



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
FACULDADE DE LETRAS

UM PANORAMA SOBRE A COMPOSIÇÃO DAS PALAVRAS:
as relações gramaticais internas aos compostos em Tupinambá e
Guarani Mbyá

Victoria Braz Souza

Rio de Janeiro
2021

VICTORIA BRAZ SOUZA

UM PANORAMA SOBRE A COMPOSIÇÃO DAS PALAVRAS:
as relações gramaticais internas aos compostos em Tupinambá e
Guarani Mbyá

Monografia submetida à Faculdade de Letras
da Universidade Federal do Rio de Janeiro,
como requisito parcial para obtenção do título
de Licenciado em Letras na habilitação
Português/Inglês.

Orientador: Prof^a Dr^a Marcia Maria Damaso Vieira

RIO DE JANEIRO

2021

Souza, Victoria Braz.

Um panorama sobre a composição de palavras: as relações gramaticais internas aos compostos em Tupinambá e Guarani Mbyá/Victoria Braz Souza.– 2021.

43f.

Orientador: Marcia Maria Damaso Vieira.

Monografia (graduação em Letras habilitação Português – Inglês) – Universidade Federal do Rio de Janeiro, Centro de Letras e Artes, Faculdade de Letras.

Bibliografia: f.95-97.

1. Linguística. 2. Morfologia Distribuída. I Souza/Victoria II - Universidade Federal do Rio de Janeiro, Faculdade de Letras, 2021 III. Título.

CDD

Índice Geral

1. Introdução.....	7
2. Sobre a composição de palavras.....	8
2.1 O que são palavras compostas?.....	8
2.2 Critérios para identificação de compostos.....	10
2.3 Diferenças entre palavras compostas e sintagmas.....	12
2.4 Os tipos de núcleos dos compostos.....	15
2.5 Relações gramaticais internas aos compostos.....	18
2.6 Compostos primários e sintéticos.....	21
3. Composição na Morfologia Distribuída.....	22
3.1 A proposta de Okubo.....	23
3.1.2. A proposta de Nóbrega (2015).....	25
4. As palavras compostas do Tupinambá e do Guarani Mbyá.....	27
4.1 Sobre as palavras em Tupinambá e Guarani Mbyá.....	27
4.1.1 Sobre os nomes.....	29
4.2 Os tipos de compostos.....	33
4.3 Um olhar interno aos compostos.....	39
5. Conclusões finais.....	41
6. Referências.....	42

RESUMO

O tema de investigação desta monografia é o processo de composição das palavras. O nosso objetivo aqui é investigar a proposta de Bisetto e Scalise (2005) sobre a universalidade das relações gramaticais entre os constituintes dos compostos, como subordinação, atribuição e coordenação. Com base na análise de dados de duas línguas da família Tupi-Guarani, o Tupinambá e o Guarani Mbyá, foi possível confirmar tal proposta. Além disso, os dados levantados forneceram evidências para a teoria da Morfologia Distribuída (OKUBO, 2013), segundo a qual a derivação das palavras compostas ocorre na sintaxe tanto no nível das raízes (*root domain*) quanto no nível das palavras (*outer domain*).

Palavras-chaves: **Composição de palavras; línguas da família Tupi-Guarani; relações gramaticais; Morfologia Distribuída.**

1. Introdução

Uma parte do trabalho aqui apresentado teve início durante o período de minha pesquisa de Iniciação Científica (08/2019 - 03/2021). O tema de investigação escolhido foi a composição de palavras.

A composição é um processo de formação de palavras que envolve muitos aspectos a serem discutidos, tais como: a sua definição; os critérios de identificação; as diferenças entre compostos e sintagmas; a natureza dos seus elementos constitutivos; os tipos de núcleos e os tipos de relações sintáticas, dentre outros.

O nosso objetivo com a escolha do tema foi investigar o que há de universal no processo de composição de palavras. Para tal, nos baseamos na proposta de Bisetto e Scalise (2005), que assume como universais as seguintes relações gramaticais entre os constituintes dos compostos: subordinação, atribuição e coordenação.

Com base nos dados das línguas Tupinambá e Guarani Mbyá (família Tupi-Guarani), foi possível não só confirmar a hipótese de Bisetto e Scalise (2005) sobre as relações gramaticais universais internas aos compostos, mas também a de Okubo (2013) sobre a existência de dois níveis sintáticos na derivação da composição.

Os dados aqui apresentados foram extraídos de fontes secundárias, como Lemos Barbosa (1956) e Rodrigues (1951) para o Tupinambá e Dooley (2013) para o Guarani Mbyá. Também alguns dos exemplos desta última língua foram coletados em trabalho de pesquisa de campo, realizado pela orientadora desta monografia.

O trabalho está estruturado da seguinte forma: a seção 2 apresenta um panorama geral sobre o processo de composição de palavras, tais como as suas definições e os seus tipos. Nesta seção está incluída ainda a proposta de Bisetto e Scalise (2005) sobre os tipos de relações gramaticais observados internamente nas palavras compostas; a seção 3 traz uma breve exposição sobre o tratamento da composição no quadro da Morfologia Distribuída, com base em Okubo (2013) e Nóbrega (2015); a seção 4 descreve alguns aspectos gramaticais do Tupinambá e do Guarani e exhibe os dados referentes aos compostos encontrados. Também discute brevemente a derivação de alguns desses compostos; a seção 5 conclui o trabalho.

2. Sobre a composição das palavras

No que diz respeito à definição de palavras compostas, podemos dizer simplesmente que são palavras formadas por dois ou mais itens lexicais e que expressam uma ideia única e autônoma, que pode ou não estar associada aos significados das partes envolvidas, como em *beija-flor* (um pássaro) e *navio-escola* (um tipo de navio que funciona como escola).

Ilustramos abaixo alguns tipos de compostos verificados em diferentes línguas naturais:

(1)

Português

a. saca-rolha

Inglês

b. *dog* “cachorro” + *house* “casa” → *doghouse* “casa de cachorro”

Tupinambá

c. *men* “marido” + *ub* “pai” → *menduba* “sogro”

Guarani

d. *eir* “mel” + *u* “pai” → *eiru* “abelha”

Nesta seção, comentaremos sobre alguns aspectos relacionados aos compostos, tais como: definições; critérios de identificação; diferenças entre eles e sintagmas; tipos de núcleos; tipos sintáticos (tipos de relações gramaticais) e a natureza morfológica dos seus constituintes.

2.1. O que são palavras compostas?

A questão da definição de palavras compostas, ou compostos, está longe de chegar a um consenso na literatura, levantando, inclusive, diversas discussões. Na busca por uma definição que abrangesse compostos de forma universal, vários autores deram as suas contribuições.

Segundo Figueiredo Silva (2011), a tradição gramatical no Brasil, representada pelos trabalhos estruturalista de Câmara Jr. (1978 apud FIGUEIREDO SILVA, 2011) e normativo de Cunha e Cintra (1985 apud FIGUEIREDO SILVA, 2011), define um composto como uma palavra única, ou seja, de significado único, cuja estrutura é formada por dois ou mais morfemas lexicais. Já Lieber (2009) postula que compostos são palavras constituídas por

duas ou mais bases, raízes ou estemas¹.

Em seu estudo sobre compostos, Altakhaineh (2016) reuniu variadas propostas para a definição desse mecanismo morfológico, conforme mencionado a seguir. Para Marchand (1960 apud ALTAKHAINEH, 2016) compostos consistem em duas ou mais palavras que são combinadas para formar uma unidade morfológica. Katamba (1993 apud ALTAKHAINEH, 2016) propõe que compostos são formados por no mínimo duas bases, que podem ser palavras ou raízes. Já segundo Fabb (1998 apud ALTAKHAINEH, 2016), um composto é uma palavra que consiste em duas ou mais palavras. Similar à proposta de Fabb, Olsen (2000 apud ALTAKHAINEH, 2016) sugere que um composto é uma combinação de duas formas livres para formar uma nova palavra complexa. Carstairs-McCarthy (2002 apud ALTAKHAINEH, 2016), por sua vez, postula que compostos são palavras formadas por raízes combinadas.

Como é possível observar acima, não há um consenso sobre quais são os elementos que constituem uma palavra composta. Seriam raízes/radicais e/ou palavras? Nenhuma das definições para compostos citadas é suficientemente satisfatória, pois, ou são muito amplas ou, então, são muito limitadas.

A definição da tradição gramatical esbarra na definição de expressão idiomática, porque descreve o composto como “um conjunto de palavras que transmite um único significado, no mais das vezes desligado do significado de suas partes” (FIGUEIREDO SILVA, 2011 p. 262). Desse modo, leva-se a crer que a diferença entre palavras compostas e expressões idiomáticas é semântica. Entretanto, palavras compostas podem tanto ter significados composicionais como não-composicionais; isto é, um significado que não se dá pela simples soma dos significados das suas partes. Este é o caso dos compostos abaixo:

(2)

Português

- a. puxa-saco “pessoa bajuladora”

Inglês

- b. *egg* “ovo” + *head* “cabeça” → *egghead* “intelectual”

Tupinambá

- c. *îuru* “boca” + *kyrá* “gorda” → *îuru-kyra* “mentiroso”

¹Segundo Bauer (1983, p. 20), estema é um dos termos utilizados para designar a parte da palavra que resta sem os afixos flexionais. Pode ou não ser complexo. Neste caso, tem afixos derivacionais ou mais de uma raiz aglutinados a ele.

Guarani

d. *pira* “peixe” + *pire* “pele” → *pirapire* “dinheiro”

Ademais, Altakhaineh, ao buscar uma definição universal para compostos, sugere a seguinte:

Um composto é uma palavra complexa que consiste de, pelo menos, dois elementos adjacentes, onde o elemento fora do núcleo normalmente é não-referencial. Cada um desses elementos é uma palavra, uma forma combinada ou um sintagma, de forma que o composto é uma combinação desses elementos (ALTAKHAINEH, 2016, p. 81, tradução própria.)²

Bisetto e Scalise (2005) observam que, nas definições dos compostos dadas na literatura, casos de composição presentes em línguas tipologicamente distintas são desconsiderados. Como há problemas com a definição de compostos relacionados à natureza de seus constituintes (raiz, radical e/ou palavra)³, fica difícil identificá-los nas línguas não indo-europeias.

Assim, para resolver esse impasse, alguns autores recorrem a um conjunto de critérios para a identificação de palavras compostas em uma dada língua, conforme aponta Nóbrega (2015).

2.2 Critérios para a identificação de compostos

Ao pensar em uma definição universal para compostos, com o intuito de diferenciá-los dos sintagmas, vários autores recorreram a uma lista de critérios. Entre os sugeridos na literatura, encontramos os de Donalies (2004 apud NÓBREGA, 2015, p. 90-91), em que está listada uma série de propriedades associadas às palavras compostas nas línguas germânicas, românicas, eslavas, fino-úgricas e no grego-moderno.

Nóbrega (2015) elenca, com base em Donalies, os critérios verificados na literatura para o reconhecimento de palavras compostas, como observado a seguir:

(3) **Critérios para identificação de compostos** (DONALIES, 2004 apud NÓBREGA, 2015)

- a. são complexos;
- b. são formados sem afixação;

²A *compound word* is a complex word that consists of at least two adjacent elements, where the non-head is normally non-referential. Each of these elements is either a word, combining form or a phrase, so that the whole compound is a combination of these elements.

³Raiz é o elemento irreduzível da palavra com conteúdo semântico. Como *cas-* em “casar”. Radical inclui a raiz e pode conter afixos derivacionais, como em *casament(o)*.

- c. são pronunciados juntos;
- d. têm um padrão acentual específico;
- e. contêm um elemento de ligação;
- f. têm núcleo à direita;
- g. são flexionados como um todo;
- h. são sintaticamente inseparáveis;
- i. são ilhas sintático-semânticas;
- j. são unidades conceituais.

Conforme propõe Nóbrega, tais critérios não se aplicam a todas as línguas. Os elementos de ligação, por exemplo, como “o” em *hidr-o-massagem*, podem estar ausentes. Outra crítica do autor é de que o padrão acentual não é uniforme nem mesmo em uma única língua. Em Português, compostos constituídos por radicais têm um único acento primário, como é o caso de *neurologia*. Já os compostos formados por palavras apresentam dois acentos, um primário e um secundário, assim como encontrado em *limpa-vidros*.

Além disso, a posição do núcleo do composto varia entre as línguas e até mesmo dentro da mesma língua. Nóbrega dá como exemplo o caso do Português em que o composto *sangue-frio* tem núcleo semântico à esquerda, ao passo que *vasoconstrição* tem núcleo semântico à direita. Sendo assim, esses critérios não podem ser tomados como universais, como sugere Donalies.

Lieber (2009), por outro lado, afirma que o teste mais confiável para identificar um composto é verificar a possibilidade ou não de se inserir um modificador na estrutura. Tal teste é desenvolvido também por Altakhaineh (2016) sob o nome de *adjacência*. No entanto, tais critérios não são universalmente comprovados, visto que a literatura apresenta diversos exemplos que configuram exceções. Alguns critérios são identificados em algumas línguas, enquanto outros não. Por isso, é importante não depender apenas desses critérios.

Para Nóbrega, o que pode ser considerado como universal no processo de composição são as relações gramaticais existentes entre os seus elementos constitutivos, conforme proposto por Bisetto e Scalise (2005)⁴ e verificado em vários outros trabalhos descritivos (CECCAGNO; BASCIANO, 2007, para o chinês e GÖKSEL, 2009, para o russo (apud NÓBREGA, 2015).

⁴Voltaremos a essa classificação em 2.5.

Tais relações gramaticais estão especificadas abaixo e são as mesmas observadas entre os constituintes dos sintagmas, o que mostra que a composição também envolve um processo sintático:

(4) Relações gramaticais internas à composição (NÓBREGA, 2015, p. 92)

- a. **Subordinação:** relação predicado-argumento (e.g., *lustra-móveis*);
- b. **Atribuição:** relação núcleo-modificador (e.g., *peixe-espada*);
- c. **Coordenação:** relação conjuntiva ou disjuntiva (e.g., *ator-diretor*).

Ainda segundo Nóbrega (p. 92): “Embora essas relações gramaticais sejam universais, o modo como as línguas naturais emolduram morfologicamente os seus compostos varia consideravelmente”. Isso quer dizer que os elementos constituintes dos compostos podem ser de diferentes naturezas: raízes ou palavras.

Na seção a seguir, comentamos brevemente sobre a diferença entre compostos e sintagmas.

2.3 Diferenças entre palavras compostas e sintagmas

Como visto na seção anterior, a distinção entre compostos e sintagmas ainda suscita debates na linguística. Ao longo do tempo, variados critérios foram sugeridos para diferenciá-los. Porém, poucos desses critérios se provaram altamente aplicáveis para esse propósito. Altakhaineh (2016) reuniu uma seleção de critérios para distinguir compostos de sintagmas. Em sua análise, focada em testes com compostos em inglês, o autor apontou a *adjacência* como o principal critério para diferenciar sintagmas de compostos. Isso significa que compostos são uma unidade fechada em si, ou seja, os seus elementos constituintes não podem ser separados. Não é possível inserir novos elementos em sua estrutura, sejam esses elementos flexão ou morfemas lexicais ou funcionais independentes. Essa restrição é vista na agramaticalidade dos exemplos abaixo:

(5) Português

- a. guarda-chuva/*guardas-chuva,
- b. porta-retrato/*porta-o-retrato

(6) Inglês

- a. *jack-in-the-box* “caixa-surpresa”/**it’s a Jack that’s in the box*

b. blackbird “pássaro preto/melro”/**Black uglybird*,

Os sintagmas, por outro lado, têm estruturas sintáticas mais livres e permitem movimentos e inserções entre os seus constituintes, como demonstra Figueiredo Silva (2011) em sua análise sobre expressões idiomáticas no Português Brasileiro (PB):

- (7) (a) Engolir sapo
 (b) Ela engoliu **um baita** sapo
 (c) **Aquele sapo** ela engoliu

Outro critério bastante relevante para a distinção entre compostos e sintagmas é a *referencialidade*, ou seja, a relação das palavras com o mundo. Em uma frase como “O gato bebeu todo o leite”/ “*The cat drank all the milk*”, *gato/cat* se refere a uma entidade específica no mundo. Altakhaineh (2016) observa que o primeiro elemento do composto à esquerda do núcleo costuma ser não-referencial, como em *catlover*, “apaixonado por gatos”, em que *cat* não se refere a nenhum gato específico. Outro fato é que, no inglês, os modificadores geralmente modificam apenas o núcleo do composto, como *these* que se refere a *accounts*, no exemplo a seguir:

- (8) *These bank accounts*
 “Estas contas bancárias”

Ainda em relação a essa questão, Bauer et al (2013, p. 464) observa que a não-referencialidade dos primeiros elementos se limita a compostos que sejam substantivos comuns. Em compostos onde o primeiro elemento é um nome próprio, como em *Beatles fan*, *Beatles* claramente se refere a uma entidade específica. No entanto, Bauer et al (2013, p. 464) apontou que a referencialidade do primeiro elemento do composto, isto é, do elemento não-nuclear, depende em grande parte do contexto. O primeiro elemento do composto pode servir como um antecedente discursivo para pronomes, como demonstra Bauer (1998, p. 72):

- (9) *You’re a cat lover. How many do you have now?*
 “Você é um apaixonado por gatos. Quantos você tem agora?”

Apesar disso, Altakhaineh (2016) ainda considera a *referencialidade* como um critério confiável para diferenciar compostos de sintagmas, pois os elementos não-nucleares de compostos em inglês costumam ser não-referenciais (Bauer et al, 2013, p. 464), assim como em árabe e hebraico.

Como critérios relativamente mais confiáveis para diferenciar compostos de sintagmas, Altakhaineh (2016) estabelece a seguinte hierarquia: adjacência > referencialidade > pluralização do elemento não-nuclear > composicionalidade > ênfase > coordenação > modificação > elipse > ortografia > substituição do segundo elemento por uma pro-forma.

As definições trazidas por Altakhaineh também não ajudam a distinguir palavras compostas de expressão idiomáticas, que por sua vez são sintagmas. Além disso, Altakhaineh aponta que elas falham em reconhecer a existência de compostos formados por sintagmas, como acontece no composto do inglês *jack-in-the-box*.

Quanto a essa distinção, ao invés de uma lista de critérios, Figueiredo Silva (2011) faz uso da proposta de Bisetto (1995 apud FIGUEIREDO E SILVA, 2011), que trabalha com compostos de forma aparente [V + N] e os divide em agentivo ou instrumental, locativo e eventivo. Sobre os compostos agentivos ou instrumentais, Bisetto nota que eles são os tipos mais produtivos em PB e os descreve como [Nome deverbal + N], pois, para a autora, o verbo é uma forma truncada do nominal agentivo ou instrumental terminado em - (d)or:

- (10) (a) abre-latas, abridor de latas
(b) puxa-saco, puxador de saco

Sendo assim, esses compostos passariam por um processo de incorporação, onde o sufixo nulo que nominaliza os verbos desencadeia a incorporação do objeto. Isso explicaria por que compostos não permitem a inserção de determinantes, como em (11) e (12), uma vez que isso impediria o movimento necessário para a incorporação do complemento (FIGUEIREDO SILVA, 2011, p. 266).

- (11) guarda-chuva/*guarda-esta-chuva
(12) porta-retrato/*porta-o-retrato

Os sintagmas, por outro lado, analisados na forma de expressões idiomáticas, sofrem processos de reanálise, mantendo ainda duas ou mais unidades morfológicas e o “congelamento” dessas unidades. Ao contrário dos compostos [Nome deverbal + N] analisados, as expressões idiomáticas são sintagmas verbais (vP), e, por isso, podem aceitar flexão verbal e permitir algumas manipulações sintáticas, como visto em (7), repetido abaixo em (13):

- (13) (a) Engolir sapo
 (b) Ela engoliu um baita sapo
 (c) Aquele sapo ela engoliu

Figueiredo Silva (2011) conclui, então, que o que diferencia um composto de um sintagma é o tipo de operação morfológica que sofrem: incorporação no primeiro caso e reanálise no segundo.

Após identificar as palavras compostas em uma dada língua, ainda é preciso agrupá-las em diferentes tipos relacionados à natureza de seus núcleos e aos tipos de relações gramaticais que exibem.

2.4 Os tipos de núcleos dos compostos

Os compostos podem ser classificados de acordo com o tipo de núcleo que possuem. É o núcleo que determina a categoria e o significado da composição. Em (14), o núcleo do composto “peixe-espada” é o nome “peixe”. Semanticamente, “peixe-espada” é um tipo de “peixe em forma de espada”, que atua sintaticamente como nome. Já em (15), “beija-flor”, palavra formada por um verbo e um nome, não há um núcleo semântico retirado de seus constituintes. O seu significado denota um tipo de “pássaro” e sintaticamente o composto se comporta como nome:

- (14) peixe_N-espada_N → N: “Um tipo de peixe”.
 (15) beija_V-flor_N → N: “Um tipo de pássaro”.

Segundo Teixeira (2009, p. 24): “quando um composto é considerado o hipônimo de seu núcleo, tem-se um composto endocêntrico.” Para Cruse (1986, p. 88 apud TEIXEIRA,

2009, p. 24), hiponímia é “a relação lexical que corresponde à inclusão de uma classe em outra”, sendo também chamada de relação “é um X” ou é “um tipo de X”. Trata-se de uma relação de classe e subclasse.

Veja-se o composto *desktop computer*. Como é possível dizer que *desktop computer* é um tipo de computador, esse composto pode ser considerado endocêntrico. A referência da palavra composta é a mesma que a de um de seus constituintes. Em alguns casos de compostos, não é possível estabelecer essa relação de classe e subclasse. Se a relação “é um” ou “é um tipo de” não pode ser identificada, a expressão composta é considerada exocêntrica, como em: *birdbrain*. A relação “é um” não se aplica nesse caso, pois não podemos afirmar que tal palavra se refere a um tipo de *bird* “pássaro” ou um tipo de *brain* “cérebro”; se chamamos alguém dessa forma, referimo-nos figurativamente a um tipo de pessoa, cujo tamanho do cérebro está sendo comparado com o de um pássaro. Trata-se de uma pessoa “estúpida/limitada mentalmente”.

Dependendo dos tipos de núcleos, os compostos podem ser divididos da seguinte forma⁵:

- (i) **Endocêntricos:** São compostos em que um dos constituintes é o núcleo semântico. O núcleo semântico se refere à mesma entidade denotada pelo composto. Também é o núcleo que determina a classe lexical da palavra.

Os núcleos dos compostos endocêntricos são responsáveis por determinar tanto a classe gramatical quanto o tipo semântico ao qual a unidade irá pertencer (LIEBER, 2009, p. 46). Tomemos como exemplos *ararinha-azul* e *doghouse* “casa de cachorro”, onde os respectivos núcleos são os nomes *ararinha* e *house* “casa”. Tanto *ararinha* quanto *house* são nomes, logo *ararinha-azul* e *doghouse* também são nomes. Da mesma forma, em termos semânticos, *ararinha-azul* é um tipo de arara e *doghouse* é um tipo de “casa”. Neste caso, podemos agrupar as seguintes palavras compostas como endocêntricas:

- (16) *peixe_N-espada_N* → N: “Um tipo de peixe”.
 (17) *balão_N-sonda_N* → N: “Um tipo de balão”.
 (18) *palavra_N-chave_N* → N: “Um tipo de palavra identificadora”.
 (19) *trem_N-balã_N* → N: “Um tipo de trem”.
 (20) *arm_N-chair_N* → N: “Um tipo de cadeira com braços”.

⁵Estamos tratando aqui apenas a parte semântica da palavra.

(21) *bear_N-skin_N* → N: “Um tipo de pele”.

(22) *book_N-shop_N* → N: “Um tipo de loja”.

(ii) **Exocêntricos:** São compostos em que nenhum dos constituintes é o núcleo semântico. A referência da palavra não é a mesma que a de um de seus constituintes. Trata-se de compostos não posicionais:

(23) *beija_V-flor_N* → N: “Um tipo de pássaro”.

(24) *saca_V-rolha_N* → N: “Instrumento para sacar rolhas”.

(25) *pão_N-duro_A* → N: “Pessoa avarenta”.

(26) *pick_Vpocket_N* → N: “Ladrão”.

(27) *cow_Nhand_N* → N: “Funcionário em fazendas responsável pelo manejo de gado”.

(28) *cut_Vthroat_N* → N: “Assassino”.

Resumindo, um composto endocêntrico tem, pelo menos, um núcleo que é ao mesmo tempo formal e semântico. Quando não há um núcleo semântico, o composto é exocêntrico.

Dentre os compostos exocêntricos, estão os chamados *Bahuvrīhi* (palavra do sânscrito que significa “tendo muito arroz”), ou compostos possessivos, em que a palavra formada denota o possuidor da especificação expressa. O composto do tipo *Bahuvrihi* se refere a uma característica da pessoa. Essa característica pode ser física ou mental, pode ser um trato profissional ou uma atitude:

(29) *colarinho_Nbranco_A* - “profissional assalariado”. Não é um tipo de colarinho ou de cor.

(30) *bird_Nbrain_N* - “pessoa estúpida”. Não é um tipo de pássaro ou de cérebro.

(31) *red_Ahead_N* - “pessoa de cabelo ruivo”. Não é um tipo de cabeça ou de cor.

(32) *white_A-collar_N* - “pessoa que trabalha em escritório”. Não é um tipo de colarinho ou de cor.

(33) *egg_Nhead_N* - “pessoa muita inteligente”. Não é um tipo de ovo ou de cabeça.

Conforme propõem Bisetto e Scalise (2005), as propriedades universais dos compostos são as relações gramaticais entre os seus constituintes, como as de subordinação, atribuição e coordenação, como mostramos a seguir.

2.5 As relações gramaticais internas ao composto

Bisetto e Scalise (2005), em sua proposta de classificação dos compostos, afirmam que as relações gramaticais entre os elementos constituintes de um composto são as mesmas que vemos entre os constituintes dos sintagmas. Essas relações podem ser descritas como de: subordinação, atribuição e coordenação.

A afirmação de Bisetto e Scalise sobre a universalidade dessas relações gramaticais foi corroborada por outros trabalhos descritivos, como Ceccagno; Basciano (2007) para o chinês, Rosenberg (2007) para o francês, Benigni; Masini (2009), para o russo, entre outros (apud NÓBREGA, 2015, p. 92).

De acordo com essa proposta de análise, os compostos podem ser agrupados em três tipos de relações sintáticas, a saber:

- (i) **Subordinação:** caracteriza a relação núcleo-argumento, verificada entre verbo e complemento e entre o elemento possuído e o possuidor, como indicam os dados a seguir:

(34) Português

- (a) puxa-saco (verbo + objeto)
- (b) saca-rolha (verbo + objeto)
- (c) língua-de-sogra (possuído + possuidor)
- (d) guarda-roupa (verbo + objeto)
- (e) paraquedas (verbo + objeto)
- (f) mão-de-vaca (possuído + possuidor)

Note-se que em Inglês, a ordem do núcleo e de seu complemento não é sempre a mesma verificada no nível do sintagma. Observa-se tanto VO quanto OV:

(35) Inglês

- (a) *heart* “coração” + *breaking* “partindo” → *heart-breaking* “desolador” (objeto + verbo)
- (b) *pick* “pegar” + *pocket* “bolso” → *pickpocket* “ladrão/batedor de carteira” (verbo +

objeto)

- (c) *taxi* “táxi” + *driver* “motorista” → *taxi driver* “motorista de táxi” (objeto + verbo nominalizado)
- (d) *cut* “cortar” + *throat* “garganta” → *cutthroat* “assassino” (verbo + objeto)
- (e) *kill* “matar” *joy* “alegria” → *killjoy* “desmancha-prazeres” (verbo + objeto)

- (ii) **Atribuição:** caracteriza a relação entre núcleo e modificador. Neste caso, um elemento (adjetivo ou nome) modifica o outro dentro do composto:

(36) Português

- (a) peixe-espada (nome + nome)
- (b) ararinha-azul (nome + adjetivo)
- (c) montanha-russa (nome + adjetivo)
- (d) palavra-chave (nome + nome)
- (e) pão-duro (nome + adjetivo)
- (f) obra-prima (nome + adjetivo)

(37) Inglês

- (a) *black* “negro” + *bird* “pássaro” → *blackbird* “pássaro preto/melro” (adjetivo + nome)
- (b) *blue* “azul” + *cheese* “queijo” → *blue cheese* “queijo Roquefort” (adjetivo + nome)
- (c) *green* “verde” + *eyed* “olho” → *green-eyed* “(pessoa de) olho verde” (adjetivo + nome)
- (d) *sword* “espada” + *fish* “peixe” → *sword-fish* “peixe-espada” (nome + nome)

- (iii) **Coordenação:** nos compostos coordenados, os elementos têm o mesmo status categorial. É como se estivessem ligados por um conectivo -“e”- abstrato. Em termos semânticos, é como se esse tipo de composto tivesse dois núcleos:

(38) Português

- (a) ator-diretor (nome + nome)

- (b) surdo-mudo (adjetivo + adjetivo)
- (c) sofá-cama (nome + nome)

(39) Inglês

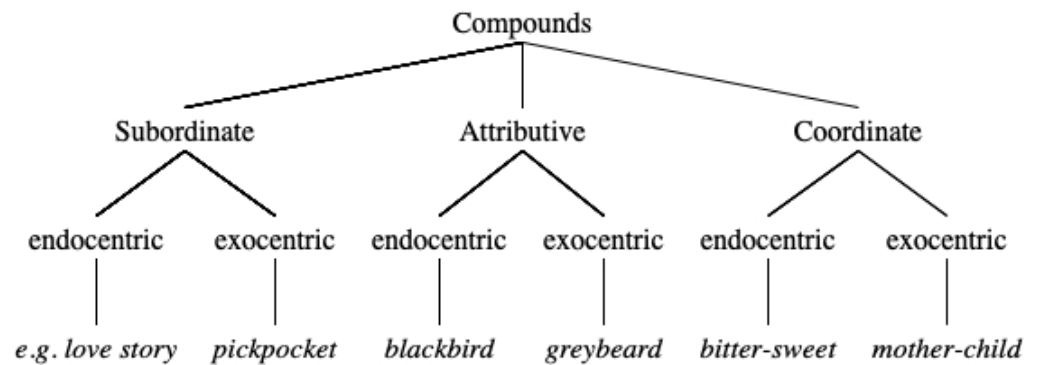
- (a) *goth* “gótico” + *folk* “folk” → *goth-folk* “gótico-folk” (nome + nome)
- (b) *parent* “pais” + *child* “criança” → *parent-child* “pais-filho (relação de)” (nome + nome)
- (c) *poet* “poeta” + *painter* “pintor” → *poetpainter* “poeta pintor” (nome + nome)

A partir dessas relações gramaticais, também é possível subdividir os compostos em *endocêntricos* e *exocêntricos*, de acordo com a referencialidade do núcleo e da unidade do composto. Um composto *endocêntrico* é aquele em que a referência do núcleo e do composto são as mesmas (LIEBER, 2009, p. 48). Já um composto *exocêntrico*, também conhecido como *bahuvrihi* por alguns autores (BAUER, 2011:700 apud BISETTO e SCALISE, 2005, p. 2), é aquele em que a referência da unidade do composto não é a mesma do núcleo (LIEBER, 2009, p. 48). O quadro (I) abaixo mostra a subdivisão dos compostos em tipos sintáticos e em tipos de núcleos. Essa mesma classificação é observada ainda na representação em (II) :

(I)

Tipos de composição	Endocêntrico	Exocêntrico
Subordinação	<i>Taxi-driver</i>	Conta-gotas
Atribuição	Tamanduá-bandeira	Dedo-duro
Coordenação	Surdo-mudo	<i>North east</i>

(II)



(LIEBER, 2009)

Além dos tipos de compostos mencionados acima, a literatura também apresenta uma divisão entre compostos primários e sintéticos, conforme se apresenta a seguir.

2.6 Compostos primários e sintéticos

Uma outra divisão comum para os compostos é entre primários e sintéticos, também conhecidos, respectivamente, como *root compounds* (“compostos de raízes”/“compostos primários”) e *synthetic compounds* (“compostos deverbais”) (BISETTO; SCALISE, 2005; LIEBER, 2009; NÓBREGA, 2015; SCHER, 2018).

Os compostos primários ou *root compounds* não envolvem um núcleo derivado de um verbo (um núcleo deverbal). Este é o caso de: *blackboard* “quadro-negro”, *bittersweet* “agridoce”, trem-bala, mesa-redonda, surdo-mudo. Eles são constituídos por dois lexemas/palavras, os quais podem ser nomes ou adjetivos. Diferentemente dos compostos sintéticos, nos compostos primários a relação entre o núcleo e o elemento não-núcleo costuma ser mais livre (LIEBER, 2009, p. 47). Os compostos do Inglês em (40) são compostos dos tipos primários:

(40) Inglês

- (a) *windmill* “moinho”
- (b) *egghead* “intelectual”
- (c) *housewife* “dona de casa”

Segundo Lieber (2009), os compostos sintéticos são aqueles em que o núcleo é derivado de um verbo, um núcleo deverbal, acompanhado de um argumento. Palavras

assim formadas envolvem tanto composição quanto derivação. De forma geral, de acordo com Scher (2018), os compostos sintéticos apresentam uma transparência de agentes e temas:

(41) Inglês

- (a) *truck driver* “motorista de caminhão”
- (b) *cat lover* “pessoa que gosta de gatos”
- (c) *babysitter* “babá”
- (d) *dishwasher* “lavadora de louça”

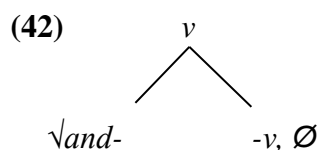
Na próxima seção, apresentamos uma breve exposição sobre o tratamento da composição no quadro da Morfologia Distribuída.

3. A composição na Morfologia Distribuída

A Morfologia Distribuída (MD) é um tipo de teoria que assume um único componente gerativo: a sintaxe. Assim, para a MD, as sentenças e as palavras são formadas na sintaxe, por meio das mesmas operações: *merge* e *move*.

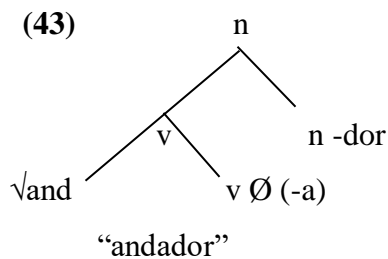
Segundo essa teoria, não há um léxico pré-estabelecido. As propriedades “lexicais” das palavras estão distribuídas em três listas em diferentes partes da gramática. A Lista 1 contém raízes acategoriais, além de morfemas funcionais abstratos. Esses elementos se combinam para formar palavras e são enviados para as Listas 2 e 3. A Lista 2 ou Vocabulário, localizada após o *spell out* sintático, é onde se dá a inserção do vocabulário. A Lista 3, ou Enciclopédia, contém os significados especiais das palavras que são geradas pela sintaxe.

As palavras são, então, derivadas através da concatenação de raízes com morfemas categorizadores. Assim, uma raiz se combina com um verbalizador (*v*) e se torna um verbo, se combina com um nominalizador (*n*) e se torna um nome e se combina com um adjetivizador (*a*) e se torna um adjetivo. Na representação em (42), por exemplo, a raiz *and-* se concatena ao verbalizador abstrato $-\emptyset$ e se converte no verbo *andar*:



“andar”

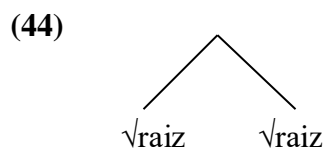
Depois de categorizada, a palavra pode ainda se recategorizar, mudando de classe, como mostra (43), em que o verbo “andar” se transforma em um nome, ao se concatenar ao sufixo nominalizador “(d)or”, formando a palavra “andador”:



3.1. A proposta de Okubo (2013)

Também no campo da Morfologia Distribuída, uma hipótese interessante para a derivação de compostos é a de Okubo (2013). Adaptando a proposta de formação de palavras de Marantz (2001 apud OKUBO, 2013), Okubo sugere que há dois tipos de processos para a formação de compostos: um é derivado no domínio da raiz e o outro, no domínio exterior.

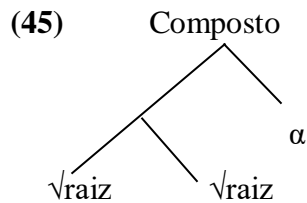
Quanto aos compostos derivados no domínio da raiz, o autor baseia-se na estrutura *root merger* sugerida por Zhang (2007), onde duas raízes são combinadas para formar o composto, como ilustra a representação a seguir:



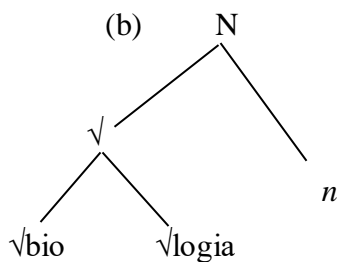
De acordo com Zhang (2007), a estrutura *root merger* é capaz de explicar a suposta exocentricidade⁶ dos compostos em chinês ao demonstrar que raízes não possuem elementos sintáticos. Na derivação, após a combinação das raízes, deve ocorrer uma categorização para

⁶Compostos sem um núcleo semântico nem sintático.

que o composto possa ser interpretado⁷. A representação em (46b) é a mesma para o composto do chinês *kai_v-xin_N* (abrir-coração), que deriva o adjetivo “feliz”. Deste modo, a estrutura resultante dessa combinação de raízes é exocêntrica, pois a categoria do composto é atribuída pelo categorizador α . (OKUBO, 2013, p. 149). Esta, então, seria a estrutura dos compostos derivados no domínio da raiz:

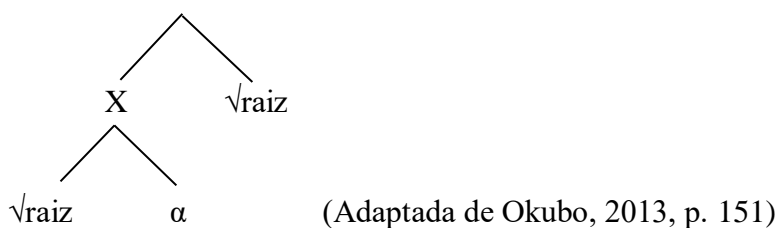


(46) (a) biologia



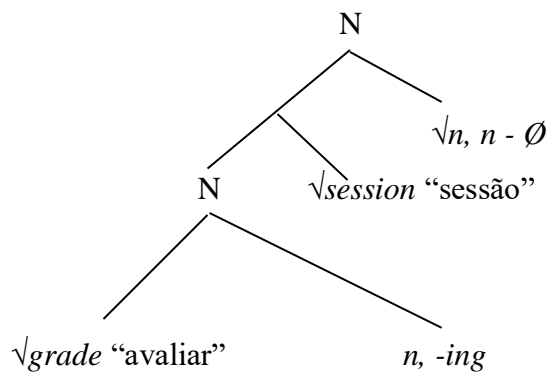
Já para os compostos derivados no domínio exterior, Okubo (2013) propõe que um elemento não-nuclear já categorizado é concatenado a uma raiz, a qual é o núcleo do composto. Em (48), o nome *grading* “avaliação” é inserido como complemento no domínio da raiz *session* “sessão”, que é categorizada pelo nominalizador abstrato. Neste caso, tem-se uma palavra no domínio da raiz:

(47) Composto derivado no domínio da raiz



⁷Em que α representa o categorizador. (Adaptada de Okubo, 2007, p. 149)

(48)

(a) *grading session* “sessão de avaliação”

Através da análise de Okubo (2013), se modifica a definição inicial de Harley (2009) sobre composição. A definição de Harley é a seguinte:

(49) Composto: uma unidade em tamanho de palavra contendo duas ou mais raízes.⁸
(Harley, 2009, p. 130, tradução própria)

Com a finalidade de dar conta de compostos que apresentam como um de seus constituintes uma raiz já categorizada, isto é, uma palavra, Okubo, sugere, então, a existência de dois tipos: um derivado em uma configuração como (45), onde as raízes podem se concatenar a categorizadores, e outro que pode ser derivado no domínio exterior, como ilustra (46). Assim, Okubo fornece a seguinte definição para os compostos:

(50) Compostos são unidades em tamanho de palavras contendo duas ou mais Raízes. As unidades são derivadas em domínios da raiz ou externos.⁹ (Okubo, 2013, p. 151, tradução própria).

⁸*Compound: a word-sized unit containing two or more Roots.*

⁹*Compounds are word-sized units containing two or more Roots. The units are derived in root or outer domains.*

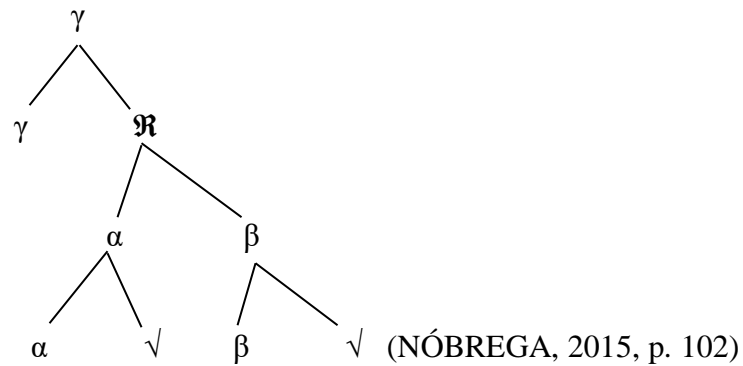
3.2. A proposta de Nóbrega (2015)

Para tratar da composição de palavras do tipo [N-N] no Português e no Grego, Nóbrega (2015, p. 88) assume os pressupostos teóricos da MD porque, segundo o autor, essa teoria: “consegue abarcar diretamente as propriedades combinatórias universais dos compostos, ou seja, derivar as variadas relações gramaticais entre seus membros constituintes (BISETTO; SCALISE, 2005; GUEVARA; SCALISE, 2009; SCALISE; BISETTO, 2009), e distribuir suas estruturas sintáticas nas variadas estruturas morfológicas atestadas empiricamente, a saber, as estruturas [radical + radical], [radical + palavra] e [palavra + palavra]”. Ainda de acordo com o autor: “...partimos do pressuposto de que as línguas naturais são uniformes do ponto de vista gerativo, e que sua variedade superficial se deve aos componentes morfológico e fonológico da gramática”.

Em outras palavras, apesar de essas relações gramaticais serem universais, pois são dadas pela operação de *Merger*, o modo como as línguas codificam morfológicamente os seus compostos está sujeito à variação mesmo dentro de uma única língua. Assim sendo, é possível encontrar combinações de raízes, radicais e palavras na geração dos compostos.

Recuperando a definição de Nóbrega (2015), um composto é formado por dois ou mais núcleos complexos em uma dada relação sintática, que são recategorizados por um núcleo definidor de categoria, podendo este ser *-n*, *v* ou *a*. De acordo com essa hipótese, núcleos complexos são equivalentes a uma raiz abstrata concatenada a um ou mais núcleos definidores de categoria. Na derivação dos compostos, os núcleos complexos são posteriormente concatenados em uma relação sintática \mathfrak{R} , podendo ser ela uma relação de subordinação, atribuição ou coordenação. Essa estrutura sintática será, então, categorizada por um outro núcleo definidor de categoria, criador de domínio e responsável por fornecer ao composto sua informação categorial, as informações flexionais e sua interpretação como uma única unidade sentencial no componente sintático. (NÓBREGA, 2015, p. 102). Nóbrega representa a sua proposta com a seguinte estrutura abstrata :

(51) (a)



(b) α , β e γ são os núcleos categorizadores e \mathfrak{R} , as relações sintáticas de coordenação, subordinação e atribuição. (NÓBREGA, 2015, p. 102).

Apesar de ser uma proposta altamente complexa para implementar, Nóbrega tenta explicar que as “relações gramaticais de subordinação, atribuição e coordenação são universais e decorrem do modo como a operação *Merge* combina objetos sintáticos” (NÓBREGA, 2015, p. 115). Além disso, a proposta também permite que a composição opere com raízes/radicais ou palavras.

Na próxima seção, apresentamos os dados de composição encontrados nas línguas Tupinambá e Guarani Mbyá.

4. As palavras compostas do Tupinambá e do Guarani Mbyá (família Tupi-Guarani)¹⁰

Através dos compostos verificados em Tupinambá e Guarani Mbyá, foi possível atestar as relações gramaticais estabelecidas na classificação de Bisetto e Scalise (2005) como universais. Como já visto anteriormente, essa classificação prevê a existência de três tipos de relações sintáticas entre os constituintes dos compostos: **subordinação, atribuição e coordenação**.

Antes de elencarmos os compostos encontrados para ilustrar esta pesquisa, apresentamos um panorama geral sobre alguns aspectos das gramáticas dessas duas línguas que serão relevantes para a compreensão dos dados analisados.

¹⁰Os dados do Tupinambá foram extraídos de Lemos Barbosa (1956). Os dados do Guarani Mbyá foram extraídos de Doodley (2013) e os que não apresentam referências foram coletados pela orientadora deste trabalho.

4.1 Sobre as palavras em Tupinambá e Guaraní Mbyá

O Tupinambá é uma língua da família Tupi-Guarani já extinta. Os seus registros, feitos pelos padres jesuítas, são datados dos séculos XVI e XVII. É uma língua classificada morfologicamente como incorporante, pois apresenta uma morfologia verbal rica onde podem figurar, além de elementos lexicais, elementos funcionais, tais como: concordância, negação, morfemas de mudança de valência e de aspecto. Em (52), o objeto aparece incorporado ao verbo¹¹:

Tupinambá

(52) *o-î-mo-pen*

3-3-CAUS-quebrar

”Ele o quebrou” (LEMOS BARBOSA, 1956, p. 128)

(53) *a-itá-potar*

1SG-pedra-querer

“Eu quero pedras” (LEMOS BARBOSA, 1956, p. 68)

O Guaraní Mbyá é um dialeto do Guaraní falado nos dias de hoje em vários estados do Brasil: Espírito Santo, Rio de Janeiro, São Paulo, Paraná, Santa Catarina e Rio Grande do Sul. Também é falado em outros países, como a Argentina e o Paraguai. É uma língua muito semelhante ao Tupinambá.

Em termos sintáticos, essas línguas são do tipo núcleo final (SOV e posposições), mas permitem ordem livre nas orações principais. Admitem também a ocorrência de sujeitos e objetos nulos. Atualmente a ordem oracional preferida pelos falantes mais jovens do Guaraní Mbyá é SVO:

Guaraní Mbyá

(54) *kunha ajaka o-japo* SOV

mulher cesta 3-fazer

“A mulher fez cesta”

¹¹Lista de abreviações usadas nos dados: CAUS = causativo; FUT = futuro; N = nominalizador; PL= plural; REL= relacional; SG = singular.

- (55) *Kunha o-japo ajaka* SVO
mulher 3-fazer cesta
“A mulher fez cesta”

No verbo, as marcas de pessoa são obrigatórias. Quanto ao tempo, somente o Guarani Mbyá desenvolveu um sufixo de futuro. O verbo sem esse sufixo é ambíguo entre uma leitura de passado e de presente:

Guarani Mbyá

- (56) *O-nha-ta*
3-correr-FUT
“Ele correrá”

- (57) *O-nha*
3-correr
“Ele corre/ ele correu”

4.1.1 Sobre os nomes

Os nomes devem sempre terminar em vogal e não há uma classe de elementos determinantes, como artigos e quantificadores. Apenas prefixos de posse são observados no nome:

Tupinambá

- (58) *Oré -pinda*
1PL-anzol
“Nosso anzol” (LEMOS BARBOSA, 1956, p. 52)

Guarani Mbyá

- (59) *i-xy*
3-mãe
“Mãe dele”

Nos sintagmas genitivos, observa-se a ordem: **possuidor-possuído**. Entre esses constituintes, ocorre uma marca relacional que pode se realizar como *R*- ou \emptyset :

Tupinambá

(60) *oré-R-uba*

1PL-REL-pai

“Nosso pai”

(LEMOS BARBOSA, 1956, p. 50)

(61) *gûyrá- \emptyset -pepó*

pássaro-REL-asa

“Asa do passarinho”

(LEMOS BARBOSA, 1956, p. 79)

(62) *pirá-R-esá*

peixe-REL-olho

“Olho de peixe”

(LEMOS BARBOSA, 1956, p. 405)

Guarani Mbyá

(63) *Xe-R-o*

1SG-REL-casa

“Minha casa”

(64) *Ara- \emptyset -kyxe*

Ara-REL-faca

“A faca da Ara”

Além de marcar a dependência entre possuído e possuidor, o morfema relacional é encontrado ainda entre o verbo e o seu argumento interno (= complemento):

Tupinambá

(65) *xe-R-ausub*

1SG-REL-amar

“Ele me ama”

(LEMOS BARBOSA, 1956, p. 130)

Guarani

- (66) *Xe-Ø-nup-a*
 1SG-REL-bater
 “Ele me bateu”

Em Tupinambá, as palavras podem terminar em vogal ou consoante. Em Guarani Mbyá, não há mais palavras com consoantes finais:

Tupinambá

- (67)
 (a) *pó* “mão” (LEMOS BARBOSA, 1956, p. 52)
 (b) *ok* “arrancar” (LEMOS BARBOSA, 1956, p. 115)

Em Tupinambá, quando raízes nominais e adjetivais terminam em consoante, elas devem receber o sufixo átono *-a*. Em (68), o sufixo *-a* é um categorizador, um nominalizador, que converte uma raiz em nome. Nessa língua, todos os nomes devem terminar em vogal:

Tupinambá

- (68) *ayr* “filho” → *ayr -a* “filho”

(LEMOS BARBOSA, 1956, p. 144)

O Guarani Mbyá perdeu as consoantes finais e não tem mais o sufixo nominalizador *-a* para categorizar os nomes. Quando um nome termina em *-a* nessa língua, a vogal já faz parte da raiz.

Os verbos podem ser recategorizados e virarem nomes. Nestes casos, em Tupinambá, pode-se utilizar um sufixo nominalizador, como *-(s)ara* ou mesmo um nominalizador abstrato, *Ø*:

Tupinambá

- (69) *îuká-sara*

matar-N
 “Matador” (LEMOS BARBOSA, 1956, p. 260)

(70) *por-ú-Ø*
 gente-comer-N
 “comedor de gente” (LEMOS BARBOSA, 1956, p. 262)
 “Canibal”

Em Guarani Mbyá, emprega-se o sufixo *-a* ou *Ø* para nominalizar um verbo:

Guarani Mbyá

(71) *karu-a*
 comer-N
 “Local de comer”

(72) *mitã juka-a*
 criança matar-N
 “Matador de criança”

Os modificadores do nome aparecem à sua direita. São adjetivos que quando terminam em consoante devem receber o sufixo *-a* em Tupinambá. Esse sufixo parece também exercer a função de adjetivizador:

Tupinambá

(73) *pirang* “vermelho” (raiz) → *pirang-a* “vermelho”
 (LEMOS BARBOSA, 1956, p. 47)

(74) *y piranga*
 rio vermelho
 (LEMOS BARBOSA, 1956, p. 58)

Guarani Mbyá

(75) *jaguakya*
 cachorro gordo
 “Cachorro gordo”

(76) *jagua piru*

cachorro magro

“Cachorro magro”

O morfema relacional também ocorre entre o nome e o adjetivo:

Guarani Mbyá

(77) *kamby-R-aku*

leite-REL-quente

“Leite quente”

4.2. Os tipos de compostos

Com base nos dados observados e na classificação de Bisetto e Scalise (2005), verifica-se a ocorrência dos seguintes tipos de compostos em Tupinambá e Guarani Mbyá:

- (i) **Compostos subordinados:** nesse tipo de composição, as relações existentes são entre possuidor e elemento possuído e entre o verbo e o seu complemento. Esses compostos podem ainda ser classificados como endocêntricos ou exocêntricos, de acordo com seus núcleos semânticos, conforme mostramos a seguir:

Exocêntricos:

Tupinambá

(78) *eir -Ø- ubá* N + N → N

mel-REL-pai

“Pai do mel” (lit.)

“Abelha”

(LEMOS BARBOSA, 1956, p. 399)

(79) *anhaga-R-atá* N + N → N

diabo-REL-fogo

“Fogo do diabo” (lit.)

“Inferno” (LEMOS BARBOSA, 1956, p. 383)

(80) *men(d)-Ø-uba* N + N → N

marido-REL-pai

“Pai do marido” (lit)

“Sogro” (LEMOS BARBOSA, 1956, p. 427)

(81) *men(d)-Ø-y* N + N → N

marido-REL-mãe

“Mãe do marido” (lit.) (LEMOS BARBOSA, 1956, p. 399)

“Sogra”

(82) *pirá-R-esá* N + N → A

peixe-REL-olho

“Olho de peixe”

“Desmaiado” (LEMOS BARBOSA, 1956, p. 405)

Guarani Mbyá

(83) *xy-R-akua* N + N → N

mãe-REL-chifre

“Chifre de mãe”

(DOOLEY, 2013, p. 85)

“Um tipo de rato”

(84) *xe-ramõi-Ø-tatu* N + N → N

1SG-avô-REL-tatu

“O tatu do meu avô”

(DOOLEY, 2013, p. 85)

“Um tipo de borboleta”

(85) *pira-Ø-pire* N + N → N

peixe-REL-pele

“Pele de peixe”

“Dinheiro”

- (86) *jaxy-Ø-tata* $N + N \rightarrow N$
 lua-REL-fogo
 “Fogo da lua”
 “**Estrela**”
- (87) *kuxa-R-akua* $N + N \rightarrow N$
 colher-REL-chifre (DOOLEY, 2013, p. 84)
 “Chifre da colher”
 “**Garfo**”
- (88) *angua-Ø-pu* $N + N \rightarrow N$
 pilão-REL-estalo (DOOLEY, 2013, p. 85)
 “Estalo do pilão”
 “**Tambor**”
- Endocêntricos:**
- Tupinambá**
- (89) *poro-ú-Ø* $N + V \rightarrow N$
 gente-comer-N
 “Comedor de gente”
 “**Canibal**” (LEMOS BARBOSA, 1956, p. 262)
- (90) *yby-ú-Ø* $N + V \rightarrow N$
 terra-comer-N
 “Comedor de terra”
 “**Um tipo de comedor**” (LEMOS BARBOSA, 1956, p. 262)
- Guarani**
- (91) *kamb-u* $N + V \rightarrow V$
 seio-comer/sugar
 “**Mamar**”

- (ii) **Compostos atributivos:** nesses tipos de composição, um constituinte do composto funciona como modificador do outro constituinte. Este modificador pode ser um nome ou um adjetivo:

Exocêntricos

Tupinambá

- (92) *pó-pindá* N + N → N
 mão-anzol (LEMOS BARBOSA, 1956, p. 406)
 “Ladrão”

- (93) *pó-etá* N + N → N
 mão-muitas
 “Muitas mãos” (LEMOS BARBOSA, 1956, p. 406)
 “Trabalhador”

- (94) *obá-îubá* N + A → N
 rosto-amarelo (LEMOS BARBOSA, 1956, p. 405)
 “Medroso”

- (95) *íta-aîuba* N + A → N
 pedra-amarela (LEMOS BARBOSA, 1956, p. 384)
 “Ouro/dinheiro”

- (96) *îuru-kyra* N + A → N
 boca-gorda (LEMOS BARBOSA, 1956, p. 406)
 “Mentiroso”

- (97) *esá-kaneõ* N + A → N
 olho-cansado (LEMOS BARBOSA, 1956, p. 404)
 “Preocupado”

Guarani

- (105) *tatu-po-ju* N+ N + A → N
 tatu-mão-amarela (DOOLEY, 1956, p. 84)
“Tatu paulistinha (um tipo de tatu)”

- (iii) **Compostos coordenados:** os constituintes destes compostos estão em uma relação de aposição. Podem ter dois núcleos semânticos ou nenhum.

Endocêntricos

Tupinambá

- (106) *abá-s-oó* N + N → N
 homem-3-bicho
“Homem-bicho” (LEMOS BARBOSA, 1956, p. 401)
- (107) *Tupã-T-ayra* N+ N → N
 deus-3-filho
“Deus-Filho” (LEMOS BARBOSA, 1956, p. 403)

Viu-se, através dos dados apresentados acima, que o Tupinambá e o Guaraní possuem em seus compostos as relações gramaticais sugeridas por Bisetto e Scalise (2005) como universais. Observou-se que os tipos mais comuns de compostos dessas línguas são os subordinados e os atributivos exocêntricos. Não foi possível identificar exemplos de compostos coordenados exocêntricos, provavelmente por falta de um conhecimento aprofundado das línguas. O quadro a seguir resume os tipos de compostos presentes no Tupinambá:

(III)

Relações gramaticais	Endocêntricos	Exocêntricos
Subordinação	<i>poro-ú-Ø</i> Gente-comer-N “Comedor de gente” “Canibal”	<i>eir -Ø- ubá</i> Mel-REL-pai “Pai do mel” “Abelha”

Atributiva	<i>guyrá-iaguára</i> Pássaro-jaguar “Pássaro-jaguar”	. <i>îuru -kyra</i> Boca -gorda “Mentiroso”
Coordenação	<i>Tupã-T-ayra</i> Deus-3-filho “Deus-Filho”	

4.3 Um olhar interno aos compostos

Tanto em Tupinambá quanto em Guaraní, observa-se a ocorrência dos dois tipos de compostos sugeridos por Okubo (2013): um derivado no domínio da raiz, e outro, constituído de uma palavra já categorizada e uma raiz, derivado no domínio exterior.

No primeiro caso, em que a composição envolve raízes lexicais, é observada uma mudança morfofonêmica com as consoantes nasais *m* e *n* em fronteira de morfema. Este é o caso do Tupinambá, como se observa em (108) onde a consoante *n* se converte em *nd*. A representação em (109) ilustra a derivação de (108b):

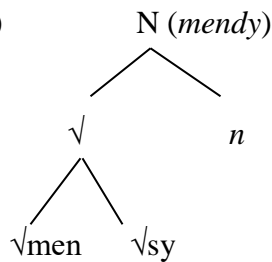
Tupinambá

(108)

(a) *men* “marido” + *ub* “pai” → *menduba* “sogro”

(b) *men* “marido” + *sy* “mãe” → *mendy* “sogra” (LEMOS BARBOSA, 1956, p. 44)

(109)



Vale salientar que a morfologia interna às palavras pode ajudar na identificação dos tipos de compostos. *Mendy* é um composto envolvendo raízes. Compare (108b) com (110), em que se tem um sintagma nominal com o significado “mãe do marido”. Neste caso, tem-se duas palavras independentes e o sentido é literal:

Tupinambá

- (110) *mena-Ø-sy*
 marido-R-mãe (LEMOS BARBOSA, 1956, p. 403)
 “Mãe do marido (de alguém)”

O mesmo padrão é verificado em compostos do Guarani Mbyá. Compare os pares abaixo:

Guarani Mbyá

- (111)
- (a) *tatá* “fogo” + *endy* “brilhante” → *tataendy*
 “lâmpião”
- (b) *tatá-r-endy*
 fogo-REL-brilhante
 “fogo brilhante”
- (112)
- (a) *Y* “água” + *aku* “quente” → *yaku*
 “café”
- (b) *y-r-aku*
 água-REL-quente
 “água quente”

Os dados em (a) são casos de composição de raízes. Nota-se que os significados dos compostos são distintos dos significados dos sintagmas em (b), em que ocorrem os prefixos relacionais. Em (a), tem-se composição, ao passo que em (b) tem-se sintagmas nominais.

Alguns dos compostos com subordinação e com atribuição permitem a ocorrência do morfema relacional (**R~Ø**) que marca a relação entre o núcleo e o seu complemento ou entre o possuído e possuidor e até entre o nome e o seu modificador (o adjetivo). Estes são os casos de (113), (114) e (115). Tem-se aqui compostos subordinados exocêntricos que envolvem uma raiz já categorizada, uma palavra:

Tupinambá

- (113) *pirá-r-esá* N + N → A
 peixe-REL-olho

Olho de peixe (LEMOS BARBOSA, 1956, p. 405)

“Desmaiado”

- (114) *anhaga-r-atá* $N + N \rightarrow A$
 diabo-REL-fogo
 “Fogo do diabo” (LEMOS BARBOSA, 1956, p. 383)
“Inferno”

Guarani Mbyá

- (115) *xy-R-akua* $N + N \rightarrow N$
 mãe-REL-chifre
 “Chifre de mãe” (DOOLEY, 2013, p. 85)
“Um tipo de rato”

A partir da observação dos dados do Guarani Mbyá e do Tupinambá, pode-se concluir que a proposta de Okubo (2013), que assume que há dois domínios na derivação dos compostos, está no caminho certo: os derivados no domínio da raiz, constituídos apenas de raízes não categorizadas, e os derivados no domínio exterior, constituídos de palavras.

5. Conclusões finais

Com base nos dados levantados para esta pesquisa, foi possível verificar e confirmar a proposta de Bisetto e Scalise (2005) sobre a existência das relações gramaticais de subordinação, atribuição e coordenação internas às palavras compostas. Tratam-se de relações sintáticas universais porque as palavras são derivadas na sintaxe pelas mesmas operações - *Merge e Move* - que formam os sintagmas. Dessa maneira, os seus constituintes podem exibir as mesmas relações gramaticais que os sintagmas.

Também foi possível confirmar a hipótese de Okubo (2013) sobre a existência de dois níveis de derivação de compostos: um envolvendo o domínio da raiz e outro, o domínio exterior que permite a ocorrência de palavras já categorizadas.

Através deste trabalho, esperamos ter podido contribuir empiricamente não só para as teorias adotadas, mas também para a área de estudos das línguas indígenas.

REFERÊNCIAS

- ALTAKHAINEH, Abdel Rahman Mitib. What is a Compound? The main criteria for compoundhood. **Explorations in English Language and Linguistics**, v.4. n.1, 201, 2016, p. 58-86.
- BAUER, Laurie. **English Word-formation**. Cambridge: Cambridge University Press, 1983.
- BAUER, L. When is a sequence of two nouns a compound in English? **English Language and Linguistics** 2, p. 65-86, 1998a
- BAUER, L. LIEBER, Rochelle Ingo Plag. **The Oxford Reference Guide to English Morphology**. Oxford: Oxford University Press, 2013
- BISETTO, Antonietta; SCALISE, Sergio. The Classification of Compounds. **Lingue e Linguaggio**, v. 4, n. 2, 2005, p. 319-332.
- DOOLEY, Robert. A. **Léxico Guarani, Dialeto Mbyá: Guarani-Português**. Brasília: SIL, 2013a
- _____. **Léxico Guarani, Dialeto Mbyá**. Introdução: informações gerais, esboço gramatical e referências. Brasília: SIL, 2013b.
- FIGUEIREDO SILVA, Maria Cristina. Compostos e Expressões Idiomáticas no Português Brasileiro. **Revista da ABRALIN**, v. Eletrônico, n. Especial, p. 261-277. 2ª parte, 2011
- HARLEY, Heidi. Compounding in Distributed Morphology. In LIEBER, Rochelle; STEKAUER, Pavol. (Eds.). **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009.
- LEMON BARBOSA, A. **Curso de Tupi Antigo: gramática, exercícios, textos**. Rio de Janeiro: Livraria São José, 1956.
- LIEBER, Rochelle. **Compounding. Introducing Morphology**. Nova York: Cambridge University Press, 2009. p.43-53.
- NÓBREGA, Vitor Augusto. Composição na Morfologia Distribuída: dos universais à variação. **ReVEL**, v. 13, n. 24, 2015.
- OKUBO, Tatsuhiro. Two Types of Compounds in Distributed Morphology. **Tsukuba English Studies**, v.32, 2013, p. 147-148.
- RODRIGUES, Aryon. A composição em Tupi. **Logos**, ano VI, n. 14, Curitiba, Separata, 1951.
- t, Ana Paula. A morfossintaxe de compostos X-V em inglês. **D.E.L.T.A.**, v. 34.2, 2018, p. 709-743

TEIXEIRA, Lílian Figueiró. **A semântica dos compostos nominais – um estudo de corpus paralelo inglês/português**. Dissertação (mestrado) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Linguística Aplicada, São Leopoldo, 2009.

ZHANG, Nina. Root Merger in Chinese Compounds. **Studia Linguistica**, v. 61, 2007, p. 170-184.