



**UNIVERSIDADE FEDERAL DA BAHIA
INSTITUTO DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LÍNGUA E CULTURA**

YASMIM CONCEIÇÃO BORGES

**COMPOSTOS [NN]_N NA LÍNGUA PORTUGUESA
(sécs. XVII E XVIII)**

Salvador
2022

YASMIM CONCEIÇÃO BORGES

**COMPOSTOS [NN]_N NA LÍNGUA PORTUGUESA
(sécs. XVII E XVIII)**

Dissertação apresentada ao Programa de Pós-Graduação em Língua e Cultura do Instituto de Letras da Universidade Federal da Bahia – UFBA como requisito parcial para obtenção do grau de Mestre em Língua e Cultura.

Linha de Pesquisa: Linguística Histórica, Filologia e História da Cultura Escrita

Orientadora: Profa. Dra. Antonia Vieira dos Santos

Salvador
2022

Ficha catalográfica elaborada pelo Sistema Universitário de Bibliotecas (SIBI/UFBA),
com os dados fornecidos pelo(a) autor(a).

Borges, Yasmim Conceição

Compostos [NN]N na língua portuguesa (sécs. XVII e XVIII) / Yasmim Conceição Borges. -- Salvador, 2022.
82 f.

Orientadora: Antonia Vieira dos Santos.

Dissertação (Mestrado - Programa de Pós-Graduação em Língua e Cultura) -- Universidade Federal da Bahia, Instituto de Letras, 2022.

1. Morfologia. 2. Composição. 3. Compostos NN. 4. Séculos XVII e XVIII. 5. História da Língua. I. Santos, Antonia Vieira dos. II. Título.

TERMO DE APROVAÇÃO

YASMIM CONCEIÇÃO BORGES

COMPOSTOS [NN]_N NA LÍNGUA PORTUGUESA (sécs. XVII E XVIII)

Dissertação apresentada ao Programa de Pós-Graduação em Língua e Cultura do Instituto de Letras da Universidade Federal da Bahia – UFBA como requisito parcial para obtenção do grau de Mestre em Língua e Cultura.

Parecer da banca: A pesquisa desenvolvida pela mestrandia, apresentada em sua dissertação, aborda um aspecto relevante no âmbito da composição de palavras em português – os compostos [NN]_N –, contribuindo para o conhecimento do seu percurso diacrônico. Destacam-se, ainda, no trabalho desenvolvido pela discente, a seleção e a apreciação crítica do lastro bibliográfico, a fundamentação teórica, a adequação metodológica, o estabelecimento do *corpus* e a análise rigorosa e pormenorizada dos dados coletados. Pelas razões mencionadas, a banca é de parecer favorável à aprovação da dissertação e recomenda a sua futura publicação.

Aprovada em 24 de fevereiro de 2022.

BANCA EXAMINADORA

ANTONIA VIEIRA DOS SANTOS (Orientadora)
Doutora em Letras e Linguística pela Universidade Federal da Bahia (UFBA)
Universidade Federal da Bahia

SÍLVIA ISABEL DO ROSÁRIO RIBEIRO (Examinadora Externa à Instituição)
Doutora em Linguística Portuguesa pela Universidade de Coimbra (UC)
Universidade de Aveiro (UA)

MAILSON DOS SANTOS LOPES (Examinador Externo ao Programa)
Doutor em Língua e Cultura pela Universidade Federal da Bahia (UFBA) / Doutor em
Linguística do Português pela Universidade de Coimbra (UC)
Universidade Federal da Bahia

*Para todas as meninas e mulheres negras que
sonham em fazer ciência.*

AGRADECIMENTOS

Achei que não conseguiria concluir este trabalho. E não foi apenas uma vez. Foram várias vezes. Duvidei constantemente da minha capacidade. Mas, durante todo o processo, tive pessoas ao meu redor que acreditaram no meu potencial. Obviamente, não poderia deixar de agradecê-las.

Aos meus pais, Bárbara e Adailton, por todo o incentivo dado ao longo da vida. Agradeço especialmente à minha mãe, que vibrou muito mais do que eu quando passei na seleção de mestrado e por sempre se preocupar com a minha saúde física e mental.

À minha avó materna, Edelzuita, por ter me permitido, ainda na infância, ser a sua professora.

À minha orientadora, Antonia Vieira, por ter me acolhido tão bem desde o segundo semestre da graduação e por ter se tornado uma segunda mãe. Agradeço, ainda, por ter despertado em mim o amor pela Linguística Histórica.

A Álvaro, meu namorado, por todo amor, cuidado, estímulo e companheirismo. Tudo isso foi mais do que essencial para que eu não desistisse.

À Wlianna, amiga e irmã quase gêmea, com quem dividi as alegrias e as frustrações relacionadas à academia, da graduação ao mestrado.

Aos amigos Rodrigo, Ícaro, Isla, Luan, Laíz, Rebeca e Atahebson, pelas conversas de apoio.

À Carla Carolina, colega querida, pelas boas risadas e informações compartilhadas.

Aos professores João Paulo Cyrino, Mailson Lopes e Sílvia Ribeiro, pelas valiosas contribuições dadas.

Por fim, agradeço à Universidade Federal da Bahia por ter proporcionado grande parte da minha formação acadêmica.

RESUMO

Este estudo, inserido no âmbito da morfologia histórica, investiga o padrão compositivo Nome-Nome, doravante [NN]_N (*peixe-anjo*), em duas sincronias pretéritas da língua portuguesa. Embora o padrão compositivo [NN]_N seja considerado um dos mais produtivos no português contemporâneo (RIBEIRO; RIO-TORTO, 2009, 2010, 2012, 2016; GONÇALVES, 2016), pesquisas sobre a composição de palavras no português arcaico (incluindo o período galego-português) têm mostrado que esse padrão não é produtivo nessa fase (SANTOS, 2009; BORGES; SANTOS, 2017, 2020). Ante o exposto, a ampliação de estudos com foco no padrão compositivo [NN]_N torna-se necessária para se ter conhecimento sobre o seu processo de desenvolvimento no decurso da história da língua portuguesa. Em função disso, este estudo busca, então, identificar, descrever e analisar compostos [NN]_N situados em um conjunto documental dos séculos XVII e XVIII. O *corpus* de análise deste estudo é constituído por 10 textos, de gêneros textuais diversos, os quais se encontram disponíveis no *Corpus Histórico do Português Tycho Brahe*. Para fundamentar a discussão, assumimos como principais aportes teóricos a tese de Santos (2009), que aborda a composição no português arcaico, incluindo o padrão compositivo [NN]_N, e o trabalho de Ribeiro e Rio-Torto (2016), no qual são apresentadas uma perspectiva mais abrangente para o fenômeno da composição e uma proposta de análise desenvolvida e aplicada notadamente aos compostos da língua portuguesa. Metodologicamente, este estudo apoia-se precipuamente na abordagem indutiva (PRODANOV; FREITAS, 2013). No que tange aos procedimentos, realizamos o levantamento dos compostos [NN]_N no *corpus* e, em seguida, a descrição e a análise linguística. Compreendemos, a partir dos resultados obtidos, que não houve um crescimento expressivo de compostos [NN]_N nas sincronias em questão e, percebendo isso, esperamos que novos estudos sobre o tema sejam desenvolvidos.

Palavras-chave: Morfologia. Composição. Compostos NN. Séculos XVII e XVIII. História da Língua.

ABSTRACT

Within the scope of historical morphology, this study investigates the Noun-Noun pattern, henceforth [NN]_N (*peixe-anjo*), in two past synchronies of the Portuguese language. Although the [NN]_N pattern is considered one of the most productive in contemporary Portuguese (RIBEIRO; RIO-TORTO, 2009, 2010, 2012, 2016; GONÇALVES, 2016), studies on composition in Old Portuguese (including the Galician-Portuguese period) have shown that this pattern is not productive at this stage (SANTOS, 2009; BORGES; SANTOS, 2017, 2020). Given the above, studies focusing on the [NN]_N pattern are necessary to obtain knowledge about its development process throughout the history of the Portuguese language. As a result, this study seeks to identify, describe and analyze [NN]_N compounds located in a document set from the 17th and 18th centuries. The *corpus* of analysis of this study consists of 10 texts from different textual genres, which are available in the *Corpus Histórico do Português Tycho Brahe*. In this discussion, the main theoretical contributions are the thesis of Santos (2009), regarding composition in Old Portuguese, including the [NN]_N pattern, as well as the work of Ribeiro and Rio-Torto (2016), presenting a more comprehensive perspective on the phenomenon of composition and an analysis proposal developed and applied for Portuguese language compounds. Methodologically, this study bases mainly on the inductive approach (PRODANOV; FREITAS, 2013). Concerning the procedures, we surveyed the [NN]_N compounds in the *corpus*, followed by a description and linguistic analysis. We understand, from the results, that there was no expressive growth of [NN]_N compounds in the synchronies in question, and, realizing this, we hope for further studies on the subject.

Keywords: Morphology. Composition. NN compounds. 17th and 18th centuries. History of Language.

LISTA DE QUADROS

Quadro 1 – Propriedades das palavras compostas	24
Quadro 2 – Esquemas compositivos de compostos morfológicos	38
Quadro 3 – Níveis de idiomaticidade	47
Quadro 4 – Compostos [NN] _N correspondentes a denominações humanas	55

LISTA DE ABREVIATURAS E SIGLAS

a) Abreviaturas não-bibliográficas e termos convencionais

AN	Adjetivo-Nome
AdvV	Advérbio-Verbo
al.	Alemão
ing.	Inglês
GEN	Genitivo
lit.	Literal
N	Nome
NA	Nome-Adjetivo
NN	Nome-Nome
NprepN	Nome-preposição-Nome
NprepdetN	Nome-preposição-determinante-Nome
NumN	Numeral-Nome
PB	Português Brasileiro
PE	Português Europeu
port.	português
Rad	Radical
sécs.	Séculos
<i>s.u.</i>	<i>sub uerbum</i> ‘sob a entrada’
V	Verbo
VN	Verbo-Nome
VV	Verbo-Verbo

b) Siglas elaboradas para os textos do *corpus*

<i>MMI</i>	<i>Manuscrito das Mãos Inábeis</i>
<i>DVP</i>	<i>Discursos Vários Políticos</i>
<i>CPAV</i>	<i>Cartas do Padre António Vieira</i>
<i>AF</i>	<i>Arte de Furtar</i>
<i>OFA</i>	<i>O Fidalgo Aprendiz</i>
<i>CJCB</i>	<i>Cartas de J.C Brochado</i>
<i>MJVE</i>	<i>Musa Jocoza de Varios Entremezes Portuguezes e Castelhanos</i>
<i>MHC</i>	<i>Vida e Morte de Madre Helena da Cruz</i>
<i>GBPE</i>	<i>Gazetas Manuscritas da Biblioteca de Évora</i>
<i>CAC</i>	<i>Cartas de António da Costa</i>

SUMÁRIO

1	INTRODUÇÃO	12
2	A COMPOSIÇÃO COMO PROCESSO DE FORMAÇÃO DE PALAVRAS	18
2.1	A DEFINIÇÃO DE PALAVRA COMPOSTA	18
2.2	A RELAÇÃO ENTRE A COMPOSIÇÃO E OUTROS FENÔMENOS LINGÜÍSTICOS	23
2.2.1	Colocações e expressões idiomáticas	26
2.2.2	Prefixação	28
3	COMPOSTOS [NN]_N	31
3.1	UM POUCO DE HISTÓRIA: ALGUMAS CONSIDERAÇÕES SOBRE O DESENVOLVIMENTO DOS COMPOSTOS [NN] _N EM PORTUGUÊS	31
3.2	CLASSIFICAÇÃO DOS COMPOSTOS [NN] _N : A PERSPECTIVA DE ANÁLISE ADOTADA	37
3.3	COMPOSTOS [NN] _N QUANTO À NATUREZA MORFOLÓGICA	37
3.3.1	Compostos morfológicos	37
3.3.2	Compostos morfossintáticos	39
3.4	COMPOSTOS [NN] _N QUANTO À NATUREZA SINTÁTICA	40
3.4.1	Compostos coordenados	41
3.4.2	Compostos modificativos	42
3.5	COMPOSTOS [NN] _N QUANTO À (IN)EXISTÊNCIA DE NÚCLEO	43
3.6	ASPECTOS SEMÂNTICOS DOS COMPOSTOS [NN] _N : COMPOSICIONALIDADE E IDIOMATICIDADE	46
4	METODOLOGIA	50
4.1	O TIPO DE PESQUISA E O MÉTODO	50
4.2	O <i>CORPUS</i> DE ANÁLISE: CONSTITUIÇÃO E BREVE CARACTERIZAÇÃO	50
4.2.1	Motivações para as escolhas	52

4.3	PROCEDIMENTOS GERAIS DE COLETA E ANÁLISE DE DADOS	54
5	APRESENTAÇÃO E ANÁLISE DOS DADOS	55
5.1	COMPOSTOS [NN] _N CORRESPONDENTES A DENOMINAÇÕES HUMANAS	55
5.2	COMPOSTOS [NN] _N CORRESPONDENTES A ESPÉCIES ZOOLOGICAS	62
5.3	OUTRAS CONSTRUÇÕES IDENTIFICADAS NO <i>CORPUS</i>	65
5.3.1	Pexe mulher e pexe homẽ	65
5.3.2	Carro-mato	66
5.3.3	Gato çapato	67
6	CONSIDERAÇÕES FINAIS	69
	REFERÊNCIAS	71

1 INTRODUÇÃO

A composição é um dos principais temas de estudo no domínio dos processos de formação de palavras, pois, além de extremamente disseminada nas línguas do mundo, apresenta uma série de desafios para as teorias linguísticas (RADIMSKÝ, 2015). Inicialmente, um dos principais desafios é a própria definição do que seja uma palavra composta. Não há um consenso entre os pesquisadores sobre a noção de composto, gerando, então, diferentes descrições do mecanismo.

As palavras compostas exibem, nas diferentes línguas naturais, diversos padrões compositivos. Desse modo, é possível observar elementos de variadas categorias lexicais no interior das formações, tais como nomes, adjetivos, preposições, verbos, numerais, advérbios e pronomes. Na língua portuguesa, no entanto, determinados esquemas de combinação estão mais bem representados. De acordo com Ribeiro e Rio-Torto (2009, 2010, 2012, 2016), quatro padrões compositivos são mais produtivos¹ no português contemporâneo: [NprepN]_N (Nome-preposição-Nome), [NA]_N (Nome-Adjetivo), [VN]_N (Verbo-Nome) e [NN]_N (Nome-Nome).

Guevara e Scalise (2009), em um trabalho sobre a existência de universais na composição, analisam algumas características desse fenômeno em 16 línguas. Com o intuito de simplificar a base de dados, tais línguas foram distribuídas em quatro grupos relacionados geneticamente, a saber: grupo do Leste Asiático (coreano, mandarim e japonês), grupo eslavo (búlgaro, polonês, russo e servo-croata), grupo germânico (alemão, holandês, inglês, norueguês e sueco) e grupo românico (catalão, espanhol, francês e italiano)². A partir da análise dos dados, constata-se que a estrutura [NN]_N é a mais recorrente em todos os grupos mencionados. Desse modo, tal esquema de combinação é descrito pelos autores como “[...] the canonical instance in compounding in the world’s languages³” (GUEVARA; SCALISE, 2009, p. 125).

¹ Mas afinal, o que seria um padrão compositivo produtivo? Segundo Fernández-Domínguez (2015, p. 220-221), há imprecisões na literatura sobre o fenômeno da produtividade. Todavia, as considerações feitas por Kastovsky (1986) a respeito do assunto parecem ser as mais influentes até então. Para ele, a produtividade possui duas dimensões: a disponibilidade, que se refere à regra, e a rentabilidade, que se refere à utilização real de uma regra. Ao pensar na composição de palavras, um padrão compositivo pode estar ou não disponível em uma dada língua. Se ele estiver disponível, a formação de novas palavras através dele pode ser rentável em um grau mais baixo ou mais alto. Considerando as noções apresentadas, aqui se reconhece, especificamente, um padrão compositivo como produtivo caso ele seja rentável em um grau mais alto, isto é, forme uma grande quantidade de palavras em uma língua.

² Embora o trabalho de Guevara e Scalise (2009) seja sobre a existência de universais na composição, considera-se que a quantidade de línguas analisadas pelos autores não é suficiente para esse tipo de investigação.

³ Tradução livre: “[...] a instância canônica na composição das línguas do mundo”.

O esquema compositivo $[NN]_N$ refere-se a um composto formado por dois substantivos com estatuto de palavra, como, por exemplo, *peixe-espada* e *auxílio-doença*. Contudo, a proposta de análise de Ribeiro e Rio-Torto (2016) também apresenta compostos formados por radicais marcados categorialmente como nome, isto é, radicais nominais, como *democracia* e *teologia*. Enquanto os compostos formados somente por substantivos são representados pelas autoras através do esquema compositivo $[NN]_N$, os compostos formados por radicais nominais são representados por $[Rad_N Rad_N]_N$. Apesar de haver essa distinção em termos de representação, é importante salientar que tanto os compostos constituídos apenas por palavras quanto os compostos constituídos por radicais correspondem, em português, a nomes.

De modo geral, as palavras compostas formadas somente por substantivos e/ou por radicais nominais do português são classificadas, especificamente, quanto à natureza morfológica e à natureza sintática. Quanto à natureza morfológica de seus constituintes, essas palavras compostas podem pertencer ao grupo dos compostos morfológicos ou ao grupo dos compostos morfossintáticos. Quanto à natureza sintática, as relações que se desenvolvem no interior desses compostos podem ser de coordenação, em que não há dependência entre os termos, ou de modificação, em que um termo atua como núcleo e o outro atua como modificador⁴.

Embora a língua portuguesa não tenha sido objeto de estudo da investigação de Guevara e Scalise (2009), o padrão compositivo $[NN]_N$ é identificado como um dos mais produtivos no português atual, como afirmam Ribeiro e Rio-Torto em diversos trabalhos (2009, 2010, 2012, 2016). Também Gonçalves (2016, p. 52) destaca que se encontra uma grande quantidade de itens lexicais de configuração $[NN]_N$ no português. Sendo assim, pode-se dizer que esses estudos corroboram, de certa forma, a pesquisa desenvolvida por Guevara e Scalise (2009).

De acordo com Giurescu (1975), todas as línguas neolatinas possuem compostos formados por dois – ou até mesmo três substantivos –, sejam eles formações antigas ou recentes. A partir daí, a indagação que deve ser feita é se esse padrão compositivo é, de fato, produtivo em todas elas.

As investigações sobre o fenômeno da composição apontam, de um modo geral, para a existência de uma alta produtividade de palavras compostas com a estrutura $[NN]_N$ nas línguas germânicas. Consoante Bustos Gisbert (1986), as línguas do ramo germânico, principalmente o inglês e o alemão, apresentam um número muito mais expressivo de compostos $[NN]_N$ em comparação com outras línguas. Para o autor, isso se deve, muito provavelmente, a questões

⁴ Essas classificações serão explicadas no capítulo 3, mais especificamente nos tópicos 3.3 e 3.4.

históricas – o que inclui a permanência, de certa forma, do caso genitivo – e à estrutura formal, na qual ocorre a anteposição do termo regente. Portanto, entende-se que o elemento determinante se encontra sempre à esquerda em alemão e em inglês (ing. *sunglasses* / al. *Sonnenbrille* ‘óculos de sol’).

William Snyder (2001) também evidencia a elevada produção de compostos [NN]_N em inglês e em alemão. Com suporte da teoria gerativa, ele apresenta o *Parâmetro de Composição*, segundo o qual a gramática de uma língua permite ou não a formação de compostos durante a derivação sintática. Tendo como base esse parâmetro, o autor investiga 19 línguas pertencentes a diversos grupos linguísticos. Conclui, então, que línguas como o inglês e o alemão possuem marcação positiva para a formação de compostos do tipo [NN]_N. Já outras línguas, como o espanhol e o francês, apresentam marcação negativa.

Para o espanhol, Bustos Gisbert (1986) salienta que os compostos [NN]_N são pouco frequentes. Entretanto, o trabalho de María Irene Moyna (2011), também sobre a composição em espanhol, mostra um panorama um pouco diferente. Ao realizar uma análise minuciosa de palavras compostas encontradas em fontes diversas – dicionários, *corpora* eletrônicos e textos do espanhol –, Moyna (2011) verifica que diferentes tipos de compostos [NN]_N são abundantes no banco de dados. Além disso, observa que esse padrão compositivo é frequente desde os primeiros períodos documentados.

Consoante Arnaud (2015), o padrão compositivo [NN]_N não é predominante para a nomeação de conceitos na língua francesa. Porém, tal esquema se encontra em processo de expansão no francês contemporâneo. Embora o autor não diga o motivo dessa expansão, Darmesteter (1875), em sua obra a respeito da composição em língua francesa, já indicava uma possível influência do inglês com relação ao surgimento de novas palavras compostas no francês. A fim de justificar essa interferência, ele diz que “a l’imitation des composés anglais, on a créé des expressions plus simples et plus rapides exigées par les besoins du commerce et des affaires, amenées par les relations avec nos voisins d’outre-Manche ou introduites par la mode et la fashion⁵” (DARMESTETER, 1875, p. 138).

Partindo da observação geral de que as línguas germânicas optam pelo uso de compostos [NN]_N, enquanto as línguas do grupo românico utilizam mais formações do tipo [NprepN]_N ou [NprepdetN]_N⁶, Schlechtweg (2018) estuda essas construções complexas em alemão, inglês e

⁵ Tradução livre: “À imitação de compostos ingleses, criamos expressões mais simples e diretas, exigidas pelas necessidades do comércio e dos negócios, provocadas por relacionamentos com os vizinhos além do Canal [da Mancha] ou introduzidas pela moda e costumes”.

⁶ No padrão compositivo [NprepdetN]_N, ocorre a combinação de uma preposição com um determinante (*café da manhã*).

francês. Nessa perspectiva, na língua francesa as formações [NN]_N são menos frequentes do que as [NprepN]_N e [NprepdetN]_N. Segundo o autor, essa tendência é sustentada pelos dados de Pius ten Hacken (2013): de 257 palavras compostas [NN]_N em inglês, somente 5% aparecem como construções [NN]_N em francês. Por outro lado, 25% dos compostos [NN]_N em inglês são encontrados como construções [NprepN]_N na língua francesa.

No caso do romeno, a baixa produtividade de compostos [NN]_N é discutida por Adina Bleotu (2010), em contraposição com a alta produtividade desse padrão na língua inglesa. Com o intuito de explicar o porquê disso acontecer, a pesquisadora, fundamentando-se no gerativismo, expõe algumas propostas existentes para tratar do assunto. Dentre os argumentos apresentados por ela, há o *Parâmetro de caso dentro dos compostos*. De acordo com esse parâmetro, determinadas línguas permitem a construção de expressões sem qualquer elemento no interior, como preposições e conjunções, enquanto outras não. Sendo assim, línguas como o inglês possuem palavras compostas [NN]_N muito semelhantes em termos estruturais, dado que o segundo substantivo exerce a função de núcleo do produto composicional e o primeiro substantivo atua como modificador do segundo (*bus driver* ‘motorista de ônibus’, *surfboard* ‘prancha de surf’). Porém, em romeno, é possível encontrar compostos [NN]_N com estruturas morfológicas diferentes. Em um artigo de 2018, Bleotu aponta que as construções [NN]_N endocêntricas do romeno, isto é, que possuem um termo nuclear, podem ser subcategorizadas em duas classes: (1) Nome-Nome no genitivo (*floarea-soarelui* ‘flower-the sun-GEN’ lit. *flor do sol* ‘girassol’) e (2) Nome + preposição + Nome no acusativo (*lapte-de-pasăre* lit. *leite de pássaro* ‘tipo de sobremesa que consiste em um merengue flutuando um creme inglês’). Desse modo, entende-se que o inglês utiliza a estrutura [NN]_N prototípica, constituída apenas por dois substantivos, enquanto que o romeno utiliza variações, tendo em vista que necessita marcar, de alguma forma, o substantivo não nuclear no interior do produto.

Diferentemente do romeno, outras línguas neolatinas, como o italiano e o catalão, não apresentam baixa produtividade de palavras compostas [NN]_N. Bisetto (2004) afirma que as construções [NN]_N são muito difundidas na língua italiana, principalmente em um gênero textual específico, o jornalístico. Inclusive, Szemberska (2013), em um trabalho sobre os compostos [NN]_N no léxico esportivo italiano, utiliza, como *corpus* de análise, o principal jornal de esportes da Itália, *La Gazzetta dello Sport*. A partir da investigação, a autora conclui que “[...] i composti Nome-Nome costituiscono un fenomeno rilevante e ben presente nell’italiano contemporaneo⁷” (SZEMBERSKA, 2013, p. 100). Além disso, Rainer (2016) destaca que entre

⁷ Tradução livre: “[...] os compostos Nome-Nome são um fenômeno relevante e bem presente no italiano contemporâneo”.

os compostos nominais, isto é, aqueles em que o produto final resulta em um substantivo, o esquema de combinação [NN]_N com núcleo à esquerda é o mais produtivo nessa língua (*parola chiave* ‘palavra-chave’, *squalo martello* ‘tubarão-martelo’).

Quanto ao catalão, Padrosa Trias (2010), em sua tese de doutorado, revela que apesar de o padrão compositivo [NN]_N ser muito mais abundante em inglês, é também bastante produtivo na língua catalã. Enquanto determinadas estruturas, como [AdvV]_V e [NumN]_N, são consideradas improdutivas na sincronia atual do catalão (BERNAL, 2012, p. 18), o esquema [NN]_N tem se tornado mais frequente. De acordo com Sentí (2014, p. 85), os dados coletados pelo projeto NEOXOC⁸ indicam que a grande maioria dos neologismos em todas as zonas catalãs corresponde a um tipo de composto [NN]_N, em que não há relação de dependência entre os elementos que formam o produto compositivo (*actor-cantant* ‘ator-cantor’, *aler-pivot* ‘ala-pivô’). Especificamente, 62,3% dos neologismos catalães formados pelo processo de composição de palavras são compostos [NN]_N.

Com relação ao português, já foi dito, inicialmente, que a estrutura [NN]_N é um dos padrões compositivos mais produtivos na fase contemporânea dessa língua. Contudo, os resultados de alguns estudos sobre a composição de palavras no período (galego-)português da língua portuguesa (sécs. XII-XVI) evidenciam a escassez de compostos [NN]_N em textos medievais (SANTOS, 2009, 2017; BORGES, SANTOS, 2018, 2020).

Segundo a eminente Carolina Michaëlis de Vasconcelos (1946 [1912-1913] p. 45), “a história da [...] composição portuguesa está por escrever”. Apesar de essa assertiva feita em *Lições de Filologia Portuguesa* ser um tanto antiga, já que possui mais de 100 anos, continua sendo extremamente válida. Até então, os estudos a respeito da composição de palavras em outros períodos da língua portuguesa ainda são poucos. Sobre a composição na chamada fase arcaica, destacam-se duas relevantes investigações: o artigo de Blanco Valdés (1985), *Palabras compostas en galego-portugués*, por causa do pioneirismo, e a tese de Santos (2009), já citada, devido à riqueza de detalhes. Todavia, salienta-se que trabalhos relativos à composição em períodos posteriores, notadamente nos séculos XVII, XVIII e XIX, são, ao que tudo indica, inexistentes.

As pesquisas de Santos (2009, 2017) e Borges e Santos (2018, 2020), coadunadas com a ainda atual declaração de Vasconcelos (1946 [1912-1913] p. 45), impulsionam uma reflexão sobre os compostos [NN]_N, tornando impositivo averiguar como o processo de expansão de palavras compostas [NN]_N ocorreu na língua portuguesa ao longo dos séculos. À vista disso,

⁸ O NEOXOC é um projeto que observa os neologismos em variedades diatópicas do catalão. Nove universidades catalãs participam desse projeto.

este trabalho tem, como principal objetivo, estudar o padrão compositivo [NN]_N em textos atinentes a diferentes gêneros textuais dos séculos XVII e XVIII, buscando observar o comportamento e o desenvolvimento desse padrão compositivo ao longo da história da língua. Como parte desse objetivo, figuram a descrição e a análise dos dados a serem identificados no *corpus* de análise.

Quanto à estrutura desta dissertação, há, além desta introdução, cinco capítulos.

No Capítulo 2, são apresentadas e discutidas algumas definições de *composição/palavra composta* encontradas na literatura, dando-se ênfase principalmente às definições adotadas neste trabalho. Logo depois, discorre-se sobre a proximidade da composição com outros fenômenos linguísticos, como as colocações, as expressões idiomáticas e a prefixação.

No Capítulo 3, realizam-se, inicialmente, algumas considerações sobre o padrão compositivo [NN]_N, concentrando-se no seu desenvolvimento no decorrer da história da língua portuguesa. Em seguida, tendo como base a proposta de análise adotada, são descritos os tipos de compostos [NN]_N do português, bem como as principais características desse padrão compositivo.

O Capítulo 4 corresponde à metodologia. Nele, são abordadas, de início, a natureza da pesquisa e a perspectiva metodológica. Em seguida, é apresentado e descrito o *corpus* de análise. Posteriormente, comentam-se as razões que influenciaram no processo de seleção dos textos e do *corpus* eletrônico que dispõe os documentos analisados. Por último, expõem-se os procedimentos utilizados para a coleta e a análise dos dados.

No Capítulo 5, apresentam-se os dados identificados no *corpus* e realizam-se a descrição e a análise relativamente a aspectos morfológicos, sintáticos e semânticos. Ainda nesse capítulo, são comentadas, de forma complementar, algumas construções encontradas no decurso da coleta de dados, mas cuja classificação como composto é dubitável.

As considerações finais constituem o Capítulo 6.

2 A COMPOSIÇÃO COMO PROCESSO DE FORMAÇÃO DE PALAVRAS

Em diversas línguas, a composição é o processo mais utilizado na formação de novos itens lexicais, sendo considerada mais simples que a derivação, uma vez que não envolve o uso de afixos específicos, mas constituintes com livre curso na língua, em geral (BOOIJ, 2007; BAUER, 2017). Embora considerada um mecanismo bastante produtivo e relativamente simples, a composição de palavras é um tema controverso no âmbito da descrição linguística, principalmente no que diz respeito à definição de composto e a sua delimitação frente a outros fenômenos linguísticos contíguos. Ainda neste capítulo, discorreremos brevemente sobre essas questões.

2.1 A DEFINIÇÃO DE PALAVRA COMPOSTA

A definição de palavra composta é facilmente encontrada em dicionários e gramáticas, como nos exemplos a seguir:

Formação de palavra pela união de elementos léxicos independentes, da qual resulta um novo conceito único e autônomo, e que pode ocorrer por justaposição (*passatempo*, *vaivém*, *amor-perfeito*) ou por aglutinação (*agricultura*, *tragicômico*) [Da composição por palavras eruditas, participam radicais gregos e latinos (*geografia*, *telefone*, *vermífugo* etc.)]. (HOUAISS; VILLAR, 2009, p. 506)

Processo de formação de palavras em que se juntam duas ou mais palavras ou radicais (Dicionário Infopédia da Língua Portuguesa)

Dá-se a *composição* [...] quando se juntam dois ou mais elementos vocabulares de significação própria, para darem a ideia de um novo ser ou objeto. O que caracteriza, em última análise, a *composição* é, além da *unidade de significação*, a existência de *mais de um radical*. (ROCHA LIMA, 2011, p. 279)

A composição [...] consiste em formar uma nova palavra pela união de dois ou mais radicais. A palavra composta representa sempre uma ideia única e autônoma, muitas vezes dissociada das noções expressas pelos seus componentes. Assim, *criado-mudo* é o nome de um móvel; *mil-folhas*, o de um doce; *vitória-régia*, o de uma planta. (CUNHA; CINTRA, 2013, p. 119)

No entanto, definir o que é um composto não é necessariamente uma tarefa simples, principalmente quando se intenta elaborar uma definição que abranja línguas tipologicamente distintas (GUEVARA; SCALISE, 2009, RADIMSKÝ, 2015).

Guevara e Scalise, em *Searching for Universals in Compounding* (2009), ao analisarem algumas definições sobre composto na literatura científica, consideram que as definições formuladas não são totalmente satisfatórias, sendo, além disso, pouco compreensíveis. Para os

autores, “[...] the many definitions of compound that one finds in the literature are tightly predetermined by the theoretical choices made by the author(s)⁹” (GUEVARA; SCALISE, 2009, p. 106). Nessa perspectiva, os compostos podem ser descritos a partir de diferentes pontos de vista, tornando, assim, muito difícil a formulação de um conceito único para o fenômeno.

O linguista Laurie Bauer, por exemplo, apresenta, em diferentes trabalhos, as seguintes definições para composto:

(1) A compound lexeme (or simply a compound) can thus be defined as a lexeme containing two or more potential stems¹⁰. (BAUER, 1983, p. 28)

(2) A compound is usually defined as being a word (in the sense of lexeme) that is made up of two other words (in the sense of lexeme)¹¹ [...] (BAUER, 1998, p. 65)

(3) Compounds are sequences of lexemes¹². (BAUER, 2003, p. 135)

Apesar de fornecer variações da definição de composto, Bauer (2017, 2020) não deixa de problematizar a natureza e a delimitação do fenômeno. Esse último aspecto é, possivelmente, um dos principais motivos de discordância, na literatura especializada, acerca da definição do termo. O autor também ressalta que a ausência de uma definição de composto que sirva para todas as línguas prejudica a realização de interpretações interlinguísticas e estudos tipológicos, pois os compostos de uma língua como o inglês, por exemplo, não correspondem precisamente aos compostos de outras línguas germânicas ou das línguas românicas.

Em *Compounds and Compounding*, de 2017, Bauer não apenas traz à tona o debate inesgotável sobre uma possível definição para o composto, mas também aborda outro ponto questionável, associado, de alguma forma, a tal definição: a utilização do termo *palavra composta* como sinônimo de *composto*. Consoante o autor (2017, p. 3), “perhaps the basic assumption underlying compounds is that they are words they even are called ‘compound

⁹ Tradução livre: “[...] as muitas definições de composto que encontramos na literatura são rigidamente predeterminadas pelas escolhas teóricas feitas pelo(s) autor(es)”.

¹⁰ Tradução livre: “Um lexema composto (ou simplesmente um composto) pode, portanto, ser definido como um lexema contendo duas ou mais bases potenciais”.

¹¹ Tradução livre: “Um composto é geralmente definido como sendo uma palavra (no sentido de lexema) composta por duas outras palavras (no sentido de lexema)”.

¹² Tradução livre: “Compostos são sequências de lexemas”.

words' on some occasions and are often defined as being words whose elements are words¹³". Entretanto, ele salienta que existem dois problemas relacionados ao uso desse termo. O primeiro problema é que sequências de palavras (sintagmas) normalmente não são estruturas morfológicas, mas sintáticas. O segundo problema, por sua vez, está ligado ao fato de também não haver uma definição consensual para palavra.

No que tange ao segundo problema, Margarida Basilio já discorreu sobre tal questão em um artigo publicado no ano de 2000. De acordo com a autora,

[o] conceito de palavra é de grande dificuldade em morfologia, dadas as múltiplas dimensões em que esta unidade pode ser enfocada, as quais nem sempre coincidem, além de apresentarem diferentes graus de relevância. Na palavra, entendida como uma unidade lexical, uma sequência fônica se associa de modo relativamente estável a (a) um significado ou conjunto de significados; (b) um conjunto de propriedades sintáticas; (c) um conjunto de propriedades morfológicas; e (d) um conjunto de determinações de uso. (BASILIO, 2000, p. 9-10)

Orientando-se, de certo modo, pelo percurso histórico dos estudos morfológicos, a autora expõe que a definição de palavra não é colocada no cerne da discussão da Gramática Tradicional, haja vista que, para o modelo clássico de descrição gramatical, a palavra é considerada a unidade mínima de análise linguística. Entretanto, com o advento da Teoria Estruturalista e a introdução da noção de morfema como a unidade básica da morfologia, a palavra deixou de ser, ao mesmo tempo, a unidade mínima de análise linguística e uma unidade importante da estrutura da língua.

Embora o conceito de morfema tenha se tornado o foco da discussão do Estruturalismo, o principal nome da vertente norte-americana dessa teoria, Leonard Bloomfield (1935 [1933]), apresenta uma definição para palavra, fazendo-a corresponder a uma forma livre mínima, isto é, “[...] uma forma que pode ocorrer isoladamente, por si só constituindo um enunciado, e não podendo ser totalmente subdividida em formas livres” (BASILIO, 2000, p. 10). Apesar de o conceito de forma livre ser até certo ponto eficiente (BASILIO, 2004, p. 17), ele não se aplica à composição, como facilmente pode ser observado na análise do composto *sofá-cama*: “[...] podemos dizer que *sofá-cama* não se subdivide em duas formas livres, *sofá* e *cama*?” (BASILIO, 2000, 11). Pelo visto, a resposta é não.

¹³ Tradução livre: “Talvez a suposição básica subjacente aos compostos seja que eles são palavras que até são chamadas de “palavras compostas” em algumas ocasiões e são frequentemente definidos como sendo palavras cujos elementos são palavras.”

Segundo a autora, há uma divergência entre a definição de composto geralmente encontrada nas gramáticas normativas – construção lexical baseada em duas ou mais palavras – e o critério utilizado por Bloomfield para definir palavra – a não subdivisão em mais formas livres. Contudo, Basilio acrescenta que a palavra não se trata apenas de uma unidade morfológica, sendo também uma unidade do léxico, entendido como um conjunto de palavras de uma dada língua. Desse modo, ao pensar em palavra como unidade lexical, ela afirma que os compostos podem ser conceptualizados como “[...] conjuntos de palavras que funcionam lexicalmente como uma palavra só” (BASILIO, 2000, p. 11).

Evidencia-se, com essa exposição, a controvérsia em torno das noções de palavra e, também, de palavra composta, tema ainda bastante atual.

Seguindo ainda nas definições de composição e composto, observa-se, agora, a noção de composição apresentada por Geert Booij no livro *The Grammar of the Words*, de 2007: “in simple cases, compounding consists of the combination of two words, in which one word modifies the meaning of the other, the head¹⁴” (2007, p. 75). Partindo dessa definição, entende-se, de início, que o autor parece considerar como compostos apenas estruturas binárias em que um elemento funciona como núcleo e o outro elemento atua como um modificador de sentido, como se constata em *cidade-dormitório*. Todavia, ele ressalta que a composição pode se valer da recursividade, gerando, então, compostos bem longos, como a formação alemã *Donaudampf-schiff-fahrtsgesellschaft*¹⁵ (lit. *Danúbio-vapor-navio-viagem-empresa*).

Mesmo que Booij (2007) considere tanto estruturas binárias quanto estruturas com mais de dois constituintes como compostos, ele chama a atenção para as chamadas *expressões lexicais frasais* das línguas românicas. Construções como *sala de jantar* e *quarto de hóspedes* são, para ele, exemplos de expressões lexicais frasais. De acordo com o autor, algumas expressões lexicais frasais, como as citadas anteriormente, são equivocadamente denominadas como compostos. Para ele, construções Nome-preposição-Nome são, na verdade, instanciações da estrutura sintática [N PP]_{NP}¹⁶, isto é, um sintagma nominal cujo núcleo nominal é seguido por um complemento preposicional. De certo modo, o autor parece não considerar estruturas do tipo Nome-preposição-Nome como compostos. No entanto, destaca-se que há pesquisadores que incluem essa estrutura no rol da composição de palavras.

¹⁴ Tradução livre: “em casos simples, a composição consiste na combinação de duas palavras, em que uma palavra modifica o significado da outra, o núcleo”.

¹⁵ Exemplo retirado de Booij (2007).

¹⁶ As siglas PP e NP, provenientes do inglês, correspondem a Prepositional Phrase (sintagma preposicional) e Noun Phrase (sintagma nominal), respectivamente.

Na sua tese, Santos (2009) apresenta as perspectivas de alguns gramáticos e linguistas, de diferentes épocas, sobre a composição de palavras na língua portuguesa. No levantamento realizado pela autora, evidenciam-se diferentes concepções do fenômeno da composição e dos padrões compositivos em autores como João de Barros, Jerônimo Soares Barbosa, Said Ali, Mattos e Silva etc. Essas concepções podem divergir ou convergir em determinados momentos. Por exemplo, enquanto para João de Barros, em sua *Grammatica da lingua portuguesa*, de 1540, os compostos da língua portuguesa correspondem a estruturas binárias, para Soares Barbosa (1881 [1822]) um composto pode conter não apenas duas, mas também três palavras, sendo a terceira palavra, muito provavelmente, um conector, como a preposição *de* (SANTOS, 2009, p. 27). Por outro lado, Said Ali (1964) inclui os esquemas de combinação NA e NprepN, também referidos por Mattos e Silva (2008a), que, como informa em seu texto, dialoga com os trabalhos de Mattoso Câmara Jr. (1976) e Rio-Torto (1998). Trata-se de padrões compositivos geralmente excluídos da composição devido à sua proximidade com outros fenômenos linguísticos. Diante de tantas perspectivas, Santos (2009) optou, na sua tese, por uma abordagem abrangente da composição, fornecida pelo trabalho inovador de Ribeiro (2006)¹⁷. No entanto, mais modernamente, Ribeiro e Rio-Torto (2016, p. 461)¹⁸ apresentam uma definição de composição que engloba alguns pontos considerados importantes nesta dissertação:

A composição é um processo de formação de palavras [...] que envolve uma relação de concatenação [...] entre pelo menos duas unidades lexicais - radicais, temas ou palavras -, cada uma das quais marcada categorialmente como Nome, Adjetivo, Verbo, Advérbio, Preposição, Numeral ou Conjunção. (RIBEIRO; RIO-TORTO, 2016, p. 461)

Em primeiro lugar, destaca-se que as autoras não levam em conta somente palavras como elementos integrantes de um composto, mas também outras unidades lexicais, como radicais e temas. Em segundo lugar, há o fato de elas considerarem a existência de uma variedade categorial, da qual resulta a formação de diferentes padrões compositivos, tais como [NN]_N, [VN]_N, [AN]_N, [NA]_N, [NprepN]_N etc. Como já foi mencionado, alguns desses esquemas, em especial [NA]_N e [NprepN]_N, costumam ser excluídos do âmbito da composição por apresentarem similaridade com sintagmas livres, notadamente quanto ao comportamento

¹⁷ A dissertação de mestrado de Ribeiro (2006), intitulada "Compostos nominais em português: as estruturas NN, VN, NprepN e NA", foi publicada em 2010 pela Lincom Europe. Contudo, é necessário salientar que não tivemos acesso a essas obras.

¹⁸ Trata-se da segunda edição da *Gramática Derivacional do Português*. A primeira edição foi publicada em 2013.

flexional. Sobre essa questão, as autoras declaram: “[a] pesar de continuarmos a assumir a dificuldade na delimitação das fronteiras dos compostos, as escassas razões elencadas nestes trabalhos para fundamentar esta decisão [de excluir determinadas estruturas] não se afiguram, em nosso entender, suficientes” (RIBEIRO; RIO-TORTO, 2016, p. 471).

Além de apresentarem uma definição para a composição, as autoras formalizam, com base em Guevara e Scalise (2009), a estrutura de um composto, que é, por definição, uma unidade plurilexêmica (RIBEIRO; RIO-TORTO, 2016, p. 461):

$$[[a]_X r [b]_Y]_Z$$

No esquema, [a] e [b] referem-se às unidades lexicais, ou seja, aos radicais, temas ou palavras que podem constituir um composto; X e Y representam as classes categoriais relacionadas a cada um desses elementos; Z representa a classe do produto composicional; e r simboliza a relação gramatical¹⁹ intracomposto.

Assim, após refletirmos sobre a ausência de definições que sejam mais satisfatórias, como expresso por Guevara e Scalise (2009), optamos, nesta dissertação, por uma definição mais ampla do fenômeno da composição e do composto, como a elaborada por Ribeiro e Rio-Torto (2016).

2.2 A RELAÇÃO ENTRE A COMPOSIÇÃO E OUTROS FENÔMENOS LINGUÍSTICOS

Além de não haver consenso sobre a definição de palavra composta, a proximidade da composição com outros fenômenos linguísticos acrescenta mais dificuldade para a sua delimitação. Não obstante, no caso específico de determinadas línguas, alguns critérios são utilizados para separar compostos de determinados fenômenos, como as construções frasais.

Conforme apontam Booij (2019), Schlücker (2019) e Iordăchioaia e Werner (2020), nas línguas germânicas, como o alemão, o holandês e o inglês, há alguns critérios para distinguir compostos e construções frasais, como a ortografia e a acentuação. Nessas línguas, os compostos geralmente são grafados sem que haja a separação através de espaços em branco e a tonicidade tende a ocorrer no primeiro constituinte do produto composicional, isto é, o elemento modificador. Essas configurações podem ser observadas em estruturas como

¹⁹ Embora Ribeiro e Rio-Torto (2016) explorem, no esquema, a relação gramatical e apresentem um modelo de análise para as relações gramaticais instituídas entre os elementos dos compostos da língua portuguesa, elas também se dedicam à observação das relações semânticas estabelecidas entre os elementos compositivos.

Frischlucht ‘ar fresco’, do alemão, *opoe fiets* ‘bicicleta retrô’, do holandês, e *greenhouse* ‘estufa’, do inglês. Em contrapartida, as construções frasais são separadas por espaços em branco e a tonicidade normalmente ocorre no segundo constituinte, o elemento nuclear. É o caso, por exemplo, dos sintagmas *frische Luft* (lit. *ar fresco*), *opoe’s fiets* (lit. *bicicleta da avó*) e *green house* (lit. *casa verde*)²⁰.

Em algumas dessas línguas germânicas, como o alemão e o holandês, outro recurso também pode ser utilizado para diferenciar compostos e construções frasais: a marcação flexional. Em alemão, por exemplo, há o composto *Weisswein* e o sintagma *weisser Wein*²¹. Enquanto o adjetivo presente no composto não está declinado, o adjetivo do sintagma está flexionado²².

Embora haja tais critérios de distinção nas línguas germânicas, eles não se adequam a algumas línguas românicas, causando, então, mais dificuldades para a delimitação de fronteiras (FINKBEINER; SCHLÜCKER, 2019). Ao avaliar, por exemplo, a aplicação de um critério como a ortografia para os compostos do português, nota-se que ele é, de certo modo, inviável, pois nessa língua ocorre oscilação quanto à representação gráfica. Inclusive, é preciso ressaltar que as regras estabelecidas pelo Acordo Ortográfico de 1990 provocam ainda mais confusões em relação à escrita dos compostos na língua portuguesa, já que determinadas estruturas devem ser hifenizadas, enquanto outras, que também se comportam como compostos, não.

Apesar dessas questões discutidas, Ribeiro e Rio-Torto (2016) sinalizam algumas propriedades que caracterizam as palavras compostas do português:

Quadro 1 – Propriedades das palavras compostas

- *são constituídas por um conjunto fixo de palavras e/ou de radicais;*
- *assentam numa forte coesão formal interna (ordem imutável, opacidade interna acentuada, total ou intensa, com grande dificuldade de inserção de novas unidades no seu interior, escassa possibilidade de extensão ou de redução do conjunto);*

²⁰ Os exemplos em alemão, holandês e inglês foram retirados, respectivamente, de Schlücker (2019), Booij (2019) e Werner e Iordăchioaia (2020).

²¹ Exemplos retirados de Werner e Iordăchioaia (2020).

²² Ainda que existam esses critérios de distinção, é necessário destacar que há um contraste que separa o inglês do alemão e do holandês. Segundo Finkbeiner e Schlücker (2019), os critérios de acentuação e de marcação flexional permitem distinguir notoriamente os compostos e as construções frasais do alemão e do holandês. Todavia, isso não ocorre no inglês. Em primeiro lugar, apesar de o critério de acentuação ser aplicável em vários compostos desta, alguns trabalhos (PLAG, 2006; KUNTER, 2011) apontam a existência de inúmeras exceções. Em segundo lugar, o critério de marcação flexional não pode ser aplicado no inglês.

• *exibem forte unicidade semântica, sendo tipicamente portadoras de um sentido unitário/holístico, umas vezes composicional, outras lexicalizado/cristalizado em graus variáveis.*

Fonte: Ribeiro e Rio-Torto (2016, p. 462)

Dentre as propriedades arroladas no quadro acima, a opacidade interna é considerada a principal propriedade definitória dos compostos, a qual permite distinguir estes dos sintagmas. Segundo essa propriedade, “[...] os compostos são unidades impermeáveis a qualquer alteração/inserção no seu interior, sendo marcados pela impossibilidade de alterar a ordem dos elementos compositivos ou de os substituir por outros” (RIBEIRO; RIO-TORTO, 2016, p. 463). É possível compreender melhor tal propriedade através do exemplo a seguir, retirado de Borges e Santos (2020), referente à aplicação de um teste de substituição:

1a. Como João se acidentou, terá direito ao *auxílio-doença*.

1b. Como João se acidentou, terá direito à **assistência doença*.

Na sentença 1a, nota-se a presença do composto *auxílio-doença*, que se trata de um benefício destinado ao segurado do Instituto Nacional do Seguro Social (INSS) em consequência de uma doença ou acidente. Em 1b, verifica-se a tentativa de modificação de um dos elementos que forma o referido composto por um sinônimo, o termo *assistência*. Ainda que *auxílio* e *assistência* sejam palavras que possuem sentido semelhante, a alteração realizada em *assistência doença* gera a perda do conjunto fixo e, conseqüentemente, ela deixa de funcionar como uma palavra composta.

Apesar de as propriedades alistadas serem utilizadas com vistas a caracterizar os compostos, ainda assim é difícil estabelecer fronteiras, dado que algumas dessas propriedades também podem caracterizar as estruturas fixas²³, como as colocações e as expressões idiomáticas.

²³ De acordo com Fulgêncio (2008, p. 100), “uma expressão fixa (abreviadamente EF) é qualquer sequência de palavras que é memorizada pelos falantes da língua como um todo unitário, sendo igualmente recuperada da memória em bloco, sem o intermédio obrigatório da aplicação de regras de valor geral”.

2.2.1 Colocações e expressões idiomáticas

Os compostos, notadamente os padrões [NA]_N e [NprepN]_N, apresentam uma interseção com as colocações e as expressões idiomáticas em razão de a estrutura formal destes coincidir superficialmente com a estrutura formal daqueles (SANTOS, 2009; RIBEIRO; RIO-TORTO, 2016). Logo, realizar uma distinção entre compostos e tais expressões fixas pode ser uma questão de difícil solução.

No que concerne às colocações, Tutin e Grossmann (2002) explicam que esse termo foi cunhado por John R. Firth, na década de 1950, com um sentido mais amplo: *palavras que tendem a “andar juntas”*. Todavia, nos últimos tempos, a definição a seguir, cuja origem está ligada aos estudos teóricos em lexicologia e lexicografia, passou a ser preferida na comunidade de lexicólogos e lexicógrafos: “la collocation est une cooccurrence lexicale privilégiée de deux éléments linguistiques entretenant une relation syntaxique²⁴” (TUTIN; GROSSMANN, 2002, p. 9). Assim sendo, nas colocações há uma associação que vai além do significado, e faz com que a utilização de uma palavra necessite de outra e só possa ocorrer junto a essa outra (FULGÊNCIO, 2008).

Embora a transparência semântica, isto é, o sentido não figurado, seja considerada como uma das principais características das colocações, nem sempre uma colocação será uma expressão com significado tão previsível. Aliás, no plano semântico, Tutin e Grossmann (2002, p. 12-13) estabelecem uma divisão das colocações em três subconjuntos: as **colocações opacas**, em que o significado do colocativo, elemento selecionado pela base, difere daquele que apresenta fora da associação (*sorriso amarelo*); as **colocações transparentes**, as quais constituem o caso prototípico de colocação, pois são compreensíveis, mas difíceis de serem previstas (*tomar banho*); e as **colocações regulares**, as quais se definem pela proximidade com a combinatória lexical livre devido à motivação e à transparência semântica (*grande tristeza*).

Quanto às expressões idiomáticas, Fulgêncio (2008, p. 101), em sua tese dedicada ao estudo dos idiomatismos, relata que elas são definidas, tradicionalmente, “[...] como um conjunto de palavras montado idiossincraticamente”. Dentre as características desse tipo de expressão fixa, destaca-se, em especial, o caráter semântico não composicional, uma vez que não se pode alcançar o significado do conjunto a partir da somatória dos significados de cada elemento que constitui uma expressão idiomática. A fim de ilustrar isso, há a expressão

²⁴ Tradução livre: “colocação é uma co-ocorrência lexical privilegiada de dois elementos linguísticos que entre si mantêm uma relação sintática”.

idiomática *quebrar a cabeça*, cujo sentido difere consideravelmente da soma das partes: ‘grande dedicação para solucionar um problema; pensar muito sobre um tema’.

Segundo Fulgêncio (2008, p. 177-178), a principal diferença entre as colocações e as expressões idiomáticas diz respeito aos traços semânticos. Nas expressões idiomáticas, nenhum dos constituintes pode ser deduzido a partir de seu significado literal (*mexer os pauzinhos* ‘recorrer a influências em favor de algo ou alguém’). Nas colocações, porém, um dos termos geralmente é utilizado no seu sentido habitual (*erro crasso*, *torrar dinheiro*). Apesar de haver essa distinção entre colocações e expressões idiomáticas, Santos (2009, p. 50) suscita uma indagação relevante: “[...] como diferenciar compostos de expressões idiomáticas e compostos de colocações?”.

Santos (2009, p. 50-51) expõe que as colocações apresentam menos restrições sintáticas do que as expressões idiomáticas. Sendo assim, as colocações tendem a permitir a inserção de material lexical entre os constituintes: *escovar os dentes – escovar bem os dentes*. Contudo, a autora salienta que, em alguns casos, os testes de inserção de material lexical podem não ser satisfatórios, já que dependem da intuição gramatical do falante.

Outro fator que, de algum modo, possibilita situar compostos e colocações em patamares distintos é a natureza prototípica. Ao considerar a organização dos compostos em um *continuum*, os compostos mais opacos sintaticamente e semanticamente são os mais prototípicos. No entanto, ocorre o contrário com as colocações: quanto menor for a opacidade sintática e semântica, mais prototípicas são (SANTOS, 2009; RIBEIRO; RIO-TORTO, 2010).

Com relação aos compostos e às expressões idiomáticas, ressalta-se que a distinção entre esses fenômenos é ainda mais complexa. Segundo Corpas Pastor (1996, p. 92-93), no espanhol, o critério ortográfico é utilizado tradicionalmente pela filologia espanhola com a finalidade de diferenciá-los. Dessa forma, consideram-se como compostos, nessa língua, as construções formadas pela união gráfica de duas ou mais bases. Por outro lado, as expressões idiomáticas não exibem união gráfica. Entretanto, como evidenciado em 2.2, o critério ortográfico não atende bem ao português.

Em sua obra, Corpas Pastor (1996, p. 95) apresenta algumas construções que são classificadas como expressões idiomáticas no espanhol, como *lágrimas de cocodrilo* (port. *lágrimas de crocodilo* ‘lágrimas falsas, fingidas’). Todavia, essa mesma construção é incluída no rol dos compostos da língua portuguesa por alguns autores, como Santos (2009) e Ribeiro e Rio-Torto (2016). A partir do que foi descrito, surge, então, um questionamento: Por qual razão tal construção é tida como uma expressão idiomática no espanhol? É somente por não exibir união gráfica entre os constituintes? Ou será que a construção é descartada do âmbito da

composição por seguir a estrutura [NprepN]_N? Apesar de ser realmente difícil distinguir compostos e expressões idiomáticas, considera-se aqui a posição de Santos (2009) e Ribeiro e Rio-Torto (2016): *lágrimas de crocodilo* trata-se de uma palavra composta.

2.2.2 Prefixação

Além de ser difícil estabelecer limites bem definidos entre os compostos, as colocações e as expressões idiomáticas, também há problemas para delimitar fronteiras entre a composição e a derivação, mais especificamente entre elementos formativos de compostos morfológicos e de produtos da recomposição (*melodrama, fumódromo, autoescola*²⁵) e elementos prefixais – afixos que se adjungem à esquerda da base (*reconduzir, infeliz*).

Consoante Rio-Torto (2016, p. 412), o critério da autonomia acentual costuma ser utilizado com certa frequência para diferenciar elementos compositivos e prefixais. Geralmente, os constituintes dos compostos têm autonomia acentual, ainda que sejam radicais não autônomos, como *bio*, em *bioesfera*. No entanto, esse critério não é preciso, pois há compostos morfológicos que apresentam apenas um acento principal, como *hemograma, fonologia, biblioteca* etc.

Devido às fronteiras maleáveis entre composição e prefixação, Lopes (2013, p. 126) sinaliza que há bastante tempo se discute, em círculos linguísticos, se a prefixação é um processo derivativo ou compositivo. Com o objetivo de evidenciar a variação relacionada ao status morfológico da prefixação, Lopes (2013, p. 128) apresenta um quadro que mostra a visão de alguns morfologistas a respeito do caráter desse mecanismo de formação de palavras na língua portuguesa. No quadro, a maioria dos morfologistas entende a prefixação como um processo derivacional. Contudo, alguns autores compreendem a prefixação como um processo composicional, como o influente Mattoso Câmara Jr.

Não obstante a discussão sobre o status derivacional ou composicional da prefixação, há um terceiro posicionamento teórico quanto ao caráter desse processo. Para alguns autores, a prefixação situa-se em um *continuum* mórfico, uma vez que compartilha propriedades da derivação e da composição (GONÇALVES, 2012; GONÇALVES; ANDRADE, 2012; LOPES, 2013; RIO-TORTO, 2016). Acredita-se aqui que esse posicionamento seja o mais coerente, tendo em vista que a prefixação não é um mecanismo uniforme e cada elemento prefixal

²⁵ A recomposição consiste na formação de um composto a partir de uma unidade truncada e de uma palavra plena. *Autoescola*, por exemplo, é um produto da recomposição, pois *auto* é resultante da redução de uma unidade já existente (*auto < automóvel*).

apresenta um comportamento específico. Além do mais, ao considerar a ideia de *continuum*, é possível observar quais são os exemplares mais prototípicos dos prefixos e dos compostos morfológicos.

Em relação ao segundo ponto, Rio-Torto (2016, p. 421-422) apresenta algumas características dos prefixos mais prototípicos e dos elementos formativos de compostos morfológicos, o que permite realizar certo nível de distinção. De acordo com a autora, os prefixos mais prototípicos caracterizam-se pelos seguintes aspectos:

- I. ocorrem apenas em posição prefixal (*desagradável*);
- II. não possuem autonomia sintática (*des-*, por exemplo, não atua livremente);
- III. possibilidade de combinação com um maior número de classes lexicais de base (*desamor*, *descolar*);
- IV. não apresentam especificação categorial sintática (*des-*, por exemplo, não é marcado como Nome, Adjetivo ou Verbo).

Por sua vez, os elementos formadores de compostos morfológicos costumam exibir as seguintes propriedades:

- I. ocorrem ora à esquerda ora à direita (*cronologia*, *diacronia*);
- II. apresentam marcação categorial (*-cron-*, por exemplo, é um Radical Nominal);
- III. em termos semânticos, referem-se a entidades ontológicas e referencialmente individualizáveis;
- IV. são marcados por restrições de seleção, já que não se combinam arbitrariamente com bases nominais, adjetivais e verbais.

Como se pôde observar neste capítulo, o estabelecimento de fronteiras bem definidas entre a composição e outros fenômenos linguísticos não está isento de dificuldades, principalmente quando se está a lidar com os compostos morfológicos e os compostos sintagmáticos, notadamente os de estrutura [NA]_N e [NprepN]_N.

Não obstante as dificuldades apontadas não afetarem, aparentemente, os compostos [NN]_N, objeto de estudo deste trabalho, esse padrão apresenta algumas complexidades. A mesma estrutura de superfície acomoda uma série de relações semânticas, nem sempre facilmente decodificadas. Isso se deve à ausência, em alguns casos, de um conectivo que

relacione os dois nomes (*sociólogo-presidente*²⁶, *seguro-saúde*²⁷, *bolsa atleta*²⁸), ou, ainda, a processos cognitivos de metáfora e metonímia, bastante comuns nesse tipo de composição (*cobra-papagaio*²⁹, *peixe-boi*³⁰, *homem-sanduíche*³¹).

A dificuldade de análise recai, principalmente, sobre o segundo termo. Em português, o segundo elemento de compostos [NN]_N é, em geral, responsável por atribuir traços ou propriedades ao primeiro constituinte. Em *sequestro-relâmpago*, por exemplo, o termo *relâmpago* qualifica metaforicamente o sequestro como um crime que ocorre em um curto espaço de tempo. Contudo, como atribuir propriedades é uma função básica dos adjetivos, é natural que surjam dificuldades na classificação de determinadas estruturas como [NN]_N ou [NA]_N. É o caso, por exemplo, de *padre procurador*, dado do nosso *corpus*, que foi analisado como um composto [NN]_N, mas que poderia, em princípio, ser analisado como um composto [NA]_N. Na nossa análise, trata-se de um composto [NN]_N, pois *procurador*, nesse contexto, não apresenta o valor semântico de 'que procura', mas indica, na verdade, uma atividade profissional exercida por um sacerdote, relacionada, mais especificamente, à administração de mosteiros.

Em decorrência do estatuto semântico fluido do segundo elemento, os compostos [NN]_N passam a manifestar variação na marcação flexional. Enquanto os compostos [NA]_N, por exemplo, exibem um único padrão flexional, em que ambos os elementos compositivos são pluralizados (*batata doce - batatas doces*), os compostos modificativos [NN]_N, isto é, aqueles em que um dos elementos atua como modificador, registram variação na marcação de plural. Exemplificando, compostos como *laranja-pera*³² e *garoto-propaganda*³³ podem manter a marcação flexional somente no constituinte nuclear (*laranjas-pera, garotos-propaganda*) ou nos dois constituintes (*laranjas-peras, garotos-propagandas*).

²⁶ Essa construção se refere a Fernando Henrique Cardoso, sociólogo e ex-presidente do Brasil (1995-2002).

²⁷ *Seguro-saúde* 'contrato com empresa, hospital etc. que implica pagamento de mensalidades e propicia atendimento na área da saúde, aí englobados consultas médicas, exames, internações, cirurgias etc.; plano de saúde' (HOUAISS; VILLAR, 2009, *s.u.* seguro-saúde)

²⁸ *Bolsa atleta* 'programa do Governo Federal, gerido pelo Ministério do Esporte, que visa garantir a manutenção pessoal aos atletas de alto rendimento que não possuem patrocínio.' (FARIA, 2011)

²⁹ *Cobra-papagaio* 'serpente amazônica, [...] que pode atingir mais de 1 m de comprimento, de dorso verde com barras transversais branco-amareladas e região ventral amarela [...]' (HOUAISS; VILLAR, 2009, *s.u.* cobra-papagaio)

³⁰ *Peixe-boi* 'denominação comum aos mamíferos aquáticos e sirênios da família dos triquecídeos, manso, de corpo roliço, de até 3 m de comprimento [...]' (CALDAS AULETE, digital, *s.u.* peixe-boi)

³¹ *Homem-sanduíche* 'pedestre que faz publicidade ambulante pelas ruas, carregando dois cartazes justapostos, presos ao ombro, um a frente e outro às costas' (HOUAISS; VILLAR, 2009, *s.u.* homem-sanduíche)

³² *Laranja-pera* 'fruto de polpa ácida e casca lisa' (CALDAS AULETE, digital, *s.u.* laranja-pera)

³³ *Garoto-propaganda* 'aquele que, nos meios visuais de comunicação, apresenta determinado produto, ressaltando suas qualidades' (HOUAISS; VILLAR, 2009, *s.u.* garoto-propaganda)

3 COMPOSTOS [NN]_N

3.1 UM POUCO DE HISTÓRIA: ALGUMAS CONSIDERAÇÕES SOBRE O DESENVOLVIMENTO DOS COMPOSTOS [NN]_N EM PORTUGUÊS

Em comparação com outras línguas indo-europeias, como o grego, o sânscrito e o armênio, o fenômeno da composição em latim – língua que deu origem às chamadas línguas românicas – é descrito como pouco produtivo. Ademais, o uso desse processo de formação de palavras se restringia a áreas bem definidas (FRUYT, 2002)³⁴. Todavia, Santos (2009, p. 67) afirma, com base em alguns estudos sobre a composição em latim, que as palavras compostas de configuração [NN]_N da língua portuguesa têm as suas raízes na língua latina, mais especificamente nos justapostos constituídos por dois substantivos em aposição.

Segundo Bader (1962, p. 296), apesar de os justapostos [NN]_N em aposição não serem numerosos no latim, quase todos surgiram através do antigo vocativo: *DYEU PATER > IŪPITER. Com relação à ordem dos constituintes, Santos (2009, p. 67) relata que esse tipo de justaposto podia apresentar ordem regressiva, em que o elemento determinante antecede o determinado, ou ordem progressiva, a mais comum, em que o elemento determinado antecede o determinante.

Na sua pesquisa, Santos (2009, p. 68) também observou a presença, embora rara, de compostos latinos [NN]_N coordenados. De fato, trabalhos sobre a composição em latim, como os de Oniga (1992) e Fruyt (2002), mencionam a escassez de construções coordenadas em latim. Segundo Oniga (1992, p. 108), os compostos latinos desenvolviam muito mais relações de dependência do que de coordenação, enquanto Fruyt (2002, p. 263), por sua vez, explica que o latim não utilizava um composto para expressar algo como 'x e y', mas sim uma frase.

De modo geral, os compostos [NN]_N latinos correspondem a “[...] expressões normais da língua, utilizadas para designar um objeto ou uma ideia, mas que já atingiram um elevado grau de ‘petrificação’, possivelmente pela frequência de uso [...]” (SANTOS, 2009, p. 69). Ainda que não produzisse tantas formas, a estrutura [NN]_N já estava presente no latim e servia, de alguma maneira, como mecanismo para nomeação das coisas.

Assim como o latim, a língua portuguesa também se valia, na fase arcaica, do uso de construções [NN]_N. Contudo, os poucos trabalhos sobre a composição de palavras em fases

³⁴ Apesar de Fruyt (2002) não citar as áreas em que a composição era produtiva no latim, ela menciona, em um tópico específico, que alguns compostos latinos foram criados no âmbito da poesia. Além disso, Detreville (2015) afirma que as palavras compostas eram abundantes na poesia latina.

pretéritas do português (SANTOS, 2009, 2017; BORGES; SANTOS, 2018, 2020) atestam que a presença desse padrão compositivo nos documentos do português arcaico é muito limitada.

A tese de doutorado de Santos (2009), referente aos compostos [VN]_N, [NN]_N, [NA]_N, [AN]_N e [NprepN]_N no português arcaico, mostra que poucas palavras pertencentes ao padrão [NN]_N foram reconhecidas no *corpus* examinado. No total, apenas 10 compostos [NN]_N foram identificados pela pesquisadora. Por outro lado, outros esquemas de combinação, como o [NA]_N e o [NprepN]_N são muito mais recorrentes nos textos analisados. No que tange à quantidade de compostos [NN]_N identificados no *corpus*, diz Santos (2009, p. 172):

[...] o número de registros, principalmente de compostos VN e NN mostrou-se bastante escasso em comparação com as inúmeras formas presentes nos dicionários modernos e com aquelas formas que surgem na contemporaneidade no âmbito de variados setores da sociedade e nos diversos tipos de produção técnica e literária.

Em investigações mais recentes, Santos (2017) e Borges e Santos (2018, 2020) também indicam uma baixa produção de compostos [NN]_N no português arcaico. Santos (2017), ao analisar a obra medievo-portuguesa *Orto do Esposo*, identificou apenas 6 compostos [NN]_N. Por sua vez, Borges e Santos, em um trabalho de 2018, identificaram somente 1 composto [NN]_N ao analisarem 34 documentos notariais da Chancelaria de D. Afonso III, pertencentes ao século XIII. Já em um trabalho de 2020, as mesmas autoras identificaram somente 4 compostos [NN]_N nas duas partes da *Crónica de D. João I* (séc. XV). Grande parte das palavras compostas encontradas por Santos (2017) e Borges e Santos (2018) corresponde a outras estruturas, principalmente à estrutura [NA]_N, tal como constatou Santos (2009).

Entretanto, quando se observam dados pretéritos do espanhol, língua também românica, nota-se uma realidade que se opõe à da língua portuguesa. Em *Compound words in Spanish*, Moyna (2011) realiza um trabalho diacrônico, em que identifica palavras compostas em fontes lexicográficas, *corpora* eletrônicos e textos da língua espanhola de diferentes períodos, e elabora um banco de dados. Após uma análise rigorosa, conclui que há dois tipos produtivos de compostos [NN]_N no espanhol desde a época medieval, os compostos [NN]_N com núcleo à esquerda e os compostos [NN]_N com núcleo à direita.

Moyna (2011, p. 164) descreve os compostos [NN]_N com núcleo à esquerda do espanhol como frequentes em registros coloquiais, portanto, típicos da linguagem cotidiana. Tais construções costumam designar nomes de animais, plantas, objetos e ferramentas artificiais. Consoante a autora (2011, p. 168), esse tipo de formação é considerado abundante, pois no banco de dados há um total de 404 compostos [NN]_N com núcleo à esquerda. Como sua

frequência é alta desde os primeiros períodos documentados, é possível encontrar esse tipo de estrutura até mesmo em textos do século XII.

Com relação às palavras compostas [NN]_N com núcleo à direita, Moyna (2011, p. 173) ressalta que esses compostos correspondem a dois níveis históricos diferentes. A primeira categoria é muito antiga e parece remontar ao latim e ao romance inicial. Nesse grupo, incluem-se nomes concretos próprios da linguagem cotidiana e do trabalho manual (*pezuña* ‘casco’). A segunda categoria, mais nova, deve a sua existência a duas tendências predominantes no léxico espanhol durante o século XX: a importação de termos técnicos e não técnicos do inglês (*tour operador* ‘operador turístico’), e a incorporação, na linguagem cotidiana, da terminologia técnico-científica que preserva a ordem dos constituintes (*avifauna* ‘conjuntos de aves de uma região’).

Os compostos [NN]_N com núcleo à direita foram marcados por momentos de perdas e ganhos na língua espanhola. Nos séculos XI e XII, há uma quantidade não negligenciável herdada de períodos mais antigos. Porém, a partir do século XVI, ocorre uma pequena queda na frequência relativa desse tipo de estrutura. Já no século XX, há um aumento abrupto no número de compostos. Apesar de seu crescimento não ser contínuo, as formações [NN]_N com núcleo à direita também são consideradas numerosas, dado que há 327 compostos desse tipo presentes no banco de dados de Moyna (2011).

Também se faz necessário abordar neste trabalho uma outra língua, que, embora não faça parte do grupo românico, sempre é referida quando se discute o fenômeno da composição: a língua inglesa. Parece ser unanimidade entre os teóricos que a composição é um mecanismo muito produtivo em inglês, sendo o padrão compositivo [NN]_N o mais comum nessa língua³⁵. No entanto, a utilização desse esquema de combinação não se restringe apenas à fase contemporânea, mas ocorre, na verdade, desde o inglês antigo.

Em um trabalho sobre a composição em fases pretéritas do inglês, Ruiz Moneva (1998) destaca que uma das principais estruturas utilizadas para formar compostos no inglês antigo foi a [NN]_N. A autora acrescenta que, muito possivelmente, duas questões foram importantes para que houvesse a criação de compostos nominais no período inicial dessa língua: as tendências sintéticas que predominavam na língua da época e a escassez de preposições. Como o inglês

³⁵ Dentre os compostos [NN]_N do inglês, destacam-se os chamados *synthetic compounds* (compostos sintéticos). Nesse tipo de composto, o segundo elemento corresponde a um substantivo ou a um adjetivo verbal e o primeiro elemento é interpretado como um argumento do segundo elemento. Alguns exemplos de compostos sintéticos [NN]_N são as formações inglesas *bus driver* (motorista de ônibus), *mountain climber* (alpinista) e *drug dealer* (traficante de drogas). Nesses casos, os elementos à esquerda têm a função de objeto direto: alguém dirige **ônibus**; alguém escala **montanha**; alguém trafica **drogas**.

antigo era uma língua sintética, as relações entre as palavras não eram feitas através de preposições, mas por meio de flexões. Desse modo, enquanto as flexões eram abundantes, as preposições eram mais escassas. Mesmo hoje não sendo considerada uma língua sintética, a produção de palavras compostas ainda é muito elevada no inglês, principalmente as de configuração [NN]_N.

Mais atual, o estudo de natureza diacrônica empreendido por Gavranović (2015) corrobora, de certo modo, o trabalho de Ruiz Moneva (1998). Ao analisar documentos referentes aos períodos do Inglês Antigo e Médio, datados de 750 a 1150 e de 1150 a 1500, respectivamente, Gavranović (2015) constata que a grande maioria dos compostos nominais pertence ao padrão compositivo [NN]_N. Das 2300 formações encontradas em textos do Inglês Antigo, aproximadamente 2000 correspondem a compostos [NN]_N. Os textos do Inglês Médio apresentaram um número total de formações menor, inferior a 600. Ainda assim, cerca de 400 compostos corresponde à estrutura [NN]_N, o que revela a predominância desse esquema de combinação ao longo da história da língua inglesa.

Apesar de, em períodos mais recuados, o português não se assemelhar ao espanhol e ao inglês quanto ao alto grau de produtividade de compostos [NN]_N, na atualidade, a situação é bem diferente, segundo alguns estudos sobre a composição de palavras (RIBEIRO; RIO-TORTO, 2009, 2010, 2012, 2016; GONÇALVES, 2016). Portanto, com base nesses estudos, é possível afirmar que o padrão compositivo [NN]_N é frequente no português contemporâneo.

As investigações realizadas por Ribeiro e Rio-Torto (2009, 2010, 2012, 2016) destacam o esquema de combinação [NN]_N como um dos mais representativos na língua portuguesa. Ainda que essas investigações se baseiem em dados do português europeu (PE), certamente a declaração das autoras portuguesas também se aplica ao português brasileiro (PB), pois, segundo Gonçalves (2016), a produção de palavras compostas [NN]_N é uma das tendências recentes do PB.

Consoante Gonçalves (2016, p. 52), um tipo específico de composto [NN]_N tem se tornado muito abundante no PB nas duas últimas décadas, o composto [NN]_N com núcleo à esquerda. Formações constituídas por elementos nucleares à esquerda tornaram-se responsáveis, de algum modo, pelo processo de expansão lexical. Parte dessas formações surgiram a fim de designar fenômenos específicos surgidos em um determinado momento. Exemplo disso é *mulher-fruta*, composto relacionado a uma série de dançarinas de funk carioca que ganharam destaque nos anos 2000 através dos meios de comunicação de massa. Outras construções estão mais associadas a programas governamentais e/ou instituições. É o caso dos compostos formados por *bolsa*, que já são considerados comuns no PB.

Segundo Faria (2011), devido a uma extensão metonímica, o nome *bolsa* passou a indicar no PB, quando unido a outro nome, um tipo de benefício social mensal, normalmente concedido pelo governo. Em sua tese de doutorado, Faria (2011) apresenta um quadro com 37 palavras compostas encabeçadas por *bolsa*. Dentre as formações que se encontram no quadro, crê-se que a mais conhecida e disseminada até hoje no Brasil é *Bolsa Família*, um programa de transferência direta de renda direcionado às famílias pobres e extremamente pobres, criado em 2003 pelo governo Lula, substituído, no governo atual, pelo *Auxílio Brasil*, também uma construção [NN]_N. No entanto, outras formações listadas, como *bolsa gargalhada*³⁶ e *bolsa palestra*³⁷, parecem não ter repercutido por um longo período, tendo, então, um caráter mais efêmero.

Alguns pesquisadores discutem o porquê do surgimento de uma quantidade expressiva de palavras compostas [NN]_N na língua portuguesa. Do ponto de vista de Sandmann (1987), o padrão compositivo [NprepN]_N deu lugar ao padrão compositivo [NN]_N em novas formações. Segundo o teórico, “[...] no passado se formaram seguro de vida, seguro contra incêndio e seguro contra acidentes, aos quais se podem opor os compostos novos do *corpus* seguro-saúde e seguro-desemprego, em que foi omitida a preposição (seguro de saúde, seguro contra desemprego)” (SANDMANN, 1987, p. 67). Entretanto, o autor ressalta que mesmo a preposição não estando presente na estrutura da superfície desses novos compostos, ela se encontra presente semanticamente, ou seja, na estrutura profunda.

Bem como Sandmann (1987), Santos (2009, 2016) também se refere à relação entre os padrões compositivos [NprepN]_N e [NN]_N. Baseando-se em dados do português arcaico, a autora considera que a possível interseção entre essas estruturas “[...] pode ser o resultado de um processo histórico, em que elementos linguísticos, como conjunções e preposições, sofrem apagamento” (SANTOS, 2009, p. 83). Identifica-se essa interseção citada pela autora através da análise diacrônica de algumas formações.

Os compostos *cobra-coral* e *cobra-cascavel*, que atualmente apresentam a estrutura [NN]_N, estão registrados em obras de séculos passados, como o *Vocabulario Portuguez, e Latino* de Bluteau (1712-1728) e a *Corografia Brasílica* (1817), com a estrutura [NprepN]_N: *cobra de cascavel* e *cobra de coral*. Muito provavelmente, em algum momento da história do

³⁶ *Bolsa gargalhada* ‘aparição na fala dos artistas do programa da TV Globo “Toma Lá, Dá Cá”. Pelo contexto da criação, deveria ser criada uma bolsa para os artistas do programa, já que eles nos fazem rir, que nos divertem. É uma forma de crítica aos benefícios criados pelo governo Lula.’ (FARIA, 2011)

³⁷ *Bolsa palestra* ‘Criação do jornalista Reinaldo Azevedo por conta do enriquecimento rápido dos ex-políticos. Na maioria dos casos, ex-presidentes, ex-ministros dizem que suas rendas aumentaram por conta das consultorias prestadas após saírem do governo.’ (FARIA, 2011)

português, a preposição *de* foi omitida dessas formações, tornando-as, então, formações [NN]_N³⁸. Ainda a respeito desse assunto, Ribeiro e Rio-Torto (2009, p. 278) acrescentam que

In Portuguese the [NprepN]_N remains a very productive template of compounding. Words such as **bolsa valores*, **fim semana*, **comboio passageiros* are unacceptable, because in Portuguese, contrarily to what has been occurring in Italian, the presence of the preposition is still obligatory in the large majority of cases. However, in contemporary Portuguese some [NN]_N compounds without the preposition have started to appear. They are some few terms with universal diffusion, such as: *efeito estufa* ‘greenhouse effect’, *indústria automóvel* ‘car industry’, *imposto automóvel* ‘car tax’, *seguro-saúde* ‘health insurance’, which show that Portuguese is also undergoing a change³⁹.

Por sua vez, Faria (2011) apresenta um argumento para explicar o crescimento de compostos [NN]_N, mais especificamente no PB, que, de algum modo, se associa aos anteriores. Para ele, a perda da preposição está a serviço da própria formação de compostos [NN]_N. Segundo o autor, “após a perda, destitui-se a ideia de lexia complexa, atribuindo à nova formação uma feição, verdadeiramente, de um composto. Quanto mais integrada for a construção, mais característica de palavra composta ela terá” (FARIA, 2011, p. 115)⁴⁰.

Embora ainda não se saiba o porquê do uso crescente do padrão [NN]_N na língua portuguesa, as justificativas apresentadas contribuem, de certa maneira, para elucidar essa questão. De fato, quando se observa o registro de determinados compostos em obras que refletem um estado de língua pretérito, percebe-se uma correlação entre as estruturas [NprepN]_N e [NN]_N. Porém, esse é apenas um dos fatores a serem considerados. A influência do inglês, uma das línguas mais faladas no mundo e grande produtor de compostos [NN]_N, também deve ser destacada.

³⁸ O mesmo processo histórico pode ser observado em espanhol, como exemplificado por *agua de miel* e *aguamiel* ‘água misturada com uma porção de mel’; ‘água preparada com cana-de-açúcar’ (Diccionario de la lengua española, versão eletrônica, *s.u.* aguamiel). O primeiro, de estrutura [NprepN]_N, dá lugar a *aguamiel*, agora um composto [NN]_N, evidenciando que compostos sintagmáticos podem estar na origem de compostos morfossintáticos (ou léxicos, na terminologia do espanhol) ou mesmo de formas aglutinadas (VAL ÁLVARO, 1999, p. 4765; SANTOS, 2016; GARCÍA GALLARÍN, 2018, p. 521). Modernamente, em português, observa-se a variação entre os termos *álcool em gel* e *álcool-gel*, instâncias das estruturas compositivas [NprepN]_N e [NN]_N, respectivamente, e *alquingel*, formado por aglutinação.

³⁹ Tradução livre: “Em português, o [NprepN]_N continua sendo um modelo muito produtivo de composição. Palavras como **bolsa valores*, **fim semana*, **comboio passageiros* são inaceitáveis, pois em português, ao contrário do que vem ocorrendo em italiano, a presença da preposição ainda é obrigatória na grande maioria dos casos. No entanto, no português contemporâneo, alguns compostos [NN]_N sem a preposição começaram a aparecer. São alguns termos de difusão universal, tais como: *efeito estufa*, *indústria automóvel*, *imposto automóvel*, *seguro-saúde*, que mostram que o português também está passando por uma mudança”.

⁴⁰ Embora não fique evidente se Faria (2011) reconhece ou não a estrutura [NprepN]_N como um padrão compositivo, salienta-se que nesta dissertação se considera legítima essa estrutura no âmbito da composição de palavras.

Dessa forma, o estudo do padrão compositivo [NN]_N, particularmente em português, tem grande relevância. Além disso, há outros aspectos relacionados aos compostos [NN]_N que também instigam a investigação dessa estrutura, como se pode observar a seguir.

3.2 CLASSIFICAÇÃO DOS COMPOSTOS [NN]_N: A PERSPECTIVA DE ANÁLISE ADOTADA

Nesta dissertação, utiliza-se como base para a classificação e análise dos compostos [NN]_N a proposta elaborada pelas pesquisadoras portuguesas Silvia Ribeiro e Graça Rio-Torto (2016). O principal motivo que influenciou essa escolha é o fato de a proposta ter sido desenvolvida e aplicada notadamente ao português. Entretanto, também é preciso dizer que as autoras se apoiam principalmente na proposta de classificação de Bisetto e Scalise (2005, 2009), cujo caráter é mais geral, isto é, pode ser utilizado, certamente, em um conjunto maior de línguas.

3.3 COMPOSTOS [NN]_N QUANTO À NATUREZA MORFOLÓGICA

Nas próximas seções, serão descritos os tipos de compostos [NN]_N quanto ao grau de autonomia dos elementos compositivos e/ou à consonância das formações aos padrões sintáticos próprios da língua portuguesa⁴¹.

3.3.1 Compostos morfológicos

Os compostos morfológicos são formados por, pelo menos, um radical não autônomo⁴², normalmente de origem grega ou latina. Ademais, caracterizam-se, sincronicamente, pela presença de uma vogal de ligação – -i- ou -o- – entre os elementos compositivos. Tendencialmente, nos compostos que possuem dois radicais eruditos, a vogal de ligação é -i- quando o segundo elemento é de origem latina ‘*agrícola*’, e é -o- quando o segundo elemento é de origem grega ‘*astrologia*’.

⁴¹ Tradicionalmente, as palavras compostas da língua portuguesa são classificadas quanto à natureza morfológica em dois tipos: compostos morfológicos e compostos morfossintáticos (VILLALVA, 2003). Todavia, Ribeiro e Rio-Torto (2016) consideram a existência de três grupos: os compostos morfológicos, os compostos morfossintáticos e os compostos sintagmáticos. No caso dos compostos [NN]_N, esses podem ser classificados como morfológicos ou morfossintáticos.

⁴² Os radicais dos compostos morfológicos são marcados categorialmente como Nome, Adjetivo ou Verbo. Contudo, esta dissertação atenta-se apenas aos radicais nominais.

Segundo Ribeiro e Rio-Torto (2016, p. 476), no âmbito dos compostos morfológicos, são possíveis vários esquemas compositivos, conforme se observa no Quadro a seguir:

Quadro 2 – Esquemas compositivos de compostos morfológicos

Esquema compositivo	Exemplo
Radical erudito + vogal de ligação + radical erudito	cardiopatia
Radical erudito + vogal de ligação + palavra vernácula	hidromassagem
Radical vernáculo + vogal de ligação + radical erudito	sambódromo
Radical vernáculo + vogal de ligação + palavra vernácula	franco-alemão

Fonte: adaptado de Ribeiro e Rio-Torto (2016)

Historicamente, os compostos morfológicos tornaram-se produtivos em diversas línguas europeias a partir dos séculos XVII e XVIII, em razão do movimento de retomada do latim como língua franca. Além disso, a evolução científica que ocorreu nos séculos XIX e XX também proporcionou o desenvolvimento da terminologia que se valia da combinação de elementos clássicos para formar palavras. Por esse motivo, uma série de compostos morfológicos está presente no vocabulário universal técnico-científico (LÜDELING, 2009; GONÇALVES, 2011a).

De acordo com Gonçalves (2011a, p. 9), os compostos morfológicos não receberam tanta atenção na literatura, sendo poucos os trabalhos dedicados exclusivamente ao assunto. Todavia, o autor ressalta que o estudo desse tipo de composto vem ganhando interesse nos últimos tempos, provavelmente devido às fronteiras entre as operações morfológicas de derivação e composição.

Sobre essa questão, Bauer (2005) afirma que o limite entre os processos de derivação e de composição é permeável em duas direções, pois há elementos que antes eram vistos como formas compostas e hoje são considerados formas afixadas, e há elementos que antes eram vistos como formas afixadas e hoje assumem traços de formas compostas. No que se refere especificamente aos compostos morfológicos, o teórico relata que um dos problemas relacionados a esse tipo de composto é o fato de não ficar claro se determinadas formações constituem-se como palavras compostas ou não, uma vez que há características que podem aproximá-las de outros mecanismos, como a derivação prefixal. Ainda que alguns estudiosos da morfologia empreguem critérios a fim de caracterizar rigorosamente os constituintes presos da composição, é muito difícil delimitar as fronteiras existentes.

Além da questão das fronteiras, um outro ponto levantado por Gonçalves (2011a) merece destaque: a incorporação da composição morfológica ao sistema de formação de palavras de diferentes línguas. Em alguns idiomas, a presença dos compostos morfológicos parece se limitar apenas ao vocabulário técnico-científico, o que seria, muito possivelmente, um resultado da influência e/ou contato com outras línguas. Todavia, outros idiomas recorrem à formação de compostos morfológicos na atualidade, mostrando, então, que esse tipo de composto pode ser produtivo.

Em alguns trabalhos de Gonçalves (2011a, 2011b, 2016), destaca-se a produção de compostos morfológicos na língua portuguesa nas últimas décadas. Segundo ele, os novos compostos morfológicos podem corresponder a formações técnicas criadas por analogia a outras já existentes (*ortorexia* ‘distúrbio psicológico que se manifesta pela preocupação obsessiva com uma alimentação saudável’) ou a formações mais inusitadas, com caráter popular (*fumódromo* ‘local, geralmente isolado e ventilado de um edifício ou de um recinto, onde se pode fumar’) (GONÇALVES, 2016, p. 60). Ademais, o autor evidencia que as novas formações do português tendem a ter o -o- como elemento de ligação. Provavelmente isso ocorre porque as construções costumam apresentar algum componente oriundo do grego. É o caso, por exemplo, de *desconfiômetro* e *olhômetro*, compostos morfológicos que possuem o radical grego *-metro*, que exprime a noção de medição ou medida⁴³.

Não obstante muitos compostos morfológicos serem constituídos por dois radicais nominais, salienta-se que este trabalho tem como foco os compostos morfossintáticos. Essa priorização ocorre pelo motivo de os compostos morfológicos presentes no *corpus* de análise serem, em geral, formações herdadas e/ou tomadas de empréstimo do latim e provenientes do grego, como *filosofia*, *geografia* etc., não sendo necessariamente percebidos como palavras compostas pelos falantes comuns.

3.3.2 Compostos morfossintáticos

Como afirmado em 2.3.1, os compostos morfológicos são aqueles formados por pelo menos um radical não autônomo. Por outro lado, os compostos morfossintáticos e os sintagmáticos consideram o grau de afastamento e proximidade em relação aos padrões sintáticos da língua portuguesa. A composição morfossintática inclui os compostos que

⁴³ Salienta-se que Gonçalves (2011b) questiona se o -o- é mesmo uma vogal de ligação, como normalmente é dito nas abordagens tradicionais, ou se trata, na verdade, de um constituinte fonológico dos elementos que se fixaram à direita (*musicólatra*, *camelódromo*).

apresentam algum desvio relativamente aos padrões sintáticos do português (*outono-inverno, mulher-objeto*)⁴⁴. Por seu turno, a composição sintagmática abarca os compostos cujas estruturas seguem os padrões sintáticos da língua (*fim de semana, sangue-frio*). Discorreremos, nesta subsecção, sobre os compostos morfossintáticos, notadamente aqueles formados por dois substantivos.

Muitas vezes, as formações morfossintáticas [NN]_N caracterizam-se pela anomalia sintática, pois não exibem um elemento de ligação, como conjunções e preposições (VILLOING, 2012). Por exemplo, em *saia-calça* é possível observar a ausência de uma conjunção (*saia e calça*), enquanto em *auxílio-doença* nota-se a ausência de outro elemento de ligação, uma preposição (*auxílio para doença*). Porém, nesse tipo de estrutura nem sempre se pressupõe a ausência de um simples conector, como em *peixe-palhaço e homem-rã*.

Além dessa particularidade dos compostos [NN]_N – a ausência de um elemento de ligação –, uma outra característica, discutida em 3.6, também torna o estudo dos compostos morfossintáticos ainda mais interessante: a produção de novos significados através da metáfora e da metonímia.

Em termos de produtividade, os compostos [NN]_N morfossintáticos do português também são considerados produtivos na sincronia atual. No que se refere especificamente ao PB, os trabalhos de Faria (2011) e Gonçalves (2016) destacam a presença de compostos desse tipo nos últimos anos. A grande maioria das formações está relacionada a fenômenos momentâneos e/ou acontecimentos (*ciclone-bomba*)⁴⁵, estereótipos (*maria-chuteira*)⁴⁶ e, principalmente, programas e benefícios governamentais e/ou institucionais (*seguro-desemprego*).

3.4 COMPOSTOS [NN]_N QUANTO À NATUREZA SINTÁTICA

Nas próximas seções, serão descritos os tipos de compostos [NN]_N quanto às relações sintáticas que podem se estabelecer entre os elementos compositivos.

⁴⁴ Val Álvaro (1999) aborda uma distinção entre compostos e construções sintáticas apositivas de caráter restritivo, como *rio Ebro*. Segundo o autor, os compostos se referem a um conceito unitário e, quando seus elementos consituíntes são considerados de forma isolada, carecem de independência sintática. Isso já não ocorre com as construções sintáticas apositivas.

⁴⁵ O termo *ciclone-bomba*, muito utilizada no Brasil entre junho e julho de 2020, refere-se a um tipo de ciclone que ganha intensidade em poucas horas.

⁴⁶ A formação *maria-chuteira* diz respeito a mulheres que costumam se relacionar com jogadores de futebol.

3.4.1 Compostos coordenados

De acordo com Shimada (*apud* BAUER, 2008, p. 2), na literatura encontram-se diferentes denominações para os compostos com significado coordenativo, tais como: compostos coordenados, compostos copulativos, estruturas de conjunção, *dvandvas* etc. Dentre essas denominações, o termo sânscrito *dvandva*⁴⁷ é bastante utilizado para se referir à relação de coordenação como um todo. No entanto, a utilização do termo não parece ser a mais adequada, pois Bisetto e Scalise (2009) apontam que a relação de coordenação que ocorre na composição de palavras pode ter diferentes resultados, podendo tratar-se de: I) associação de dois elementos individuais sem que haja referência a nenhum deles como entidades separadas, como o composto sânscrito *mātā-pitarau*⁴⁸ ‘pais’; II) associação de duas propriedades ligadas a uma entidade, como *padre-frei*.

Além das divergências relacionadas à terminologia, há outra questão que envolve os compostos coordenados: alguns pesquisadores não consideram formações coordenadas como palavras compostas. Para Adams (2001), as expressões que contêm elementos coordenados são frases. Inclusive, salienta que “Noun + noun examples in which both elements denote roles, like *author-illustrator*, *producer-director* are arguably also phrases, not complex words⁴⁹” (ADAMS, 2001, p. 82). Assim como Adams (2001), Padrosa Trias (2009) também nega a existência dos chamados compostos coordenados. Em seu artigo, argumenta que esse tipo de construção é, na verdade, um caso de coordenação sintática assindética.

Embora haja discordâncias quanto à existência de compostos que instituem relação de coordenação e à terminologia empregada para designá-los, salienta-se que no presente trabalho considera-se a existência desse tipo de composto, seguindo-se, assim, os estudos de Bisetto e Scalise (2009) e Ribeiro e Rio-Torto (2016), e opta-se pelo uso do termo *composto coordenado*.

Em sua proposta de classificação, Bisetto e Scalise (2009, p. 46) fornecem uma definição mais geral para os compostos coordenados: “[...] formations whose constituents are connected by the conjunction *e*⁵⁰”. Contudo, Radimský (2015) chama atenção para o fato de

⁴⁷ Oriundo da tradição gramatical indiana, o termo *dvandva* pode ser entendido como uma estrutura formada a partir de duas entidades nomeadas (SHIMADA *apud* BAUER, 2008, p. 2). É o caso, por exemplo, do composto sânscrito *vātavarṣaḥ* ‘vento e chuva’, em que os elementos constituintes da formação se encontram relacionados semanticamente, todavia designam aspectos distintos (SANTOS, 2009; ARCODIA, 2018).

⁴⁸ Em tradução literal: *pai e mãe*.

⁴⁹ Tradução livre: “Exemplos Nome + Nome, nos quais os dois elementos denotam papéis, como autor-ilustrador, produtor-diretor também são, sem dúvida, frases, não palavras complexas”.

⁵⁰ Tradução livre: “[...] formações cujos constituintes estão conectados pela conjunção *e*”.

que a definição elaborada pelos autores deve ser entendida no sentido de que a conjunção *e* corresponde a uma possível paráfrase do composto e não à sua forma real.

Por sua vez, Ribeiro e Rio-Torto (2016, p. 490) apresentam uma definição que parece refletir melhor a natureza morfológica e semântica desses compostos:

Os compostos coordenados caracterizam-se pela presença obrigatória de dois elementos com a mesma categoria gramatical, entre os quais se estabelece uma relação de adição. Entre tais elementos compositivos existe uma evidente proximidade semântico-referencial, pertencendo ambos, com frequência, a campos semânticos em relação de ‘pareceza de família’.

Segundo Rio-Torto (2013) e Ribeiro e Rio-Torto (2016), os compostos coordenados pertencem, majoritariamente, ao padrão compositivo [NN]_N. No que se refere à língua portuguesa, as formações que estabelecem relação de coordenação costumam designar agentes (*autor-intérprete*), locais (*padaria-pastelaria*), eventos (*jantar-comício*) e objetos (*saia-casaco*)⁵¹.

3.4.2 Compostos modificativos

Os compostos modificativos são formados por um elemento modificado, o qual atua como núcleo do produto composicional, e por um elemento modificador, cuja função é atribuir algum traço ou alguma propriedade ao núcleo. No que se refere aos compostos modificativos [NN]_N do português, o elemento modificador se localiza quase sempre à direita (*palavra-chave*, *padre-cura*)⁵².

Com relação à classificação dos compostos modificativos, destaca-se que as propostas de Bisetto e Scalise (2009) e Ribeiro e Rio-Torto (2016) diferem entre si em alguns pontos. Embora a definição geral apresentada por esses autores seja muito semelhante, alguns aspectos considerados e os termos utilizados para as subcategorizações são diferentes.

Segundo a proposta de Bisetto e Scalise (2009), a relação de modificação/atribuição é apenas o macronível de classificação dos compostos. É preciso ainda considerar um segundo nível de classificação, o qual compreende as diferentes relações semânticas que se estabelecem entre os constituintes dos compostos. Sendo assim, o grupo dos compostos modificativos, chamado pelos autores de classe ATAP, abarca as formações atributivas (AT) e as formações

⁵¹ Dados do PE extraídos de Ribeiro e Rio-Torto (2016).

⁵² No *Vocabulario Portuguez, e Latino*, de Bluteau (1712-1728), são encontrados alguns compostos modificativos [NN]_N, como *gallacrista* ‘erva com inflorescências que se assemelham à crista de um galo’ (TRIEBIG; SANTOS, 2020, p. 1064).

apositivas (AP). As formações atributivas são aquelas em que o elemento modificador corresponde a um adjetivo, cujo papel é o de expressar a qualidade do elemento nuclear/modificado (*arroz doce*). Já as formações apositivas possuem como elemento modificador um substantivo, o qual expressa uma propriedade do elemento nuclear (*homem-bomba*).

Voltada para o português, a proposta de Ribeiro e Rio-Torto (2016) também considera que há diferentes relações modificativas, entretanto, a divisão não se baseia necessariamente em classes de palavras. De acordo com as autoras, há duas relações de modificação: a modificação restritiva e a modificação qualificativa. A modificação restritiva ocorre quando o elemento modificador especifica ou restringe o elemento modificado, funcionando, então, como um subclassificador do núcleo (*auxílio-creche*). Já na modificação qualificativa, o elemento modificador caracteriza o núcleo, proporcionando algum tipo de identificação (*tatu-bola*). Ribeiro e Rio-Torto (2016, p. 496) destacam que “alguns compostos [NN]_N de tipo modificativo qualificativo assentam numa leitura figurada, sendo construídos com base em processos de alteração do sentido, como a metáfora ou a metonímia”. Por exemplo, em *tatu-bola* o núcleo *tatu* é descrito metaforicamente pelo modificador *bola*, uma vez que há relação de similaridade entre o referente e o objeto.

Apesar de Ribeiro e Rio-Torto (2016) organizarem a relação de modificação em dois grupos – a modificação de restrição e a modificação de qualificação –, o que parece, na verdade, é que toda modificação qualificativa também assenta em uma modificação restritiva. Portanto, aqui se assume que a relação de modificação consiste, em geral, na atribuição, por parte do elemento modificativo, de alguma propriedade ao elemento nuclear, havendo, então, a delimitação desse núcleo.

3.5 COMPOSTOS [NN]_N QUANTO À (IN)EXISTÊNCIA DE NÚCLEO

Como aponta Jackendoff (2016), a designação do núcleo é um dos fatores que determinam a estrutura conceptual de um composto [NN]_N⁵³. O outro fator, que será abordado adiante, é o estabelecimento da relação semântica entre N1 e N2⁵⁴. O núcleo de uma palavra composta pode ser entendido como o elemento responsável por determinar a(s) propriedade(s)

⁵³ Segundo Buenafuentes de la Mata (2016, p. 10), nesse tipo de composto, constituído por unidades de igual categoria sintática, é difícil determinar que nome é responsável pela categoria do produto compositivo. Acrescenta que a aplicação dos critérios semântico e morfológico permite identificar a posição do núcleo, como se pode observar nos exemplos do espanhol (o) *coche cama* (*vagão-leito*) e (o) *perro policía* (*cão polícia*).

⁵⁴ N1 corresponde ao primeiro elemento e N2, ao segundo elemento.

do produto composicional. Contudo, é possível encontrar, na literatura, diferentes abordagens sobre o assunto, principalmente no que se refere às propriedades categoriais, morfológicas e semânticas.

Apontado como um dos primeiros teóricos a empregar a noção de núcleo, Bloomfield (1935 [1933]) baseia a sua concepção apenas na dimensão semântica. Outros autores, como Nakimi e Kageyama (2016), focalizam apenas na dimensão categorial. Entretanto, o trabalho de Scalise, Fábregas e Forza (2009) considera que para adquirir uma melhor compreensão da noção de núcleo, é preciso levar em conta o papel desempenhado por três tipos de características que parecem ser relevantes em várias línguas: a dimensão categorial, a dimensão morfológica e a dimensão semântica.

Tendo apoio no trabalho de Scalise, Fábregas e Forza (2009), Ribeiro e Rio-Torto (2016) apresentam as definições de núcleo seguindo cada uma das três dimensões. Na dimensão categorial, entende-se que o núcleo da palavra composta é o termo que fornece a categoria gramatical do produto. Na dimensão morfológica, o núcleo é o termo que fornece os traços morfológicos do produto, mais especificamente os de gênero e número. Em *homem-rã*, é o elemento à esquerda que transmite as informações de gênero e número de todo o composto (*o homem-rã ~ os homens-rã*⁵⁵). Por fim, na dimensão semântica, o núcleo é o termo que atua como hiperônimo do produto, ou seja, como um termo mais abrangente. Desse modo, compostos como *tubarão-tigre*, *tubarão-martelo* e *tubarão-baleia* compartilham o mesmo núcleo semântico, pois os três compostos se referem a espécies de tubarão.

Ao considerar as noções de núcleo apresentadas, é possível distinguir os compostos em endocêntricos e exocêntricos. Segundo Benczes (2013), essa distinção é originária de Bloomfield (1933), sendo aplicada do ponto de vista sintático e semântico. Utilizada principalmente na classificação das palavras compostas, tal distinção permaneceu na literatura morfológica ao longo dos anos e foi adotada, de diferentes formas, por uma série de estudiosos.

De modo geral, os compostos endocêntricos compreendem construções marcadas pela presença de um núcleo na sua estrutura interna, enquanto os compostos exocêntricos

⁵⁵ Em um trabalho sobre os padrões flexionais dos compostos [NN]_N, Rio-Torto (2013) observa que enquanto a marcação flexional dos compostos coordenados ocorre nos dois elementos (*saias-calças*), a marcação flexional dos compostos modificativos tende a variar, podendo ocorrer no núcleo ou nos dois elementos (*homens-rã ~ homens-rãs*). Segundo a autora, os dados mostram que o padrão mais antigo (flexão somente no núcleo) ainda é o mais recorrente tanto em Portugal como no Brasil. Por outro lado, os dados também mostram que essa variação no plural é maior agora do que no passado. Essa variação também se observa em espanhol: alguns compostos [NN]_N, como *casa cuartel* ‘posto permanente ou destacamento de corpo armado provido de alojamento para os indivíduos deste corpo e suas famílias’ (Diccionario de la lengua española, versão eletrônica, *s.u.* casa cuartel) e *coche cama* (vagão-leito) podem ter o plural incidindo apenas sobre o núcleo (*casas cuartel, coches cama*) ou sobre os dois nomes (*casas cuarteles, coches camas*) (BUENAFUENTES DE LA MATA, 2014, p. 4).

caracterizam-se pela ausência de algum elemento que atue como núcleo do produto composicional, falhando, neste caso, o teste da hiponímia (BAUER, 2017, p. 37). De acordo com esse critério, *cavalo-marinho*, por exemplo, não é um tipo de cavalo que vive no mar, não sendo o nome *cavalo*, portanto, o núcleo semântico do composto. Assim como o núcleo, a identificação dos compostos em endocêntricos e exocêntricos também depende das dimensões categorial, morfológica e semântica. A análise dos compostos através desses critérios possibilita notar a existência de diferentes níveis de endocentrismo e exocentrismo, que vai do [+ endocêntrico] ao [+ exocêntrico].

De acordo com Ribeiro e Rio-Torto (2016, p. 502-503), uma palavra composta é considerada totalmente endocêntrica quando um dos seus termos é, ao mesmo tempo, núcleo categorial, morfológico e semântico. No que se refere aos compostos [NN]_N da língua portuguesa, o elemento nuclear tende a se localizar à esquerda do produto (*peixe-espada*). Também há compostos [NN]_N que possuem dois núcleos (*trabalhador-estudante*)⁵⁶.

Além dos compostos totalmente endocêntricos, há compostos que não exibem núcleo semântico, mas exibem núcleo categorial e morfológico. Tomando-se como exemplo a formação *banho-maria*, verifica-se que ela é semanticamente exocêntrica, uma vez que não exhibe núcleo semântico. *Banho-maria* não se refere a um tipo de banho, mas sim a um ‘modo de aquecer, derreter, cozinhar ou evaporar qualquer substância em que o recipiente que a contém é colocado dentro de outro recipiente com água fervente ou quente’ (HOUAISS; VILLAR, 2009, *s.u.* banho-maria). Apesar disso, o termo *banho* atua como núcleo categorial e morfológico.

Por fim, há os compostos totalmente exocêntricos, em que nenhum dos elementos assume a função de núcleo categorial, morfológico ou semântico. Normalmente, as construções [NN]_N do português apresentam, pelo menos, núcleo categorial e morfológico. Contudo, outros padrões compositivos, como o [VN]_N (*vira-casacas*) e o [VV]_N (*corre-corre*), configuram-se como totalmente exocêntricos, pois as formações não apresentam nenhum núcleo na sua estrutura interna.

⁵⁶ O composto *trabalhador-estudante* parece não ser muito utilizado no Brasil, todavia, é comum em Portugal.

3.6 ASPECTOS SEMÂNTICOS DOS COMPOSTOS [NN]_N: COMPOSICIONALIDADE E IDIOMATICIDADE

As unidades lexicais das línguas naturais são caracterizadas pelos fenômenos de composicionalidade e idiomatidade semânticas. A propriedade de composicionalidade diz respeito à definição do significado de uma unidade lexical a partir do significado dos seus constituintes. A propriedade de idiomatidade, por outro lado, define que o significado de uma unidade lexical não resulta da soma do significado das partes que a compõem. Em suma, enquanto a composicionalidade está relacionada à transparência semântica, a idiomatidade está relacionada à opacidade semântica.

Com relação à idiomatidade, Leme (2008, p. 49) afirma que os estudos sobre esse fenômeno ganharam relevância em diversas áreas, como a semântica, a sintaxe, a psicolinguística, a psicologia e, até mesmo, a ciência da computação. Para a autora, esse interesse se justifica pelo fato de que as línguas em geral possuem milhares de expressões que apresentam significado figurado. E tal significado figurado é considerado fundamental para que um falante seja capaz de captar as nuances da significação de uma língua. Também Barreto, Marcilese e Oliveira (2018, p. 120) destacam a importância do significado não literal, já que ele não é mais associado somente ao discurso poético, mas é reconhecido como instrumento de uso cotidiano.

Apesar de a idiomatidade não ser uma propriedade exclusiva da composição de palavras, muitos compostos apresentam casos de opacidade semântica, inclusive os de configuração [NN]_N. Geralmente, a significação figurada de muitas palavras compostas tem a sua origem na metáfora e na metonímia, dois processos cognitivos que contribuem bastante para a formação de novos compostos (GASSENFERTH, 1990; SILVA, 2011; RIBEIRO; RIO-TORTO, 2016).

Gassenferth (1990), em sua dissertação de mestrado sobre os compostos metafóricos e metonímicos do português, descreve a metáfora como um recurso que se fundamenta em uma relação de similaridade e que possui caráter referencial, pois se conecta com o mundo a que se refere de um modo mais direto. Segundo a autora, “a metáfora se faz presente quando a lacuna que deveria ser preenchida por uma palavra no seu sentido literal não satisfaz as expectativas do locutor e, então, há a opção pela linguagem figurada” (GASSENFERTH, 1990, p. 25).

Por sua vez, a metonímia é um recurso que assenta em uma relação de contiguidade, implicando na transferência do nome de uma coisa para outra. Em termos de produção de palavras, Gassenferth (1990) relata que a metonímia parece ser mais obscura e desafiadora para

o falante. Enquanto a metáfora tende a facilitar o entendimento devido ao uso do critério de semelhança, a metonímia, “[...] por trabalhar com relação de contiguidade, exige um processo mental do falante mais elaborado, tornando mais dificultosa a formação e, em certos contextos, a percepção de compostos metonímicos” (GASSENFERTH, 1990, p. 23).

Embora não seja tão fácil para o falante reconhecer um composto metonímico, Basilio (2016) aborda, no artigo *Metáfora e metonímia em nomes compostos em português*, a existência de uma interação entre os processos de metáfora e metonímia nas palavras compostas. Do seu ponto de vista, os compostos classificados como metafóricos também envolvem, de alguma forma, a metonímia, pois quando uma palavra é utilizada para estabelecer algum tipo de conexão com outra, nem todas as suas propriedades são evocadas, mas somente aquelas que importam ao contexto. Desse modo, ao analisar uma estrutura como *trem-bala*, percebe-se que somente as propriedades relacionadas à rapidez/velocidade do termo *bala* foram selecionadas para a formação do composto.

No que diz respeito especificamente aos compostos [NN]_N da língua portuguesa, Gassenferth (1990, p. 51) ressalta que esse tipo de composto exibe uma tendência: somente o segundo elemento do produto composicional contém a carga metafórica ou metonímica (*cidade-fantasma*). Apesar de o trabalho de Gassenferth não ser tão recente e incluir estruturas que se encontram dicionarizadas, como *peixe-agulha*, algumas formações criadas ou popularizadas em algum momento do século XXI permitem observar que a tendência mencionada pela autora ainda parece se manter: *mulher-fruta*, *ciclone-bomba*, *bueiro-bomba*⁵⁷ etc.

Conforme apontam Ribeiro e Rio-Torto (2016, p. 514), os processos de metáfora e metonímia atuam de diferentes formas na produção de significação figurada. Desse modo, as palavras compostas podem apresentar diferentes níveis de idiomaticidade semântica. Segundo as autoras, ao considerar o *continuum* [+composicional | +idiomático], é possível organizar as palavras compostas em três níveis distintos: nível mínimo de idiomaticidade, nível intermediário de idiomaticidade e nível máximo de idiomaticidade. No Quadro a seguir, são apresentados exemplos de compostos [NN]_N de acordo com os seus níveis de idiomaticidade:

Quadro 3 – Níveis de idiomaticidade

Níveis de idiomaticidade	Compostos [NN] _N
mínimo	sofá-cama

⁵⁷ A formação *bueiro-bomba* surgiu devido a uma série de explosões de bueiros da cidade do Rio de Janeiro.

intermediário	peixe-serra
máximo	unha-gata

Fonte: a autora – com base em Ribeiro e Rio-Torto (2016)

Os compostos inseridos no nível mínimo de idiomaticidade são aqueles cuja significação consiste no resultado da soma dos significados dos elementos. Logo, os compostos pertencentes a esse nível possuem um caráter [+ composicional]. É o caso de *sofá-cama*, cuja acepção encontrada em dicionários reflete a articulação dos constituintes: ‘sofá que tem embaixo uma cama dobrável, que se abre quando se quer deitar’.

Já os compostos de nível intermediário de idiomaticidade apresentam uma significação que não resulta necessariamente na soma dos significados dos elementos composicionais, contudo, alguns traços semânticos dos elementos são conservados. De acordo com Ribeiro e Rio-Torto (2016, p. 515), “[...] ainda que muitas vezes assente em processos de metaforização e/ou de metonimização, o significado destes produtos compositivos é, em princípio, descodificável por um falante de português língua-materna”. Como exemplo, há o composto *peixe-serra*: ‘tipo de peixe caracterizado por uma maxila alongada e dentes também alongados inseridos nos bordos exteriores’. Com base nessa acepção, percebe-se que o *peixe-serra* possui esse nome por ter uma parte do corpo que se assemelha a um objeto, a serra.

Finalmente, os compostos de nível máximo de idiomaticidade são aqueles cuja significação não se estabelece, de modo algum, através da soma dos significados dos elementos. Sendo assim, não é possível deduzir a acepção do produto composicional através dos constituintes. Isso é observado, por exemplo, no composto *unha-gata*, pois há uma grande diferença entre o significado das partes que compõem a formação e o significado do produto final. *Unha-gata* não se refere a um tipo de unha ou a um tipo de gato, mas sim a uma ‘planta subarborescente, espinhosa e pubescente, da família das Leguminosas, frequente em Portugal’⁵⁸ (Infopédia, digital, *s.u.* unhagata).

A partir da análise desses níveis de idiomaticidade, deduz-se que o conhecimento do falante é crucial para a interpretação dos compostos. Caso o grau de proximidade com os significados literal e figurado das palavras seja maior, é possível que isso favoreça, de alguma forma, o processo de decodificação. Caso o grau de proximidade seja menor – o que costuma ocorrer principalmente com falantes não nativos –, o processo de decodificação será, muito

⁵⁸ Triebig e Santos (2020, p. 1064) relatam que a planta recebeu esse nome devido à sua semelhança com a unha de um gato.

provavelmente, mais difícil e exigirá mais atenção (RIBEIRO; RIO-TORTO, 2016; BARRETO; MARCILESE; OLIVEIRA, 2018).

4 METODOLOGIA

4.1 O TIPO DE PESQUISA E O MÉTODO

A presente dissertação de mestrado é resultado de uma pesquisa de natureza predominantemente descritiva. Conforme apontam Prodanov e Freitas (2013, p. 52), a pesquisa descritiva tem, como principais objetivos, observar, registrar, ordenar e analisar dados. Além disso, busca verificar a frequência com que um determinado fenômeno ocorre, a sua natureza, as suas principais características e, até mesmo, as relações que estabelece com outros fenômenos. Sendo assim, neste trabalho propõe-se descrever as palavras compostas [NN]_N em duas sincronias pretéritas da língua portuguesa (séculos XVII e XVIII) através do levantamento de dados em um *corpus* específico, a ser apresentado em 4.2.

No que se refere à perspectiva metodológica de abordagem, a presente dissertação ampara-se principalmente no método indutivo, o qual se baseia na observação de fenômenos para inferir conclusões. De acordo com Prodanov e Freitas (2013, p. 29), o método indutivo tem grande importância, pois surgiu e serviu para que os estudiosos se afastassem da especulação e se aproximassem da observação meticulosa dos fenômenos. Ademais, a partir da influência desse método, foram definidas técnicas relativas à coleta de dados. Tendo em vista isso, este estudo se baseia na descrição e análise de dados, isto é, de palavras compostas [NN]_N, coletadas em um conjunto documental dos séculos XVII e XVIII, para que seja possível detectar fatos.

Além das reflexões de Prodanov e Freitas (2013), aqui também se considera o argumento exposto por Lopes (2013) sobre a utilização do método indutivo em trabalhos voltados para o estudo de fenômenos linguísticos em fases pretéritas. Segundo o pesquisador:

Fatos de língua de estágios pretéritos, cuidadosamente coletados e descritos, conduzem a análises acuradas, que fazem emergir teorizações e generalizações, que, se não reproduzem fielmente o sistema linguístico do período estudado, aproximam-se, de forma significativa, da realidade a que se pretende descrever. (LOPES. 2013. p. 191).

4.2 O *CORPUS* DE ANÁLISE: CONSTITUIÇÃO E BREVE CARACTERIZAÇÃO

Para o estudo do padrão [NN]_N, e das formas por ele instanciadas, utilizou-se um *corpus* de análise constituído por 10 textos, de diferentes gêneros textuais, situados entre os séculos

XVII e XVIII. Todos os textos se encontram disponíveis no *Corpus Histórico do Português Tycho Brahe*, um *corpus* eletrônico anotado de textos históricos escritos em língua portuguesa.

Abaixo, são descritos brevemente os textos que constituem o *corpus* deste estudo:

- 1) *Manuscritos das mãos inábeis* – coletânea composta por 27 cartas que se encontram nos arquivos da Inquisição Portuguesa. As cartas, que possuem caráter de denúncia, foram escritas no século XVII por indivíduos semi-alfabetizados. A edição utilizada é a de Rita Marquilhas (1996), pesquisadora responsável pela reunião das cartas em sua tese de doutorado.
- 2) *Discursos vários políticos* – obra redigida, em 1624, por Manuel Severim de Faria, considerado o primeiro jornalista português⁵⁹. A obra contém textos sobre as vidas de Luís de Camões, João de Barros e Diogo do Couto, e discursos sobre a língua portuguesa e o exercício da caça. A edição utilizada é a de Maria Leonor Soares Albergaria Vieira, de 1999.
- 3) *Cartas do Padre António Vieira* – coletânea composta por várias cartas escritas pelo Padre António Vieira, datadas entre 1648 e 1670. As cartas possuem diferentes destinatários, como o Marquês de Niza, o Procurador da Província do Brasil, o Rei D. João IV, o Rei D. Afonso VI e Duarte Ribeiro de Macedo. A edição utilizada é a de João Lúcio de Azevedo, publicada pela Imprensa da Universidade de Coimbra, em 1925.
- 4) *Arte de Furtar* – redigida em 1652, a *Arte de furtar* é hoje considerada uma obra clássica da literatura barroca portuguesa. Embora a autoria tenha sido atribuída, inicialmente, ao Padre António Vieira, atualmente se acredita que a obra foi escrita pelo Padre Manuel da Costa. A edição utilizada é a de Roger Bismut, publicada pela Imprensa Nacional Casa da Moeda, em 1991.
- 5) *O fidalgo aprendiz* – peça de teatro cuja autoria é do lisboeta Francisco Manuel de Melo. A primeira publicação de *O fidalgo aprendiz* ocorreu em 1665, nas *Obras métricas de Dom Francisco Manuel*. O texto que se encontra no *Corpus Histórico do Português Tycho Brahe* é uma reprodução fac-similada da edição de 1665.

⁵⁹ Informação obtida no artigo de Sousa *et al.* (2006).

- 6) *Cartas de J.C. Brochado* – coletânea composta por 92 cartas escritas por José da Cunha Brochado, datadas entre 1696 e 1703. Brochado foi considerado um dos indivíduos mais notáveis do cenário político de Portugal durante os séculos XVII e XVIII⁶⁰. A edição utilizada é a de António Álvaro Dória, de 1944.
- 7) *Musa Jocoza de varios entremezes portuguezes, & castelhanos* – peças teatrais curtas (entremezes) cuja autoria é de Nuno Nisceno Sutil. A coletânea foi publicada originalmente em 1709, na Oficina de Miguel Manescal. O texto que se encontra no *Corpus Histórico do Português Tycho Brahe* é uma reprodução fac-similada da edição de 1709.
- 8) *Rellação da Vida, e Morte da Serva de Deos a veneravel Madre Elenna da Crus Religiosa no Convento da Esperança desta cidade de Lisboa* – biografia sobre a Madre Helena da Cruz, escrita em 1721 pela poetisa e religiosa franciscana Sórora Maria do Céu⁶¹. A edição utilizada é a de Filomena Belo, de 1993.
- 9) *Gazetas manuscritas da Biblioteca Pública de Évora* – coletânea composta por diários redigidos por D. Francisco Xavier de Meneses e publicados entre 1729 e 1731. Esses diários apresentam notícias diversificadas, referentes a casamentos, despachos, roubos, notas de falecimento etc. A edição utilizada é a de João Luís Lisboa, Tiago dos Reis Miranda e Fernanda Olival, de 2002.
- 10) *Cartas do Abade António da Costa* – coletânea composta por 13 cartas escritas pelo Abade António da Costa, datadas entre 1750 e 1780. Ainda hoje, as poucas informações que se tem sobre a vida desse religioso devem-se à publicação de suas cartas. A edição utilizada é a de Fernando Lopes Graça, de 1946.

4.2.1 Motivações para as escolhas

O processo de constituição do objeto observacional deste estudo, isto é, o *corpus* de análise, fundamentou-se, de certa maneira, nas colocações feitas por Maia (2002, 2012), Mattos e Silva (2008b, 2008c) e Sacramento e Nascimento (2019). De modo geral, esses autores

⁶⁰ Informação obtida no artigo de Gomes (2016).

⁶¹ Informação obtida no site da Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB).

abordam as relações existentes entre a Linguística Histórica e a Filologia, ressaltando a importância desta última para os estudos linguísticos de fases pretéritas de qualquer língua.

Mattos e Silva (2008b, p. 10) afirma, com veemência, que a Linguística Histórica depende estritamente da Filologia, a qual “[...] abrange práticas de crítica, interpretação e edição diversas” (SACRAMENTO; NASCIMENTO, 2019, p. 474). Tal dependência é esmiuçada por Clarinda Maia (2012, p. 537), que declara: “[...] para a investigação em perspectiva diacrônica, é absolutamente necessária uma consistente infraestrutura filológica e uma sólida preparação por parte do investigador que lhe permita uma interpretação dos textos escritos que constituem o seu *corpus* de análise [...]”. Sendo assim, a Filologia “[...] possui um conjunto de recursos interpretativos e explicativos que a Linguística Histórica não pode e não deve desprezar (MAIA, 2012, p. 539).

Haja vista que as fontes textuais atuam como “[...] testemunhos de língua a partir dos quais a própria língua seria investigada” (SACRAMENTO; NASCIMENTO, 2019, p. 480), o processo de constituição de um *corpus* é considerado imprescindível para os trabalhos desenvolvidos no âmbito da Linguística Histórica (MAIA, 2002, 2012; MATTOS E SILVA, 2008c). Com relação a essa constituição, Bandeira (2011) julga a definição do *corpus* de uma pesquisa como uma das maiores dificuldades que o pesquisador precisa enfrentar, pois o *corpus* a ser elegido deve se adequar ao tema que se deseja estudar. Além disso, Maia (2012, p. 538-539) acrescenta que para se definir um *corpus*, alguns requisitos são cruciais para garantir a qualidade dos resultados a serem obtidos, como a autenticidade e a riqueza textual.

No que diz respeito à autenticidade – fator relacionado à fiabilidade –, é possível afirmar que este estudo buscou elaborar um conjunto documental de análise que obedecesse a um certo rigor filológico. Para isso, elegeram-se textos em um *corpus* eletrônico fidedigno, o *Corpus Histórico do Português Tycho Brahe*. Sobre esse *corpus* eletrônico, Paixão de Sousa (2014, p. 56-57) apresenta algumas informações que lhe conferem mais confiabilidade:

O *Corpus Anotado do Português Histórico Tycho Brahe*, construído a partir do projeto de Galves (1998), foi uma iniciativa pioneira no âmbito da língua portuguesa, e permanece hoje como o maior *corpus* eletrônico anotado de textos históricos em português. [...] Hoje, a coleção inclui textos escritos por autores portugueses, brasileiros e africanos, nascidos entre 1380 e 1845. [...] O conjunto de textos anotados que forma o *Corpus Tycho Brahe* (doravante CTB) tem como objetivo principal possibilitar de forma ampla a recuperação de informações filológicas e linguísticas dos textos.

Com relação à riqueza textual, Maia (2012, p. 538-539) destaca que o pesquisador deve se apoiar em um *corpus* textual rico caso tencione tirar conclusões sobre uma língua em uma

dada sincronia ou sobre a história de um fenômeno específico de mudança no decorrer do tempo. Ao se referir à riqueza textual, a autora reporta-se tanto à amplitude quanto à diversidade das fontes escritas. Baseando-se nessas considerações, este estudo procurou selecionar um *corpus* diversificado do ponto de vista da abrangência temporal e dos gêneros textuais. Desse modo, os registros que compõem o *corpus* foram escritos em décadas distintas dos séculos XVII e XVIII e correspondem a diferentes gêneros textuais, como se observou em 4.2.

4.3 PROCEDIMENTOS GERAIS DE COLETA E ANÁLISE DE DADOS

Com a finalidade de atingir o objetivo geral desta dissertação – investigar o padrão compositivo [NN]_N em sincronias pretéritas da língua (sécs. XVII e XVIII), em especial no que concerne à sua produtividade –, foram adotados os seguintes procedimentos:

1. Leitura pormenorizada de todos os textos que fazem parte do *corpus*;
2. Levantamento das potenciais palavras compostas pertencentes exclusivamente ao padrão compositivo [NN]_N⁶²;
3. Realização de consultas em obras lexicográficas físicas e digitais, principalmente o *Google Books*⁶³, com o intuito de conhecer e/ou confirmar as acepções das palavras compostas [NN]_N extraídas do *corpus*;
4. Descrição e análise dos aspectos morfológicos, sintáticos e semânticos envolvidos na estrutura das palavras compostas [NN]_N identificadas. No que se refere à descrição e análise, este estudo baseia-se principalmente nos trabalhos desenvolvidos por Santos (2009) e Ribeiro e Rio-Torto (2016)⁶⁴ sobre os compostos da língua portuguesa.

⁶² Para identificar as potenciais palavras compostas no *corpus*, foram observadas algumas propriedades definitórias que caracterizam uma palavra composta, apontadas por Ribeiro e Rio-Torto (2016). Tais propriedades são: (i) estabelecimento de um conjunto fixo de palavras e/ou radicais, (ii) opacidade interna e (iii) forte unicidade semântica.

⁶³ Segundo Maroneze (2019), o *Google Books* trata-se de uma excelente fonte de dados para pesquisas de cunho etimológico, pois é uma biblioteca digital com mais de 25 milhões de títulos.

⁶⁴ Enfatizamos, no entanto, o trabalho pioneiro de Ribeiro (2006) sobre os compostos nominais em português, publicado em 2010 pela LINCOM, e os vários artigos que a autora desenvolveu em coautoria com Rio-Torto (2009, 2010, 2012).

5 APRESENTAÇÃO E ANÁLISE DOS DADOS

Para a descrição e a análise dos dados identificados no *corpus*, optamos por organizar as subsecções 5.1 e 5.2 a partir das classes semânticas a que pertencem os potenciais compostos [NN]_N. Justifica-se essa preferência por uma parte do inventário de formações estar associada à classe de denominações humanas, ao passo que a outra parte está associada à classe de espécies zoológicas. Na subsecção 5.3, discute-se sobre algumas construções Nome + Nome registradas no *corpus*, mas cujo estatuto de composto não é categórico.

5.1 COMPOSTOS [NN]_N CORRESPONDENTES A DENOMINAÇÕES HUMANAS

No *corpus*, 14 potenciais compostos [NN]_N correspondentes a denominações humanas foram identificados, os quais são apresentados no Quadro 4 a seguir:

Quadro 4 – Compostos [NN]_N correspondentes a denominações humanas

COMPOSTOS	XVII					XVII- XVIII	XVIII			
	M M I	D V P	C P A V	A F	O F A	C J C B	M J V E	M H C	G B P E	C A C
<i>bispo</i>										X
<i>governador</i>										
<i>conde barão</i>								X		
<i>conde regedor</i>				X						
<i>marquês</i>						X				
<i>embaixador</i>										
<i>mestre-escola</i>										X
<i>oficiais gerais</i>						X				
<i>padre assistente</i>			X							
<i>padre confessor</i>			X							
<i>padre cura</i>							X			
<i>padre frei</i>		X	X					X	X	
<i>padre mestre</i>							X			
<i>padre prior</i>	X									
<i>padre procurador</i>			X							

<i>padre visitador</i>			X						
------------------------	--	--	----------	--	--	--	--	--	--

Fonte: a autora

Ao se observar o Quadro 4, nota-se, de imediato, a existência de 8 construções cujo elemento situado à esquerda é a palavra *padre*. A maioria dessas construções apareceu na coletânea de cartas do Padre António Vieira (séc. XVII), o que se justifica em virtude dos assuntos tratados nesses documentos.

Com exceção de *padre frei* e *padre mestre*⁶⁵, nas demais construções o elemento *padre* atua sozinho como núcleo, já que é responsável por designar a categoria gramatical e as propriedades morfológicas de gênero e de número dos produtos composicionais. Por exemplo, em *padre visitador*, o termo à esquerda determina a classe a que pertence o produto composicional (substantivo) e os traços morfológicos (masculino e singular). Além disso, observa-se, nesses compostos, que o elemento *padre* também é núcleo semântico, pois funciona como um hiperônimo do produto composicional, ou seja, como um termo mais abrangente. Desse modo, entende-se que um composto como *padre visitador* corresponde a um padre que cumpre uma função específica: a pessoa responsável por verificar, outrora, as atividades de outros padres nas colônias⁶⁶.

Quanto à natureza morfológica, as 8 construções encabeçadas por *padre* inserem-se na classe dos compostos morfossintáticos, uma vez que exibem desvios em relação aos padrões sintáticos da língua portuguesa. No caso dos compostos [NN]_N, o desvio normalmente está relacionado à ausência de conectores entre os elementos dos produtos composicionais. Embora não seja possível deduzir qual conector está ausente entre os constituintes de quase todos os compostos mencionados, em *padre frei* pressupõe-se a ausência da conjunção *e* entre os constituintes do produto composicional por causa do significado veiculado pelo composto: *padre frei* é um membro religioso que acumula os títulos eclesiásticos de *padre* e *frei*.

Aliás, esse último composto chama a atenção no que tange ao número de ocorrências. No âmbito dos compostos associados a denominações humanas, *padre frei* é o que mais se sobressai no *corpus*, surgindo 30 vezes, em 5 textos diferentes. Destaca-se, ainda, a variabilidade gráfica com que se registra, como se verifica nos excertos a seguir:

«com este seu intento deu motivo a que esta história se compusesse depois pelo **Padre Fr.**
Bernardo de Brito» [DVP, capítulo *Vida de João de Barros*]

⁶⁵ Ao tratar das relações sintáticas intracomposto, discutir-se-á, de forma mais detalhada, o caso de *padre mestre*.

⁶⁶ Informação obtida através do texto de Wrege (2014).

«o **Pe. Frei** Antonio da Piadade se foi para Varatojo deixando feita a renúncia da Abbadeça de Santa Clara.» [GBPE, *Diário de Lisboa de 9 de agosto de 1729*]

«Aqui esteve alguns dias o **Pe. Fr.** Antonio da Piedade filho do Conde da Ericeira recebendo delRey muitas honras.» [GBPE, *Diário de 20 de dezembro de 1729*]

«Sahiu academico Alexandre Ferreira Deputado da Meza da Conçiença em lugar do **Pe. Frej** Jozeph da Purificação» [GBPE, *Diário de 17 de abril de 1731*]⁶⁷

No tocante à estrutura dos constituintes, as palavras que compõem o composto podem ser simples ou complexas. *Padre, cura, frei*⁶⁸, *mestre* e *prior* são designados como palavras simples, já que possuem como constituintes morfológicos apenas o radical e a vogal temática (*padre*) ou apenas o radical (*prior*). Já *assistente, procurador* e *visitador* são consideradas palavras complexas, pois apresentam morfemas derivacionais adicionais, como os sufixos nominais *-nt(e)* e *-dor*, cujo valor semântico é, nesses casos, de ‘agentivo que porta uma dada função’. Sincronicamente, considera-se que esses nomes foram formados através da combinação dos sufixos com temas verbais (*assistir* → *assistente*, *procurar* → *procurador*, *visitar* → *visitador*)⁶⁹.

Na forma *confessor*, por sua vez, apresenta-se o sufixo *-sor*, variante alomórfica de *-dor*⁷⁰. Em sua *Gramática Histórica Portuguesa*, Nunes (1956, p. 368) faz referência à vitalidade do sufixo *-dor* (< lat. *-tor*), utilizado para formar nomes relacionados não só a agentes humanos (*procurador, visitador* etc.), mas também a instrumentos (*ralador, escorredor* etc.). Nas formações cultas ou populares, em que o abrandamento do *-t-* foi impedido por causa de uma consoante precedente, houve a manutenção da forma originária, como se observa em *escritor, agricultor, progenitor* etc. Já o *-s-*, que se originara, no latim, por meio da combinação das dentais *-d-* e *-t-*, passou para o português com alguns nomes: *ofensor, defensor, impressor, confessor* etc.

Não obstante a constituição mórfica de *confessor* orientar, primeiramente, para a leitura de ‘aquele que confessa’, o sentido desse termo no composto *padre confessor* é outro:

⁶⁷ Nos excertos, é possível observar a existência das formas abreviadas Pe. e Fr., as quais correspondem, respectivamente, a *padre* e *frei*. Enquanto Pe. é uma abreviatura por síncope (quando faltam letras do meio do vocábulo), Fr. é uma abreviatura por apócope (quando falta o final da palavra) (FLEXOR, 2008).

⁶⁸ O substantivo *frei* corresponde à forma apocopada de *freire* (CUNHA, 2013, s.u. frade).

⁶⁹ Salienta-se que a maioria das palavras complexas que constituem os compostos do *corpus* foi formada no latim. Todavia, faz-se aqui uma análise sincrônica, buscando observar os afixos que se mantêm ativos no português.

⁷⁰ Também está presente, em *confessor*, o prefixo *con-*, de sentido opaco. Para uma descrição mais detalhada deste prefixo (e de outros prefixos) em termos morfolexicais, semânticos e etimológicos, remetemos o leitor para a tese de doutorado de Lopes (2018).

«O **Padre Confessor** da Princesa, nossa senhora, me dá boas esperanças do que tanto se deseja.» [CPAV, *Carta CLXI*]

Neste excerto, faz-se a leitura de que o *padre confessor* não é aquele que confessa, mas sim aquele que ouve as confissões, as confidências de alguém, ou aquele que ouve em confissão.

Quanto à análise sintática, as construções *padre assistente*, *padre confessor*, *padre cura*, *padre prior*, *padre procurador* e *padre visitador* constituem-se como compostos modificativos. Esse tipo de composto caracteriza-se pela presença de um elemento modificador ou determinante, cuja função é atribuir propriedades ao núcleo. Nos compostos acima referidos, o elemento modificador situa-se à direita, enquanto o elemento modificado está localizado à esquerda. Em *padre assistente*, por exemplo, o elemento à direita atua como modificador, tendo em vista que delimita a referência do elemento à esquerda a que se associa: um *padre assistente* é aquele que assiste o bispo nas funções pontifícias ou que assiste o pároco no cuidado pastoral dos paroquianos⁷¹.

No *corpus*, é possível observar, ainda, um fenômeno relacionado a alguns compostos modificativos: a possibilidade de redução do composto ao elemento modificador. Esse fenômeno ocorre, mais especificamente, com *padre cura*, isto é, o padre encarregado de uma paróquia (*padre cura* → *cura*):

«E me atrevo hir bufcallo a caza do **Cura**/ Se me favorecer minha ventura» [MJVE, *Entremez intitulado da vossa farinha*]

«Acode o **Cura** à porta, & metea dentro.» [MJVE, *Entremez intitulado da vossa farinha*]

Segundo Bustos Gisbert (1986, p. 85), algumas construções podem ser reduzidas ao modificador, pois esse elemento representa o específico no conteúdo semântico do composto. Em relação a *padre cura*, deduz-se que o elemento situado à direita pode aparecer sozinho devido ao significado que apresenta: *cura* corresponde a *pároco*, o sacerdote responsável por uma paróquia.

Além de *padre cura*, há outros compostos que se encontram reduzidos ao modificador. Ainda que isso não ocorra no *corpus* de análise, constata-se esse fenômeno em outras fontes, como as obras lexicográficas contemporâneas. A título de exemplo, no Dicionário Houaiss da

⁷¹ Significado elaborado a partir de Goggin (1911) e de informações obtidas na enciclopédia *Wikipédia*.

Língua Portuguesa (edição física) e no Dicionário Caldas Aulete (digital), as construções *padre confessor* e *padre prior* não foram registradas. Entretanto, os modificadores foram identificados apresentando acepções muito similares às dos compostos: o *padre confessor*, ou apenas *confessor*, é o padre que ouve confissões; o *padre prior*, ou simplesmente *prior*, é o religioso superior de algumas ordens.

No caso de *padre frei*, não se trata de um composto modificativo, mas de um composto coordenado. Nessa formação, os dois constituintes atuam simultaneamente como núcleo categorial, morfológico e, ainda, semântico⁷², estabelecendo uma relação de adição, pois o *padre frei* é, ao mesmo tempo, *padre* e *frei*.

A respeito de *padre mestre*, há duas interpretações possíveis para o tipo de relação sintática existente entre os constituintes do composto. A partir das acepções fornecidas por dicionários, crê-se logo que se trata de uma estrutura coordenada, pois o composto é descrito como uma pessoa que exerce duas funções: ‘sacerdote que se ocupa também do magistério’; ‘tratamento dado ao sacerdote que ao mesmo tempo é professor’⁷³. Contudo, também é possível considerar a existência de uma relação de modificação entre os constituintes, dado que *padre mestre* diz respeito a um tipo específico de padre – é um sacerdote que se ocupa da tarefa de ensinar. Saliente-se, no entanto, que na sua análise como modificativo, somente o elemento *padre* pode atuar como núcleo.

Quanto à semântica, os compostos que têm como primeiro elemento a palavra *padre* exibem caráter [+ composicional], uma vez que as acepções correspondem, em algum grau, à soma dos significados dos elementos compositivos. Exemplificando, *padre procurador* apresenta uma acepção resultante da soma dos significados das partes: ‘**sacerdote** que **administra** as propriedades dos mosteiros’⁷⁴.

Ainda no que tange às observações semânticas, salienta-se que as inferências feitas sobre as relações semânticas instituídas entre os constituintes têm como base o significado do produto composicional. A respeito disso, Ribeiro e Rio-Torto (2016, p. 512) apontam que a identificação do tipo de relação semântica instaurado entre os constituintes de um composto costuma ser mais fácil quanto menos opaca for a sua significação global. Todavia, nos compostos cujo semantismo global se afasta consideravelmente da soma dos significados dos constituintes, torna-se muito difícil identificar o tipo de relação semântica estabelecida. Mesmo

⁷² Declara-se a existência de dois núcleos semânticos em compostos coordenados tendo como base Bisetto e Scalise (2009).

⁷³ Acepções retiradas, respectivamente, de Houaiss e Villar (2009, *s.u.* padre-mestre) e do Dicionário Caldas Aulete (digital, *s.u.* padre-mestre).

⁷⁴ Acepção retirada de Gleig (1845).

que as acepções dos compostos encabeçados por *padre* sejam deduzidas, de alguma forma, a partir dos elementos formativos, admite-se que as análises a serem feitas acerca das relações semânticas são passíveis de interpretações distintas tendo em vista a dificuldade na identificação do tipo de relação envolvida.

Nas construções *padre assistente*, *padre confessor*, *padre cura*, *padre mestre*⁷⁵, *padre prior*, *padre procurador* e *padre visitador*, analisamos como sendo de *agente-atividade* a relação semântica instaurada entre os respectivos elementos compositivos. Segundo Adams (2001, p. 86), os compostos cujos núcleos correspondem a um referente humano provavelmente indicam algo relacionado à ocupação ou à identidade da pessoa referida. Isso se verifica nas combinações elencadas, pois enquanto os constituintes da margem esquerda se referem a um indivíduo, os modificadores indicam a atividade/ocupação com a qual esse indivíduo está envolvido.

Em *padre frei*, é possível inferir a existência de uma relação semântica de *multifuncionalidade* (RENNER, 2008). Segundo Renner (2008, p. 608), os compostos multifuncionais podem ser reconhecidos através do seguinte teste: (um) X.Y é (um) X que também é (um) Y. Observa-se que esse teste pode ser aplicado a *padre frei*, uma vez que esse composto denota um indivíduo que é um padre e também é um frei. Logo, entende-se que um *padre frei* desempenha, de alguma forma, dupla função.

No *corpus*, além das formações encabeçadas por *padre*, há outros compostos associados a denominações humanas, correspondentes a ofícios desempenhados no âmbito da Igreja Católica (*bispo governador*, *mestre-escola*), a títulos nobiliárquicos e/ou cargos importantes (*conde regedor*, *conde barão*, *marquês embaixador*) ou a grau hierárquico militar (*oficiais gerais*). Esses compostos também não seguem os padrões sintáticos típicos, fazendo pressupor a existência, implícita, de um conectivo: *conde e regedor*, *conde e barão*, *marquês e embaixador* e *mestre de escola*.

A análise da estrutura mórfica das formas que constituem os compostos leva em consideração a sua composicionalidade nas sincronias em questão, aspecto que pode ser observado nas acepções fornecidas por dicionários produzidos no período. No que se refere aos elementos formadores desses compostos, há constituintes simples e complexos. Nomes como *bispo*, *mestre*, *escola*, *conde*, *barão*, *marquês* e *embaixador* são classificados como simples⁷⁶.

⁷⁵ Nesta análise, considera-se *padre mestre* um composto modificativo.

⁷⁶ *Marquês* e *embaixador* tratam-se de casos mais difíceis de serem analisados sincronicamente, pois não se encaixam nas regras de derivação sufixal da língua portuguesa. Etimologicamente, *marquês* originou-se no português através do provençal *marques*, derivado do germânico *marka* 'limite, fronteira' (CUNHA, 2013, *s.u.*)

Por outro lado, *governador*, *regedor*, *oficiais* e *generais* correspondem a nomes complexos. *Governador* e *regedor* são formados pelo sufixo *-dor* (*governar* → *governador*, *reger* → *regedor*). Já *oficiais* (singular *oficial*) e *generais* (singular *general*) correspondem a casos de recategorização: são adjetivos que passaram por um processo de substantivação⁷⁷.

Figuram, nessas construções, relações sintáticas intracomposto de coordenação e de modificação. Os compostos *conde barão*, *conde regedor* e *marquês embaixador* são binucleares, isto é, incluem dois elementos nucleares que instituem relação de adição: um *conde barão* é uma pessoa que detém, concomitantemente, os títulos de conde e barão; um *conde regedor* é uma pessoa que possui o título de conde, e também desempenha o papel de regedor; e um *marquês embaixador* é uma pessoa que possui o título de marquês, e também desempenha o papel de embaixador.

Os compostos *bispo governador*, *mestre escola* e *oficiais gerais*, por sua vez, estabelecem relação de modificação, posto que os elementos da margem direita se comportam como determinantes do núcleo (morfológico, categorial e semântico) situado à esquerda, restringindo-os: *bispo governador* é um tipo de bispo; *mestre escola* é um tipo de mestre (professor); *oficial general* é um tipo de oficial. Aliás, este último composto é o único que aparece somente na sua forma pluralizada – *oficiais gerais* –, correspondendo a ‘oficial de grau hierárquico superior a coronel ou a capitão de mar e guerra’:

«me parece dizer a V. Ex.a que os **oficiais gerais** emprestados não servem mais»
[CJCB, Carta 84]

Os compostos modificativos [NN]_N apresentam variação na marcação de plural, que pode incidir somente no núcleo ou em ambos os termos. É essa segunda possibilidade que se observa em *oficiais gerais*, comportamento registrado em Houaiss e Villar (2009) e Caldas Aulete (digital).

Da mesma maneira que os compostos encabeçados por *padre*, os demais compostos associados a denominações humanas apresentam a propriedade [+ composicional], pois a interpretação semântica do produto condiz com a soma do significado dos constituintes. Por exemplo, o *bispo governador* é o **bispo** responsável por **governar** um bispado. Desse modo,

marca). Já *embaixador* provém do francês *ambassadeur*, o qual é um empréstimo do italiano *ambasciata*, *-atore* (CUNHA, 2013, s.u. embaixador; BLOCH; WARTBURG, 2008, s.u. ambassade).

⁷⁷ Etimologicamente, *oficial* provém do latim *officialis* e *general* é um caso de empréstimo, provavelmente do castelhano *general* (CUNHA, 2013, s.u. oficial).

nesse composto, vê-se que o semantismo global coincide com o significado dos elementos combinados.

Quanto à natureza da relação semântica instituída entre os constituintes do composto, interpretamos que ela é do tipo *agente-atividade*, no caso dos compostos *bispo governador*, *oficiais gerais* e *mestre-escola*, dado que os referentes humanos exercem funções indicadas, de alguma forma, pelos modificadores. Além disso, faz-se a leitura de que os constituintes da formação *mestre-escola*⁷⁸ estabelecem uma relação semântica mais específica, a de *locativo*, posto que a escola é um dos locais em que um mestre, isto é, um professor leciona.

Por seu turno, interpreta-se que em *conde barão*, *conde regedor* e *marquês embaixador* a relação semântica estabelecida entre os dois nomes é a de *multifuncionalidade* (RENNER, 2008), pois ocorre a designação de indivíduos pela sua dupla atribuição: a de conde e barão, a de conde e regedor, a de marquês e embaixador.

Em síntese, a análise linguística dos compostos morfossintáticos [NN]_N associados a denominações humanas mostrou, do ponto de vista morfológico, que os elementos compositivos formadores das construções podem ser simples ou complexos. Do ponto de vista sintático, as relações estabelecidas entre os constituintes podem ser de modificação ou de coordenação. Por fim, do ponto de vista semântico, observou-se a existência de estruturas [+composicionais], as quais instituem relações semânticas de agente-atividade, de multifuncionalidade ou, em um caso particular, de locativo.

5.2 COMPOSTOS [NN]_N CORRESPONDENTES A ESPÉCIES ZOOLÓGICAS

Apesar de ser bastante comum o uso do padrão [NN]_N para nomear espécies zoológicas, situação facilmente verificável em obras lexicográficas, no *corpus*, que abrange textos dos séculos XVII e XVIII, foram identificados apenas quatro compostos dessa natureza, destacados em negrito nos fragmentos a seguir:

«tambem he certo, que dizem, e se escreve, que levarãõ para Castella o livro do **Porco espim**»
[AF, *Reposta da Senhora Dona Catharina contra as razoens delRey D. Filippe*]

«Este peixe ha no bayrro da Pampulha, Enfia-os muy bem o **peixe agulha**.»
[MJVE, *Entremez dos pexes*]

⁷⁸ Em Santos (2009), apresentam-se as seguintes acepções para *mestre-escola*: ‘dignidade de algumas igrejas catedrais, a cujo cargo estava antigamente o ensino das ciências eclesiásticas’; ‘dignidade em cabido, de graduação inferior’..

«Vede a laftima, olhay o defarranjo, Tal eftà que dirà que he **pexe Anjo**.»
[MJVE, *Entremez dos pexes*]

«Precatarme hey, ã eftou ameaçada? Com rifco ferirme o **pexe espada**⁷⁹.»
[MJVE, *Entremez dos pexes*]

O composto *porco espim* está registrado na *Arte de Furtar* (1652), enquanto os compostos com o núcleo *peixe* à esquerda registram-se na obra *Musa Jocoza de Varios Entremezes Portuguezes e Castelhanos* (1709). No caso dessa última obra, trata-se de um conjunto de entremezes, isto é, de peças teatrais de curta duração e com tom jocoso. Um dos entremezes que faz parte dessa coletânea é intitulado *Entremez dos pexes*, o que justifica, de antemão, o registro dos compostos destacados nos fragmentos acima.

No caso de *porco espim*, presume-se a ausência de um conector: *porco com espim*. Em relação à estrutura dos elementos compositivos, considera-se, na sincronia, que todos os termos correspondem a palavras simples, embora uma análise etimológica permita identificar formas historicamente complexas, como a palavra *agulha*, derivada em latim (*acus* + *-ula* (sufixo diminutivo) > *acucula* ‘agulha’). Na atualidade, o falante comum muito provavelmente desconhece o étimo desta e de outras palavras que contêm sufixos diminutivos eruditos, como *apícula*, *flâmula*, *testículo* etc., pois dificilmente enxergaria a noção diminutiva que essas palavras apresentavam quando foram construídas (CARVALHO, 2012, p. 692).

Todos os compostos [NN]_N associados a animais instituem uma relação sintática de modificação. Nessas construções, observa-se que o elemento modificador, situado à direita, atribui traços específicos ao núcleo nominal, através de processos de metaforização e/ou metonimização.

Consoante Benczes (2006a, 2006b), os compostos [NN]_N que utilizam metáfora e/ou metonímia podem ser chamados de *compostos criativos*, uma vez que se valem de associações baseadas em semelhança e em contiguidade. Para a autora (2006a, p. 59), os compostos criativos são produzidos porque o pensamento metafórico e metonímico trata-se de uma habilidade natural dos seres humanos. Desse modo, esse pensamento é usado para interpretar e/ou conceituar o mundo de um modo particular. Ademais, o conhecimento cultural dos falantes também pode favorecer o surgimento de compostos criativos.

Nas formações *peixe agulha*, *peixe anjo* e *peixe espada*, o primeiro elemento,

⁷⁹ Nesta dissertação, mantêm-se as formas ortográficas encontradas no *corpus* para os compostos *porco-espinho*, *peixe-agulha*, *peixe-anjo* e *peixe-espada* a fim de evidenciar a grafia do período.

correspondente ao núcleo morfológico, categorial e semântico, permanece com o seu significado literal, sendo, portanto, descritivo, enquanto o segundo elemento apresenta uma carga semântica de natureza metafórica. Nesses casos, a especificação dos tipos de peixe é feita por meio de uma associação por semelhança, pois o segundo constituinte é, em algum aspecto, comparado ao primeiro. A formação do composto *pexe agulha* se deve, certamente, à similaridade existente entre uma parte do corpo desse peixe, o bico, e a agulha, objeto longo e estreito. O *pexe anjo*, por sua vez, provavelmente foi assim designado por causa de sua forma: as nadadeiras parecem asas de anjo. E quando esse peixe é visto sob um determinado ângulo – com a cabeça para baixo –, é possível notar a similitude de seu corpo com a silhueta de um anjo. Finalmente, o composto *pexe espada* corresponde a um peixe cuja cauda, por ser pontiaguda, assemelha-se a uma espada. Tendo em vista a associação por semelhança evidenciada nesses três compostos, depreende-se uma relação semântica do tipo ‘forma/similitude’ entre os elementos compositivos dessas estruturas.

Embora os compostos mencionados sejam designados como formações metafóricas, é necessário salientar que essas construções também possuem base na metonímia, já que esta pode desempenhar uma função essencial na origem de construções resultantes de processos de metaforização (SILVA, 2011, p. 19). Quando o segundo elemento se associa ao primeiro elemento, somente algumas propriedades daquele são selecionadas, havendo, então, uma relação do tipo ‘todo pela parte’. Em *pexe agulha* e *pexe espada*, somente as características relacionadas ao formato dos objetos foram evocadas; em *pexe anjo*, houve a seleção apenas das características ligadas à aparência corpórea de um anjo.

No caso de *porco espim*, a análise que se faz desse composto leva à identificação de processos metonímicos e metafóricos. Entende-se que o segundo elemento – *espim* – instancia um processo metonímico, uma vez que aciona a relação semântica de ‘parte pelo todo’, pois os espinhos que recobrem todo o dorso⁸⁰ servem como ponto de referência para designar o animal: o *porco* (tem) *espinho*. O processo metafórico pode ser observado no núcleo – *porco* –, uma vez que não se trata de um tipo de porco – um animal da família dos suídeos –, mas sim de um mamífero roedor. Apesar de não serem muito similares aos porcos, os *porcos-espim* (ou *porcos-espinhos*, designação mais comum na atualidade) possuem um corpo mais atarracado, pernas curtas e emitem grunhidos; talvez resida nessas propriedades a motivação para a designação atribuída a essa espécie de roedor.

⁸⁰ Em zoologia, espinhos são pelos grossos, ásperos e pontiagudos que revestem o corpo de alguns mamíferos.

Resumidamente, os compostos morfossintáticos [NN]_N associados a animais são constituídos por palavras simples e instauram relação sintática intracomposto de modificação. Quanto a aspectos semânticos, os compostos apresentam um grau de composicionalidade/idiomaticidade mediano, tendo em vista a ação de processos cognitivos metafóricos e metonímicos.

5.3 OUTRAS CONSTRUÇÕES IDENTIFICADAS NO *CORPUS*

5.3.1 Peixe mulher e peixe homem

As construções *peixe mulher* e *peixe homem* foram identificadas na composição dramática *Entremez dos peixes*, a qual faz parte da coletânea *Musa Jocoza de Varios Entremezes Portuguezes e Castelhanos* (1709). Nesse texto, há algumas construções Nome + Nome que têm como primeiro elemento a palavra *peixe*, como as formações mencionadas.

A análise que se faz a respeito de *peixe mulher* e *peixe homem* é a de que ambos não são compostos [NN]_N correspondentes a tipos de peixe, como *peixe-agulha*, *peixe-anjo*, *peixe-espada* etc., mas sim construções utilizadas como uma espécie de recurso próprio do gênero textual *entremez*:

«Eftas couzas me confomem,/ Viftes **peixe mulher**, nũa **peixe homẽ**?»
[MJVE, *Entremez dos peixes*]

Como esse gênero textual é marcado pelo caráter cômico e irônico, observa-se, através da leitura do excerto, que *peixe mulher* e *peixe homem* promovem um tom jocoso à composição dramática. Ademais, inferimos que essas construções foram empregadas estrategicamente para dar continuidade à sequência de rimas emparelhadas que aparecem ao longo de todo o texto.

No caso de *peixe mulher*, alguns dicionários, como o de Houaiss e Villar (2009, *s.u.* peixe-mulher) e o Caldas Aulete (digital, *s.u.* peixe-mulher)⁸¹, fornecem a seguinte acepção para a construção⁸²: ‘a fêmea do peixe-boi’. No entanto, a analista ambiental Fábila Luna afirma,

⁸¹ Em Bluteau (1712-1728), dentro do verbete *peixe*, está registrado *peixe mulher*, mas a sua acepção é remetida para o verbete *molher*, da qual reproduzimos o seguinte trecho: ‘Nas Ilhas Bociças, quinze légoas de Sofala [Moçambique], ao longo da Costa, para a parte do Sul, há muito peixe, com feições humanas, do ventre até o pescoço. Destes peixes, os que têm mais semelhança de mulher, que de homem, chamam-lhe peixe mulher [...]’ (BLUTEAU, 1712-1728, *s.u.* molher).

⁸² Nesses dicionários, *peixe mulher* é registrado, modernamente, como *peixe-mulher*.

em uma entrevista concedida a um portal de notícias⁸³, que essa estrutura nunca foi utilizada como o feminino de peixe-boi por especialistas em animais marinhos e nem por pescadores e populações ribeirinhas. Aliás, o Instituto Chico Mendes de Conservação e Biodiversidade (ICMBio), órgão ambiental do governo brasileiro, já enviou ofícios a um dicionário solicitando a correção. A construção *pexe homẽ*, por sua vez, não foi encontrada nas obras lexicográficas consultadas⁸⁴.

Embora não se depreendam os significados de *pexe mulher* e *pexe homẽ* no contexto do entremez, acreditamos que as palavras *mulher* e *homẽ* foram utilizadas com a finalidade de realizar contrastes de gênero⁸⁵.

5.3.2 Carro-mato

Nas *Cartas de Antonio da Costa*, obra que constitui o *corpus* de análise, identificou-se a construção *carro-mato*:

«e as cem léguas que lhe faltam desde ali a Madrid, o levam em um **carro-mato**»
[CAC, *Carta VI*]

Essa construção, grafada sem hífen, está registrada em alguns dicionários contemporâneos digitais, os quais apresentam acepções que coincidem, em certo grau:

- (1) ‘carro de rodas grandes, com tabuleiro formado de cordas entrançadas’ (Dicionário Caldas Aulete);
- (2) ‘carroça cujo leito é formado por cordas entrançadas’ (Dicionário Priberam da Língua Portuguesa)
- (3) ‘carroça cuja base é formada por uma trama de cordas’ (Dicionário Infopédia da Língua Portuguesa)

⁸³ Entrevista disponível em: <https://ne10.uol.com.br/canal/noticias/ciencia-e-vida/noticia/2013/11/21/nem-peixe-mulher-nem-peixe-vaca-o-feminino-de-peixe-boi-e-455975.php>. Acesso em: 22 dez. 2021.

⁸⁴ Consultamos as seguintes obras lexicográficas: Bluteau (1712-1728), Houaiss e Villar (2009), Dicionário Caldas Aulete (digital), Dicionário Infopédia da Língua Portuguesa (digital) e Dicionário Priberam da Língua Portuguesa (digital). Também consultamos o *Google Books*.

⁸⁵ Villalva e Silvestre (2014, p. 174), ao abordarem os contrastes de gênero em português, apresentam *águia-macho / águia-fêmea* e *crocodilo-macho / crocodilo-fêmea* como exemplos de contraste por composição, análise que poderia ser estendida às expressões *pexe mulher* e *pexe homẽ* do *corpus*.

Observando somente as acepções encontradas, considera-se *carro-mato* um composto [NN]_N, formado a partir da junção dos substantivos *carro* e *mato*. Contudo, ao se verificar a etimologia dessa construção, percebe-se que *carro-mato* não se trata de uma estrutura constituída por dois substantivos.

Segundo o *Diccionario de la lengua española* (versão eletrônica, *s.u.* carromato), essa forma é oriunda do italiano antigo *carro matto*. Em italiano, *carro* pode significar *carro*, *carroça*, *carreta*. Em contrapartida, o significado de *matto* não corresponde de forma alguma aos sentidos da palavra *mato* da língua portuguesa. De acordo com Sabino (2011, p. 89), *mato* (pt.) e *matto* (it.) são falsos cognatos, pois apesar de serem vocábulos ortograficamente e fonologicamente muito semelhantes, seus valores semânticos são bem distintos. Enquanto *mato*, em português, pode apresentar os sentidos de ‘planta’, ‘terreno’ ou ‘lugar afastado de uma cidade’, em italiano, *matto* é um adjetivo cuja acepção é ‘doido, maluco, louco’. Sendo assim, *carro-mato* não pode ser analisado como um composto [NN]_N.

Segundo Roberts (2014, p. 325), no caso de *carro-mato*, o adjetivo italiano *matto* é empregado a fim de indicar algo que não é apropriado, como esse tipo de veículo. Consequentemente, entende-se que um *carro-mato* não oferece a segurança devida.

5.3.3 Gato çapato

Na peça teatral *O Fidalgo Aprendiz*, foi encontrada a construção *gato çapato*⁸⁶. A princípio, presumimos que essa formação fosse um composto [NN]_N. No entanto, ao analisarmos atentamente o contexto, percebemos que *gato çapato* integra uma expressão idiomática verbal, como se constata no fragmento a seguir:

«porlhe não dar outro meo **faz dellegato çapato**» [OFA, Primeira Jornada]

Quanto ao surgimento dessa expressão, registrada em dicionários, mas com algumas variações, há algumas explicações. Para João Ribeiro (1908, p. 132), autor que se dedicou ao estudo da fraseologia portuguesa, *fazer de gato sapato* trata-se de um erro de quem lia ou escrevia no período em que o uso constante de abreviaturas poderia provocar enganos. Outrora

⁸⁶ As formas ortográficas atuais dessa construção são *gato-sapato* e *gato e sapato*.

escrevia-se a palavra *sapato* com *ç* e a palavra *gato* podia ser lida como *çapato* devido à abreviatura *çato*. É possível, então, que essa confusão tenha dado origem à expressão.

Cascudo (1986), por sua vez, descreve *gato-sapato* como um jogo infantil muito antigo, uma modalidade da cabra-cega. Nesse jogo, uma criança, de olhos vendados, é batida pelos companheiros com sapatos e chinelos, até que consiga agarrar a um deles. A criança que for agarrada passará a ficar com os olhos vendados. Destaca-se que as interações que ocorrem nessa brincadeira parecem ter relação, de certa forma, com os significados da expressão idiomática *fazer de gato sapato*. Nessa expressão, não há intenção de se fazer alusão a um gato ou a um sapato. Ela possui, na verdade, valor semântico depreciativo, sendo utilizada pelos falantes com os sentidos de tratar com desprezo, explorar, humilhar etc.

Neste capítulo, descrevemos e analisamos os compostos [NN]_N extraídos do *corpus* quanto a aspectos morfológicos, sintáticos e semânticos. Também questionamos o estatuto de composto [NN]_N das construções *peixe mulher*, *peixe homem*, *carro-mato* e *gato çapato*, registradas durante o processo de coleta de dados. A análise empreendida permitiu concluir que essas construções não correspondem a compostos [NN]_N. Insta salientar, no entanto, *que peixe mulher* e *peixe homem*, expressões que evidenciam um contraste de gênero através das palavras *mulher* e *homem*, podem eventualmente ser consideradas formas compostas, se considerarmos a abordagem de Villalva e Silvestre (2014, p. 174) para o gênero em português (veja-se a nota 72).

6 CONSIDERAÇÕES FINAIS

Nesta dissertação, estudamos o padrão compositivo [NN]_N em duas sincronias pretéritas da língua portuguesa (séculos XVII e XVIII), com a finalidade de observar o seu desenvolvimento ao longo do período considerado. Para realizar esta investigação, tivemos como base um *corpus* representativo do ponto de vista da abrangência temporal e dos gêneros textuais.

Com relação ao aporte teórico, apoiamo-nos em alguns estudos, mas é necessário destacar novamente, por algumas razões, os trabalhos das portuguesas Silvia Ribeiro e Graça Rio-Torto sobre a composição, principalmente o trabalho publicado em 2016. As colocações feitas por essas pesquisadoras a respeito do padrão compositivo [NN]_N foram fundamentais para que refletíssemos sobre o processo de desenvolvimento dessa estrutura na língua portuguesa. Além disso, é preciso ressaltar que acreditamos, de fato, na visão abrangente de composição difundida pelas pesquisadoras e, também, na sua proposta de classificação voltada especificamente ao português, a qual foi utilizada na descrição e na análise dos dados coletados no *corpus*.

Quanto aos resultados, identificamos, no total, 18 palavras compostas [NN]_N correspondentes a denominações humanas e a espécies zoológicas. Esses compostos dividiram-se entre 14 modificativos (*mestre-escola*, *porco espin*) e 4 coordenados (*padre frei*, *conde barão*). Embora o número de ocorrências obtido por Santos (2009) em sua tese de doutorado seja menor – a autora identificou 10 compostos [NN]_N em um *corpus* que abrange os séculos XIII, XIV, XV e XVI –, não evidenciamos, em termos quantitativos, um aumento significativo de compostos [NN]_N durante os séculos XVII e XVIII. Logo, considerando o *corpus* de análise, não consideramos que esse padrão compositivo estivesse em crescente desenvolvimento nesse período.

Os resultados alcançados nesta dissertação provocam alguns questionamentos acerca da produtividade do padrão compositivo [NN]_N na língua portuguesa: (1) Em que momento ocorre o processo de expansão de compostos [NN]_N no português? (2) A produção de compostos [NN]_N está relacionada sobretudo à formação de esquemas⁸⁷ em contextos específicos? Obviamente, ainda não temos resposta para a primeira pergunta, todavia, é possível realizar algumas considerações sobre a segunda indagação.

⁸⁷ Utilizamos o termo *esquema* para fazer referência às formações em série, como os diversos compostos cujos elementos situados à esquerda são as palavras *peixe*, *padre*, *bolsa* etc.

Ao compararmos, por exemplo, parte dos dados deste estudo com os dados de Faria (2011)⁸⁸, observamos a presença de alguns esquemas, como *padre-X* e *peixe-X*, neste trabalho, e *auxílio-X*, *bolsa-X*, *seguro-X* e *vale-X*, no trabalho de Faria (2011). Apesar de esses esquemas terem se tornado mais numerosos nas últimas duas décadas, principalmente no PB, este estudo já mostra a existência dessas formações em série. Enquanto os compostos constituídos por ‘auxílio’, ‘bolsa’, ‘seguro’ e ‘vale’ surgiram, conforme aponta Faria (2011, p. 113), devido a questões sociais, supomos que os compostos formados a partir de *padre* e *peixe* surgiram por outros motivos. No caso do esquema *padre-X*, percebemos que os compostos foram gerados para designar e especificar as múltiplas atividades que poderiam ser desempenhadas pelos sacerdotes. O esquema *peixe-X*, por sua vez, nomeia tipos de peixe. Aliás, no português atual, há uma quantidade considerável de nomes de espécies zoológicas que segue o padrão compositivo [NN]_N. A consulta a alguns dicionários, ainda que de maneira superficial, permite observar várias instanciações desse padrão, podendo ser depreendido, em muitos casos, um esquema específico de construção, como, por exemplo, *cobra-X* (*cobra-cipó*, *cobra-papagaio*, *cobra-jabuti* etc.), *gavião-X* (*gavião-pato*, *gavião-relógio*, *gavião-tesoura* etc.), *tubarão-X* (*tubarão-baleia*, *tubarão-martelo*, *tubarão-tigre* etc.), entre outros.

Destacamos, ainda, a atuação de processos cognitivos de metáfora e metonímia na formação de compostos [NN]_N, como observado nos compostos do *corpus* correspondentes a espécies zoológicas – *peixe-agulha*, *peixe-anjo*, *peixe-espada* e *porco-espinho* – e em formações recentes como *mulher-fruta*, *ciclone-bomba*, *bueiro-bomba*, mencionadas em 3.6.

Finalizamos este estudo com a expectativa de que ele contribua, de alguma forma, para os estudos de morfologia histórica, em especial da composição na língua portuguesa. Como perspectivas consequentes da elaboração desta dissertação, acreditamos que outras investigações sobre o padrão compositivo [NN]_N sejam necessárias, tanto nos séculos XVII e XVIII, com base em outros conjuntos documentais, quanto nos séculos XIX e XX. A partir disso, será possível estabelecer o percurso diacrônico desse padrão no português e, também, ter um conhecimento maior sobre a sua produtividade.

⁸⁸ A tese de Faria (2011) analisa a motivação morfossemântica de compostos [NN]_N no PB.

REFERÊNCIAS

a) Referências gerais

ADAMS, Valerie. **Complex Words in English**. Londres: Routledge, 2001.

ARCODIA, Giorgio Francesco. Coordinating nominal compounds: Universal vs. areal tendencies. **Linguistics**, De Gruyter, v. 56, n. 6, p. 1197–1243, 2018.

ARNAUD, Pierre. Noun-Noun compounds in french. *In*: MÜLLER, P.; OHNHEISER, I.; OLSEN, S.; RAINER, F. (ed.). **Word-Formation**: an internacional handbook of the languages of Europe. v. 1. Berlim; Boston: De Gruyter Mouton, 2015. p. 673-688. <https://books.google.com.br/>. Acesso em: 22 mar. 2020.

AULETE, Caldas. **Aulete Digital** – Dicionário Contemporâneo da Língua Portuguesa. Versão *online*. Disponível em: <http://www.aulete.com.br/>.

BADER, Françoise. **La formation des composés nominaux du latin**. Paris: Les Belles Lettres, 1962.

BANDEIRA, Joalêde Gonçalves. **Carteando e dialogando com o pretérito mais-que-perfeito**: caminhos trilhados do século XVI. 2011. 333 f. Tese (Doutorado em Letras e Linguística) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2011. Disponível em: <https://repositorio.ufba.br/ri/handle/ri/8399>. Acesso em: 23 ago. 2020.

BARRETO, Sara de Oliveira Gomes; MARCILESE, Mercedes; OLIVEIRA, Ágata Jéssica Avelar de. Idiomaticidade, familiaridade e informação prévia no processamento de expressões idiomáticas do PB. **Letras de Hoje**, v. 53, n. 1, p. 119-129, jan./mar. 2018. Disponível em: <https://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/28680>. Acesso: 20 jul. 2020.

BASILIO, Margarida. Em torno da palavra como unidade lexical: palavras e composições. **Veredas** – Revista de Estudos Linguísticos, Juiz de Fora, v. 4, n. 2, p. 9-18, jul-dez 2000. Disponível em: <https://periodicos.ufjf.br/index.php/veredas/article/view/25314>. Acesso em: 26 jul. 2021.

BASILIO, Margarida. **Formação e classes de palavras no português do Brasil**. São Paulo: Contexto, 2004.

BASILIO, Margarida. Metáfora e metonímia em nomes compostos em português: um estudo de construções S-ADJ. **Revista Linguística**, Rio de Janeiro, v. 12, n. 1, p. 91-105, jan./jun. 2016. Disponível em: <https://revistas.ufrj.br/index.php/rl/article/view/4521>. Acesso em: 21 jul. 2020.

BAUER, Laurie. Some basics concepts. *In*: BAUER, Laurie. **English Word-formation**. Cambridge: Cambridge University Press, 1983. p. 7-41.

BAUER, Laurie. When is a sequence of two nouns a compound in English? **English Language and Linguistics**, v. 2, n. 1, p. 65-86, maio 1998.

BAUER, Laurie. **Introducing Linguistic Morphology**. 2. ed. Edimburgo: Edinburgh University Press, 2003.

BAUER, Laurie. The Borderline between Derivation and Compounding. *In*: DRESSLER, W; KASTOVSKY, D; PFEIFFER, O; RAINER, F. (ed.). **Morphology and its Demarcations**. Amsterdam, Philadelphia: John Benjamins Publishing Company, 2005. p. 97-108. Disponível em: <https://books.google.com.br/>. Acesso em: 03 jun. 2020. Disponível em:

BAUER, Laurie. **Compounds and Compounding**. (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press, 2017.

BAUER, Laurie. Compounds. *In*: AARTS, Bas; BOWIE, Jil; POPOVA, Gergana (ed.). **The Oxford Handbook of English Grammar**. Oxford: Oxford University Press, 2020. p. 262-280.

BENCZES, Réka. Analysing metonymical noun-noun compounds: The case of freedom fries. *In*: BENCZES, Réka; CSÁBI, Szilvia (ed.). **The Metaphors of Sixty: Papers Presented on the Occasion of the 60th Birthday of Zoltán Kövecses**. Budapeste: Eötvös Loránd University, 2006a. p. 53-61.

BENCZES, Réka. **Creative Compounding in English: The Semantics of Metaphorical and Metonymical Noun-Noun Combinations**. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2006b.

BENCZES, Réka. Are exocentric compounds really exocentric? **SKASE Journal of Theoretical Linguistics**, v. 12, n. 3, p. 54-73, 2015. Disponível em: http://www.skase.sk/Volumes/JTL29/pdf_doc/04.pdf. Acesso em: 10 jul. 2020.

BERNAL, Elisenda. Catalan compounds. **Probus**, De Gruyter, 24 (1), p. 5-27, 2012. Disponível em: https://repositori.upf.edu/bitstream/handle/10230/34348/bernal_probus_catcomp.pdf?sequence=1&isAllowed=y. Acesso em: 25 abr. 2020.

BISETTO, Antonietta. Composizione. *In*: GROSSMANN, M.; RAINER, F. (ed.). **La formazione delle parole in italiano**. Tübingen: Max Niemeyer Verlag, 2004. p. 31-95.

BISETTO, Antonietta; SCALISE, Sergio. The classification of compounds. **Lingue e Linguaggio**, v. 4, n. 2, p. 319-332, 2005.

BISETTO, Antonietta; SCALISE, Sergio. The classification of compounds, *In*: LIEBER, Rochelle; ŠTEKAUER, Pavol. (ed.). **The Oxford Handbook of Compounding**, Oxford University Press: Oxford, 2009. p. 34-53.

BLANCO VALDÉS, Xoán. Palabras compostas en galego-portugués. **Verba**, 12, 199-252, 1985.

BLEOTU, Adina Camelia. **The Productivity of NN Compounds in English and Romanian**. 2010. 97 f. Dissertação (Mestrado em Linguística Inglesa) – Faculdade de Línguas Estrangeiras, Universidade de Bucareste, Bucareste, 2010. Disponível em: <https://ling.auf.net/lingbuzz/003144>. Acesso em: 22 mar. 2020.

BLEOTU, Adina Camelia. The acquisition of root NN compounds in Romanian. **Bucharest Working Papers in Linguistics**, XX (2), p. 77-96, 2018. Disponível em: <http://www.diacronia.ro/en/indexing/details/A29088>. Acesso em: 23 mar. 2020.

BLOCH, Oscar; WARTBURG, Walther von. **Dictionnaire étymologique de la langue française**. 3. ed. Paris: Presses Universitaires de France, 2008.

BLOOMFIELD, Leonard. **Language**. Londres: George Allen & Unwin, 1935.

BLUTEAU, Raphael. **Vocabulário português, e latino, aulico, anatomico, architectonico, bellico, botanico, brasilico, comico, critico, chimico, dogmatico, dialectico, dendrologico, ecclesiastico, etymologico, economico, florifero, forense, fructifero... autorizado com exemplos dos melhores escritores portugueses, e latinos**. Coimbra: no Collegio das Artes da Companhia de Jesus, 1712-1728. 10 v. Disponível em: <http://purl.pt/13969>. Acesso em: 10 maio 2020.

BOOIJ, Geert. Compounding. *In*: BOOIJ, Geert. **The Grammar of Words: An Introduction to Morphology**. 2. ed. Oxford: Oxford University Press, 2007. p. 75-95.

BOOIJ, Geert. Compounds and multi-word expressions in Dutch. *In*: SCHLÜCKER, Barbara (ed.). **Complex lexical units: compounds and multiword expressions**. Berlim; Boston: De Gruyter Mouton, 2019. p. 95-126.

BORGES, Yasmim Conceição; SANTOS, Antonia Vieira dos. O estudo de palavras compostas em documentos portugueses da Chancelaria de D. Afonso III (séc. XIII). *In*: ATAÍDE, Cleber; AZEVEDO, Isabel; FREITAG, Raquel (org.). **Linguística e literatura: teoria, análises e aplicações**. Recife: Pipa Comunicação, 2018. p. 249-260.

BORGES, Yasmim Conceição; SANTOS, Antonia Vieira dos. Breve estudo sobre as palavras compostas de configuração Nome-Nome na Crónica de D. João I (Séc. XV). *In*: FIGUEIREDO, Cristina; GAYER, Juliana Ludwig; SOUSA, Lílían Teixeira de; PINTO, Carlos Felipe (org.). **Lingua em movimento: história e funcionamento das línguas naturais**. v.1. Salvador: EDUFBA, 2020, p. 373-387. Disponível em: <https://repositorio.ufba.br/ri/handle/ri/32460>. Acesso em: 10 set. 2021.

BUENAFUENTES DE LA MATA, Cristina. Compounding and variational morphology: the analysis of inflection in spanish compounds. **Borealis – An International Journal of Hispanic Linguistics**, v. 3, n. 1, p. 1-21, 2014. Disponível em: <https://septentrio.uit.no/index.php/borealis/article/view/2828>. Acesso em: 21 jan. 2022.

BUSTOS GISBERT, Eugenio de. **La composición nominal en español**. Salamanca: Ediciones Universidad de Salamanca, 1986.

CARVALHO, Wandercy de. -Inho e as relações semântico-funcionais: a estilística na sala de aula. *In*: CONGRESSO NACIONAL DE LINGUÍSTICA E FILOLOGIA, 16., 2012, Rio de

Janeiro. **Anais eletrônicos** [...] Rio de Janeiro: CiFEFiL, 2012. p. 691-705. Disponível em: http://www.filologia.org.br/xvi_cnlf/tomo_1/tomo_1.pdf. Acesso em: 04 nov. 2021.

CASAL, Manuel Aires de. **Corografia brasílica ou relação historico-geografica do Reino do Brazil**. Tomo I. Rio de Janeiro: na Impressão Régia, 1817. Disponível em: <http://www.bibliotecadigital.puc-campinas.edu.br/>. Acesso em: 10 maio 2020.

CASCUDO, Luís da Câmara. **Locuções tradicionais no Brasil**. São Paulo: Editora da Universidade de São Paulo, 1986.

CORPAS PASTOR, Gloria. **Manual de fraseologia española**. Madrid: Gredos, 1996.

CUNHA, Antônio Geraldo da. **Dicionário etimológico da língua portuguesa**. Rio de Janeiro: Lexikon, 2010.

CUNHA, Celso; CINTRA, Lindley. **Nova gramática do português contemporâneo**. 6. ed. Rio de Janeiro: Lexicon, 2013.

DARMESTER, Arsène. **Traité de la formation des mots composés de la langue française**: comparée aux autres langues romanes et au latin. Paris: Librairie A. Franck, 1875. Disponível em: <https://books.google.com.br/>. Acesso em: 22 mar. 2020.

DETREVILLE, Eleanor. **An overview of latin morphological calques on greek technical terms**: formation and success. 2015. 134 f. Dissertação (Mestrado em Artes) – Universidade da Geórgia, Athens, 2015. Disponível em: https://getd.libs.uga.edu/pdfs/detreville_eleanor_201505_ma.pdf. Acesso em: 05 out. 2020.

DICIONÁRIO Infopédia da Língua Portuguesa [em linha]. Porto: Porto Editora, 2003-2019. Disponível em: <https://www.infopedia.pt/>.

DICIONÁRIO Priberam da Língua Portuguesa [em linha]. Disponível em: <https://dicionario.priberam.org>.

FARIA, André Luíz. **Motivação morfossemântica das construções compostas N-N no português brasileiro**. 2011. 189 f. Tese (Doutorado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.

FERNÁNDEZ-DOMÍNGUEZ, Jesús. The syntagmatic and paradigmatic axes of productivity: an application to the onomasiological model of word-formation. **SKASE Journal of Theoretical Linguistics**, v. 12, n. 3, p. 218-237, 2015. Disponível em: http://www.skase.sk/Volumes/JTL29/pdf_doc/10.pdf. Acesso em: 27 jul. 2020.

FINKBEINER, Rita; SCHLÜCKER, Barbara. Compounds and multi-word expressions in the languages of Europe. In: SCHLÜCKER, Barbara (ed.). **Complex lexical units**: compounds and multiword expressions. Berlin; Boston: De Gruyter Mouton, 2019. p. 1-43.

FLEXOR, Maria Helena Ochi. **Abreviaturas**: manuscritos dos séculos XVI ao XIX. 3. ed. rev. aum. Rio de Janeiro: Arquivo Nacional, 2008.

FRUYT, Michèle. Constraints and productivity in latin nominal compounding. **Transactions of the Philological Society**, v. 100, n. 3, p. 259-287, 2002.

FULGÊNCIO, Lúcia. **Expressões fixas e idiomatismos do Português Brasileiro**. 2008. 506 f. Tese (Doutorado em Linguística e Língua Portuguesa) – Faculdade de Comunicação e Artes – Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, 2008. Disponível em: http://www.biblioteca.pucminas.br/teses/Letras_FulgencioLM_1.pdf. Acesso em: 21 set. 2021.

GARCÍA GALARÍN, Consuelo. **Diccionario histórico de la morfología del español**. Madrid: Guillermo Escolar, 2018.

GASSENFERTH, Denise. **Produtividade lexical**: compostos metafóricos e metonímicos. 1990. 67 f. Dissertação (Mestrado em Letras) – Universidade Federal do Paraná, Curitiba, 1990. Disponível em: <https://acervodigital.ufpr.br>. Acesso em: 20 jul. 2020.

GAVRANOVIĆ, Valentina. Compound nouns in old and middle english. **Zbornik za Jezike i Književnosti Filozofskog Fakulteta u Novom Sadu**, v. 5. n. 5, p. 57-67, 2015. Disponível em: <http://zjik.ff.uns.ac.rs/index.php/zjik/article/view/1572/1575> Acesso em: 17 maio 2020.

GIURESCU, Anca. **Les mots composés dans les langues romanes**. Paris: Mouton, 1975.

GLEIG, George Robert. **Things Old and New**: Being a Sequel to "the Chronicles of Waltham". Londres: J. C. Moore, 1845. Disponível em: <https://books.google.com.br>. Acesso em: 08 out. 2021.

GOGGIN, John. "Assistant Priest." *In*: HERBERMANN, Charles George (ed.). **The Catholic Encyclopedia**. v. 12. New York: Robert Appleton Company, 1911. Disponível em: <http://www.newadvent.org/cathen/12407a.htm>. Acesso em: 08 out. 2021.

GOMES, Ana Luiza de Castro Pereira. José da Cunha Brochado: de secretário de embaixada a embaixador extraordinário. **Revista Tempo**, Niterói, v. 22, n. 39, jan./abr. 2016. Disponível em: https://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-77042016000100072. Acesso em: 15 ago. 2020.

GONÇALVES, Carlos Alexandre. Compostos neoclássicos: estrutura e formação. **Revista Virtual de Estudos da Linguagem**, v. 9, p. 6-39, 2011a. Disponível em: http://www.revel.inf.br/files/artigos/revel_esp_5_compostos.pdf. Acesso em: 03 jun. 2020.

GONÇALVES, Carlos Alexandre. Composição e derivação: polos prototípicos de um continuum? Pequeno estudo de casos. **Domínios de Linguagem**, v. 5, p. 62-89, 2011b. Disponível em: <http://www.seer.ufu.br/index.php/dominiosdelinguagem/article/view/13644>. Acesso em: 03 jun. 2020.

GONÇALVES, Carlos Alexandre. Prefixação: composição ou derivação? Novos enfoques sobre uma antiga polêmica. **Matraga**, Rio de Janeiro, v. 19, n. 30, p. 142-167, jan./jun. 2012. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/matraga/article/view/22625>. Acesso em: 26 set. 2021.

GONÇALVES, Carlos Alexandre. **Atuais tendências em formações de palavras**. São Paulo: Contexto, 2016.

GONÇALVES, Carlos Alexandre; ANDRADE, Katia Emmerick. El status de los componentes morfológicos y el continuum composición-derivación en portugués. **Lingüística**, Montevideu, v. 28, p.119-145, dez./2012. Disponível em: http://www.scielo.edu.uy/scielo.php?script=sci_arttext&pid=S2079-312X2012000100008. Acesso em: 26 set. 2021.

GUEVARA, Emiliano; SCALISE, Sergio. Searching for Universals in Compounding. *In*: SCALISE, Sergio; MAGNI, Elisabetta; BISETTO, Antonietta (ed.). **Universals in Language Today**. Dordrecht: Springer, 2009. p. 101-128.

HERCULANO, Alexandre. **História de Portugal**: desde o começo da Monarquia até o fim do Reinado de Afonso III. Tomo quarto. Lisboa: Em Casa da Viúva Bertrand e Filhos, 1853. Parcialmente disponível em: <https://books.google.com.br>. Acesso em: 30 de set. 2021.

HOUAISS, Antônio; VILLAR, Mauro de Salles. **Dicionário Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva, 2009.

IORDĂCHIOAIA, Gianina; WERNER, Martina. Introduction: Compounds between words and phrases. **Word Structure**, v. 13, n. 2, p 127–141, 2020.

JACKENDOFF, Ray. English noun-noun compounds in Conceptual Semantics. *In*: TEN HACKEN, Pius (ed.). **The semantics of compounding**. Cambridge: Cambridge University Press, 2016. p. 15-37.

KASTOVSKY, Dieter. The problem of productivity in word formation. **Linguistics**, v. 24, p. 585-600, 1986.

LEME, Andreza. **Idiomaticidade e composicionalidade das expressões idiomáticas da língua inglesa**: o significado na interface semântico-pragmática-etimológica. 2008. 166 p. Tese (Doutorado em Letras) – Faculdade de Letras – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2008. Disponível em: <http://tede2.pucrs.br/tede2/>. Acesso em: 20 jul 2020.

LOPES, Mailson dos Santos. **A prefixação na primeira fase do português arcaico**: descrição e estudo semântico- morfolexical-etimológico do paradigma prefixal da língua portuguesa nos séculos XII, XIII e XIV. 2014. 2 v. 943 f. Dissertação (Mestrado em Língua e Cultura) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2013. Disponível em: <https://repositorio.ufba.br/ri/handle/ri/15537>. Acesso em: 27 set. 2021.

LOPES, Mailson dos Santos. **Estudo histórico-comparativo da prefixação no galego-português e no castelhano arcaicos (séculos XIII A XVI)**: aspectos morfolexicais, semânticos e etimológicos. 2019. 2433 f. Tese (Doutorado em Língua e Cultura) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2018. Disponível em: <https://repositorio.ufba.br/handle/ri/29879>. Acesso em: 25 jan. 2022.

LÜDELING, Anke. **Neoclassical word-formation**. Berlin: Universität zu Berlin, 2009.

MAIA, Clarinda de Azevedo. Dos textos escritos à história da língua. *In*: HEAD, Brian F.; TEIXEIRA, José; LEMOS, Aida Sampaio; BARROS, Anabela Leal de; PEREIRA, António (org.). **História da Língua e História da Gramática**. Actas do Encontro. Braga: Centro de Estudos Humanísticos da Universidade do Minho, 2002. p. 231-249.

MAIA, Clarinda de Azevedo. Linguística histórica e filologia. *In*: LOBO, Tânia, CARNEIRO, Zenaide; SOLEDADE, Juliana; ALMEIDA, Ariadne; RIBEIRO, Silvana (org.). **Rosae: linguística histórica, história das línguas e outras histórias**. Salvador: EDUFBA, 2012. p. 533-542. Disponível em: <http://books.scielo.org/id/67y3k/pdf/lobo-9788523212308-38.pdf>. Acesso em: 23 ago. 2020.

MARONEZE, Bruno. A história da pétala: etimologia de um termo científico. **Linha D'Água**, São Paulo, v. 32, n. 3, p. 159-176, set.-dez. 2019. Disponível em: <https://www.revistas.usp.br/linhadagua/article/view/159835>. Acesso em: 10 set. 2020.

MATTOS E SILVA, Rosa Virgínia. **O português arcaico: uma aproximação**. v. 1. Léxico e Morfologia. Lisboa: Imprensa Nacional-Casa da Moeda, 2008a.

MATTOS E SILVA, Rosa Virgínia. **Caminhos da linguística histórica: ouvir o inaudível**. São Paulo: Parábola, 2008b.

MATTOS E SILVA, Rosa Virgínia. Teorias da mudança lingüística e a sua relação com a(s) história(s) da(s) língua(s). **Linguística**, Porto, v. 3, p. 39-53, 2008c. Disponível em: <https://ojs.letras.up.pt/index.php/EL/article/view/2805>. Acesso em: 23 ago. 2020.

MATTOS E SILVA, Rosa Virgínia. **O português arcaico: fonologia, morfologia e sintaxe**. 2. ed. São Paulo: Contexto, 2013.

MOYNA, María Irene. **Compound words in Spanish: theory and history**. Amsterdã, Filadélfia: John Benjamins Publishing Company, 2011.

NUNES, José Joaquim. **Compêndio de gramática histórica portuguesa (fonética e morfologia)**. 5. ed. Lisboa: Livraria Clássica Editora, 1956.

PAIXÃO DE SOUSA, Maria Clara. O Corpus Tycho Brahe: contribuições para as humanidades digitais no Brasil. **Filologia e Linguística Portuguesa**, São Paulo, v. 16, n. spe, p. 53-93, dez. 2014. Disponível em: <http://www.revistas.usp.br/flp/article/view/88404>. Acesso em: 23 ago. 2020.

PRIEST. *In*: Wikipedia, the free encyclopedia. Disponível em: <https://en.wikipedia.org/wiki/Priest>. Acesso em: 08 out. 2021.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de. **Metodologia do Trabalho Científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico**. 2. ed. Novo Hamburgo, Feevale, 2013. *E-book*. Disponível em: <http://www.feevale.br/Comum/midias/8807f05a-14d0-4d5b-b1ad-1538f3aef538/E-book%20Metodologia%20do%20Trabalho%20Cientifico.pdf>. Acesso em: 21 ago. 2020.

NAMIKI, Takayasu; KAGEYAMA, Taro. Word structure and headedness. *In*: KAGEYAMA, Taro; KISHIMOTO, Hideki. (ed.). **Handbook of Japanese Lexicon and Word Formation**. Londres; Berlim: De Gruyter Mouton, 2016. p. 201-235.

ONIGA, Renato. Compounding in Latin. **Rivista di Linguistica**, v. 4, n. 1, p. 97-116, 1992.

PADROSA TRIAS, Susanna. Are there coordinate compounds? **Mediterranean Morphology Meetings**, v. 7, p. 98-111, 2009. Disponível em: <https://mmm.library.upatras.gr/mmm/article/view/2417>. Acesso em: 26 jun. 2020.

PADROSA TRIAS, Susanna. **Complex Word-Formation and the Morphology-Syntax Interface**. 2010. 334 f. Tese (Doutorado em Linguística) – Faculdade de Filosofia e Letras – Universidade Autônoma de Barcelona, Barcelona, 2009. Disponível em: <https://www.tdx.cat/bitstream/handle/10803/32103/spt1de1.pdf>. Acesso em: 23 mar. 2020.

RADIMSKÝ, Jan. **Noun+Noun Compounds in Italian: a corpus-based study**. České Budějovice: Jihočeská univerzita, 2015.

RAINER, Franz. Italian. *In*: MÜLLER, Peter; OHNHEISER, Ingeborg; OLSEN, Susan; RAINER, Franz (ed.). **Word-Formation: an international handbook of the languages of Europe**. v. 4. Berlim; Boston: De Gruyter Mouton, 2015. p. 2712-2730. Disponível em: <https://books.google.com.br/>. Acesso em: 23 mar. 2020.

REAL Academia Española. Diccionario de la lengua española. 23. ed. [versão 23.4 em linha]. Disponível em: <https://dle.rae.es>.

RENNER, Vincent. On the Semantics of English Coordinate Compounds, **English Studies**, v. 89, n. 5, p. 606-613, 2008. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/00138380802011982?scroll=top&needAccess=true>. Acesso em: 16 nov. 2021.

RIBEIRO, João. **Frazes feitas: estudo conjectural de locuções, ditados e provérbios**. Rio de Janeiro: Livraria Francisco Alves, 1908.

RIBEIRO, Sílvia; RIO-TORTO, Graça. Denominações compositivas de estrutura VN, NN, NprepN e NA: nexos intralexicais. *In*: ILIESCU, Maria; SILLER-RUNGGALDIER, Heidi; DANLER, Paul (ed.). **Actes du XXV Congrès International de Philologie et de Linguistique Romanes**. Tome VII. Berlim: Mouton de Gruyter, 2010. p. 477-487.

RIBEIRO, Sílvia; RIO-TORTO, Graça. Composição. *In*: RIO-TORTO, Graça *et al.* **Gramática derivacional do Português**. 2. ed. Coimbra: Imprensa da Universidade de Coimbra, 2016. p. 461-520.

RIO-TORTO, Graça; RIBEIRO, Sílvia. Compounds in Portuguese. **Lingua e Linguaggio**, Bolonha, v. 8, n. 2, p. 269-288, 2009. Disponível em: https://ria.ua.pt/bitstream/10773/9056/1/riotorto&ribeiro_portuguese.pdf. Acesso em: 10 jul. 2019.

RIO-TORTO, Graça; RIBEIRO, Sílvia. Portuguese compounds. **Probus**, De Gruyter, v. 24, n. 1, p. 119-145, 2012.

RIO-TORTO, Graça. Nouns in apposition: Portuguese data. **Revista de Estudos Linguísticos da Universidade do Porto**, v. 8, p. 17-38, 2013.

RIO-TORTO, Graça. Prefixação. *In*: RIO-TORTO, Graça *et al.* **Gramática derivacional do Português**. 2. ed. Coimbra: Imprensa da Universidade de Coimbra, 2016. p. 411-459.

ROBERTS, Edward A. **A Comprehensive Etymological Dictionary of the Spanish Language with Families of Words Based on Indo-European Roots**. v. 1. Bloomington: Xlibris, 2014.

ROCHA LIMA, Carlos Henrique. **Gramática normativa da língua portuguesa**. 49. ed. Rio de Janeiro: José Olympio, 2011.

RUIZ MONEVA, Maria-Angeles. Compound nouns in the old english period: functional and pragmatic approaches in Ælfric's Lives of Saints: functional and pragmatic approaches. **Salim**, n. 8, p. 239-258, 1998.

SABINO, Marilei Amadeu. **Dicionário italiano-português de “falsos cognatos” e “cognatos enganosos”**. Subsídios teóricos e práticos ao ensino/aprendizagem de línguas, à lexicografia pedagógica e à tradução. São Paulo: Editora Unesp, 2011. Disponível em: <https://repositorio.unesp.br/bitstream/handle/11449/113689/ISBN8539302123.pdf?sequence=1&isAllowed=y>. Acesso em: 25 nov. 2021.

SACRAMENTO, Arivaldo; NASCIMENTO, Hérvickton Israel. Entre a Filologia e a Linguística Histórica: o texto como artefato histórico. **Macabéa – Revista Eletrônica do NETLLI**, Crato, v. 8, n. 2, 2019, p. 473-487. Disponível em: <http://periodicos.urca.br/ojs/index.php/MacREN/article/view/1968>. Acesso em: 23 ago. 2020.

SANDMANN, Antônio. Novidades do “front” de formação de palavras. **Revista Letras**, Curitiba, v. 36, p. 54-68, 1987. Disponível em: <https://revistas.ufpr.br/letras/article/view/19252> Acesso em: 20 jul. 2020.

SANTOS, Antonia Vieira dos. **Compostos sintagmáticos nominais VN, NN, NA, AN e NprepN no português arcaico (sécs. XIII-XVI)**. 2009. 2 v. 241 f. Tese (Doutorado em Letras) – Instituto de Letras – Universidade Federal da Bahia, Salvador, 2009. Disponível em: <https://repositorio.ufba.br/ri/handle/ri/11221>. Acesso em: 15 nov. 2021.

SANTOS, Antonia Vieira dos. Compostos NprepN no português arcaico: aspectos morfossintáticos. **Confluência**, n. 50, 1º semestre de 2016, p. 191-214.

SANTOS, Antonia Vieira dos. Compostos e mecanismos de composição de palavras no Orto do Esposo. *In*: Congresso Internacional da Abralín, 10., 2017. **Anais eletrônicos [...]** Niterói: Letras da UFF, 2017. p. 2313-2324. Disponível em: <https://www.abralin.org/site/wp-content/uploads/2019/06/4-7-PB-1.pdf>. Acesso em: 26 jul. 2020.

SCALISE, Sergio; FÁBREGAS, Antonio; FORZA, Francesca. Exocentricity in Compounding. **Genko Kenkyu**, v. 135, p. 49-84, 2009.

SCHLÜCKER, Barbara. Compounds and multi-word expressions in German. *In*: SCHLÜCKER, Barbara (ed.). **Complex lexical units: compounds and multiword expressions**. Berlin; Boston: De Gruyter Mouton, 2019. p. 69-94.

SCHLECHTWEG, Marcel. **Memorization and the Compound-Phrase Distinction: an investigation of complex constructions in German, French and English**. Berlin; Boston: De Gruyter Mouton, 2018.

SENTÍ, Andreu. Composició patrimonial i sintagmació. *In*: CABRÉ, M. T.; DOMÈNECH, Ona; ESTOPÀ, R. (ed.). **Mots nous em catalã: uma panorâmica geolectal**. Amsterdã, Filadélfia: John Benjamins Publishing Company, 2014. p. 85–114. Disponível em: <https://books.google.com.br/>. Acesso em: 23 mar. 2020.

SHIMADA, Masaharu. Coordinated compounds: Comparison between English and Japanese. **SKASE Journal of Theoretical Linguistics**, v. 10, n. 1, p. 77-96, 2013. Disponível em: http://www.skase.sk/Volumes/JTL22/pdf_doc/05.pdf. Acesso em: 26 jun. 2020.

SILVA, Alba Valéria Tinoco Alves. Metáfora e metonímia: o traço-de-união entre os compostos. **RBLA**, Belo Horizonte, v. 11, n. 1, p. 27-45, 2011. Disponível em: <https://www.scielo.br/>. Acesso: 20 jul. 2020.

SILVA, Neide Higino da. **Metáfora e metonímia nas construções com ‘pé’**: uma abordagem cognitivista. 2011. 111 f. Dissertação (Mestrado em Letras Vernáculas – Língua Portuguesa) – Faculdade de Letras – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011. Disponível em: <https://silo.tips/download/universidade-federal-do-rio-de-janeiro-faculdade-de-letas-metafora-e-metonimia>. Acesso em: 08 nov. 2021.

SILVA, Gabriel; DELICATO, Mônica Oréfica; SILVA, Nair; SOUSA, Jorge Pedro. Manuel Severim de Faria: vida de um noticiarista interventor. **RFCHS**, Porto, n. 3, 2006, p. 167-177. Disponível em: <https://bdigital.ufp.pt/handle/10284/599>. Acesso em: 15 ago. 2020.

SNYDER, William. 2001. On the nature of syntactic variation: Evidence from complex predicates and complex word-formation. **Language**, v. 77, n. 2, 2001, p. 324–342.

SZEMBERSKA, Anna. I composti nome-nome coordinati e subordinati nel lessico sportivo italiano come prova delle tendenze in atto nell’italiano contemporâneo. **Studia Romanica Posnaniensia**, v. 40, n. 3, Poznań, p. 95-103, 2013.

TRIEBIG, Victória; SANTOS, Antonia Vieira dos. Compostos [NN]_N no *Vocabulário Português, e Latino...* de Bluteau (séc. XVIII). *Miguilim - Revista Eletrônica do Netli, Crato*, v. 9, n. 3, p. 1054-1072, set.-dez. 2020. Disponível em: <http://periodicos.urca.br/ojs/index.php/MigREN/article/view/2586>. Acesso em: 29 set. 2021.

TUTIN, AGNÈS; GROSSMANN, Francis. Collocations régulières et irrégulières: esquisse de typologie du phénomène collocatif. **Revue française de linguistique appliquée**, v. VII, p. 7-25, 2002. Disponível em: <https://hal.archives-ouvertes.fr/hal-01336278/document>. Acesso em: 21 set. 2021.

VAL ÁLVARO, José. La composición. *In*: BOSQUE, Ignacio; DEMONTE, Violeta (dir.). **Gramática Descriptiva de la Lengua Española**. Vol. III: Entre la oración y el

discurso/Morfologia. Madrid: Editorial Espasa/Calpe, 1999. p. 4757-4842.

VASCONCELOS, Carolina Michaëlis de. **Lições de filologia portuguesa** (segundo as preleções feitas aos cursos de 1911/12 e de 1912/13) seguidas das lições práticas de português arcaico. Lisboa: Nova Edição da Revista de Portugal, 1946.

VIGÁRIO. *In*: Wikipédia: a enciclopédia livre. Disponível em: <https://pt.wikipedia.org/wiki/Vigário>. Acesso: 08 out. 2021.

VILLALVA, Alina. Formação de palavras: composição. *In*: MATEUS, Maria Helena Mira *et al.* **Gramática da Língua Portuguesa**. Lisboa: Editorial Caminho, 2003. p. 970-983.

VILLALVA, Alina; SILVESTRE, João Paulo. **Introdução ao estudo do léxico**: descrição e análise do português. Petrópolis: Vozes, 2014.

VILLOING, Florence. French compounds. **Probus**, De Gruyter, v. 24, n. 1, p. 29-60, 2012.

WREGGE, Rachel Silveira. A organização das escolas dos jesuítas no Brasil-Colônia: burocracia, estrangeirismo e competitividade. *In*: CONGRESSO INTERNACIONAL DE HISTÓRIA: “Entre Europa, África e América: o Império Português no Atlântico Sul”, 2., 2014, Goiânia. **Anais eletrônicos** [...] Goiânia: PUC-Goiás, 2014. p. 392-402. Disponível em: <https://congressointernacionaldehistoria.files.wordpress.com/2014/09/textos-completos4.pdf>. Acesso em: 08 out. 2021.

b) Referências dos textos do *corpus*

BROCHADO, José da Cunha. **Cartas** (selecção, prefácio e notas de António Álvaro Dória). Lisboa: Editora Livraria Sá da Costa, 1944. Disponível em: <http://www.tycho.iel.unicamp.br/corpus>. Último acesso em: 18 jan. 2022.

CÉU, Maria do. **Rellação da Vida e Morte da Serva de Deos a Venerável Madre Elenna da Crus** (transcrição do Códice 87 da Biblioteca Nacional precedida de um estudo histórico, por Filomena Belo). Lisboa: Quimera, 1993. Disponível em: <http://www.tycho.iel.unicamp.br/corpus>. Último acesso em: 18 jan. 2022.

COSTA, António da. **Cartas do Abade António da Costa** (introdução e notas de Fernando Lopes Graça). Lisboa: Cadernos da Seara Nova, 1946. Disponível em: <http://www.tycho.iel.unicamp.br/corpus>. Último acesso em: 18 jan. 2022.

COSTA, Manuel da. **Arte de Furtar** (selecção, cointrodução e notas de Roger Bismut). Lisboa: Imprensa Nacional-Casa da Moeda, 1991. Disponível em: <http://www.tycho.iel.unicamp.br/corpus>. Último acesso em: 18 jan. 2022.

FARIA, Manuel Severino de. **Discursos vários políticos** (introdução, actualização e notas de Maria Leonor Soares Albergaria Vieira). Lisboa: Imprensa Nacional-Casa da Moeda, 1999. Disponível em: <http://www.tycho.iel.unicamp.br/corpus>. Último acesso em: 18 jan. 2022.

LISBOA, João Luis; MIRANDA, Tiago dos Reis; OLIVAL, Fernando. **Gazetas manuscritas da Biblioteca Pública de Évora**. v. 1 (1729-1731). Lisboa: Edições Colibri, 2002. Disponível em: <http://www.tycho.iel.unicamp.br/corpus>. Último acesso em: 18 jan. 2022.

MARQUILHAS, Rita. **A faculdade das letras: Leitura e escrita em Portugal no século XVII**. 1996. Tese de Doutoramento apresentada à Universidade de Lisboa. Disponível em: <http://www.tycho.iel.unicamp.br/corpus>. Último acesso em: 18 jan. 2022.

MELO, Francisco Manuel de. **Obras Métricas de Don Francisco Manuel al Sereníssimo Señor Infante Don Pedro**. Léon: Horacio Boechat & George Remeus, 1665. Disponível em: <http://www.tycho.iel.unicamp.br/corpus>. Último acesso em: 18 jan. 2022.

SUTIL, Nuno Nisceno. **Musa Jocoza de varios entremezes portuguezes, & castelhanos**. Offerecidos ao excellentissimo senhor Joseph Miguel Joam de Portugal [,] Primogenito do Excellentissimo Senhor Conde de Vimioso, etc. e compostos por Nuno Nisceno Sutil. Lisboa: Oficina de Miguel Manescal, 1709. Disponível em: <http://www.tycho.iel.unicamp.br/corpus>. Último acesso em: 18 jan. 2022.

VIEIRA, António. **Cartas do Padre António Vieira**. (coordenadas e anotadas por J. Lúcio d'Azevedo). Tomo I. Coimbra: Imprensa da Universidade de Coimbra, 1925. Disponível em: <http://www.tycho.iel.unicamp.br/corpus>. Último acesso em: 18 jan. 2022.