Competição entre motivações e variação linguística: artigo definido em sintagmas nominais com topônimo na *Peregrinação*

César Nardelli Cambraia¹

Resumo: Este estudo teve como objetivo analisar o comportamento linguístico do artigo definido em sintagmas nominais com topônimo na *Peregrinação*, obra escrita no terceiro quarto do séc. XVI por Fernão Mendes Pinto. A análise se fundamentou no modelo tipológico-funcional givóniano e na sociolinguística variacionista laboviana. Foram testadas treze variáveis linguísticas, das quais apenas nove se mostraram estatisticamente relevantes nos casos de variação, na seguinte ordem decrescente: *polivalência, estrutura do SN, indicador de entidade geográfica, menção, estrutura genitiva, função sintática, quantificador, tradição e coordenação*. Realizou-se também um confronto com resultados relativos a sintagmas nominais com antropônimo na mesma obra e se identificaram semelhanças e diferenças.

Palavras-chave: Funcionalismo. Variação Linguística. Linguística Histórica. Artigo Definido. Topônimos.

Competition between motivations and linguistic variation: definite article in noun phrases with toponym in the *Peregrinação*

Abstract: This study aimed to analyze the linguistic behavior of definite article in noun phrases with toponym in *Peregrinação*, a work written in the third quarter of the 16th century by Fernão Mendes Pinto. The analysis was based on the Givónian typological-functional model and Labovian variationist sociolinguistics. Thirteen linguistic variables were tested, of which only nine were statistically relevant in the cases of variation, in the following descending order: *versatility, SN structure, indicator of geographic entity, mention, genitive structure, syntactic function, quantifier, tradition and coordination*. A comparison was also carried out with results relating to noun phrases with anthroponyms in the same work and similarities and differences were identified.

Keywords: Functionalism. Linguistic Variation. Historical Linguistics. Definite article. Toponyms.

1. Introdução²

No latim, língua da qual derivaram as línguas românicas, não existia artigo definido. No curso da diferenciação da língua latina em seus rebentos românicos, essa categoria emergiu a partir do demonstrativo de 3ª pessoa e, com o passar do tempo, foi ampliando seus contextos de uso. No século XVI, seu emprego se encontrava em um estágio intermediário: era empregado em mais contextos do que no português arcaico, mas em menos contextos do que no português

¹ Professor na Universidade Federal de Minas Gerais. Doutor em Filologia e Língua Portuguesa pela Universidade de São Paulo. Orcid ID: <https://orcid.org/0000-0002-2403-3021>. E-mail: nardelli@ufmg.br.

² Este estudo faz parte dos resultados do projeto “Para uma gramática do português clássico: o sintagma nominal e suas funções na *Peregrinação*, de Fernão Mendes Pinto [Fase I]”, Produtividade em Pesquisa, CNPq, 2021-2024.

moderno. Justamente por isso, é de grande interesse examinar como era o comportamento linguístico nessa época intermediária.

No presente estudo, apresenta-se uma análise linguística do comportamento do artigo definido em sintagmas nominais com topônimo na obra *Peregrinação*, de Fernão Mendes Pinto (ca. 1510-1583), publicada pela primeira vez somente em 1614, mas elaborada no terceiro quarto do séc. XVI. Esta análise contempla especialmente o fenômeno da variação linguística no emprego do artigo definido em sintagmas nominais com topônimo, mas aborda também, comparativamente, o emprego dessa forma em sintagmas nominais com antropônimo na mesma obra³.

2. Artigos definidos e topônimos na língua portuguesa

Para Cunha e Cintra (1985, p. 199), o artigo definido é uma forma que é anteposta ao substantivo “quando se trata de um ser já conhecido do leitor ou ouvinte, seja por ter sido mencionado antes, seja por ser objeto de um conhecimento de experiência”. De acordo com Cunha e Cintra (1985, p. 204), existe uma gradação em termos de precisão expressa por determinantes, na ordem que se segue: o *artigo indefinido* indica a espécie do substantivo, o *artigo definido* restringe a extensão do significado do substantivo e o *pronome demonstrativo* limita mais o sentido do substantivo ao situá-lo no espaço e no tempo.

No que diz respeito aos nomes próprios, Cunha e Cintra (1985, p. 216) consideram que, por serem por definição individualizantes, eles deveriam tornar desnecessário o artigo definido. Entretanto, os autores esclarecem:

Mas, no curso da história da língua, *razões diversas concorreram* para que esta norma lógica nem sempre fosse observada e, hoje, há mesmo um grande número de nomes próprios que exigem obrigatoriamente o acompanhamento do artigo definido. (CUNHA; CINTRA, 1985, p. 216, *itálicos nossos*)

No que se refere a nomes geográficos (i. é, topônimos), Cunha e Cintra (1985, p. 219-222) assinalam três padrões:

- a) normalmente se emprega artigo definido:
 - (i) com os nomes de países, regiões, continentes, montanhas, vulcões, desertos, constelações, rios, lagos oceanos, mares e grupos de ilhas; e
 - (ii) com nomes dos pontos cardeais e colaterais.

³ Os dados referentes à análise dos antropônimos constam todos de Cambraia (2025), razão pela qual, em nome da concisão, não se apresentará a referência desse estudo no presente trabalho cada vez que for mencionado o tema.

- b) geralmente não se emprega artigo definido:
 - (i) com os nomes de cidades, de localidades e da maioria das ilhas; e
 - (ii) com os nomes de planetas e estrelas.
- c) não se emprega artigo definido de maneira uniforme com os nomes de estados brasileiros e de províncias portuguesas.

No entanto, também chamam a atenção para alguns casos como *Portugal*, *França* e *Inglaterra*, dentre outros, casos em que, ao contrário da norma geral, o topônimo pode ocorrer sem artigo definido, sobretudo quando regido de preposição. Note-se que os padrões assinalados são o comportamento geral, pois há tantos casos de nomes de país e regiões que rejeitam artigo (cf. *Castela*) quanto casos de nomes de cidades que exigem artigo (cf. *o Porto*). Por fim, os estudiosos registram que, como no caso de nomes de pessoas, os nomes geográficos admitem artigo definido quando acompanhados de determinação ou qualificação (cf. *A Lisboa dos Poetas Cavaleiros*).

Diferentemente dos antropônimos, cujo comportamento linguístico em relação aos artigos definidos tem sido discutido no âmbito da linguística moderna já há décadas (cf., p. ex., CALLOU; SILVA, 1997; AMARAL, 2003; MORAES; LIMA, 2021), os topônimos foram pouco contemplados.

O único trabalho mais sistemático a lidar com esse tema sob a perspectiva da variação foi o de Mendes (2009). Uma vez que esse trabalho não apenas contemplou a questão dos topônimos como ainda a analisou comparativamente com antropônimos (que são os dois objetivos do presente estudo), é imperativo dar especial atenção a ele.

Mendes (2009) analisou os sintagmas nominais contendo antropônimos e topônimos a partir de entrevistas realizadas na zona rural das cidades de Abre Campo e Matipó, localidades vizinhas no leste do Estado de Minas Gerais. Foram entrevistados quatro informantes em cada cidade, todos com nível de escolaridade baixo (analfabetos ou com até o ensino fundamental completo), dois de cada gênero e dois de cada uma das duas faixas etárias consideradas (faixa 1 = 18 a 30 anos; faixa 2 = acima de 70 anos). A variável *dependente* foi a *presença ou ausência* de artigo definido e as variáveis *independentes* foram: (a) *localidade* (Abre Campo ou Matipó); (b) *gênero* (masculino ou feminino); (c) *idade* (18 a 30 ou acima de 70); (d) *tipo de antropônimo* (nome próprio de pessoa [prenome], sobrenome, apelido, nome completo ou hipocorístico); (e) *grau de intimidade do informante em relação ao antropônimo a que se refere* (pessoa mais próxima ou pessoa mais distante); (f) *estrutura de genitivo* [no caso de antropônimos] (presente

ou ausente); (g) *topônimos da cidade natal* (cidade natal ou outras localidades) e (h) *tipo de topônimo* (país, região, cidade/distrito, bairro, rua, serra, córrego, fazenda/sítio/casa, igreja ou estabelecimento comercial).

Foram excluídos da análise de Mendes (2009) os seguintes casos: (a) usos não-referenciais do artigo, como nos casos de vocativo e de nomenclatura didática e performativa; (b) usos metonímicos; (c) antropônimos e topônimos precedidos por demonstrativo; (d) antropônimos e topônimos precedidos pelas expressões *um(a) tal de*, *o(a) tal de*; (e) antropotopônimos usados no plural com referência a nomes de família; (f) estruturas em que os antropônimos ou os topônimos são iniciados por vogais homófonas aos artigos ou o termo anterior ao artigo definido termine em vogal homófona; (g) antropônimos precedidos por pronome indefinido; e (h) repetição do termo falado pelo pesquisador. Ao final, os dados dos *corpora* analisados ficaram constituídos da seguinte maneira: (a) Abre Campo, 206 com antropônimo e 115 com topônimo; e (b) Matipó: 414 com antropônimo e 113 com topônimo.

Os resultados obtidos por Mendes (2009) são apresentados a seguir de forma sintética em duas tabelas:

Tabela 1 – Presença ou ausência de artigo definido: Abre Campo/MG

Variáveis	Antropônimo		Total ₁	Topônimo		Total ₂	
	Presença	Ausência		Presença	Ausência		
Gênero	Masculino	54 (43%)	72 (57%)	126 (100%)	35 (53%)	31 (47%)	66 (100%)
	Feminino	45 (56%)	35 (44%)	80 (100%)	25 (51%)	24 (49%)	49 (100%)
Idade	18-30	48 (50%)	48 (50%)	96 (100%)	25 (51%)	24 (49%)	49 (100%)
	+70	51 (46%)	59 (54%)	110 (100%)	35 (53%)	31 (47%)	66 (100%)
Tipo	Prenome	60 (46%)	70 (54%)	130 (100%)	—	—	—
	Sobrenome	1 (100%)	—	1 (100%)	—	—	—
	Apelido	—	—	—	—	—	—
	Nom. compl.	6 (46%)	7 (54%)	13 (100%)	—	—	—
Grau de intimid.	Hipocorístico	—	—	—	—	—	—
	Mais próxima	68 (47%)	76 (53%)	144 (100%)	—	—	—
Estr. gen.	Mais distante	31 (50%)	31 (50%)	62 (100%)	—	—	—
	Presente	18 (54%)	15 (45%)	33 (100%)	—	—	—
Cidade	Natal	—	—	—	22 (56%)	17 (44%)	39 (100%)
	Outra	—	—	—	39 (51%)	37 (49%)	76 (100%)
	País	—	—	—	—	—	—
	Região	—	—	—	1 (100%)	—	1 (100%)
	Cid./distr.	—	—	—	43 (45%)	52 (55%)	95 (100%)
	Bairro	—	—	—	2 (50%)	2 (50%)	4 (100%)
Tipo de topônimo	Rua	—	—	—	—	—	—
	Faz./sít./casa	—	—	—	3 (100%)	—	3 (100%)
	Córrego	—	—	—	6 (86%)	1 (14%)	7 (100%)
	Serra	—	—	—	2 (100%)	—	2 (100%)
	Igreja	—	—	—	—	—	—
	Est. comerc.	—	—	—	2 (100%)	—	2 (100%)
Totals		99 (48%)	107 (52%)	206 (100%)	60 (52%)	55 (48%)	115 (100%)

Fonte: Adaptado de Mendes (2009)

Tabela 2 – Presença ou ausência de artigo definido: Matipó/MG

Variáveis	Antropônimo		Total ₁	Topônimo		Total ₂	
	Presença	Ausência		Presença	Ausência		
Gênero	Masculino	125 (90%)	14 (10%)	139 (100%)	20 (62,5%)	12 (37,5%)	32 (100%)
	Feminino	219 (80%)	56 (20%)	275 (100%)	37 (46%)	44 (54%)	81 (100%)
Idade	18-30	143 (90%)	16 (10%)	159 (100%)	8 (50%)	8 (50%)	16 (100%)
	+70	201 (79%)	54 (21%)	255 (100%)	49 (50%)	48 (50%)	97 (100%)
Tipo	Prenome	168 (88%)	23 (12%)	191 (100%)	—	—	—
	Sobrenome	14 (93%)	1 (7%)	15 (100%)	—	—	—
	Apelido	12 (71%)	5 (29%)	17 (100%)	—	—	—
	Nom. compl.	25 (66%)	13 (34%)	38 (100%)	—	—	—
	Hipocorístico	127 (84%)	24 (16%)	15 (100%)	—	—	—
Grau de intimid.	Mais próxima	227 (89%)	28 (11%)	255 (100%)	—	—	—
	Mais distante	112 (71%)	45 (29%)	157 (100%)	—	—	—
Estr. gen.	Presente	62 (91%)	6 (9%)	68 (100%)	—	—	—
Cidade	Natal	—	—	—	35 (71%)	14 (29%)	49 (100%)
	Outra	—	—	—	15 (23%)	49 (77%)	64 (100%)
Tipo de topônimo	País	—	—	—	1 (50%)	1 (50%)	2 (100%)
	Região	—	—	—	—	—	—
	Cid./distr.	—	—	—	28 (35%)	53 (65%)	81 (100%)
	Bairro	—	—	—	6 (86%)	1 (14%)	7 (100%)
	Rua	—	—	—	2 (100%)	—	2 (100%)
	Faz./sít./casa	—	—	—	9 (64%)	5 (36%)	14 (100%)
	Córrego	—	—	—	2 (50%)	2 (50%)	4 (100%)
	Serra	—	—	—	—	—	—
	Igreja	—	—	—	2 (67%)	1 (33%)	3 (100%)
Est. comerc.	—	—	—	—	—	—	
Totals		344 (83%)	70 (17%)	414 (100%)	57 (50%)	56 (50%)	113 (100%)

Fonte: Adaptado de Mendes (2009)

Convém ressaltar alguns aspectos de interesse dos resultados de Mendes (2009) para a discussão a ser realizada no presente estudo

Primeiramente, embora se trate de localidades vizinhas, constataram-se diferenças: em Abre Campo, a presença de artigo definido é menos frequente junto a antropônimos (48%) e mais frequente junto a topônimos (52%), mas, em Matipó, a presença é mais frequente junto a antropônimos (83%) e há empate junto a topônimos. Chama a atenção que, para topônimos, os valores são próximos demais entre as localidades (52% e 50%), o que faz imaginar que a diferença entre elas não seria realmente significativa nesse caso. Por outro lado, no caso de antropônimos, a diferença é bem substantiva.

Em segundo lugar, examinando os topônimos segundo as duas outras variáveis extralinguísticas (gênero e de idade), têm-se novamente valores muitos próximos, predominando a presença de artigo, mas com diferença percentual mais acentuada (62,5% frente a 37,5%) apenas para informantes do gênero masculino em Matipó.

A impressão que se tem é a de que, para topônimos, variáveis extralinguísticas têm pouco peso (ou quicá nenhum), sendo mais importantes as variáveis intralinguísticas. Para os

antropônimos, aquelas parecem ser relevantes, já que, em Abre Campo, o gênero masculino prefere ausência de artigo (57%) e o feminino a presença (56%). Em Matipó, todas as variáveis apresentam predominância de presença do artigo no caso de antropônimos, havendo, porém, diferença em termos de pontos percentuais em cada caso.

Como os dados de Mendes (2009) não foram objeto de uma análise multivariada, que se costuma realizar em estudos sociolinguísticos, então a avaliação do peso relativo de cada variável não foi apurada. Considerando o baixo número de ocorrências de dados por valor de algumas variáveis, é possível estimar que a aplicação desse tipo de análise não seria realmente viável, porque acabaria resultando em um número excessivo de dados com comportamento categórico (*knockout*): há, p. ex., apenas um dado para o valor *região* junto a topônimo por informantes de Abre Campo e apenas dois para o valor *país* de Matipó.

Um aspecto bastante interessante do estudo de Mendes (2009) foi evidenciar a existência de variação na presença ou ausência de artigo definido diante não apenas de antropônimos, mas também de topônimos. Considerando especialmente estes últimos, convém lembrar que Cunha e Cintra (1985, p. 219-222) asseveraram que, com nomes de cidades, geralmente não se emprega artigo definido na língua portuguesa, mas, nos dados de Mendes (2009), 45% dos casos são de presença de artigo para cidades ou distritos em Abre Campo e 35% em Matipó, ou seja, valores muito altos. Isso revela que é necessário aprofundar a análise linguística do comportamento do artigo definido em sintagmas nominais com topônimo de forma geral, contemplando novas variáveis, que é o que se pretende aqui.

3. Fundamentação teórica⁴ e hipótese de trabalho

O presente estudo se fundamenta no modelo tipológico-funcional givóniano e na sociolinguística variacionista laboviana.

De acordo com Neves (1997, p. 15), o funcionalismo constitui em uma teoria da organização gramatical das línguas naturais enquadrada em uma teoria global da interação social, que assume que a gramática se sujeita a pressões do uso. São pressupostos dessa teoria segundo Neves (1997, p. 46): (a) a língua é definida como instrumento de interação social; (b) a principal função da língua é a comunicação; (c) o correlato psicológico da língua é a

⁴ Como o presente estudo segue as mesmas diretrizes de estudo correlato em relação a antropônimos na *Peregrinação* (CAMBRAIA, 2025), então o instrumental teórico é basicamente o mesmo, que aqui se retoma.

competência comunicada, entendida como a habilidade de interagir socialmente por meio da língua; (d) o sistema linguístico deve ser analisado com base no seu uso; (e) a descrição linguística deve apresentar dados para dar conta de seu funcionamento num dado contexto; (f) a aquisição da linguagem é feita com a ajuda de um *input* extenso e estruturado de dados apresentados no contexto natural; (g) os universais linguísticos são explicados com base em restrições comunicativas, biológicas/psicológicas e contextuais; e (h) a pragmática é prioritária, aspecto dentro do qual a semântica e sintaxe são consideradas.

O modelo tipológico-funcional de Givón (2001) é bastante elucidativo para os estudos de variação e mudança linguística, uma vez que combina uma orientação *funcionalista*, com ênfase na função comunicativa da linguagem baseada no estudo da língua no seu contexto de uso, com uma orientação *tipológica*, que procura compreender a diversidade linguística. Esse modelo, de acordo com Martelotta e Areas (2003, p. 28), tem os seguintes pressupostos: (a) a linguagem consiste em uma atividade sociocultural; (b) a estrutura linguística se presta a funções cognitivas e comunicativas; (c) a estrutura linguística é não arbitrária, motivada, icônica; (d) a mudança e a variação linguística sempre se fazem presentes; (e) o sentido é contextualmente dependente e não atômico; (f) as categorias linguísticas não são discretas; (g) a estrutura linguística é maleável e não rígida; (h) as gramáticas são emergentes; e (i) as regras da gramática permitem exceções.

Especialmente importante para a discussão sobre o processo de variação e mudança no quadro do funcionalismo givóniano é a noção de *compromisso adaptativo*:

O fato de a gramática das orações codificar simultaneamente informação semântico-proposicional e discursivo-pragmática tem grandes consequências. Uma vez que as exigências de codificação das duas estão frequentemente em conflito, a estrutura resultante é um *compromisso adaptativo* entre as pressões funcionais em competição (GIVÓN, 2001, v. 1, p. 19, grifos do autor, tradução nossa).

No modelo givóniano, sintagmas nominais são descritos como compostos de núcleos (ing. *heads*) e modificadores (ing. *modifiers*). A principal função dos modificadores é a de especificar mais ou restringir o domínio da referência dos núcleos nominais que acompanham (GIVÓN, 2001, v. II, p. 1). Os modificadores se distribuem em duas grandes classes: (a) pré-nominais (quantificadores, determinantes, adjetivos ou sintagmas adjetivos e substantivos modificadores [aspecto este próprio da língua inglesa]) e (b) pós-nominais (orações relativas, complementos nominais e sintagmas preposicionados) (GIVÓN, 2001, v. II, p. 4-10). Os modificadores podem também ser classificados em: (a) *restritivos* (usados para restringir o

domínio da referência) e (b) *não restritivos* (empregados para enriquecer a descrição do referente com mais atributos característicos, mas sem restringir o domínio da referência) (GIVÓN, 2001, v. II, p. 10).

De acordo com Givón (2001, v. I, p. 459), os falantes codificam um referente nominal como definido quando assumem que seja *identificável* ou *acessível* ao ouvinte, ou seja, partem do pressuposto de que o referente esteja presente em alguma representação mental pré-existente do ouvinte. Por disso, interpreta-se que a definitude seja uma questão profundamente pragmática, relacionada à avaliação do falante quanto ao estado corrente de conhecimento do ouvinte em um dado momento da comunicação.

É de grande interesse em relação à definitude a questão da sua previsibilidade em termos de funções ou papéis: (a) os papéis de sujeito, de dativo/beneficiário e de associativo (isto é, adjunto de companhia) são majoritariamente definidos, por causa de sua topicalidade, humanidade ou de ambos; (b) os papéis de lugar e de tempo são majoritariamente definidos, por serem elementos de enquadre no discurso natural, normalmente estabelecidos antes de os principais participantes serem introduzidos; e (c) os papéis de instrumento e de modo são majoritariamente indefinidos, por causa de sua não topicalidade e não referencialidade. O único papel com definitude imprevisível é a de objeto/paciente. (GIVÓN, 2001, v. I, p. 473-474).

Como já mencionado, no modelo de Givón, a variação e a mudança linguística estão sempre presentes e a dinamicidade da língua pode ser entendida com base no compromisso adaptativo que decorre da atuação de pressões funcionais em competição. A compreensão da interação entre as diversas variáveis que determinam formas específicas de expressão linguística tem sido buscada desde o trabalho de Weinreich, Labov e Herzog (1968). No âmbito dos estudos sociolinguísticos desenvolvidos por Labov (1994, 2001, 2005), assume-se que a língua apresenta heterogeneidade ordenada e que a forma de compreender essa sistematicidade é analisar os fenômenos (as variáveis *dependentes*) com base em fatores intralinguísticos e extralinguísticos (as variáveis *independentes*). No presente trabalho, dá-se atenção primordialmente a fatores intralinguísticos, já que se trata de um *corpus* vinculado a um mesmo falante (Fernão Mendes Pinto), a uma mesma época (séc. XVI) e a um mesmo gênero textual (narrativa histórica). Avaliando-se as características sócio-históricas do autor do texto, apurou-se que a obra “provavelmente revela o uso linguístico escrito, mas com traços de oralidade, de um falante culto da variante da região de Lisboa (do dialeto padrão, portanto) na faixa etária de 60 anos, em um estilo mais ou menos informal” (CAMBRAIA, 2000, p. 1359).

A hipótese de trabalho que será testada no presente estudo é a de que *fatores de naturezas diversas atuam na determinação da presença ou ausência de artigo definido em sintagmas nominais com topônimo no português do séc. XVI (especificamente, na Peregrinação)*. Esse comportamento multifatorial decorre do compromisso adaptativo entre pressões funcionais em competição mencionado por Givón (2001, v. 1, p. 19).

4. Metodologia

O *corpus* adotado para o presente estudo foi o texto completo da *Peregrinação* de Fernão Mendes Pinto (ca. 1510-1583) presente na versão do primeiro estado da sua edição de 1614. Para o processamento do texto, o autor deste estudo realizou uma edição diplomática, ainda inédita, feita a partir do fac-símile digital do exemplar de cota 393924-C da Biblioteca Nacional Austríaca (PINTO, 1614). Ela se fez necessária diante da constatação da inexistência de edição digital fidedigna (CAMBRAIA, 2023) para estudos linguísticos e seguiu os critérios propostos por Cambraia (2005, p. 93-95), mas com a particularidade de realização de conjecturas sempre assinaladas em nota. Para a coleta de dados, empregou-se o programa *AntConc* versão 4.3.1 (ANTONY, 2024), tomando como primeira referência a listagem de topônimos e gentílicos apresentada na base *Pucau, Pucau?*⁵, elaborada a partir de Canosa Rodríguez (2017, p. 223-268).

Há, na *Peregrinação*, um sistema muito complexo de itens empregados para se referir a lugares, sobretudo os de origem estrangeira (que são a maioria absoluta). Esclarecem-se a seguir os principais problemas relativos à sua classificação e a solução adotada.

Um topônimo pode ocorrer sozinho em um SN (cf. ... *do esprital de [Lisboa]*) ou acompanhado de outros elementos (cf. ... *Lusões d[a ilha Borneo]*). Nos casos em que o indicador de entidade geográfica estivesse separado do topônimo por preposição, considerou-se, para a classificação da presença ou não de artigo definido, apenas o SN interno ao SP (cf. *o reyno d[o Chiammay]*).

Os SPs com topônimo podem ter função *indicativa* ou *possessiva*, como nos casos abaixo:

- (01) ... he hũa seita gentilica das principais *deste reyno da China*, como adiante direy quando vier a tratar do labarinto das trinta & duas leys que ha nelle. (f. 99rb17-21)

⁵ Disponível em: <https://www.pucau.org>.

- (02) ... & nos mandou logo dar embarcação para *a costa da China*, onde nos pareceo que achassemos nauios nossos em que nos fossemos para Malaca, ... (f. 157ra24-28)

No dado em (01), a preposição *de* na expressão *este reyno da China* tem a função de indicar o nome do reino (“o reino com nome de China”), enquanto essa mesma preposição na expressão *a costa da China* do dado em (02) tem a função de assinalar que a costa pertence à China (“a costa que está no território do reino da China”).

Considerou-se com função *indicativa* todo SP que acompanhasse os substantivos *império, reino, monarchia, anchacilado, senhorio(s), estado, prouincia, comarca, lugar, terra, cidade, villa, aldea, pouoaçaõ(oês), ilha, ilheo(s), lago e rio*; e com função *possessiva* todo SP que acompanhasse os substantivos *angra, bahia, barra, boqueyraõ, caez, calheta, campo, chifanga, confins, costa, enseada, esteiro, estreito, mar, mina, monte, morro, parcel, parte(s), picos, ponta, porto, praia, serra e tranqueyra*. As únicas exceções foram as expressões *ciudades da China* e *lugares da China*, cujo SP foi classificado como com função possessiva não só pelo substantivo estar no plural como também pelas informações disponíveis na própria obra. Nos casos em que havia combinação do primeiro tipo com o segundo (p. ex., *costa & ilha de Ainaõ*), deu-se prioridade à função indicativa. Essa decisão teve impacto também na análise de estrutura genitiva, já que os SPs encabeçados pela preposição *de* só foram considerados genitivos quando em função possessiva.

Há uma certa fluidez categorial na *Peregrinação*, segundo a qual determinadas formas são usadas seja como topônimo seja como gentílico:

- (03) E como *esta nação Iaoa* he grandissimamente cobiçosa, como lhe tratamos de seu interesse, ... (f. 231ra37-40)
- (04) ... nos deu hũ tempo tão forte ã não o podêdo payrar, nos foy forçado arribarmos *â Iaoa*, onde dos sete nauios de remo ã eramos se perderaõ os seys, ... (f. 226rb4-8)
- (05) ... onde o Rey de Demaa Emperador *desta ilha Iaoa* então se estaua fazendo prestes com hum exercito de oitocentos mil homês ... (f. 226vb1-5)

No dado em (03), está claro o fato de se tratar de gentílico, porque a forma *Iaoa* (feminino de *Iao*) está especificando *nação*. Já em (04), está claro consistir em topônimo, pois a forma *Iaoa* está indicando localidade, como se percebe por ser complemento do verbo de movimento *arribarmos*. Mas, no dado em (05), há certa ambivalência, porque se pode

interpretar *Iaoa* como topônimo ("ilha com o nome de Iaoa") ou como gentílico ("ilha pertencente à nação Iaoa").

A solução adotada foi a de considerar como topônimo toda forma que também é usada como gentílico adjunta a substantivo locativo (*cidade, ilha, império e reino*), a saber: *Chaleu, Pegú* (e variantes gráficas), *Sauady, Borneo, Iaoa, Iapaõ, Calaminhan, Arracão, Bata, Champaa e Tanguu*. Assim, no dado em (05), interpretou-se como topônimo (*Iaoa*) a forma adjunta a substantivo locativo (*ilha*). Para que essa ambivalência pudesse ser também considerada na análise, os dados também foram classificados levando em conta a forma ter sido ou não usada como gentílico na mesma obra.

Foram consideradas sempre como genticos, e não como topônimos, todas as formas no plural: p. ex., *ilha dos Lequios; terra dos Lauhòs, & Pafuaas, & Gueos; Rey dos Batas* etc. Também foram consideradas como genticos as formas presentes nas expressões *Raja Benão, Raja Sauady*, etc.

Não se consideraram dados relativos a nomes de casas reais ou religiosas (*casa de Meca; casa de Buatêdoo; casa Lechune; grande casa da Fiancima, & Tosa, & Bandou; casa do Fucheo*), por se tratar de uso metonímico, nem de edificações (*fortaleza Ternate; castello de Nixiamcoo; pagode de Tinagoogoo*; etc).

Aplicados esses critérios, o *corpus* para análise de topônimos ficou constituído de 2808 ocorrências distribuídas entre 598 topônimos⁶ (com as respectivas variantes formais): cf. a lista completa de formas consideradas na análise no anexo ao final deste trabalho.

Para analisar o comportamento dos topônimos foram consideradas as seguintes variáveis⁷:

- a) *Determinante*⁸: artigo definido, artigo indefinido, demonstrativo ou nenhum.
- b) *Tradição do topônimo*: referente à Europa, à América, à África ou à Ásia.
- c) *Indicador de entidade geográfica*: presente (no SN com o topônimo) ou ausente.

⁶ A base *Pucau, Pucau?* apresenta 688 itens. A diferença na quantificação se deve ao fato de essa base conter muitas ocorrências de genticos e de nomes de edificações, como pagodes, templos, ermidas, fortalezas etc, que não fizeram parte do presente estudo.

⁷ No estudo correlato sobre antropônimos na *Peregrinação*, foram consideradas dezesseis variáveis independentes (as onze últimas também foram contempladas no presente estudo, com as devidas adaptações): *tipo de antropônimo, gênero do antropônimo, forma de tratamento, qualificador, referenciador, tradição do antropônimo, indicador de posição/função social, adjetivo, possessivo, estrutura genitiva, construção apresentativa, estrutura do SN, coordenação, menção, função sintática e localização*. Não se constataram dados de sintagma nominal com topônimo e referenciador (*mesmo* ou *próprio*) na *Peregrinação*.

⁸ Adota-se aqui o termo *determinante* para designar uma subclasse dos especificadores formada por artigo definido, artigo indefinido e demonstrativo: eles apresentam a especificidade de nunca coocorrerem no mesmo SN no português.

d) *Polivalência*: item polivalente (usado na obra ora como gentílico ora como topônimo) ou não (usado apenas como topônimo).

e) *Adjetivo*: presente ou ausente.

f) *Possessivo*: presente ou ausente. Foram consideradas possessivas apenas as formas pronominais, e não as construções genitivas preposicionadas (*delle, della, etc*).

g) *Quantificador universal*: presente ou ausente.

h) *Estrutura genitiva*: presente ou ausente. Refere-se a todo SP encabeçado por *de*, mas com valor de posse.

i) *Construção apresentativa*: presente ou ausente. Foram consideradas como construções apresentativas aquelas compostas dos verbos *chamar, dizer e nomear* ou da expressão *por nome*.

j) *Estrutura do SN*: preposicionada ou não. Considerou-se como preposicionada quando há preposição ou locução prepositiva imediatamente antes do SN com topônimo, mesmo que a preposição introduzisse oração.

k) *Coordenação*: ausente ou presente como primeiro elemento ou como elemento subsequente.

l) *Menção*: única, primeira ou subsequente.

m) *Função sintática*: sujeito, predicativo do sujeito, objeto direto, objeto indireto, predicativo do objeto direto ou indireto (complementos de verbos de movimento foram classificados como objeto indireto), aposto, vocativo, agente da passiva, complemento nominal ou adjunto adnominal.

n) *Localização*: título de capítulo ou texto de capítulo⁹.

Maiores explicações sobre essas categorias serão fornecidas no comentário respectivo na análise dos dados.

Para calcular o peso relativo de cada variável, empregou-se o programa de análise multivariada *GoldVarb X* (SANKOFF; TAGLIAMONTE; SMITH, 2005). Esse programa informa também quando uma variável independente é estatisticamente significativa ou não.

⁹ Esta variável foi considerada para avaliar a hipótese de que o autor do título dos capítulos seria diferente do autor do texto dos capítulos (Francisco de Andrada para aquele e Fernão Mendes Pinto para este), mas três outros estudos (CAMBRAIA; CUNHA, 2023; CAMBRAIA; LEITE, 2023; CAMBRAIA, 2025) já ofereceram subsídios para a interpretação de ser apenas um o autor do título e o do texto (Fernão Mendes Pinto), embora com algumas poucas intervenções de um terceiro (Francisco de Andrada).

5. Descrição e análise dos dados

Considerando inicialmente os tipos de determinantes, os resultados obtidos foram os seguintes:

Tabela 3 – Topônimos por tipo de determinante

	Total	
	n	%
Sem nenhum determinante	1926	68.6
Com artigo definido	825	29.4
Com artigo indefinido	7	0.2
Com demonstrativo	50	1.8
Total	2808	100

Como se pode ver pela tabela 3, *o comportamento prototípico em relação a topônimos é não ocorrerem acompanhados de determinantes*. Essa mesma tabela permite perceber que a presença de artigo indefinido ou de demonstrativo é bastante rara com topônimos (esses casos perfazem apenas 2,1% das ocorrências). Uma vez que o foco do presente estudo recai sobre os artigos definidos, não se abordarão esses dois casos minoritários, mas exemplos de sua ocorrência podem ser visualizados nos dados abaixo:

- (06) ... porque se não ha de imaginar que he ella [hũa Roma], [hũa Constantinopla], [hũa Veneza], [hum Paris], [hum Londres], [hũa Seuilha], [hũa Lisboa], nem nenhũa de quantas cidades insignes ha na Europa por mais famosas & populosas que seião, ... (f. 125rb14-20)
- (07) ... em algũas partes vy grandissimas abundancias de diuersissimos mantimentos que não ha n[esta nossa Europa], mas em verdade affirmo, que não digo eu o que ha em cada hũa dellas, ... (f. 114va32-37)

Os sintagmas nominais com topônimos e artigo definido ou sem determinante são conjuntamente a grande maioria dos casos: 2751 ocs. A análise dos dados evidenciou que há alguns contextos exclusivos para cada caso, o que exige sua exclusão da análise multivariada (apresenta-se a seguir apenas um exemplo de cada caso).

Há, por um lado, a presença categórica de artigo definido nos dois seguintes contextos:

a) topônimo com adjetivo (10 ocs.):

- (08) E desta maneyra se acabaraõ de aueriguar todas as discordias & enfadamentos que [a triste Malaca] teue naquelle tempo. (f. 293ra20-23)

b) topônimo com possessivo (2 ocs.):

- (09) ... fique entendida a razão do pronostico, & do receyo em ã tantas vezes cõ gemidos & suspiros tenho apontado por parte d[*a nossa Malaca*], tão importãte ao estado da India, quanto (ao ã parece) esquecida daquelles de quẽ com razão diuera ser mais lãbrada, ... (f. 26vb9-16)

Por outro lado, há a ausência categórica de artigo definido na função de aposto (46 ocs.):

- (10) Seguindo daquy nosso caminho para diante por espaço de mais sete dias, chegamos a hum lugar, por nome [*Caleypute*], no qual os moradores delle nos não consentirão sayr em terra ... (f. 153va36-153vb1)

Pareceu conveniente excluir ainda alguns outros poucos casos, dada sua idiossincrasia. Não se considerou o caso único em que o topônimo ocorre na função de agente da passiva, com artigo, por se tratar de uso metonímico de um estado para um órgão estatal (*fazer ressenha geral da gente que pelo Vagaruu lhe he taxada*). Igualmente foram excluídos da análise dois tipos especiais de topônimos: (i) as duas ocorrências de topônimos formados por substantivo comum (nome de entidade geográfica) e adjetivo (*islas platarias* e *Villanoua*), que sempre ocorrem desacompanhados de artigo definido; e (ii) as dezesseis ocorrências de topônimos formados por substantivo comum (nome de entidade geográfica) e SP com substantivo comum, que se dividem em dois grupos, sempre com artigo (*ilha do fogo, ilha do ouro, ilha dos ladroẽs, ilha dos cocos, ilha da Madeyra, ilha das naos, rio do sal e rio das serpes*) e nunca com artigo (*Villa de Conde*). O fato de o substantivo comum ser inerente ao topônimo diferencia esses casos dos demais; além disso, não há variação quanto à presença ou não de artigo nesses casos.

Excluindo, portanto, as 76 ocs.¹⁰ referentes aos casos acima, obteve-se um *corpus* de 2675 ocs. com variação entre presença e ausência de artigo definido. Esses dados foram processados no GoldVarbX e duas das onze variáveis independentes restantes¹¹ não foram selecionadas como significativas: *construção apresentativa* e *localização*. As demais foram selecionadas na seguinte ordem decrescente: *polivalência, estrutura do SN, indicador de entidade geográfica, menção, estrutura genitiva, tradição, função sintática, quantificador universal* e *coordenação*.

No caso da variável *localização*, obteve-se nova evidência de que os títulos de capítulo terão sido redigidos pelo próprio Fernão Mendes Pinto, já que não há diferença linguística

¹⁰ Em uma ocorrência havia simultaneamente adjetivo e possessivo.

¹¹ As variáveis *adjetivo* e *possessivo* já tinham sido excluídas em função de presença categórica de artigo definido.

relevante entre eles e os textos de capítulo, diferença que se deveria esperar se tivessem sido escritos por um terceiro.

Quanto à variável *construção apresentativa*, causou surpresa não ter sido considerada significativa, já que há um padrão bem claro de predomínio de ausência de artigo nesse caso: 99,5% de ausência no caso desse tipo de construção contra 67,5% de ausência quando fora desse tipo. A prevalência da ausência se deve ao fato de que, nesse tipo de construção, o topônimo não tem propriamente uma função referencial (apontar para um referente) mas sim denominativa (nomear um referente), atuando quase como um adjetivo, e não como um substantivo, tal como se passa com os antropônimos. O único dado com presença de artigo nesse tipo de construção apresenta também dado correlato com ausência, respectivamente:

- (11) ... sendo tanto auante como o rio a que os naturaes da terra chamão Tinacoreu, & os nossos [*a varella*], lhe pareceo bem por conselho de algũs entrar dentro nelle ... (ff. 42ra35-42rb4)
- (12) Como Antonio de Faria chegou ao rio de Tinacoreu, a que os nossos chamão [*Varella*] ,... (f. 42ra19-21)

É interessante assinalar que o dado em (11) ocorre no texto de capítulo, com artigo, como nas demais oito ocorrências nesse contexto, mas o dado em (12) ocorre no título de capítulo, sendo a única ocorrência sem artigo junto a *Varella* em toda a obra. Aparentemente, isso deveria ter chamado a atenção na análise multivariada, mas não chamou, talvez por interferência de alguma das outras variáveis. Na análise dos antropônimos na mesma obra também se constatou a presença, igualmente rara, de artigo diante de nome próprio em construções apresentativas (5 dados perfazendo 2,6% do total de dados nesse contexto), mas a diferença é que o Goldvarb X selecionou como significativa essa variável no caso dos antropônimos (CAMBRAIA, 2025). No estudo de Mendes (2009), dados dessa natureza, chamados de *nominação didática e performativa*, foram excluídos da análise.

A exposição a seguir aborda, portanto, apenas as nove variáveis independentes consideradas significativas (a rodada eleita como a melhor apresentou o nível de significância 0.021).

A variável independente *tradição do topônimo* foi estabelecida para verificar se há comportamento diferenciado entre topônimos referentes à Europa (p. ex., *Toscana*), à América (p. ex. *Brasil*), à África (p. ex., *Cayro*) e à Ásia (p. ex., *China*). Essa variável tem certa compatibilidade com a variável *intimidade do falante em relação à pessoa mencionada em*

estudos de antropônimos e com a variável *cidade natal* no estudo de Mendes (2009). A expectativa era que os topônimos apresentariam uma gradação decrescente de favorecimento da presença de artigo na ordem *Europa > África > América > Ásia*, ou seja, quanto mais próximo do ponto de referência do autor (Portugal), mais conhecido seria o topônimo e mais frequente seria a presença do artigo. Os resultados obtidos foram:

Tabela 4 – Atuação da variável *tradição do topônimo* no comportamento de artigo definido na *Peregrinação*

		Presença	Ausência	N	%
Europa	N	9	83	92	3.4
	%	9.8	90.2		
	PR	0.164			
África	N	22	30	52	2.0
	%	42.3	57.7		
	PR	0.720			
América	N	1	1	2	0.1
	%	50.0	50.0		
	PR	0.921			
Ásia	N	768	1761	2529	94.5
	%	30.4	69.6		
	PR	0.509			
Total	N	800	1875	2675	
	%	29.9	70.1		

Como se vê, os dados apresentaram resultado inverso ao esperado, com os topônimos da Europa desfavorecendo fortemente (PR 0.164) a presença de artigo definido e os da África e da América favorecendo-a fortemente (PR 0.720 e 0.921, respectivamente), além de os da Ásia apresentarem influência quase nula (PR 0.509), sendo que deveriam desfavorecer mais do que os da África e da América, por se referirem a regiões menos conhecidas aos europeus do séc. XVI. Curiosamente, também na análise dos antropônimos dessa obra se identificou comportamento inesperado, com os antropônimos de tradição consolidada (judaico-cristã) desfavorecendo fortemente (PR 0.248) a presença de artigo definido e os de não consolidada (afro-asiática) favorecendo-a fortemente (PR 0.927). A explicação aventada para esse comportamento peculiar dos antropônimos talvez também se aplique aos topônimos, ou seja, a presença incomum de artigo definido em sintagmas nominais com topônimos afro-asiáticos seria em razão de maior sensibilidade do autor (entendida como o que faz sair de um estado de apatia) em relação a eles, já que foi justamente a viagem para o oriente que o tornou famoso (mesmo que certos elementos sejam ficcionais). Para os topônimos, não se pode descartar a hipótese de alguma influência da língua de origem do topônimo, já que, p. ex., o topônimo

Cayro na *Peregrinação* sempre ocorre acompanhado de artigo, tal como se dá no árabe (cf. القاهرة = al-Qāhirah), mas é possível que isso explique apenas alguns casos, já que o artigo definido não é categoria linguística universal e os topônimos da *Peregrinação* têm origem em diversas línguas.

Para a variável independente *indicador de entidade geográfica*, a expectativa era que a presença do nome de entidade geográfica favorecesse a presença de artigo definido, porque confere mais definitude ao referente. A distribuição dos dados quanto a este fator foi:

Tabela 5 – Atuação da variável *indicador de entidade geográfica* no comportamento de artigo definido na *Peregrinação*

		Presença	Ausência	N	%
Indicador presente	N	67	3	70	2.7
	%	95.7	4.3		
	PR	0.989			
Indicador ausente	N	733	1872	2605	97.4
	%	28.4	71.6		
	PR	0.470			
Total	N	800	1875	2675	
	%	29.9	70.1		

Os resultados foram o esperado: quando a entidade geográfica está presente (cf. *o reyno Champaa*), há um forte favorecimento da presença do artigo (PR 0.989), mas, quando está ausente, há desfavorecimento (PR 0.470), apesar de leve. A razão para esse favorecimento é o fato de o indicador aumentar a definitude, justificando mais a presença de artigo definido.

Para a variável independente *polivalência*, a expectativa era que itens polivalentes deveriam favorecer a presença de artigo definido, porque este, no singular, contribuiria para a identificação do emprego do item especificamente como topônimo. Os resultados obtidos foram:

Tabela 6 – Atuação da variável *polivalência* no comportamento de artigo definido na *Peregrinação*

		Presença	Ausência	N	%
Polivalente	N	339	238	577	21.6
	%	58.8	41.2		
	PR	0.752			
Não polivalente	N	461	1637	2098	78.4
	%	22.0	78.0		
	PR	0.424			
Total	N	800	1875	2675	
	%	29.9	70.1		

Os dados confirmaram a expectativa, pois topônimos polivalentes favorecem claramente a presença do artigo definido (PR 0.752), mas os não polivalentes a desfavorecem (PR 0.424).

Para a variável independente *quantificador universal*, a expectativa era que a presença do quantificador deveria favorecer a presença de artigo definido, porque a quantificação torna o referente mais definido. Os resultados obtidos foram:

Tabela 7 – Atuação da variável *quantificador universal* no comportamento de artigo definido na *Peregrinação*

		Presença	Ausência	N	%
Com quantificador	N	10	2	12	0.4
	%	83.3	16.7		
	PR	0.900			
Sem quantificador	N	790	1873	2663	99.6
	%	29.7	70.3		
	PR	0.498			
Total	N	800	1875	2675	
	%	29.9	70.1		

Os dados confirmaram a expectativa, pois a presença de quantificador favorece fortemente a presença do artigo definido (PR 0.900), enquanto sua ausência tem influência quase nula (PR 0.498).

Para a variável independente *estrutura genitiva*, a expectativa era que estar nesse tipo de estrutura favorecesse a presença de artigo, tal como já se havia constatado para antropônimos, mas também para topônimos (MENDES, 2009). A distribuição dos dados quanto a este fator foi:

Tabela 8 – Atuação da variável *estrutura genitiva* no comportamento de artigo definido na *Peregrinação*

		Presença	Ausência	N	%
Em estrutura genitiva	N	411	536	947	35.4
	%	43.4	56.6		
	PR	0.625			
Em estrutura indicativa	N	152	394	546	20.4
	%	27.8	72.2		
	PR	0.493			
Fora de estrutura genitiva ou indicativa	N	237	945	1182	44.2
	%	20.1	79.9		
	PR	0.402			
Total	N	800	1875	2675	
	%	29.9	70.1		

Os resultados foram o esperado, pois estar em uma estrutura genitiva favorece a presença do artigo definido (PR 0.625), enquanto estar em estrutura indicativa (PR 0.493) e

estar fora de qualquer uma delas (0.402) a desfavorece. É possível que o favorecimento tenha relação com um pressuposto assumido pelo falante, quando usa uma estrutura genitiva, de que o referente é já conhecido, ou seja, estruturas genitivas não seriam contextos comuns para a primeira menção de um topônimo.

Para a variável independente *estrutura do SN*, a expectativa era que a preposição favorecesse a presença de artigo, já que se trataria de uma questão rítmica, tal como aventado por Callou e Silva (1997) na análise de antropônimos. Os resultados obtidos foram:

Tabela 9 – Atuação da variável *estrutura do SN* no comportamento de artigo definido na *Peregrinação*

		Presença	Ausência	N	%
Precedido de preposição	N	785	1471	2256	84.3
	%	34.8	65.2		
	PR	0.565			
Não precedido de preposição	N	15	404	419	15.7
	%	3.6	96.4		
	PR	0.195			
Total	N	800	1875	2675	
	%	29.9	70.1		

Encontrou-se o esperado, com a existência de preposição precedendo um SN com topônimo favorecendo a presença de artigo definido (PR 0.565), embora de forma bastante sutil, e a inexistência desfavorecendo-a (PR 0.195). Este resultado é compatível com a afirmação de Cunha e Cintra (1985) de que a preposição tem influência no comportamento do artigo definido diante de topônimo, mas esses autores afirmaram, ao contrário, que ela desfavoreceria sua presença diante de nome de alguns países (*Portugal, França, Inglaterra*) e o que os dados acima mostram é que ela favorece. Não se pode descartar a possibilidade de que a ausência do artigo quando o SN não estava precedido de preposição se deva a casos de elisão com a vogal da palavra precedente (como no caso de substantivos, verbos etc.) não marcados como tal na escrita: como as preposições mais comuns não terminam em *a* ou *o*, os casos de elisão são evidentes com elas (*do = de + o*), mas não com outros itens (*todo = apenas todo* ou *tod'o*, isto é, *todo + o?*). Mendes (2009), p. ex., optou por eliminar da análise dados em que houvesse homofonia entre a vogal final de um item lexical e a forma do artigo, justamente por não poder saber com certeza, nos dados orais, se tinha havido elisão ou não.

Para a variável independente *coordenação*, a expectativa era que ser elemento subsequente de coordenação deveria desfavorecer a presença de artigo, porque a existência de

artigo nos itens antecedentes da coordenação já serviria para marcar a definição (haveria um “agasalhamento” da marcação de definitude). Os resultados obtidos foram:

Tabela 10 – Atuação da variável *coordenação* no comportamento de artigo definido na *Peregrinação*

		Presença	Ausência	N	%
Fora de coordenação	N	774	1640	2414	90.2
	%	32.1	67.9		
	PR	0.537			
Primeiro elemento de coordenação	N	23	56	79	3.0
	%	29.1	70.9		
	PR	0.480			
Elemento subsequente de coordenação	N	3	179	182	6.8
	%	1.6	98.4		
	PR	0.126			
Total	N	800	1875	2675	
	%	29.9	70.1		

Os resultados foram o esperado, com o contexto de ser elemento subsequente de coordenação desfavorecendo fortemente a presença de artigo definido (PR 0.126). Para explicar isso, convém considerar a noção de economia: se a definitude já está marcada no primeiro elemento, ela se estenderia aos demais sem a necessidade de se marcarem formalmente esses casos.

Para a variável independente *menção*, a expectativa era que ser menção subsequente deveria favorecer a presença de artigo definido. Os resultados obtidos foram:

Tabela 11 – Atuação da variável *menção* no comportamento de artigo definido na *Peregrinação*

		Presença	Ausência	N	%
Menção única	N	19	273	292	10.9
	%	6.5	93.5		
	PR	0.231			
Primeira menção	N	39	213	252	9.4
	%	15.5	84.5		
	PR	0.389			
Menção subsequente	N	742	1389	2131	79.7
	%	34.8	65.2		
	PR	0.554			
Total	N	800	1875	2675	
	%	29.9	70.1		

Os resultados foram o esperado, com o contexto de ser menção subsequente favorecendo a presença de artigo definido (PR 0.554), embora com valor não tão alto. Uma explicação para dar conta da atuação desse fator é o fato de um topônimo ter sido mencionado antes o tornar mais definido a partir de então, favorecendo a presença de artigo definido nas ocorrências seguintes.

Para a variável independente *função sintática*, a expectativa aqui era a de que as funções sintáticas de sujeito e objeto direto, que desempenham papel importante na coerência referencial segundo Givón (2001, v. I, p. 476), favoreceriam a presença de artigo definido. Os resultados obtidos foram:

Tabela 12 – Atuação da variável *função sintática* no comportamento de artigo definido na *Peregrinação*

		Presença	Ausência	N	%
Sujeito (S)	N	7	34	41	1.5
	%	17.1	82.9		
	PR	0.940			
Objeto direto (OD)	N	5	19	24	0.9
	%	20.8	79.2		
	PR	0.843			
Adjunto adverbial (AV)	N	65	169	234	8.7
	%	27.8	72.2		
	PR	0.566			
Objeto indireto (OI)	N	128	345	473	17.7
	%	27.1	72.9		
	PR	0.533			
Adjunto adnominal (AN)	N	583	1076	1659	62.0
	%	35.1	64.9		
	PR	0.497			
Complemento nominal (CN)	N	11	41	52	1.9
	%	21.2	78.8		
	PR	0.457			
Predicativo do objeto (PO)	N	1	191	192	7.2
	%	0.5	99.5		
	PR	0.214			
Total	N	800	1875	2675	
	%	29.9	70.1		

Os dados confirmaram a pertinência das funções de sujeito (PR 0.940) e de objeto direto (PR 0.843) no favorecimento da presença do artigo definido em sintagmas nominais com topônimo. Na análise de antropônimos na mesma obra, verificou-se que, de forma geral, funções argumentais favorecem a presença do artigo definido, mas, no caso de topônimos, não se encontra claramente o mesmo padrão, pois tanto a função de adjunto adnominal, não argumental, quanto a de complemento nominal, argumental, desfavorecem a presença do artigo (PR 0.497 e 0.457, respectivamente). Há até mesmo padrão oposto: nos antropônimos, a função de complemento nominal favorece (PR 0.615), enquanto, nos topônimos, desfavorece (PR 0.457). Com exceção das funções de sujeito, objeto direto e objeto indireto, que favorecem em ambos os casos, e das de adjunto adnominal e de predicativo do objeto, que desfavorecem também em ambos os casos, as demais comuns (complemento nominal e adjunto adverbial) não apresentam uma relação simétrica entre topônimos e antropônimos.

6. Antropônimos e topônimos na *Peregrinação*: semelhanças e diferenças

Comparando-se os resultados referentes ao comportamento de artigo definido em sintagmas nominais com topônimos ou antropônimos na *Peregrinação*, é possível verificar semelhanças e diferenças.

Os antropônimos e os topônimos na *Peregrinação* ocorrem predominantemente com ausência de artigo definido no sintagma nominal (64,2% e 68,4%, respectivamente). Considerando a média dos valores igualmente com ausência para as cidades mineiras de Abre Campo e de Matipó (28,5% para antropônimos e 48,7% para topônimos), fica evidente a existência de mudança linguística. Do séc. XVI para o séc. XXI houve um aumento na presença de artigo definido em sintagmas nominais, não apenas com antropônimos, como já tinham assinalado Callou e Silva (1997), mas também com topônimos: curiosamente, topônimos apresentam mais resistência ao processo de expansão da presença de artigo definido que os antropônimos, pois já no séc. XVI era mais comum a ausência naquele caso e ainda no séc. XXI, em Minas Gerais, se dá o mesmo.

Considerando as variáveis independentes comuns à análise de antropônimos e topônimos, constataram-se os seguintes padrões para a presença do artigo definido no sintagma nominal (que constituiu uma inovação do ponto de vista diacrônico):

Quadro 1 – Favorecimento da presença de artigo definido em sintagmas nominais com antropônimo ou topônimo na *Peregrinação*¹²

Domínio	Variável	Antropônimo	Topônimo
Pragmático	Tradição	(1º) afro-asiático	(8º) americano-afro-asiático
Semântico	Indicador	[função/posição social] (2º) presente	[entidade geográfica] (3º) presente
	Adjetivo	(4º) presente	presente*
	Possessivo	(11º) presente	presente*
	Estrutura genitiva	***	(5º) presente
Sintático	Construção apresentativa	(3º) ausente	***
	Função	(7º) argumental (S, OD, OI, PS, CN, AP) ¹³	(6º) argumental/não argumental (S, OD, OI) + (AV)
	Coordenação	(9º) fora ou início	(8º) fora
	Estrutura do SN	***	(2º) precedido de preposição
Discursivo	Menção	(8º) subsequente	(4º) subsequente
	Localização	***	***

¹² O número ordinal entre parênteses indica a ordem de seleção da variável pelo GoldVarb X: as posições ausentes do quadro se referem a variáveis privativas de antropônimos ou de topônimos. Um asterisco indica comportamento categórico e três asteriscos indicam variável estatisticamente não significativa.

¹³ PS = Predicativo do sujeito; AP = Agente da passiva.

Os dados do quadro acima permitem ver aspectos muito interessantes.

Primeiramente, a variável *localização* foi considerada não significativa para ambas as categorias, reforçando a tese de que os títulos de capítulos foram realmente de autoria do próprio Fernão Mendes Pinto (ainda que se saiba de evidência de intervenção eventual de terceiro).

Em segundo lugar, há, curiosamente, uma diferença bem nítida: preposições influenciam a presença de artigo definido em sintagma nominal com topônimo, mas não com antropônimo. Isso parece debilitar a tese de Callou e Silva (1997) de que se trate de uma questão rítmica: se se tratasse de uma questão puramente formal, deveria ocorrer em ambos os casos. O fato de a diferença também se relacionar à questão da estrutura genitiva (com preposição), irrelevante para os antropônimos e relevante para os topônimos, reforça que não se trata de questão puramente formal.

Terceiramente, a ordem de relevância na seleção das variáveis não é a mesma para as duas categorias. Por um lado, isso deve decorrer de as análises também terem considerado variáveis que eram privativas de cada uma dessas categorias (como forma de tratamento apenas para antropônimo e polivalência apenas para topônimo). Por outro lado, mesmo desconsiderando as posições que não constam do quadro (por se referirem às variáveis privativas), diferenças persistem: basta assinalar que adjetivo e possessivo desencadeiam presença categórica de artigo junto a topônimos, mas apenas favorecem junto a antropônimos.

7. Considerações finais

A análise multivariada do comportamento linguístico do artigo definido em sintagmas nominais com topônimos na *Peregrinação* permitiu constatar que se está diante de um fenômeno bem complexo, regulado por diferentes variáveis linguísticas.

A hipótese de que *fatores de naturezas diversas atuam na determinação da presença ou ausência de artigo definido em sintagmas nominais com topônimo no português do séc. XVI (especificamente, na Peregrinação)* foi confirmada, já que se constatou a atuação de fatores de natureza pragmática (*tradição do topônimo*), semântica (*indicador de entidade geográfica, adjetivo, possessivo, quantificador universal e estrutura genitiva*), sintática (*função, coordenação e estrutura do SN*) e discursiva (*menção*).

Tal como no caso dos antropônimos, percebe-se um processo de competição entre pressões funcionais, cujo resultado é fruto de um jogo complexo. Cunha e Cintra (1985, p. 216) já tinham chamado a atenção para a existência de razões diversas para o fato de nem sempre nomes próprios ocorrerem com artigo definido, o que seria o esperado, por se tratar de itens individualizantes e, portanto, definidos.

Referências

- AMARAL, E. T. R. *A ausência/presença de artigo definido diante de antropônimos em três localidades de Minas Gerais: Campanha, Minas Novas e Paracatu*. Dissertação (Mestrado em Estudos Linguísticos) — Universidade Federal de Minas Gerais, Belo Horizonte, 2003. Disponível em: <https://repositorio.ufmg.br/handle/1843/LHAM-5SMGZ9>. Acesso em: 10 jan. 2025.
- ANTHONY, L. *AntConc (version 4.3.1)*. Tokyo: Waseda University, 2024. Disponível em: <https://www.laurenceanthony.net/software>. Acesso em: 10 jan. 2025.
- CALLOU, D.; SILVA, G. M. O. O uso do artigo definido em contextos específicos. In: HORA, D. da (org.) *Diversidade linguística no Brasil*. João Pessoa: Ideia, 1997. p. 11-27.
- CAMBRAIA, C. N. Contributo para uma gramática do português clássico: a linguagem da *Peregrinação* de Fernão Mendes Pinto. In: CONGRESSO NACIONAL DA ABRALIN, II, Florianópolis, 25-27 fevereiro 1999. *Anais...* Florianópolis: Abralín, 2000. p. 1355-1362. Cd-rom.
- CAMBRAIA, C. N. *Introdução à crítica textual*. São Paulo: Martins Fontes, 2005.
- CAMBRAIA, C. N. Editometria: mensurando conjecturas nas edições da *Peregrinação*. *Filologia e Linguística Portuguesa*, São Paulo, v. 25, p. 9-30, 2023. DOI: <http://dx.doi.org/10.11606/issn.2176-9419.v25i1p9-30>. Acesso em: 10 jan. 2025.
- CAMBRAIA, C. N. Mecanismos funcionais e variação linguística no português clássico: artigo definido em sintagmas nominais com antropônimo na *Peregrinação*. *Laborhistórico*, Rio de Janeiro, v. 11, n. 1, 2025 (no prelo).
- CAMBRAIA, C. N.; CUNHA, E. L. T. P. Atribuição de autoria em discussão: o caso dos títulos dos capítulos da *Peregrinação*. *Confluência*, Rio de Janeiro, v. 64, p. 65-130, 2023. DOI: <https://doi.org/10.18364/rc.2023n64.1311>. Acesso em: 10 jan. 2025.

CAMBRAIA, C. N.; LEITE, R. C. S. Variação linguística no português quinhentista: contração entre *com* e *o(s)* na *Peregrinação*. *Laborhistórico*, Rio de Janeiro, v. 9, n. 2, p. e59885, 2023. DOI: <https://doi.org/10.24206/lh.v9i2.59885>. Acesso em: 10 jan. 2025.

CANOSA RODRÍGUEZ, A. J. *A identificação e referenciação de entidades geográficas mencionadas: o caso da Peregrinação de Fernão Mendes Pinto*. Tese (Doutorado em Geografia) — Universidade de Santiago de Compostela, Santiago de Compostela, 2017. Disponível em: https://citius.gal/static/a8ef53665d03c5f9387aeab30b830fe8/tese_final_xaviercanosa_20210618112119669_a169c7fa12.pdf. Acesso em: 10 jan. 2025.

CUNHA, C.; CINTRA, L. F. L. *Nova gramática do português contemporâneo*. 2. ed. 5. reimpr. Rio de Janeiro: Nova Fronteira, 1985.

GIVÓN, T. *Syntax: an introduction*. Amsterdam/Philadelphia: John Benjamins, 2001. 2 v.

LABOV, W. *Principles of linguistic change: internal factors*. Oxford/Cambridge: Blackwell, 1994.

LABOV, W. *Principles of linguistic change: social factors*. Oxford/Cambridge: Blackwell, 2001.

LABOV, W. *Principles of linguistic change: cognitive and cultural factors*. Oxford/Cambridge: Blackwell, 2005.

MARTELOTTA, M. E.; AREAS, E. K. A visão funcionalista da linguagem no século XX. In: CUNHA, M. A. F. da; OLIVEIRA, M. R. de; MARTELOTTA, M. E. *Linguística funcional: teoria e prática*. Rio de Janeiro: FAPERJ/DP&A, 2003.

MENDES, A. A. *A ausência ou a presença de artigo definido diante de nomes próprios na fala dos moradores da zona rural de Abre Campo e Matipó – M.G.* Dissertação (Mestrado em Estudos Linguísticos) — Universidade Federal de Minas Gerais, Belo Horizonte, 2003. Disponível em: <https://repositorio.ufmg.br/handle/1843/ARCO-7X3N4Z>. Acesso em: 03 jan. 2025.

MORAES, R. N. de; LIMA, A. F. de. Variação do artigo definido diante de nome próprio nas capitais do norte e nordeste do Brasil. In: LIMA, A. F. de; RAZKY, A.; OLIVEIRA, M. B. de (Orgs.). *Estudos geossociolinguísticos*. Campinas: Pontes, 2021. V. 3, p. 157-185.

NEVES, M. H. de M. *A gramática funcional*. São Paulo: Martins Fontes, 1997.

PINTO, F. M. *Peregrinação de Fernam Mendez Pinto* [...]. Lisboa: Pedro Crasbeeck, 1614. Disponível em: <http://data.onb.ac.at/rep/104A70DB>. Acesso em: 10 jan. 2025.

SANKOFF, D.; TAGLIAMONTE, S. A.; SMITH, E. *Goldvarb X: a variable rule application for Macintosh and Windows*. Toronto: Department of Linguistics – University of Toronto, 2005. Disponível em: <http://individual.utoronto.ca/tagliamonte/goldvarb.html>. Acesso em: 10 jan. 2025.

WEINREICH, U.; LABOV, W.; HERZOG, M. Empirical foundations for theory of language change. In: LEHMANN, W. P.; MALKIEL, Y. (Eds.) *Directions for historical linguistics*. Austin: University of Texas Press, 1968. p. 95-195.

Recebido em: 12 de janeiro de 2025.

Aceito em: 15 de fevereiro de 2025.

Anexo

Topônimos da *Peregrinação* no presente estudo¹⁴

I – Europa

Alcouchete, Alemanha, Alfama, Bargança, Buda, Castella, Cerdenha, Cezimbra, Cordoua, Couilham, Dinamarca, Espanha, Euora, Europa, Florença, França, Freixo de espada cinta, Guimaraës, (ilha da) Madeyra, Lisboa, Londres, Mediterraneo, Melides, Montemór o velho, Moscouia ~ Muscoo, Nazarè, Obidos, Oeyras, Orliens, Paris, Penamacor, Põte de Lima, Portugal, Rates, Roma, Santarem, Santiago ~ Santiago de Galiza, Santiago de Cacem, Setuual ~ Setuuel, Seuilha, Toscana, Veneza, Villa de Conde e Villanoua.

II – América

Brasil, noua Espanha e Panamá.

III - África

Africa, Arquico, Babylonia, Bitonto, Cairo ~ Cayro, Cambaya, Egypto, Ethiopia ~ Ethyopia ~ Etiopia, Gileytor, Gotor, Larache, Manamotapa, Massuaa, Melinde, Moçambique, S. Tome, Sinay, Suez, Tigremahom, Toro e Vangaleu.

IV – Ásia

Aachem ~ Achẽ ~ Achem, Aapessumhee, Aarù ~ Aarû, Abedalcuria, Adẽ ~ Adem, Agimpur, Ainão ~ Ainaõ, Alcocer, Amadabad, Amboyno, Anapleu, Anay, Anchepisaõ, Ancolaa, Andraguire ~ Andraguire ~ Andraguiree, Angegumaa, Angenia ~ Angenio, Angitur, Angunee, Ansedaa, Apefingau, Arabia Felix, Arimaa, Arissumhee, Armenia, Arracão, Asia, Auaa, Azebibe, Baarròs ~ Baarrós, Baçaim, Bãda ~ Banda, Bagou, Baguetor;

¹⁴ No caso de homônimos, colocou-se aqui numeração subscrita no seu final para diferenciá-los.