

DANDO NOME AOS NOMES: NOTAS SOBRE A FORMA E O SIGNIFICADO DOS NOMES PRÓPRIOS

HOW NOUNS GET THEIR NAMES: REMARKS ON THE FORM AND THE MEANING OF PROPER NAMES

Maurício Resende¹ (UNICAMP)

RESUMO

O presente artigo trata dos problemas estritamente linguísticos que subjazem à caracterização dos mecanismos de derivação e dos diferentes empregos dos nomes próprios. Primeiramente, este artigo mostra que há diferentes tipos de nomes próprios e que eles apresentam propriedades distintas. Em segundo lugar, o presente trabalho levanta uma série de questões empíricas (sobretudo, morfológicas, sintáticas e semânticas) que põe em xeque a abordagem/assunção de tratar os nomes próprios como entradas lexicais. Por fim, este estudo propõe um tratamento sintático para a formação de palavras, assumindo o quadro da Morfologia Distribuída (HALLE & MARANTZ, 1993), e defende que os fatos empíricos levantados são descritivamente mais bem explicados à luz desse modelo.

PALAVRAS-CHAVE: Nomes próprios; Morfologia Distribuída, Formação de palavras.

ABSTRACT

This paper deals with strictly linguist issues underlying the description of proper names in relation to both their mechanisms of formation and their different uses. Firstly the present work shows that there are different types of proper names and these exhibit different properties. Secondly, this paper raises a range of empirical issues (mostly morphological, syntactic, and semantic ones), which challenges the approach/assumption of treating proper names as lexical entries. Finally, this case-study proposes a syntactic treatment for word formation, assuming Distributed Morphology framework (HALLE & MARANTZ, 1993) and argues that the raised empirical facts are descriptively better explained in the light of this model.

KEYWORDS: Proper names; Distributed Morphology; Word formation.

INTRODUÇÃO

Ainda que a caracterização dos NOMES PRÓPRIOS – ou, de maneira mais geral, das EXPRESSÕES REFERENCIAIS – tenha sido alvo de grande interesse por parte dos filósofos (da linguagem) desde o trabalho de Frege (1892), nos trabalhos em Linguística, esse fenômeno (ou esse objeto) recebeu muito menos atenção. Não há, na literatura sobre o português, uma vasta gama de trabalhos que investiguem, de maneira sistemática, as propriedades formais dos nomes próprios.

Por “não sistemático” se quer dizer que os poucos estudos que, de fato, discutem problemas diretamente relacionados à forma e à interpretação dos diferentes empregos dos nomes próprios estão, na verdade, se debruçando sobre outros fenômenos e tratam, então, os nomes próprios ou como problema residual ou como evidência independente para

¹ Gostaria de agradecer ao Renato Basso, por seus comentários e sugestões quando este trabalho era apenas um conjunto de ideias, e também ao Rodolfo Ilari, por todas as conversas sobre Frege, que me fizeram ter uma ideia mais clara do problema semântico. Um agradecimento especial ainda a um dos pareceristas anônimos por suas perguntas e observações, que ajudaram levar este artigo à sua versão final. Finalmente, agradeço ao CNPq por ter financiado a minha pesquisa de Doutorado, processo 141644/2016-8. Contato: mauricio_resende@hotmail.com

algum fenômeno diferente. Nesse sentido, os fenômenos subjacentes aos processos de formação e ao licenciamento dos diferentes empregos dessas expressões carecem de uma reflexão mais direcionada.

Isso quer dizer que, embora já se saiba alguma coisa sobre os nomes próprios – Longobardi (1994), Brito (2003), Matushansky (2008), Massani-Cagliari (2011), Saab & Lo Guercio (2019), Santana (2019) – é verdade que os mecanismos responsáveis pela sua formação ainda são pouco explorados – mas cf. Monteiro (2002) para algumas ideias. Seja como for, a “não sistematicidade” e/ou a “superficialidade” no estudo desse fenômeno têm a ver, sobretudo, com a ausência de respostas para algumas questões importantes.

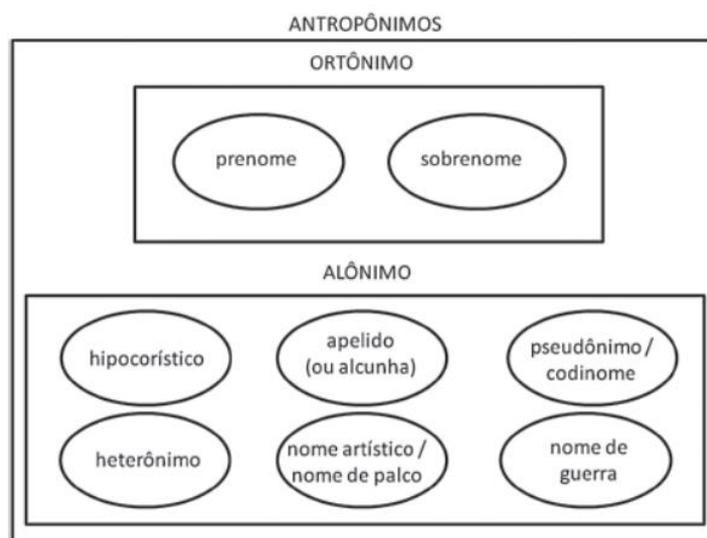
Mais especificamente, tais questões versam sobre (i) quais são as operações formais responsáveis pela derivação dos nomes próprios e em que medida elas diferem das operações que formam outras palavras da língua, (ii) o que há na composição dos nomes próprios que licencia sua leitura referencial e de que modo isso se relaciona aos seus outros empregos, (iii) como relacionar as características sintáticas e semânticas específicas a essa classe aos mecanismos formais responsáveis pela sua formação.

Ao formular o problema dessa maneira, através dessas três questões norteadoras, o presente artigo defende que os nomes próprios não têm nada de especial do ponto de vista formal (isto é, fonológico, morfológico e sintático) em relação a quaisquer outras “(sub)classes” de palavras da língua; suas particularidades residem no domínio da semântica – especificamente, na sua leitura referencial –, e mesmo nesses casos, é possível derivar esse tipo de leitura a partir de traços formais, que estão presentes em sua composição.

Dadas essas considerações, o objetivo central deste artigo não é o de propriamente apresentar uma proposta de como tratar os fenômenos que subjazem à ocorrência dos nomes próprios, mas antes o de levantar quais são os problemas semânticos, sintáticos e morfológicos relevantes para a caracterização do problema empírico e, adicionalmente, mostrar de que forma eles se relacionam, tendo em vista aquilo que já aparece discutido na literatura. Antes disso, porém, convém tecer algumas considerações sobre o que, de fato, são – ou pelo menos, o que está sendo entendido como – nomes próprios.

A maioria dos trabalhos (sobre o português) cujo foco são os nomes próprios se debruça, via de regra, sobre sua classificação. Amaral (2011), por exemplo, entende que os nomes próprios constituem uma classe bastante heterogênea e, a rigor, dizem respeito a ANTROPÔNIMOS (nomes de pessoas), TOPÔNIMOS (nomes de lugares) e outros, como os nomes de organizações sociais, títulos de obras, marcas etc. Pode-se acrescentar ainda a esse grupo, nomes de entidades abstratas, tais como *Tornado El Reno*, *Furacão Katrina*, *El Niño* etc. Amaral observa que os próprios antropônimos não constituem uma (sub)classe homogênea e, portanto, propõe uma tipologia para esses casos, a qual aparece reproduzida no Quadro 1 – retirado de Amaral (2011, p. 76).²

² Conforme a autora, “ortônimos” são nomes normais (ou comuns) de uma pessoa, e “alônimos” são os nomes não oficiais (ou “diferentes”).



Quadro 1: classificação dos antropônimos, segundo Amaral (2011, p. 76)

Não é objetivo deste trabalho discutir, corroborar ou contestar a tipologia proposta pela autora, a ideia é apenas mostrar que mesmo “nome próprio” ou “antropônimo” são rótulos bastante vagos para nomear essas expressões. Seja como for, o presente artigo se detém apenas nos antropônimos comuns (*João, Maria* etc.) e nos topônimos mais usuais, como nomes de cidades, estados e países (*Curitiba, Paraná, Brasil* etc.). A hipótese é a de que qualquer nome de entidade abstrata ou qualquer um dos subtipos dos antropônimos propostos por Amaral acomode as generalizações (formais) observadas.

Assim sendo, este artigo aparece dividido da seguinte maneira: a seção 1 apresenta a agenda de questões semânticas, sintáticas e morfológicas que subjazem aos diferentes empregos dos nomes próprios, reunindo inclusive tópicos já levantados na literatura, e a seção 2 defende que todos esses fenômenos podem ser acomodados em um modelo sintático para a formação de palavras, como a MORFOLOGIA DISTRIBUÍDA (HALLE & MARANTZ, 1993). Este trabalho explora uma visão ALOSSÊMICA das raízes (SAAB & LO GUERCIO, 2019) e propõe que os nomes próprios são formados na sintaxe e que suas diferentes leituras são o resultado de diferentes configurações estruturais e/ou conjuntos de traços.

1 Nomes próprios: a agenda de problemas formais

Do ponto de vista da descrição empírica e dos desafios que os nomes próprios colocam para a teoria linguística, indiscutivelmente, os fatos mais conhecidos jazem no domínio da semântica (e da ontologia). Desde o trabalho de Frege (1892), linguistas e filósofos (da linguagem) têm se debruçado sobre questões subjacentes à referencialidade, à definitude etc. Não é objetivo deste trabalho reproduzir ou alimentar essa discussão, mas é importante ter claro de que forma (e por que) certos empregos dos nomes próprios colocam problemas interessantes para uma teoria de gramática – cf. Brito (2003) e Matushansky (2008) para um panorama, discussão e referências; este mais voltado para questões estritamente linguísticas, aquele mais detido em problemas filosóficos. Para Frege (1892), assim como para a LÓGICA CLÁSSICA, os nomes próprios são CONSTANTES, ou seja, são unidades que denotam indivíduos e designam um único indivíduo no mundo (ou no contexto). Esse é o caso dos exemplos em (1).

- (1) (a) João é o irmão de Ana.
(b) Florianópolis fica entre Curitiba e Porto Alegre.

Em (1a), *João* denota um único indivíduo, João, assim como *Florianópolis* denota um único indivíduo em (1b), Florianópolis – o mesmo se aplica a *Ana*, *Curitiba* e *Porto Alegre*. Nesses casos, a referência é sempre a mesma – é constante – e toma, no mundo, o indivíduo de carne e osso ou o ponto exato da localização de um lugar específico. Por mais elementar e intuitiva que seja essa noção, alguns exemplos bastante corriqueiros colocam em xeque essa definição, como indicam as sentenças em (2).

- (2) (a) Harry Potter tem uma cicatriz em formato de raio.
(b) Papai Noel mora no Polo Norte.

Embora qualquer falante que conheça um pouco da narrativa de Harry Potter e/ou do Papai Noel tenda a classificar como verdadeiras as afirmações em (2), não é muito claro que se possa apontar no mundo (ou no contexto) algum indivíduo de carne e osso que se refira a esses nomes, ainda que seja possível fazer afirmações a respeito deles e, inclusive, atribuir a tais afirmações valores de verdade. De acordo com Frege (1892), esses são casos de NOMES PRÓPRIOS APARENTES e dizem respeito a nomes que não têm REFERÊNCIA (apenas SENTIDO).

A relação unívoca entre o emprego de um nome próprio e a designação direta a um indivíduo no mundo fica ainda mais enfraquecida com sentenças do tipo de (3). O fato de um predicado como *nunca existiu* ser compatível com um nome próprio deveria criar uma contradição, uma vez que se é o caso de um indivíduo nunca ter existido, não seria possível usar, à primeira vista, um nome próprio para se referir especificamente a ele – cf. Brito (2003) para uma discussão desses problemas e também para o problema de NOMES PRÓPRIOS COM REFERÊNCIA MÓVEL.

- (3) Moisés nunca existiu.

Esses fenômenos têm uma relação mais estreita com a ontologia, pois têm a ver com como o mundo é, com o que existe etc. Pondo de lado esses casos, em termos de descrição (e simplicidade) gramatical, seria possível apenas definir “nomes próprios” como um tipo de unidade linguística que tem conteúdo referencial (sem se comprometer com questões ontológicas e/ou metafísicas do que é “referencial”). Ainda assim, alguns empregos dos nomes próprios, como os que aparecem em (4), colocam em xeque a sua própria referencialidade.

- (4) (a) Todo Carlos é engraçado.
(b) Tem duas Marias na minha turma de Sintaxe.
(c) Nasce uma nova Madona em Nova Iorque.
(d) Lá vai o João da Silva dando uma de João da Silva.
(e) Encontrei o Pedro na rua hoje. – Quem? O Pedro *Pedro*?

É bastante claro que nenhum dos exemplos em (4) pode ser explicado unicamente lançando-se mão da noção de referencialidade dos nomes próprios. Nos exemplos (4a) e (4b), o nome próprio claramente está se comportando como um predicado – o predicado de “ser chamado de X”. A interpretação que *todo Carlos é engraçado* tem é que o conjunto dos indivíduos que têm a propriedade de ser chamado de “Carlos” está contido no conjunto de coisas que têm a propriedade de ser engraçado – isto é, $\forall x (\text{CARLOS}(x) \rightarrow \text{ENGRAÇADO}(x))$, para todo x , se x é Carlos, então x é engraçado. Seguindo o mesmo raciocínio, o que (4b) denota é que há na minha turma de Sintaxe dois indivíduos que têm a propriedade de ser chamado de “Maria” – $\text{MARIA} \cap \text{ESTAR NA TURMA DE SINTAXE} = 2$, ou seja, a

intersecção entre o conjunto “ser chamado de Maria” e o conjunto “estar na minha turma de Sintaxe” é igual a 2.

A intuição por trás de (4c) talvez seja menos simples de descrever. A ideia é a de que “Madona” não é unicamente uma propriedade, mas um conjunto de propriedades – tais como ser uma grande artista, causar polêmica etc. – e que o tal indivíduo que nasceu em Nova Iorque compartilha um certo conjunto de propriedades com Madona, *exceto* a de ser chamado de “Madona”, isto é, ao proferir (4c), o falante não quer dizer que nasceu um outro indivíduo que tem a propriedade de ser chamado de “Madona”, mas que nasceu um indivíduo que compartilha certas propriedades com outro indivíduo que tem a propriedade de ser chamado de “Madona”, ainda que o nome possa não ser (geralmente não é) um desses pontos de intersecção.

Em (4d), a situação parece ser mista: a primeira ocorrência está fazendo referência ao indivíduo João da Silva e a segunda, ao predicado (à propriedade ou ao conjunto de propriedades) de ser *João da Silva* que, assim como em (4c), não parece estar fazendo referência à propriedade de “ser chamado de João da Silva”, isto é, há uma leitura de que o indivíduo que tem a propriedade de ser chamado de “João da Silva” apresenta um certo conjunto de propriedades associada a esse nome. Uma vez que esses casos não são nem contraditórios nem redundantes, uma conclusão razoável seria a de que um nome próprio não está em uma relação unívoca e/ou direta com sua referência.

Finalmente, (4e) também parece apresentar um contexto misto: ao passo que no primeiro uso, *Pedro* se refere ao indivíduo de carne e osso; no segundo, parece que o falante está repetindo o nome como uma forma de confirmação da pressuposição de que é do mesmo indivíduo Pedro de que se está falando. A ideia de que os nomes próprios podem carregar pressuposições já aparece discutida na literatura – cf. Brito (2003) para detalhes – e versa sobre o fato de que, ao empregar um nome próprio, o falante assume – ou pressupõe – que o ouvinte sabe quem é o seu referente.

Discutir cada uma dessas questões com profundidade iria muito além do propósito deste trabalho. De todo modo, os exemplos em (4) servem para ilustrar o fato de que, independentemente dos problemas ontológicos que subjazem à interpretação dos nomes próprios e à determinação da sua referência – como exemplificado em (1), (2) e (3) – usos mais corriqueiros podem colocar problemas para o próprio tratamento desses itens em termos de competência linguística e licenciamento estrutural/semântico etc.

Dito de outro modo, o comportamento dos nomes próprios – ora como predicado ora como indivíduo³ – faz surgir a questão de qual deve ser o estatuto categorial desses itens e o que exatamente os licencia e/ou os distingue configuracionalmente. À primeira vista, uma hipótese plausível seria considerar os nomes próprios como entradas lexicais do tipo *João, Maria, Carlos* etc. – similarmente à proposta de Matushansky (2008). No entanto, como mostrado a seguir, além dos problemas semânticos, certos empregos dos nomes próprios atestam convincentemente contra o engessamento estrutural – ou seja, morfológico e sintático – dessas unidades, de forma que um tratamento em termos de RAÍZES (ANTROPONÍMICAS/TOPONÍMICAS) seja descritivamente mais adequado – cf. § 3.

Desde Longobardi (1994), os linguistas têm utilizado os nomes próprios tanto como evidência para uma certa configuração estrutural dentro do DP quanto para a postulação de certos movimentos sintáticos (com consequências semânticas, tais como a leitura de definitude). Especificamente, a ideia é a de que em *João saiu, João*, sendo um nome, ocupa a posição de N e se move para D, onde recebe (/checa) sua leitura definida. Isso levou alguns autores a defenderem que o artigo definido que ocorre com o nome próprio é expletivo, dada a ausência de contraste semântico entre *João saiu* e *o João saiu*, por exemplo.

³ Sem mencionar seu emprego como confirmador de uma pressuposição e/ou os usos em que o nome próprio parece denotar não um indivíduo ou uma propriedade (a de “ser chamado de”), mas um conjunto de propriedades.

A ocorrência de determinantes com nomes próprios é um fenômeno com muita variação em português. Há variação não somente entre o português brasileiro (PB) e o português europeu (PE) ou entre dialetos do PB,⁴ mas também existe variação interna, qual seja: ao passo que os antropônimos se combinam bem (a depender do dialeto) com artigos, nos topônimos, há uma subgeneralização de que nomes de países e estados se combinam com artigos, mas nomes de cidades não, como indicam os exemplos em (5).

- (5) (a) o Brasil, a Argentina, a França, a Alemanha, o Uruguai...
 (b) o Paraná, o Rio Grande do Norte, o Espírito Santo, a Bahia...
 (b) *a Curitiba, *a Florianópolis, *o Porto Alegre, *o São Paulo...

Há exceções para essas subgeneralizações, tais como **o/a Cuba*, **o/a Portugal* (para países); **o/a Santa Catarina*, **o/a São Paulo* (para estados); *o Rio de Janeiro* (para cidades). A despeito de haver alguns casos excepcionais, pode-se observar que, nos topônimos, a combinação com artigo é muito mais rígida e, claramente, o sistema deve ser capaz de diferenciar cidades, de estados e países, de antropônimos. Ainda no domínio da sintaxe dos topônimos, dada a incompatibilidade dos nomes de cidades (e de alguns países e estados) com artigos, a pergunta que emerge é de qual informação a gramática dispõe para estabelecer a relação de concordância com gênero, como nos exemplos em (6).

- (6) (a) Fortaleza foi invadida/*invadido pelos guerrilheiros.
 (b) Brasília foi saqueada/*saqueado pelos corruptos.
 (c) Porto Alegre foi colonizado/?colonizada pelos tropeiros.
 (d) Amparo é bonito/?bonita no inverno.

É possível que os julgamentos de gramaticalidade sobre os dados em (6) variem. Independentemente disso, mantém-se a pergunta sobre o lugar de onde vem a informação responsável por determinar a flexão do adjetivo, dada a incompatibilidade com o artigo. A tradição gramatical sugere que casos como (6) devem fazer concordância com o feminino, dada a presença do nome *cidade* (feminino) que está subentendido (ou elíptico). Porém, não é difícil encontrar contraexemplos para essa afirmação. O primeiro vem de casos como (6c-d), em que a concordância com o feminino não é imediatamente óbvia, a despeito de serem cidades. Em segundo lugar, os exemplos em (7) não podem ser explicados por essa hipótese.

- (7) (a) Cuba foi invadida/*invadido pelos inimigos.
 (b) Portugal foi conquistado/?conquistada pela Gália.

A incompatibilidade de *Cuba* (um topônimo que não aceita artigo) com um adjetivo masculino mostra que não pode ser o caso de a palavra *país* estar elidida, forçando a concordância com o gênero masculino. Como aparece em § 3, a hipótese deste artigo é a de que os nomes próprios não adentram a derivação sintática como unidades atômicas e opacas (categorizadas); elas são construídas na sintaxe. Uma consequência direta disso é a de que o mecanismo de concordância é capaz de acessar a estrutura interna da palavra,

⁴ Um exemplo de variação entre PB e PE está no nome *França*: ao passo que em PB, esse topônimo é feminino e precedido por *a* (*a França é o maior exportador...*), em PE, esse nome próprio ocorre sem determinante (*França é o maior exportador...*). Quanto à variação dialetal do PB, enquanto as variedades do Sul e do Sudeste empregam artigo antes de antropônimos (*eu fui na casa da Maria*), algumas variedades do Norte e do Nordeste utilizam-nos sem artigo (*eu fui na casa de Maria*), mesmo em contextos sintáticos mais marcados, como complemento de preposição.

fazendo com que a gramática acesse a informação de gênero em seus usos não toponímicos (quando houver) e se valha dela para flexionar o adjetivo.

A hipótese de que as unidades primitivas que servem de base para a formação de nomes são raízes, e não entradas lexicais, levanta outros questionamentos e requer um maior número de fenômenos empíricos como base. Ainda acerca dos nomes de lugares, as raízes que derivam topônimos também formam adjetivos pátrios (ou relativos a cidades e estados). Os exemplos mais claros aparecem em (8a). Já os exemplos em (8b) mostram que, às vezes, apenas uma “parte do topônimo” serve como base para o adjetivo, o que constitui evidência de que os nomes próprios não são unidades privilegiadas não internamente analisáveis.

- (8) (a) brasileiro, argentino, francês, paranaense, são-carlense, baiano...
 (b) paulistano/paulista, catarinense, rio-pretense...

Em (8), os adjetivos pátrios formados mantêm uma relação semântica com os topônimos aos quais eles estão morfológicamente associados. Contudo, em (8b), se *São José do Rio Preto* ou *São Paulo* fossem itens primitivos, alguma operação morfológica ou lexical teria que eliminar parte do nome próprio (no caso, *são* e *santa*) de modo a formar o adjetivo. Não somente isso, a derivação do adjetivo pátrio a partir do topônimo criaria uma certa dependência semântica, no que tange àquilo que o adjetivo denota, no sentido de que não se esperaria que fosse possível utilizar adjetivos que derivam de topônimos sem que esses fizessem, de fato, referência a um lugar, como pode ser observado nos exemplos em (9).

- | | |
|---------------------|------------------|
| (9) montanha russa | sair à francesa |
| pão francês | sonho americano |
| batata inglesa | fazer a egípcia |
| jeitinho brasileiro | negócio da China |

O que os dados em (9) mostram é que há adjetivos (ou locuções) pátrios que não estabelecem mais uma relação direta com os topônimos dos quais eles deveriam derivar – o que acarretaria uma dependência semântica. Naturalmente, alguém poderia argumentar que, por exemplo, *batata inglesa* ou *pão francês* remetem a produtos que foram fabricados ou cultivados no país ao que o adjetivo se refere; todavia, quando um cliente pede ao vendedor algum desses produtos, é muito provável que o atendente não imagine que o cliente queira comprar algum produto importado, mas sim algum *tipo específico* de pão ou batata que, apenas metafórica ou historicamente, se relaciona ao lugar de origem. O mesmo se aplica às demais expressões.

Obviamente alguém pode sempre sugerir que algum mecanismo morfológico ou lexical dê conta de armazenar ou gerar as formas com significado/estrutura cristalizado. Porém, a falta de correspondência morfológica em (8b) e semântica em (9) enfraquece a dependência entre esses itens lexicais. Pondo de lado esses casos, é na formação dos antropônimos que a criatividade e produtividade morfológica do falante revelam o ápice de sua dinamicidade. Primeiramente, há vários nomes próprios no PB (assim como em qualquer outra língua) que são, além de antropônimos, nomes comuns, como em (10).

- | | |
|------------|---------------------------|
| (10) Linda | Vitória |
| Branca | Margarida |
| Esmeralda | Aparecida |
| Rosa | Júlio (cf. <i>julho</i>) |

Observando os dados em (10), entre muitos outros, seria difícil sustentar que para além de uma entrada lexical antroponímica, o léxico deveria dispor de uma entrada distinta para cada nome comum, não antroponímico. A dificuldade para sustentar essa hipótese advém, em adição aos casos já arrolados, de mais três argumentos. O primeiro deles é o de que o falante dispõe de recursos morfológicos “usuais” para formar nomes próprios, tais como SUFIXAÇÃO – em (11a) – e MESCLAGEM, em (11b). Nesse sentido, não há nada na formação dos antropônimos que os privilegie em relação aos outros processos de formação de palavras.

- (11) (a) Tereza/Terezinha, setembro/Setembrino, Angélica/angelical.
 (b) Jomar (José + Maria), Claudionor (Cláudio + Leonor), Silvanir (Sílvio + Nair).

O segundo argumento é que não só é possível “transformar” nomes comuns em nomes próprios, mas também é possível utilizar nomes próprios para gerar nomes ou adjetivos comuns. Evidentemente, o emprego desses antropônimos nas expressões em (12) não parece estar associado a nenhum indivíduo particular, da mesma forma que mostrado para os topônimos em (9).

- | | |
|-----------------|-------------------------|
| (12) mauricinho | zé-droguinha |
| patricinha | maria-mole |
| banho-maria | maria-vai-com-as-outras |
| joão-ninguém | maria-fumaça |
| joão-bobo | maria-chuteira |
| zé-ruela | maria-gasolina |
| zé-ninguém | ricardão |

O terceiro argumento – e talvez o mais difícil de contestar sob a ótica da descrição empírica – é o de que, do ponto de vista morfológico, via de regra, é possível criar nomes próprios a partir de qualquer item lexical – ou combinação de itens lexicais – já existente. Monteiro (2002, p. 206) apresenta um levantamento dos nomes mais incomuns já registrados em cartórios no Brasil. Os casos em (13) são alguns de seus exemplos.

- (13) Restos Mortais de Catarina
 Bem Vindo O Dia Do Meu Nascimento
 Céu Azul Do Céu Poente
 Finadina Defuntina Da Boa Morte
 José Casou De Calças Curtas
 Himineu Casamentício Das Dores Conjugais
 Estrada De Ferro Brasileira
 Um Dois Três De Oliveira Quatro
 Rolando de Alto A Baixo Da Escada

Os dados em (13) exemplificam a alta produtividade (e criatividade) dos falantes na criação de antropônimos. Diante de tal dinamicidade, a hipótese de que nomes próprios constituem entradas lexicais independentes, com propriedades “especiais”, como propõe Matushansky (2008), parece bem menos convincente. Seja como for, independentemente da produtividade morfológica, há certas restrições que os falantes parecem *não* violar, tais como as restrições fonotáticas.

Massini-Cagliari (2011) mostra que os empréstimos de antropônimos estrangeiros são geralmente suscetíveis a adaptações, de modo que a gramática fônica do PB possa acomodar nomes próprios que, na sua origem, violariam algum princípio fonológico da

língua – o que é visto em nomes como em *Washington* (inglês) → *Uósto* (PB) e *Giovanni* (italiano) → *Diovani* (PB). Segundo a autora (p. 799), esses seus exemplos indicam que a mudança ortográfica é sinalizadora de uma adaptação fonológica.

À luz das propriedades morfológicas do PB, há também generalizações. Uma delas é mostrada por Santana (2019). Em primeiro lugar, a autora advoga em favor da existência de raízes antroponímicas, ou seja, de unidades formais menores que entradas lexicais que estão na base da criação de nomes próprios do PB, como indica o Quadro 2, retirado de Santana (2019, p. 70). Adicionalmente, a autora defende uma sistematicidade no que tange às vogais temáticas que aparecem com essas raízes, de tal modo que *-a* e *-e* formem nomes femininos e *-o* e \emptyset , nomes masculinos.

Contexto fonológico	“raiz”	-o	\emptyset	-e	-a
/l/	Rafael-	---	Rafael	Rafaele	Rafaela
	Daniel-	---	Daniel	Daniele	Daniela
	Gabriel-	---	Gabriel	Gabriele	Gabriela
	Marcel-	Marcelo	Marcel	Marcele	Marcela
	Manuel-	---	Manuel	Manuele	Manuela
/z/	Luiz-	---	Luiz	Louise	Luíza
	Denis-	---	Dênis	Denise	---
/r/	Artur-	Arturo	Artur	---	---
	Vilmar-	---	Vilmar	---	Vilmara
	Valdir-	---	Valdir	---	Valdira
/N/	Alan-	Alano	Alan	Alane	Alana
	Jean-	---	Jean	Jeane	Jeana
	Adrian-	Adriano	Adrian	Adriane	Adriana
	Cristian-	Cristiano	Cristian	Cristiane	Cristiana
	Julian-	Juliano	Julian	Juliane	Juliana
	Luan-	---	Luan	Luane	Luana
	Rosan-	Rosano	---	Rosane	Rosana
C não licenciada em coda	Felip-	Felipo	Felip(e)	---	Felipa
	Alexandr-	---	Alexandr(e)	---	Alexandra
	Erik-	Erico	Erick(e)	---	Erica
	Enrik-	Enrico	Enriqu(e)	---	Enrica

Quadro 2: distribuição das vogais temáticas nos antropônimos, retirada de Santana (2019, p. 70)

Em relação ao Quadro 2, é verdade que a oposição entre masculino em *-o* e feminino em *-a* traz pouca novidade; contudo, a oposição sistemática entre *-e* e \emptyset mostra que a marcação de gênero é bastante previsível nos antropônimos. Mais do que isso, o Quadro 2 constitui evidência em favor de raízes antroponímicas, das quais as diferentes “famílias de nomes próprios” emergem. Não obstante, uma análise que defende a hipótese de raízes como primitivos deve tomar certas decisões com relação a “quanto material” a raiz apresenta. Em relação aos topônimos, os dados em (8b) sugerem que a raiz seja somente $\sqrt{\text{PAUL}}$ e não $\sqrt{\text{SÃO-PAUL}}$, dada a ausência de *são* em algumas palavras derivadas – mas não em outras, como *são-paulino*. Fenômeno semelhante – ou decisão semelhante – versa sobre os antropônimos em (14), cujo *-s* final não aparece em todas as “formas” do nome.

- (14) (a) Carlos / Carlinhos / Carlão, mas *Carlãos
- (b) Douglas / Douglinhas / Douglão, mas *Douglãos
- (c) Jonas / Joninhas / Jonão, mas *Jonãos
- (d) Lucas / Luquinhas / Lucão, mas *Lucãos
- (e) Marcos / Marquinhos / Marcão, mas *Marcãos

Os dados em (14) ilustram um problema⁵ – uma assimetria no que tange à presença de -s nas diferentes marcações de grau – com que uma análise que trata da morfologia dos nomes próprios tem que lidar. Feitas essas considerações, o objetivo desta seção foi não o de propor uma análise para acomodar os fenômenos que subjazem à formação e aos diferentes empregos dos nomes próprios, mas o de mostrar que, de fato, existe uma agenda de problemas empíricos subjacentes à sua formação e uso.

2 Uma abordagem não lexicalista para os nomes próprios: alguns caminhos

Ainda que – do ponto de vista estritamente linguístico – se saiba relativamente pouco acerca dos usos e dos mecanismos de formação dos nomes próprios, o conjunto de fatos empíricos a serem explicados, apresentado em § 1, atesta contra a postulação de entradas lexicais antroponímicas/toponímicas. Essa é uma ideia defendida por Matushansky (2008). Os detalhes da análise da autora não são relevantes para os presentes propósitos; de todo modo, a essência de sua proposta aparece em (15) – de Matushansky (2008, p. 592).

$$(15) \parallel \text{Alice} \parallel = \lambda x \in D_e. \lambda R_{\langle e, \langle n, t \rangle \rangle}. R(x) (/a'lise/)$$

em que n é a CADEIA FONOLÓGICA de e , e R é a CONVECÇÃO DE NOMEAÇÃO.

Por questão de espaço e de escopo, convém deixar de lado o problema da postulação de um argumento de natureza puramente fonológica na forma lógica em (15). Ademais, é justo observar que Matushansky mobiliza argumentos empíricos consistentes em favor de um tratamento dos antropônimos como predicados.⁶ De qualquer forma, à luz do que foi apresentado em § 1, convém tecer algumas considerações sobre por que esse tratamento é pouco adequado. Muito embora uma proposta como (15) seja capaz de lidar com a alternância argumento/predicado, vista nos diferentes empregos dos antropônimos, e com alguns outros problemas semânticos (como quando os antropônimos aparecem como complemento de verbos de nomeação, como em: *por favor, me chame de “João”*), do ponto de vista estrutural, essa análise enfrenta muitos problemas.

Primeiramente, a análise de Matushansky, que trata apenas de antropônimos comuns, tem como consequência a assunção de que, para cada nome próprio que a gramática gera, deve haver uma entrada lexical diferente e independente. Assim, se o falante cria “Rubi” ou “Estrada de Ferro Brasileira” como antropônimos, é necessário inchar o léxico com itens idênticos, à exceção de que um tem uma leitura referencial; e o outro, não. Além disso, foi mostrado não só que se podem gerar nomes próprios a partir de nomes comuns, mas também que é possível gerar nomes próprios através de recursos usuais de formação de palavras, como em *Natalino* e *Celestino* – adicionalmente a (11). Uma regra como (15) não prevê que a sufixação de nomes comuns possa gerar uma leitura referencial.

Outro ponto contra o engessamento estrutural (ou “atomicidade”) dos nomes próprios diz respeito ao fato de que topônimos e antropônimos são internamente analisáveis, mesmo quando eles não derivam de nomes comuns, ou seja, a possibilidade de formar adjetivos pátrios que contêm apenas parte do topônimo (*Santa Catarina* → *catarinense*) e de identificar uma sistematicidade nas vogais temáticas dos antropônimos em

⁵ Adicionalmente a isso, como um dos pareceristas anônimos apontou, casos como *Robson / Robinho / Robão* (mas não **Robsoninho* e **Robsonão*) e também *César / Cesinha / Cesão* (mas não **Cesarinha* e **Cesarão*) mostram que o truncamento da raiz parece ocorrer com outros nomes, que não os terminados em -s. A derivação das formas de grau em nomes como *César* e *Robson* também deve levar em conta a preservação da vogal aberta da raiz. Um melhor entendimento de como derivar a estrutura desses casos, acomodando esses fenômenos empíricos aguarda trabalhos futuros.

⁶ Na proposta da autora, quando um nome próprio funciona como um predicado de uma MINIORAÇÃO no complemento de um verbo de nomeação, é o próprio verbo que “preenche” a CONVECÇÃO DE NOMEAÇÃO. Por outro lado, quando um nome próprio aparece na posição de argumento, a posição para a convenção de nomeação é saturada por uma variável livre, contextualmente dependente (p. 592).

relação à marca de gênero (*Julian, Juliano, Juliana, Juliane*) revela que os nomes próprios não são unidades atômicas. Adicionalmente ao fato de que os topônimos parecem se valer do gênero gramatical do nome comum – quando for o caso – para estabelecer a relação de concordância, como mostrado em (6) e (7).

Finalmente, o fato de ser possível encontrar expressões contendo nomes e adjetivos – que morfologicamente deveriam derivar de nomes próprios – que não fazem referência a algum lugar ou a algum indivíduo específico mostra a independência semântica entre esses itens, para além da oposição argumento/predicado, isto é, em *tem dois maurícios na minha turma de Semântica*, o antropônimo denota um predicado: o de “ter o nome Maurício”, mas em *tem dois mauricinhos na minha turma de Semântica*, o emprego do antropônimo, mesmo como predicado, nada tem a ver com o nome “Maurício”.

Dadas essas considerações, alternativamente ao tratamento lexicalista para os nomes próprios (proposto por alguns autores e assumido por muitos sintaticistas e semanticistas), convém tecer alguns comentários e/ou apontar algumas direções a respeito de uma análise sintática para a formação de palavras, como a da Morfologia Distribuída, que advoga em favor de tratar as raízes como primitivos, mesmo para a derivação dos nomes próprios, seguindo uma visão alossêmica das raízes, como proposto por Saab & Lo Guercio (2019) e inicialmente desenvolvida por Resende & Santana (2019) para o PB.

De maneira geral, a ALOSSEMIA tem a ver com a determinação do significado de uma raiz (ou mesmo de uma unidade maior) com base em um contexto x, da mesma forma que a ALOFONIA determina a realização de um dado fonema em um contexto y e a ALOMORFIA determina a forma de um morfema em um contexto z. Por exemplo, na maioria dos dialetos do PB, o fonema /t/ tem duas realizações, a saber, [t] antes de /i/ e [ʈ] nos demais ambientes. Assim, diz-se que [t] e [ʈ] são alofones – são formas do mesmo fonema – e que sua alofonia é determinada contextualmente (isto é, no contexto de /i/).

Do mesmo modo, o PB dispõe de um prefixo de negação *iN-* que se manifesta de duas⁷ formas, quais sejam: *i-* antes de *-l* e *-r* e *iN-* nos demais ambientes. Por essa razão, é dito que *i-* e *iN-* são alomorfes – são formas do mesmo morfema – e que sua alomorfia é determinada contextualmente (ou seja, no contexto do tipo de material subsequente). À luz do mesmo raciocínio, este artigo defende a alossemia das raízes antroponímicas: $\sqrt{\text{JOAN}}$ tem uma interpretação referencial no contexto de um traço [\pm HUMANO] e uma leitura predicativa nos demais ambientes, como mostrado a seguir.

Em linhas gerais, o modelo da Morfologia Distribuída (MD) defende que a sintaxe é o único componente gerativo da gramática e que toda derivação, de palavras a sentenças, é feita pelo mesmo conjunto de operações. Assim, o elemento básico de uma estrutura – isto é, o primitivo de análise – é a raiz. A ela se adjungem os núcleos funcionais que, com seus traços, vão dar conta de gerar a estrutura correta com suas respectivas forma fonológica e interpretação semântica. Na MD, o componente fonológico é apenas interpretativo (e não gerativo) e acionado apenas pós-sintaticamente; desse modo, qualquer referência à forma fonológica antes da sintaxe, como em (15), é desmotivada nesse modelo.

Assumindo o quadro teórico da Morfologia Distribuída, na esteira de Santana (2019), este artigo defende que o LÉXICO ESTRITO (a LISTA 1) dispõe de raízes antroponímicas como $\sqrt{\text{JOAN}}$, $\sqrt{\text{MARI}}$, $\sqrt{\text{DANIEL}}$ etc. que são os átomos da derivação dos nomes próprios comuns. Vale notar que, nesse ponto da derivação (ou seja, na seleção das raízes), não há nada de “antroponímico” nelas, de modo que o mesmo tipo de derivação se aplique a $\sqrt{\text{LIND}}$, $\sqrt{\text{ESMERALD}}$ ou $\sqrt{\text{JULI}}$ se os nomes próprios forem cognatos a nomes comuns (como *linda*, *esmeralda*, *julho* etc.). Sejam nomes comuns sejam próprios, a ideia é que essas estruturas são *nomes* e, portanto, a derivação de sua estrutura vai depender da concatenação

⁷ Em prol da simplicidade do exemplo, não se está considerando suas variantes alomórficas fonologicamente condicionadas.

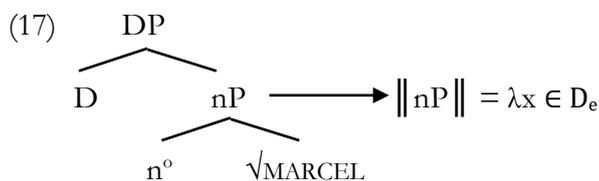
da raiz a um nominalizador (um n^o), e o tipo de leitura que o nome apresenta (referencial ou não) vai depender dos traços do núcleo nominalizador.

Como defendido, a interpretação das raízes deve ser determinada contextualmente e, além disso, as estruturas que se referem a indivíduos específicos são formalmente idênticas às estruturas que se referem a predicados (ou a nomes comuns), *exceto* no que tange à leitura referencial. Na mesma linha de Saab & Lo Guercio (2019) e Resende & Santana (2019), o presente trabalho propõe que a diferença formal entre esses dois tipos de leitura advém de um traço [\pm HUMANO]. Na verdade, esse traço não tem a ver, de fato, com ser ou não *humano*, mas sim com ser ou não *referencial*, e a oposição “+” e “-” visa lidar com o comportamento estrutural distinto entre antropônimos ([+HUMANO].) e topônimos ([-HUMANO].),⁸ como exemplificado a seguir. Com base nisso, o algoritmo de alossemia é dado em (16).

$$(16) \sqrt{\text{RAIZ}} = \langle e \rangle / [\pm\text{HUMANO}]$$

$$\sqrt{\text{RAIZ}} = \langle e, t \rangle / \text{demais ambientes}$$

O que (16) determina é que qualquer raiz (e, por extensão, qualquer estrutura) terá uma interpretação de argumento – isto é, referencial – no contexto do traço [\pm HUMANO], como ilustrado em (17). Esse traço é alocado em n^o, porque é no ambiente nominal que se deriva(m) a(s) sintaxe(s) e a(s) semântica(s) de um nome próprio. Em LF, esse traço vai ser lido como “interprete como referencial tudo aquilo que estiver sendo c-comandado por [\pm HUMANO]”, como formulado a seguir. Assim sendo, mesmo antropônimos como *Restos Mortais de Catarina* serão interpretados como não composicionais, dada a leitura referencial. Ainda que os detalhes a respeito da composição de nomes próprios “mais complexos” não seja explorada, a ideia é que qualquer nome com leitura referencial seja encabeçado por um n^o e que, na presença do traço [\pm HUMANO], qualquer informação “interpretável” abaixo do nominalizador seja neutralizada; entretanto, as informações estruturais devem permanecer “ativas” para operações sintáticas e morfológicas, como concordância etc.



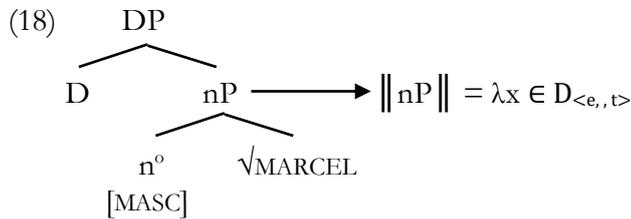
[MASC, +**HUMANO**] → Se [\pm HUMANO] c-comanda α , então α é referencial.

A estrutura em (17) – ou uma com o traço [-HUMANO] para derivar um topônimo – tem a vantagem de capturar o fato de que qualquer estrutura linguística – e não apenas antropônimos comuns como *João, Maria* etc. – podem denotar um nome próprio. Dessa forma, o percurso derivacional deve ser o mesmo, sobretudo porque há nomes próprios que são formados por sufixação, composição, mesclagem etc. Em adição a isso, uma abordagem sintática para a formação dessas palavras tem a vantagem de explicar por que (e como) os nomes próprios podem ser internamente analisáveis.

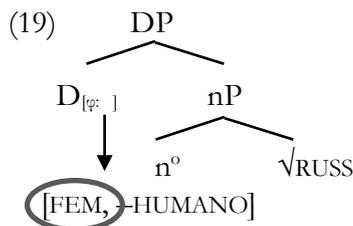
É um fato conhecido sobre o PB – e outras línguas românicas – que os nomes têm, como propriedade inerente, gênero e classe. Assim, há um requerimento de boa formação nominal que manipula/acomoda essas duas propriedades, e isso nada tem a ver com referencialidade, ou seja, o traço de gênero em n^o – cf. Resende & Santana (2019) para discussão e referências – independe de a leitura ser ou não referencial. A ideia, então, é a de

⁸ Para Saab & Lo Guercio (2019), trata-se de um traço privativo.

que a estrutura que subjaz ao emprego predicativo dos antropônimos não contém o traço $[\pm\text{HUMANO}]$, mas somente a informação de gênero e classe, pois delas dependem a correta realização fonológica do nome, como mostrado em (18) para *Marcelo*, em um contexto de *todo Marcelo é um ótimo engenheiro*.



A ausência de $[\pm\text{HUMANO}]$ implica na leitura predicativa do nome e captura o fato de que a única diferença entre uma estrutura que denota um único indivíduo da que denota um conjunto de indivíduos é um traço de referencialidade. Finalmente, é preciso dizer algo em relação à concordância com topônimos. Do ponto de vista descritivo, a concordância nominal diz respeito à dependência do determinante (e do adjetivo) em relação aos traços flexionais do nome. Em termos mais técnicos, Kučerová (2018) e outros defendem que o núcleo D realiza uma operação de SONDA sobre o núcleo n° , de modo a valorar seus traços- ϕ . Dessa maneira, em uma estrutura como (19), a sonda de D vai valorar seus traços, de acordo com a especificação de n° independentemente da leitura, e inserir em PF o expoente fonológico apropriado para esse(s) traço(s). Em (19), uma estrutura para *Rússia*.



Sobre (19), alguns comentários. Primeiramente, a operação de sonda se aplica não só a determinantes, mas também a adjetivos. Em segundo lugar, a sonda de D sobre n° se aplica independentemente da presença do traço $[-\text{HUMANO}]$, ou seja, a despeito de haver leitura referencial. Uma terceira observação versa sobre a realização do núcleo D. Como mostrado em (5), ao passo que nomes de países são compatíveis com determinantes, nomes de cidades e estados (e de países como *Cuba* e *Portugal*) não são.⁹

Portanto, algum mecanismo adicional deve restringir a realização de D quando a especificação do traço for $[-\text{HUMANO}]$, muito embora esse mesmo mecanismo deva ser capaz de diferenciar os diferentes tipos de topônimos, à luz das generalizações em (5). É possível que se trate do mesmo mecanismo que restringe o emprego de artigos com antropônimos em alguns dialetos do PB. Todos esses desdobramentos aguardam trabalhos futuros.

CONSIDERAÇÕES FINAIS

O presente artigo teve como objetivo principal apresentar os problemas empíricos estreitamente linguísticos (isto é, semânticos, sintáticos, morfológicos e até fonológicos, em

⁹ Dentro do conjunto de ferramentas de que a teoria da Morfologia Distribuída dispõe, seria possível sugerir que algumas raízes sejam marcadas com algum diacrítico ou que algum (feixe de) traço adicional deve co-ocorrer com a especificação $[-\text{HUMANO}]$. Entretanto, no que tange à melhor maneira de acomodar algumas dessas generalizações empíricas a uma modelagem formal, esse é um problema que permanece em aberto.

adição aos ontológicos já bastante discutidos na literatura) que subjazem à caracterização da forma e dos diferentes empregos dos nomes próprios. A partir dos problemas levantados, este trabalho advogou contra o estatuto atômico dos nomes próprios (isto é, da hipótese de que eles constituem entradas lexicais independentes) e a favor de uma abordagem sintática de formação de palavras, como a da Morfologia Distribuída.

Para ilustrar como esse tratamento – e como a abordagem alossêmica para as raízes – funciona(m), este trabalho discutiu a oposição argumento/predicado no que concerne ao emprego dos nomes próprios e ao caráter criativo e produtivo da formação/criação desses itens. Adicionalmente, este artigo teceu algumas considerações sobre a concordância nos topônimos e levantou problemas (teóricos e empíricos) a respeito de (i) as diferenças entre antropônimos e topônimos e entre seus diferentes tipos, (ii) a restrição de determinantes com alguns topônimos, (iii) o lugar que aloca a informação de concordância dos adjetivos.

Em síntese, este artigo tentou reunir os problemas que constituem a agenda de questões empíricas que cercam esse fenômeno e mostrou que uma análise sintática para a formação dessas unidades é descritivamente mais adequada do que a que propõe/assume a existência de entradas lexicais antroponímicas. Naturalmente, como adiantado, nem todos os problemas levantados puderam ser explicados – alguns apenas mencionados –, mas é possível que as ferramentas introduzidas neste trabalho auxiliem trabalhos futuros, que visam tratar esses e outros fenômenos.

REFERÊNCIAS

AMARAL, Eduardo Tadeu Roque. Contribuições para uma tipologia de antropônimos do português brasileiro. **Alfa**. São Paulo, v. 55, n. 1, p. 63-82, 2011.

BRITO, Adriano Neves de. **Nomes próprios: semântica e ontologia**. Brasília: UnB, 2003.

FREGE, Gotlob. Sobre o sentido e a referência. In: _____. **Lógica e filosofia da linguagem**. 2. ed. São Paulo: USP, 2009 [1892]. p. 129-158.

HALLE, Morris; MARANTZ, Alec. **A Morfologia Distribuída e as peças da flexão**. Curitiba: UFPR, 2020 [1993].

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

LONGOBARDI, Giuseppe. Reference and proper names: a theory of N-movement in syntax and Logical Form. **Linguistic Inquiry**. Cambridge, v. 25, n. 4, p. 609-665, 1994.

MASSINI-CAGLIARI, Gladis. Adaptação fonológica de nomes próprios de origem estrangeira: comparação entre português arcaico e português brasileiro. **Estudos Linguísticos**. São Paulo, v. 40, n. 2, p. 795-807, 2011.

MATUSHANKSY, Ora. On the linguistic complexity of proper names. **Linguistic and Philosophy**. v. 21, p. 573-627, 2008.

MONTEIRO, José Lemos. **Morfologia Portuguesa**. 4. ed. Campinas: Pontes, 2002.

RESENDE, Maurício Sartori; SANTANA, Beatriz Pires. A relação entre raízes, gênero, classe e significado. **Revista da ABRALIN**. s. l., v. 18, n. 1, p. 2-55, 2019.

SAAB, Andrés. LO GUERCIO, Nicolás. No name: the alossemy view. **Studia Linguistica**. Oxford, v. 74, n. 1, p. 1-38, 2019.

SANTANA, Beatriz Pires. **Morfologia ornamental**: as vogais temáticas do português brasileiro. 2019. 337f. Tese (Doutorado em Linguística) – Universidade Federal do Paraná, Curitiba.