

A relação entre raízes, gênero, classe e significado

Maurício RESENDE¹

Beatriz Pires SANTANA²

¹ Universidade Estadual de Campinas (UNICAMP), mauri_cio_resende@hotmail.com, orcid.org/0000-0001-7487-5043, Papéis: conceptualização, análise formal, investigação, escrita – rascunho original, escrita – análise e edição.

² Universidade Federal do Paraná (UFPR), beatrizpiressantana@gmail.com, orcid.org/0000-0001-8976-3001, Papéis: conceptualização, análise formal, investigação, escrita – rascunho original, escrita – análise e edição.

RESUMO O presente trabalho analisa, à luz do quadro da Morfologia Distribuída, o comportamento das vogais temáticas verbais e das vogais temáticas nominais do português brasileiro e sua relação com as raízes a que se anexam. Perseguimos uma análise que busca unificar o comportamento desses dois tipos temáticos, mostrando que o que obscurece tal unificação é a interação das vogais nominais com o valor de gênero. Uma vez que se abstraem os efeitos causados pela manipulação de gênero, vemos que as vogais temáticas verbais e nominais mantêm um mesmo grau de relação com as raízes e que ambas são computadas da mesma forma. Neste artigo, nós mostramos que é possível também a unificação daquilo que se chama de “gênero semântico” (interpretável) e “gênero gramatical” (não interpretável).

ABSTRACT The present paper analyses, in the light of Distributed Morphology framework, the behavior of both verbal and nominal theme vowels from Brazilian Portuguese and their relation to the roots they merge with. We propose treatment which unifies the behavior of these two theme units, by showing that what makes this unification less clear is the interaction between nominal theme vowels and gender values. By abstracting from the effects of gender manipulation, we see that both verbal and nominal theme vowels keep

the same kind of relation with the roots and both are captured by the same mechanism. We also show in this paper that it is possible to unify what is called “semantic gender” (interpretable) and “grammatical gender” (non-interpretable).

PALAVRAS-CHAVE Vogais temáticas.
Gênero. Morfologia Distribuída.

KEYWORDS Theme vowels. Gender.
Distributed Morphology.

Introdução

O presente trabalho investiga a relação entre as raízes da língua e as vogais temáticas que as caracterizam em contextos nominais e verbais. Ao analisar uma raiz que gera um verbo e sua relação com a classe verbal (conjugação) a que pertence, vemos uma relação bastante estreita, em que uma raiz pode se combinar com apenas uma das três vogais temáticas verbais disponíveis. Quando aquilo que aparenta ser uma mesma raiz se anexa a duas vogais temáticas verbais diferentes, o resultado é uma mudança drástica de significado da base, o que parece mostrar que estamos perante duas raízes homófonas diferentes (cf., e.g., *podar-poder*). Por outro lado, ao analisar uma raiz que forma um nome e sua relação com a classe nominal a que pertence, a relação é bem menos estreita. Pares de substantivos como *menino-menina* mostra que uma mesma raiz pode se ligar a duas vogais finais diferentes, no âmbito da flexão nominal de gênero. Mesmo, no entanto, quando não se está perante o que se costuma tratar como flexão de gênero (que envolve o sexo dos referentes), como no par *barco-barca*, uma mesma (forma de) raiz ainda assim pode aparecer no contexto de duas vogais nominais, diferentes mantendo-se um significado em comum, o que nos faz supor que a identidade morfológica da raiz é, de fato, a mesma. No presente trabalho, nós intentamos mostrar que, em realidade, os dois tipos de vogais temáticas mantêm uma relação igualmente estreita com a raiz, e o que faz os nomes superficialmente se mostrarem mais propensos a combinar uma mesma raiz com diferentes vogais finais decorre do fato de a informação de classe interagir com a propriedade *gênero* nos nomes, mas não nos verbos. Mais particularmente, investigamos a hipótese de que uma raiz pode ser inerentemente associada a *no máximo* uma vogal temática verbal e a *no máximo* uma vogal temática nominal. Como o gênero não é uma propriedade dos verbos, a relação entre a raiz e a vogal temática verbal é superficialmente cristalina, mas, no caso dos nomes, a possibilidade de uma raiz ser associada a

valores de gênero diferentes obscurece a relação estreita entre raiz e informação de classe. Uma vez que nós entendemos o papel da valoração de gênero na mediação entre a raiz e sua vogal nominal final, vemos que as raízes se comportam da mesma forma com relação a vogais temáticas nominais e vogais temáticas verbais. Isso nos permite unificar, de um lado, informação de classe (nominal e verbal) e, de outro, gênero interpretável e não interpretável.

Para investigar os fenômenos supracitados, este artigo está dividido da seguinte maneira: na seção 1, são apresentados os principais pressupostos da Morfologia Distribuída, enfatizando, em § 1.1, a discussão sobre a acategorialidade das raízes e sobre de que maneira as informações de classe estão codificadas na gramática; em § 1.2, as questões que subjazem às classes verbais e nominais e a sua relação com as raízes; em § 1.3, o fenômeno da marcação de gênero e as propostas encontradas na literatura a respeito de como implementá-lo, sobretudo ao se levar em conta a questão da sua interpretabilidade.

Na seção 2, são apresentadas e discutidas as diferentes propostas para o fenômeno das vogais temáticas e da expressão de gênero para o português dentro da Morfologia Distribuída – mais especificamente Alcântara (2003, 2010) e Armelin (2015) – apontando vantagens e problemas dessas análises. Na seção 3, nós apresentamos a nossa proposta, em § 3.1, para a relação entre as raízes e a classe verbal e, em § 3.2, para a relação entre as raízes, o gênero (interpretável e não interpretável) e a classe nominal, focando nas propriedades morfofonológicas.

Finalmente, na seção 4, nós mostramos a nossa análise para a interpretação das raízes nos ambientes sintáticos em que elas podem ocorrer e, posteriormente, discutimos a questão da interpretabilidade de gênero e de que maneira ela aparece codificada na nossa proposta; por fim, nós abordamos os casos de coerção (alguns deles já discutidos na literatura) e mostramos como eles podem ser capturados de acordo com a proposta defendida neste trabalho.

1. Antecedentes teóricos

1.1. O modelo da Morfologia Distribuída

Diferentemente de modelos lexicalistas baseados em palavras, isto é, de modelos que assumem a existência de um componente gerativo pré-sintático – um Léxico – formador dos itens lexicais que já entram na derivação com suas propriedades (fonológicas, morfológicas, sintáticas e semânticas) determinadas, a Morfologia Distribuída (MD) defende que o único componente gerativo da gramática é a sintaxe, e que tanto palavras quanto sintagmas/sentenças são formados pelo mesmo conjunto de operações (concatenação e movimento).

Assim, dentro do quadro da MD, as informações fonológicas, sintáticas e semânticas dos então chamados “itens lexicais” estão distribuídas em três listas, que são acessadas em diferentes momentos da derivação; mais especificamente, conforme Marantz (1997), o *Léxico estrito* (a Lista 1) alimenta a sintaxe com raízes e feixes de traços morfossintáticos/semânticos abstratos, isto é, desprovidos de conteúdo fonológico e de conteúdo semântico não composicional. Essa primeira lista alimenta o sistema computacional (a sintaxe), e este, por meio de concatenação e movimento, gera todas as estruturas – palavras, sintagmas e sentenças.

Nessa perspectiva, mesmo as estruturas morfológicas mais simples – ou “morfologicamente básicas” em termos lexicalistas – são o resultado da concatenação de uma raiz acategorial com um núcleo categorizador. Posteriormente, as estruturas geradas pela sintaxe são simultaneamente enviadas para PF e LF. Em PF, a *estrutura morfológica* (MS) realiza operações adicionais (fusão, fissão, empobrecimento, inserção de morfemas dissociados), atendendo, assim, aos requerimentos específicos de boa formação morfológica das línguas.

Em seguida, uma operação em MS denominada *inserção de Vocabulário* atribui às estruturas geradas pelo sistema computacional a sua informação fonológica, a qual aparece listada no *Vocabulário* (a Lista 2) juntamente com a informação contextual para a sua inserção. Simultaneamente, na ramificação em LF, a *Enciclopédia* (a Lista 3) fornece às estruturas o seu conteúdo semântico não composicional, extralinguístico, por meio de instruções contextuais para a sua interpretação.

Portanto, diferentemente de modelos lexicalistas nos quais os itens lexicais (as palavras) são os átomos da derivação sintática, na MD os primitivos sintáticos são raízes (morfemas lexicais) desprovidas de categoria e feixes de traços gramaticais abstratos (morfemas funcionais). Assim, a derivação sintática articula desde morfemas até sentenças – propriedade que ficou conhecida como *estrutura sintática hierárquica por toda derivação*, conforme Halle & Marantz (1994).

Se, por um lado, a presença de traços gramaticais na derivação já era moeda corrente nas teorias sintáticas e morfológicas, por outro, a nova articulação dos primitivos lexicais em termos de raízes acategoriais abriu um novo leque de investigações teóricas. Recentemente, por exemplo, há muita discussão na literatura a respeito da identidade das raízes, isto é, de que modo as raízes (unidades desprovidas de categoria) são identificadas na Lista 1. Na versão clássica do modelo – isto é, Halle & Marantz (1993, 1994) e Marantz (1997) as raízes apresentam tanto conteúdo semântico mínimo quanto forma fonológica, cabendo a inserção tardia somente aos itens funcionais. Contudo, ambos os pressupostos têm sido questionados por alguns teóricos do modelo³.

Discutir questões empíricas e teóricas subjacentes a cada posição iria muito além do escopo do presente trabalho; por essa razão, nós vamos simplesmente assumir a inserção tardia das raízes, isto é,

³ Cf. Harley (2014) e trabalhos relacionados.

assumir que as raízes – assim como os itens funcionais – adentram o sistema computacional desprovidos de conteúdo fonológico. Além disso, este trabalho parte do princípio de que as raízes possuem um conteúdo semântico, o qual permite identificá-las na Lista 1. Mais especificamente, no que diz respeito às consequências dessa postulação para a presente análise, assumir que as raízes contêm um conteúdo semântico mínimo levaria à conclusão de que o que permite ao falante identificar a diferença, por exemplo, entre $\sqrt{\text{LEIT}}$ DE *leito* e $\sqrt{\text{LEIT}}$ DE *leite* é a sua semântica, e não sua forma fonológica, suas propriedades morfológicas (como informação de gênero) e, tampouco, informações estruturais, uma vez que essas duas raízes podem figurar em um mesmo contexto sintático. O mesmo vale para casos como $\sqrt{\text{PONT}}$ DE *ponte* e $\sqrt{\text{PONT}}$ DE *ponta*.

Sob esse mesmo critério, pares como *barco/barca* e *caneco/caneca* derivam da mesma raiz ($\sqrt{\text{BARC}}$ e $\sqrt{\text{CANEC}}$ respectivamente), uma vez que apresentam um significado relacionado, ou seja, possuem uma semântica mínima comum. Uma das principais críticas à identificação das raízes por meio de uma semântica inerente é apresentada em Harley (2014); para a autora, o fato de existirem raízes cujo significado depende crucialmente do contexto (morfo) sintático em que elas aparecem é uma evidência contra a semântica (pré-sintática) das raízes; exemplos desse caso seriam $\sqrt{\text{MIT}}$, como em *demitir*, *remitir*, *permitir*, *omitir* e $\sqrt{\text{CEB}}$, visto em *conceber*, *perceber*, *receber*.

Tais casos, que ficaram conhecidos na literatura como instâncias do “morfema *cranberry*”, imporiam uma dificuldade na determinação do significado da raiz, já que ele dependeria do ambiente estrutural em que a raiz ocorre; entretanto, Minussi & Bassani (2017) e Resende (no prelo), na tentativa de enfraquecer os argumentos de Harley, defendem que é possível identificar um conteúdo semântico comum às raízes, já que mesmo para casos como $\sqrt{\text{GRED}}$ (*agredir*, *progredir*,

transgredir) e $\sqrt{\text{TRA}}$ (*distrair, extrair, retrair*, os verbos derivados dessas raízes envolvem um significado relacionável, por exemplo, ligado conceitualmente à ideia de movimento (o que, em última análise, atribuiria aos verbos derivados uma semântica composicional prefixo + raiz).

Para Harley, a não identidade semântica desses casos seria compensada pela identidade morfofonológica, observada na alternância sistemática entre /mit/~ /mis/, vista em *remitir~remissão, demitir~demissão, permitir~permissão*. Por esse motivo, a autora considera que $\sqrt{\text{MIT}}$ é uma única raiz e que, portanto, deve-se abrir mão do critério semântico para a identificação de uma raiz; todavia, segundo Minussi & Bassani (2017), pares como *perceptível/percebível* e *aberto/abrido* mostram que não há correspondência necessária entre identidade formal e semântica; o mesmo é mostrado por Resende (2018) para pares como *cabelo/capilar, mão/manual, fiel/fidelidade*.

Assim, o que esses dados sugerem é que o que permite relacionar essas palavras derivadas é a semântica mínima comum de suas raízes e não a forma fonológica da raiz, que pode ser determinada contextualmente. Evidência adicional é fornecida por Lemle (2008), para quem as raízes de pares como *restaurar/restaurante* e *apartar/apartamento*, embora apresentem a mesma forma fonológica, são analisadas pelos falantes como sendo ocorrências de raízes diferentes – em virtude do que elas denotam. Dadas essas considerações, este trabalho assume que as raízes contêm um conteúdo semântico mínimo que permite a sua identificação.

Adicionalmente a essas questões, com relação à semântica não composicional das estruturas, assumindo Marantz (2001) e Arad (2003), é a concatenação da raiz ao seu primeiro núcleo categorizador que determina o domínio para a semântica não composicional de uma dada derivação. Para citar um exemplo, por hipótese, $\sqrt{\text{PONT}}$ tem um conteúdo semântico mínimo que permite sua identificação, e a

sua primeira concatenação é o que vai determinar a sua semântica “concreta”, não composicional; no caso da concatenação com um categorizador nominal *no*, o resultado seria *ponto* ou *ponta* que, ainda que com significados relacionados, designam elementos distintos; da mesma forma, no caso de uma concatenação com um verbalizador *vo*, o resultado poderia ser *apontar* (*o lápis*) ou *pontuar* (*no jogo*).

De qualquer forma, o significado não composicional seria negociado no primeiro ciclo, isto é, na primeira concatenação da raiz; no entanto, qualquer categorização adicional resultaria em um elemento com uma semântica composicional a partir do significado definido no primeiro ciclo; no caso, *pontuação* (*do jogo*) só poderia ser derivado de *pontuar* da mesma sorte que *apontação* (*do lápis*) só poderia derivar de *apontar*, no que se refere à semântica dessas unidades.

Nesse sentido, buscaremos mostrar que o significado composicional do gênero interpretável (*menino/menina*) em oposição ao significado não composicional do gênero não interpretável (*ponto/ponta*) é decorrente de, no primeiro caso, a valoração do traço de gênero se dar acima do primeiro categorizador (em *D*, mais precisamente) e, no segundo, a valoração se dar no primeiro categorizador (em *no*). A apresentação dessa proposta é esboçada na seção 1.3 e os detalhes da análise aplicada aos dados do português encontram-se nas seções 3.2 e 4.

1.2. Morfemas de classe e diacríticos de raiz

Conforme discutido na seção anterior, para a Morfologia Distribuída, as raízes são primitivos acategoriais, sendo sua categoria determinada pelo núcleo a que se adjunge em uma estrutura: caso se adjunja a *vo*, a derivação resultará em um verbo; caso se adjunja a *no*, o elemento formado será um nome, e assim por diante. Embora a hipótese da acategorialidade das raízes pareça funcionar bem para

línguas que não são morfológicamente ricas, como o inglês, o fato de as línguas românicas agruparem verbos e nomes em diferentes classes conjugacionais determinadas pela vogal temática que segue a raiz complexifica a questão.

A informação da categoria de um núcleo associado a uma raiz não é suficiente para que a estrutura receba a realização morfofonológica apropriada. O componente morfofonológico deve, de alguma forma, ter acesso à informação de a qual classe conjugacional uma dada raiz está arbitrariamente associada. Uma das soluções comumente encontradas na literatura – mas não por isso isenta de críticas⁴ – é assumir que as raízes carregam diacríticos que apontam para a classe conjugacional a que pertencem. Para ilustrar, o pertencimento de raízes verbais do português às três diferentes classes conjugacionais poderia ser representado como (1):

(1) Relação entre raízes e classes verbais

$\sqrt{\text{ADOR}}_{\text{[I]}}$

$\sqrt{\text{BEB}}_{\text{[II]}}$

$\sqrt{\text{PART}}_{\text{[III]}}$

Oltra-Massuet (1999), ao analisar a flexão verbal do catalão, admite que as raízes carregam diacríticos de classe verbal, mas que esses diacríticos se relacionam entre si por meio de uma hierarquia de marcação. Nessa hierarquia, quanto mais traços de valor positivo, mais marcada é a classe. Teixeira (2012) e Santana (2016) transpõem essa hierarquia às classes verbais do português, conforme mostrado em (2).

(2) Hierarquia de classes verbais do português

(a) Conjugação I ('amar'): $[-\alpha]$

(b) Conjugação II ('beber'): $[+\alpha, +\beta]$

(c) Conjugação IIIa ('unir') $[+\alpha, -\beta]$

⁴Cf. Acquaviva (2009), Bermúdez-Otero (2013), entre outros.

Dentre outras propriedades, a especificação de traços de conjugação conforme (2) captura o fato de a 1ª conjugação, em *-a*, ser a menos marcada (o que é capturado pela ausência de traços com valores positivos em sua especificação) e o fato de a 2ª e 3ª conjugações serem sincréticas em algumas formas verbais (o que é capturado pelo compartilhamento do traço $[+\alpha]$).

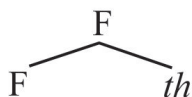
Como o pertencimento a uma classe conjugacional nas línguas românicas está diretamente relacionado com o morfema de vogal temática, que não carrega nenhuma informação (sintática ou semântica) a não ser o de classe conjugacional, a autora admite que a posição de vogal temática na estrutura não está presente na derivação sintática, mas é inserida pós-sintaticamente, no componente morfológico, para satisfazer uma condição de boa formação morfológica que exige que todo núcleo funcional projete uma posição temática, conforme (3).

(3) Condição de boa formação morfológica

Sintaxe

F

Morfologia



Ultra-Massuet (1999) sugere que o mesmo tratamento pode ser oferecido para as categorias não verbais da língua, observando que a projeção de uma posição temática, responsável pela indicação de classe, não é exclusividade do sistema verbal, já que ela ocorre também no domínio não verbal (que abarca nomes, adjetivos, advérbios, pronomes). Harris (1999) adota a proposta da autora e oferece um tratamento para a categoria nominal do espanhol, mostrando que, assim como os verbos, nomes também podem ser agrupados em classes de acordo com a vogal final que aparece adjungida à raiz,

mas que tal distribuição não deve ser confundida com a marcação de gênero, uma vez que todas as classes congregam elementos tanto masculinos quanto femininos, o que é ilustrado em (4) com dados do português.

(4) Classes nominais do espanhol

	Classe	Vogal	Gênero	Exemplos
(a)	I	<i>o</i>	Feminino-x	<i>tribo, libido</i>
			Masculino	<i>livro, corpo</i>
(b)	II	<i>a</i>	Feminino	<i>mesa, porta</i>
			Masculino-x	<i>planeta, sistema</i>
(c)	III	$\emptyset \sim e$	Masculino	<i>amor, lote</i>
			Feminino	<i>dor, ponte</i>
(d)	IV	<i>e</i>	Masculino	<i>vale, doce</i>
			Feminino	<i>classe, pele</i>

Nas classes em (4), “x” equivale a “excepcional”, já que os (poucos) membros femininos que compõem a Classe I e masculinos que compõem a Classe II fogem à tendência comum do espanhol (e do português) de que nomes terminados em *o* são masculinos e nomes terminados em *a*, femininos. Dada a arbitrariedade também da informação de gênero, o autor admite que as raízes vêm especificadas não só para classe, conforme os verbos, mas também para gênero.

No entanto, nem todas as especificações precisam vir marcadas; apenas aquelas que não correspondem ao padrão geral. Mais particularmente, quanto ao gênero, apenas o traço de gênero feminino precisa vir marcado, uma vez que o autor considera que o masculino é o gênero default. Quanto à classe, apenas raízes que não se conformam ao padrão geral mencionado necessitam vir com a informação de classe. A fim de explicar a grande afiliação de palavras femininas à classe II, Harris propõe a regra de redundância em (5),

responsável por encaminhar raízes femininas subespecificadas para classe nominal à classe II.

(5) Regra de redundância

$[_{FEM}] \rightarrow II$

Adicionalmente, do mesmo modo que Oltra-Massuet (1999), Harris entende que a melhor maneira de expressar as informações de classe é em termos de feixes de traços, o que pode ser observado em (6).

(6) Hierarquia de traços nominais do espanhol

- (a) Classe I: $[\alpha]$
- (b) Classe II: $[\alpha, \beta]$
- (c) Classe III: $[\alpha, \beta, \gamma]$
- (d) Classe IV: $[\alpha, \beta, \gamma, \delta]$

Nesse esquema, a Classe I é a classe não marcada, isto é, a classe default, a Classe II levemente marcada, a Classe III mais marcada e a Classe IV maximamente marcada. Como explica Harris, essa análise assume que os identificadores das classes têm o estatuto formal de traços (decomponíveis) e, conforme já mencionado, são codificados como parte da informação das entradas de Vocabulário das raízes, do mesmo modo que os traços de gênero. Alguns dos exemplos apresentados em (4) são representados em (7) de acordo com a codificação de gênero e classe proposta por Harris (1999).

(7) Raízes nominais e traços intrínsecos

- | | |
|--|--|
| (a) ‘livro’: \sqrt{LIVR} | ‘tribo’: $\sqrt{TRIB}_{[\alpha, FEM]}$ |
| (b) ‘planeta’: $\sqrt{PLANET}_{[\beta]}$ | ‘mesa’: $\sqrt{MES}_{[FEM]}$ |
| (c) ‘amor’: $\sqrt{AMOR}_{[\gamma]}$ | ‘dor’: $\sqrt{DOR}_{[\gamma, FEM]}$ |
| (d) ‘vale’: $\sqrt{VAL}_{[\delta]}$ | ‘classe’: $\sqrt{CLASS}_{[\delta, FEM]}$ |

A discussão a respeito dos problemas relacionados ao pressuposto de que tanto gênero quanto classe são codificados na raiz aparece na seção 2, onde são apresentadas as análises de Alcântara (2003, 2010) e Armelin (2015) para o português, a primeira delas assumindo os mesmos pressupostos de Harris e a segunda, assumindo uma posição contrária, qual seja, a de que nenhuma dessas informações pode fazer parte da raiz. Conforme mencionado na introdução, o presente trabalho adota uma análise intermediária, em que as raízes albergam a classe, porém, não o gênero. Dadas essas considerações, a próxima subseção apresenta uma análise alternativa para o núcleo que aloca os traços de gênero.

1.3. Nominalizadores e traços de gênero

Conforme mencionado brevemente na seção anterior, o trabalho de Harris (1999), dedicado à derivação morfofonológica das classes não verbais do espanhol, admite que tanto os traços de classe quanto os traços de gênero são inerentes às raízes. No entanto, muitos trabalhos desde Ritter (1993), que se dedicam à explicação de aspectos sintáticos e semânticos do gênero, admitem que os traços de gênero são carregados por núcleos funcionais que estão acima da raiz. Para Ritter (1993), a determinação de qual núcleo deve albergar os traços de gênero é uma variação paramétrica e, portanto, translinguística. A autora admite que, nas línguas românicas, gênero é um traço do núcleo de NumP e que, no hebraico, gênero é um traço do nominalizador que se liga à raiz.

Muitos autores que seguem a intuição de Ritter, no entanto, admitem que a variação do núcleo que alberga os traços de gênero pode ser uma variação intralinguística, determinada pela interpretabilidade ou não interpretabilidade do gênero. A intuição geral compartilhada por todos esses trabalhos é a de que gênero não interpretável está em um núcleo mais baixo do que gênero

interpretável, mas a implementação específica costuma ser bastante variável. A título de exemplo, para Atkinson (2015), o gênero não interpretável é propriedade das raízes e gênero interpretável é propriedade dos nominalizadores; conforme Panagiotidis (no prelo), gênero não interpretável está no nominalizador e gênero interpretável está em AnimP e, para Kučerová (2018), gênero não interpretável está no nominalizador e o interpretável está em D.

Seguindo a ideia central dos trabalhos recém mencionados, o presente estudo assume que os traços de gênero não interpretável não são inerentes às raízes, mas sim, constituem uma propriedade dos nominalizadores, e que as condições de licenciamento em LF são responsáveis por convergir as estruturas bem formadas e implodir as malformadas.

Para os propósitos do presente trabalho, o principal motivo para assumir que o gênero não interpretável não é inerente à raiz, mas sim propriedade do nominalizador, é o fato de que nem todas as raízes figuram em contextos com apenas um valor de gênero; raízes como $\sqrt{\text{BARC}}$, por exemplo, podem receber o valor feminino ou o valor masculino de gênero – em outras palavras, podem se concatenar a nominalizadores que carregam [MASC] ou [FEM]. Muito embora a mesma raiz possa apresentar também vogais temáticas diferentes, admitimos que isso só ocorre como consequência da mudança do valor de gênero, de uma forma que ficará mais clara na seção 3.2. Assim, defendemos que a especificação de vogais temáticas é inerente a uma raiz, mas que a afirmação de que o gênero é intrínseco às raízes é imprecisa e pouco econômica.

Quanto ao valor do gênero interpretável, nós adotaremos a análise de Kučerová (no prelo), que admite que o núcleo onde o traço de gênero semântico é valorado é D; entretanto, não discutiremos a diferença entre essa e as outras análises que admitem que o gênero interpretável é valorado, por exemplo, em NumP ou AnimP, uma vez que não nos deteremos sobre questões sintáticas e semânticas

envolvidas no gênero interpretável. Assim, o que é relevante para a presente análise é o pressuposto de que o núcleo responsável pela atribuição de gênero semântico está acima do núcleo nominalizador e acima, portanto, do primeiro núcleo categorizador, contribuindo composicionalmente com o significado final da estrutura. Adicionalmente, ainda conforme Kučerová (no prelo), admitimos que a valoração do gênero interpretável é realizada por avaliação do contexto extralinguístico.

Para concluir esta seção, retomamos que, seguindo as intuições de trabalhos como o de Oltra-Massuet (1999) e Harris (1999), o presente trabalho investiga a hipótese de que diacríticos de classe (tanto verbal quanto nominal) são inerentes às raízes e computados apenas em PF, não sendo, portanto, uma informação disponível em LF. Por outro lado, seguindo as intuições de trabalhos discutidos na presente seção, investigamos a hipótese de que a informação de gênero gramatical é uma propriedade do núcleo nominalizador e está presente na sintaxe, sendo, portanto, acessível a LF e, finalmente, a informação de gênero semântico é atribuída acima do nominalizador (mais particularmente, em D).

Se essas hipóteses estiverem na direção correta, espera-se que uma raiz só possa receber interpretações diferentes e não composicionais em um ciclo nominal caso se associe a diferentes matizes de nominalizador, isto é, a nominalizadores com diferentes traços de gênero. O mesmo não é possível quando uma raiz se associa a diferentes diacríticos de classe nominal ou a diferentes diacríticos de classe verbal, uma vez que essa informação não está disponível em LF. Dito de outro modo, uma mesma raiz associada a diferentes vogais temáticas só poderá receber interpretações diferentes à medida que informações computáveis em LF, como gênero, também forem manipuladas, mas, crucialmente, não quando *apenas* a informação de classe difere.

Ademais, quando as diferentes interpretações de uma mesma raiz com diferentes valores de gênero forem composicionais, isso decorre de a raiz ter sido concatenada a um mesmo nominalizador, mas ter recebido diferentes valores de gênero em D, via avaliação do contexto extralinguístico. As consequências das primeiras considerações são examinadas na seção 3 e, das últimas, na seção 4. A seção 2 a seguir retoma as análises disponíveis na literatura para o português brasileiro, com base no quadro da Morfologia Distribuída.

2. A marcação do gênero em português

Assumindo o modelo da MD – e mais especificamente com base na análise de Harris (1999) para o espanhol, discutida na seção 1 – Alcântara (2010) propõe um tratamento para as classes formais nominais do português. Diferentemente da classificação sugerida por Camara Jr., baseada em gênero, a autora admite que nomes (e adjetivos) estão distribuídos em quatro classes formais determinadas pela vogal final, sendo, portanto, heterogêneas com respeito ao gênero.

Conforme assume a autora, a “Classe I” é identificada pelo morfema *-o* e reúne majoritariamente palavras masculinas, como *lobo*, *figo* e *jato* – sendo a classe não marcada para esse gênero – e um pequeno número de palavras femininas, tais como *libido* e *tribo*. A “Classe II” – cujo morfema de classe é *-a* – é considerada a classe não marcada para feminino (como em *casa*, *menina* e *girafa*) ainda que agrupe algumas palavras do gênero masculino, tais como *sistema*, *planeta* e *cinema*.

A autora formaliza a não marcação dessa classe para o feminino por meio de uma regra de redundância semelhante à de Harris, apresentada em (5), que encaminha raízes marcadas para gênero feminino, mas não para classe, à Classe II. A “Classe III”, por sua vez, é aquela composta por elementos que possuem a vogal final *-e* no

contexto de plural, podendo ser tanto \emptyset quanto *-e* no singular, como é o caso de *pente*, *ponte*, *amor* e *autor*. Finalmente, a “Classe IV” contém palavras que são identificadas pela presença do morfema \emptyset tanto no plural quanto no singular, como é o caso de *bagagem*, *museu* e *sofá*⁵.

Ainda seguindo a proposta de Harris, Alcântara admite que tanto a classe formal quanto o valor de gênero, por serem imprevisíveis, devem aparecer listados com as raízes como, por exemplo, $\sqrt{\text{COLHER}}_{[\text{III}, \text{FEM}]}$. Não são todas as raízes da língua que devem vir marcadas tanto com a informação de classe quanto com a de gênero. Como o traço masculino é o traço de gênero não marcado, e o morfema *-o* é a realização default para palavras masculinas, as raízes que formam nomes masculinos em *-o* não vêm marcadas com nenhum diacrítico, o que reflete a não marcação desse grupo de palavras e o número extenso de nomes que se enquadram nesse padrão na língua.

As palavras femininas terminadas em *-a* vêm marcadas apenas para o traço feminino, sendo o pertencimento de palavras femininas à Classe II previsível, não necessitando, portanto, que essa informação seja carregada pela raiz – a regra de redundância já mencionada as encaminha automaticamente para a Classe II. Já palavras femininas em *-o* necessitam da dupla marcação, masculinas em *-a* apenas marcação de classe, e assim por diante.

A proposta de Alcântara (2010) é bastante abrangente e consegue dar conta de muitos fenômenos do sistema nominal da língua, inclusive no que tange a propriedades (morfo)fonológicas; contudo, especificamente com relação à marcação de gênero, sua proposta falha em dois aspectos fundamentais. Em primeiro lugar, sua análise

⁵ Em trabalho anterior (ALCÂNTARA, 2003), os membros da Classe III estão distribuídos em duas classes distintas, totalizando, assim, cinco classes. Uma das classes - Classe III - é formada por nomes que apresentam a vogal final *-e* no plural e, no singular, \emptyset ou *-e*, desde que, no caso dos nomes em *-e*, a última consoante da raiz não seja licenciada em coda silábica (sendo o *-e* do singular, assim, uma vogal epentética). A outra classe - Classe IV - é formada por nomes que terminam em *-e* tanto no singular quanto no plural desde que a última consoante da raiz seja licenciada em coda silábica (sendo o *-e* do singular uma verdadeira vogal temática, e não uma epêntese). Para os presentes propósitos, a escolha por uma ou outra proposta não é relevante.

se concentra sobretudo em nomes não variáveis, não havendo uma discussão explícita a respeito de uma possível operação de flexão de gênero que permite gerar pares como *menino/menina* e que faça referência ao gênero natural.

Além disso, sua análise não explica de que maneira se dá a relação entre palavras como *barco/barca*, que parecem apresentar a mesma raiz, dada a semelhança fonológica e semântica. A maneira como o sistema da autora permite que se lide com tais pares é admitindo que as raízes $\sqrt{\text{BARC}}$ e $\sqrt{\text{MENIN}}$ sejam listadas duas vezes na Lista 1, com as informações necessárias para gerar cada uma das formas. No caso desses dois pares, a Lista 1 seria provida com $\sqrt{\text{BARC}}$, $\sqrt{\text{BARC}}_{[\text{FEM}]}$, $\sqrt{\text{MENIN}}$ e $\sqrt{\text{MENIN}}_{[\text{FEM}]}$.

Dessa forma, um par como *barco/barca* – com uma “flexão de gênero não interpretável” – e um como *menino/menina* – com uma “flexão de gênero interpretável” – são não somente não relacionáveis, mas também indistinguíveis de um par como *livro/mesa*, cujos membros não apresentam qualquer relação. A semelhança fonológica e semântica entre *barco* e *barca* e entre *menino* e *menina*, assim, se perde. O segundo problema, consequência do primeiro, é que há uma proliferação de raízes, o que incha a Lista 1, levando a uma maior implausibilidade psicológica da análise

Convém observar que o objetivo da autora não é descrever pares de nomes variáveis ou o mecanismo de flexão; logo, sua análise, apesar de incompleta, fornece um instrumental para descrever os padrões encontrados no português, que se baseia na premissa de que quaisquer informações imprevisíveis de classe e de gênero são carregadas pelas raízes.

Operando igualmente com a maquinaria da Morfologia Distribuída, Armelin (2015) parte da distribuição em classes formais assumida por Alcântara (2003, 2010) e, na tentativa de sanar o problema da proliferação de raízes e de diacríticos de raiz, a autora alega, na esteira de trabalhos como o de Acquaviva (2009), que as

raízes não carregam nenhum diacrítico, nem de gênero nem de classe.

Diferindo-se ainda mais radicalmente da literatura de gênero e classe formal, a autora afirma que a distinção entre gênero e classe não é teoricamente motivada e que só há gênero; a aparente não isomorfia entre o valor de gênero e expressão morfofonológica é capturada por meio dos próprios itens de Vocabulário que estão em competição para preencher o núcleo de gênero (Gen) que, nos casos não default, carregam a informação de a qual raiz devem se anexar – ou seja, são apenas diferentes formas de realizar um mesmo núcleo funcional e um mesmo conjunto de traços.

Mais particularmente, a autora apresenta duas versões dos itens de Vocabulário responsáveis por realizar fonologicamente o gênero, que aparecem reproduzidas em (8) (exemplos (12) e (14) da autora) e (9) (exemplos (12) e (24) da autora), como segue. Mesmo que a autora opte pelo segundo conjunto, acreditamos que ambos enfrentam dificuldades que são discutidas a seguir, o que nos levará ao retorno do tratamento das vogais temáticas nominais por meio de traços de classe.

- (8) (a) $\text{Gen}_{[f]} \leftrightarrow /o/ / \sqrt{\text{TRIB}}, \sqrt{\text{LIBID}} \dots$
(b) $\text{Gen}_{[m]} \leftrightarrow /a/ / \sqrt{\text{PLANET}}, \sqrt{\text{MAP}} \dots$
(c) $\text{Gen}_{[f]} \leftrightarrow /-a/$
(d) $\text{Gen}_{[m]} \leftrightarrow /o/$
- (9) (a) $\text{Gen} \leftrightarrow /o/ / \sqrt{\text{TRIB}}, \sqrt{\text{LIBID}}, \sqrt{\text{MODEL}}$
(b) $\text{Gen} \leftrightarrow \emptyset / \sqrt{\text{MAR}}, \sqrt{\text{CAFÉ}}, \sqrt{\text{FÂ}}$
(c) $\text{Gen} \leftrightarrow /a/ / \sqrt{\text{PLANET}}, \sqrt{\text{MAP}}, \sqrt{\text{COLEG}}$
(d) $\text{Gen}_{[f]} \leftrightarrow /a/$
(e) $\text{Gen}_{[m]} \leftrightarrow /o/$

No primeiro conjunto de itens de Vocabulário – propositalmente incompleto para gerar todos os padrões nominais da língua, já que a versão final dos itens é, para a autora, os itens em (9) – todos os itens são especificados para o valor de gênero ao qual deverão ser inseridos, e apenas os itens não default (8a-b) apresentam em seu contexto de inserção a listagem de raízes. O problema enfrentado por (8), conforme a própria autora menciona, tem relação com os nomes com gênero variável, como *menino/menina*, *professor/professora*, *o/a artista*, *o/a modelo*. Afinal, a partir do momento em que se contemplam as vogais temáticas *-e* e *ø*, perde-se a generalização de que, em pares de gênero variável, ou as duas formas são iguais (*o/a artista*, *o/a modelo*) ou a forma feminina é necessariamente em *-a* (*menino/menina*, *professor/professora*, *mestre/mestra*). Em outras palavras, não há pares como *menino/*menine* ou *artista/*artiste*, em que a vogal final varia sem que a vogal da forma feminina seja *-a*.

O motivo pelo qual os itens em (8) obliteram a generalização mencionada anteriormente é o de que as vogais temáticas não listadas *e* e *ø* necessitariam ser listadas duas vezes: uma para garantir sua inserção no contexto do traço masculino e outra para garantir sua inserção no contexto do traço feminino – e, por serem sempre imprevisíveis em relação ao valor de gênero, teriam as raízes listadas conforme (8a-b).

Assim, havendo seis itens de Vocabulário com especificação contextual, nada impediria que uma raiz como $\sqrt{\text{COLEG}}$, por exemplo, estivesse listada no item que realiza *-a* no masculino e no item que realiza *-e* no feminino em vez de apenas no primeiro item mencionado, ou que uma raiz como $\sqrt{\text{PARENT}}$ estivesse listada no item que realiza *-e* no masculino e no item que realiza *-o* no feminino, em vez de nos dois itens que realizam *-e* (no masculino e no feminino).

Por conta dessa perda de generalização, a autora opta pelo conjunto de itens de Vocabulário em (9)⁶, em que os itens que carregam a lista de raízes em sua especificação contextual não carregam valor de gênero. Essa formulação permite dar conta tanto dos pares gerais da língua (*menino/menina*) quanto dos pares em que a forma masculina e feminina são iguais (*o/a artista, o/a estudante*), mas não dá conta de pares como *mestre/mestra, presidente/presidenta* e todos os agentivos em *-(d)or*, como *professor/professora*.

Outro problema enfrentado pelo modelo de itens de Vocabulário em (9) é que ele viola o princípio do subconjunto, conforme formulado por Halle (1997) e adotado pela própria autora. Esse princípio, que rege a operação de inserção de Vocabulário, determina que quando há uma relação de subconjunto entre dois candidatos em potencial para um complexo de traços, o mais altamente especificado, ou seja, aquele que contém o maior número de traços, ganha a competição. Sob essa formulação do princípio do subconjunto, os itens em (9d-e) são mais específicos do que os itens em (9a-c) e, portanto, devem ganhar a competição sempre quando há um traço masculino ou feminino envolvido, o que impede a inserção dos itens (9a-c) em quaisquer contextos, gerando resultados agramaticais, como **triba* e **planeto*.

Por fim, ao admitir que as vogais temáticas nominais são sempre a realização do nó de gênero, impede-se uma análise unificada da expressão morfofonológica das outras classes não verbais que não têm gênero, como preposições e advérbios, que também apresentam as mesmas vogais finais. Como exemplo de preposições, há *para, entre, com*ø, e como exemplo de advérbios tem-se *dentro, fora, hoje* e *após*ø.

Quanto à diferença entre gênero gramatical e gênero semântico, se, por um lado, Alcântara (2003, 2010) põe a discussão de lado, por

⁶ A autora sugere que todas as instâncias de *-e* final são parte da raiz, e não vogais temáticas, limitando o conjunto de vogais finais a *-o, -a* e *ø*.

outro, para dar conta dessa diferença, Armelin (2015) propõe que o núcleo Gen pode albergar três configurações dos traços [MASC] e [FEM]: apenas [MASC], apenas [FEM] ou a marcação dupla [MASC, FEM]. Quando o núcleo apresenta apenas um dos dois traços isoladamente, o resultado é a manifestação do gênero não interpretável. Quando, por outro lado, o núcleo apresenta a dupla especificação [MASC, FEM], apenas um deles se sujeita à inserção de Vocabulário, e o traço não pronunciado se mantém disponível como informação de fundo em LF, o que permite a interpretação contrastiva de gênero, relacionada ao gênero biológico. Isso é ilustrado em (10) para o par de gênero não interpretável *barco/barca* e para o par de gênero interpretável *menino/menina*. O tachado simboliza a não realização morfofonológica do traço, mas sua permanência como pano de fundo para a interpretação em LF.

(10) Marcação de gênero conforme Armelin (2015)

- (a) ‘barco’: [[√BARC] Gen_[MASC]]
- (b) ‘barca’: [[√BARC] Gen_[FEM]]
- (c) ‘menino’: [[√MENIN] Gen_[MASC, FEM]]
- (d) ‘menina’: [[√menin] Gen_[MASC, FEM]]

Assim, para a autora, gênero é sempre interpretável, e a diferença entre o que se costuma chamar de gênero interpretável e não interpretável é reavaliada em termos de leitura contrastiva e leitura não contrastiva: quando o traço aparece isoladamente no núcleo, não é codificado contraste e, assim, não se pode ter uma interpretação pareada masculino/feminino, o que gera uma interpretação não composicional; quando, por outro lado, há uma especificação bivalente do núcleo Gen, a estrutura codifica contraste e, portanto, admite a formação de pares masculino/feminino, cuja interpretação relaciona-se ao gênero biológico e é composicional.

Uma limitação dessa análise é que ela parece não ser capaz de capturar a leitura ampla do gênero masculino, ou seja, o fato de, por exemplo, *aluno* poder se referir tanto a homens quanto a mulheres, ao contrário de *aluna*, que só pode se referir a mulheres. Afinal, se o núcleo Gen na derivação de *aluno* chega em LF especificado como [MASC, FEM], não é claro de que maneira a forma masculina pode ser interpretada como se referindo a seres dos dois sexos, já que a representação é tal que exclui a interpretação feminina. Se, por outro lado, essa interpretação for, de alguma forma, acessível a LF, não seria claro de que maneira a mesma interpretação ampla é bloqueada no caso da especificação de traços [MASC, FEM]. Dito de outro modo, essa análise propõe uma representação simétrica para um fenômeno que é assimétrico.

Contudo, o insight da autora de separar composicionalidade de interpretabilidade no que diz respeito à contribuição da especificação de gênero – ou seja, ao afirmar que gênero é sempre interpretável, mas nem sempre composicional – é explorada no presente trabalho. Entretanto, devido às limitações mencionadas nesta seção, a presente análise se distancia da análise de Armelin (2015) no que se refere à distinção entre gênero e classes nominais e na maneira como a gramática codifica a composicionalidade e a não composicionalidade do valor de gênero.

3. Raízes e vogais temáticas: expressão morfofonológica

Esta seção apresenta a análise defendida no presente trabalho para a expressão morfofonológica das vogais temáticas verbais e nominais e, no caso dos nomes, de sua relação com a especificação de gênero. Conforme já mencionado, nossa proposta mescla diferentes intuições disponíveis na literatura mencionada nas seções anteriores,

tendo como ponto de partida a tentativa de capturar tanto os padrões existentes quanto os não existentes, segundo nosso levantamento, da relação entre vogais temáticas e raízes (e gênero, no caso dos nomes).

No caso da classe verbal, parece não haver na língua pares de verbos que pertençam a diferentes classes conjugacionais e que compartilhem a mesma raiz. Se tomados critérios semânticos, e não apenas fonológicos, para determinar a identidade de uma raiz, conforme sugerido na seção 1.1, vê-se que a mudança de conjugação, por mudar drasticamente o significado, indica que se está diante de duas raízes diferentes, porém homófonas, conforme ilustrado em (11). Assim, é razoável afirmar que uma mesma raiz só pode pertencer a uma classe conjugacional.

(11) Relação entre raízes e vogais temáticas verbais

- | | |
|---------------------|-------------------------|
| (a) podar - poder | (g) cobrar - cobrir |
| (b) sentar - sentir | (h) vender - vender |
| (c) ver - vir | (i) cegar - seguir |
| (d) fundar - fundir | (j) gerar - gerir |
| (e) falar - falir | (k) doar - doer |
| (f) parar - parir | (l) consumir - consumir |

No domínio nominal, não encontramos pares de nomes que pertencem a diferentes classes sem que (i) o gênero mude juntamente com a classe nominal e (ii) ao menos a forma feminina do par de nomes pertença à classe nominal default (o que é bem menos comum do que quando ambas as formas pertencem à classe default). Isso vale tanto para pares de nomes com gênero interpretável quanto para pares com gênero não interpretável. Conforme ilustrado nos exemplos em (12), quando o gênero do nome se mantém, mas a vogal final é modificada, a mudança drástica de significado indica que se

está perante raízes diferentes (diferentemente de pares em que a vogal muda juntamente com o gênero, como *barco-barca*, *trilho-trilha*, *fruto-fruta*, *jarro-jarra* etc.).

(12) Relação entre raízes e vogais temáticas nominais

- | | |
|-------------------|-------------------|
| (a) leite - leito | (e) porte - porto |
| (b) mente - menta | (f) saque - saco |
| (c) ponte - ponta | (g) posse - poça |
| (d) posto - poste | |

Sugerimos que a impossibilidade de uma única raiz pertencer a classes verbais e nominais diferentes (sem a simultânea mudança de gênero no caso dos nomes) pode ser capturada ao assumir que traços de gênero são traços ativos na sintaxe e, portanto, disponíveis em LF, mas que traços de classe só estão disponíveis para leitura no componente morfofonológico (em PF) e, portanto, não estão presentes em LF – o que impede que a mudança *apenas* de traços de classe modifique a interpretação de uma mesma raiz. Nas próximas duas subseções, é mostrada a implementação do sistema proposto, iniciando com classes verbais (seção 3.1) para em seguida apresentar o tratamento para as classes nominais (seção 3.2).

3.1. Verbos

No caso dos verbos, seguindo a análise de Teixeira (2012) baseada na proposta de Oltra-Massuet (1999), propomos que as raízes podem estar associadas ou ao traço $[+\alpha]$, quando pertencente à 3ª conjugação, ou ao traço $[+\beta]$, quando pertencente à 2ª conjugação. Pelo fato de as classes conjugacionais se relacionarem entre si por meio de uma hierarquia de marcação, uma raiz associada ao traço $[+\beta]$ é automaticamente também associada ao traço $[+\alpha]$ – e uma raiz $[+\alpha]$ é automaticamente interpretada como contendo $[-\beta]$.

Quando a raiz não está associada a nenhuma classe verbal, ela é automaticamente encaminhada à classe $[-\alpha]$, relativa à 1ª conjugação, conforme esquematizado em (13). Os itens de Vocabulário para as vogais temáticas verbais podem ser conferidos em (14) – em que o ordenamento reflete o grau de marcação do traço de classe:

(13) Hierarquia de marcação para vogais temáticas verbais

- | | |
|---------------------------------------|---------------------------|
| (a) 1ª conjugação: $[-\alpha]$ | $\sqrt{AM}_{[]}$ |
| (b) 2ª conjugação: $[\alpha, +\beta]$ | $\sqrt{BEB}_{[+\beta]}$ |
| (c) 3ª conjugação: $[\alpha, -\beta]$ | $\sqrt{PART}_{[+\alpha]}$ |

(14) Itens de Vocabulário para vogais temáticas verbais

- (a) $/e/ \leftrightarrow Th_{[+\beta]}$
 (b) $/i/ \leftrightarrow Th_{[+\alpha]}$
 (c) $/a/ \leftrightarrow$ demais ambientes

Por ser sintática e semanticamente inerte, a posição preenchida pela vogal temática só é inserida no componente morfológico, por meio de uma condição de boa formação que exige que todo categorizador projete uma posição temática⁷. Após a inserção desse nó, a operação de inserção de Vocabulário ocorre ciclicamente da raiz para fora, inserindo corretamente os expoentes temáticos quando no contexto das raízes apropriadas.

É interessante notar que, em algumas formas verbais flexionadas, a 2ª e a 3ª conjugações sofrem neutralização em proveito da 3ª conjugação. O caso prototípico são as formas de particípio, como *bebido* (de *beber*) e *partido* (de *partir*), em que a esperada vogal temática *-e-* aparece como *-i-* sem que haja nenhuma motivação fonológica para tal. Teixeira (2012) propõe que essa neutralização seja alcançada por meio de uma operação de empobrecimento no

⁷ Nesse ponto, nos divergimos da proposta de Ultra-Massuet (1999), segundo a qual não são somente os núcleos categorizadores que projetam uma posição temática, mas sim, qualquer núcleo funcional.

componente morfológico (cf. BONET, 1991; HALLE & MARANTZ, 1993; HALLE, 1997), responsável por apagar o traço $[+\beta]$ no contexto dos traços de participio passado. Como a 2ª e a 3ª conjugações compartilham o traço $[+\alpha]$, o empobrecimento de $[+\beta]$ tem como efeito uma mudança de classe (da 3ª para 2ª conjugação) no contexto especificado pela regra em (15).

(15) Regra de empobrecimento

$$[+\beta] \leftrightarrow \emptyset / [+PART, +PASS]$$

Crucialmente, por ser uma operação puramente morfológica, que se aplica no caminho para PF após a estrutura já ter sido enviada para LF, essa operação em nada interfere na interpretação das formas de participio. Sua função é simplesmente dar conta da expressão morfofonológica da vogal temática de 2ª conjugação em determinados contextos.

Além disso, por mais que aparente ser o caso de uma mesma raiz, em contextos diferentes, estar associada a traços de classe diferentes, a raiz não está *intrinsecamente* associada a diferentes traços de classe. A realização de uma vogal temática de 3ª conjugação em um verbo de 2ª conjugação não significa que essa raiz verbal pertença a duas classes diferentes. A raiz pertence a uma única classe, que pode adquirir o aspecto de outra classe por meio de uma operação de empobrecimento, cujo contexto de aplicação é morfossintaticamente previsível. Outros casos de neutralização entre diferentes classes verbais não serão tratados no presente trabalho, por questões de espaço. Para tanto, referimos o leitor a Teixeira (2012) e Santana (2019). Como mostrado na subseção 3.2, uma operação semelhante está ativa no sistema nominal.

3.2.Nomes

Como já afirmado, a expressão morfofonológica dos nomes em português envolve, principalmente, duas propriedades, a saber, gênero e classe. Sendo assim, o objetivo da presente subseção é apresentar uma análise da expressão morfofonológica das vogais nominais que leve em conta três fatos do sistema nominal do português, a saber, (i) a natureza idiossincrática das vogais temáticas em muitos nomes da língua, (ii) a arbitrariedade da marcação de gênero na maioria (mas não em todos) os nomes do português, (iii) a diferença de interpretação do gênero semântico e do gênero gramatical ao lado da semelhança na maneira como a expressão morfofonológica desses dois tipos de gênero se dá.

Com relação ao comportamento idiossincrático da informação de classe dos nomes, nós vamos assumir, na esteira de Harris (1999) e Alcântara (2003, 2010), que ela é uma propriedade inerente às raízes, da mesma forma como o proposto para os verbos na subseção anterior. Esses traços também podem ser descritos por meio de uma hierarquia de marcação, como mostrado em (16)⁸. Os itens de Vocabulário podem ser conferidos em (17).

(16) Hierarquia de marcação para classes nominais do PB

- (a) Classe I (-o): $[-\alpha]$
- (b) Classe II (-a): $[\alpha, -\beta]$
- (c) Classe III (\emptyset / -e): $[\alpha, \beta]$ ⁹

⁸ Diferentemente do que nós propomos em (16), tanto Harris (1999) quanto Alcântara (2003) utilizam traços privativos, e não binários, na descrição das diferentes classes nominais, conforme ilustrado em (6). A fim de manter o paralelismo com a análise proposta para os verbos com base em Oltra-Massuet (1999) e Teixeira (2012), analisamos a relação entre as classes nominais por meio de traços binários.

⁹ Por questão de espaço, não exploraremos a distinção entre as classes III, IV e V sugeridas por Alcântara (2010) e agrupamos todas as palavras em \emptyset ou em -e na mesma classe.

(17) Itens de Vocabulário para vogais nominais

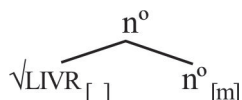
- (a) /e ~ ø/ ↔ Th_[+β]
- (b) /a/ ↔ Th_[+α]
- (c) /o/ ↔ demais ambientes

Todavia, assim como nos verbos, não é necessário que todas as raízes carreguem esses traços, mas apenas aquelas cuja vogal final não é a realização default do traço de gênero do nominalizador (-o para masculino e -a para feminino); portanto, antes de ilustrar como o sistema atribui a realização das vogais temáticas nominais, é necessário entender como se dá a atribuição de gênero, o que nos leva ao segundo ponto elencado.

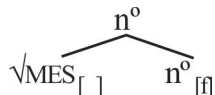
Seguindo, dentre outros (cf. § 1.3), Kučerová (no prelo) e Panagiotidis (no prelo), nós assumimos que gênero não interpretável é uma propriedade do nominalizador ao qual a raiz se anexa – e não das próprias raízes, como defende Alcântara (2003, 2010). Logo, assim como uma mesma raiz pode se associar a diferentes núcleos categorizadores, pode também se associar a diferentes matizes de um núcleo no, quais sejam, um subespecificado para gênero ($n^{\circ}_{[\]}$), restrito a raízes que formam pares de nome com gênero interpretável, um que carrega o traço feminino ($n^{\circ}_{[f]}$), restrito a raízes que formam nomes femininos com gênero não interpretável, e um que carrega o traço masculino ($n^{\circ}_{[m]}$), restrito a raízes que formam nomes masculinos com gênero não interpretável. Uma mesma raiz carregando qualquer traço de classe pode, em princípio, se unir a quaisquer matizes de nominalizador. Os exemplos em (18) ilustram essas combinações para as raízes que se ligam unicamente a um nominalizador.

(18) Palavras invariáveis para gênero

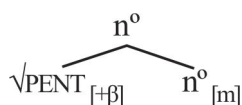
(a) Masculinas terminadas em -o (“livro”)



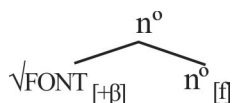
(b) Femininas terminadas em -a (“mesa”)



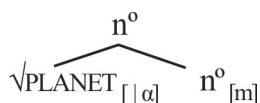
(c) Masculinas terminadas em -e (“pente”)



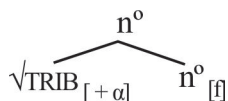
(d) Femininas terminadas em -e (“fonte”)



(e) Masculinas terminadas em -a (“planeta”)



(f) Feminina terminada em -o (“tribo”)



Como a realização default do gênero masculino é -o e de gênero feminino é -a, as raízes em (18a-b), que apresentam esse padrão, não vêm marcadas para classe – e a regra de redundância apresentada em (5) ao lado da subespecificação do item de Vocabulário em (17c)

dão conta de inserir corretamente, respectivamente, *-a* e *-o* em (18a-b) – entretanto, conforme mostrado na seção 4, é possível que os falantes interpretem a raiz de, por exemplo, *garrafa* tanto como marcada quanto como não marcada para classe [+α], quando perante fenômenos que forçam essa decisão.

Quando o nome apresenta quaisquer outras combinações de valor de gênero e vogal temática, a informação de classe deve ser parte da informação intrínseca da raiz, conforme ilustram (18c-f). Nesse sentido, a presente análise em muito se assemelha à análise de Harris (1999) e Alcântara (2003, 2010). A diferença fundamental recai sobre o núcleo que alberga o traço de gênero, que não é a raiz, e sim o nominalizador, o que permite evitar a proliferação de raízes para casos como *barco/barca*, conforme discutido na seção 2. Na seção 4, é discutida a maneira como a gramática bloqueia derivações em que as raízes se concatenam a nominalizadores indesejados, gerando, por exemplo, **livra*, em vez de *livro*.

No caso de pares de nomes que compartilham a mesma raiz, mas que podem ocorrer tanto em contextos masculinos quanto femininos, já foi mencionado que só pudemos encontrar pares que fazem masculino em *-o* e feminino em *-a* (como *mato/mata*) ou pares cuja realização da classe não varia com a mudança de valor de gênero (como (o) rádio/(a) rádio). Isso porque, quando uma raiz não é marcada, as regras de redundância encaminham-na para diferentes classes a depender do valor de gênero e, quando é marcada, a realização da vogal independe do valor de gênero e é, portanto, sempre a mesma. Um exemplo do primeiro tipo é dado em (19a) e um do segundo tipo é dado em (19b)

(19) Palavras variáveis com gênero não interpretável

(a) Pares masculino em -o/feminino em -a (“mato” e “mata”)



(b) Pares invariáveis para classe (“(o) rádio” e “(a) rádio”)



Finalmente, a respeito de (iii), seguindo Kučerová (no prelo), nós admitimos que, no caso de gênero interpretável, a informação de gênero não vem do nominalizador, mas é valorada no núcleo D que o c-comanda. Mais especificamente, o sistema de Kučerová que adotamos nesta análise é composto das seguintes operações sintáticas, aplicadas nesta ordem: (i) concatenação de D, (ii) os traços- ϕ em D sondam os traços- ϕ de no, (iii) os traços- ϕ dos dois núcleos ficam em correspondência, (iv) valoração por concordância (operação *Agree*) usando os traços já presentes na estrutura, (v) valoração de D por avaliação do contexto extralinguístico para os traços que ainda não foram valorados e (vi) valoração automática por concordância (operação *Agree*) dos traços remanescentes dentro da cadeia de correspondência já estabelecida.

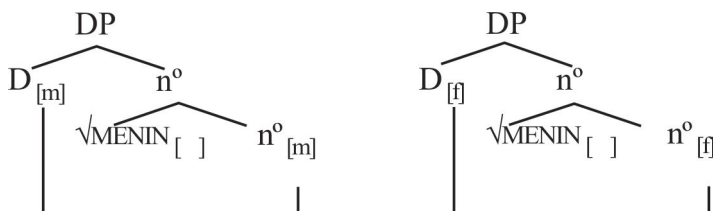
Assim, no caso de, por exemplo, *menino/menina*, a raiz $\sqrt{\text{MENIN}}$ se concatena com um nominalizador do tipo $n^\circ_{[]}$. Após D se concatenar à estrutura e combinar seus traços- ϕ com os traços- ϕ de no, D avalia o contexto extralinguístico para valorar o traço não valorado de gênero que, após essa valoração via contexto extralinguístico, é automaticamente valorado pela operação de *Agree* em n° . Após valorados todos os traços, a estrutura é encaminhada para PF, onde

têm vez as operações morfológicas, como regras de redundância e a operação de inserção de Vocabulário, que operam de forma indistinta tanto nos casos em que o gênero é interpretável quanto nos casos em que não é.

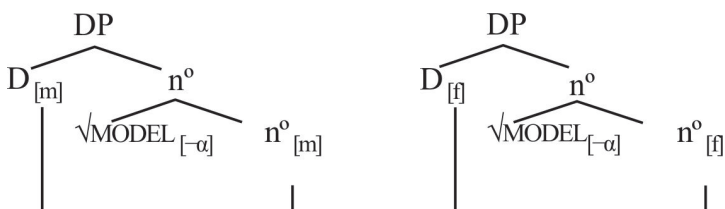
Assim, da mesma forma como ocorre com os pares de nome vistos em (19), espera-se que pares de nomes que compartilham a mesma raiz e cuja valoração de gênero se dá em D também apresentem os mesmos padrões: aqueles que se conformam com o padrão geral (masculino em -o, feminino em -a) e aqueles que, por terem a raiz especificada para classe, são invariáveis em forma, o que pode ser visto em (20).

(20) Pares variáveis com gênero interpretável

- (a) Pares masculinos em -o e femininos em -a (“menino” e “menina”)



- (b) Pares invariáveis (“(o) modelo” e “(a) modelo”)



Um último padrão a ser explicado é o de pares como *mestre-mestra*, em que a forma masculina pertence à classe mais marcada /-e ~ ø/ e apenas a forma feminina se adequa ao padrão geral. Embora sejam poucos os pares que seguem esse padrão, eles devem ser explicados. Sugerimos, conforme faz Harris (1999), que o mesmo tipo de regra de empobrecimento – utilizado para explicar a neutralização de conjugações nas formas de particípio – proposto para verbos esteja em jogo nos nomes. Particularmente, uma regra que apaga o traço [+β] no contexto de um nominalizador feminino concatenado a certas raízes é responsável por reencaminhar essas raízes para a classe [+α], conforme (21).

(21) Regra de empobrecimento

[+β] ↔ ø / X_[f], em que X = √_{MESTR}, √_{CHEF}, [n^o, -or]...

Nesta seção, buscou-se mostrar que, ao retirar o traço de gênero da raiz e alocá-lo no nominalizador, é possível capturar o mesmo conjunto de dados capturado por Alcântara (2010), mas evitando a proliferação de raízes devido à assunção de que gênero é propriedade dos nominalizadores. Ao mesmo tempo, ao assumir que traços de classe são um primitivo da gramática, eliminam-se os problemas da realização morfofonológica de gênero enfrentada pelos itens de Vocabulário propostos por Armelin (2015). Finalmente, esta proposta, que mescla intuições das duas autoras, busca capturar a diferença entre a aparente visibilidade dos traços de gênero em LF e a aparente invisibilidade dos traços de classe – tanto nominal quanto verbal – no mesmo componente. Na próxima seção, exploramos as consequências da proposta para a interpretabilidade do gênero.

4. Raízes e gênero: interpretação e interpretabilidade

Na seção anterior, foi mostrado como a presente proposta lida com a relação entre as raízes e a informação de classe (que tange à realização das vogais temáticas nominais e verbais) e, adicionalmente, com o traço de gênero no caso dos nomes. No que diz respeito às informações de classe, buscamos mostrar que essa é uma informação intrínseca às raízes e que só é relevante em PF, uma vez que essas informações não só não têm relevância para a computação sintática, como também não parecem estar disponíveis para a interpretação dos vocábulos em LF. Assim, só entram na computação posteriormente à derivação sintática, para dar conta da boa formação morfofonológica das palavras.

Diferentemente, a informação de gênero – propriedade do nominalizador, conforme a presente análise – está ativa na computação sintática e, logo, é relevante para a interpretação em LF. Sendo esse o caso, o objetivo desta seção é tecer algumas considerações sobre a interpretação dessas estruturas pela Enciclopédia e, além disso, explicitar como a presente análise lida com a questão da interpretabilidade de gênero e com alguns casos de coerção discutidos na literatura.

Marantz (1996) entende que a Enciclopédia relaciona as peças do output da derivação sintática aos seus significados não composicionais extralinguísticos; por exemplo, da mesma forma que no Vocabulário há uma entrada que relaciona a raiz $\sqrt{\text{FAC}}$, de *faca*, ao seu expoente fonológico, na Enciclopédia há uma entrada que relaciona essa mesma raiz ao seu conteúdo semântico extralinguístico; para Marantz, a interpretação de uma dada estrutura depende do contexto sintático em que ela ocorre. Assim, o mecanismo de interpretação de, por exemplo, *faca* e *carro* seria como ilustrado em (22).

(22) Instruções na Enciclopédia

- (a) $\sqrt{\text{FAC}} \leftrightarrow$ “utensílio de cozinha utilizado para cortar” / [[—] $\sqrt{\text{n}^\circ[\text{FEM}]}$]
- (b) $\sqrt{\text{CARR}} \leftrightarrow$ “veículo de quatro rodas movido a combustível” / [[—] $\sqrt{\text{n}^\circ[\text{MASC}]}$]

O que os exemplos em (22) esquematizam é que a entrada enciclopédica de uma dada raiz contém instruções para a sua interpretação, a saber, a raiz $\sqrt{\text{FAC}}$ no contexto de um nominalizador com o traço feminino é interpretada como sendo um utensílio de cozinha utilizado para cortar; da mesma forma, a raiz $\sqrt{\text{CARR}}$ no contexto de um nominalizador com o traço masculino é interpretada como um automóvel de quatro rodas movido a combustível.

No caso de raízes como $\sqrt{\text{FAC}}$, o fato de não haver na Enciclopédia uma entrada para sua interpretação no contexto de um nominalizador masculino (gerando **faco*) é tão acidental quanto não haver uma entrada para sua interpretação em contexto verbal (gerando **facar*); entretanto, existe a possibilidade para tal entrada e é, de fato, o que ocorre no caso de raízes que formam pares masculino/feminino não interpretáveis. Dito de outro modo, o mesmo procedimento se aplica a quaisquer casos em que há especificação de gênero no nominalizador, mesmo para raízes que possuem diferentes interpretações a depender da especificação de gênero contida no no com que ela se concatena. Por exemplo, como mostrado em (19a), a raiz $\sqrt{\text{MAT}}$ pode se concatenar tanto a um nominalizador com o traço de gênero masculino quanto a um nominalizador com o traço de gênero feminino. Isso equivale a dizer que, na Enciclopédia, haverá duas entradas diferentes para determinar a interpretação da raiz em cada um dos contextos, como esquematizado em (23).

(23) Instruções na Enciclopédia para a raiz $\sqrt{\text{MAT}}$

- (a) $\sqrt{\text{MAT}} \leftrightarrow$ “vegetação constituída por plantas não cultivadas” / [[—] $\sqrt{\text{n}^\circ[\text{MASC}]}$]
- (b) $\sqrt{\text{MAT}} \leftrightarrow$ “área coberta por plantas silvestres de tipos variados” / [[—] $\sqrt{\text{n}^\circ[\text{FEM}]}$]

Nesse sentido, seguindo Marantz (1996), uma entrada enciclopédica contém as instruções para a interpretação de uma dada raiz em um dado contexto, fornecendo o seu conteúdo semântico não composicional e, nesse caso, “não composicional” se refere ao fato de que a especificação de gênero no nominalizador não contribui semanticamente para o significado global do que o nome denota, apenas atua como contexto sintático para a sua interpretação. Um outro exemplo disso é o da expressão *chutar o balde*, em que *balde* não ‘compõe’ o significado de “perder o controle”, apenas serve de contexto sintático para uma interpretação não composicional da raiz $\sqrt{\text{CHUT}}$. Na presente proposta, então, gênero não interpretável é aquele em que o traço de gênero não compõe o significado do nome, apenas serve de contexto gramatical para a sua interpretação na Enciclopédia; logo, é “interpretável”, embora não composicional (cf. ARMELIN, 2015)¹⁰.

No entanto, existem casos em que o significado da estrutura [raiz + informação de gênero] é composicional, como pode ser visto no contraste entre *menino* e *menina*. Na seção anterior, foi mostrado

¹⁰ Contudo, mesmo nesses casos, podem-se encontrar situações em que o gênero não interpretável é tangencial ou metaforicamente *interpretável*. Por exemplo, em contextos de fábula, falantes nativos do PB tendem a antropomorfizar *lua* como feminino e *sol* como masculino; diferentemente, falantes nativos do alemão parecem fazer uma antropomorfização na direção inversa: *lua* como masculino e *sol* como feminino. Ainda que em um nível especulativo, isso pode ser reflexo da marcação de gênero gramatical que as palavras que nomeiam esses astros têm nessas línguas – em alemão, *die Sonne* (‘a sol’) e *der Mond* (‘o lua’). Assim, a interpretabilidade “literal” da informação de gênero gramatical para seres inanimados, na ausência de outras evidências ou de outras explicações mais sistemáticas, dá suporte à presente análise.

que a derivação desses nomes envolve a concatenação da raiz $\sqrt{\text{MENIN}}$ a um nominalizador não especificado para gênero, e que o gênero será valorado em D por meio da avaliação do contexto (KUČEROVÁ, no prelo). Adicionalmente, como já mencionado, nós estamos assumindo que as raízes são portadoras de um conteúdo semântico mínimo e, no caso da raiz $\sqrt{\text{MENIN}}$, vamos supor que seja algo como “criança”; logo, a concatenação de “criança” ao lado de “feminino”, por meio da atribuição do traço feminino via avaliação do contexto extralinguístico gera composicionalmente o significado de “menina”.

Interessante notar, contudo, que esse mecanismo de interpretação não está ligado à expressão morfofonológica do nome, ou seja, à oposição entre *-o* e *-a*. Afinal, como já mostrado, o nome *modelo* é invariável, mas o fato de o nominalizador vir não marcado para gênero, impelindo à valoração por meio da avaliação do contexto extralinguístico, leva à interpretabilidade do gênero – e, nessa perspectiva, a um significado composicional.

Adicionalmente, casos como *homem* e *mulher*, que denotam seres com gênero biológico, não constituem exemplos de gênero interpretável, embora o gênero gramatical coincida com o gênero biológico, já que a informação interpretável como “gênero biológico” não se dá por meio de uma valoração de D a partir do contexto extralinguístico, uma vez que o significado relativo ao gênero biológico é parte do significado da raiz. Nesse sentido, embora a palavra *mulher* seja feminina e ao mesmo tempo se refira a seres do sexo feminino, essa relação, segundo nossa proposta, é apenas coincidente¹¹.

¹¹ A respeito disso, Legate (2002) entende que “considerações de aquisição e talvez de diacronia vão favorecer a compatibilidade entre traços semânticos e morfossintáticos, mas isso será uma tendência, e não uma imposição do sistema, que permite discrepâncias”. Evidência disso é encontrada em alemão, no nome *Mädchen* (‘menina’) que, embora denote um indivíduo com gênero biológico feminino, possui um gênero gramatical neutro. Além disso, no próprio PB há a palavra “bicha” que, embora seja feminina, se refere exclusivamente a seres do sexo masculino – que são homossexuais.

Assim, se a Enciclopédia for o lugar para relacionar as estruturas ao(s) seu(s) significado(s) não composicional(is), não há necessidade de os traços de gênero associados a $\sqrt{\text{MENIN}}$ virem listados na Enciclopédia, já que suas derivações têm um significado composicional, qual seja, raiz + informação de gênero. Além disso, a ideia de que a interpretabilidade de gênero vem da não especificação para gênero no nominalizador seguida da avaliação do contexto extralinguístico em D – núcleo em que, por hipótese, a possibilidade de significados não composicionais não está mais disponível – captura mais um fenômeno encontrado em português.

A oposição entre masculino e feminino pode ser dependente do contexto extralinguístico¹² mesmo para raízes que não denotam entidades com sexo biológico; por exemplo, em fábulas ou histórias de ficção, os objetos assumam propriedades e/ou ações atribuídas a humanos, é possível que se profira a sentença *o garrafo e a garrafa se casaram* – como observado por Armelin (2015). Nesse contexto, a oposição de gênero torna-se ‘relevante’ e, conforme a nossa proposta, isso significa que a raiz $\sqrt{\text{GARRAF}}$ se concatenou a um nominalizador não especificado para gênero – cuja valoração se deu contextualmente.

De todo modo, é sabido que a interpretação default de *garrafa* se dá no contexto de um nominalizador especificado com o traço feminino, logo, na Enciclopédia, há uma entrada com a instrução de que a raiz $\sqrt{\text{GARRAF}}$ no contexto de um nominalizador feminino significa “recipiente de gargalo e boca estreitos, utilizado para armazenar líquidos”. Vale a pena notar que a forma *garrafo* só ocorre quando for condicionada pelo contexto extralinguístico; do contrário, a derivação implode, dado que, na Enciclopédia, só há instrução para a interpretação de *garrafa* no contexto de um nominalizador com a especificação de gênero feminino.

¹² Haspelmath & Sims (2011) fazem referência à “relevância pragmática da marcação de gênero”, isto é, a situações em que o contexto extralinguístico condiciona uma marcação de gênero não default; para esses autores, trata-se de casos como *mosquita* ou *formigo*, para os quais a oposição de gênero é normalmente não relevante, diferentemente de *cachorra* ou *gata*.

Além disso, no caso de *garrafo/garrafa*, vê-se a ocorrência das vogais default. Armelin (2015) atenta para o fato de que há a possibilidade de se proferirem sentenças como *a garrafa e o garrafo se casaram*, em que a vogal final é a mesma para os dois valores de gênero. A autora atribui esse uso a uma discrepância de gênero entre D e o nome (ou seja, em *o garrafo*, o artigo *o* está realizando um traço de gênero masculino, enquanto a vogal temática *-a* está realizando um traço de gênero feminino), ao passo que afirma que em *o garrafo* tal discrepância não ocorre (tanto determinante quanto vogal temática realizam um traço masculino).

Entendendo que se esperaria que houvesse consequências para a interpretação caso *o garrafo* e *o garrafa* tivessem configurações de traços de gênero diferentes, o que não parece ocorrer, nós propomos que a diferença entre *o garrafo* e *o garrafa* está restrita à informação de classe: ao proferir *o garrafo* e *a garrafa*, o falante está considerando que a raiz $\sqrt{\text{GARRAF}}$ não vem marcada com nenhum traço de classe (da mesma forma que $\sqrt{\text{MENIN}}$), o que faz emergir as vogais default quando do uso coercivo da mudança de gênero. Quando, por outro lado, o falante profere *o garrafa* e *a garrafa*, entende que a raiz $\sqrt{\text{GARRAF}}$ vem marcada para Classe II (ou $[\alpha]$, cf. § 1.2, 2 e 3), o que ocasiona a manutenção da vogal final (conforme ocorre em *o, a artista*). As estruturas em (24) ilustram a diferença entre *o garrafo* e *o garrafa*.

(24) Estruturas de “o garrafo” e “o garrafa”

- (a) “o garrafo”: $[D_{[\text{MASC}]} [n^{\circ}_{[]} [\sqrt{\text{GARRAF}}_{[]}]]]$
 (b) “o garrafa”: $[D_{[\text{MASC}]} [n^{\circ}_{[]} [\sqrt{\text{GARRAF}}_{[\alpha]}]]]$

A esta altura, é importante resgatar a questão mencionada na seção 2 a respeito da interpretação default do gênero masculino. Nossa proposta é a de que, quando a avaliação do contexto extralinguístico é desnecessária – ou seja, quando a intenção do falante é se referir à

propriedade descrita pelo nome independentemente do sexo dos referentes ou, ainda, quando no contexto extralinguístico há seres de ambos os sexos –, o traço de gênero se mantém subespecificado. Como o item de Vocabulário menos marcado é o item *-o*, na ausência de quaisquer outras informações (gênero ou classe), é esse o item inserido. Nesse sentido, embora morfofonologicamente seja acurado dizer que o gênero masculino coincide com a leitura ampla, em termos de traços, a leitura ampla é capturada pela ausência de traços de gênero, e não pelo traço de gênero masculino. A coincidência na morfofonologia desses dois conjuntos de traços decorre da subespecificação dos itens de Vocabulário.

Uma evidência de que a interpretação ampla de gênero decorre da ausência de traços de gênero, e não da presença de um dos traços de gênero especificamente marcado pode ser conferida na comparação das sentenças em (25).

- (25) (a) O garrafo e a garrafa se casaram e, quando eles se beijaram, todas as garrafas aplaudiram.
- (b) O garrafo e a garrafa se casaram e, quando eles se beijaram, todos os garrafos aplaudiram.

Nos exemplos em (25), há um contraste entre as leituras disparadas por *garrafas* e *garrafos*, a saber, em (25b) a única interpretação disponível é aquela em que apenas as garrafas “do sexo masculino” aplaudiram, o que é capturado pela não especificação de gênero no nominalizador que faz com que o traço masculino seja valorado por avaliação do contexto extralinguístico, gerando um significado composicional. No entanto, quando houver a especificação feminina para o traço, a estrutura será enviada para LF, componente no qual, há uma entrada que dá instruções para a interpretação de $\sqrt{\text{GARRAF}}$ no contexto de um nominalizador feminino, ou seja, uma interpretação não composicional, o que faz com que *garrafas* tenha, no caso de coerção, uma interpretação default, não especificada para gênero.

Diferentemente do fenômeno recém abordado, em que há uma possibilidade de mudança da vogal final quando da mudança coerciva de gênero (a que atribuímos à maneira como o falante codifica a relação entre a raiz e o traço de classe), há um caso de discrepância entre o gênero biológico do referente e gênero gramatical, discutido por Lazzarini-Cyrino, Armelin & Minussi (2015), em que essa mudança parece não ser possível. Esses autores abordam, igualmente operando com o modelo da Morfologia Distribuída, o licenciamento de sentenças como (26) e (27). Crucialmente, no entanto, não mencionam que as sentenças em (28) parecem não ser possíveis.

(26) (a) A garrafa está na minha casa.

(b) O Garrafa está na minha casa.

(27) (a) A bola está lá no quintal.

(b) O Bola está lá no quintal.

(28) (a) *O Garrafo está na minha casa.

(b) *O Bolo está lá no quintal. (de *bola*)

Para os autores, casos como (26b) e (27b), em que há discrepância de gênero, ou seja, há ocorrência do gênero masculino (no determinante) com uma palavra feminina, não poderiam ser explicados por propostas que se valessem de diacríticos de gênero para as raízes ou de condições de licenciamento, pois dada a ausência de instruções de interpretação na Enciclopédia, (26b) e (27b) seriam sentenças agramaticais, o que não parece ser o caso do PB. Assim, conforme os autores, a discrepância de gênero leva à atribuição de animacidade ao nome relevante; mais especificamente, segundo eles, o licenciamento do gênero é um epifenômeno de interpretação enciclopédica, isto é, as informações de gênero devem ser interpretadas como conjuntos na Enciclopédia.

No caso do PB, estes seriam dois: um de entidades masculinas e outro de entidades femininas. Sempre que uma associação estrutural entre uma raiz e os traços de gênero é enviada para a Enciclopédia, ocorre uma checagem, cujo papel é verificar se a entidade referida está incluída no conjunto de gênero correspondente. Portanto, na proposta desses autores, quando há incompatibilidade entre o gênero gerado na estrutura sintática e o gênero esperado na Enciclopédia, a associação é considerada falsa, levando a uma violação da máxima da qualidade – nos termos de Grice (1975). Tal violação desencadeia uma operação de resgate semântico que é, por exemplo, a mesma disparada na interpretação de metáforas ou expressões idiomáticas.

Entretanto, ao contrário do que propõem os autores, em casos como (26b) e (27b) não se parece estar diante de mudança de (apenas) animacidade. Em primeiro lugar, porque, como mostrado anteriormente, a interpretabilidade do gênero pode ser capturada pela não especificação do traço de gênero, fazendo com que mesmo nomes como *garrafa* recebam – por meio de uma varredura contextual – uma interpretação animada; em segundo lugar, porque esse tipo de “coerção” não está restrito a casos isolados de discrepância de gênero, mas sim a qualquer nome comum (ou adjetivo) que pode ocorrer como nome próprio, tais como *Rosa*, *Esmeralda*, *Violeta*, *Linda* etc.

Nesse sentido, nós defendemos que com casos como *o Bola* ou *o Garrafa* (que teriam uma interpretação de nome próprio) está-se diante de um tipo de “coerção” não de [–ANIMADO] para [+ANIMADO], mas sim de [–REFERENCIAL] (ou descritivo) para [+REFERENCIAL]. A respeito disso, Saab (2016) entende que, dentro dos contextos sintáticos apropriados, uma mesma raiz pode ter tanto uma interpretação referencial quanto uma interpretação descritiva¹³, como é o caso de nomes próprios que funcionam como predicadores, o que pode ser visto em (29).

¹³ “Predicativa”, nos termos do autor.

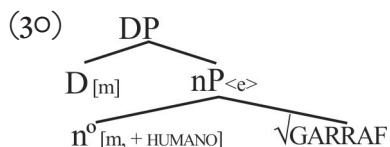
- (29) (a) Há duas Marias na minha turma de Linguística Geral.
(b) Nunca vi um João que seja um bom professor.

O que as sentenças em (29) mostram é que os nomes próprios, *Maria* e *João*, não possuem uma interpretação referencial, ou seja, não referem indivíduos específicos no mundo; pelo contrário, disparam uma leitura descritiva, no sentido de que descrevem um predicado que poderia ser parafraseado por “ser chamado de”. Então, a sentença (29a) é interpretada como *há dois indivíduos com a propriedade de ser chamado Maria na minha turma de Linguística Geral*, e a sentença (29b), *nunca vi um indivíduo com a propriedade de ser chamado de João que seja um bom professor*. Nesse sentido, a interpretação de (27b) poderia ser algo como *um indivíduo que tem a propriedade de ser chamado Bola está na minha casa*.

Vale a pena notar que, conforme a nossa proposta, quando um nome sofre “coerção” de animacidade, o predicado descrito pelo nome não sofre alterações; no caso de *o garrafo e a garrafa se casaram* ainda que os nomes passem a ser interpretados como sendo animados, eles não perdem a propriedade de ser garrafa (isto é, seu conteúdo descritivo). Diferentemente, em casos como *o Garrafa está na minha casa* ou *a Esmeralda chegou*, o nome perde totalmente a propriedade de ser garrafa/ser esmeralda, e isso vale para qualquer nome próprio; portanto, há evidências de que a coerção é, de fato, de [-referencial] para [+referencial] e, como nome referencial, perde sua leitura descritiva.

Além disso, diferentemente do que alegam Lazzarini-Cyrino, Armelin & Minussi (2015), a nossa proposta para a interpretação de gênero – a qual assume condições de licenciamento na Enciclopédia – pode dar conta desses casos como (26b) e (27b). Na esteira de Saab (2016), nós vamos assumir que, adicionalmente a um traço de gênero, o nominalizador pode carregar um traço [+HUMANO], e que a

presença desse traço faz com que o nome formado naquela estrutura seja interpretado referencialmente¹⁴, como ilustrado em (30).



Nesse sentido, assim como nós propomos que a não especificação para gênero no nominalizador faz com que D avalie o contexto extralinguístico para valorar o traço de gênero, nós vamos propor que sempre que houver um traço [+HUMANO] no nominalizador, a avaliação do contexto extralinguístico vai determinar também o valor da referência do nome: se for um indivíduo masculino, o traço de gênero em D será masculino; se o referente for feminino, a valoração de gênero em D será igualmente de feminino. Assim, nesse caso, como a valoração do traço de gênero é realizada contextualmente, desencadeada pelo traço [+HUMANO], não se espera que haja necessariamente coincidência entre a informação de gênero presente no nominalizador e a informação de gênero em D.

Portanto, retomando os exemplos já discutidos, a formação de *Garrafa* em *o Garrafa está na minha casa* envolve a concatenação da raiz √GARRAF a um nominalizador especificado com o gênero feminino, o qual vai gerar a forma feminina, e igualmente especificado com o traço [+HUMANO], que vai gerar a interpretação referencial. Assim, se o referente de *Garrafa* for masculino, em virtude da avaliação do contexto extralinguístico realizada por D, o traço de gênero especificado em D vai ser masculino, gerando a forma discrepante.

¹⁴ Além disso, na proposta do autor, a leitura descritiva/predicativa de um nome próprio como em (24) é capturada pela ausência do traço [HUMANO] EM ADIÇÃO A UMA PROJEÇÃO DE NÚMERO – QUE ALOCA O PLURAL, COMO PODE SER VISTO EM (24A).

Logo, esse é um caso que envolve, de fato, discrepância de gênero na estrutura, ou seja, em que não está valorado com um traço e D, com outro. Conforme já mencionado, a (não) ocorrência desse tipo de discrepância, por envolver traços sintáticos, implica em consequências para a interpretação. Por esse motivo, a interpretação referencial de nomes simples não permite formas como (28), em que a realização morfofonológica do nome muda juntamente com a variação no gênero do determinante, uma vez que o nome mantém a valoração de gênero conforme listado na Enciclopédia, porém com o acréscimo do traço de referencialidade, essa discrepância é, em última análise, licenciada. Desse modo, é possível capturar todos esses fenômenos estruturalmente, sem lançar mão das implicaturas, como fazem Lazzarini-Cyrino, Armelin & Minussi.

Ainda que a proposta de Saab (2016) postule um traço [humano] privativo, nós vamos estender sua análise e propor que esse traço seja binário, ou seja, [\pm HUMANO]. Assim, ao lado da especificação [+HUMANO] recém apresentada, a especificação [-HUMANO] desencadeará igualmente a interpretação referencial dos nomes gerados pela estrutura; porém, dado que são nomes com interpretação não humana, não haverá valoração de traço a partir do contexto extralinguístico, e o gênero desse tipo de nome próprio (quando houver¹⁵) vai ser aquele listado na Enciclopédia – ou seja, um gênero não interpretável – como é o caso de (o) *Brasil*, (a) *França*, (o) *Cristo Redentor*, (a) *Estátua da Liberdade* etc.

¹⁵ Sabe-se que alguns nomes próprios, tais como os que nomeiam cidades, estados e países, por exemplo, não ocorrem com determinantes, como é visto em (*o/*a) *Brasília*, (*o/*a) *Santa Catarina* e (*o/*a) *Portugal*, da mesma forma que alguns dialetos do português brasileiro não aceitam nomes determinantes co-ocorrendo com nomes próprios que nomeiam pessoas. Todavia, esse tipo de restrição sintática não obsta a proposta defendida neste trabalho e, por essa razão, ficará para trabalhos futuros.

5. Considerações finais

No presente trabalho, procurou-se explorar um sistema capaz de descrever a relação entre, de um lado, raízes, traços de classe e gênero e, de outro, gênero e interpretação, assumindo-se o quadro teórico da MD, que fornece às raízes um estatuto teórico particular. Buscamos motivar e mesclar diferentes aspectos de diferentes análises que contemplam esse tipo de fenômeno, a fim de chegar a um sistema que dê conta de diferentes conjuntos de dados. Conforme apresentadas na introdução, as nossas hipóteses iniciais eram:

H1 – Traços de classe são primitivos da gramática inerentes às raízes e computáveis apenas em PF (nunca podendo, portanto, ser a propriedade definidora de diferentes interpretações em LF);

H2 – Gênero é um traço do nominalizador, disponível em LF e o nominalizador vem em três diferentes matizes: $n^o_{[]}$, $n^o_{[FEM]}$ e $n^o_{[MASC]}$. Quando subespecificado, a valoração de gênero fica sob a responsabilidade de D, que faz uma avaliação do contexto extralinguístico;

H3 – Quando há uma avaliação do contexto extralinguístico para a valoração do traço de gênero, está-se diante do que se costuma chamar de “gênero interpretável”, e essa avaliação do contexto extralinguístico pode devolver o valor de gênero masculino, quando se está referindo a seres do sexo masculino, gênero feminino, quando se está referindo a seres do sexo feminino, ou ainda se abster de valorar, quando a especificação de gênero é irrelevante ou quando a avaliação falha em devolver um valor de gênero único.

Com a primeira hipótese, em certa medida implícita em propostas como a de Oltra-Massuet (1999), Harris (1999) e Alcântara (2003, 2010), buscou-se explicar o fato de que, segundo nosso levantamento de dados, parece não ser possível formar pares de

nomes ou pares de verbos que compartilham a mesma raiz e cuja única diferença reside na mudança de classe formal (sem haver, simultaneamente, mudança no valor do gênero, no caso dos nomes). Com a segunda hipótese, bastante articulada em Kučerová (no prelo), buscou-se explicar a diferença entre pares como *barco/barca* e *menino/menina*, que envolvem a mesma raiz, mas cujo valor de gênero resulta em diferentes tipos de interpretação (especificamente, não interpretável no primeiro par e interpretável no segundo).

Finalmente, a terceira hipótese não só articula melhor a interpretabilidade do gênero dito “semântico/interpretável”, mas também busca capturar a interpretação ampla do gênero masculino, bem como diferentes usos coercivos de gênero – particularmente, casos em que a mudança de gênero resulta na animacidade ou relevância pragmática de nomes codificados para um gênero único na gramática e casos em que a mudança de gênero resulta na interpretação referencial de um nome com função, em princípio, unicamente predicativa/descriptiva.

Tendo a consciência de que o presente trabalho se limitou ao domínio quase que exclusivo da palavra, esperamos que, em pesquisas futuras, o sistema proposto seja avaliado com relação à sua adequação para fenômenos sintáticos que envolvem, por exemplo, concordância, elipse – sobretudo no que toca à identidade de gênero entre o antecedente e constituinte elidido – e para fenômenos morfológicos adicionais da marcação de gênero, como aqueles que envolvem sufixos derivacionais.

Agradecimentos

Agradecemos ao CNPq pelo auxílio (processo 424025/2016-7), concedido pelo Edital Universal 01/2016.

Agradecemos aos dois pareceristas anônimos pelas sugestões e comentários que contribuíram para o formato final do presente

trabalho, ainda que nem todas as sugestões puderam ser contempladas. Qualquer problema ou erro que tenha eventualmente permanecido é de nossa inteira responsabilidade.

Referências

ACQUAVIVA, Paolo. Roots and lexicality in Distributed Morphology. In: GALANI, Alexandra; REDINGER, Daniel; YEO, Norman (Eds.). *YPL2, issue 10 – York Essex Morphology Meeting*. University of York, 2009.

ALCÂNTARA, Cíntia da Costa. *As classes formais do português e sua constituição: um estudo à luz da teoria da Morfologia Distribuída*. 2003. 204f. Tese (Doutorado em Linguística) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre.

ALCÂNTARA, Cíntia da Costa. As classes formais do português brasileiro. *Letras de Hoje*. Porto Alegre, v. 45, n. 1, p. 5-15, 2010.

ARAD, Maia. Locality constraints on the interpretation of roots. *Natural Language and Linguistic Theory*, v. 21, p. 737-778, 2003.

ARMELIN, Paula Roberta Gabbai. *A relação entre gênero e morfologia avaliativa nos nominais do português brasileiro: uma abordagem sintática da formação de palavras*. 2015. 247f. Tese (Doutorado em Linguística) – Universidade de São Paulo, São Paulo.

ATKINSON, Emily. Gender features on *n* & the root: An account of gender in French. In: SMITH, Jason; IHSANE, Tabea (Eds.). *Romance linguistics 2012: Selected papers from the 42nd linguistic symposium on romance languages*. Amsterdam: Benjamins. p. 229-244, 2015.

BERMÚDEZ-OTERO, Ricardo. The Spanish lexicon stores stems with theme vowels, not roots with inflectional class features. *Probus*, v. 25, n. 1, p. 3-103, 2013.

BOBALJIK, Jonathan David; ZOCCA, Cynthia Levart. Gender markedness: the anatomy of a counterexample. *Morphology*, v. 21, n. 2, p. 141-166, 2011.

BONET, Eulalia Alsina. *Morphology after syntax*: pronominal clitics in Romance. 1991. 240f. Tese (Doutorado em Linguística) – Instituto de Tecnologia de Massachussts, Cambridge.

CAMARAJR, Joaquim Mattoso. *Estrutura da língua portuguesa*. 27. ed. Petrópolis: Vozes, 2007 [1970].

GRICE, Paul. Lógica e conversação. In: DASCAL, Marcelo (org.). *Fundamentos metodológicos da linguística*: pragmática. Edição do organizador: Campinas, 1982 [1975]. p. 81-103. v. 4.

HALLE, Morris. Distributed Morphology: impoverishment and fission. *MIT Working papers in Linguistics*. Cambridge, v. 30, p. 425-449, 1997.

HALLE, Morris; MARANTZ, Alec. Distributed morphology and the pieces of inflection. In: HALE, Keneeth; KEYSER, Jay (Eds.). *View from the word building* 20. Cambridge: MIT Press. p. 111-176, 1993.

HALLE, Morris; MARANTZ, Alec. Some key features of Distributed Morphology. *MIT Working papers in Linguistics*: papers in Phonology and Morphology. Cambridge, v. 21, p. 275-288, 1994.

HARLEY, Heidi. On the identity of roots. *Theoretical linguistics*, v. 40, n. 3-4, p. 225-276, 2014.

HARRIS, James. Nasal depalatalization no, morphological well-formedness si: the structure of Spanish word classes. *MIT Working papers in Linguistics*, Cambridge, v. 33, p. 47-82, 1999.

HASPELMATH, Martin; SIMS, Andrea. *Understanding morphology*. 2. ed. Londres: Hodder Education, 2010.

KUČEROVÁ, Ivona. Φ -features at the syntax-semantics interface: evidence from nominal inflection. *Linguistic Inquiry*, Cambridge, v. 49, n. 4, p. 813-845, 2018.

LAZZARINI-CYRINO, João Paulo; ARMELIN, Paula; MINUSSI, Rafael. On the encyclopedic knowledge of gender. *MMM9 Online proceedings*, p. 77-87, 2015.

LEGATE, Julie Anne. *Phases in “Beyond Explanatory Adequacy”*. Manuscript, MIT, Cambridge, 2002.

LEMLE, Miriam. *Arbitrariedade saussuriana: saltos e sobressaltos*. 2008. 32f.

MARANTZ, Alec. ‘Cat’ as a phrasal idiom: consequences of late insertion in Distributed Morphology. 1996. 25f.

MARANTZ, Alec. Sem escapatória da sintaxe: não tente fazer análise morfológica na privacidade do seu próprio léxico. *ReVEL*, Porto Alegre, v. 13, n. 24, p. 8-33, 2015 [1997].

MARANTZ, Alec. *Words*. 2001. 29f.

MINUSSI, Rafael; BASSANI, Indaiá. Em favor do conteúdo semântico das raízes. *Letras*, Curitiba, n. 96, p. 152-173, 2017.

OLTRA-MASSUET, Isabel. *On the notion of theme vowel: a new approach to Catalan verbal morphology*. 1999. 89f. Tese (Doutorado em Linguística) – Instituto de Tecnologia de Massachusetts, Cambridge.

PANAGIOTIDIS, Phoevos. *(Gramatical) gender troubles and the gender of pronouns* (no prelo).

RESENDE, Maurício Sartori. Competência morfológica e gênese histórica: limites entre diacronia e sincronia. *Filologia e Linguística portuguesa*, São Paulo, v. 20, n. 1, p. 61-76, 2018.

RESENDE, Maurício Sartori. *Mudança semântica diacrônica no domínio intravocabular: o caso das raízes cranberry do português* (no prelo).

RITTER, Elizabeth. Where’s gender? *Linguistic Inquiry*, v. 24, n. 4. p. 795-803, 1993.

ROCHA, Luiz Carlos Assis. *Estruturas morfológicas do português*. 2. ed. São Paulo: Martins Fontes, 2008 [1998].

SAAB, Andreas. *No name: the alossemy view*. 2016. 54f.

SANTANA, Beatriz Pires. Os futuros do indicativo: por uma análise sintática para a flexão verbal do português brasileiro. *Caderno de Squibs*, Brasília, v. 2, n. 1, p. 43-53, 2016.

SANTANA, Beatriz Pires. Morfologia Ornamental: as vogais temáticas do português brasileiro. Tese (Doutorado em Linguística) – Universidade Federal do Paraná, 2019.

TEIXEIRA, Gabriel Sanches. *As formas verbais regulares e simples do português brasileiro: uma proposta à luz da Morfologia Distribuída*. 2012. 141f. Tese (Doutorado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis.

ZOCCA, Cynthia Levar. *O que não está lá? Um estudo sobre morfologia flexional em elipses*. 2003. 175f. Dissertação (Mestrado em Linguística) – Universidade Estadual de Campinas, Campinas.

Como citar

RESENDE, Maurício; SANTANA, Beatriz Pires. A relação entre raízes, gênero, classe e significado. *Revista da Abralin*, v. , n. , p. 01-55, 2019.

Recebido em 01/05/2019 e aceito em 13/08/2019