicas:

|       |                            |                            |                   |            |
|-------|----------------------------|----------------------------|-------------------|------------|
| (165) | a. <i>filo-Israélisch</i>  | (* <i>filisraelisch</i> )  | ‘filo-israelense’ | (Holandês) |
|       | b. <i>áuto-immunitéit</i>  | (* <i>autimmuneit</i> )    | ‘autoimunidade’   |            |
|       | c. <i>hydro-eléktrisch</i> | (* <i>hydrelektrisch</i> ) | ‘hidroelétrica’   |            |

Não nos aprofundaremos nas discussões sobre os compostos do Tipo 3 feitas em Nespor (1999), pois a autora apenas amplia empiricamente a discussão elaborada em Nespor e Ralli (1996). O que devemos ter em mente com relação a esses dados é que compostos [palavra + palavra] formam duas palavras fonológicas, haja vista que se equiparam, fonologicamente, a um sintagma (e.g., compostos V-N, N-N, N-A, N-prep-N, dentre outros).

A partir do conjunto de informações elencadas até aqui, trazemos o resumo da correspondência entre a estrutura morfológica do composto e o seu mapeamento fonológico:

**Quadro 5** - Constituição morfológica e fonológica dos compostos (NESPOR, 1999)

| [radical + radical] | [radical + palavra]    | [palavra + palavra] |
|---------------------|------------------------|---------------------|
| 1 $\omega$          | 1 $\omega$ /2 $\omega$ | 2 $\omega$          |

A generalização elaborada por Nespor (1999), de acordo com a tabela acima, diz que dadas as estruturas morfológicas dos diferentes tipos de compostos, o elemento X<sup>o</sup> mais abaixo na estrutura em (167) é mapeado em uma palavra fonológica (1999, p. 139), seguindo o Princípio da Palavra Mínima, de Nespor e Ralli (1996):

(166) *Princípio da Palavra Mínima (NESPOR; RALLI, 1996, p. 378, tradução nossa)*

No mapeamento dos compostos em palavras fonológicas, escolha o elemento X<sup>o</sup> mais baixo na estrutura<sup>169</sup>.

Com base na estrutura geral em (167), temos, para os compostos do Tipo 1, que o nó mais baixo X<sup>o</sup> corresponde ao nó que domina os dois radicais – uma vez que radicais não são elementos do nível X<sup>o</sup> – e isso, segundo as autoras, prevê que esse tipo de composição formará uma única palavra fonológica. Essa generalização se estende a todas as línguas europeias (cf. NESPOR, 1999), incluindo o grego, o italiano, o holandês, o islandês, o espanhol, o francês, o romeno e o português.

Nos compostos do Tipo 2, parece haver uma variação dependendo da relevância atribuída ao membro da direita ou da esquerda no composto. No grego, em algumas

<sup>169</sup> Trecho original (NESPOR; RALLI, 1996, p. 378): “Minimal Word Principle: Choose the lowest X<sup>o</sup> in the mapping of compounds onto phonological words.”



variedades do italiano e no islandês, o composto todo é mapeado em uma única palavra fonológica. Por outro lado, em outras variedades do italiano e no holandês, como vimos em (165), esse tipo de composto é mapeado em duas palavras fonológicas. Outras línguas além do italiano e do holandês também apresentam compostos do Tipo 2 mapeados em duas palavras fonológicas, tal como o catalão, o romeno e o servo-croata (cf. NESPOR, 1999).

Por fim, nos compostos do tipo 3, os níveis mais baixo de  $X^0$  correspondem cada um a um membro do composto, e, portanto, cada elemento é mapeado em uma palavra fonológica separada. Nespor (1999) defende, ainda, que compostos do Tipo 3 são sempre mapeados em duas palavras fonológicas, uma hipótese que encontra evidências em muitas línguas românicas, germânicas e eslavas.

As diferentes possibilidades de prosodização dos compostos estão representadas abaixo, a partir da adaptação elaborada por Vigário (2003) das generalizações feitas por Nespor (1999), resumidas no Quadro 5. Nelas, o  $X^0$  relevante para a construção do domínio da palavra fonológica de cada estrutura está representado em negrito e entre parênteses:

Com relação ao português, a distribuição foi abordada de diferentes formas em suas duas variedades principais: o PB e o PE. No que concerne à primeira variedade, queremos salientar que os Tipos 1 e 2 foram frequentemente negligenciados nas análises acentuais. Lee (1995), como veremos na próxima subseção, nem menciona esse tipo de composição diretamente. Moreno (1997) também isolou esse tipo de dados de sua discussão, e Bopp da Silva (2010), apesar de discutir esses dados, não explicita sua relevância estrutural dentro da língua.

No PE, por outro lado, o trabalho de Vigário (2003) contempla os três tipos estruturais, fornecendo um detalhamento da atribuição acentual de acordo com a estrutura morfológica do composto. A autora ainda questiona a posição de Villalva (1994), de que compostos do Tipo 2 inexistem nessa língua, e traz um conjunto variado de dados e argumentos para motivar a sua presença. Por esse motivo, desenvolvemos o mesmo percurso para descrever os dados do PB

na subseção 6.3.3, visto que as análises que temos até hoje não consideraram, diretamente, a estrutura morfológica na distribuição acentual, o que fez com que compostos do Tipo 1 e do Tipo 2 fossem deixados de lado nessas abordagens.

De modo geral, Toneli (2011) comenta que, embora muitos trabalhos tenham feito referência ao estatuto prosódico das palavras funcionais (cf. BISOL, 2000, 2004), das palavras lexicais (cf. LEIRIA, 2000) e dos prefixos (cf. SCHWINDT, 2000), ainda não há um estudo que englobe sistematicamente o domínio formado pelas palavras compostas no PB. Neste capítulo, contribuímos com a descrição das palavras compostas (i.e., com mais de uma palavra fonológica), mais especificamente dos compostos, ao discutir a influência das diferenças estruturais desse modo de formação de palavras e sua caracterização prosódica no PB.

### **6.3.2. Níveis lexicais governando a aplicação de acento**

Esta subseção, apesar de apresentar um trabalho que dialoga com a divisão operacional da fonologia em níveis lexicais e pós-lexicais, tal como os trabalhos apresentados na última subseção, propõe uma tipologia distinta para a definição dos compostos. Nessa visão, delineada nos trabalhos de Lee (1995, 1997), compostos são divididos em lexicais (CL) e pós-lexicais (CPL), como vimos na seção 3.2, do Capítulo 3, com base não em sua estrutura morfológica, mas nos processos fonológicos e morfossintáticos presentes nos compostos do PB.

Essa divisão proposta pelo autor pode ser vista, em certa medida, como semelhante à de Nespor e Ralli (1996) e Nespor (1999), pois ilustra o posicionamento mais comum das visões lexicalistas de gramática sobre a composição, em que os compostos são explicados basicamente de duas formas: (i) há compostos que são analisados como formados no léxico e são sintaticamente opacos, ou seja, compostos que se comportam como uma unidade em relação a processos morfossintáticos, pois não permitem flexão, derivação, nem concordância; e (ii) há compostos que são analisados como formados no componente pós-lexical e, portanto, são sintaticamente transparentes, permitindo flexão, derivação e concordância.

A diferença crucial, no entanto, é que, enquanto as autoras apresentadas na última subseção apontam a relevância da estrutura morfológica no mapeamento prosódico, a visão de Lee (1995, 1997) baseia-se apenas nas distinções processuais que circundam a composição, o que leva o autor a desconsiderar outros conjuntos de compostos no PB, como veremos nas descrições feitas na subseção 6.3.3.

No que diz respeito ao domínio prosódico, a distinção estabelecida por Lee (1995) não apresenta impacto direto, uma vez que tanto os CL, como aqueles em (169a) e (169b), quanto os CPL, em (169c), podem ser formados por mais de uma palavra fonológica, e suas construções dão-se como representado no exemplo (170), abaixo:

- (169) a.  $\alpha$  composição I:  $\langle \_\_\_\_\_\_ \rangle_{m\alpha}$  .  $\rightarrow$  [  $\_\_\_\_\_\_$  ]  $p\alpha$   
 e.g. [ferrovia] $\omega$
- b.  $\alpha$  composição II:  $\langle \_\_\_\_\_\_ \rangle_{m\alpha}$   $\rightarrow$  [  $\_\_\_\_\_\_$  ]  $p\alpha$  [  $\_\_\_\_\_\_$  ]  $p\alpha$   
 e.g. [puxa] $\omega$  [saco] $\omega$
- c. pós-lexical:  $\langle \_\_\_\_\_\_ \rangle_{\gamma}$  [  $\_\_\_\_\_\_$  ]  $\gamma$   $\rightarrow$  [  $\_\_\_\_\_\_$  ]  $\gamma$  [  $\_\_\_\_\_\_$  ]  $\gamma$   
 e.g. [presidente] $\omega$  [ministro] $\omega$
- 
- (170) a.  $\langle \text{ferro} \rangle_m$  [feR] $_{p\alpha}$  o MCF e PCF<sup>170</sup>  
 $\langle \text{via} \rangle_m$  [vi] $_{p\alpha}$  a MCF e PCF  
 $\langle \text{ferrovia} \rangle_m$  [feRovi] $_{p\alpha}$  a composição de  $\alpha$  e PCF  
 [feRoví] $_{p\alpha}$  a acento  
 :  
 [feRovía]
- b.  $\langle \text{puxa} \rangle_m$  [puš] $_{p\alpha}$  a MCF e PCF  
 $\langle \text{saco} \rangle_m$  [sak] $_{p\alpha}$  o MCF e PCF  
 $\langle \text{puxasaco} \rangle_m$  [puš] $_{p\alpha}$  a [sak] $_{p\alpha}$  o composição de  $\alpha$  e PCF  
 $\langle \text{puxasaco} \rangle_m$  [púš] $_{p\alpha}$  a [sák] $_{p\alpha}$  o acento  
 :  
 [pùšasáko]
- c.  $\langle \text{presidente} \rangle_m \langle \text{ministro} \rangle_m$  [prezidênte] $_{p\alpha}$  [minístro] $_{p\alpha}$

Obviamente, os problemas em sua tipologia, vistos no Capítulo 3, podem acarretar problemas na sua proposta acentual para os compostos do PB, uma vez que o autor

<sup>170</sup> MCF: Formação de Constituinte Morfológico; PCF: Formação de Constituinte Prosódico.

desconsidera compostos que são formados por radicais nessa língua. No entanto, deixando de lado essas questões por um momento, focalizemos a proposta de acento elaborada por Lee (1995), contrastando-a com a proposta de Bisol (1992), as quais são centrais para o debate sobre a aplicação de acento no PB.

### 6.3.3. Os algoritmos acentuais no PB: entre Bisol (1992) e Lee (1995)

Lee (1995), em linhas gerais, procura se distanciar de propostas para o acento que assumem a palavra como domínio de aplicação, tal como faz Bisol (1992)<sup>171</sup>, argumentando em favor de que o domínio de aplicação das regras de acento no PB levam em consideração o radical. Ambos os autores fazem uso da Fonologia Métrica (cf. LIBERMAN; PRINCE, 1977; HALLE; VERGNAUD, 1987; HAYES, 1995), teoria que propõe que o acento não se relaciona diretamente com a vogal, mas provém de uma relação que se estabelece entre as sílabas, formando um contorno de proeminência<sup>172/173</sup>. Dessa forma, torna-se necessário verificar como as sílabas se organizam em constituintes prosódicos maiores, a saber: em pés métricos.

Bisol (1992) vale-se da proposta métrica de Halle e Vergnaud (1987) para propor o algoritmo acentual do PB. Sua análise não faz distinção entre categorias lexicais para a aplicação da regra, e o pé métrico, para a autora, é construído a partir de grades parentetizadas<sup>174</sup>. No entanto, o domínio de aplicação da regra é diferenciado, pois, enquanto nos nomes, a regra se aplica na palavra derivacional (i.e., radical + vogal temática), ciclicamente, nos verbos, a regra se aplica na palavra pronta, ou seja, na palavra lexical de uma só vez, e, como o acento se aplica na palavra pronta, a regra não é cíclica. Dessa forma, temos o seguinte algoritmo acentual:

<sup>171</sup> Citamos Bisol (1992) por ser a proposta mais comentada pelo autor. No entanto, em seu trabalho de (1995), Lee menciona uma lista de outros autores que tomam a palavra como domínio para a aplicação do acento no português.

<sup>172</sup> Em modelos lineares para a fonologia, como em SPE, o acento era tratado como uma propriedade de segmentos individuais (vogais), fazendo com que a atribuição de acento fosse determinada por sequências particulares de segmentos, a partir de traços com valores binários, e.g. [+acentuado], [-acentuado].

<sup>173</sup> Para uma ampla revisão da literatura acerca das propostas de aplicação de acento, cf. Magalhães (2004), cap. 1; Collischonn (2010).

<sup>174</sup> A parentetização da grade métrica, proposta por Halle e Vergnaud (1987), é uma modificação inserida no modelo de grades para a aplicação do acento (cf. PRINCE, 1983 *apud* MAGALHÃES, 2004), em que se adicionam parênteses à grade a fim de delimitar constituintes, não necessariamente binários, de modo que cada constituinte projete no nível subsequente o elemento com maior proeminência, i.e. o núcleo.

(171) *Regra do acento primário (BISOL, 1992)*

Domínio: a palavra lexical

- i. Atribua um asterisco (\*) à sílaba pesada final, ou seja, sílaba de rima ramificada;
- ii. Nos demais casos, forme um constituinte binário (não-iterativamente), com proeminência à esquerda, do tipo (\* .), junto à borda direita da palavra.

Nessa proposta, a regra de acento é sensível ao peso silábico, fazendo com que o acento seja atribuído às oxítonas terminadas em consoante ou ditongo, como em (172a) e (172b). Caso a primeira sílaba, da direita para a esquerda, não seja pesada, a regra determina que o acento cairá na segunda sílaba, marcando os acentos das palavras paroxítonas, como em (172c).

(172) a. /pomar/                      b. /trofeu/                      c. /kas + a/

                  (\*)                                      (\*)                                      (\*) .

                  (                      \*)                      (                      \*)                      (                      \*)

A extrametricidade, ou seja, a possibilidade de tornar unidades invisíveis às regras de construção dos constituintes métricos (e.g., uma sílaba, uma mora, ou um segmento), e, consequentemente, de isentá-las na aplicação do acento, é importante para abarcar casos que fogem à generalização. A localização do constituinte extramétrico deve ser inicial ou final no ponto da derivação onde a construção dos constituintes métricos ocorre, sendo, no PB, categórica nos verbos e marcada lexicalmente nos não-verbos. Dessa forma, palavras com acento na terceira sílaba, ou seja, proparoxítonas, e palavras terminadas em consoante ou ditongo com acento não-final devem marcar a sílaba final como extramétrica:

(173) a. /numer + o/                      b. /arvor + e/                      c. /fosfor + o/

<ro>                                      <re>                                      <ro>

(\*     .)                                      (\*     .)                                      (\*     .)

(\*             )                                      (\*             )                                      (\*             )

Alguns autores, como Lee (1995), criticam a proposta de Bisol (1992) pelo fato de a autora (i) fazer uso da extrametricidade, (ii) por adotar a sensibilidade ao peso e (iii) por unificar as categorias lexicais em uma única regra. Lee (1995), ao contrário, propõe que existem diferentes regras de aplicação de acento para verbos e não-verbos no PB. Baseando-se no modelo de Hayes (1991), o autor considera que o acento é sensível à categoria lexical,

insensível ao peso silábico, e, como já apontamos antes, distribui-se em dois níveis de aplicação: no nível ( $\alpha$ ), para o não-verbo, e no nível ( $\beta$ ), para o verbo.

(174) *Regra do acento primário para não-verbo (LEE, 1995)*

Domínio: nível  $\alpha$  (radical derivacional)

i. Casos não-marcados

- a. constituinte binário;
- b. núcleo à direita;
- c. direção: direita para a esquerda;
- d. não-iterativo.

ii. Casos marcados

- a. Constituintes binários;
- b. Núcleo à esquerda;
- c. Direção: direita para a esquerda;
- d. Não-iterativo.

(175) *Regra do acento primário para verbos (LEE, 1995)*

Domínio: nível  $\beta$  (palavra)

i. Casos não-marcados

- a. Constituintes binários;
- b. Núcleo à esquerda;
- c. Direção: direita para a esquerda;
- d. Não-iterativo.

ii. Casos marcados

- a. Constituintes binários;
- b. Núcleo à direita;
- c. Direção: direita para a esquerda;
- d. Não-iterativo.

Aplicando as respectivas regras, temos:

## (176) Não-verbos

|           |          |         |            |                         |
|-----------|----------|---------|------------|-------------------------|
| a. almoço | b. tonel | c. café | d. abóbora |                         |
| /almoço/o | /tonel/  | /café/  | /abobor/a  | domínio: nível $\alpha$ |
| (. *)     | (. *)    | (. *)   | _____      | (174-i)                 |
| _____     | _____    | _____   | (* .)      | (174-ii)                |

## (177) Verbos

|            |            |           |                        |
|------------|------------|-----------|------------------------|
| a. computo | b. falamos | c. baterá |                        |
| /computo/  | /falamos/  | /baterá/  | domínio: nível $\beta$ |
| (* .)      | (* .)      | _____     | (175-i)                |
| _____      | _____      | (. *)     | (175-ii)               |

Notamos, claramente, que há um aumento no número de regras para a aplicação do acento na proposta de Lee (1995). No entanto, essa proposta permite que não recorramos à extrametricidade, recurso que parece não ser explicado de modo trivial em modelos de inserção tardia para a fonologia, pois a demarcação do segmento extramétrico depende da quantidade de informação sintática enviada pelo domínio da fase, a fim de receber informações fonológicas.

Em decorrência disso, assumimos que a proposta acentual de Lee (1995) reduz a necessidade de demarcação das idiossincrasias extramétricas para a aplicação da regra, além de sugerir uma melhor adequação à arquitetura da gramática empregada nesta dissertação. Porém, para proceder a uma tentativa de aplicação da proposta de Lee (*op. cit.*) na MD, temos, ainda, que equacionar a hipótese das raízes acategoriais com a noção de fase, com o intuito de garantir que a informação categorial esteja visível para a fonologia após o *Spell-Out* sintático.

#### 6.3.4. Quantas palavras fonológicas são necessárias para um composto no PB?

O que vimos até aqui é que as propostas acentuais para a composição dividem-se em dois grandes grupos: estruturais e processuais. Não podemos dizer que há semelhanças entre as duas propostas, uma vez que para o segundo grupo as informações estruturais dos compostos são irrelevantes. O que propomos nesta seção é uma divisão dos compostos do PB seguindo a visão estrutural para a distribuição do acento, como utilizadas em Nespor e Ralli (1996) e Nespor (1999).

Dando continuidade à defesa da presença de compostos do Tipo 1 [radical + radical] e Tipo 2 [radical + palavra] no PB, mostramos, nas subseções seguintes, suas diferenças prosódicas, evidenciando, através de um conjunto de processos fonológicos restritos à palavra fonológica, em quais níveis prosódicos os compostos do PB são mapeados.

#### 6.3.4.1. Compostos formados por radicais

No PB, tal como nas demais línguas românicas e germânicas, compostos formados por radicais contêm, no mínimo, um radical preso de origem grega ou latina (e.g., *hidr-*, *neur-*, *psic-*, *-log*, *-metr*), e suas formações são denominadas de compostos neoclássicos, tal como vimos no Capítulo 3. Esse tipo de composição apresenta uma característica em particular, comum em compostos formados por radicais, que é a presença de um elemento de ligação, sendo *-o-* ou *-i-* a depender do constituinte que ocupa a segunda posição.

Dito isso, temos, como exemplos de compostos formados por radicais os dados em (178), abaixo:

- |                           |                      |                       |
|---------------------------|----------------------|-----------------------|
| (178) a. cervej-ó-log-o/a | e. cervej-o-log-ía   | i. ----               |
| b. palpit-ó-metr-o        | f. palpit-o-metr-ía  | j. ----               |
| c. psic-ó-log-o/a         | g. psic-o-log-ía     | k. psic-o-lóg-ico     |
| d. histori-ó-graf-o/a     | h. histori-o-graf-ía | l. histori-o-gráf-ico |

Assumindo-se que a forma básica desses compostos seja a concatenação dos dois radicais somados a um marcador de classe em sua borda direita, como nos dados (178a-d), notamos que há apenas um acento primário localizado na antepenúltima sílaba do composto. Assim, fazendo um contraponto com o que foi discutido ao longo do capítulo, podemos tomar essa informação como evidência para que compostos do Tipo 1 [radical + radical], no PB, formem uma única palavra fonológica, como ocorre nos dados de mesma estrutura no grego moderno e nas demais línguas românicas e germânicas.

Outros processos fonológicos restritos ao domínio da palavra fonológica também são atestados nesse tipo de composição, dando suporte para a evidência acentual apontada. O primeiro processo fonológico a ser mencionado é a neutralização da pretônica, que ocorre quando o /ε/ ou /ɔ/ perdem o acento, seja por derivação morfológica ou por adição de morfemas flexionais, levando a um traço de abertura na vogal, como ilustrado nos dados em (179):



- (179) a. belo > beleza  
 b. sol > solação  
 c. leve > leveza  
 d. rolo > rolamos

O mesmo processo ocorre nos compostos formados por radicais, caso o afixo adicionado desencadeie o movimento do acento para a borda direita da palavra. Por exemplo, tomando os dados (178e-h), a adição do sufixo *-ia* neutraliza a vogal acentuada, da mesma forma que ocorre nos dados em (179):

- (180) a. cervejologo > cervejología  
 b. palpímetro > palpitometría  
 c. psicologo > psicología  
 d. historiografo > historiografía

Por outro lado, se adicionamos o sufixo *-ico*, como nos dados em (178i-l), não há a alteração na posição do acento, e, portanto, a regra de neutralização da pretônica não se aplica. Porém, nos casos em que a adição de *-ico* altera, consideravelmente, o número de sílabas da palavra, como em (178l), e em *tragediográfico*, *musicográfico* e *epistolográfico*, o acento deve se mover da posição em que é encontrado na forma básica do composto a fim de permanecer na antepenúltima sílaba da janela trissilábica de acento, o que, por sua vez, desencadeia a neutralização da pretônica (e.g., *historiografo* > *historiográfico*; *tragediografo* > *tragediográfico*).

Rondinini e Gonçalves (2006) e Rondinini (2009) propuseram que formas como *-ólogo* e *-ógrafo* são vistas, sincronicamente, como sufixos derivacionais no PB. Os autores argumentam que o uso amplo e categórico da vogal *-o-* nesse contexto, e o fato dela apresentar exclusivamente sua realização como uma vogal média baixa /ɔ/, seriam evidências para a gramaticalização. Os autores, no entanto, não notaram que a vogal *-o-* é a vogal menos marcada para ocorrer em ambientes de conexão de radicais no PB, como argumentamos no Capítulo 3 (e.g., *dent-o-bucal*, *kassab-o-sserrista*, *dilm-o-fóbica*, *cam-o-terapia*, *cervej-o-chato*), e que, em especial nas formas *-log* e *-graf*, devido à sua origem grega, o elemento de ligação a ser utilizado é a vogal *-o-*, o que foi mantido na sincronia.

Outro ponto contra a proposta sufixal dos autores supramencionados é que a vogal não se mantém categoricamente como /ɔ/, visto que ela é neutralizada quando sufixos derivacionais são concatenados ao composto, o que pode ser verificado nos dados em (180), passando a ser uma vogal /o/. Se seguirmos a proposta dos autores, deveríamos ter uma gramática que contivesse vários sufixos para abarcar as formações com *-log* e *-graf*, a saber: *-ólogo*, *-ógrafo*, *-ologia*, *-ografia*, *-ológico*, *-olológico*, *-ográfico*, o que não é nada econômico.

Outro processo fonológico que evidência a formação de uma única palavra fonológica nesses dados é a impossibilidade de ocorrer alçamento de vogal nos pontos de junção – processo comum na frase fonológica (181) – já que a vogal de ligação pertence à última sílaba do primeiro radical, como pode ser visto em (182):

- (181) a. cantolírico      \*[o]/[w]  
       b. plantio de cana    \*[o]/[w]

- (182) a. cervejologia      [o]/\*[w]  
       b. psicologia        [o]/\*[w]

O que foi argumentado com relação a esses dados promove uma alteração na listagem de unidades que constituem uma palavra fonológica individual no PB (cf. MORENO, 1997, p. 24), a saber: a adição dos compostos formados por radicais. Dessa forma, trazemos a listagem reformulada em (183):

- (183) *Formações que constituem uma palavra fonológica individual no PB (reformulada)*
- a. radical + marcador (se houver)
  - b. radical + sufixos derivacionais (o produto também é um radical)
  - c. cada membro de um composto **formado por palavras**
  - d. prefixos tônicos (pré-, anti, pós, etc.)
  - e. sufixos especiais: *-zinhV*, *-mente* e *-íssimo*
  - f. compostos formados por radicais**

Passemos, agora, aos compostos de estrutura [radical + palavra].

### 6.3.4.2. Compostos formados por radical e palavra

Compostos formados por [radical + palavra] são semelhantes aos compostos formados por radicais, pois desencadeiam a inserção de um elemento de ligação, no caso, a vogal -o-, exclusivamente. Exemplos desse tipo de composição no PB estão listados em (184):

- |                          |                      |
|--------------------------|----------------------|
| (184) a. neur-o-cirurgia | e. kassab-o-sserista |
| b. psic-o-afetivo        | f. dente-o-bucal     |
| c. hidr-o-ginástica      | g. eletr-o-choque    |
| d. heter-o-machista      | h. psic-o-genético   |

Alguns dados, no entanto, são difíceis de serem definidos entre compostos [radical + radical] e compostos [radical + palavra], pois sugerem uma recursividade na concatenação. Casos como esses são exemplificados em (185), em que o segundo e o terceiro radicais podem ser vistos ou (i) como unidades dissociadas formando um composto com três radicais (185a-c), ou (ii) como um composto [radical + radical] que tem, posteriormente, um terceiro radical concatenado à sua borda esquerda (185a'-c'):

- |                             |                          |
|-----------------------------|--------------------------|
| (185) a. psic-o-bi-o-log-ia | a'. psic-o-biologia      |
| b. neur-o-psic-o-log-ia     | b'. neur-o-psicologia    |
| c. eletr-o-glott-o-graf-ia  | c'. eletr-o-glottografia |

Partindo para o critério acentual, percebemos que, nos compostos em (184), há a presença de dois acentos principais: (a) o acento principal preservado da palavra na segunda posição, e (b) um acento no radical, como pode ser visto nos dados repetidos em (186):

- |                          |                      |
|--------------------------|----------------------|
| (186) a. néur-o-cirurgia | e. kassáb-o-sserista |
| b. psíc-o-afetivo        | f. dént-o-bucal      |
| c. hídr-o-ginástica      | g. elétr-o-chóque    |
| d. héter-o-machista      | h. psíc-o-genético   |

O mesmo ocorre com os dados em (185), em que o acento dos dois últimos radicais está localizado na mesma posição em que ele é atribuído quando um composto [radical + radical] é formado. Essa semelhança na aplicação do acento sugere que os dois últimos

radicais formam um composto antes da adição do terceiro radical na borda esquerda, de acordo com a divisão em (185a'-b'), configurando-se, portanto, como compostos [radical + palavra]:

- (187) a. psíc-o-biología  
       b. néur-o-psicología  
       c. élétr-o-glología

Passando para outros processos fonológicos, percebemos que compostos [radical + palavra] permitem o apagamento da vogal no ponto de junção, quando a palavra na segunda posição inicia-se em vogal. Esse tipo de apagamento da vogal não é permitido na palavra fonológica, como verificamos em (188), mas é permitido na frase fonológica, como em (189), o que indica, além da evidência acentual, que compostos [radical + palavra] formam duas palavras fonológicas:

- |                                |                      |
|--------------------------------|----------------------|
| (188) a. <u>to</u> alha        | *0/[o]               |
| b. <u>do</u> ação              | *0/[o]               |
|                                |                      |
| (189) a. espa <u>ço</u> aberto | 0/[o] <sup>175</sup> |
|                                |                      |
| (190) a. psic-o-afetivo        | 0/[o]                |
| b. psic-o-análise              | 0/[o]                |
| c. neur-o-anatomia             | 0/[o]                |
| d. gastr-o-intestinais         | 0/[o]                |

Outro processo que pode servir como evidência para o mapeamento em duas palavras fonológicas dessa estrutura é a possibilidade de alçamento da vogal de ligação /o/ para /w/, processo fonológico que foi apontado como ausente em compostos [radical + radical], no exemplo (182). Apesar do alçamento da vogal ser possível, os falantes não foram unânimes na avaliação dos dados, o que pode ser visto nas possibilidades marcadas com uma interrogação em (191):

<sup>175</sup> Nos exemplos em que não há marcação de agramaticalidade (\*), ambas as formas são encontradas.

- |                                  |          |
|----------------------------------|----------|
| (191) a. neuro <u>o</u> cirurgia | [o]/[w]  |
| b. psic <u>o</u> biologia        | [o]/[w]  |
| c. psic <u>o</u> afetivo         | [o]/[w]  |
| d. neuro <u>o</u> ciência        | [o]/?[w] |
| e. hidro <u>o</u> ginástica      | [o]/?[w] |

Portanto, apesar dos testes não apresentarem uma uniformidade de julgamento, o que encontramos, até o presente momento, é um conjunto de evidências que nos levam a considerar os compostos [radical + palavra] do PB como sendo mapeados em duas palavras fonológicas.

#### 6.3.4.3. Compostos formados por palavras

Os compostos formados por palavras são a grande maioria dos compostos do PB, como mencionamos no Capítulo 3. Como já vimos, a variedade de combinações categorias presentes nessas formações é bem diversificada (e.g., V-N, N-N, A-N, N-A, A-A), diferentemente dos demais tipos estruturais vistos acima que se restringem às combinações N-N quase que exclusivamente. Abaixo, trazemos um conjunto de dados em (192), a fim de endossar as evidências sobre seu mapeamento prosódico:

- (192) a. límpa-vídro  
 b. péixe-espáda  
 c. férro-vélho  
 d. lívre-comércio  
 e. pórtá-espáda  
 f. pédra-infernál

Câmara Jr. (2010 [1971]), como mencionado na seção 6.2, já enfatizava a presença de dois acentos principais nos compostos formados por palavras no PB. Essa característica parece ser bem atestada translinguisticamente, se tomarmos os dados discutidos em Nespor (1999), o que induz à conclusão de que esses compostos formam duas palavras fonológicas independentes.

Outros processos fonológicos podem ser citados a fim de evidenciar a formação de duas palavras fonológicas. O primeiro é a neutralização da átona final, que ocorre no primeiro membro de cada um dos compostos, exemplificado nos dados em (193), abaixo:

- (193) a. ferro-velho           \*[o]/[w]  
       b. livre-comércio       \*[i]/[j]  
       c. bate-boca           \*[i]/[j]

Bopp da Silva (2010) traz outro argumento, a ocorrência de sândi vocálico externo, mais especificamente a elisão, a qual promove a reestruturação de sílabas e pode ser visto como uma evidência para mostrar a prosodização desses compostos em duas palavras fonológicas, já que a elisão é bloqueada no interior de palavras e só ocorre entre palavras fonológicas independentes.

- (194) a. cara estranha   → car[i]stranha  
       b. esvaecer       → \*esv[i]cer
- (195) a. pedra infernal   → pedr[i]nfernal  
       b. peixe espada   → peix[i]spada

Em resumo, podemos dizer que o PB apresenta as seguintes estruturas morfológicas e distribuições acentuais na constituição de seus compostos. Seus diferentes padrões acentuais são evidenciados pela variação no número de acento primário com relação à estrutura morfológica e pela presença de processos fonológicos restritos ao domínio da palavra fonológica ou da frase fonológica:

**Quadro 6** - Constituição morfológica e fonológica dos compostos do PB

| <b>[radical + radical]</b> | <b>[radical + palavra]</b> | <b>[palavra + palavra]</b> |
|----------------------------|----------------------------|----------------------------|
| 1ω                         | 2ω                         | 2ω                         |

Será a partir das informações resumidas no Quadro 6 que iremos nos pautar para sugerir uma análise da aplicação de acento nos compostos do PB, assumindo que a formação de domínios prosódicos constitui parte do ramo de PF no arcabouço teórico da MD assumido nessa dissertação.

#### 6.4. Morfologia Distribuída: em busca de evidências para um mapeamento indireto dos constituintes sintáticos

Como pôde ser notado, a MD não apresenta um primitivo que seja a “palavra” ou o “radical”. No entanto, a teoria se vale de um conjunto de primitivos – raízes e morfemas abstratos – e um conjunto de procedimentos para sua combinação a fim de gerar núcleos complexos. Portanto, quando o assunto é a formação de palavras, não estamos tratando de apenas um nó terminal sintático  $X^0$ , mas de um conjunto de nós terminais sintáticos responsáveis por constituir o que é, grosso modo, definido como palavra.

Dessa forma, o que a fonologia prosódica vinha questionando acerca do isomorfismo ou não entre o componente sintático e o componente fonológico, ou seja, especificar se um único nó terminal sintático é mapeado em uma ou duas palavras fonológicas, passa a requerer uma revisão dentro de uma visão de gramática como a da MD, visto que o número de nós terminais que constituem a “palavra” é mais complexo. Ao mesmo tempo, o modelo não conta, a priori, com a noção de palavra fonológica – e dos demais domínios prosódicos – como parte de sua constituição ontológica, e parte do esforço empreendido neste capítulo visa sugerir uma adequação à criação desses domínios prosódicos como parte dos mecanismos presentes no caminho para PF.

Com isso, antes de adentrarmos o caminho para PF, é necessário ter em mente o que já foi descrito como parte de sua constituição e qual a sua relevância em um modelo de inserção tardia para a fonologia. Desde o trabalho seminal de Halle e Marantz (1993), a MD conta com um componente pós-sintático, o componente morfológico, ou, simplesmente, a morfologia, que toma como *input* as estruturas geradas pelo componente sintático, com o intuito de aplicar sobre elas operações responsáveis por garantir a distribuição correta dos expoentes fonológicos.

Por conseguinte, a morfologia é vista como uma série de operações pós-sintáticas que ocorrem no ramo de PF, as quais são responsáveis por remodelar a estrutura sintática e sua matriz de traços – se necessário – com vias a permitir a inserção de vocabulário de acordo com a disponibilidade de IV presentes em dada língua. É dentro dessa perspectiva que o caminho para PF contém determinações que são específicas à língua, e permite que a sintaxe opere de maneira independente, sem recorrer a um conjunto de operações a fim de resolver problemas morfofonológicos e garantir a ordem linear da língua em questão.

Com isso, obtemos um sistema dinâmico de PF, no qual o componente fonológico consiste de um conjunto (ordenado) de computações que se aplicam no *output* sintático (cf. EMBICK, 2007). Essas operações morfológicas são, principalmente, as seguintes:

(196) *Operações morfológicas presentes no caminho para PF*

- a. Fissão
- b. Fusão
- c. Empobrecimento
- d. Abaixamento
- e. Deslocamento local
- f. Inserção de morfemas dissociados

As três primeiras operações restringem-se aos traços presentes nos nós terminais sintáticos e ocorrem antes da inserção de vocabulário. Podemos dizer que as duas primeiras operações – fissão e fusão – são estruturais, pois rearranjam o número de nós terminais derivados pela sintaxe de acordo com os IVs que uma dada língua contém. Como a relação entre os traços presentes nos nós terminais e sua expoência fonológica não é de um para um, essas operações podem ou fissionar o nó terminal em dois, quando determinada informação sintática contiver dois expoentes fonológicos, ou seja, dois IVs, ou fundir dois nós terminais em um quando a língua detiver apenas um IV para as informações sintáticas em questão.

Essas operações serão relevantes na distribuição dos traços presentes nos núcleos categoriais dentro do nosso sistema, pois embora eles estejam representados como um único feixe de traços na sintaxe, o componente morfológico especificará se todos esses traços – por exemplo, os traços presentes no categorizador nominal e verbal, como os traços  $\phi$  – serão fissionados em nós terminais distintos a serem preenchidos por diferentes IVs, ou se serão preenchidos por um único IV (e.g., número e pessoa nos verbos do PB). Essa distribuição, no entanto, é determinada pela língua.

O empobrecimento, por sua vez, corresponde à subespecificação de traços sintáticos presentes em determinado nó terminal, alterando o valor de sua marcação ou o apagando. Ao operar sobre a matriz de traços de um nó terminal, o Empobrecimento pode, por exemplo, apagar um determinado traço sintático, a fim de permitir que a inserção de vocabulário requerida por esse nó não ocorra, e, então, um IV menos especificado é inserido em seu lugar.



Essa operação é relevante para os contextos de sincretismos, ou seja, casos em que um único IV é inserido em todo o paradigma – ou entre paradigmas – sem levar em consideração os traços específicos a cada um deles.

As operações de abaixamento e deslocamento local representam uma subdivisão da operação *Merger* Morfológico (cf. MARANTZ, 1988) proposta por Embick e Noyer (2001). Em linhas gerais, são operações de movimento pós-sintático, em que constituintes são deslocados a fim de garantir a ordem linear esperada. A subdivisão do *Merger* Morfológico foi baseada na temporalidade em que os movimentos ocorrem, ou seja, se antes ou depois da inserção de vocabulário.

O abaixamento, tal como as três primeiras operações descritas acima, é um tipo de *Merger* Morfológico que ocorre antes da inserção de vocabulário, pois deve ter acesso à estrutura sintática para que ocorra. O deslocamento local e a inserção de morfemas dissociados, tais como vogal de ligação dos compostos, são operações que ocorrem durante a linearização, pois independem de informação estrutural para a sua efetuação.

A MD se concentrou nos processos fonológicos segmentais que ocorrem simultaneamente à inserção de vocabulário, como pode ser visto nos trabalhos de Halle (1997), Embick e Noyer (2001), Embick (2007, 2010), dentre outros. Poucos trabalhos, como veremos na próxima subseção, detiveram-se na discussão suprasegmental, mais especificamente na atribuição de acento, a saber: Marvin (2003) e Lowenstamm (2010).

Apesar disso, é apenas no trabalho de Embick e Noyer (2001) que vemos um primeiro posicionamento acerca da localização da formação de domínios prosódicos na MD, tal como representado na Figura 2, repetida do Capítulo 2, e empregada nesta dissertação com as devidas adaptações:

**Figura 2** - O ramo de PF da gramática adaptado de EMBICK e NOYER (2001, p. 566)

Nessa visão, os domínios prosódicos são construídos tardiamente, após a inserção de vocabulário e a linearização da estrutura sintática. Embick e Noyer (2001) argumentam que a inserção de vocabulário ocorre antes da linearização, pois alguns processos morfológicos, como deslocamento local, necessitam da informação fonológica e da estrutura sintática para ocorrer.

O que vamos endossar nesse ponto é que a aplicação de acento na estrutura linearizada sinaliza o domínio da palavra prosódica, e, portanto, precisa ocorrer concomitantemente à sua formação, a fim de que outros processos fonológicos se apliquem. A distinção entre o número de acentos varia de acordo com a presença ou não do marcador de classe, que será introduzido ou não antes da linearização, a depender da presença do traço de classe C no núcleo categorizador, como veremos detalhadamente na seção 6.5.

#### **6.4.1. A noção de fase na palavra**

Os desenvolvimentos ocorridos na teoria sintática, particularmente aqueles associados ao Programa Minimalista (cf. CHOMSKY, 1993, 1995, e trabalhos subsequentes), passaram a assumir que a estrutura gerada no componente sintático deveria ser enviada para as interfaces, fonológica (PF) e semântica (LF), em diferentes ciclos, comumente denominados de fases (cf. CHOMSKY, 2001).

Essa visão cíclica assumida para os constituintes sentenciais foi expandida por Marantz (2000, 2001) ao domínio da palavra. O autor propunha que o primeiro núcleo categorizador forma um domínio com a raiz, de acordo com a configuração estrutural em (197a), fazendo com que o segundo categorizador concatenado à estrutura não tenha acesso direto à raiz, mas apenas à combinação raiz e primeiro categorizador, como em (197b), desencadeando, assim, consequências semânticas e fonológicas na interpretação da estrutura:

Tal assunção é evidenciada nos trabalhos de Arad (2003, 2005), a partir de um conjunto de argumentos semânticos. A autora mostra que há uma diferença entre a (i) criação de palavras a partir de raízes, ou seja, a partir da concatenação de núcleos categorizadores diretamente à raiz, e a (ii) criação de palavras a partir de palavras já existentes, ou seja, a partir da concatenação de um segundo categorizador a uma raiz já categorizada.

Essa diferença reside na seguinte distribuição empírica: em (i), há a possibilidade da ocorrência de idiossincrasias na negociação do significado da raiz, ao passo que, em (ii), o segundo núcleo categorizador, por não ter acesso direto à raiz, não influenciará no seu significado, podendo apenas somar-se ao significado já negociado da raiz com o primeiro categorizador. Um exemplo desse fato pode ser dado com a raiz do hebraico,  $\sqrt{xšb}$ , em (198):

|          |              |           |                  |                                                          |            |
|----------|--------------|-----------|------------------|----------------------------------------------------------|------------|
| (198) a. | $\sqrt{xšb}$ | CaCaC     | <i>xašav</i>     | ‘pensar’                                                 | (Hebraico) |
| b.       |              | CiCCeC    | <i>xišev</i>     | ‘calcular’                                               |            |
| c.       |              | hiCCiC    | <i>hexšiv</i>    | ‘considerar’                                             |            |
| d.       |              | hitCaCCeC | <i>hitxašev</i>  | ‘ser considerado’                                        |            |
| e.       |              | taCCiC    | <i>taxšiv</i>    | ‘cálculo’                                                |            |
| f.       |              | maCCaCa   | <i>maxšava</i>   | ‘pensamento’                                             |            |
| g.       |              | maCCeC    | <i>maxšev</i>    | ‘computador’                                             |            |
| h.       |              | CiCCeC    | <i>mixšev</i>    | ‘computar’, de <i>maxšev</i> ‘computador’                |            |
| i.       |              | CiCCon    | <i>xešbon</i>    | ‘explicação, aritmética, cálculo, conta’                 |            |
| j.       |              | hitCaCCeC | <i>hitxašben</i> | ‘acertar as contas com alguém’, de <i>xešbon</i> ‘conta’ |            |

É importante ter em mente que, no hebraico, os núcleos categorizadores são padrões vocálicos (e.g., **CaCaC**), os quais preenchem raízes tri-consonantais impronunciáveis, fornecendo a elas tanto informação categorial e gramatical (e.g., alteração de voz), quanto a possibilidade de realização fonológica. Como podemos ver em (198), a formação de (198a) a (198g) ocorre com a inserção de um padrão vocálico – ou seja, de um núcleo categorizador – diretamente à raiz, possibilitando, na visão de Arad (2003, 2005), a multiplicidade de significados verificada.

Por outro lado, a formação dos verbos *mixšev* ‘computar’, em (198h), e *hitxašben* ‘acertar as contas’, em (198j), ocorre a partir dos nomes *maxšev* ‘computador’, em (198g), e de *xešbon* ‘conta’, em (198i), respectivamente, o que faz com que seus significados sejam baseados no significado já negociado da raiz em sua primeira categorização. Essa definição de localidade relaciona-se, intimamente, com a noção de fase no domínio da palavra, algo que foi tomado como evidência para o desencadeamento do *Spell-Out* nesse ponto da derivação, tal como ilustrado em (199):

Marvin (2003), contemporaneamente à Arad (2003), endossa a visão de Marantz (2000, 2001) pelo lado fonológico, afirmando que ao se considerar o primeiro categorizador como um núcleo de fase é possível mapear diretamente questões acentuais e outros processos fonológicos na raiz. Sua proposta de derivação por fases é representada da seguinte maneira:

Sua perspectiva fonológica, no entanto, baseia-se em uma teoria de acesso direto, visando demonstrar que o primeiro categorizador não é responsável somente pelo significado idiossincrático da raiz, mas também por sua pronúncia idiossincrática. Assim, se o primeiro categorizador  $x_1$ , na configuração em (201a), induz a aplicação de uma regra fonológica, então essa regra terá um efeito na realização fonológica da raiz e isso será mantido por toda a derivação. Por outro lado, se o mesmo categorizador for concatenado a uma raiz já categorizada, como em (201b), ele não terá acesso direto à raiz, pois essa já terá sido enviada à interface fonológica, e o potencial que  $x_1$  tem para alterar o material fonológico da raiz não se aplica devido à sua posição estrutural.

Um exemplo ilustrativo da análise de Marvin (2003) pode ser dado com a inserção do schwa e a silabificação em *metric* e *metering*. Em (202a), o *Spell-Out* da raiz ocorre no domínio de  $nP$ , estando a raiz ainda acessível a  $n$ . A sequência  $[\sqrt{\phantom{x}} + n]$  está sujeita a restrições fonológicas no inglês, o que desencadeia a silabificação da consoante final com a inserção do schwa, gerando /mi:tər/.

Em (202b), o *Spell-Out* da raiz ocorre no domínio de  $aP$ , com  $a$  sendo capaz de influenciar o *Spell-Out* da raiz. Isso faz com que a sequência  $[\sqrt{\phantom{x}} + a]$  se silabifique sem a inserção do schwa, uma vez que a consoante final de *metr-* pode se transformar no ataque da sílaba seguinte, ocasionando /metrik/.

Em (202c), a raiz sofre o *Spell-Out* em  $vP$ . Com isso, as regras fonológicas já terão se aplicado e o schwa já terá sido introduzido, o que faz com que o sufixo *-ing* tenha acesso ao  $v$  de  $vP$ , mas não à raiz, de acordo (201), o que impossibilita alterar sua silabificação para que a raiz se transforme em \*/metriŋ/.

- (202) a.  $nP$   
 $\begin{array}{cc} n & \checkmark \\ \emptyset & metr- \end{array}$   
 → silabificação e inserção do schwa dentro do  $nP$   
 → *Spell-Out* da  $\checkmark$  /mi:tər/
- b.  $aP$   
 $\begin{array}{cc} a & \checkmark \end{array}$   
 → silabificação dentro do  $aP$  sem inserção do schwa  
 → *Spell-Out* da  $\checkmark$  /mɛtrik/
- c.  $nP$   
 $\begin{array}{cc} n & vP \\ & \begin{array}{cc} v & \checkmark \\ \emptyset & metr- \end{array} \end{array}$   
 → inserção de schwa já negociada na fase  $vP$ /  $n$  não pode influenciar o *Spell-Out* da  $\checkmark$ , \*/mɛtriŋ/, /mi:təriŋ/.  
 → inserção de schwa e silabificação de  $vP$ / *Spell-Out* da  $\checkmark$  /mi:tər/

Entretanto, temos um problema se assumirmos a proposta acentual de Lee (1995) e a proposta de fase na palavra tal como definida na abordagem Marantz-Arad-Marvin, uma vez que o componente fonológico não terá acesso à informação categorial e aos traços presentes no feixe de traços do núcleo categorizador, caso o *Spell-Out* ocorra no complemento do primeiro núcleo de fase. Isso, conseqüentemente, impedirá a aplicação do acento, visto que, para Lee (*op. cit.*), a informação categorial é imprescindível a fim de que as regras de acento sejam desencadeadas<sup>176</sup>.

Uma alternativa, nesse sentido, seria a proposta de Lowenstamm (2010), que, com intuito de demarcar a posição correta da aplicação de acento nas palavras, considera que afixos derivacionais são raízes, e que toda estrutura sintática correspondente à palavra é categorizada uma única vez por um núcleo categorizador fonologicamente nulo, enviando, dessa forma, a estrutura sintática da palavra de uma só vez para as interfaces, tal como representado em (203):

<sup>176</sup> Além disso, a hipótese de que o primeiro categorizador é um núcleo de fase encontra alguns contra-argumentos na literatura sintática. Borer (2009b, 2009c, 2012), por exemplo, traz uma série de dados evidenciando que o domínio do primeiro categorizador com a raiz não é responsável por atribuir significados idiossincráticos, pois tais significados também ocorrem quando um segundo categorizador se concatena à estrutura (e.g., a raiz do hebraico  $\checkmark xrb$  ‘relacionada à destruição, desolação’, da qual um nome derivado *xerbon* ‘desastre, falha’ tem derivações verbais subsequentes contendo significados idiossincráticos, tais como: (a) *xirben* i. causou um desastre, uma falha/ ii. defecado, e (b) *hitxarben* i. sofreu um desastre/ uma falha/ ii. comeu a menstruar).

O que temos, de acordo com a proposta de Lowenstamm (2010), é a combinação direta entre raízes, visto que a raiz, enquanto um objeto sintático, não desencadeia o *Spell-Out* da estrutura. Assim, a combinação de raízes deve aguardar a concatenação de um núcleo categorizador para que, então, seja enviada às interfaces. Essa proposta, segundo Lowenstamm, é capaz de explicar a mudança de acento decorrente da concatenação de afixos derivacionais, por exemplo, *átom*, que tem seu acento deslocado para a borda direita da palavra quando concatenado a *-ic*, formando *atómic*, e outra vez quando concatenado a *-ity*, formando *atomíc**ity*.

Como aponta Lowenstamm, essa mudança na posição do acento não é permitida na abordagem Marantz-Arad-Marvin, pois o acento seria atribuído em *átom* e não poderia ser deslocado, visto que a concatenação do categorizador, ao desencadear o *Spell-Out* da raiz, faz com que a regra de acento seja aplicada, ocasionando, assim, o estabelecimento de suas marcas suprasegmentais. Dessa forma, mesmo com a adição do sufixo *-ic*, o acento não seria deslocado para a borda direita da palavra, pois o acento inicial fica congelado no ponto em que foi estabelecido, ou seja, em *\*átomic*, e não há possibilidade de movê-lo a uma fase seguinte a fim de gerar *atómic*.

Contudo, a proposta de Lowenstamm tampouco é satisfatória em termos operacionais, já que, além de conter um *look-ahead* ao demarcar na sintaxe o que deve ocorrer na fonologia, ao assumir que afixos derivacionais são raízes, há ainda uma exigência de que raízes se concatenem umas às outras na sintaxe, embora, ontologicamente, elas não contenham nenhum traço gramatical que permita tal concatenação, e sejam consideradas unidades impossibilitadas de ser operadas no espaço computacional por si só, tal como discutido na subseção 5.2.2, do Capítulo 5.

Além disso, a proposta de Lowenstann torna nebulosa a divisão entre morfemas derivacionais e raízes, principalmente, no que concerne à presença de conteúdo lexical nos últimos itens. Essa proposta não explica também por que uma raiz  $\sqrt{\text{IC}}$  não pode ser utilizada como uma forma livre na língua, caso seja concatenada a um núcleo definidor de categoria. É

por esses motivos que acreditamos que a sugestão de Lowenstamm não é aceitável para a criação dos domínios de fase na formação de palavras.

Recorremos, para tanto, à proposta de Embick (2010), que assume uma abordagem diferente daquelas supramencionadas, no que diz respeito ao ponto da estrutura sintática em que o *Spell-Out* da estrutura é desencadeado. Em sua teoria C<sub>1</sub>-LIN, o autor assume que núcleos categorizadores, tais como *n*, *v* ou *a*, são núcleos de fase, do mesmo modo que as últimas abordagens. No entanto, é apenas quando um núcleo categorizador é concatenado à estrutura que o que o núcleo de fase presente em seu complemento sofre *Spell-Out*. Para ilustrar essa proposta, tomemos a estrutura em (204):

Consideremos *x* e *y* como núcleos categorizadores, e, portanto, cíclicos, e *W* e *Z* como núcleos funcionais não-cíclicos. Sendo *x* um núcleo de fase, seu complemento ou domínio é o material em  $\sqrt{P}$ , enquanto que o núcleo de fase *x* na fase *xP*, tal como os elementos adjungidos a *xP*, e seus especificadores, são definidos como a borda da fase. Como no domínio de *x*, em (204), não há um núcleo cíclico, mas apenas a raiz, seu complemento não sofrerá *Spell-Out*. Será apenas quando um segundo núcleo cíclico, no caso *y*, concatenar-se à estrutura, que o complemento do núcleo *x* (i.e., a raiz,  $\sqrt{P}$ ), o núcleo *x* ele mesmo, e o material presente em sua borda (i.e., *W* e *Z*) sofrerão *Spell-Out*.

Nesse sentido, o domínio imediato de um categorizador com uma raiz não desencadeia o *Spell-Out* da estrutura sintática, pois não há um núcleo cíclico no complemento do primeiro categorizador, mas somente a raiz. Dentro dessa teoria de derivação por fases, garantimos que a informação categorial e os demais traços presentes no feixe de traços do categorizador sejam enviados juntamente com a raiz às interfaces, tornando possível que a regra de acento proposta por Lee (1995) seja aplicada.



Como veremos, na próxima seção, a possibilidade de mudança de acento, apontada por Lowenstamm (2010) como um problema para a abordagem Marantz-Arad-Marvin, não será um problema para C<sub>1</sub>-LIN, visto que, propomos, adicionalmente a essa teoria de fase, que IVs são marcados para a reaplicação da regra de acento, permitindo, assim, que o acento fornecido à raiz após o *Spell-Out* não fique congelado, mas se desloque para a borda direita da palavra fonológica quando necessário.

### 6.5. Mapeamento dos constituintes sintáticos em constituintes prosódicos com a aplicação do acento

Como definido no Capítulo 4, a estrutura de um composto é dada como em (205), na qual dois núcleos complexos, ou seja, duas raízes categorizadas,  $\alpha$ P e  $\beta$ P, concatenados em uma relação gramatical – seja ela de subordinação, atribuição ou coordenação – são recategorizados por um terceiro núcleo categorizador,  $\gamma$ , responsável por tornar a combinação sintática em um composto.

Assim, de acordo com a derivação por fases definida em C<sub>1</sub>-LIN, nem  $\alpha$  e nem  $\beta$ , embora gerados paralelamente no espaço computacional, desencadeiam o *Spell-Out* de seus complementos, ou seja, das raízes. Será apenas quando  $\gamma$  se concatenar à estrutura que tanto  $\alpha$  quanto  $\beta$ , sendo núcleos cíclicos, sofrerão *Spell-Out* juntamente com seus complementos e elementos de borda (caso houver):

Garantimos, dessa forma, que a informação categorial, tal como os demais traços presentes nos núcleos categorizadores sejam enviados em conjunto com a raiz às interfaces. Isso permite não apenas que a informação categorial esteja presente para determinar qual regra de acento aplicar (e.g., se marcada ou não-marcada), mas também permite que o traço de classe esteja presente para sinalizar a presença de uma palavra, desencadeando, assim, a aplicação da regra de acento sobre o conteúdo fonológico da raiz. Caso o traço de classe esteja





(207) a. Casos não-marcados: constituintes binários; núcleo à direita; direção: direita para esquerda; não-iterativo.

b. *antropomorf-*

( . \*)

Com a aplicação de acento, temos a demarcação de uma palavra fonológica, correspondente ao composto todo, o que deflagra a isomorfia entre a estrutura sintática e os domínios para aplicação de regras fonológicas advogada pela fonologia prosódica, pois será internamente a esse domínio prosódico criado que os processos fonológicos relativos à palavra fonológica poderão ocorrer posteriormente.

Vale mencionar, ainda, que a ausência do traço de classe no categorizador  $n_1$  é o que impede o preenchimento fonológico dos traços  $\varphi$  em seu feixe, visto que sem o marcador de classe não é possível concatenar demais afixos ao radical, tal como discutido no Capítulo 5. É por esse motivo que nos compostos [radical + radical] a realização fonológica dos traços  $\varphi$  ocorre apenas na borda direita do composto, já que é apenas nesse ponto que há um marcador de classe realizado; no caso de (206a), /e/.

Caso um categorizador como *-ia* seja inserido na sequência, o IV desse categorizador carrega uma informação capaz de desencadear a reaplicação da regra de acento, seja ela marcada ou não-marcada, fazendo com que o acento já atribuído à sequência linearizada (i.e., *antropomórf-*) seja movido da posição onde foi aplicado pela primeira vez para a borda direita da palavra fonológica. Essa hipótese permite uma ciclicidade pós-sintática, contrária ao congelamento do acento apontado por Lowenstamm (2010), e pode ser formulada como na hipótese em (208), abaixo:

(208) *Hipótese da ciclicidade pós-sintática*

Um IV marcado desencadeia a aplicação da regra de acento deslocando o acento para a borda direita da palavra fonológica.

O sufixo *-ia*, por exemplo, é marcado para reaplicação da regra de acento, e desencadeia o algoritmo para não-verbos não-marcado, fazendo com que o acento se desloque para a borda direita, como mencionado:

- (209) a. /-ia/: N \_\_\_ → \*reaplicação do acento, de acordo com o algoritmo não-marcado: constituinte binário; núcleo à direita; direção: direita para a esquerda; não-iterativo.

b. *antropomorfia*

[ . \*]

No composto [radical + palavra], em (206b), a presença de um traço de classe no categorizador da raiz de segunda posição, ou seja,  $\sqrt{\text{GINAST}}$ , permite a inserção de um marcador de classe, e, nesse ambiente, dá-se a aplicação do acento, gerando *ginást-*, o qual é preservado após a concatenação de *hidr-*, de modo semelhante à solução fornecida por Nespor e Ralli (1996), ao assumir o princípio da preservação do acento de Burzio (1994). No entanto, o primeiro radical em compostos [radical + palavra], como vimos em 6.3.3.2, apesar de não ter a inserção de um marcador de classe, recebe a aplicação de acento.

Para garantir a aplicação de acento no radical de compostos [radical + palavra], sugerimos que o categorizador de suas raízes contém um traço de classe [*i*C: 4], utilizado em nomes como *mar*, *jovem*, *homem*, que não apresentam um marcador de classe realizado fonologicamente (i.e., *hidr-∅*), mas são visíveis para a regra de acento. Isso explicaria por que os radicais na primeira posição em compostos [radical + palavra], no PB, recebem um acento, formando um composto com duas palavras fonológicas, a saber: *hidr-o-ginástica*.

Nos compostos [palavra + palavra], tal como *peixe-espada* em (206c), a aplicação das regras de acento e, conseqüentemente, da criação da palavra fonológica, segue de maneira direta, uma vez que cada um dos categorizadores de suas raízes apresentam traços de classe e são interpretados como palavras para o desencadeamento da regra, formando, desse modo, duas palavras fonológicas independentes, como representado em (210):

- (210) *peix-e* *espada*

/peix/ /espada/

[\*] [ . \*]

Nesse tipo de compostos, ambos os acentos são primários no nível da palavra fonológica (i.e., *péixe*, *espáda*), enquanto a proeminência do segundo membro é determinada no nível da frase fonológica (i.e., *pèixe espáda*), assemelhando-se, assim, a um sintagma como “*mesa redonda*”. Nespor e Vogel (1996) argumentam que a eventual leitura não-

composicional presente em um composto não será determinada pela fonologia, mas sim pelo seu comportamento morfossintático e o seu significado particular.

Para nós, no entanto, o responsável por atribuir a diferença interpretativa entre *mesa-redonda* (i.e., reunião entre pessoas para discussão de um assunto) e *mesa redonda* (i.e., uma mesa com formato redondo) é a Enciclopédia (ou Lista 3), a qual irá atribuir o significado não-composicional do item composto com base no contexto em que essa combinação está inserida, visto que a estrutura sintática que caracteriza o composto já não está mais disponível nesse ponto da derivação.

## 6.6. Conclusões preliminares

Neste capítulo, demonstramos como ocorre a variação superficial nos compostos do PB, através do modo como os traços de sua estrutura sintática são lidos pelo componente morfológico da gramática. Fornecemos, ao mesmo tempo, uma descrição mais precisa dos aspectos fonológicos relativos às estruturas morfológicas do PB, argumentando que enquanto a estrutura [radical + radical] é mapeada em uma única palavra fonológica, as estruturas [radical + palavra] e [palavra + palavra] são mapeadas em duas palavras fonológicas, o que é explicado através da distribuição dos traços de classe nos núcleos categorizadores.

Sugerimos também que não há um mapeamento direto da estrutura sintática em uma estrutura fonológica, e que a construção dessa última é regulada de acordo com as informações presentes nos IV. Isso significa que, enquanto as regras de linearização distribuem os nós terminais nas posições encontradas na superfície, a inserção de IV ocorre simultânea e ciclicamente na atribuição do acento na estrutura linearizada, e, conseqüentemente, na construção da hierarquia prosódica.

## Considerações Finais

---

O objetivo geral desta dissertação foi fornecer uma explicação sobre quais são as informações mínimas necessárias para a formação sintática de um composto nas línguas naturais e quais informações presentes nesse processo de formação de palavras são específicas a uma língua particular. A explicação fornecida a essas questões foi elaborada a partir da discussão de quatro tópicos principais, a saber:

- (i) Caracterização do processo de composição em uma língua particular, o PB;
- (ii) Reunião das unidades mínimas da composição a serem abarcadas em um sistema gerativo universal;
- (iii) Explicação de como ocorre a derivação sintática dos compostos translinguisticamente;
- (iv) Explicação fonológica dos padrões acentuais dos compostos do PB, como um exemplar de externalização do processo de composição.

O tópico (i) foi desenvolvido no Capítulo 3, ao apresentar como a composição se distribui em um sistema linguístico específico, o PB, através da revisão das classificações propostas para os seus dados. A partir dessas classificações, filtramos quais delas explicitam propriedades intrínsecas ao processo de composição de um ponto de vista translinguístico, nomeadamente, (a) as combinações de categorias lexicais e (b) as relações gramaticais presentes entre seus membros, e quais delas descrevem propriedades específicas aos dados dessa língua, a saber, (c) a distribuição dos compostos em estruturas morfológicas.

Dentro de um modelo realizacional não-lexicalista, a MD, argumentamos que as estruturas morfológicas estão relacionadas ao modo como uma estrutura sintática é externalizada na interface fonológica, e que noções como “radical” e “palavra” dizem respeito às formas como uma raiz sintática é realizada fonologicamente no ramo de PF. Com base nessa assunção, apontamos que o PB apresenta três possibilidades de realização fonológica para as estruturas sintáticas de seus compostos, a saber: (i) [radical + radical], (ii) [radical + palavra] e (iii) [palavra + palavra].

Expandimos a hipótese supracitada à formação dos compostos das línguas naturais de modo geral, tomando por base o Princípio da Uniformidade (cf. CHOMSKY, 2001), o qual prediz que as línguas são uniformes de um ponto de vista gerativo e que sua variação se deve

ao componente fonológico (e morfológico, dentro da visão de gramática proposta pela MD). Desse modo, sugerimos que a formação de qualquer composto ocorre a partir de operações sintáticas gerais que geram estruturas sintáticas, presumivelmente, universais, ao passo que a variação translinguística atestada é resultado do modo como essas estruturas sintáticas serão externalizadas fonologicamente, através de diferentes “estruturas morfológicas”, como rotula a literatura tradicional.

Nesse sentido, o Capítulo 3 insere o leitor em uma perspectiva diferente de encarar a composição, pois foge do modo como as gramáticas tradicionais e os trabalhos morfológicos sobre a composição mais conhecidos concebe a distribuição desses dados no PB. Ao mesmo tempo, ele deflagra a amplitude do fenômeno ao explicitar que determinadas combinações categoriais e que um conjunto mínimo de relações gramaticais se estabelecem de modo uniforme translinguisticamente, e podem ser abarcadas da mesma forma a partir de um sistema gerativo universal.

A contribuição dada por essa proposta é a integração dos compostos de uma língua particular, principalmente, daqueles que são tratados como periféricos – no caso do PB, os compostos formados por radicais – como parte intrínseca do sistema computacional responsável por criar objetos complexos. Além disso, acreditamos que há ganhos descritivos quando iniciamos a descrição dos compostos de uma língua tendo em mente que há um conjunto variado de estruturas morfológicas disponíveis para a realização dessas palavras, pois isso tende a evitar qualquer negligência na descrição de alguns tipos de dados.

O tópico (ii), por sua vez, foi desenvolvido no Capítulo 4, onde reunimos as unidades mínimas para a formação sintática dos compostos das línguas naturais, ao discutir o processo de composição de modo amplo. Nosso intuito era elencar quais são as informações computacionais mínimas que devem ser explicadas no componente sintático para se gerar um composto. Dessa busca, chegamos a duas unidades mínimas: (a) as relações gramaticais, divididas em subordinação, atribuição e coordenação, e (ii) o domínio categorial, o qual deve ser formado acima de uma estrutura sintática com vias a garantir a interpretação desse complexo como uma única unidade.

Como resultado, chegamos à conclusão de que a formação sintática de um composto deve ser definida da seguinte maneira:



(211) *Composto*

Um composto é formado quando dois ou mais núcleos complexos, em determinada relação sintática, são recategorizados por um núcleo definidor de categoria – *n*, *v* ou *a*.

A partir dessa conclusão, passamos a discutir o tópico (iii), no Capítulo 5, com vias a formalizar as unidades mínimas supramencionadas. Admitimos que a natureza da operação *Merge*, tal como definida em Chomsky (2000, 2004), é capaz de abarcar como se processa o estabelecimento das relações gramaticais entre dois núcleos complexos, bem como a concatenação de um núcleo definidor de categoria acima dessa relação. Para tanto, assumimos que *set-Merge* é a variedade da operação *Merge* capaz de concatenar os membros de compostos subordinados, e que *pair-Merge* é a variedade responsável por concatenar os membros de compostos atributivos, sem ter que recorrer a projeções funcionais, como faz Di Sciullo (2005b, 2009).

Por fim, no Capítulo 5, exploramos o tópico (iv), ao tratar do modo como se dá a externalização dos compostos em uma língua particular. Nesse capítulo, além de mostrar como se dá a variedade superficial encontrada nos compostos do PB, argumentamos que a estrutura sintática da palavra não é mapeada diretamente em uma estrutura fonológica, e que, portanto, a formação de domínios prosódicos deve ser vista como parte dos processos no ramo de PF. Admitimos que a teoria de fases proposta por Embick (2010) é capaz de demarcar a quantidade ideal de estrutura sintática a ser interpretada pelas interfaces na formação dos compostos, e, em seguida, para proceder à atribuição do acento e criação do domínio da palavra fonológica.

## Referências Bibliográficas

---

ABELS, K. **Successive cyclicity, anti-locality, and adposition stranding**. 2003. Tese (Doutorado em Linguística) – University of Connecticut, Storrs, 2003.

ACQUAVIVA, P. Roots and Lexicality in Distributed Morphology. In: GALANI, A.; REDINGER, D.; YEO, N. (eds.) **6<sup>th</sup> Décembrettes Morphology in Bordeaux**, 2009, p. 50-62.

ADGER, D. **Core Syntax: A Minimalist Approach**. Oxford: Oxford University Press, 2003.

\_\_\_\_\_. A Minimalist Theory of Feature Structure. In: KIBORT, A.; CORBETT, G. (eds.) **Features: Perspectives on a Key Notion in Linguistics**. Oxford: Oxford University Press, 2010, p. 185-218.

ALCÂNTARA, C. da C. **As classes formais do português e sua constituição**: um estudo a luz da teoria da Morfologia Distribuída. 2003. Tese (Doutorado em Letras) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2003.

\_\_\_\_\_. As Classes Formais do Português Brasileiro. **Letras de Hoje**, v. 45, n. 1, 2010, p. 5-15.

ALEXIADOU, A. Inflection class, gender and DP internal structure. In: MÜLLER, G. et al (eds.) **Exploration in Nominal Inflection**. Berlim/Nova Iorque: Mouton de Gruyter, 2004, p. 21-50.

ALLEN, M. R. **Morphological Inverstigations**. 1978. Tese (Doutorado em Linguística) – University of Connecticut, Storrs, 1978.

ALVES, I. M. **Neologismo**: criação lexical. 3<sup>a</sup> ed. São Paulo: Ática, 2007.

AMIOT, D.; DAL, G. Integrating neoclassical combining forms into a lexeme-based morphology. In: BOOIJ, G. et al (eds.) **On-line Proceedings of the Fifth Mediterranean Morphology Meeting**, Universidade de Bolonha, 2007, p. 323-336.

ARAD, M. Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. **NLLT**, v. 21, 2003, p. 737-778.

\_\_\_\_\_. **Roots and Patterns**: Hebrew Morpho-syntax. Studies in Natural Language and Linguistic Theory. Amsterdam: Springer, 2005.

ARONOFF, M. **Word-Formation in Generative Grammar**. Cambridge, MA: MIT Press, 1976.

ARREGI, C.; NEVINS, A. **Morphotactics**: Basque Auxiliaries and the Structure of Spellout. Dordrecht: Springer, 2012.

AZEREDO, J. C. de. **Gramática Houaiss de Língua Portuguesa**. 2<sup>a</sup> ed. São Paulo: Publifolha, 2008.

BAESKOW, H. **Lexical Properties of Selected Non-native Morphemes of English**. Tübingen: Gunter Narr, 2004.

BAKER, M. C. **Incorporation**: a theory of grammatical function changing. Chicago: The University of Chicago Press, 1988.

\_\_\_\_\_. On agreement and its relation to case: Some generative ideas and results. **Lingua**, v. 130, 2013, p. 14-32.

BARBOSA, J. W. C.; MINUSSI, R. D. Compositionality in nouns and sentences: analyzing compounds and complex predicates. **Actas del V Encuentro de Gramática Generativa**, v. 1, 2009, p. 43-62.

BASCIANO, B.; KULA, N.; MELLONI, C. Modes of compounding in Bantu, Romance and Chinese. **Rivista di Linguistica**, v. 23, n. 2, 2011, p. 203-249.

BASSANI, I. de S. **Uma abordagem localista para a morfologia e estrutura argumental dos verbos complexos (parassintéticos) do português brasileiro**. 2013. Tese (Doutorado em Letras) – USP, São Paulo, 2013.

BAUER, L. Is there a class of neoclassical compound, if so, is it productive? **Linguistics**, v. 36, n. 3, 1998, p. 403-422.

\_\_\_\_\_. Compounding. In: HASPELMATH, M. et al. (eds.) **Language Typology and Language Universals**. Vol. 1. Berlin: Walter de Gruyter, 2001, p. 695-707.

\_\_\_\_\_. Typology of Compounds. In: LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, p. 345-356.

BECHARA, E. **Moderna Gramática Portuguesa**. 37<sup>a</sup> ed. Rio de Janeiro: Editora Nova Fronteira/Editora Lucerna, 2009.

BÉJAR, S. **Phi-Syntax: a Theory of Agreement**. 2003. Tese (Doutorado em Linguística) – Universidade de Toronto, Toronto, 2003.

BENIGNI, V.; MASINI, F. Compounds in Russian. **Lingue e Linguaggio**, v. 8, n. 2, 2009, p. 171-194.

BENVENISTE, É. **Problemas de Linguística Geral II**. Tradução: Eduardo Guimarães et. Al, 2<sup>a</sup> ed. Campinas: Pontes, 2006 [1974].

BERWICK, R. C.; CHOMSKY, N. The Biolinguistic Program: The Current State of its Development. In: DI SCIULLO, A-M.; BOECKX, C. (eds.) **The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty**. Oxford: Oxford University Press, 2011, p. 19-41.

BISETTO, A. Note sui Composti VN dell'Italiano. In: BENINCA, P.; MIONI, A.; VANELLI, L. **Fonologia e Morfologia dell'Italiano e dei Dialecti d'Italia**. Roma: Bulzoni, 1999, p. 555-589.

\_\_\_\_\_. Note su alcuni composti nominali dell'italiano. In: BISETTO, A.; IACOBINI, C.; THORNTON, A-M. (eds.) **Scritti di morfologia in onore di Sergio Scalise in occasione del suo 60° compleanno**. Cesena: Caissa Italia, 2003, p. 29-46.

\_\_\_\_\_. (ed.) Compounds Crosslinguistically. Edição Especial de **Lingue e Linguaggio**, v. 8, n. 2, 2009.

\_\_\_\_\_. Recursiveness and Italian compounds. **SKASE Journal of Theoretical Linguistics**, v. 7, n. 1, 2010.

\_\_\_\_\_. **Linking Elements and Recursive Compounds**. Alma Mater Studiorum – Università di Bologna. Manuscrito. 2013.

BISETTO, A.; MELLONI, C. Parasynthetic Compounds. **Lingue e Linguaggio**, v. 7, n. 2, 2008, p. 233-259.

BISETTO, A.; SCALISE, S. The Classification of Compounds. **Lingue e Linguaggio**, v. 4, n. 2, 2005, p. 319-332.

BISOL, L. **O Acento**: Duas Alternativas de Análise. UFRGS/PUCRS. Manuscrito. 1992.

\_\_\_\_\_. (org.) **Introdução a Estudos de Fonologia do Português Brasileiro**. Porto Alegre: EDIPUCRS, 1999.

\_\_\_\_\_. O clítico e seu status prosódico. **Revista de estudos da Linguagem**. Belo Horizonte, v. 9, n. 1, 2000, p. 5-30.

\_\_\_\_\_. Mattoso Câmara Jr. e a palavra prosódica. **D.E.L.T.A.**, v. 20, especial, 2004, p. 59-70.

\_\_\_\_\_. O clítico e seu hospedeiro. **Letras de hoje**. Porto Alegre, v. 40, n. 3, 2005, p. 163-184.

BLOOMFIELD, L. **Language**. Chicago: University of Chicago Press, 1933.

BOBALJIK, J. Distributed Morphology. In: SPENCER, A. (ed.) **Handbook of Morphology**. Londres: Wiley-Blackwell, no prelo.

BOECKX, C. **Defeating Lexicocentrism**: Outline of Elementary Syntactic Structure. ICREA – Universidad Autónoma de Barcelona. Manuscrito, 2010. [Disponível online em: <http://ling.auf.net/lingbuzz/001130>].

BOK-BENNEMA, R.; KAMPERS-MANHE, B. Taking a closer look at Romance VN compounds. In: NISHIDA, C.; MONTREUIL, J-P. Y. (eds.) **New Perspectives on Romance Linguistics – Vol. 1**: Morphology, Syntax, Semantics and Pragmatics. Amsterdam: John Benjamins Publishing, 2006, p. 13-27.

BOOIJ, G. **The Grammar of Words**. New York: Oxford University Press, 2005.

BOOIJ, G.; RUBACH, J. Postcyclic versus lexical rules in lexical phonology. **Linguistic Inquiry**, v. 18, 1987, p. 1-44.

BOPP DA SILVA, T. **Formação de Palavras Compostas em Português Brasileiro**: uma Análise de Interfaces. 2010. Tese (Doutorado em Letras) – UFRGS, Porto Alegre, 2010.

BORER, H. Morphology and Syntax. In: SPENCER, A.; ZWICKY, A. F. (eds.) **The Handbook of Morphology**. Oxford/Malden, MA: Blackwell, 1998, p. 151-190.

\_\_\_\_\_. Exo-skeletal vs. endo-skeletal explanations: Syntactic projections and the lexicon. In: MOORE, J.; POLINSKY, M. (eds.) **The Nature of Explanation in Linguistic Theory**. Chicago: University of Chicago Press, 2003, p. 31-66.

\_\_\_\_\_. **In Name Only**: Structuring Sense. Vol. 1. Oxford: Oxford University Press, 2005.

\_\_\_\_\_. Afro-Asiatic, Semitic: Hebrew. In: LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009a, p. 491-511.

\_\_\_\_\_. **Roots and Categories**. Trabalho apresentado no 19<sup>th</sup> Colloquium on Generative Grammar, University of Basque Country, 2009b.

\_\_\_\_\_. **Root Bound**. Trabalho apresentado no CASTL, University of Tromsø, 2009c.

\_\_\_\_\_. **Taking Form: Structuring Sense**. Vol. 3. Oxford: Oxford University Press, 2012.

BRESNAN, J. Polyadicity. In: \_\_\_\_\_. **The Mental Representation of Grammatical Relations**. Cambridge, MA: MIT Press, 1982, p. 140-172.

BROADBENT, J. Linking and Intrusive-r in English. **University College London Working Papers in Linguistics**, n. 3, 1991, p. 281-302.

BURZIO, L. **Principles of English Stress**. Cambridge: Cambridge University Press, 1994.

CÂMARA JR., J. M. **Problemas de Linguística Descritiva**. 20<sup>a</sup> ed. Petrópolis: Vozes, 2010 [1969].

\_\_\_\_\_. **Dicionário de Linguística e Gramática**. 25<sup>a</sup> ed. Petrópolis: Vozes, 2004[1977].

CARNIE, A. **Syntax: A Generative Introduction**. 3<sup>a</sup> ed. Oxford: Wiley-Blackwell, 2013.

CARSTENS, V. DP in Bantu and Romance. In: DE CAT, C. C.; DEMUTH, K. (eds.) **The Bantu-Romance Connection: a comparative investigation of verbal agreement, DPs and information structure**. Amsterdam e Filadélfia: John Benjamins, 2008, p. 131-166.

CECCAGNO, A.; BASCIANO, B. Compound headedness in Chinese: an analysis of neologisms. **Morphology**, v. 17, 2007, p. 207-231.

\_\_\_\_\_. Sino-Tibetan: Mandarin Chinese. In: LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, p. 478-490.

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

\_\_\_\_\_. A Minimalist Program for Linguistic Theory. In: HALE, K.; KEYSER, S. J. (eds.) **The view from building 20**. Cambridge, MA: MIT Press, 1993, p. 1-52.

\_\_\_\_\_. Bare Phrase Structure. In: **MIT Occasional Papers in Linguistics 5**, 1994 [Reimpresso em WEBELHUTH, G. (ed.) **Government and Binding Theory and the Minimalist Program**. Oxford: Blackwell, 1995, p. 383-439].

\_\_\_\_\_. **The Minimalist Program**. Cambridge, MA: MIT Press, 1995.

\_\_\_\_\_. Minimalist Inquiries: The Framework. In: MARTIN, R.; MICHAELS, D.; URIAGEREKA, J. **Step by Step**: Essays on Minimalist Syntax in Honor to Howard Lasnik. Cambridge, MA: MIT Press, 2000, p. 89-155.

\_\_\_\_\_. Derivation by Phase. In: KENSTOWICZ, M. (ed.) **Ken Hale: A life in language**. Cambridge, MA: MIT, 2001, p. 1-52.

\_\_\_\_\_. Beyond Explanation Adequacy. In: BELLETTI, A. (ed.) **Structures and Beyond**. Oxford: Oxford University Press, 2004, p. 104-131.

\_\_\_\_\_. Three Factors in Language Design. **Linguistic Inquiry**, v. 36, n. 1, 2005, p. 1-22.

\_\_\_\_\_. Problems of Projection. **Lingua**, v. 130, 2013, p. 33-49.

CHOMSKY, N.; HALLE, M. **The Sound Pattern of English**. Nova Iorque: Harper and How, 1968.

COLLINS, C. Eliminating Labels. In: EPSTEIN, D.; SEELY, D. (eds.) **Derivation and Explanation in the Minimalist Program**. Malden, MA: Blackwell, 2002, p. 42-64.



COLLISCHONN, G. O acento em português. In: BISOL, L. (org.) **Introdução a Estudos de Fonologia do Português Brasileiro**. Porto Alegre: EDIPUCRS, 1991.

CONTINI MORAVA, E. Swahili morphology. In: KAYE, A. S. (ed.) **Morphology of Asia and Africa**. Winona Lake, Indiana: Eisenbrauns, 2007, p. 1129-1158.

CONTRERAS, H. Spanish Exocentric Compounds. In: NUESSEL, F. (ed.) **Current Issues in Hispanic Phonology and Morphology**. Bloomington: Indiana University Linguistics Club, 1985, p. 14-27.

CORBETT, G. **Gender**. Cambridge: Cambridge University Press, 1991.

COWAN, M. M. **Tzotzil Grammar**. Norman: Summer School of Linguistics of the University of Oklahoma, 1969.

CUNHA, C.; CINTRA, L. **Nova gramática do português contemporâneo**. 4ª ed. Rio de Janeiro: Lexicon Editora Digital, 2007.

CYRINO, S.; ESPINAL, M. T. Bare nominals in Brazilian Portuguese: more on the DP/NP analysis. **Natural Language and Linguistic Theory**, no prelo.

DE BELDER, M.; VAN CRAENENBROECK, J. **How to Merge a Root**. Utrecht University. Manuscrito. 2011. [Disponível em: <http://ling.auf.net/lingbuzz/001226>].

\_\_\_\_\_. **On Vocabulary Insertion**. Utrecht University. Manuscrito. 2013. [Disponível em: <http://ling.auf.net/lingbuzz/001837>].

DE DREU, M. **The internal structure of the Zulu DP**. 2008. Dissertação (Mestrado em Linguística) – Universidade de Leiden, Leiden, 2008.

DELFITTO, D.; FÁBREGAS, A.; MELLONI, C. Compounding at the interfaces. **Proceedings of the 39º NELS**, v. 39, n. 1, 2011, p. 253-268.

DELFITTO, D.; MELLONI, C. Compounds don't come easy. **Lingue e Linguaggio**, v. 8, n. 1, 2009, p. 75-104.

DESMETS, M.; VILLOING, F. French VN lexemes: morphological compounding in HPSG. In: MÜLLER, S. (ed.) **Proceedings of the HPSG09 Conference**. Stanford: CLSI Publications, 2009, p. 89-109.

DI SCIULLO, A-M. Deverbal compounds and the external argument. In: ROCA, I. (ed.) **Thematic Structure: Its Role in Grammar**. Dordrecht: Foris, 1992, p. 65-78.

\_\_\_\_\_. **Asymmetry in Morphology**. Cambridge, MA: MIT Press, 2005a.

\_\_\_\_\_. Decomposing Compounds. **SKASE Journal Of Theoretical Linguistics**, v. 2, n. 3, 2005b, p. 14-33.

\_\_\_\_\_. Why are compounds a part of human language? A view from asymmetry theory. In: LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, p. 145-177.

DI SCIULLO, A-M.; WILLIAMS, E. **On the Definition of Word**. Cambridge, MA: MIT Press, 1987.

DON, J. IE, Germanic: Dutch. In: LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, p. 370-385.

DONALIES, E. **Grammatik des Deutschen im europäischen Vergleich: Kombinatorische Begriffsbildung – Vol. 1: Substantivkomposition**. Mannheim: Institut für deutsche Sprache, 2004.

DRESSLER, W. U. Compound types. In: LIBBEN, G.; JAREMA, G. (eds.) **The Representation and Processing of Compound Words**. Oxford: Oxford University Press, 2006, p. 23-44.

DRYER, M. S. The Greenbergian word order correlations. **Language**, v. 68, n. 1, p. 81-138.

EMBICK, D. On the Structure of Resultative Predicates in English. **Linguistic Inquiry**, v. 35, 2004, p. 355-392.

\_\_\_\_\_. Linearization and Local Dislocation: Derivational mechanics and interactions. **Linguistic Analysis**, v. 33, n. 3-4, 2007, p. 2-35.

\_\_\_\_\_. **Localism versus Globalism in Morphology and Phonology**. Cambridge, MA: MIT Press, 2010.

EMBICK, D.; NOYER, R. Movement operations after syntax. **Linguistic Inquiry**, v. 32, 2001, p. 555-595.

\_\_\_\_\_. Distributed Morphology and the Syntax/Morphology Interface. In: RAMCHAND, G.; REISS, C. (ed.) **The Oxford Handbook of Linguistic Interfaces**. Oxford: Oxford University Press, 2007, p. 289-324.

FABB, N. Compounding. In: SPENCER, A; ZWICKY, M. (eds.) **The Handbook of Morphology**. Malden: Blackwell Publishers, 2001, p. 66-83.

FÁBREGAS, A. **The definition of the grammatical category in a syntactically oriented morphology**: the case of nouns and adjectives. 2005. Tese (Doutorado em Linguística) – Universidad Autónoma de Madrid, Madrid, 2005.

FRANCO, L. **Romance VN compounds, Phrasal Spell-Out and Rebooting**. Università degli Studi di Firenze. Manuscrito. 2012. [Disponível em: <http://ling.auf.net/lingbuzz/001662>].

FROMKIN, V.; RODMAN, R.; COLLINS, P.; BLAIR, D. **An Introduction to Language**. 3<sup>a</sup> ed. Sydney: Harcourt Brace, 1996.

GAETA, L.; GROSSMANN, M. (eds.) Compounds between Syntax and the Lexicon. Edição Especial da **Rivista di Linguística**, v. 21, n. 1, 2009.

GAETA, L.; RICCA, D. Composita solvantur: Compounds as lexical units or morphological objects?. **Rivista di Linguistica**, v. 21, n. 1, 2009, p. 35-70.

GONÇALVES, C. A. V. Cruzamento Vocabular em Português: A questão das fronteiras com outros processos de criação de palavras. III Congresso Internacional da ABRALIN. **Anais do III Congresso da ABRALIN**. Niterói: UFF – Centro de Estudos Gerais, v. 1, 2003a, p. 824-831.

\_\_\_\_\_. *Blends* lexicais em português: não-concatenatividade e correspondência. **Veredas UFJF**, v. 7, n. 1-2, 2003b, p. 149-167.

\_\_\_\_\_. Compostos neoclássicos: estrutura e formação. **ReVEL – Revista Virtual de Estudos da Linguagem**, v. 9, n. 5, 2011a, p. 6-39.

\_\_\_\_\_. Composição e derivação: polos prototípicos de um continuum? Pequeno estudo de casos. **Domínios de Lingu@gem**, v. 5, 2011b, p. 62-89.

GÖSKEL, A. Compounds in Turkish. **Lingue e Linguaggio**, v. 8, n. 2, 2009, p. 213-238.

GRAČANIN-YUKSEK, M. V-N Compounds in Italian: a case of agreement in word formation. NISHIDA, C.; MONTREUIL, J-P. Y. (eds.) **New Perspectives on Romance Linguistics – Vol. 1: Morphology, Syntax, Semantics and Pragmatics**. Amsterdam: John Benjamins Publishing, 2006, p. 113-126.

GREENBERG, J. H. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In: \_\_\_\_\_. (ed.) **Universals of Language**. Cambridge, MA: The MIT Press, 1963, p. 73-113.

GUEVARA, E.; SCALISE, S. Searching for universals in compounding. In: SCALISE, S.; BISETTO, A.; MAGNI, E. **Universals of Language Today**. Amsterdam: Springer, 2009, p. 101-128.

HALE, K.; KEYSER, S. J. **Prolegomena to a Theory of Argument Structure**. Cambridge, MA: MIT Press, 2002.

HALLE, M. Prolegomena to a theory of word-formation. **Linguistic Inquiry**, v. 4, n. 1, 1973, p. 3-16.

\_\_\_\_\_. On Stress and Accent in Indo-European. **Language**, v. 73, 1997, p. 275-313.

\_\_\_\_\_. Distributed Morphology: Impoverishment and fission. **MIT Working Papers in Linguistics 30: PF – Papers at the interface**. Cambridge, MA: MIT Press, 1997, p. 425-449.

HALLE, M.; MARANTZ, A. Distributed Morphology and the pieces of inflection. In: HALE, K.; KEYSER, S. J. (eds.) **The view from Building 20**. Cambridge, MA: MIT Press, 1993, p. 111-176.

\_\_\_\_\_. Some Key Features on Distributed Morphology. In: CARNIE, A.; HARLEY, H.; BURES, T. (eds.) **Papers on Phonology and Morphology**, v. 21. Cambridge, MA: MIT Papers in Linguistics MITWPL, 1994, p. 275-288.

HALLE, M.; VERGNAUD, J. R. **An Essay on Stress**. Cambridge, MIT Press, 1987.

HARÐARSON, G. R. **Peeling away the layers of the onion: on layers, inflection and domains in Icelandic compounds**. University of Connecticut. Manuscrito. 2013.

HARLEY, H. **Semantics in Distributed Morphology**. University of Arizona. Manuscrito. 2008. [Disponível online em: <http://ling.auf.net/lingbuzz/001528>].

\_\_\_\_\_. Compounding in Distributed Morphology. In: LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, p. 130-144.

\_\_\_\_\_. On the identity of roots. **Theoretical Linguistics**, no prelo.

HARLEY, H.; NOYER, R. Distributed Morphology. **GLOT International**, v. 4, n. 4, 1999, p. 3-9.

HARRIS, J. W. The exponence of gender in Spanish. **Linguistic Inquiry**, v. 22, 1991, p. 27-62.

\_\_\_\_\_. Nasal depalatalization no, morphological well-formedness sí: the structure of Spanish word classes. In: ARREGI, K. et al. (eds.) **MIT Working Papers in Linguistics 33: Papers on Morphology and Syntax – Cycle One**. Cambridge, MA: MIT Press, 1999, p. 47-82.

HASPELMATH, M. **Understanding Morphology**. Londres: Arnold, 2002.

HASPELMATH, M; SIMS, A. D. **Understanding Morphology**. Londres: Hodder Education – An Hachette UK Company, 2010.

HAYES, B. The Prosodic Hierarchy in Meter. In: KIPARSKY, P.; YOUNG, G. (eds.) **Rhythm and Meter**. Orlando: Academic Press, 1989, p. 220-249.

\_\_\_\_\_. **Metrical Stress Theory: Principles and Case Studies**. Los Angeles, University of California, 1991.

\_\_\_\_\_. **Metrical Stress Theory: Principles and Case Studies**. Chicago: University of Chicago Press, 1995.

HORNSTEIN, N.; NUNES, J. Adjunction, Labeling and Bare Phrase Structure. **Biolinguistics**, v. 2, n. 1, 2008, p. 57-86.

INKELAS, S. **Prosodic constituency in the lexicon**. 1989. Tese (Doutorado em Linguística) – Stanford University, Stanford, 1989.

\_\_\_\_\_. Deriving Cyclicity. In: HARGUS, S.; KAYSSE, E. M. (eds.) **The Studies in Lexical Phonology**. San Diego: Academic Press, 1993.

IRURTZUN, A.; GALLEGU, A. J. Consequences of pair-Merge (at the interfaces). **ASJU**, **XLI-2**, 2007, p. 179-200.

IRWIN, P. Intransitive Sentences, Argument Structure, and the Syntax-Prosody Interface. In: WASHBURN, M. B. et al (eds.) **Proceedings of the 28<sup>th</sup> West Coast Conference on Formal Linguistics (WWCFL)**, 2011, p. 275-284.

JACKENDOFF, R. Compounding in the Parallel Architecture and Conceptual Semantics. In: LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, p. 105-128.

KAHNEMUYIPOUR, A. **The Syntax of Sentential Stress**. 2004. Tese (Doutorado em Linguística) – University of Toronto, Toronto, 2004.

\_\_\_\_\_. **The Syntax of Sentential Stress**. Oxford, UK: Oxford University Press, 2009.

KATAMBA, F. **Morphology**. Londres: The Macmillan Press, 1993.

KEHDI, V. **Formação de Palavras em Português**. São Paulo: Ática, 2002.

KIPARSKY, P. Word-Formation and the Lexicon. In: INGEMANN, F. (ed.) **Proceedings of the Mid-America Linguistics Conference**. Lawrence, KA, 1982.

KLEIN, W.; PERDUE, C. The Basic Variety (or: Couldn't natural languages be much simpler?). **Second Language Research**, v. 13, 1997, p. 301-347.

KORNFELD, L. M. IE, Romance: Spanish. LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, p. 436-452.

KRATZER, A; SELKIRK, E. Phase theory and prosodic spellout. **The Linguistic Review**, 2007, p. 93-135.

LADD, D. R. **Intonational Phonology**. Cambridge: CUP, 1996.

LANGENDOEN, T. D. Merge. In: CARNIE, A.; WILLIE, M.; HARLEY, H. (eds.) **Formal Approaches to Functional Phenomena: In Honor of Eloise Jelinek**. Amstedam: John Benjamins Publishing, 2003, p. 307-318.

LAZZARINI, J.P.; ARMELIN, P.; MINUSSI, R. Sobre o conhecimento enciclopédico de gênero. **Cadernos do NEMP**, v. 4, n. 1, 2013, p. 105-117.

LEBEAUX, D. Relative clauses, licensing, and the nature of derivation. In: ROTHSTEIN, S. (ed.) **Perspectives on phrase structure: heads and licensing**. San Diego: Academic Press, 1991, p. 209-240.

LEIRIA, L. L. **Em busca da palavra prosódica**. 2000. Tese (Doutorado em Letras) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2000.

LEE, S. H. **Morfologia e Fonologia Lexical do Português do Brasil**. 1995. Tese (Doutorado em Linguística) – UNICAMP, Campinas, 1995.

\_\_\_\_\_. Sobre os compostos do PB. **D.E.L.T.A.** São Paulo, v. 13, n. 1, 1997.

LEES, R. B. **The Grammar of English Nominalizations**. Bloomington: Indiana University Press/The Hague: Mouton, 1960.

LEGATE, J. A. **The structure of agents in implicit passives**. Trabalho apresentado em 41<sup>st</sup> annual meeting of the North East Linguistic Society, Philadelphia, 2010.

\_\_\_\_\_. Subjects in Achenese and the nature of the passive. **Language**, v. 88, n. 3, 2012, p. 495-525.

LEMLE, M. Resenha de “Problems of Projection”. **Linguística**, v. 9, n. 1, 2013, p. 29-30.

LEVI, J. N. **The Syntax and Semantics of Complex Nominals**. Nova York: Academic Press, 1978.

LIBBEN, G. Why study compound processing? An overview of the issues. In: LIBBEN, G.; JAREMA, G. (eds.) **The Representation and Processing of Compound Words**. Oxford: Oxford University Press, 2006, p. 1-22.



LIBERMAN, M.; PRINCE, A. On stress and linguistic rhythm”. **Linguistic Inquiry**, v. 8, 1977, p. 249-336.

LIEBER, R. **On the Organization of the Lexicon**. 1980. Tese (Doutorado em Linguística) – MIT, Cambridge-MA, 1980.

\_\_\_\_\_. Argument linking and compounds in English. **Linguistic Inquiry**, v. 14, 1983, p. 251-285.

\_\_\_\_\_. **Deconstructing Morphology**. Chicago: The University of Chicago Press, 1992.

\_\_\_\_\_. **Morphology and Lexical Semantics**. Cambridge: Cambridge University Press, 2004.

LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009a.

\_\_\_\_\_. Introduction: Status and Definition of Compounding. In: \_\_\_\_\_. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009b, p. 3-18.

LIMA, R. **Gramática Normativa da Língua Portuguesa**. 51<sup>a</sup> ed. Rio de Janeiro: José Olympio, 1972/2013.

LOWENSTAMM, J. **Derivational affixes as roots (Phrasal spellout meets English stress shift)**. CNRS. Manuscrito. 2010.

MAGALHÃES, J. S. **O Plano Multidimensional do acento na Teoria da Otimidade**. 2004. Tese (Doutorado em Letras) – UFRGS, Porto Alegre, 2004.

MARANTZ, A. Clitics, Morphological Merger, and the mapping to phonological structure. In: HAMMOND, M.; NOONAN, M. (eds.) **Theoretical Morphology**. San Diego: Academic Press, 1988, p. 253-270.

\_\_\_\_\_. **Cat as a phrasal idiom**. MIT. Manuscrito. 1996.

\_\_\_\_\_. No Escape from Syntax: Don't try Morphological Analysis in the privacy of your own Lexicon. In: DIMITRIADIS, L. S.; SUREK-CLARK, C.; WILLIAMS, A. **Proceedings of the 21<sup>st</sup> Penn Linguistics Colloquium**. Philadelphia: UPenn Working Papers in Linguistics, 1997, p. 201-225.

\_\_\_\_\_. **Roots**: The universality of roots and pattern morphology. Trabalho apresentado em Afro-Asiatic languages, University of Paris VII, 2000.

\_\_\_\_\_. **Words**. Trabalho apresentado em West Coast Conference on Formal Linguistics, University Southern California, 2001.

MARCHAND, H. **The Categories and Types of Present-Day English Word-Formation**. Wiesbaden: Otto Harrassowitz, 1960.

\_\_\_\_\_. **The categories and types of present-day English word-formation**: a synchronic-diachronic approach. 2<sup>a</sup> ed. Munich: Beck, 1969.

MARTINS, S. A. **Fonologia e Gramática Dâw**. Vol. 1. Utrecht: Netherlands Graduate School of Linguistics, 2004.

MARVIN, T. **Topics in the Stress and Syntax of Words**. 2002. Tese (Doutorado em Linguística) – MIT, Cambridge, 2003.

MCHOMBO, S. **The Syntax of Chichewa**. Cambridge: Cambridge University Press, 2004.

MINUSSI, R. D. Os nomes compostos no português brasileiro: uma análise morfossintática. **ReVEL**, edição especial n. 5, 2011.

\_\_\_\_\_. **Os sabores do nome**: um estudo sobre a seleção de argumentos e as nominalizações do hebraico. 2012. Tese (Doutorado em Letras) – USP, São Paulo, 2012.

MINUSSI, R. D.; NÓBREGA, V. A. A interface sintaxe-pragmática na formação de palavras: avaliando os pontos de acesso da Enciclopédia na arquitetura da gramática. **Veredas (UFJF)**, v. 18, n. 1, 2014, p. 161-184.

MOHANAN, K. P. **Lexical Phonology**. 1981. Tese (Doutorado em Linguística) – MIT, Cambridge, MA, 1982.

MORENO, C. **Morfologia nominal do português – um estudo de fonologia lexical**. 1997. Tese (Doutorado em Letras) – UFRGS, Porto Alegre, 1997.

\_\_\_\_\_. A formação de compostos no português. **Letras de Hoje**. Porto Alegre, v. 37, n. 1, 2002.

MUKAI, M. Recursive Compounds and Linking Morpheme. **International Journal of English Linguistics**, v. 3, n. 4, 2013, p. 36-49.

MÜLLER, G. On Decomposing Inflection Class Features: Syncretism in Russian Noun Inflection. In: GUNKEL, L.; MÜLLER, G.; ZIFONUN, G. (eds.) **Explorations in Nominal Inflection**. Berlin: Mouton de Gruyter, 2002.

MUNN, A. **Topics in the syntax and semantics of coordination structures**. 1992. Tese (Doutorado em Linguística) – University of Maryland, Maryland, 1992.

NAMER, F.; VILLOING, F. Saxifage et casse-pierre: quelles propriétés distinctives des mots composés VN et NV en français? **XXIX Romanistentages in Saarbrücken**, Alemanha, 2006, p. 1-17.

NEEF, M. IE, Germanic: German. In: LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, p. 386-399.

NESPOR, M. Word stress. In: van der HULST, H. (ed.) **Word Prosodic Systems in the Languages of Europe**. Berlim, Nova Iorque: Mouton de Gruyter, 1999, p. 1-115.

NESPOR, M.; RALLI, A. Morphology-Phonology Interface: Phonological domains in Greek Compounds. **The Linguistic Review**, v. 13, p. 357-382.

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

\_\_\_\_\_. **Prosodic Phonology**. Dordrecht-Holland: Foris Publications, 2007 [1986].

NIEMI, J. Compounds in Finnish. **Lingue e Linguaggio**, v. 8, n. 2, 2009, p. 237-256.

NÓBREGA, V. A. Explorando o emprego dos elementos sino-japoneses na formação de palavras do japonês. **ReVEL**, edição especial n. 5, 2011, p. 40-60.

\_\_\_\_\_. A Morfossintaxe da Composição Neoclássica. **Anais do VIII Congresso Internacional da Associação Brasileira de Linguística**, no prelo.

\_\_\_\_\_. **Derivando a Estrutura dos Compostos V-N**. Universidade de São Paulo. Manuscrito. 2013.

NTHELITEOS, D. A syntactic analysis of synthetic and phrasal compound formation in Greek. **Issues in Mediteranean Syntax**. Leiden: Brill Publications, 2013.

NUNES, J. Triangulismos e a sintaxe do português brasileiro. In: CASTILHO, A.T. de et al. (eds.) **Descrição, História e Aquisição do Português Brasileiro**. São Paulo: Fapesp; Campinas: Pontes, 2007, p. 25-35.

\_\_\_\_\_. Sideward Movement: Triggers, Timing, and Outputs. In: URIBE-ETXEBARRIA, M.; VALMALA, V. (eds.) **Ways of Structure Building**. Oxford: Oxford University Press, 2012.

NUÑES CEDEÑO, R. A. Headship assignment resolution in Spanish compounds. In: CAMPOS, H.; MARTÍNEZ-GIL, F. (eds.) **Current Studies in Spanish Linguistics**. Washington: Georgetown University Press, 1991, p. 537-598.

OLSEN, S. Composition. In: BOOIJ, G.; LEHAMAM, M.; MUGDAN, J. (eds.) **Morphology: An International Handbook on Inflection and Word-Formation**. Berlin: Mouton de Gruyter, 2000, p. 897-915.

\_\_\_\_\_. Copulative compounds: a closer look at the interface between morphology and syntax. In: BOOIJ, G.; van MARLE, J. (eds.) **Yearbook of Morphology 2000**. Dordrecht: Kluwer, 2001, p. 279-320.

PACKARD, J. **The morphology of Chinese**. Cambridge: Cambridge University Press, 2000.

PADROSA-TRIAS, S. Catalan Verbal Compounds and the Syntax-Morphology Competition. In: MONTERMINI, F.; BOYÉ, G.; HATHOUT, N. (eds.) **Selected Proceedings of the 5<sup>th</sup> Décembrettes: Morphology in Toulouse**. Sommerville, MA: Cascadilla Proceedings Project, 2007, p. 91-107.

PEPPERKAMP, S. **Prosodic Words**. Den Haag: Holland Academic Grafics (HIL Dissertations, 34), 1997.

PERKLES, M. **Compounds in English and in Russian: A comparative analysis**. 2008. Dissertação (Mestrado em Letras) – Universitat Wien, Viena, 2008.

PESETSKY, D. **Zero Syntax**. Cambridge, MA: MIT, 1995.

PESETSKY, D.; TORREGO, E. Tense, Case and the nature of syntactic categories. In: GUÉRON, J.; LERCARME, J. (orgs.) **The Syntax of Time**. Cambridge, MA: MIT Press, 2004, p. 495-537.

\_\_\_\_\_. Probes, Goals and Syntactic Categories. In: **Proceedings of the 7<sup>th</sup> annual Conference on Psycholinguistics**. Keio University, Japan, 2006.

\_\_\_\_\_. The syntax of valuation and interpretability of features. In: KARIMI, S.; SAMIAN, V.; WILKINS, W. (eds.) **Phrasal and clausal architecture**. Amsterdam: John Benjamins Publishing, 2007, p. 262-294.

PETROPOULOU, E. On the parallel between neoclassical compounds in English and Modern Greek. **Patras Working Papers in Linguistics**, v. 1, 2009, p. 40-58.

PFAU, R. **Grammar as a Processor**: a distributed morphology account of spontaneous speech errors. Amsterdam: John Benjamins Publishing, 2009.

PIRANI, L. Bound Roots in Mandarin Chinese and Comparison with European “Semi-Words”. In: CHAN, M.K.M.; KANG, H. (eds.) **Proceedings of the 20th North American Conference on Chinese Linguistics**. Columbus, Ohio: The Ohio University, 2008, p. 261-277.

PLAG, I. Morphology in pidgins and creoles. In: BROWN, K. (ed.) **Encyclopedia of Language and Linguistics**. 2<sup>a</sup> ed.. Oxford: Elsevier, Vol. 8, 2006, p. 304-308.

PRINCE, A. Relating to the grid. **Linguistic Inquiry**, v. 14, n. 1, 1983, p. 19-100.

RAINER, F.; VARELA, S. Compounding in Spanish. **Rivista di Linguistica**, v. 4, n. 1, 1992, p. 117-142.

RALLI, A. Compounds in Modern Greek. **Rivista di Linguistica**, v. 4, n. 1, 1992, p. 143-174.

\_\_\_\_\_. Compound markers and parametric variation. **Language Typology and Universals**, v. 61, n. 1, 2008a, p. 19-38.

\_\_\_\_\_. Greek Deverbal Compounds with Bound Stems. **Southern Journal of Linguistics**, v. 29, n. 1-2, 2008b, p. 150-173.

\_\_\_\_\_. I.E. Hellenic: Modern Greek. In: LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, p. 453-463.

\_\_\_\_\_. Compounding vs. Derivation. In: SCALISE, S.; VOGEL, I. (eds.) **Cross disciplinary issues in compounding**. Amsterdam: John Benjamins, 2010, p. 57-76.

\_\_\_\_\_. **Compounding in Modern Greek**. Nova Iorque/Londres: Springer, 2013.

RALLI, A.; KARASIMOS, A. The Bare-Stem Constraint in Greek Compound Formation. **Gengo Kenkyuu**, v. 135, 29-48.

REFSING, K. **The Ainu Language**. Aarhus: Aarhus University Press, 1986.

RICE, K. Athapaskan: Slave. In: LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, p. 542-563.

RIZZI, L. Null Subjects in Italian and the Theory of pro. **Linguistic Inquiry**, v. 17, 1986, p. 501-557.

ROCHA, L. C. de A. **Estruturas Morfológicas do Português**. 2ª ed. São Paulo: Editora WMF Martins Fontes, 2008.

ROEPER, T.; SIEGEL, M. E. A. A lexical transformation for verbal compounds. **Linguistic Inquiry**, v. 9, p. 109-128.

RONDININI, R. B. Análise das formações com -logo e -grafo segundo a morfologia derivacional. **ReVEL**, v. 7, n. 12, 2009, p. 1-29.

RONDININI, R. B.; GONÇALVES, C. A. V. Formações X-logo e X-grafo: um caso de deslocamento da composição para a derivação? In: **Textos Selecionados do XXII Encontro Nacional da Associação Portuguesa de Linguística (APL)**. Lisboa: Colibri, v. 22, 2006, p. 533-545.

RIO-TORTO, G. M.; RIBEIRO, S. Compounds in Portuguese. **Probus**, v. 24, n. 1, 2012, p. 119-145.

ROSENBERG, M. Classification, headedness and pluralization: corpus evidence from French compounds. **Acta Linguistica Hungarica**, 2007, p. 341-360.

SAID ALI, M. **Gramática Histórica da Língua Portuguesa**. 7ª ed. Rio de Janeiro: Edições Melhoramentos, 1971.

SANDLER, W.; LILLO-MARTIN, D. **Sign language and linguistic universals**. Cambridge: Cambridge University Press, 2006.

SANDMANN, A. J. **Morfologia Lexical**. São Paulo: Contexto, 1990.

\_\_\_\_\_. **Morfologia Geral**. São Paulo: Contexto, 1991.

SCALISE, S. **Generative Morphology**. Dordrecht: Foris, 1984.

\_\_\_\_\_. (ed.) The Morphology of Compounding. Edição Especial da **Rivista di Linguistica**, v. 4, n. 1, 1992.

SCALISE, S.; BISETTO, A. Composizione. In: \_\_\_\_\_. **La struttura delle parole**. Bologna: Il Mulino, 2008, p. 117-141.

\_\_\_\_\_. The Classification of Compounds. In: LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, p. 34-53.

SCALISE, S.; FÁBREGAS, A.; FORZA, F. Exocentricity in Compounding. **Genko Kenkyu**, v. 135, 2009, p. 49-84.

SCALISE, S.; MASINI, F. (eds.) On Romance Compounds. Edição Especial da **Probus**, v. 24, n. 1, 2012.

SCALISE, S.; VOGEL, I. (eds.) **Cross-disciplinary issues in compounding**. Amsterdam: John Benjamins Publishing, 2010a.

\_\_\_\_\_. Why compounding? In: \_\_\_\_\_. **Cross-disciplinary issues in compounding**. Amsterdam: John Benjamins Publishing, 2010b, p. 1-19.

SCHER, A. P.; NÓBREGA, V. A. Unifying neoclassical and stem-based compounds: a non-lexicalist approach. **Revista Linguística**, v. 10, n. 1, 2014, p. 74-98.



SCHROTEN, J. The head of Verb+Noun compounds in the Romance languages. **Iberia**, v. 2, 2010, p. 85-98.

SCHWINDT, L. C. **O prefixo do português brasileiro: análise morfofonológica**. 2000. Tese (Doutorado em Letras) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2000.

SELKIRK, E. Prosodic domains in phonology: Sanskrit revisited. In: ARONOFF, M.; KEEN, M.L. **Juncture: a collection of original papers**. Saratoga: Anma Libri, 1980, p. 107-129.

\_\_\_\_\_. **The Syntax of Words**. Cambridge, MA: MIT Press, 1982.

\_\_\_\_\_. **Phonology and Syntax: the relation between sound and structure**. Cambridge, MA: The MIT Press, 1984.

SHIBATANI, M. **The Languages of Japan**. Cambridge: Cambridge University Press, 1990.

SHIBATANI, M.; KAGEYAMA, K. Word formation in a modular theory of grammar: a case of post-syntactic compounds in Japanese. **Language**, v. 64, 1988, p. 451-584.

SIDDIQI, D. **Syntax within the Word: economy, allomorphy, and the argument selection in Distributed Morphology**. Amsterdam: John Benjamins Publishing, 2009.

SPENCER, A. **Morphological Theory: An Introduction to Word Structure in Generative Grammar**. Oxford: Blackwell, 1991.

ŠTEKAUER, P.; VALERA, S. **Universals, tendencies and typology in word-formation: cross-linguistic research**. Manuscrito. 2007.

ŠTEKAUER, P.; VALERA, S.; KÖRTVÉLYESSY, L. **Universals, tendencies and typology in word-formation: a cross-linguistic study**. Universidade Josef Šafarik. Manuscrito. 2008.

\_\_\_\_\_. **Word-Formation in the World's Languages: a typological survey**. Cambridge: Cambridge University Press, 2012.

STEPANOV, A. Late adjunction and minimalist phrase structure. **Syntax**, v. 4, n. 2, 2001, p. 94-125.

ŠTICHAUER, P. Compounds in Czech. **Lingue e Linguaggio**, v. 8, n. 2, 2009, p. 293-314.

ten HACKEN, P. Early generative approaches. In: LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, p. 54-77.

TONELI, P. M. A palavra prosódica no português brasileiro. **Anais do Seta**, v. 5, 2011.

WATANABE, A. **Case absortion and wh-agreement**. Dordrecht: Kluwer, 1996.

WILLIAMS, E. On the notions of “Lexically Related” and “Head of a Word”. **Linguistic Inquiry**, v. 12, 1981, p. 45-74.

WURMBRAND, S. The Merge Condition: a syntactic approach to selection. In: KOSTA, P.; SCHÜRCKS, L.; FRANKS, S.; RADEV-BORK, T. (eds.) **Minimalism and Beyond: Radicalizing the interfaces**. Amstedam: John Benjamins Publishing, no prelo.

VERCELOTTI, M. L.; MORTENSEN, D. R. A classification of compounds in American Sign Language: an evaluation of the Bisetto and Scalise framework. **Morphology**, v. 22, 2012, p. 545-579.

VIGÁRIO, M. **The Prosodic Word in European Portuguese**. Berlin: Mouton de Gruyter, 2003.

VILLALVA, A. Compounding in Portuguese. **Rivista di Linguistica**, v. 4, n. 1, 1992, p. 201-219.

\_\_\_\_\_. **Estruturas Morfológicas: Unidades e Hierarquias nas Palavras do Português**. 1994. Tese (Doutorado em Letras) – Universidade de Lisboa, Lisboa, 1994.

\_\_\_\_\_. Configuração Não-Binárias em Morfologia. In: Direcção da Associação Portuguesa de Linguística (org.) **Actas do X Encontro Nacional da Associação Portuguesa de Linguística**. Lisboa: Edições Colibri, 1995, p. 583-597.

\_\_\_\_\_. **Estruturas Morfológicas**: Unidades e Hierarquias nas Palavras do Português. Lisboa: Fundação Calouste Gulbenkian, 2000.

VILLOING, F. French Compounds. **Probus**, v. 24, n. 1, 2012, p. 29-60.

ZAMPONI, R. Arawakan: Maipure-Yavitero. In: LIEBER, R.; ŠTEKAUER, P. (eds.) **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, p. 584-593.

ZEIJLSTRA, H. There is only one way to Agree. **The Linguistic Review**, no prelo.

ZHANG, N. Root Merger in Chinese Compounds. **Studia Linguistica**, v. 61, n. 2, 2007, p. 170-184.

ZWART, J-W. Structure and order: asymmetric merge. In: BOECKX, C. (ed.) **Oxford Handbook of Linguistic Minimalism**. Oxford: OUP, 2010, p. 96-118.

ZWITSERLOOD, I. Morphology below the level of the sign: “frozen” forms and classifier predicates. In: QUER, I. (ed.) **Proceedings of the 8<sup>th</sup> Conference on Theoretical Issues in Sign Language Research**. Hamburg: Signum Verlag, 2008, p. 251-272.