Revista GaLEGa de filoloxía
monografía 11

Aproximacións á variación lexical
no dominio galego-portugués

María Álvarez de la Granja
Ana Boullón Agrelo
Ernesto González Seoane (editores)

Revista GaLEGa de filoloxía
monografía 11

María Álvarez de la Granja
Ana Boullón Agrelo
Ernesto González Seoane (editores)

Aproximacións á variación lexical no dominio galego-portugués

ÁREA DE FILOLOXÍAS GALEGA E PORTUGUESA

Director: Xosé Ramón Freixeiro Mato (Universidade da Coruña)

Subdirector: Xosé Manuel Sánchez Rei (Universidade da Coruña)

Secretario: Xoán López Viñas (Universidade da Coruña)

Edita: Área de Filoloxías Galega e Portuguesa
Departamento de Galego-Portugués, Francés e Lingüística

Depósito Legal: C 1237-2017

ISBN: 978-84-9749-669-8

Distribúe: Consorcio Editorial Galego: pedimentos@coegal.com

Deseño: Torné Asociados

Maquetación: Antonio Souto

A edición e preparación deste libro foi financiada parcialmente cunha axuda da Secretaría Xeral de Universidades (Xunta de Galicia - Fondo Europeo de Desenvolvemento Rexional) ao grupo de investigación Filoloxía e Lingüística Galega (USC).

Índice

Limiar
María Álvarez de la Granja / Ana Boullón Agrelo / Ernesto González Seoane 7-9

Variação lexical açoriana: estudo dialetométrico do Atlas Linguístico-Etnográfico dos Açores
Fernando Brissos / Raïssa Gillier / João Saramago 11-28

Outros *Caminhos de Santiago*... os da Bahia (Brasil)
Suzana Alice Marcelino Cardoso / Ana Regina Torres Ferreira Teles 29-42

O campo semântico *corpo humano* no *Atlas Geossociolinguístico Quilombola do Nordeste do Pará* e no *Atlas Linguístico do Pará*
Marcelo Pires Dias / Marilúcia Barros de Oliveira 43-67

Vesgo, zarolho ou estrábico? O que dizem os dados do ALiB das regiões Norte e Sul do Brasil
Juliany Fraide Nunes e Aparecida Negri Isquerdo 69-88

Novos caminhos na morfologia lexical do Português: perspetiva histórica
Maria José Carvalho 89-114

O vocabulário dos tecidos no banco de textos do *Dicionário Histórico do Português do Brasil – séculos XVI, XVII e XVIII*: um recorte lexicográfico
Clotilde de Almeida Azevedo Murakawa 115-131

Arabismos no galego?
Juan A. Thomas 133-147

Notas sobre toponímia portuguesa medieval
Maria Alice Fernandes / Esperança Cardeira 149-174

Hagiotoponímia no Brasil: resultados de pesquisa
Ana Paula Mendes Alves de Carvalho / Maria Cândida Trindade Costa de Seabra 175-194

Limiar

Este volume reúne unha selección de nove contribucións que analizan, desde perspectivas diferentes e con metodoloxías así mesmo diversas, aspectos relacionados coa variación lexical no galego, no portugués europeo e no portugués brasileiro. Os textos comparten, ademais, outro elemento común: o feito de partiren de traballos presentados en *Gallaecia - III Congresso Internacional de Linguística Histórica*, que se celebrou entre os días 27 e 30 de xullo de 2015 en Santiago de Compostela en homenaxe aos profesores Ramón Lorenzo e Antón Santamarina.

Aínda que o volume non está formalmente dividido en seccións, as contribucións que o integran preséntanse agrupadas e ordenadas en función da súa afinidade temática. Un primeiro grupo está constituído por traballos do ámbito da xeografía lingüística, un deles encadrable na dialectometría, os outros tres na dialectoloxía clásica.

No primeiro, "Variação lexical açoriana: estudo dialetométrico do *Atlas Linguístico-Etnográfico dos Açores*", elaborado por Fernando Brissos, Raïsa Gillier e João Saramago, preséntase unha proposta de clasificación dialectal dos Azores a través dunha análise dialectométrica. Esta baséase nos datos lexicais recollidos nos dezasete puntos que constitúen a rede do *Atlas Linguístico-Etnográfico dos Açores*. Os resultados permiten establecer unha primeira división lingüística e unha subdivisión en nove grupos, que coinciden con cada unha das illas que conforman o arquipélago.

Os outros tres traballos que integran este bloque adoptan un enfoque máis tradicional ou clásico. Así, o asinado polas profesoras Suzana Cardoso e Ana Regina Teles, "Outros *Caminhos de Santiago*... os da Bahia (Brasil)", analiza a distribución no estado da Bahia dos resultados obtidos para o concepto 'vía láctea' en dúas obras capitais da xeografía lingüística brasileira: o *Atlas Prévio dos Falares Baianos* (publicado en 1963 sobre datos obtidos nos comezos dos anos 60) e o *Atlas Linguístico do Brasil* (cuxos dous primeiros volumes viñeron a lume en 2014 con materiais colectados nos comezos do século XXI).

A seguir, Marcelo Pires Dias e Marilúcia Barros de Oliveira analizan e comparan en "O campo semântico *corpo humano* no *Atlas Geossociolinguístico Quilombola do Nordeste do Pará* e no *Atlas Linguístico do Pará*" os resultados obtidos para este ámbito temático en ambos os dous atlas. A primeira das obras citadas recolle as variedades lingüísticas dos quilombolas, descendentes dos africanos escravizados no Brasil durante os séculos XVII, XVIII e XIX. O obxectivo fundamental é contrastar a presenza de formas de orixe africana en comunidades quilombolas e non quilombolas.

O grupo de traballos sobre xeografía lingüística péchase coa contribución de Juliany Fraide Nunes e Aparecida Negri Isquerdo, "*Vesgo*, *zarolho* ou *estrábico*? O que dizem os dados do ALiB das regiões Norte e Sul do Brasil". As autoras analizan as diferentes designacións encontradas para as persoas estrábicas nas rexións brasileiras sinaladas. Os datos, inéditos e extraídos do *Atlas Linguístico do Brasil*, analízanse desde o punto de vista distribucional, etimolóxico, semántico e lexicográfico.

O segundo grupo reúne tres traballos elaborados desde unha perspectiva diacrónica, pois están centrados no estudo histórico ou etimolóxico do léxico.

En primeiro lugar, a profesora Maria José Carvalho ofrécenos en "Novos caminhos na morfologia lexical do Português: perspetiva histórica" un documentado estudo sobre a evolución e produtividade dunha serie de sufixos, a maior parte deles formadores de substantivos e adxectivos. Parte para iso de formas derivadas extraídas de textos datados entre 1289 e 1565 pertencentes ao corpus documental do mosteiro de Santa Maria de Alcobaça.

En "O vocabulário dos tecidos no banco de textos do *Dicionário Histórico do Português do Brasil – séculos XVI, XVII e XVIII*: um recorte lexicográfico", a profesora Clotilde de Azevedo Murakawa parte dun conxunto de oitenta e tres unidades lexicais designadoras de tecidos ou panos, extraídas do corpus documental compilado para a elaboración do citado dicionario histórico, e centra a súa análise nunha selección dalgunhas destas voces, con especial atención á súa fortuna lexicográfica.

O terceiro traballo encadrable neste grupo, "Arabismos no galego?", asinado por Juan A. Thomas, analiza desde unha perspectiva crítica unha listaxe de dezasete formas exclusivas do galego consideradas arabismos no *Diccionario de arabismos y voces afines en iberorromance* (Corriente, 1999). As voces, das cales só sete son clasificadas polo autor como arabismos directos, son estudadas etimolóxica e semanticamente e nesta análise póñense en relación con outras formas das restantes linguas peninsulares, especialmente do castelán e do portugués.

O último dos grupos toma como base o material toponímico, adoptando unha perspectiva histórica e analizando os resultados dialectais en territorio portugués e brasileiro, respectivamente.

No traballo de Maria Alice Fernandes e Esperança Cardeira, "Notas sobre toponímia portuguesa medieval", realízase un achegamento á paisaxe toponímica da Idade Media portuguesa, poñéndoa en relación cos fenómenos de repoboación na época da Reconquista. A tal fin, as autoras estudan a toponimia dos asentamentos rurais (exemplificando cos descendentes e derivados de diferentes voces, como as latinas QUINTANA ou CASALE), así como os topónimos con orixe en (MALA) MATTIANA. É de destacar o uso de programas de xeorreferenciación toponímica, que permite a elaboración de mapas que ilustran eficazmente o traballo.

Finalmente, Ana Paula Mendes Alves de Carvalho e Maria Cândida Trindade Costa de Seabra presentan en "Hagiotoponímia no Brasil: resultados de pesquisa" as conclusións dun estudo descritivo dos haxiotopónimos no estado brasileiro de Minas Gerais, vinculado ó proxecto *Atlas Toponímico do Estado de Minas Gerais*. No traballo téñense conta os datos actuais e os fornecidos por mapas históricos dos séculos XVIII e XIX, o que permite descubrir que os haxiotopónimos constitúen máis do 6% do total de voces toponímicas actuais, así como a súa diminución a través dos tempos. Analízanse tamén as distintas proporcións de nomes de santos, santas e advocacións marianas.

Os nove traballos que conforman este volume constitúen unha mostra evidente da pluralidade de perspectivas (e métodos) con que se pode abordar a variación lexical, así como da diversidade de criterios selectivos que se poden empregar, illadamente ou en combinación, para o establecemento de obxectos de estudo abarcables polo investigador: o criterio semántico, distribucional, designativo, etimolóxico, formal, textual... Son nove interesantes aproximacións que nos abren novos camiños no terreo do léxico galego e portugués e que nos revelan o carácter poliédrico das palabras, a súa riqueza e a súa complexidade.

María Álvarez de la Granja
Ana Boullón Agrelo
Ernesto González Seoane

Variação lexical açoriana: estudo dialetométrico do Atlas Linguístico-Etnográfico dos Açores

Fernando Brissos[1]
fernando.brissos@gmail.com

Raïssa Gillier
raissa.gillier@gmail.com

João Saramago
japsaramago@gmail.com

Universidade de Lisboa, Faculdade de Letras, Centro de Linguística, Complexo Interdisciplinar da Universidade de Lisboa

Resumo:

Este estudo apresenta uma primeira proposta de classificação dialetal do espaço linguístico açoriano através de uma abordagem dialetométrica ao material lexical recolhido nos 17 pontos de inquérito que formam a rede do Atlas Linguístico-Etnográfico dos Açores. O material que possibilitou este trabalho é o que foi obtido como resposta a 257 conceitos na totalidade dos pontos e a metodologia de análise, inserida no paradigma da Escola Dialetométrica de Salzburgo, é a preconizada pelo programa Diatech, que apresenta uma solução inovadora para o problema das respostas múltiplas. Através das diferentes análises permitidas pela Diatech, foi possível determinar: (i) uma primeira divisão linguística, por grupos de ilhas, que corresponde aos 3 grupos geográficos que formam o arquipélago – oriental, central e ocidental; (ii) uma subdivisão linguística dos 17 pontos em 9 grupos que correspondem exatamente às 9 ilhas do arquipélago. A análise sinótica, realizada pela aplicação do coeficiente de assimetria de Fisher, mostrou que a ilha que mais se distingue das restantes é a Terceira e que as ilhas que mais semelhanças apresentam relativamente às outras são o Corvo e Santa Maria. Esta primeira proposta só ficará completa e definitiva com idêntico estudo de outras áreas gramaticais dos dialetos açorianos.

Palavras-chave:

Dialetometria, dialetologia portuguesa, dialetos açorianos, léxico.

Sumário:

1. Introdução. 2. Metodologia. 3. Resultados. 4. Conclusão.

1 Financiado pela Bolsa de Pós-doutoramento da Fundação para a Ciência e a Tecnologia (Governo de Portugal) com a referência SFRH/BPD/78479/2011.

Azorean lexical variation: dialectometrical study of the Linguistic and Ethnographic Atlas of the Azores

Abstract:

This paper presents a first proposal for a dialectal classification of the Azorean linguistic space through a dialectometrical analysis on lexicon collected in the 17 enquiry points of the Linguistic and Ethnographic Atlas of the Azores. The material enabling this study was taken from a total of 257 concepts covering all geo-points. The data were analysed according to the Salzburg Dialectometrical School paradigm, in Diatech, a new tool that offers an innovative solution for the problem posed by multiple responses. Through different analyses enabled by Diatech, it was possible to determine: (i) a primary linguistic division according to groups of islands, corresponding to the 3 geographical groups that constitute the archipelago – oriental, central and occidental; (ii) a linguistic subdivision of the 17 points into 9 different groups, referring exactly to the 9 islands of the archipelago. Taking Fisher's skewness coefficient, the comparative analysis between localities shows that Terceira is the most distinctive of all islands, while Corvo and Santa Maria are the ones that display greater similarities to the entire set of islands. In order for this proposal to be complete and definitive, however, it is still necessary to carry out an identical study on other grammatical areas of the Azorean dialects.

Key words:

Dialectometry, Portuguese dialectology, Azorean dialects, lexicon.

Contents:

1. Introduction 2. Methodology 3. Results 4. Conclusion.

1. Introdução

Ao longo do tempo, a dialetologia portuguesa focou-se no estabelecimento de fronteiras dos dialetos continentais, menosprezando, de certo modo, as características particulares dos dialetos insulares (Açores e Madeira). De facto, assumiu-se que os 'falares' dos arquipélagos correspondiam a um "prolongamento do grupo dos dialetos centro-meridionais" (Cunha, & Cintra, 1984: 19; veja-se também Segura, 2013: 105). O pressuposto de que os dialetos insulares pertenciam ao conjunto dos dialetos do sul do país assentou no facto de os primeiros coincidirem com os segundos relativamente a alguns traços fundamentais de ambos. No quadro da classificação dialetal do português proposta por Cintra (1983b), os dialetos insulares têm mesmo uma concordância quase total com o sul do país no que toca aos traços diferenciadores que permitiram estabelecer a oposição entre os dois grupos principais de dialetos (os setentrionais e os centro-meridionais) considerados pelo autor. Contudo, a especificidade de alguns traços presentes nos dialetos dos arquipélagos afasta-os dos dialetos continentais, devendo ser considerados um grupo principal dos dialetos do português europeu, a par dos dialetos continentais setentrionais e dos dialetos continentais centro-meridionais (Segura, & Saramago, 1999; Segura, 2013). De facto, tal como qualquer falante nativo tem sensibilidade para reconhecer, de forma impressionista, que uma pessoa é proveniente do norte ou do sul, por exemplo, não é menos sensível à proveniência de alguém dos arquipélagos, embora algumas ilhas sejam mais 'marcadas' que outras, facto que sustenta a peculiaridade destes dialetos.

Temos portanto que pelo menos a presença de fenómenos fonéticos específicos, inexistentes no território continental, confere aos dialetos insulares um estatuto singular e independente que não deve ser ignorado. Não obstante, os dialetos açorianos ainda carecem de uma descrição aprofundada das suas principais características, bem como de uma análise de conjunto, no sentido de se poder chegar a uma proposta de classificação do sistema dialetal dos Açores. Com efeito – e enfatizamos esse aspeto –, apesar dos avanços significativos que se têm levado a cabo no conhecimento desses dialetos nas últimas duas ou três décadas, ainda não foi possível concretizar uma classificação dialetal do sistema referido. A esquematização mais apurada é a de Segura (2013: 110-116), que, a partir de traços fonéticos, identifica dois grupos dialetais que, "pela especificidade e regularidade de alguns dos seus traços" (pág. 110), são os mais demarcados dos restantes. São eles os dialetos micaelenses e terceirenses, correspondentes às ilhas de São Miguel e Terceira, respetivamente[2]. Nos primeiros, destaca-se

2 Deve notar-se que a pertença destes dialetos às ilhas mencionadas não deve ser entendida em sentido restrito e exclusivo, uma vez que, na sequência do que dissemos, não é ainda possível delimitar com precisão as fronteiras dialetais açorianas.

um ‘deslocamento em cadeia’ do vocalismo tónico, ao passo que nos segundos o fenómeno mais distintivo é a harmonização que as vogais tónicas sofrem por influência da qualidade da vogal ou semivogal pretónicas. Remetemos para a obra citada de Segura para a descrição desses e dos outros traços fonéticos que sustentam a visão estabelecida dos dialetos portugueses em geral e açorianos em particular, relembrando os seus principais eixos: (i) traços fonéticos descritos pela metodologia da dialetologia tradicional; (ii) ao contrário do que ocorreu durante a maior parte do tempo (tendo em conta apenas, naturalmente, o que respeita à dialetologia portuguesa *científica*, iniciada no final do séc. XIX), considera-se hoje que os dialetos insulares não são um mero prolongamento dos dialetos do sul de Portugal continental; (iii) são de destacar, no caso dos Açores, duas variedades especialmente idiossincráticas (S. Miguel e Terceira); (iv) inexistência de uma classificação dialetal de nenhum dos arquipélagos.

Tendo em consideração as principais virtudes (conhecimento fonético razoável; estudos continuados ao longo de décadas) e as principais lacunas (estudo quase exclusivo da fonética, no paradigma metodológico tradicional; ausência de, se não uma classificação dialetal, pelo menos uma sistematização detalhada) do estado da arte do estudo dos dialetos açorianos, o presente trabalho analisa a variação lexical desses dialetos a partir de uma abordagem dialetométrica. O uso dessa abordagem procura, paralelamente, colmatar a escassez de trabalhos do tipo desenvolvidos para o português europeu. De facto, apesar dos excelentes resultados atingidos no estudo de várias línguas e da tradição importante que já tem em vários países (vejam-se p. ex. Goebl, 2006 e Wieling, & Nerbonne, 2015), a análise dialetométrica tem ainda uma tradição diminuta no caso português; no que toca à dialetometria pós-clássica, i.e. a dialetometria desenvolvida a partir dos anos 1990’ com recurso a *software* específico, a tradição é mesmo inexistente. Pretendemos também, por outro lado, analisar a variação dialetal numa perspetiva lexical, uma vez que, apesar de a classificação dos dialetos portugueses ter assentado sempre em traços fonéticos e/ou fonológicos, sabemos que a variação existe, e se faz sentir decisivamente na perceção dos falantes, não apenas a nível fonético, mas também lexical, sintático, etc.

Nessa sequência, o nosso trabalho girará em torno de duas questões principais: *a*) descrição dos aspetos mais importantes da variação lexical açoriana no espaço, particularmente no que respeita à identificação dos principais subgrupos dialetais da região; *b*) comparação desses resultados com os dados atuais relativos à subdivisão dialetal dos Açores. Esta análise resultará numa contribuição para o desenvolvimento dos estudos dialetais no sentido de se poder chegar a uma proposta de classificação do sistema dialetal dos Açores que não abranja exclusivamente a fonética nem parta apenas da metodologia da dialetologia tradicional.

2. Metodologia

Utilizamos como *corpus* o Atlas Linguístico-Etnográfico dos Açores (ALEAç), parte integrante do atlas nacional, i.e. o Atlas Linguístico-Etnográfico de Portugal e da Galiza (ALEPG)[3]. O ALEAç tem por isso o perfil de informante clássico da geografia linguística europeia: conservador, com pouca ou nenhuma escolarização, acima dos 50 anos de idade, ligado aos ofícios tradicionais, nascido e com poucas ou nenhumas ausências do ponto de inquérito respetivo, etc. Trata-se, em suma, do perfil de informante que melhor representa o estado de língua mais ligado a cada ponto de inquérito, permitindo-nos, ponto de inquérito a ponto de inquérito, catalogar de forma eficiente a variação linguística ao longo da área pretendida. O mapa 1 apresenta os 17 pontos de inquérito do ALEAç, cuja abrangência é significativa: nenhuma ilha está ausente, tendo as ilhas maiores mais de um ponto inquirido.

Também os conceitos do questionário utilizado pelo ALEAç se identificam com os procedimentos estabelecidos na geografia linguística europeia, na medida em que dizem respeito, maioritariamente, ao léxico tradicional: agricultura, ofícios, fauna e flora, etc. Para a listagem dos conceitos do ALEAç, cf. a página *web* do projeto (ver Referências Bibliográficas).

O *corpus* do ALEAç foi utilizado neste trabalho de forma exaustiva, tanto no que respeita aos pontos de inquérito como aos conceitos. Dos 955 conceitos que preenchem os volumes I a VIII do atlas[4], resultaram, preenchidos os critérios de exclusão, 257 conceitos. A determinação dos conceitos a utilizar foi efetuada a partir de três princípios *standard*: *a*) os conceitos sem resposta em pelo menos um dos pontos de inquérito foram excluídos, de modo a não distorcer a análise estatística; *b*) igualmente foram excluídos os conceitos mononímicos (i.e. os conceitos com a mesma resposta em todos os pontos de inquérito), dado que não acrescentam à identificação de agrupamentos dialetais; *c*) o critério de diferenciação de cada forma é o etimológico, ficando, assim, excluídos os aspetos fonéticos e morfossintáticos (ex. *cão*, *cãozinho*, *cachorro* (3 respostas) → *cão-cãozinho* (que se pode agrupar no lema *cão*) e *cachorro* (2 formas diferentes)).

A análise dialetométrica foi efetuada no *software* Diatech (Aurrekoetxea *et al.*, 2013; veja-se a página *web* do *software* nas Referências Bibliográficas), que, para além de permitir a criação do mapa-base dos pontos de inquérito diretamente na aplicação,

3 Veja-se, sobre o ALEAç, a sua página *web*, que disponibiliza todos os materiais (cf. Referências Bibliográficas). Sobre o ALEPG, a breve apresentação constante da página *web* do Centro de Linguística da Universidade de Lisboa (cf. Referências Bibliográficas), Saramago (2006) e o vol. I do questionário dos inquéritos (cf. Referências Bibliográficas). Os dados do ALEPG foram recolhidos entre 1973 e 2004, tendo os materiais relativos aos Açores sido obtidos entre 1977 e 1996.

4 O volume restante, o n.º IX, não foi utilizado no presente estudo porque a sua rede de pontos de inquérito não é idêntica à dos outros volumes, dado que trata o campo semântico dos animais e plantas marinhos.

aborda de forma inovadora o problema (clássico e delicado em dialetometria) das respostas múltiplas (i.e. a existência de mais de uma resposta a um dado conceito em um dado ponto de inquérito), propondo um algoritmo melhorado (cf. Aurrekoetxea *et al.*, 2013: 26-27). Não são difíceis de encontrar os atlas linguísticos, sobretudo entre os menos recentes, em que a existência de respostas múltiplas é diminuta, sendo por isso possível descartá-las e utilizar uma matriz de dados binária, ou seja, uma matriz simples que apenas considera uma variante por localidade. É este, de resto, o princípio da mais importante tradição dialetométrica, a da Escola de Dialetometria de Salzburgo (EDS), fundada e desenvolvida por Hans Goebl (vejam-se p. ex. Goebl, 2006; Goebl, 2010). No caso do ALEAç, contudo, as respostas múltiplas têm uma dimensão que não permite que sejam rejeitadas. Dois índices mostram-no para lá de qualquer dúvida:

(i) O peso percentual de respostas múltiplas nos nossos dados é de 25,77%. (25,77% = 1126 / 4369, em que 1126 é o número de respostas múltiplas e 4369 o número total de respostas; 4369 é atingido multiplicando 257 (número de conceitos) por 17 (número de pontos de inquérito). Ou seja, das 4369 respostas dos nossos dados, 1126 têm mais de uma forma – sejam duas, três, etc.)

(ii) 89,11% dos conceitos utilizados têm pelo menos uma resposta múltipla. (89,11% = 229 / 257, em que 229 é o número total de conceitos com mais de uma forma como resposta em pelo menos um ponto de inquérito e 257 é o número total de conceitos.)

Na análise dos dados, que passamos a ver, seguimos a metodologia da EDS.

3. Resultados

Há duas forças principais, em parte contraditórias mas que não deixam de ser de natureza fundamentalmente geográfica, que explicam a variação lexical nos dialetos açorianos: uma *superficial* e previsível, a outra latente e imprevisível. A primeira força localiza-se no plano da geografia *macro* e é a região em que os locais estão inseridos: as semelhanças/diferenças lexicais entre locais estão indexadas, em primeiro lugar, às relações de proximidade geográfica, na medida em que os locais próximos geograficamente tendem a ser mais próximos linguisticamente (entenda-se, no nosso caso, lexicalmente) do que os locais distantes[5]. A segunda força está situada

5 A noção de que as proximidades lexicais entre os dialetos açorianos decrescem, em geral, com o aumento das distâncias geográficas tinha já sido verificada, a partir de dados dialetométricos obtidos num *corpus* mais reduzido do ALEAç, por Saramago, & Gonçalves 2003. Os dados desse estudo apresentam algumas diferenças menos relevantes com o presente, mas devem ser tidos em conta dois factos: (i) ao contrário do nosso, o estudo referido segue uma abordagem interpontual (i.e. a análise dos contactos entre pontos confinantes, e não uma síntese de todas as semelhanças e diferenças entre todos os pontos); (ii) ao passo que o nosso *corpus* aproveita exaustivamente o ALEAç, o do estudo referido utiliza menos conceitos

no plano da geografia *micro* e pode ser definida como a autonomia ou isolamento que a insularidade confere a cada ilha, individualidade essa que pode, naturalmente, variar com o contexto histórico, social, etc., e permite à ilha exceder a coerência geográfica em que está inserida seguindo um caminho evolutivo independente. Analisaremos seguidamente exemplos de cada uma dessas forças, começando pela identificação dos principais grupos dialetais do arquipélago.

A técnica dialetométrica por excelência que nos permite identificar ou decompor[6] uma dada área em grupos dialetais é a *análise cluster* ou *dendrográfica*, que segmenta, num diagrama em árvore, os núcleos dialetais num nível crescente de homogeneidade a partir do tronco da árvore. Assim, quanto mais avançada na ramificação da árvore a subdivisão é, mais homogeneidade existe entre os elementos que a compõem. O mapa 2 (ver Anexo) apresenta o resultado da aplicação da análise dendrográfica[7] aos nossos dados, com uma linha de corte (subdivisão dos grupos a representar espacialmente no mapa) de três grupos. O estabelecimento da linha de corte em três grupos é, de todas as hipóteses (dois a dezasseis grupos, dado que utilizamos dezassete pontos de inquérito), a mais interessante, por três razões: *a*) corresponde ao número de agrupamentos dialetais primários dos Açores que, em sentido lato, podemos identificar na visão estabelecida na literatura, construída a partir de dados fonéticos tal como analisados pela metodologia tradicional[8]; *b*) corresponde ao número de grupos geográficos em que os Açores convencionalmente são divididos (o grupo ocidental, o central e o oriental); *c*) corresponde igualmente à divisão administrativa anterior à autonomia dos arquipélagos, que se baseava em três distritos (Ponta Delgada, a que pertenciam as ilhas de Santa Maria e São Miguel; Angra do Heroísmo, com as ilhas da Terceira, Graciosa e São Jorge; e Horta, com Faial, Pico, Flores e Corvo).

Os dados são claros: a subdivisão dos Açores que vemos no mapa 2 coincide totalmente com a subdivisão geográfica tradicional – e diverge, por isso, da visão dialetológica tradicional, tal como da divisão administrativa pré-autonomia. As ilhas do Corvo e Flores, que constituem o grupo geográfico do ocidente, formam um

(100, contra os nossos 257) e menos campos semânticos, para além de que faz um aproveitamento diferente das respostas múltiplas. Isso faz com que uma comparação dos resultados dos dois estudos não seja praticável por agora, mas, como vimos, o nexo geográfico geral está verificado em ambos.

6 Não "delimitar", no sentido clássico da dialetologia; isso é feito, acima de tudo, pela análise interpontual (Goebl, & Smečka, 2014: 468).

7 O algoritmo que aplicamos à análise dendrográfica – i.e. o algoritmo que classifica os pontos de inquérito numa hierarquia aglomerativa – neste trabalho é o chamado Ward, de acordo com o seu autor: Joe Ward, Jr. Trata-se do algoritmo cujos resultados têm sido mais satisfatórios nos trabalhos da EDS; vejam-se sobre o assunto, entre outros, Álvarez, Dubert, & Sousa (2006: 485-487), Goebl (2012: 153) e Goebl, & Smečka (2014: 460).

8 Recorde-se o que foi dito na secção 1: na visão referida, são destacados dois grupos de dialetos, os micaelenses e os terceirenses; por omissão, resta outro conjunto de dialetos, que, mesmo não formando um verdadeiro grupo dialetal, não deixa de representar o terceiro elemento.

grupo dialetal; as ilhas de São Miguel e Santa Maria, o grupo oriental, são também um grupo dialetal; e o grupo central, constituído pelas ilhas do Faial, Pico, São Jorge, Terceira e Graciosa, corresponde ao restante segmento dialetal.

Ao concordar de forma tão próxima com a geografia, o léxico açoriano diverge, como se disse, da fonética (tal como ela nos tem sido descrita, é importante ter presente, pela dialetologia tradicional). Sem embargo, se analisarmos mais proximamente a árvore dos dialetos açorianos encontramos um aspeto que implica uma concordância com a fonética que não deve ser ignorada: o grupo oriental é o mais individualizado do conjunto, i.e. o primeiro entroncamento da árvore dá-se na divergência centro e oeste, por um lado, e oriente, por outro. Ora, é um facto bem conhecido que é na zona oriental, concretamente em São Miguel, que, de acordo com a perceção geral do falante português (linguista ou não), existem os dialetos mais destacados do sistema açoriano, tal como do português europeu em geral. O mapa 3 representa essa subdivisão.

Um aparente problema da subdivisão em dois grupos é que ela emparelha a ilha de Santa Maria com a de São Miguel, o que não sucede, de acordo com os dados conhecidos, a nível fonético. Só dispondo de dados fonéticos dos Açores trabalhados dialetometricamente poderemos fazer uma comparação perfeita léxico-fonética, mas dois fatores podem facilmente explicar essa divergência. Divergência, note-se, que não está pouco sustentada, pois os traços fonéticos tradicionalmente utilizados para caracterizar São Miguel – não registados em Santa Maria – são extremamente salientes no panorama dialetal português, conferindo à maior ilha do grupo oriental a posição destacada que referimos acima.

O primeiro fator que pode explicar a eventual divergência léxico-fonética no grupo oriental é que, como se sabe, léxico e fonética são áreas gramaticais que apresentam frequentemente comportamentos geográficos diferentes. Tenha-se presente, por exemplo, o caso do português continental, em que as áreas lexicais definidas por Cintra não conferem na sua plenitude com as áreas que o mesmo autor definiu a partir de traços de ordem fonética[9]. Mais do que isso, aliás, no léxico, como refere Goebl (2010:13), "la intensidad de un contacto lingüístico es normalmente mayor […] que en la fonética". E, de facto, é precisamente no léxico que estamos a constatar uma maior uniformidade no grupo oriental açoriano.

Por outro lado, a atuação da segunda força variacional que referimos, a unidade insular, pode sempre gerar divergências como a considerada: a nível fonético, São Miguel e Santa Maria, com a autonomia conferida pela sua insularidade, terão um distanciamento

9 Vejam-se por exemplo os dois trabalhos basilares de Cintra sobre os temas: Cintra (1983a) e Cintra (1983b). No caso do galego, a mesma divergência foi verificada já com base em dados dialetométricos: como se pode ver em Álvarez *et al.* (2006: 482-485), o comportamento areal do léxico no galego é marcadamente diferente do da fonética e da morfologia.

que não realizaram a nível lexical. São Miguel, note-se, é toda ela um grupo uno na variação lexical (cf. mapas 2 e 3 e ainda o 4, comentado seguidamente).

Um exemplo muito claro da existência primacialmente insular dos dialetos açorianos é dado no mapa 4, que apresenta uma subdivisão dos dezassete pontos de inquérito utilizados em tantos grupos quanto o número de ilhas do arquipélago: nove. Podemos verificar que os grupos coincidem na perfeição com as ilhas, ou seja, não existe *overlapping* entre ilhas.

Outro parâmetro dialetométrico, a *distribuição skewness* ou *de assimetria*, mostra-nos mais um exemplo de como os dialetos açorianos podem evoluir com relativa autonomia e de como, adicionalmente, a variação lexical não parece corresponder à fonética. Esse parâmetro parte do facto de que nem todos os locais têm o mesmo grau de semelhança com os restantes locais do conjunto analisado e apresenta, a partir do coeficiente de assimetria de Fisher (CAF), uma síntese relacional dos locais mais e menos semelhantes com o conjunto. Por isto mesmo, é um dos parâmetros dialetométricos mais úteis, como refere Goebl (2012:133): "Il s'est avéré que le meilleur indice pour saisir la symétrie d'une distribution de similarité et, avec elle, *sa vertu communicative*, est le "coeficiente d'asymétrie de Fisher"" (itálico nosso).

Tomemos como exemplo os mapas 5-1 e 5-2, que são, respetivamente, as cartas de similaridade de Fontinhas (Terceira) e Santo Espírito (Santa Maria), isto é, as cartas que apresentam a hierarquia de semelhança/diferença desses locais com cada um dos restantes locais. Pode verificar-se, nos histogramas das cartas, que as Fontinhas são um local linguisticamente mais afastado do conjunto do que Santo Espírito, visto que o número de pontos geográficos que apresentam maiores diferenças (representados a azul) é, na primeira localidade, superior àqueles que apresentam mais semelhanças (representados a vermelho), ao contrário de Santo Espírito, que tem uma maior similitude com o conjunto. Na distribuição de assimetria, discrimina-se, em mapa sinótico, a hierarquia correspondente entre pontos indicada pelas cartas de similaridade.

A distribuição de assimetria dos dialetos açorianos está representada no mapa 6 e deve ser entendido, uma vez que se utiliza o CAF, que os locais com valores mais baixos (cor vermelha) são os mais similares com o conjunto, e os com valores mais altos (cor azul) os menos similares. Nessa carta pode-se ver que (i) os grupos mais distintos não formam uma coerência geográfica, surgindo antes dispersos e (ii) numa distribuição geográfica mais uma vez diferente da assumida na visão tradicional (a Terceira é a ilha mais distinta de todas e, por via disso, tem uma posição destacada do conjunto, tal como na classificação referida, mas São Miguel, pelo contrário, não se destaca mais do que o Pico ou as Flores). Por outro lado, (iii) os locais que se destacam pela semelhança com o conjunto são precisamente os pontos extremos (i.e. os mais periféricos), Corvo e Santa Maria, numa distribuição particularmente

relevante, pois (iii') estão próximos de ilhas que se destacam do conjunto (Flores e S. Miguel, respetivamente), notando-se aqui mais uma vez a autonomia insular dos dialetos açorianos, e (iii'') a larga maioria dos pontos está ao redor da média (7 amarelos + 7 verdes = 14/17), facto que sugere alguma homogeneidade lexical do todo da região. Contudo, esse é um assunto que deverá ser analisado com dados comparativos de outras regiões do português, algo de que ainda não dispomos.

Notamos ainda o interesse histórico do ponto (iii), que sugere a existência de uma maior uniformidade original no conjunto dialetal açoriano, com evoluções específicas posteriores, pois só assim se explica o facto de os pontos mais periféricos serem uma espécie de *máximo comum*, numa distribuição, como se referiu, centrada na média. Esse aspeto é, portanto, um contributo para a muito discutida questão da origem e evolução dos dialetos açorianos, que, devido à falta de dados pormenorizados sobre o povoamento histórico das ilhas (e.g. Meneses, 2008), permanece ainda muito especulativa.

4. Conclusão

Os principais pontos evidenciados no nosso estudo podem ser sistematizados do seguinte modo:

a) A segmentação dialetal do léxico açoriano, tal como obtida a partir dos dados dialetométricos, é significativamente diferente da obtida a partir dos dados fonéticos tradicionais;

b) Aquela segmentação é, na mesma medida, diferente da visão de conjunto dos dialetos açorianos legada pela tradição dialetológica portuguesa, que foi construída a partir dos dados fonéticos tradicionais. Convém ressaltar, porém, que essa visão é ainda deliberadamente incompleta, pois não dispomos, até ao presente, de uma proposta de classificação dialetal do arquipélago (no sentido da sua subdivisão, e não da interpretação da sua posição no macroconjunto galego-português, algo já feito e praticável a partir de uma abordagem não estritamente dialetológica);

c) A divergência referida na alínea *a* consubstancia-se no facto de que o léxico segue um paralelo geográfico na sua segmentação dialetal: os principais grupos dialetais açorianos, no que concerne à variação lexical, coincidem na perfeição com os principais grupos geográficos;

d) Em termos lexicais, a divisão dialetal açoriana simplificada, e por isso a que deve ser privilegiada para efeitos de classificação dialetal, é a tripartição apresentada no mapa 2 (devido sobretudo às analogias que podem ser estabelecidas com a geografia do arquipélago, com os restantes dados dialetais de conjunto de que dispomos sobre a região e com a divisão administrativa pré-autonomia);

e) O fator autonomia, ou isolamento insular, tem sido fundamental no desenvolvimento das personalidades dialetais da região. Por muito que cada local esteja inserido numa coerência geográfica *macro*, a sua posição *micro* reside na ilha de que faz parte e é sempre a primeira identidade. Mais do que isso, o desenvolvimento mais ou menos autónomo que cada ilha tem confere um elevado grau de imprevisibilidade e desordenamento espacial aos padrões de variação da região. Exemplos desse *desordenamento* são as diferenças de comportamento perante o conjunto que se verificam dentro dos grupos ocidental e oriental: como verificámos na análise ao mapa 6, Corvo (grupo ocidental) e Santa Maria (grupo oriental) destacam-se pela semelhança com o conjunto, totalmente ao contrário dos pares dos grupos respetivos (Flores e S. Miguel);

f) Fica assim evidente que o estudo dialetométrico de áreas descontínuas – no caso, um arquipélago com quase seis séculos de povoamento humano continuado –, que tem pouca tradição na literatura, coloca desafios que o de áreas contínuas (tipicamente países), largamente praticado, não coloca – ou coloca menos. Ali há que lidar, em grau exponencial, com uma espécie de *dimensão de incomunicabilidade* (lembrando os tradicionais conceitos de *função* simultaneamente *comunicativa* e *discriminatória* dos dialetos) que condiciona o desenvolvimento dialetal dos locais; quer dizer, uma dimensão, neutra, de abstração de cada dialeto do conjunto de que faz parte. Cada ilha dos Açores, vimo-lo bem, funciona como um bloco e os exemplos de divergência da ilha vizinha não são pouco significativos.

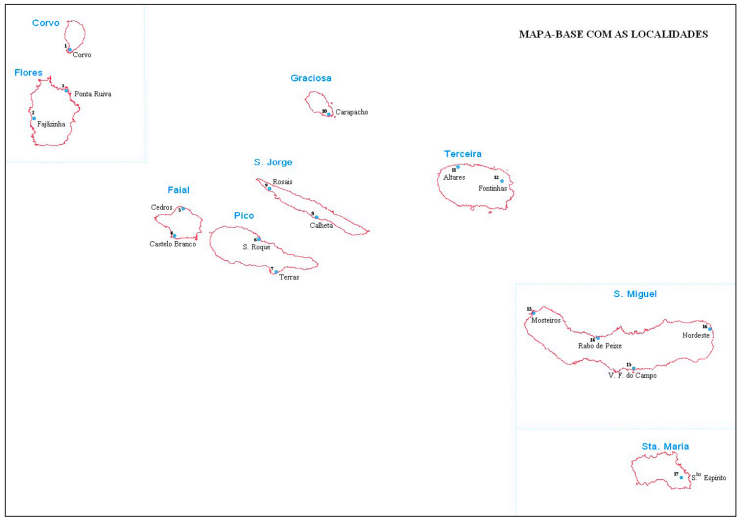
Coloca-se portanto a necessidade de estudar dialetometricamente outras áreas gramaticais açorianas (fonética, morfologia, etc.), após o que chegaremos a uma proposta robusta de classificação dos dialetos açorianos, pois assentará em dados quantitativos (por via da análise dialetométrica) e compreenderá vários domínios linguísticos (e não, ao contrário do procedimento tradicional, apenas um, seja ele a fonética, o léxico, etc.). Os procedimentos dialetométricos, por via precisamente da sua componente simultaneamente quantitativa e espacial (dado que o foco de cada análise dialetométrica é a possibilidade de representação cartográfica eficiente), proporcionam a capacidade de síntese necessária para efetuar abordagens tão abrangentes.

Em todo o caso, dispomos já de uma primeira classificação dialetal dos Açores a partir de dados lexicais.

Agradecimentos

Os autores desejam expressar o seu forte agradecimento a Gotzon Aurrekoetxea e a Hans Goebl pelas ajudas dadas no decurso deste trabalho (o primeiro sobretudo com o uso do *software* de análise dialetométrica, o segundo com os aspetos teóricos da dialetometria).

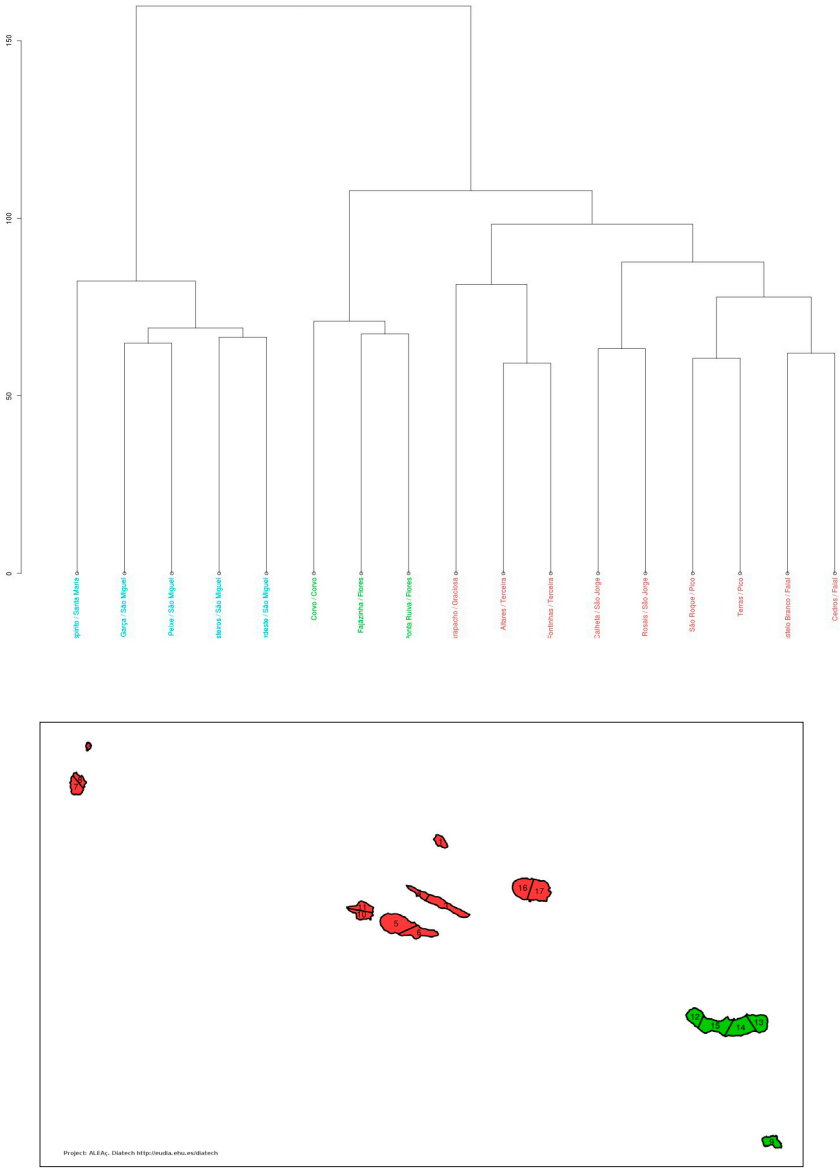
Anexo



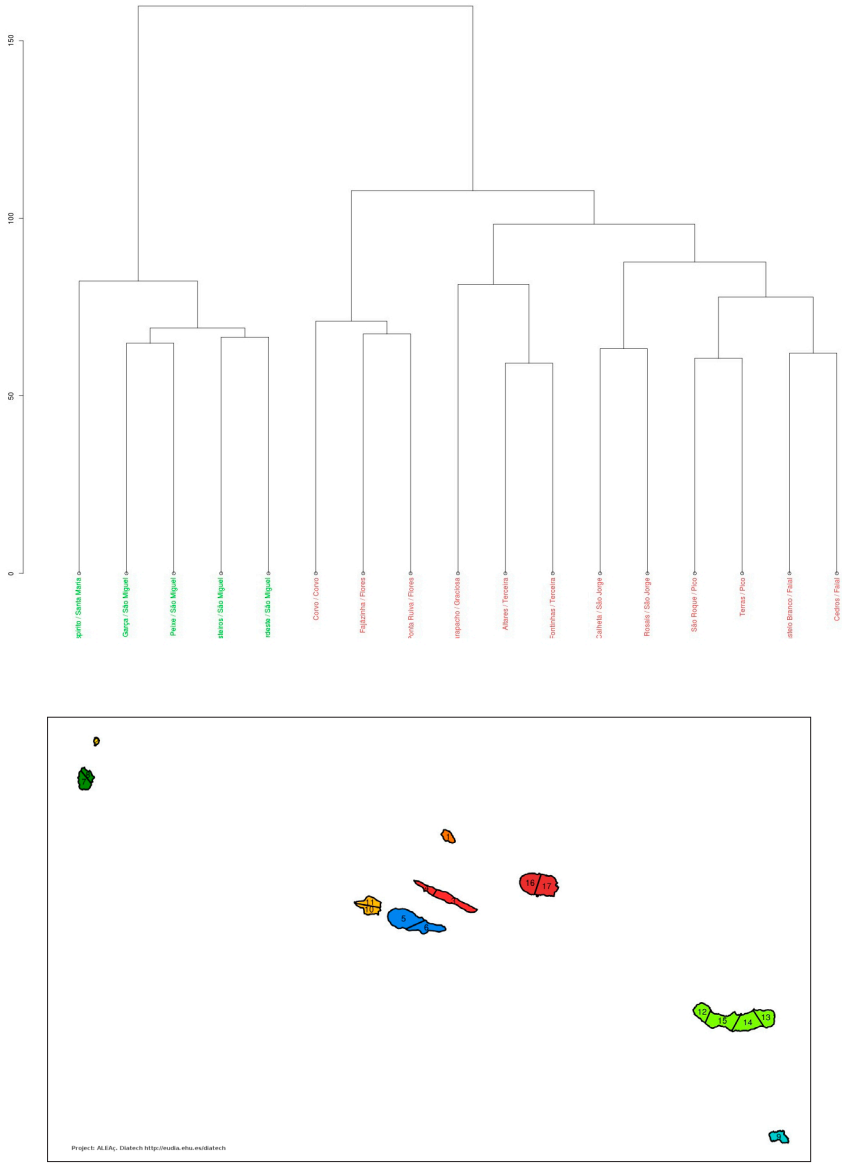
Mapa 1. Pontos de inquérito do ALEAç

Mapa 2. Análise dendrográfica do ALEAç
Corpus: léxico, totalidade (257 conceitos); índice de similaridade: IRI[10]; algoritmo: Ward; número de agrupamentos: 3.

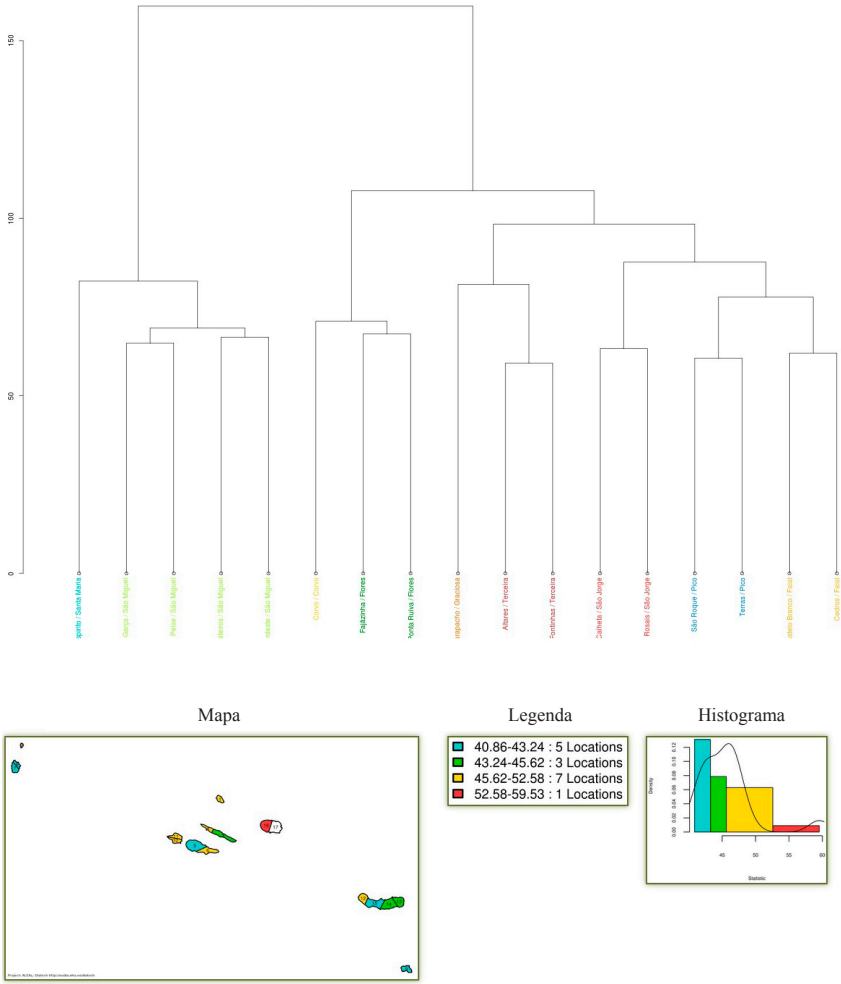
10 IRI = Índice Relativo de Identidade, o principal índice dialetométrico da EDS (e.g. Goebl, 2010: 10).



Mapa 3. Análise dendrográfica do ALEAç
Corpus: léxico, totalidade (257 conceitos); índice de similaridade: IRI; algoritmo: Ward; número de agrupamentos: 2.

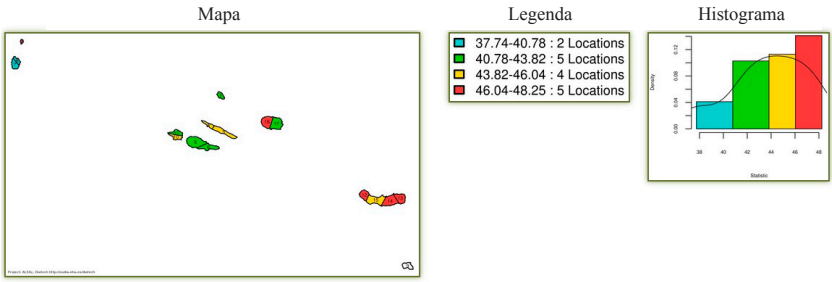


Mapa 4. Análise dendrográfica do ALEAç
Corpus: léxico, totalidade (257 conceitos); índice de similaridade: IRI; algoritmo: Ward; número de agrupamentos: 9.

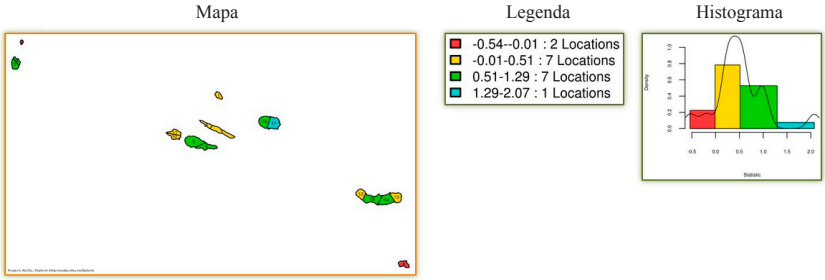


Mapa 5-1. Carta de similaridade do ponto 17 do ALEAç, Fontinhas (Terceira)
Corpus: léxico, totalidade (257 conceitos); índice de similaridade: IRI; algoritmo de visualização: MINMWMAX[11] 4-tuplo.

11 MINMWMAX é um dos algoritmos de visualização comummente utilizados pela EDS. O seu procedimento consta da divisão do conjunto de dados a partir da média artimética (MW) em escalões simétricos para baixo (MIN = mínimo) e para cima (MAX = máximo), tendo em conta o número de intervalos de visualização pretendidos. Se pretendermos quatro intervalos de visualização, aplicamos as seguintes operações: (i) determina-se a média aritmética da totalidade dos valores; (ii) divide-se por 2 a diferença entre a média aritmética e o valor mais baixo; (iii) divide-se por 2 a diferença entre



Mapa 5-2. Carta de similaridade do ponto 9 do ALEAç, Santo Espírito (Santa Maria)
Corpus: léxico, totalidade (257 conceitos); índice de similaridade: IRI; algoritmo de visualização: MINMWMAX 4-tuplo.



Mapa 6. *Distribuição de assimetria* (coeficiente de assimetria de Fisher) do ALEAç
Corpus: léxico, totalidade (257 conceitos); índice de similaridade: IRI; algoritmo de visualização: MINMWMAX 4-tuplo.

---

o valor mais elevado e a média. Disto resulta que o número de pontos de inquérito pertencente a cada intervalo é independente do número de intervalos (embora, naturalmente, os intervalos extremos da distribuição tenham sempre pelo menos um ponto de inquérito): ocorrendo *n* pontos de inquérito no primeiro intervalo, no segundo podem ocorrer *n* + 1 ou *n* - 1, etc. A escolha do número de intervalos é arbitrária, mas deve ajustar-se o melhor possível ao conjunto de dados e ao número de pontos de inquérito de que se dispõe. No caso do presente trabalho, utilizamos sempre quatro intervalos, na medida em que dois seriam, naturalmente, pouco e, a partir dos seis (inclusive), ocorreriam intervalos sem pontos de inquérito.

Referências bibliográficas

Álvarez Blanco, Rosario, Dubert García, Francisco, & Sousa Fernández, Xulio (2006). "Aplicación da análise dialectométrica aos datos do *Atlas Lingüístico Galego*". Em Álvarez Blanco, Rosario, Dubert García, Francisco, & Sousa Fernández, Xulio (eds.), *Lingua e territorio*, 461-493. Santiago de Compostela: Consello da Cultura Galega / Instituto da Lingua Galega.

Atlas Linguístico-Etnográfico de Portugal e da Galiza, página *web*: http://www.clul.ulisboa.pt/en/10-research/680-alepg-linguistic-and-ethnographic-atlas-of-portugal-and-galicia?showall=1&limitstart= .

Atlas Linguístico-Etnográfico de Portugal e da Galiza. Questionário linguístico, 3 vols. (1974). Lisboa: Instituto de Linguística.

Atlas Linguístico-Etnográfico dos Açores, página *web*: htttp://www.culturacores.azores.gov.pt/alea/ .

Aurrekoetxea, Gotzon *et alii* (2013). "'DiaTech': A new tool for dialectology", *Literary and Linguistic Computing, 28(1)*, 23-30. DOI: 10.1093/llc/fqs049 .

Cintra, Luís Filipe Lindley (1983a). "Áreas lexicais no territorio português". Em Luís Filipe Lindley Cintra, *Estudos de Dialectologia Portuguesa*, 55-94. Lisboa: Sá da Costa.

Cintra, Luís Filipe Lindley (1983b). "Nova proposta de classificação dos dialectos galego-portugueses". Em Luís Filipe Lindley Cintra, *Estudos de Dialectologia Portuguesa*, 119-163. Lisboa: Sá da Costa.

Cunha, Celso, & Cintra, Luís Filipe Lindley (1984). *Nova gramática do português contemporâneo*. Lisboa: Edições Sá da Costa.

Diatech, página *web*: http://eudia.ehu.es/diatech/index/ .

Goebl, Hans (2006). "Recent advances in Salzburg dialectometry", *Literary and Linguistic Computing, 21(4)*, 411-435. DOI: 10.1093/llc/fql042 .

Goebl, Hans (2010). "Introducción a los problemas y métodos según los principios de la Escuela Dialectométrica de Salzburgo (con ejemplos sacados del "Atlante Italo-Svizzero", AIS)". Em Aurrekoetxea, Gotzon, & Ormaetxea, José (eds.), *Tools for linguistic variation*, 3-39. Universidad del País Vasco.

Goebl, Hans (2012). "Introduction aux problèmes et méthodes de l'"École dialectométrique de Salzboug" (avec des exemples gallo-, italo- et ibéroromans)". Em Álvarez Perez, Xosé, Carrilho, Ernestina, & Magro, Catarina (eds.), *Proceedings of the International Symposium on Limits and Areas in Dialectology (LimiAr), 2011*, 117-166. Lisboa: Centro de Linguística da Universidade de Lisboa.

Goebl, Hans, & Smečka, Pavel (2014). "L'analyse dialectométrique des cartes de la série B de l'ALF", *Revue de Linguistique Romane, 78* (311-312), 439-497.

Meneses, Avelino de Freitas de (2008). "O povoamento". Em Matos, Artur Teodoro de, Meneses, Avelino de Freitas de, & Leite, José Guilherme Reis (orgs.), *História dos Açores*, vol. I, 63-109. Angra do Heroísmo: Instituto Açoriano da Cultura.

Saramago, João (2006). "O Atlas Linguístico-Etnográfico de Portugal e da Galiza (ALEPG)", *Estudis Romànics, XXVIII*, 281-298.

Saramago, João & Gonçalves, José Bettencourt (2003). "Diferenciação lexical interpontual nos Açores (estudo dialectométrico aplicado em materiais do ALEAç)". Em Rita Caprini (ed.), *Parole romanze. Scritti per Michel Contini*, 421-440. Alessandria: Edizioni dell'Orso.

Saramago, João, & Segura (da Cruz), (Maria) Luísa (1999). "Açores e Madeira: autonomia e coesão dialectais". Em Hub Faria, Isabel (org.), *Lindley Cintra: homenagem ao homem, ao mestre e ao* cidadão, 707-738. Lisboa: Cosmos / Faculdade de Letras da Universidade de Lisboa.

Segura, Luísa (2013). "Variedades dialetais do português europeu". Em Paiva Raposo, Eduardo *et alii* (orgs.), *Gramática do português*, vol. I, 85-142. Lisboa: Fundação Calouste Gulbenkian.

Wieling, Martijn, & Nerbonne, John (2015). "Advances in dialectometry", *Annual Review of Linguistics, 1*, 243-264. DOI: 10.1146/annurev-linguist-030514-124930.

Outros *Caminhos de Santiago*... os da Bahia (Brasil)

Suzana Alice Marcelino Cardoso
Universidade Federal da Bahia / Conselho Nacional de Desenvolvimento Científico e Tecnológico

Ana Regina Torres Ferreira Teles
Universidade Federal da Bahia

Resumo:

Neste artigo consideram-se as denominações, na Bahia, para *via láctea*, um dos itens semântico-lexicais constantes dos questionários linguísticos aplicados para o *Atlas Prévio dos Falares Baianos* (APFB) e para o Projeto *Atlas Linguístico do Brasil* (ALiB). Examina-se o rol de variantes localizadas no estado da Bahia em duas sincronias distintas, separadas por cerca de 50 anos. No que se refere ao APFB, a documentação traz dados de 50 localidades, não incluída a capital do estado, e de um total de 100 informantes. Quanto ao Projeto ALiB, os materiais foram coletados em 22 localidades, que constituem a rede no estado, e a um total de 88 informantes, equitativamente distribuídos em duas faixas etárias e pelos dois sexos. Na análise das variantes documentadas, são consideradas a distribuição diatópica, a natureza da motivação que refletem os itens lexicais registrados e as diferenças que podem ser sentidas entre os usos e preferências nas duas sincronias consideradas. Complementarmente, acrescenta-se uma breve notícia do que ocorre nos demais estados dos quais já se publicaram atlas linguísticos e nos quais figuram informações sobre a via láctea. Observa-se, por fim, que a denominação *Caminho de Santiago*, com suas variantes fonéticas, apresenta o maior índice de ocorrência, verificando-se também uma presença significativa da motivação de base cristã no conjunto de denominações registradas.

Palavras-chave:

Léxico, Atlas Linguístico do Brasil, Atlas Prévio dos Falares Baianos.

Sumário:

1. Introdução. 2. A *via láctea* na Bahia: fontes. 3. Os dados considerados. 4. Via láctea: natureza da motivação. 5. Um breve passeio diacrônico. 6. À guisa de conclusão.

Other Caminhos de Santiago... *those from Bahia (Brazil)*

Abstract:

In this paper, the denominations in Bahia for Milky Way *are considered as one of the Semantic-Lexical items included in the linguistic questionnaires applied for the* Atlas Prévio

dos Falares Baianos (APFB) *and for the* Atlas Linguístico do Brasil *(ALiB). The list of variants located in the state of Bahia is examined in two distinct synchronicities, separated by around 50 years. In relation to the APFB, the documentation provides information from 50 localities, not including the state capital, and of a total of 100 informants. In relation to the ALiB Project, the materials were collected in 22 localities, which form the network in the state, and a total of 88 informants, equally distributed in two age groups and in both genders. The analysis of the documented variants considered the following: diatopic distribution, the nature of the motivation behind the registered lexical items and the differences that can be perceived amongst the uses and preferences of both synchronicities examined. Moreover, brief information on what occurs in other states that have already published linguistic atlases where information on the* Milky Way *is also included. Finally, it should be observed that the denomination Caminho de Santiago, with its phonetic variants, presents a greater rate of occurrence, verifying also a significant presence of Christian based motivation behind the set of denominations recorded.*

Key words:

Lexicon, Brazilian Linguistic Atlas, Atlas Prévio dos Falares Baianos.

Contents:

1. Introduction. 2. The Milky Way *in Bahia: sources. 3. The data applied. 4. Milky Way: nature of motivation. 5. A brief diachronic tour. 6. By way of conclusion.*

1. Introdução

O firmamento, o espaço sideral, tudo o que nos envolve com um sabor de impenetrável, infinito, exerce sobre o homem um grande fascínio e lhe povoa a mente, sobretudo na busca de denominações que respondam ao seu imaginário. Com esse entendimento, busca-se examinar as denominações para *via láctea* ocorrentes no estado da Bahia (Brasil), partindo dos dados recolhidos na pesquisa de campo efetuada para a elaboração do *Atlas Prévio dos Falares Baianos* (APFB) (Rossi, 1963). Com base nos princípios da Geolinguística Pluridimensional, examina-se o referencial do APFB numa perspectiva léxico-semântica. Ao percorrer os caminhos do APFB, objetiva-se (i) identificar a variedade de nomes registrados e encontrar as raízes da motivação para as denominações dadas à *via láctea*; (ii) verificar o caráter areal de que podem se revestir algumas das variantes documentadas; (iii) definir elementos que permitam qualificar, do ponto de vista diageracional e diassexual, as preferências de uso, e (iv) relacionar os resultados dessa sincronia –aqui denominada sincronia 1– aos daquela a que dizem respeito os dados do Projeto ALiB –sincronia 2–, e, na medida do possível, a outros alcançados pela Geolinguística no Brasil. O estudo parte, fundamentalmente, do que se registra no estado da Bahia (Figura 1) a partir do que documentam as pesquisas para o APFB.

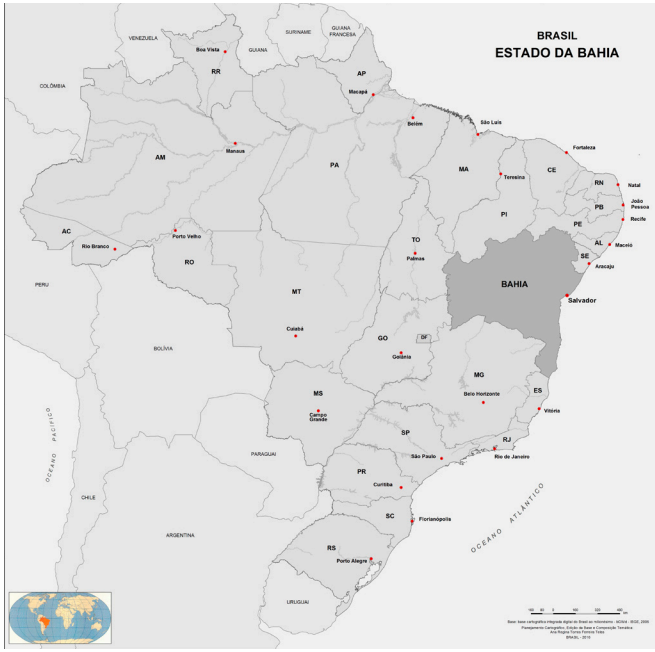


Figura 1. A Bahia no Brasil

2. A *via láctea* na Bahia: fontes

As denominações ocorrentes para *via láctea* na Bahia foram documentadas de forma sistemática em todo o estado, mediante a aplicação de inquéritos linguísticos, *in loco*, em dois momentos distintos. O primeiro –aqui denominado sincronia 1–, nos começos da década de 1960, século XX, com as pesquisas para o *Atlas Prévio dos Falares Baianos*; o segundo –sincronia 2–, no início do século XXI, com a constituição do *corpus* do Projeto *Atlas Linguístico do Brasil* (Projeto ALiB). Uma breve palavra sobre esses dois *corpora*, de modo a caracterizar cada uma dessas recolhas e fornecer o perfil dos informantes, se faz necessária.

O *Atlas Prévio dos Falares Baianos*, de autoria de Nelson Rossi, com a colaboração especial de Dinah Isensee e Carlota Ferreira, foi feito e publicado entre 1960 e 1963 e recobre todo o Estado da Bahia. Trata-se do primeiro atlas linguístico brasileiro que inaugura a produção geolinguística no Brasil, abordando dados regionais, focalizando o estado da Bahia. O autor, Nelson Rossi, teve, entre outros, o mérito de implantar a pesquisa dialetal direcionada à produção de atlas no país, estimulando assim trabalhos de idêntica natureza que passaram a se desenvolver em vários outros estados.

Conta com uma rede de 50 localidades que se distribuem pelas diferentes áreas geográficas e culturais, recobrindo todas as mesorregiões, e traz a documentação de 100 informantes, de ambos os sexos e, quanto ao grau escolaridade, analfabetos ou semialfabetizados. Constitui-se de um conjunto de 209 cartas, das quais 198 são cartas linguísticas, 44 são resumos das cartas fonéticas, e 11 são cartas introdutórias que fornecem dados complementares de caráter geral. As cartas linguísticas vêm acompanhadas de notas que ampliam os dados linguísticos não apenas no nível do léxico ou da fonética, mas também da morfossintaxe, e transmitem melhor o ambiente cultural em que vive o informante, trazendo esclarecimentos e descrições apresentados pelos próprios informantes e/ou comentários dos pesquisadores (inquiridores ou transcritores).

Para este artigo, recorre-se a informações que, embora não tendo gerado carta linguística, encontram-se disponíveis nas respostas recolhidas pelo questionário aplicado e passíveis de serem retomadas posteriormente, possibilidade que antevia o autor ao anunciar nos comentários registrados no volume de *Introdução*: "Os materiais recolhidos foram cuidadosamente examinados e estão convenientemente catalogados para aproveitamento posterior" (Rossi, 1965: 55). Assim, são examinadas as respostas à pergunta 2 (Constelações) do Extrato de Questionário do APFB, resultantes da formulação indireta e transcritas diretamente, no momento do inquérito, pela inexistência, àquela época, de aparelhos de gravação portáteis e com autonomia de corrente elétrica.

O Projeto *Atlas Linguístico do Brasil*, projeto nacional implantado em 1996, objetiva descrever a realidade linguística do país, documentando e analisando o português

brasileiro, considerando os vários níveis de abordagem da língua e a diversidade de usos refletida pelos informantes estratificados conforme a faixa etária, a escolaridade e o sexo. Utilizando uma metodologia pluridimensional, investiga os vinte e seis estados brasileiros, fornecendo cartas linguísticas em que figuram dados diatópicos e, conforme o item considerado, informações de caráter diageracional, diastrático e diassexual.

Para a constituição do *corpus* foram realizadas gravações com 1100 informantes distribuídos por 250 localidades espalhadas pelos 8.500.000 km^2 do país, aí incluídas todas as capitais (com exceção de Palmas-TO e Brasília-DF por serem cidades novas, a primeira criada em 1989 e a segunda fundada em 1960, e não preencherem os requisitos metodológicos estabelecidos para a pesquisa), perfazendo um total de aproximadamente 3300 horas de gravação. A rede de pontos atende à realidade do país e conta com a inclusão de localidades que contemplam a diversidade geográfica, os aspectos históricos e culturais e as características sociais e linguísticas de cada área. Os informantes são estratificados pelos sexo –homens e mulheres–, em duas faixas etárias (18 a 30 anos e 50 a 65 anos) e por dois níveis de escolaridade (fundamental e universitário). Quanto às entrevistas, foram aplicados um questionário fonético-fonológico (incluindo questões de prosódia), um questionário semântico-lexical (versando sobre várias áreas temáticas) e um questionário morfossintático, além de questões de pragmática, de discurso semidirigido (com relatos pessoais), de perguntas metalinguísticas e de um texto para leitura. Os primeiros resultados encontram-se nos dois volumes, recentemente publicados (Cardoso *et al.*, 2014a, 2014b). O volume 1 (*Introdução*) traz a história do projeto, a metodologia utilizada e um informe sobre a cartografia dos dados. O volume 2 (*Cartas linguísticas 1*) apresenta um conjunto de cartas linguísticas, com resultados das capitais brasileiras concernentes à fonética, incluindo dados prosódicos, ao léxico e à morfossintaxe, precedidas por cartas introdutórias.

Para a análise que se está propondo neste artigo, foram consideradas as respostas à pergunta 33 do Questionário semântico-lexical nos 22 pontos que compõem a rede no estado, assim formulada: "Numa noite bem estrelada, aparece uma banda ou faixa que fica no céu de fora a fora, onde tem muitas estrelas muito perto uma das outras. Como chamam essa banda ou faixa?" (Comitê Nacional do Projeto ALiB, 2001: 24).

3. Os dados considerados

O APFB não apresenta uma carta para *via láctea*, mas os dados coletados no campo encontram-se devidamente arquivados, reunindo todas as respostas documentadas. Nesse conjunto há respostas a perguntas que, naquele então, por razões diversas não chegaram a ser cartografadas. Dirimidas as dificuldades de interpretação com

a realização de pesquisas posteriores, sobretudo as relativas ao *Atlas Linguístico de Sergipe* (Ferreira *et al*., 1987), examinam-se as denominações ocorrentes, naquela sincronia, para a *via láctea*.

Na Bahia, no que estamos denominando de sincronia 1, a representada pela documentação do APFB, registram-se as seguintes variantes lexicais para *via láctea*:

- cama de boi
- caminho de Deus
- caminho de Santiago
- caminho de santiágua(s)
- Santiago
- caminho de São Felipe
- caminho de São José
- caminho de São Pedro
- guardanapinho
- guardanapo de Nossa Senhora
- guardanapo de Santa Polônia (Apolônia?)

Registre-se que, dos 50 pontos da rede, em três localidades –pontos 11, 19 e 36– não foram anotadas respostas à pergunta, como também não se fornecem os dados referentes aos pontos 4, 8, 10 e 50 em virtude de não terem sido localizadas as fichas de campo relativas a esses pontos. O Quadro 1 apresenta o panorama das ocorrências localizadas no estado da Bahia.

Variantes registradas	Total de ocorrências	
	Número de localidades	Número de informantes
cama de boi	2	2
caminho de Deus	1	1
caminho de Santiago / São Tiago	36	49
caminho de santiágua(s)	8	9
caminho de São Felipe	3	4
caminho de São José	1	1
caminho de São Pedro	1	1
guardanapinho	1	1
guardanapo de Nossa Senhora	1	1
guardanapo de Santa Polônia	1	3
Santiago	1	1

Quadro 1. Total de ocorrências por localidade/informante

Observando-se o que exibem o Quadro 1 e as Figuras 2 e 3, (i) verifica-se que a variante *caminho de Santiago*, única do conjunto que se encontra registrada com a

acepção de *via láctea* nos dicionários Houaiss e Villar (2001) e Ferreira (2004), surge como a de maior índice de ocorrência, presente em 36 das 50 localidades da rede, a que se acrescentam as variantes fonéticas *caminho de santiágua* (em 6 localidades), *caminho de santiaua* (em uma localidade) e *caminho de santiáguas* (em uma localidade); (ii) nota-se a não ocorrência da forma padrão *via láctea* e (iii) destaca-se a manifestação presente, na quase totalidade das denominações, do sentimento religioso-cristão, com evocação de nomes de santos da Igreja católica, além de Santiago, forma majoritária na área, tais como São Felipe, São José, São Pedro, Nossa Senhora, Santa Polônia (Apolônia?). As Figuras 2 e 3 mostram a distribuição da lexia *caminho de Santiago* e de suas variantes fonéticas que se configuram como casos de perda da motivação principal, a própria relação com o santo da Igreja Católica, São Tiago/Santiago, e o surgimento de uma nova relação motivadora, no caso, com a água. Há de se perguntar: como explicar a perda de motivação? Sem querer, nem poder, responder de forma precisa, algumas considerações se põem.

No que concerne à variação diassexual, verifica-se que os registros se distribuem equitativamente, ou seja, 50% na fala tanto dos homens como das mulheres. Quanto à variação diageracional, tomando-se como base da divisão etária com base na idade igual ou superior a 50 anos, os resultados podem, ainda que levemente, sugerir certa relação idade-escolha do item lexical. Os dados mostram que as variantes fonéticas de *caminho de Santiago*, com perda da motivação, têm, nesse pequeno conjunto, uma maior presença nos informantes de idade mais avançada, pois, dos oito casos registrados, cinco integram a fala de informantes de idade igual ou superior a 50 anos e apenas três estão na fala dos mais jovens –informantes com 31, 36 e 41 anos–. Os dados, portanto, de cunho diassexual e diageracional não explicam, pelo menos nesse caso específico, a natureza da escolha, apenas deixam a indicação de um leve caminho a ser aprofundado –o da variação diageracional– na busca de uma explicação para o fato.

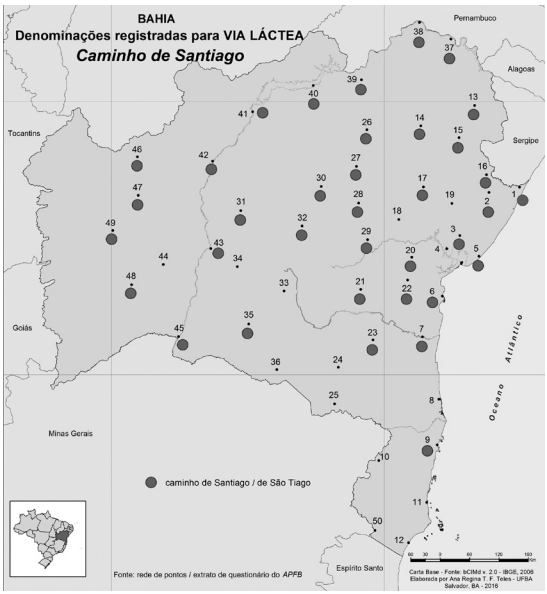


Figura 2. *Caminho de Santiago*

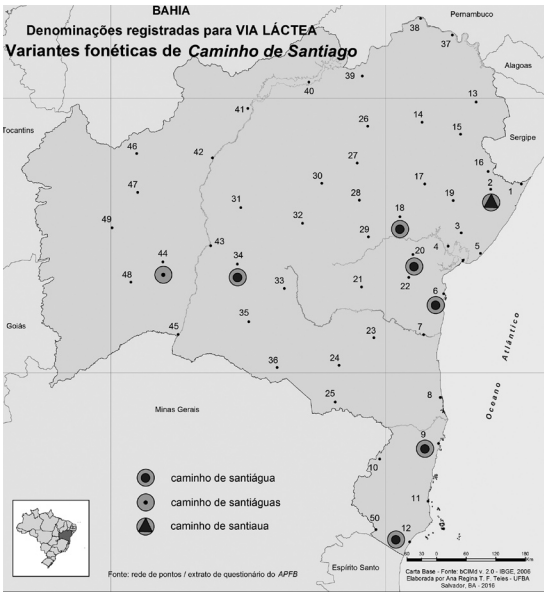


Figura 3. *Caminho de Santiago*: variantes fonéticas

O Quadro 2 fornece os totais de ocorrências de outras denominações, sem referência a *Santiago* ou variantes, verificando-se que se trata de menor número de formas. A Figura 4 reúne os dados gerais, permitindo verificar-se a distribuição das variantes e a presença significativa daquelas que envolvem o nome de Santiago.

Variantes	Total de ocorrencias
cama de boi	2
caminho de Deus	1
caminho de São Felipe	3
caminho de São José	1
caminho de São Pedro	1
guardanapinho	1
guardanapo de Nossa Senhora	1
guardanapo de Santa Polônia	1

Quadro 2. Outras denominações

Do ponto de vista diatópico, como mostra a Figura 4, *caminho de Santiago* e variantes fonéticas recobrem, praticamente, todo o estado, ficando as demais ocorrências restritas a uma área geográfica especifica, exatamente uma pequena faixa na parte leste, com dois pontos mais ao sul. *Caminho de Santiago* configura-se como uma denominação de uso comum pela população, cujo índice de ocorrência atesta a vitalidade da forma e a sua atualidade na área.

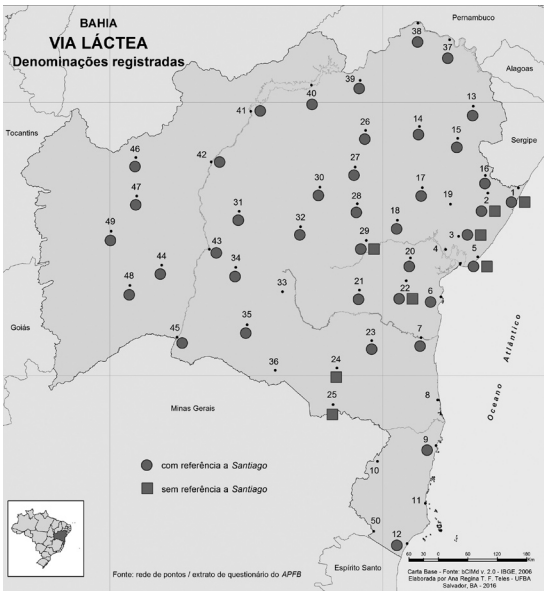


Figura 4. Via láctea: denominações registradas

Não se tem elementos para avaliar o grau de consciência dos informantes no que se refere a estabelecer relação entre a denominação que adotam –*caminho de Santiago*– e a existência dos vários caminhos de Santiago que conduzem a Santiago de Compostela ou do papel da Via Láctea na indicação do local onde a tradição diz encontrar-se o corpo de São Tiago. Muito provavelmente essa forma foi trazida pelos portugueses, no período da colonização, em mais uma das transposições de seus traços culturais às terras conquistadas, como sucedeu com *sarolha* 'terra úmida', ausente dos dicionários atuais, mas presente de forma representativa em todo o estado da Bahia. Se nos começos é possível que a relação fosse feita e assim se explicasse a seleção do nome a dar a essa galáxia, com o correr do tempo se perdeu a motivação, que, por certo, fez surgir a forma em questão, o que se evidencia nas variantes *caminho de santiaua*, *caminho de santiágua*, *caminho de santiáguas*, registradas na Bahia.

Por outro lado, as razões da denominação *caminho de Santiago* assumida na Bahia para Via Láctea não comparecem nos comentários dos informantes. Citam-na, apenas, como nome que registram para essa constelação e, na maioria dos casos, dado como primeira resposta. O *Atlas Linguístico do Paraná* (Aguilera, 1994), cujas notas à carta Via Láctea são abundantes, fornece, apenas, comentários interpretativos dos informantes, como os que, a seguir, se transcrevem:

> Ele [o caminho de Santiago] dá um vurto branco assim né, que passa. O pessoar fala que quano o craro passa em cima da casa duma moça, vai casá logo (*Atlas Linguístico do Paraná*, ponto 6, informante A).

> Quando ele [o caminho de Santiago] vinha fácil, certinho por riba da casa, diz que ia casá um filho (*Atlas Linguístico do Paraná*, ponto 33, informante B).

Câmara Cascudo (1954), uma das fontes relativas a estudos regionais e folclóricos, registra, apenas, *carreiro de Santiago* para Via Láctea, não fazendo menção a *caminho de Santiago* e sem fornecer dados etimológicos ou explicações para a denominação.

Similarmente, não se dispõe de elementos que justifiquem, de forma transparente a escolha de nomes de outros santos da Igreja católica (Quadro 2) para identificação da Via Láctea nos três pontos da rede do APFB onde se registram. C*aminho de São Felipe* está documentado nos pontos 3 (Rio Fundo) e 5 (Abrantes). O primeiro é uma vila do município de Santo Amaro, cujo nome anterior foi Rio Fundo de São Paulo. O segundo, Abrantes, teve como nome anterior Espírito Santo de Nova Abrantes, vila do Município de Camaçari. Neste último também se registra *caminho/guardanapo de santa Pelônia (Apolônia)*. No ponto 25 (Encruzilhada), do qual não há documentação de nome anterior, registraram-se *caminho de são José* e *caminho de são Pedro*. Os nomes dos santos evocados nas denominações desses três pontos, até onde se pode apurar, não refletem nenhuma ligação direta com situação especifica dessas localidades, como, por exemplo, denominação de santo padroeiro, nome da freguesia, festa típica da localidade.

4. *Via láctea:* natureza da motivação

O conjunto de variantes para denominar *via láctea* constitui-se de nomes compostos por dois elementos que refletem diferenciada motivação. Para a análise desse aspecto, considera-se a natureza da motivação do primeiro e do segundo elemento.

No que diz respeito ao primeiro elemento, registram-se: *caminho* "porção mais ou menos estreita de terreno entre dois lugares por onde alguém pode seguir" (Houaiss, & Villar, 2001: 586); *guardanapo* "pequena toalha de pano ou de papel" (Houaiss,

& Villar, 2001: 1493); *cama* "lugar onde uma pessoa ou animal pode deitar-se e/ou dormir com relativo conforto" (Houaiss, & Villar, 2001: 579).

A motivação se apresenta de forma transparente, uma vez que a conformação da via láctea, que se encontra descrita como "faixa luminosa larga, tênue e irregular, visível nas noites claras, que se estende pela linha do equador galáctico da esfera terrestre, constituída por um imenso número de estrelas e de outros corpos celestes" (Houaiss, & Villar, 2001: 2855), reflete continuidade, espraiamento e certo repouso.

O segundo elemento se constitui dos segmentos: *de boi*, *de Deus, de Santiago / São Tiago, de São Felipe, de São José, de São Pedro, de Nossa Senhora* e *de Santa Polônia*. À exceção de *de boi*, os demais elementos são marcadamente de caráter religioso e refletem a profunda relação estabelecida pelo falante entre esse fenômeno da natureza –a *via láctea*– e as questões relacionadas a coisas divinas. Observa-se, assim, um traço predominante do antropomorfismo cristão e uma tênue presença do elemento zoomórfico.

5. Um breve passeio diacrônico

O desenvolvimento do Projeto *Atlas Linguístico do Brasil*, com recolha de dados no período de 2001 a 2013, enseja o confronto entre sincronias distintas a partir da relação com outros *corpora* anteriormente constituídos. Dessa forma, no que concerne às denominações para *via láctea* temos elementos que permitem uma análise na perspectiva diacrônica que pode apontar as mudanças ocorridas e/ou constatar a manutenção de denominações existentes na primeira das sincronias consideradas.

Aos dados do primeiro momento, trazidos do que se registra com a aplicação do questionário APFB, somam-se os que a seguir se apresentam, relativos ao *corpus* do Projeto ALiB.

A rede de pontos do Projeto ALiB, no estado da Bahia, constitui-se de 22 localidades, nove das quais coincidentes com aquelas que integram a rede do APFB, pontos esses assinalados na Figura 5. Registram-se as variantes *caminho de Santiago, caminho do céu, carreiro de Santiago, cordão de São Francisco e via láctea*, denominações que assinalam: (i) a presença, também majoritária como na primeira sincronia considerada, da variante *caminho de Santiago*; (ii) a ocorrência, única, da forma tida como padrão, *via láctea*, em Salvador, capital do estado; (iii) o registro em apenas uma localidade de cada uma das três outras variantes.

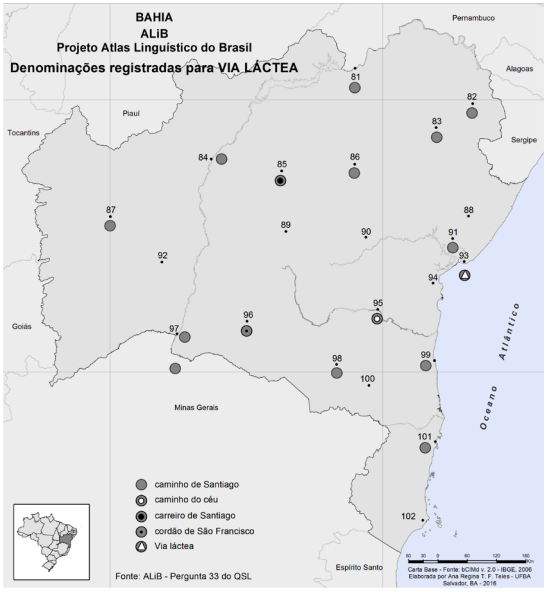


Figura 5. Denominações para *via láctea* segundo dados do ALiB

Como se vem de indicar, a forma tida como do padrão da língua, *Via Láctea*, teve um só registro documentado na denominada, neste trabalho, como sincronia 2, no *corpus* do Projeto ALiB, e exatamente na fala de informante masculino, de faixa etária II (ou seja entre 50-65 anos) e de nível de escolaridade universitário. *Via Láctea*, do latim VIA LACTEA 'caminho de leite', 'caminho leitoso', é a denominação geral na língua portuguesa para a galáxia descrita como "faixa luminosa larga, tênue e irregular, visível nas noites claras" (Houaiss, & Villar, 2001: 2855, s.v. *via*). Em numerosas línguas encontra-se a mesma base de formação: juntam-se "caminho", que representa a configuração que assume a galáxia, e uma adjetivação à base de "leite", que descreve o aspecto apresentado no firmamento. Assim, e para exemplificar, cita-se a denominação em algumas línguas, particularmente nas românicas, que seguem esse mesmo processo de formação: esp. *vía láctea*, cat. *via làctia*, fr. *voie lactée*, gal. *vía láctea*, it. *via lattea*, rom. *calea laptelui*, ing. *Milky way*, al. *Milchstrasse*. A história justifica a lexia a partir da mitologia grega, que explica a denominação como decorrente do fato de o leite dos seios da deusa Hera ter-se derramado nos céus, deixando uma mancha esbranquiçada, quando sugado por Hércules, filho de Zeus com uma mortal (cf. Andrade, & Henrique, s.d).

Caminho de Santiago, nos dados do ALiB, está registrado a 13 informantes: um homem da faixa etária I, seis homens da faixa etária II e seis mulheres da faixa etária II,

configurando-se como forma da preferência de falantes do nível de escolaridade Fundamental e daqueles da segunda faixa etária, ou seja, entre 50 e 65 anos.

Ampliando esse passeio e verificando o que registram atlas regionais, publicados em diferentes momentos, que incluem nos seus resultados denominações para *via láctea*, como o *Atlas Linguístico do Paraná* (Aguilera, 1994) e *Atlas Linguístico de Mato Grosso do Sul* (Oliveira, 2007), observamos que a variante *caminho de Santiago* se faz presente nesses dois estados: no Paraná, em 40 dos 65 pontos, e em Mato Grosso do Sul, em 11 das 32 localidades que constituem a rede do atlas. Esses registros assinalam a presença da denominação em outras áreas e, de certo modo, a vitalidade da denominação *caminho de Santiago,* assegurando que uma forma documentada, pelo menos em pesquisa geolinguística, nos anos 1961/1963 vem transpondo décadas e se mantendo no uso dos falantes do português no Brasil.

6. À guisa de conclusão

A referência no título a "outros caminhos de Santiago... os da Bahia" objetivou chamar a atenção para a presença representativa da designação *caminho de Santiago*, para *via láctea*, nas suas diversas variantes fonéticas, no Estado da Bahia, e, ao mesmo tempo indicar e qualificar as demais variantes registradas para, por fim, relacionar duas sincronias com vistas a retratar a realidade em dois momentos, apontando a vitalidade, ou não, das formas documentadas. A esse objetivo se acrescentou uma breve visita a atlas regionais brasileiros que trazem nos seus dados a informação sobre as denominações para *via láctea*, o que veio a confirmar vitalidade de *caminho de Santiago.*

Das reflexões trazidas, podemos destacar alguns aspectos:

1. As duas sincronias consideradas, com dados da Bahia, mostram panoramas distintos quanto à pluralidade de variantes registradas. Assim, na primeira delas, estão documentadas 11 denominações para Via Láctea, enquanto na segunda registram-se apenas cinco. A forma do padrão da língua comparece apenas no segundo momento considerado e com registro único, tomado a um informante de nível universitário e da capital do Estado. De outro lado, ambas as sincronias mostram a vitalidade da denominação *caminho de Santiago,* uma vez que em mais de 50% da rede de pontos, nos dois momentos, se registra a forma.

2. A vitalidade de *caminho de Santiago* para *via láctea* se atesta com dados de dois outros atlas regionais relativos aos estados de Paraná e Mato Grosso do Sul, mostrando que não se trata de uma preferência localizada, mas de uso mais amplo, uso este cuja real dimensão se virá a ter quando se concluir a análise do vasto *corpus* do Projeto ALiB e se derem a conhecer os resultados.

3. Na escolha de nomes para designar a via láctea predomina a preferência pela seleção de denominações com sentido religioso-cristão de que resulta a invocação de diferentes santos da Igreja Católica.

E, assim, vamos seguindo nós, na Bahia, no Brasil, e a nosso modo, o *caminho de Santiago*!

Referências bibliográficas

Aguilera, Vanderci de Andrade (1994). *Atlas Lingüístico do Paraná*, 2 vols. Curitiba: Imprensa Oficial do Estado.

Andrade, Victória Flório Pires de, & Henrique, Alexandre Bagdonas (s.d.). *Uma breve história da galáxia*. Minicurso UNICAMP / USP / SAB. www.sbpcnet.org.br/livro/61ra/minicursos/MC_VictoriaAndrade.pdf(Consultado em 27.10.2016).

Cascudo, Luís da Câmara (1954). *Dicionário do folclore brasileiro*. Rio de Janeiro: Ministério da Educação e Cultura / Instituto Nacional do Livro.

Cardoso, Suzana Alice Marcelino *et alii* (2014a). *Atlas Linguístico do Brasil. 1. Introdução*. Londrina: EDUEL.

Cardoso, Suzana Alice Marcelino *et alii* (2014b). *Atlas Linguístico do Brasil. 2. Cartas linguísticas 1*. Londrina: EDUEL.

Comitê Nacional do Projeto ALiB (2001). *Atlas Lingüístico do Brasil. Questionários 2001*. Londrina: Universidade Estadual de Londrina.

Ferreira, Aurélio Buarque de Holanda (2004). *Novo Dicionário da língua portuguesa.* Rio de Janeiro: Nova Fronteira.

Ferreira, Carlota *et alii* (1987). *Atlas Lingüístico de Sergipe*. Salvador: Universidade Federal da Bahia / Fundação Estadual de Cultura de Sergipe.

Houaiss, Antônio, & Villar, Mauro de Salles (2001). *Dicionário Houaiss da Língua Portuguesa*. Rio de Janeiro: Objetiva.

Oliveira, Dercir Pedro (org.) (2007). *Atlas Lingüístico de Mato Grosso do Sul*. Campo Grande: Universidade Federal de Mato Grosso do Sul.

Rossi, Nelson (1963). *Atlas Prévio dos Falares Baianos*. Rio de Janeiro: Ministério de Educação e Cultura / Instituto Nacional do Livro.

Rossi, Nelson (1965). *Atlas Prévio dos Falares Baianos*. *Introdução. Questionário Comentado. Elenco das respostas transcritas.* Rio de Janeiro: Ministério de Educação e Cultura / Instituto Nacional do Livro.

O campo semântico *corpo humano* no *Atlas Geossociolinguístico Quilombola do Nordeste do Pará* e no *Atlas Linguístico do Pará*

Marcelo Pires Dias
Marilúcia Barros de Oliveira
Universidade Federal do Pará

Resumo:

Este trabalho tem como objetivo descrever a variedade lexical em comunidades quilombolas, a partir de seis cartas que integram o campo semântico *corpo humano* do *Atlas Geossociolinguístico Quilombola do Nordeste do Pará* (AGQUINPA), e compará-la com a apresentada no *Atlas Geossociolinguístico do Pará* (ALIPA). Consideramos, para o presente estudo, as dimensões diatópica, diassexual e diageracional. O AGQUINPA busca mapear a variedade linguística do português falado em comunidades quilombolas da região Nordeste do Pará, por meio de inventário lexical, e construir um banco de dados com base nos dados coletados. As comunidades consideradas neste trabalho estão situadas na região Nordeste do Estado do Pará (Brasil), são elas: a) Comunidade do Cacau (Colares); b) Comunidade América (Bragança); c) Comunidade Rio Acaraqui/Campompema (Abaetetuba); d) Comunidade Taperinha (São Domingos do Capim); e) Comunidade África e f) Comunidade Laranjituba (Moju). Para a pesquisa, utilizamos o Questionário Semântico-lexical do *Projeto Atlas Linguístico do Brasil* (ALiB), assim como os pressupostos teórico-metodológicos da Geossociolinguística. Os resultados apontam a presença de étimos de origem bantu ou africanias e variantes não dicionarizadas, bem como semelhanças entre os dois atlas mencionados.

Palavras-chave:

Quilombolas, Geossociolinguística, léxico.

Sumário:

1. Introdução. 2. Dialetologia no Brasil. 3. O AGQUINPA. 4. Metodologia. 4.1. A rede de pontos. 4.2. Perfil dos informantes. 4.3. O instrumento de coleta de dados. 4.4. Transcrição e cartografia dos dados. 4.5. Comparação entre o AGQUINPA e o ALIPA. 5. O campo semântico *corpo humano*: resultados. 6. Considerações finais.

The human body *semantic field in the* Atlas Geossociolinguístico Quilombola do Nordeste do Pará *and the* Atlas Geossociolinguístico do Pará

Abstract:

This paper aims to describe the lexical variety observed in quilombola communities in Northeastern Pará, in the Brazilian Amazon Region, and to compare it with the variety

showed in the Atlas Geossociolinguístico do Pará *(ALIPA). To this end, a set of six maps that comprises the* human body *semantic field of the* Atlas Geossociolinguístico Quilombola do Nordeste do Pará *(AGQUINPA) was analysed. The diatopic and diastratic (sex and age) dimensions of variation were considered in the present study. The AGQUINPA aims to map the linguistic variety of spoken Portuguese in quilombola communities through a lexical inventory and to build a database based on the information collected. The quilombola communities studied are the following: a) Cacau (Colares); b) América (Bragança); c) Rio Acaraqui/Campompema (Abaetetuba); d) Taperinha (São Domingos do Capim); e) África (Moju) and f) Laranjituba (Moju). Data were collected by the means of the Semantic-Lexical Questionnaire of the* Brazilian Linguistic Atlas *(ALiB) Project. The theoretical-methodological principles of Geo-sociolinguistics grounded the research. The findings indicate the presence of Bantu root words or Africanities and non-dictionarised variants in the studied communities. They also reveal similarities between the two atlases.*

Key words:

Quilombolas, Geo-sociolinguistics, lexicon.

Contents:

1. Introduction. 2. Dialectology in Brazil. 3. The AGQUINPA. 4. Methodology. 4.1. The network points. 4.2. Profile of informants. 4.3. The instrument of data collection. 4.4. Transcription and mapping of data. 4.5. Comparison between AGQUINPA and ALIPA. 5. Semantic field human body: *results. 6. Final considerations.*

1. Introdução

Os quilombolas constituem um dos povos tradicionais do Brasil, descendentes dos negros escravizados durante os séculos XVII, XVIII e XIX, vindos de diferentes regiões do continente africano. Grande parte das comunidades quilombolas atuais foram formadas entre os séculos XVIII e XIX, a partir de fugas ou pela ocupação das antigas fazendas após a decadência da cana-de-açúcar, ocasionada pelo desmantelamento do sistema escravagista. No Brasil, o maior quilombo de que se teve notícia ficava localizado na Capitania de Pernambuco e resistiu por mais de um século sob a liderança de Zumbi dos Palmares até a sua destruição, em 1695, no período do Brasil Colônia.

Segundo Bezerra Neto (2012: 46), "entre 1708 e 1750, com poucos anos sem a entrada de nenhuma embarcação, 46 navios chegaram ao Estado do Maranhão e Grão-Pará. A partir de 1730, houve um aumento significativo do número de navios negreiros ingressados".

A baixa utilização de mão de obra escrava se dava em virtude do emprego de mão de obra indígena (que passou a ser ilegal posteriormente); desse modo, somente na segunda metade do século XVIII esse cenário sofreu alteração com o comércio de cativos na região. Grande parte dos negros escravizados eram utilizados na região que corresponde, hoje, a áreas localizadas na região Nordeste do Pará, *locus* de nossa pesquisa.

Dentre as principais atividades econômicas dos quilombolas, podemos apontar o agroextrativismo e a pesca. A maioria das comunidades quilombolas usa transporte fluvial. A infraestrutura dessas comunidades é muitas vezes precária (sem sistema de esgotamento sanitário, água encanada e energia elétrica). Os quilombolas entrevistados se autodeclaram negros, a maioria professa a religião católica e, em alguns casos, religião de matriz africana.

De acordo com Castro (2011: 4):

> Calcula-se que dos 75% dos quatro milhões de indivíduos trazidos em escravidão para o Brasil a maioria era proveniente dos reinos do Congo e do Ndongo, de falas quimbundo e quicongo, e foram distribuídos por todo o território brasileiro do começo do tráfico, no século XVI, ao seu final, no século XIX, aos quais vieram se juntar, em número relativamente menor, mas igualmente importante, a partir do século XVIII, os de fala umbundo de Benguela, que foram dirigidos em grande parte para os trabalhos de garimpagem em Minas Gerais.

Esse contingente de negros escravizados não só deixou de falar sua língua de origem, como o quimbundo ou o quicongo, mas também teve que aprender a língua portuguesa

irregularmente[1] (Lucchesi, Baxter, & Ribeiro, 2009), propiciando a interpenetração cultural e linguística, que pode ser observada na fonética, na fonologia, na morfologia e no léxico sob forma de africanias[2], como aponta Castro (2011: 2):

> São marcas lexicais portadoras de elementos culturais compartilhados por toda a sociedade brasileira e que transitam no âmbito da recreação (*samba*, *capoeira*, *forró*, *lundu*, *maculelê*), dos instrumentos musicais (*berimbau*, *cuíca*, *agogô*, *timbau*), da culinária (*mocotó*, *moqueca*, *mungunzá*, *canjica*), da religiosidade (*candomblé*, *macumba*, *umbanda*), das poéticas orais (os *tutus* dos acalantos, o *tindolelê* das cantigas de roda), das doenças (*caxumba*, *tunga*), da flora (*dendê*, *maxixe*, *jiló*, *andu*, *moranga*), da fauna (*camundongo*, *minhoca*, *caçote*, *marimbondo*), dos usos e costumes (*cochilo*, *muamba*, *catimba*), dos ornamentos (*miçanga*, *balangandã*), das vestes (*tanga*, *sunga*, *canga*), da habitação (*cafofo*, *moquiço*), da família (*caçula*, *babá*), do corpo humano (*bunda*, *corcunda*, *banguela*, *capenga*), dos objetos fabricados (*caçamba*, *tipóia*, *moringa*), das relações pessoais de carinho (*xodó*, *dengo*, *cafuné*), dos insultos (*sacana*, *xibungo*, *lelé*), do mando (*bamba*, *capanga*), do comércio (*quitanda*, *bufunfa*, *muamba*, *maracutaia*).

De acordo com Bonvini (2014), as obras *A Arte da língua de Angola*, de autoria do Padre Pedro Dias, e *A língua geral de Mina*, de autoria de António da Costa Peixoto, indicam a situação linguística e provável uso de línguas francas entre os séculos XVII e XVIII, línguas essas que foram gradualmente abandonadas à medida que a língua portuguesa se tornava hegemônica.

Nos anos 70, na cidade de São Paulo, Vogt e Fry (1996) identificaram em uma comunidade quilombola (comunidade do Cafundó[3]), uma série de étimos oriundos do grupo linguístico bantu, o que viria a ser chamado pelos pesquisadores de "língua

1 Lucchesi, Baxter e Ribeiro (2009: 68) defendem o ponto de vista de que ocorreu a transmissão irregular da língua portuguesa, oriunda de um contato abrupto entre línguas africanas, línguas indígenas e a língua geral amazônica; ou seja, o português falado por índios, negros e mestiços provocou alterações consideráveis na estrutura da língua portuguesa, dando espaço à crioulização incipiente e não à crioulização plena, suficiente para formar uma língua crioula.

2 Segundo Castro (2012: 15), "podemos entender africanias como a bagagem cultural submergida no inconsciente iconográfico dos negroafricanos entrados no Brasil em escravidão e que se faz perceptível na língua, na música, na dança, na religião, no modo de ser e de ver o mundo, e, no decorrer dos séculos, como forma de resistência e de continuidade na opressão, transformaram-se e converteram-se em matrizes partícipes da construção de um novo sistema cultural e linguístico que nos identifica como brasileiros".

3 De acordo com Petter (2005: 196), a "língua do Cafundó" é "uma variedade do português regional, um dialeto rural, caracterizada por um léxico reduzido de origem bantu (quimbundo, em particular), com estrutura morfossintática do português. O léxico de origem africana contém cerca de 160 itens, com 15 verbos e dois advérbios. A maioria de seus locutores possui um conhecimento passivo desse repertório, visto que seu uso efetivo vem diminuindo, mantendo-se apenas na comunicação de alguns adultos. As crianças, hoje, aprendem esporadicamente alguns vocábulos, como o nome de alimentos e de alguns animais. Para os cafundoenses, a mais importante função da 'língua', ou da cupópia, como a identificam, é a de código secreto, restrito a membros da comunidade. O uso secreto dessa língua cumpre, na verdade, uma função lúdica, pois agrada-lhes enganar os desavisados. Dessa forma, os falantes se distinguem como descendentes de africanos, superiores a toda degradação social e econômica de que são vítimas".

do Cafundó". Outras situações linguísticas semelhantes foram identificadas em Minas Gerais, no município de Bom Despacho, onde um grupo de negros utilizou uma suposta língua chamada de língua tabatinga (Queiroz, 1998); e em Patrocínio, também no referido estado, nos povoados Capela Nova e Milho Verde (Vogt, & Fry, 1996).

O presente artigo expõe os primeiros resultados referentes à variação no campo semântico *corpo humano* no *Atlas Geossociolinguístico Quilombola do Nordeste do Pará* (doravante AGQUINPA). Tem como objetivo descrever a variedade lexical a partir de seis cartas linguísticas e compará-las com os resultados mapeados, na mesma mesorregião, porém em comunidades não quilombolas do *Atlas Geossociolinguístico do Pará* (ALIPA). Com isso, pretendemos apontar se prevalece, em comunidades quilombolas, a presença de africanias em relação ao seu uso em comunidades não quilombolas. Para a construção do atlas, partimos dos pressupostos da Dialetologia e de seu método, a Geolinguística.

Com o AGQUINPA, pretendemos debater acerca da contribuição das línguas trazidas pelos escravos e se, do ponto de vista lexical, ainda podemos encontrar traços no léxico coletado nas comunidades quilombolas.

2. Dialetogia no Brasil

De acordo com Cardoso (2010), atualmente, o Brasil conta com 45 atlas produzidos. Dez estão publicados (ver quadro 1). Fato curioso é que a última década, que compreende 2001 a 2010, foi considerada a mais produtiva para a Dialetologia no Brasil, em termos de atlas publicados.

Atlas	Ano
Atlas Prévio dos Falares Baianos	1963
Esboço de um Atlas Linguístico de Minas Gerais	1977
Atlas Linguístico da Paraíba	1984
Atlas Linguístico de Sergipe	1987
Atlas Linguístico do Paraná	1994
Atlas Linguístico de Sergipe II	2002
Atlas Linguístico Etnográfico da Região Sul	2002
Atlas Linguístico Sonoro do Estado do Pará	2004
Atlas Linguístico de Mato Grosso do Sul	2007
Atlas Linguístico do Ceará	2010

Quadro 1. Atlas linguísticos publicados no Brasil

Também temos atlas concluídos, porém não publicados, que são resultados de Dissertações de Mestrado e Teses de Doutorado. Vejamos, no quadro 2, os atlas nessa situação:

Atlas	Ano
Atlas Linguístico do Amazonas (tese)	2004
Atlas Linguístico de Adrianópolis (dissertação)	2004
Atlas Fonético do Entorno da Baía da Guanabara (dissertação)	2006
Atlas Linguístico Rural do Município de Ponta-Porã – Mato Grosso do Sul (dissertação)	2006
Atlas Geolinguístico do Litoral Potiguar (tese)	2007
Atlas Linguístico do Paraná II (tese)	2007
Atlas Linguístico-etnográfico da Região Oeste do Paraná	2007
Atlas Linguístico de São Francisco do Sul	2007
Atlas Semântico-Lexical da Região do Grande ABC (tese)	2007
Micro-atlas fonético do Estado do Rio de Janeiro – Micro AFERJ (tese)	2008
Atlas Linguístico de Iguatu/CE (dissertação)	2009
Atlas Linguístico da Mesorregião Sudeste do Mato Grosso (dissertação)	2009
Atlas Linguístico da Mata Sul de Pernambuco (dissertação)	2009
Atlas Semântico-lexical de Caraguatatuba, Ilha Bela, São Sebastião e Ubatuba – Litoral Norte de São Paulo (tese)	2010
Atlas Linguístico de Buíque (monografia)	2011
Atlas Linguístico de Capistrano (dissertação)	2011
Atlas Linguístico dos Falares do Baixo Amazonas (dissertação)	2011
Atlas Linguístico do Oeste Potiguar (tese)	2012
Atlas Linguístico dos falares do Alto Rio Negro (tese)	2012
Atlas Semântico Lexical da Região Norte do Alto Tietê (tese)	2012
Atlas Geossociolinguístico de Londrina: um estudo em tempo real e tempo aparente (dissertação)	2012
Atlas Semântico Lexical do Estado de Goiás (tese)	2012
Atlas Linguístico de Pernambuco (tese)	2013
Atlas Linguístico de Corumbá e Ladário (dissertação)	2013
Atlas Linguístico Etnográfico da fronteira Brasil/Paraguai: um registro das línguas em contato (tese)	2013

Quadro 2. Atlas linguísticos concluídos, mas não publicados

Além dos atlas regionais e de pequeno domínio linguístico, o Brasil também possui um atlas nacional: o *Atlas Linguístico do Brasil*[4] (Cardoso *et alii*, 2014). O *Atlas Linguístico do Brasil* (ALiB) comporta 250 localidades, das quais os dados já foram coletados, e 1.100 informantes estratificados socialmente, levando em consideração: a) faixa etária de 18 a 30 anos, e de 50 a 65 anos; b) sexo/gênero: masculino e feminino; c) nível de escolaridade: ensino fundamental e ensino superior (apenas para as capitais).

Ao lado desses atlas, ainda temos aqueles que estão em desenvolvimento, como o *Atlas Linguístico Sonoro das Línguas Indígenas do Brasil* e o *Atlas Linguístico do Maranhão*. Na próxima seção, abordaremos os procedimentos adotados na elaboração do AGQUINPA.

3. O AGQUINPA

O levantamento dos dados para a constituição do AGQUINPA foi realizado nas comunidades denominadas de remanescentes de quilombo, localizadas na mesorregião Nordeste do Pará.

Esta região foi escolhida como foco do AGQUINPA pelo fato de ter o maior adensamento populacional de comunidades[5] quilombolas no Estado (Salles, 2005). Outros fatores sócio-históricos como ocupação territorial, migração, participação de negros escravizados na lavoura da região e que atuaram na Revolução Cabana, além da proximidade em relação ao principal centro urbano que compõe a Região Metropolitana de Belém também foram considerados.

Tendo em vista o panorama apresentado anteriormente, podemos dizer que o AGQUINPA pretende: a) realizar levantamento histórico e geossociolinguístico das comunidades pesquisadas; b) descrever as variedades linguísticas usadas nas comunidades quilombolas, obtidas por meio da aplicação de Questionário

4 O projeto de construção do atlas linguístico nacional retoma o decreto de 1952, no entanto, só a partir de 1996 (44 anos depois), na Universidade Federal da Bahia, foi criado um Comitê Nacional com esta finalidade. O Comitê Nacional definiu a rede de pontos, que incluiu capitais, com exceção do Distrito Federal e Palmas, e não capitais.

5 De acordo com o levantamento de Gomes (2015: 145), o Pará possui 523 comunidades remanescentes de quilombo reconhecidas e certificadas, das quais 133 são tituladas. Outras fontes apontam que há 398 comunidades remanescentes de quilombo no Pará (Bentes, 2013: 82), com dados da FCP, UnB, INCRA, SEPPIR, Programa Raízes e CEDENPA. Segundo Bentes (2013: 83), "pode haver alguma diferença (de quantitativo de comunidades) [...] Existem também casos de comunidades que já são praticamente cidades e por isso deixaram de ser 'quilombola'; outras, mesmo tendo características, não querem ser 'quilombolas'".

Semântico-lexical (QSL)[6]; e c) elaborar cartas linguísticas semântico-lexicais pluridimensionais do português falado nas comunidades quilombolas pesquisadas.

4. Metodologia

Nesta seção, apresentaremos a metodologia utilizada no AGQUINPA, desde a escolha da rede de pontos até o processo de elaboração computadorizado das cartas linguísticas.

4.1. A rede de pontos

Compuseram a rede de pontos do AGQUINPA apenas comunidades com terras tituladas e de população autodenominada quilombola. Entendemos que o critério que toma por base comunidades com terras tituladas pelo Instituto Nacional de Colonização e Reforma Agrária (INCRA) foi o mais adequado, em virtude do estudo antropológico prévio a que cada comunidade é submetida. Ele é usado para determinar se aquela terra ocupada pertence a remanescentes de quilombos. No Pará, atualmente, há em torno de 133 comunidades quilombolas tituladas[7] e muitas outras em processo de titulação.

Essas comunidades localizam-se predominantemente em áreas rurais dos municípios que compõem a mesorregião Nordeste do Pará. A seguir, apresentamos o quadro de municípios e pontos de inquérito:

Microrregião	Município	Comunidade	Nª do ponto
Salgado	Colares	Cacau	1
Bragantina	Bragança	América	2
Cametá	Abaetetuba	Rio Acaraqui/Campompema	3
Guamá	São Domingos do Capim	Taperinha	4
Tomé-Açu	Moju	África	5
Tomé-Açu	Moju	Laranjituba	6

Quadro 3. Lista de pontos de inquérito por microrregião e município pertencentes à Mesorregião Nordeste do Pará

6 O referido questionário é instrumento de coleta de dados idealizado e usado no âmbito do *Projeto Atlas Linguístico do Brasil*. No AGQUINPA, ele é usado com adaptações.

7 Para detalhes sobre os critérios adotados para titulação das terras, consultar *O caminho da titulação* http://www.cpisp.org.br/terras/html/comosetitula_caminho.aspx (consulta em 23.01.2017).

4.2. Perfil dos informantes

Os informantes foram selecionados considerando-se a faixa etária/geração (18 a 30 e 50 a 65 anos) e sexo/gênero (homens e mulheres). Não consideramos a escolaridade pelo fato de que nessas comunidades a maioria dos informantes cursou até o Ensino Fundamental e os que possuem mais escolarização geralmente vivem na sede do Município, ou seja, residem a maior parte do tempo fora da comunidade e em áreas urbanizadas. Foram alvo dos inquéritos falantes nascidos e criados na localidade, que dela não tenham se afastado por período maior que três anos, filhos de pais também nascidos e criados nas comunidades quilombolas. Também levamos em consideração boas condições de fonação, disposição e tempo para as gravações.

4.3. O instrumento de coleta de dados

Para a elaboração do AGQUINPA, buscamos no ALiB os avanços acumulados em quase duas décadas de discussões acerca da Geolinguística, o que motivou a utilização do Questionário Semântico-lexical (doravante QSL) do atlas nacional. Isso também propiciaria a comparação de dados entre os atlas que o utilizaram, inclusive com os dados do próprio ALiB.

O QSL do ALiB é composto por 202 itens lexicais distribuídos em 14 campos lexicais, a saber: a) acidentes geográficos; b) fenômenos atmosféricos; c) astros e tempo; d) atividades agropastoris; e) fauna; f) corpo humano; g) ciclos da vida; h) convívio e comportamento social; i) religião e crenças; j) jogos e diversões infantis; k) habitação; l) alimentação e cozinha; m) vestuário e acessórios; e n) vida urbana. O QSL figura, dentre os questionários do ALiB, como o de maior extensão e tem regularmente sido adotado na coleta de dados para a construção de vários atlas linguísticos no Brasil.

4.4. Transcrição e cartografia dos dados

De posse das gravações obtidas por meio da aplicação do QSL, foi realizada a transcrição grafemática dos itens lexicais a partir do uso do programa computacional de anotação linguística ELAN 4.6.1[8]. Após a transcrição, criamos um banco de dados semântico-lexical que serviu de base para a cartografia dos dados.

Para a elaboração das cartas semântico-lexicais, utilizamos o programa computacional de processamento geoespacial QGIS 2.6 *Wien*[9]. Esse *software* proporcionou a composição das cartas de modo claro e elegante, com o objetivo de facilitar a leitura e a posterior exegese.

---

8 Programa disponível em https://tla.mpi.nl/tools/tla-tools/elan/

9 *Quantum Geographic Information System* (QGIS), disponível em http://www.qgis.org/en/site/

4.5. Comparação entre o AGQUINPA e o ALIPA

De acordo com Guedes (2012), o *Atlas Geossociolinguístico do Pará* (ALIPA) foi construído em 1996, com o objetivo de mapear a diversidade linguística no Estado do Pará, a partir de dados de 50 localidades em seis mesorregiões, dentre as quais a mesorregião Nordeste do Pará (também mapeada pelo AGQUINPA). O ALIPA considera localidades rurais (não quilombolas) e urbanas, informantes estratificados socialmente por sexo, faixa etária, escolaridade e renda.

Escolhemos os resultados do ALIPA para a comparação com os mapeados no AGQUINPA por conta da coincidência relativa à região mapeada, de alguns fatores de estratificação social e da escolha metodológica. Os dois atlas mapeiam a mesorregião Nordeste do Pará e consideram, em alguns casos, os mesmos pontos de inquérito na região. O ALIPA é formado pelos seguintes pontos: Abaetetuba, Moju, Maracanã e Viseu; os dois primeiros, Abaetetuba e Moju, também integram o AGQUINPA (ver quadro 3). Além disso, as questões estudadas são comuns em ambos os atlas.

No presente artigo, além de fazermos a referida comparação, investigaremos a dicionarização dos étimos registrados no AGQUINPA, especificamente daqueles que apresentam influência africana. Tomaremos como base os seguintes dicionários: *Dicionário Caldas Aulete* (Aulete, 2016), *Dicionário Houaiss da Língua Portuguesa* (Houaiss, 2010), além dos dicionários de étimos bantu: *Dicionário Kimbundu-Português* (Assis Júnior, 1947) e *Novo Dicionário banto do Brasil* (Lopes, 2012).

5. O campo semântico *corpo humano:* resultados

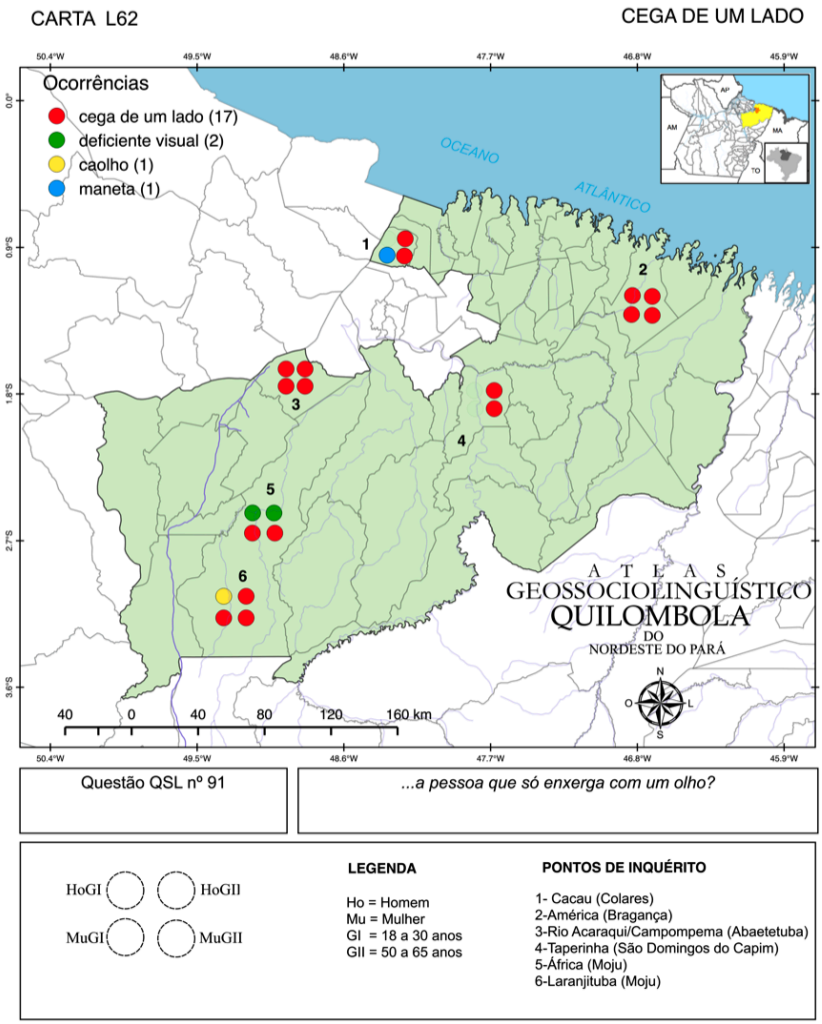
O campo semântico *corpo humano* apresenta 32 questões que englobam desde estruturas anatômicas do corpo humano até patologias que afetam essas partes do corpo. Dessas 32 questões, 14 cartas linguísticas foram geradas. Optamos por descrever apenas aquelas que apresentam alguma variante considerada africania.

A seguir, apresentaremos a descrição das cartas referentes às questões nº 91 *(... a pessoa que só enxerga com um olho*), nº 100 *(A pessoa que não tem dentes),* nº 102 *(... a sujeirinha dura que se tira do nariz com o dedo),* nº 104 *(... esta parte atrás do pescoço),* nº 107 *(... a pessoa que tem um calombo nas costas)* e nº 109 *(... o mau cheiro embaixo dos braços).*

a) *A pessoa que só enxerga com um olho*

A questão nº 91, do QSL, que inquiriu sobre a pessoa que só enxerga com um olho, apresentou as seguintes variantes: *cega de um lado* (dezessete ocorrências),

deficiente visual (duas ocorrências), *caolho* (uma ocorrência) e *maneta* (uma ocorrência) como podemos observar no mapa que segue:



Mapa 1. Carta L62 (questão nº 91)

De acordo com as respostas registradas, podemos afirmar que prevaleceu a variante *cega de um lado*. As variantes *maneta* e *caolho* aparecem em menor número, com uma ocorrência cada. A variante *cega de um lado* ocorreu em todos os pontos

e prevaleceu em quase todos eles: 1 (Cacau), 2 (América), 3 (Rio Acaraqui/ Campompema), 4 (Taperinha) e 6 (Laranjituba). No ponto de inquérito 5 (África), tivemos *deficiente visual*, uma variante de caráter mais generalizante, pois não especifica o tipo de deficiência, com 50% das ocorrências. Podemos afirmar que a norma lexical para este item nos pontos de inquérito do AGQUINPA é pouco heteroléxica, se comparada às duas últimas a serem apresentadas, L71 e L73. Todas as variantes registradas nesta questão são dicionarizadas.

Segundo Lopes (2012: 76), a variante *caolho* é considerada léxico híbrido no português: *olho*, com a aposição do prefixo diminutivo do quimbundo *ka*.

Os resultados do ALIPA, mapeados por Guedes (2012), comprovam que nesse atlas prevalece a variante *cego de um lado* (treze ocorrências) e, em menor número, as variantes *defeituoso* (duas ocorrências) e *caolho* (uma ocorrência). São resultados parecidos com os mapeados no AGQUINPA. Portanto, podemos dizer que o AGQUINPA e o ALIPA possuem usos que se aproximam e distinguem-se em alguns aspectos. No ALIPA, tivemos o registro de *defeituoso* e no AGQUINPA de *deficiente*.

Cabe notar também que no AGQUINPA a variante *cega de um lado* apresenta-se, prioritariamente, na forma feminina, já no ALIPA sob a forma masculina.

No gráfico 1, temos os resultados referentes às dimensões diassexual e diageracional:

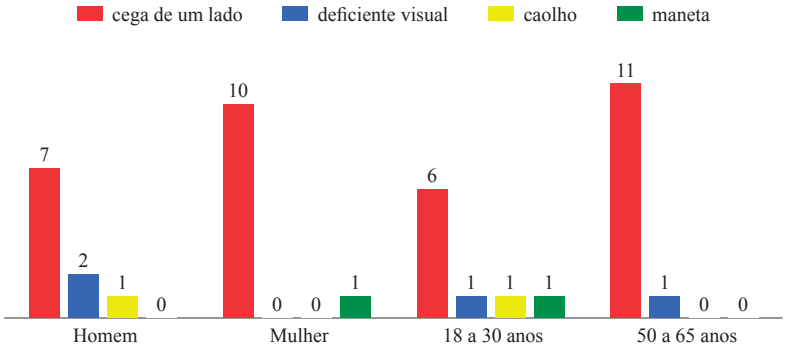
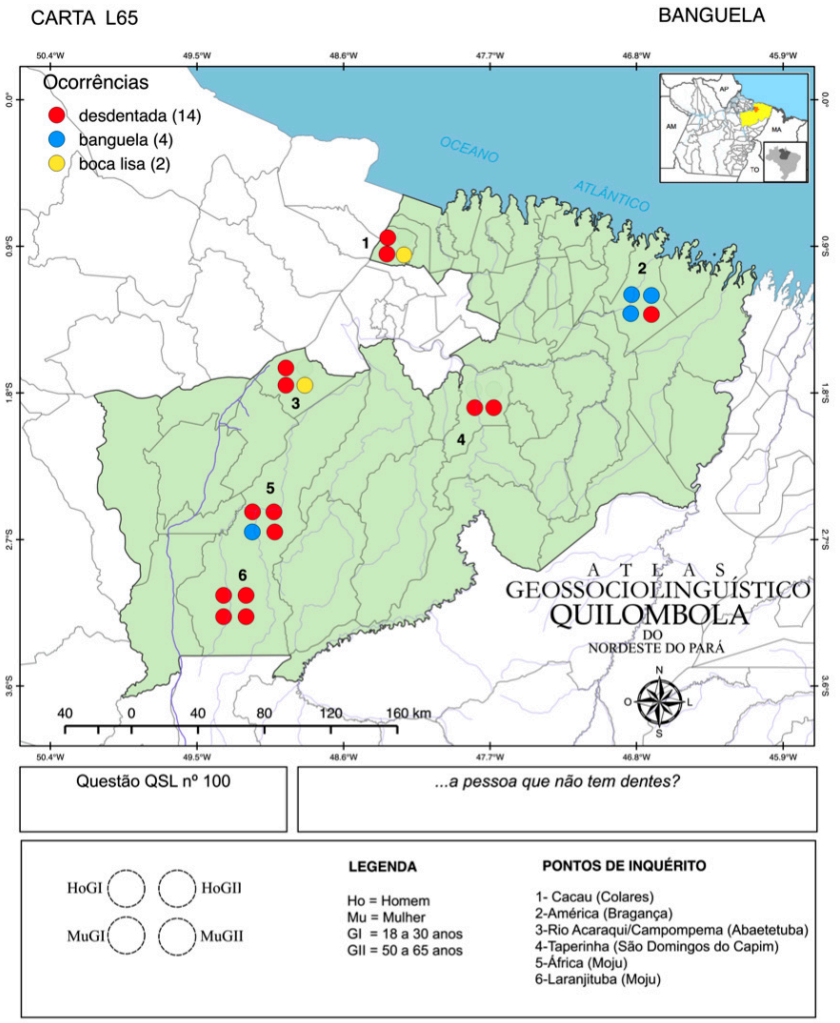


Gráfico 1. Resultados diassexuais e diageracionais para a questão nº 91 do QSL

No gráfico 1, podemos visualizar a prevalência da variante *cega de um lado* entre as mulheres (dez ocorrências) e também entre os informantes de 50 a 65 anos (onze ocorrências). Também foi possível constatar que os informantes de 18 a 30 anos apresentaram o maior número de variantes, ou seja, quatro, a saber: *cega de um lado*, *deficiente visual*, *caolho* e *maneta*. Curiosamente, a forma de influência africana é usada por falante da primeira faixa, do sexo masculino.

b) *A pessoa que não tem dentes*

A questão nº 100, do QSL, apresentou as seguintes variantes: *desdentada* (quartorze ocorrências), *banguela* (quatro ocorrências) e *boca lisa* (duas ocorrências). Vejamos os resultados dispostos no mapa a seguir:



Mapa 2. Carta L65 (questão nº 100)

Como podemos observar no mapa 2 (carta L65), a variante *desdentada* ocorreu em todos os pontos. Só não prevaleceu no ponto de inquérito 2 (América), onde dominou a variante *banguela*. Em contrapartida, foi a única usada no ponto 6 (Laranjituba). Podemos dizer que esta carta apresenta norma pouco heteroléxica. É importante destacar que todas as variantes registradas nesta questão estão dicionarizadas.

No ALIPA, tivemos, para a mesma questão, também a prevalência da variante *desdentada* (dez ocorrências), seguida de *banguela* (quatro ocorrências), *sem dente* (três ocorrências) e *boca de poraquê*[10] (uma ocorrência). Tanto no ALIPA quanto no AGQUINPA a variante *desdentada* predomina, com o maior número de ocorrências, seguida de *banguela*, variante considerada uma africania. A forma tem origem no étimo bantu *ba-ngela*, que possui a acepção de dente limado. A seguir, apresentamos os resultados diassexuais e diageracionais:

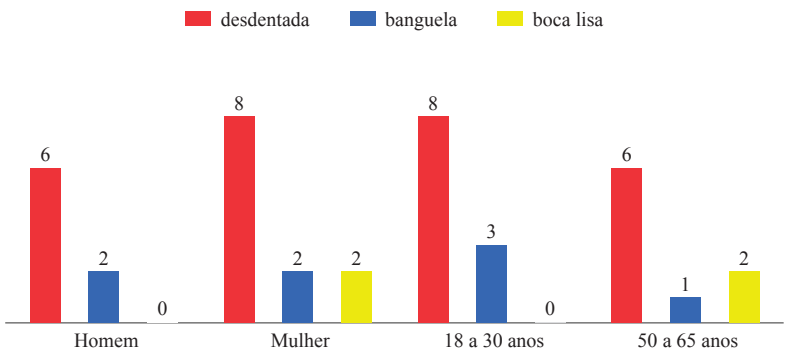


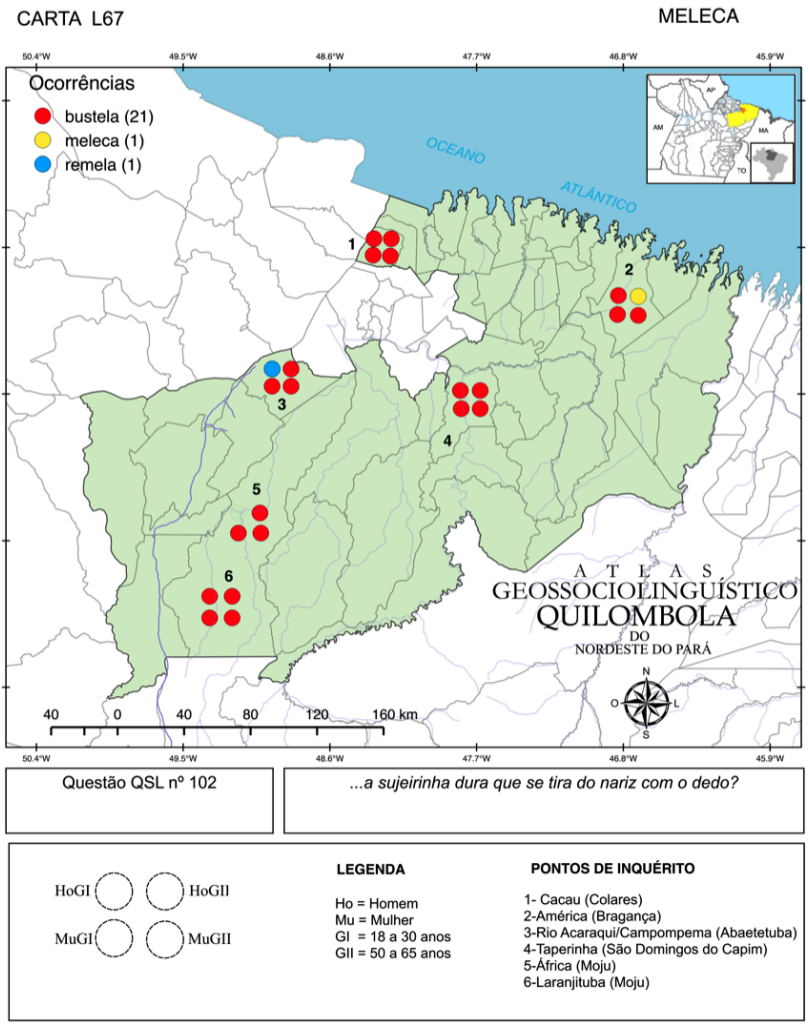
Gráfico 2. Resultados diassexuais e diageracionais para a questão nº 100 do QSL

No gráfico 2, podemos observar que a variante *desdentada* prevaleceu entre as mulheres (oito ocorrências) e entre os informantes de 18 a 30 anos (oito ocorrências). As mulheres de 50 a 65 anos apresentaram o número máximo de variantes (três no total, a saber, *desdentada*, *banguela* e *boca lisa*). Nessa questão, a africania ocorre igualmente em ambos os sexos e prevalece na faixa etária de 18 a 30 anos.

c) *A sujeirinha dura que se tira do nariz com o dedo*

As variantes registradas para a questão nº 102, do QSL, foram *bustela* (vinte e uma ocorrências), *meleca* (uma ocorrência) e *remela* (uma ocorrência). Vejamos os resultados na carta a seguir:

10 Peixe elétrico característico da região Amazônica *(Electrophorus electricus)*.



Mapa 3. Carta L67 (questão nº 102)

No mapa 3, é possível observar que a variante *bustela* ocorreu em todos os pontos de inquérito e teve uso categórico nos pontos 1 (Cacau), 4 (Taperinha), 5 (África) e 6 (Laranjituba). No ALIPA, a variante *bustela* também foi predominante; obteve quinze ocorrências. A variante *meleca* apresentou uma única ocorrência e também ocorreu no ALIPA. A forma é considerada uma africania na língua portuguesa; tem origem no quimbundo (*mele adikeleka*), que tem a acepção de leite coagulado

ou coalhado. Para esta questão, todas as variantes registradas estão devidamente dicionarizadas. A seguir, apresentamos os resultados diassexuais e diageracionais para a questão:

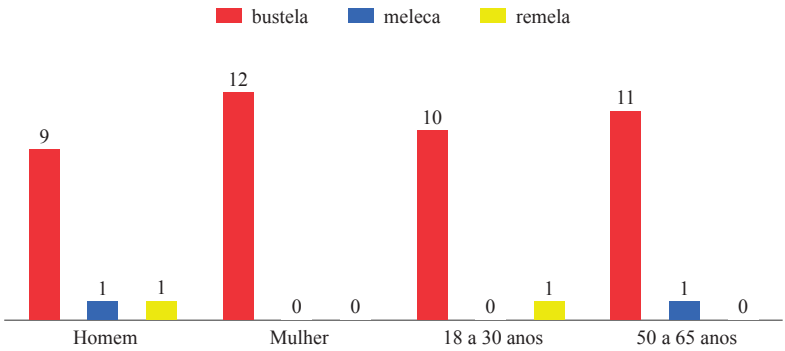


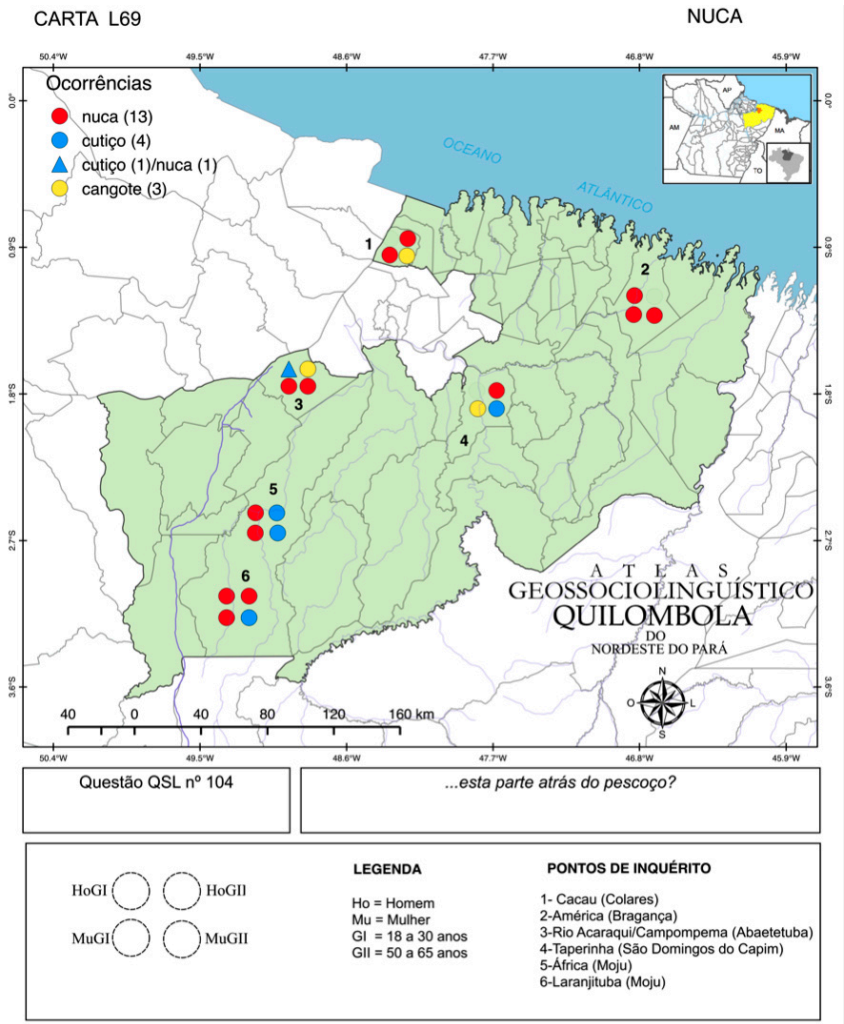
Gráfico 3. Resultados diassexuais e diageracionais para a questão nº 102 do QSL

Do ponto de vista diassexual, os resultados comprovam que a variante *bustela* prevaleceu entre as mulheres, na verdade, nenhuma outra variante foi usada por elas. Do ponto de vista diageracional, a variante apresentou frequência muito aproximada, sem diferença significativa. A variante *meleca* ocorreu exclusivamente na segunda faixa e *remela* na primeira. A forma de influência africana ocorreu na fala de um homem da segunda faixa.

d) *Esta parte por trás do pescoço*

A questão de nº 104, que inquiriu a respeito da parte de trás do pescoço, apresentou as seguintes respostas: *nuca* (quatorze ocorrências), *cutiço* (cinco ocorrências) e *cangote* (três ocorrências). Vejamos a distribuição diatópica das variantes:

No mapa 4, podemos observar que *nuca* ocorre em todos os pontos. É a única usada no ponto 2. Ela prevalece nos pontos de inquérito 1 (Cacau), 2 (América), 3 (Rio Acaraqui/Campompema) e 6 (Laranjituba). Nos pontos 3 (Rio Acaraqui/Campompema) 4 (Taperinha), tivemos o registro de três das quatro variantes identificas (*nuca*, *cutiço* e *cangote*). No ponto 5 (África), as variantes *nuca* e *cutiço* apresentaram a mesma frequência. Os resultados mostram que a norma representada nesta região figurou como pouco heteroléxica. Assim como no AGQUINPA, o ALIPA teve a prevalência da variante *nuca* (oito ocorrências), seguido da variante *cangote* (sete ocorrências). Dentre as respostas registradas, apenas a variante *cutiço* não está dicionarizada e não identificamos a sua origem.



Mapa 4. Carta L69 (questão nº 104)

De acordo com Lopes (2012), o étimo *cangote*[11] tem origem na palavra *en-goti*, da língua cuanhama, com acepção de pescoço, portanto, podemos considerá-lo como uma africania presente na língua portuguesa. Curiosamente, essa forma é mais

11 De acordo com Nascentes (1966), *cangote* tem origem no étimo espanhol *cogote*.

frequente no ALIPA do que no AGQUINPA.

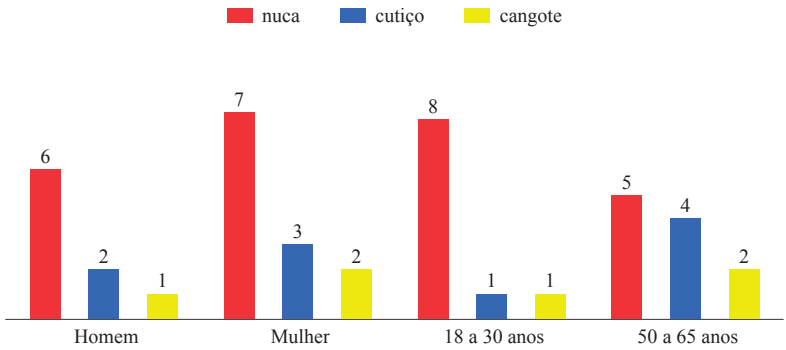


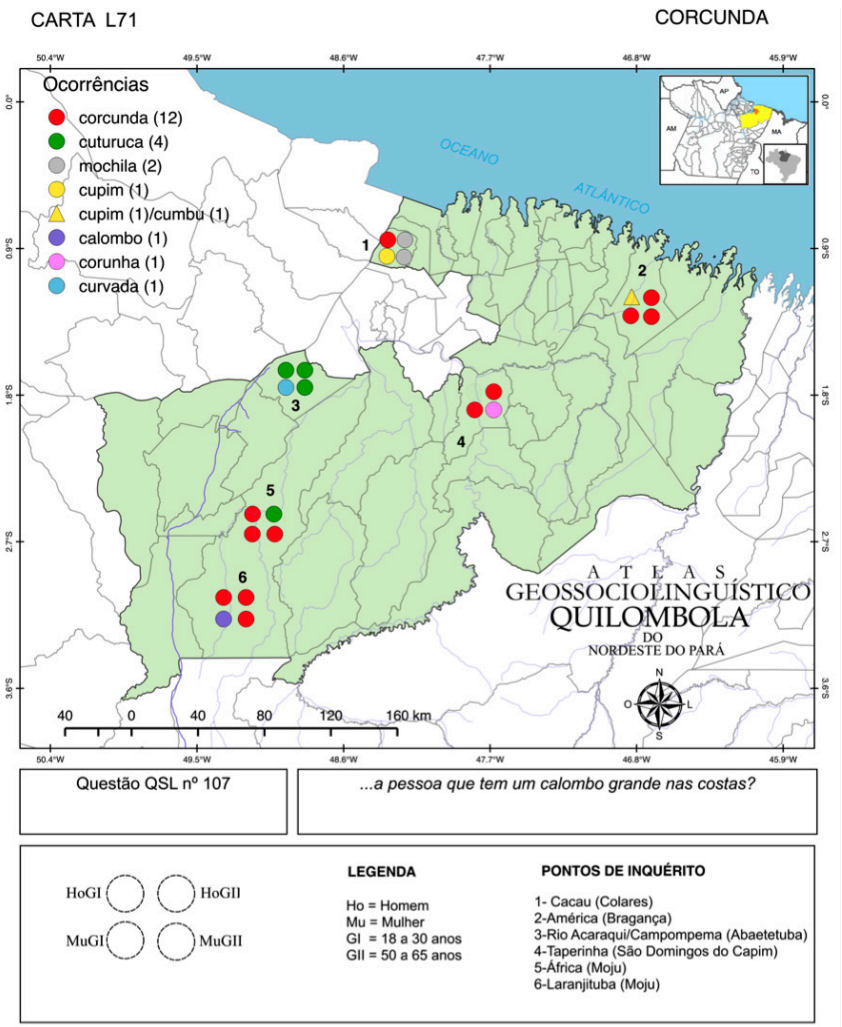
Gráfico 4. Resultados diassexuais e diageracionais para a questão nº 104 do QSL

No gráfico 4, é possível observar que para a variante *nuca* não houve diferença significativa entre homens e mulheres, enquanto que do ponto de vista diageracional essa variante prevaleceu entre os informantes de 18 a 30 anos, com oito ocorrências no total, o que pode indicar sua tendência ao uso. A variante considerada como influência africana ocorreu nos dois sexos e nas duas faixas etárias, no entanto, é mais usada pelas mulheres e pelos que apresentam mais idade.

e) *A pessoa que tem um calombo grande nas costas*

A questão nº 107, do QSL, apresentou as seguintes variantes: *corcunda* (doze ocorrências), *cuturuca* (quatro ocorrências), *mochila* (duas ocorrências), *cupim* (duas ocorrências), além de *cumbu*, *calombo*, *corunha* e *curvada*, com uma ocorrência cada, como podemos observar no mapa 5:

Essa carta é a mais heteroléxica das apresentadas até agora, ou seja, a que possui o maior número de variantes, oito ao todo. Mas, diferentemente do que detectamos para algumas cartas, nessa a variante mais frequente, *corcunda*, não ocorreu em todos os pontos, isto é, não foi verificada no ponto 3 (Rio Acaraqui/Campompema). Esta variante prevalece nos pontos de inquérito 2 (América), 4 (Taperinha), 5 (África) e 6 (Laranjituba). No ponto de inquérito 1 (Cacau), prevaleceu a variante *mochila* e, no ponto de inquérito 3 (Rio Acaraqui/Campompema), a variante de maior ocorrência foi *cuturuca*.



Mapa 5. Carta L71 (questão nº 107)

Há alta produtividade lexical para essa questão. Vejamos algumas observações sobre elas. Apenas as variantes *cuturuca* e *cumbu* não estão dicionarizadas. Destacamos para esta questão, a presença das variantes *corunha*[12] e *cumbu*. A primeira, de origem galego-portuguesa, tem a acepção de caroço (Aulete, 2016; Vasconcelos, 1910),

12 Corunha é a segunda maior cidade da Galícia. É um topônimo.

e foi registrada no ponto 4 (Taperinha). A segunda, de origem bantu, tem acepção dicionarizada bem distinta do que encontramos aqui, pois em quimbundo, o étimo *cumbu* ou *kumbu* possui acepção de dinheiro (Assis Júnior, 1947). Cabe frisar, aqui, no entanto, que a noção de 'pessoa que tem um calombo grande nas costas' envolve uma protuberância. Quando uma pessoa está com o bolso ou a carteira cheios de dinheiro, não raro também se alude a uma protuberância, usando-se, inclusive, as denominações como bolso cheio, gordo, avantajado. Talvez, a forma tenha sido usada em alusão a isso. Entendimento semelhante vale para as formas, aqui mencionadas, dicionarizadas com acepções diferentes das que foram usadas na presente pesquisa.

Com base em nosso levantamento, a variante *cuturuca* remete à palavra *kuturuka*, presente no vocabulário da língua chichewa ou nianja, que faz parte do grupo de línguas bantu e que é falada em Moçambique, Zâmbia e Zimbabwe. De acordo com Laws (2012), tem a acepção de 'para fora'.

De acordo com Lopes (2012: 79), a variante *corcunda* tem origem no étimo quimbundo *karikunda* e tem a acepção de protuberância deforme nas costas, enquanto que a variante *calombo* não tem origem confirmada, mas remete ao étimo quicongo *lombolo* e quimbundo *mulumbu*, ambos com a acepção de tumefação cutânea, de corcova ou de excrecência.

Na comparação com os dados do ALIPA, foi possível constatar a prevalência da variante *corcunda* (nove ocorrências), assim como no AGQUINPA. Naquele atlas, foram registradas também as variantes *aleijado* e *lombinho*, com duas ocorrências cada, além das variantes *torto*, *cotoco, cupim nas costas* e *calombinho*, que aparecem em menor número, com uma ocorrência cada. Destacamos no ALIPA o registro da variante *caturupa* em dois pontos de inquérito. Suspeitamos que essa forma seja uma variante fonética de *cuturuca*, que teve registro no AGQUINPA. Nessa carta, especificamente, o número de variantes de origem africana é bem superior ao encontrado no ALIPA.

A seguir, apresentamos os resultados diassexuais e diageracionais para a questão:

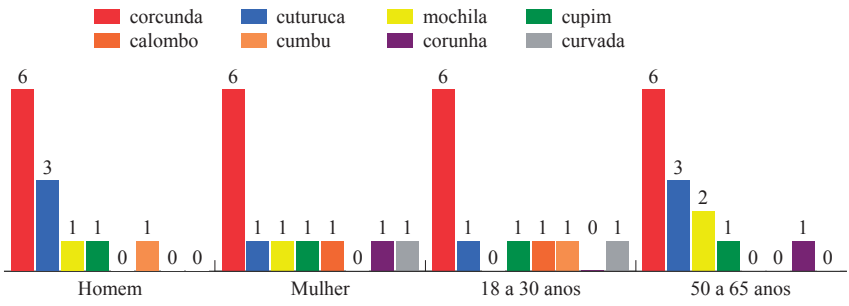
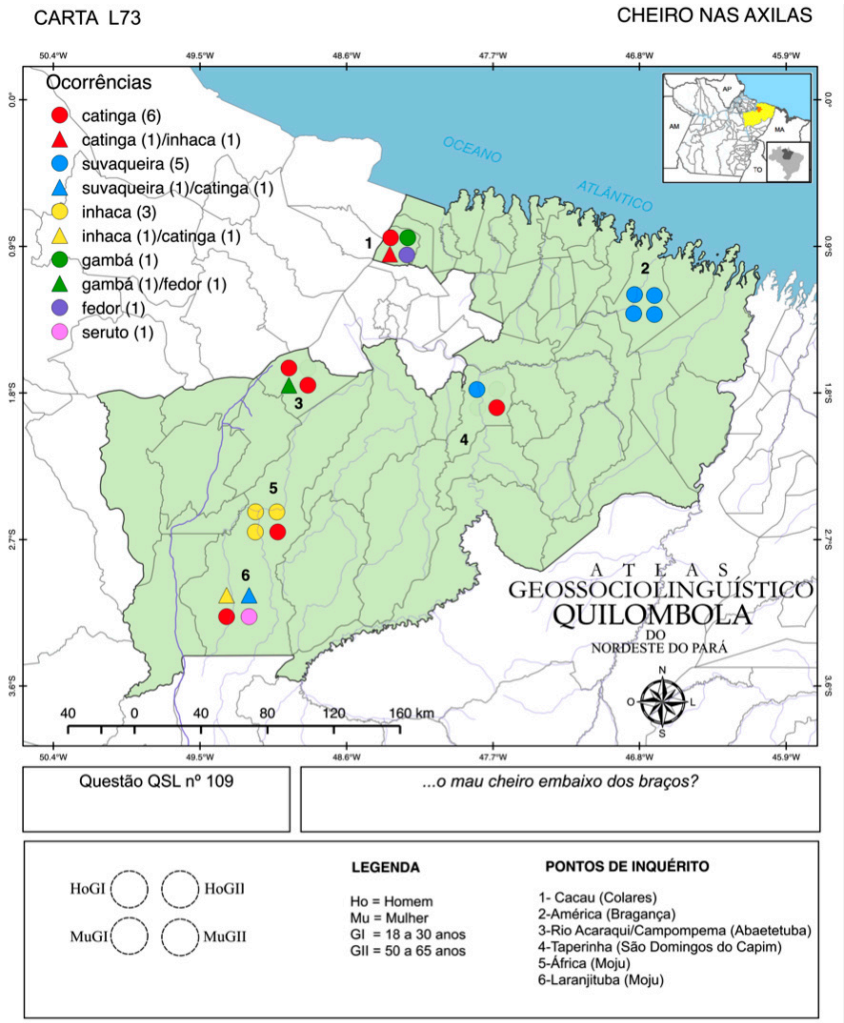


Gráfico 5. Resultados diassexuais e diageracionais para a questão nº 107 do QSL

Como é possível observar no gráfico 5, a variante *corcunda* apresentou a mesma frequência tanto para homens quanto para mulheres. Isso ocorreu também para a faixa etária. As mulheres apresentaram o maior número de variantes em sua fala, ou seja, sete no total, a saber: *corcunda*, *cuturuca*, *mochila*, *cupim*, *calombo*, *corunha* e *curvada*. As formas apontadas como de origem africana aparecem nos dois sexos e nas duas faixas etárias. A variante *corunha* ocorreu na fala de uma mulher da segunda faixa etária.

f) *O mau cheiro embaixo dos braços*



Mapa 6. Carta L73 (questão nº 109)

A questão nº 109 do QSL apresentou as variantes *catinga* (nove ocorrências), *suvaqueira* (seis ocorrências), *inhaca* (cinco ocorrências), *gambá* (duas ocorrências), *fedor* (duas ocorrências) e *seruto* (uma ocorrência). Veja-se a distribuição diatópica das variantes no mapa 6.

Como é possível observar no mapa 6, a variante *catinga* ocorreu em cinco dos seis pontos investigados. Prevaleceu nos pontos de inquérito 1 (Cacau), 3 (Rio Acaraqui/Campompema) e 6 (Laranjituba). No ponto 2 (América), prevaleceu a variante *suvaqueira*, sendo a única usada nesse ponto. No ponto 5 (África), tivemos a prevalência da variante *inhaca* e, no ponto de inquérito 4 (Taperinha), as variantes *catinga* e *suvaqueira* tiveram a mesma quantidade de ocorrências (uma para cada variante). Do ponto de vista da dicionarização, apenas a variante *seruto* não se encontra dicionarizada.

Nos dados do ALIPA, tivemos maior ocorrência da variante *catinga* (seis ocorrências), seguido de *gambá* (cinco ocorrências), *inhaca* (duas ocorrências), além de *suvaqueira*, *fedorendo* e *suor forte*, com uma ocorrência cada. Como podemos observar, a variante *catinga* se destacou nos dois atlas em termos de frequência. Os resultados de um e de outro atlas são bem parecidos.

Segundo Lopes (2012), do ponto de vista etimológico, a variante *catinga* tem origem na língua quicongo *katinga* ou *okatinga*, com a acepção de cheiro repugnante e, portanto, pode ser considerada uma africania, assim como a variante *inhaca*, que tem origem no étimo bantu *yaka*.

A seguir, comentaremos os resultados diassexuais e diageracionais para a questão:

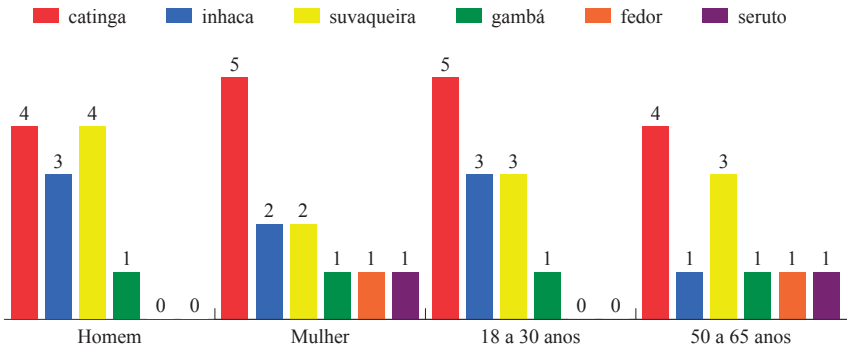


Gráfico 6. Resultados diassexuais e diageracionais para a questão nº 107 do QSL

No gráfico 6, podemos observar que a variante *catinga* apresentou pouca diferença entre homens e mulheres, assim como entre as duas faixas etárias. Entre os homens, as variantes *catinga* e *suvaqueira* apresentaram a mesma quantidade de ocorrências (quatro ocorrências). As mulheres apresentaram o maior número de variantes (seis variantes no total), ou seja, todas as variantes registradas ocorrem no falar feminino. Do ponto de vista diageracional, os informantes da faixa etária de 50 a 65 também apresentaram o número máximo de variantes registradas.

6. Considerações finais

Este artigo buscou apresentar variedades lexicais encontradas no AGQUINPA e compará-las com os resultados do ALIPA.

Na comparação com os dados do ALIPA, específicos da mesorregião Nordeste do Pará, pudemos observar muitas semelhanças, pois em todas as seis cartas comparadas, as variantes de maior ocorrência coincidiram nos dois atlas e há usos de formas africanas num e noutro, o que nos leva a afirmar, com base nos dados colhidos até aqui, que não há diferenças elevadas entre os usos lexicais de comunidades quilombolas do AGQUINPA e de comunidades não quilombolas do ALIPA, para as questões comparadas; mas elas existem. A carta L71 (Mapa 5) foi a que apresentou caráter mais heteroléxico e formas africanas não encontradas no ALIPA. Por outro lado, esses resultados apontam para um contínuo entre os usos lexicais dessas comunidades.

Talvez devamos falar de um português rural. Nesse português, há influências recíprocas entre as comunidades que as integram. É possível afirmar que em ambos os atlas há, no léxico, a presença de formas africanas que resistem e que já não são tão usuais nos centros urbanos da região.

As cartas para as questões 107 e 109 foram as que se apresentaram mais heteroléxicas. Dentre as respostas registradas, destacaram-se variantes que podem ser consideradas africanias, pertencentes principalmente do grupo linguístico bantu, como *caolho*, *banguela*, *meleca*, *cumbu*, *inhaca*, *catinga*, *cangote*, *calombo*, *corcunda* e *cuturuca*.

Levando em consideração as dimensões diassexual e diageracional, tivemos variantes prevalecendo entre as mulheres (*cega de um lado*, *desdentada* e *bustela*), entre os mais jovens (*desdentada* e *nuca*), além de variantes que apresentaram pouca diferença diassexual e diageracional, como *nuca*, *corcunda*, *catinga* e *bustela*.

Agradecimentos

Deixamos aqui o registro de agradecimento ao Instituto da Lingua Galega pelo acolhimento durante o estágio doutoral, especialmente ao prof. Dr. Xulio Sousa e César Osorio que nos auxiliaram na elaboração informatizada dos mapas, assim como a profª. Dra. Rosário Alvarez que nos orientou na organização e constituição do banco de dados do AGQUINPA. Agradecemos à Coordenação de apoio ao Pessoal de Pós-Graduação (CAPES) pela concessão da bolsa, ao Programa de Pós-Graduação em Letras da Universidade Federal do Pará pela aprovação do projeto e também aos líderes das comunidades quilombolas Raimundo Magno (Quilombo África e Laranjituba), Rosete (quilombo América), Sr. Avelino (quilombo Cacau), a Associacão das Comunidades Remanescentes de Quilombos das Ilhas de Abaetetuba (ARQUIA) e a assistente social da Prefeitura de Castanhal Siane Karla.

Referências bibliográficas

Assis Júnior, António de (1947). *Dicionário Kimbundu-Português*. Luanda: Argente, Santos & Cia.

Aulete, Caldas (2016). *Dicionário Aulete digital*. Rio de Janeiro: Lexikon. http://www.aulete.com.br/site.php?mdl=aulete_digital (Consulta em 19.1.2017).

Bentes, Nilma (2013). *Aspectos da trajetória da população negra no Pará*. Belém: GEAM / UFPA.

Bezerra Neto, José Maia (2012) [2001]. *Escravidão Negra no Grão-Pará (séculos XVII-XIX)*. Belém: Paka-Tatu.

Bonvini, Emilio (2014). "Línguas africanas e português falado no Brasil". Em Fiorin, José Luis, & Petter, Margarida (orgs.), *África no Brasil: a formação da língua portuguesa,* 15-62. São Paulo: Contexto.

Cardoso, Susana Alice Marcelino (2010). *Geolinguística: tradição e modernidade*. São Paulo: Parábola.

Cardoso, Susana Alice Marcelin *et alii* (2014). *Atlas Linguístico do Brasil*. Londrina: EDUEL.

Castro, Yeda Pessoa de (2011). "Marcas de africania no português brasileiro", *Africanias.com*, *1*. http://www.africaniasc.uneb.br/pdfs/n_1_2011/ac_01_castro.pdf (Consulta em 19.1.2017).

Castro, Yeda Pessoa de (2012). "Prefácio - Renato Mendonça e 'A influência africana no português do Brasil', um estudo pioneiro de africanias no português brasileiro". Em Mendonça, Renato, *A influência africana no português do Brasil*, 15-28. Brasília:

FUNAG. https://repositories.lib.utexas.edu/bitstream/handle/2152/24719/983-Influencia_Africana_no_Portugues_do_Brasil_A.pdf (Consulta em 19.1.2017).

Gomes, Flávio dos Santos (2015). *Mocambos e Quilombos: uma história do campesinato negro no Brasil*. São Paulo: Claro Enigma.

Guedes, Regis José da Cunha (2012). *Estudo Geossociolinguístico da variação lexical na zona rural do Estado do Pará*. Belém: PPGL / UFPA.

Houaiss, Antônio (2010). *Dicionário Houaiss da Língua Portuguesa*. São Paulo: Objetiva.

Laws, Robert (2012). *An English-Nyanja Dictionary of the Nyanja Language Spoken in British Central Africa*. New York: James Thin.

Lopes, Nei (2012). *Novo dicionário banto do Brasil*. Rio de Janeiro: Pallas.

Lucchesi, Dante, Baxter, Alan, & Ribeiro, Ilza (orgs.) (2009). *O português afro-brasileiro*. Salvador: EDUFBA.

Nascentes, Antenor (1966). *Dicionário etimológico resumido*. Rio de Janeiro: INL / MEC.

Petter, Margarida Maria Taddoni (2005). "Línguas Africanas no Brasil", *Gragoatá, 19*, 193-217. DOI: http://dx.doi.org/10.11606/issn.2526-303X.v0i27-28p63-89 (Consulta em 19.1.2017).

Queiroz, Sônia (1998). *Pé preto no barro branco: a língua dos negros da Tabatinga*. Belo Horizonte: UFMG.

Salles, Vicente (2005) [1971]. *O negro no Pará sob o regime da escravidão*. Belém: IAP.

Vasconcelos, José Leite de (dir.) (1910). "Vocabulário trasmontano", *Revista Lusitana XIII*, 110-126.

Vogt, Carlos, & Fry, Peter (1996). *Cafundó: a África no Brasil*. São Paulo: Companhia das Letras.

Vesgo, zarolho ou *estrábico*? O que dizem os dados do ALiB das regiões Norte e Sul do Brasil

Juliany Fraide Nunes
Aparecida Negri Isquerdo
Universidade Federal de Mato Grosso do Sul

Resumo:

Os estudos dialetais registram a língua em uso em um espaço geográfico e, assim, fornecem dados linguísticos que refletem condicionantes históricos relativos a uma sociedade. Dentre esses dados, situa-se o léxico, acervo vocabular que os falantes utilizam para nomear referentes relacionados ao seu meio, razão pela qual os estudos sobre o léxico acabam por se configurarem como uma rica fonte sobre a cultura, a identidade e as tradições de uma comunidade de falantes. Este trabalho discute resultados de estudo sobre o vocabulário dialetal relacionado à área semântica do *corpo humano* com base em dados inéditos do Projeto ALiB (*Atlas Linguístico do Brasil*). Para tanto, foi selecionada a pergunta 092 do QSL – Questionário Semântico-Lexical/ALiB que apura designações para a 'pessoa que tem os olhos voltados para direções diferentes'. Foram analisadas as respostas fornecidas pelos informantes do Projeto ALiB naturais de 59 localidades do interior que compõem a rede de pontos pertencentes às regiões Norte e Sul do Brasil, coletadas *in loco* junto a 286 informantes. Os dados foram examinados sob duas perspectivas: i) produtividade das variantes segundo fatores espaciais (diatópicos) e sociais (idade, escolaridade...), cotejando os dados das duas regiões; ii) análise semântica das variantes, com destaque para as diferenças em termos da norma regional e para a questão dos tabus linguísticos. Os dados coletados como denominação do conceito selecionado resultou em seis variantes (e respectivas variantes fonéticas): *vesgo*, *zarolho*, *estrábico*, *caolho*, *olho torto*, *birolho*. Dentre os itens lexicais documentados, *zarolho* foi o mais produtivo na região Norte e *vesgo* o mais citado na região Sul, o que, em um primeiro plano, já apresenta indícios de diferenças na norma regional veiculada em cada região.

Palavras-chave:

Atlas linguístico do Brasil, corpo humano, léxico dialetal.

Sumário:

1. Considerações iniciais. 1.1. Fundamentação teórica. 1.2. Pressupostos metodológicos. 2. Análise dos dados. 2.1. Distribuição geográfica das variantes léxicas. 2.1.1. Região Norte. 2.1.2. Região Sul. 2.2. Análise léxico-semântica. 2.2.1. Tabus linguísticos. 3. Considerações finais.

Vesgo, zarolho or estrábico *('cross-eyed')? What ALiB data say about Northern and Southern regions of Brazil*

Abstract:

Dialectal studies record the language in use in a geographic space and thus provide linguistic data that reflects historical conditions for a society. Among these data lies the lexicon, a collection of vocabulary that speakers use to name referents related to their environment, and the reason why studies on lexicon represent a rich source of culture, identity and traditions of a community of speakers. In this paper, the results of a study on dialectal vocabulary related to the semantic field of the human body *based on unpublished AliB Project data (Linguistic Atlas of Brazil) are discussed. Therefore, question 092 QSL - Semantic-Lexical Questionnaire Alib, which seeks designations for the 'person whose eyes are looking in different directions', was selected. The responses provided by 286 informants from 59 interior localities of network points in Northern and Southern Brazil were analysed. The data were examined from two perspectives: i) the productivity of variants based on spatial factors (diatopical) and social (age, education) in both regions; ii) the semantic analysis of variants, highlighting the differences in the regional standard and the topic of linguistic taboos. All the data provided for the select designation produced six variants (and their phonetic variants):* vesgo, zarolho, estrábico, caolho, olho torto *and* birolho. *Among the documented lexical items,* zarolho *(one-eyed) was the most productive in the Northern region and, by contrast,* vesgo *(cross-eyed) was the most referenced in the Southern region, showing signs of different regional standards conveyed in each of the regions.*

Key words:

Linguistic Atlas of Brazil, human body, dialectal lexicon.

Contents:

1. Initial considerations. 1.1. Theoretical basis. 1.2. Methodological assumptions. 2. Data analysis. 2.1. Geographical distribution of lexical variants. 2.1.1. Northern region. 2.1.2. Southern region. 2.2. Lexicon-semantics analysis. 2.2.1. Linguistic taboos. 3. Final considerations.

1. Considerações iniciais

O estudo da língua portuguesa no Brasil precisa considerar a dimensão continental do país, tendo em vista as particularidades que singularizam cada uma das cinco grandes regiões brasileiras: Norte, Nordeste, Sudeste, Sul e Centro-Oeste. As duas regiões geográficas selecionadas para este estudo (Norte e Sul), além de serem localizadas em dois extremos do Brasil, em termos geográficos, também sofreram processos de colonização e povoamento bastante distintos. De um lado, a região Norte que, apesar de concentrar a maior área territorial do país (45% da área geográfica do Brasil) e ser composta por sete estados (Amazonas, Amapá, Acre, Pará, Rondônia, Roraima, Tocantins), configura-se como a região brasileira com menor densidade demográfica (15.864.454 habitantes) e, por outro lado, a região Sul que, em termos geográficos, representa 7% do território nacional e reúne três estados (Paraná, Santa Catarina, Rio Grande do Sul) e uma população média de 27.386.891 habitantes[1].

A região Norte abriga a Amazônia, estado que concentra a maior população indígena do Brasil. Esse território ficou conhecido principalmente pelo comércio do látex da seringueira, que atraiu para a região Norte brasileiros de diferentes regiões do país, sobretudo dos estados do Nordeste, no final do século XIX e início do século XX, nos cognominados 'ciclos da borracha'. Já a região Sul foi povoada principalmente por grupos europeus (italianos, alemães, poloneses, holandeses, dentre outros) que imigraram para o Brasil a partir do século XVIII. Assim, situadas nos dois extremos do território brasileiro, essas regiões se individualizam em termos de formação étnica da população, de características físicas e econômicas e, consequentemente, de traços linguísticos.

Nesse sentido, ao serem comparados dados lexicais recolhidos nessas duas regiões geográficas, tem-se a expectativa de identificar a presença de duas normas distintas, motivadas por fatores geográficos, sociais, históricos e culturais de cada uma delas. Toma-se aqui o termo norma na acepção solidificada por Louis Hjelmslev (1943[2]) e sistematizada e aplicada pelo linguista romeno Eugenio Coseriu (1979): norma como o conjunto de escolhas legitimadas pelo uso, aceitas por um grupo de falantes.

Em um país como o Brasil, devido a sua extensa área territorial, embora haja uma norma da língua portuguesa que aproxime todos os falantes, há, também, realizações regionais que evidenciam particularidades motivadas por fatores extralinguísticos, como a cultura e a geografia, entre outros. Partindo desse pressuposto, este estudo analisa a norma regional a partir de uma amostra do acervo vocabular de falantes das

1 Fonte: Isquerdo, & Teles (2014: 31-77).

2 A versão utilizada neste trabalho é a de 1991, publicada pela editora Perspectiva.

regiões Norte e Sul do Brasil[3]. Para este estudo foram selecionadas as denominações fornecidas pelos informantes do Projeto ALiB para o conceito 'pessoa que tem os olhos voltados para direções diferentes' (Comitê Nacional do Projeto ALiB 2001). Os dados foram coletados *in loco* por meio das entrevistas realizadas pelos pesquisadores do Projeto nas duas regiões estudadas. Ao todo, foram catalogadas 237 ocorrências fornecidas como respostas pelos 236 informantes naturais das 59 localidades selecionadas, 18 da região Norte e 41 da região Sul.

1.1. Fundamentação teórica

A cultura de uma sociedade é manifestada, principalmente, pela linguagem, à medida que, a partir de atos de comunicação do ser humano, é possível depreender condicionantes sociais e geográficos que o rodeiam. Assim, sociedade / cultura / linguagem formam uma tríade capaz de traduzir aspectos das vivências do ser humano dentro de uma comunidade. Nesse sentido, entende-se que a cultura perpassa o homem em todas as suas escolhas e atos, isso acarreta que o estudo de uma sociedade não pode ser independente da estrutura cultural: 'la cultura, por ser contenido más amplio de la conducta humana, es tan importante para el psicólogo como para el investigador de lo social, para el historiador como para el lingüista' (Malinowski, 1948:16).

E uma das formas de estudo dos processos culturais de uma sociedade é por meio do léxico, o conjunto de palavras disponíveis no sistema linguístico do qual o falante se vale para interagir com os demais membros de uma comunidade de fala. Entende-se, pois, que esse nível linguístico é o que melhor evidencia a identidade de um grupo social, uma vez que plasma peculiaridades motivadas por fatores extralinguísticos, como a cultura e a história, por exemplo. Com razão Matoré (1953: 43) defendeu que as palavras são "témoins de l'histoire", "ils sont le reflet d'un état de sociéte".

Nessa perspectiva, o conjunto de palavras que os indivíduos de uma comunidade utilizam representa uma relevante fonte, não só para pesquisas linguísticas, mas também para os estudos das ciências humanas. Sapir, por exemplo, ao relacionar léxico e ambiente, defende que, dentre os níveis da língua, o léxico "é o que mais nitidamente reflete o ambiente físico e social dos falantes. O léxico completo de uma língua pode se considerar, na verdade, como o complexo inventário de todas as ideias, interesses e ocupações que açambarcam a atenção da comunidade" (Sapir, 1969: 45).

3 Projeto "Vocabulário do *corpo humano* nas regiões Norte e Sul do Brasil: perspectiva geossociolinguística" em desenvolvimento por Juliany Fraide Nunes, no Programa de Pós-Graduação em Estudos de Linguagens, sob a orientação de Aparecida Negri Isquerdo.

Ao tomar como pressuposto que o exame do acervo lexical de um grupo de falantes favorece a recuperação de fatores extralinguísticos, o estudioso do léxico também analisa aspectos particulares da língua de cada região, ou seja, os regionalismos veiculados no âmbito de um grupo social. Sob esse ponto de vista, Isquerdo (2003: 166) alerta que, para o estudo da norma linguística no nível lexical, deve-se considerar, sobretudo, a questão da variação, ao classificar um item léxico regionalismo em detrimento de outro. Isso se deve ao fato de a norma lexical de uma comunidade estar imbricada na variação espacial da língua, analisada, por exemplo, a partir de dados dialetais. O exame da norma regional deve, pois, considerar os eixos do espaço e do tempo: "o primeiro, ligado às especificidades regionais [...] e o segundo, relacionado à tendência conservadora da língua, que normalmente se manifesta de forma mais acentuada em regiões menos susceptíveis a influências dos meios de comunicação de massa" (Isquerdo, 2003: 166).

A coleta dos dados lexicais aqui analisados foi orientada pelos pressupostos da Dialetologia, ramo da Linguística que se ocupa do estudo das variedades da língua dentro de um espaço geográfico. Como produto dos estudos dialetológicos, podem-se construir atlas linguísticos, que são constituídos por uma reunião de mapas que distribuem em um espaço geográfico a realização de fatos linguísticos. Mouton (1990: 28) argumenta que a Geografia Linguística se constitui "un complemento ideal, para el dialectólogo que, serviéndose de ella, puede abordar el estudio comparado de las variedades, deducir cómo se relacionan y se estructuran". Assim, um atlas linguístico, além de registrar as escolhas dos falantes de uma região geográfica, fornece dados para estudos contrastivos acerca de diferentes aspectos da estrutura e do funcionamento da língua. Na interpretação dos mapas pode-se, por exemplo, examinar aspectos como *zonas fronteiriças*, *influências de umas línguas sobre outras, revitalização de palavras, focos inovadores, vias de irradiação* (cf. Meya Llopart, 1975).

Mantendo uma estreita relação com a Dialetologia, a Geolinguística fornece o método para a representação de dados dialetológicos por meio de mapas linguísticos. Coseriu (1982) esclarece que a Geolinguística "[...] pressupõe o registro em mapas especiais de um número relativamente elevado de formas linguísticas (fônicas, lexicais ou gramaticais) comprovadas mediante pesquisa direta e unitária numa rede de pontos de um determinado território" (Coseriu, 1982: 79).

Retomando ainda o pensamento de Mouton (1990: 30), acerca da função dos atlas linguísticos, registre-se que, do ponto de vista dialetológico, "el atlas no pretende informar teóricamente, pero sus materiales son fundamentales para actualizar los conocimientos y tienen un claro aprovechamiento didáctico". Nesse sentido a autora, recuperando Coseriu (1981), menciona diferentes vertentes da Dialetologia:

"la dialectología comparada, la dialectología de las isoglosas, la dialectología contrastiva, la dialectología y la mejor base para hacer dialectología estructural" (Coseriu 1981 *apud* Mouton, 1990: 29).

Ainda é preciso considerar que os dados mapeados pelos atlas linguísticos deixam transparecer condicionantes históricos e culturais de uma sociedade. Em suma:

> [...] a história tem mostrado que o interesse pela variedade dialetal ou a preocupação com ela não se constitui tema unicamente da ciência da linguagem, nem se torna, por isso, patrimônio e apanágio dos linguistas. As diferenças dialetais têm sido explicitadas em diversificadas manifestações, em momentos distintos e motivadas por razões as mais variadas e se têm constituído em alvo de interesses políticos, lato sensu, os mais diferenciados (Cardoso, 2010: 27).

Os estudos geolinguísticos, para garantirem a representatividade dos dados mapeados, seguem uma metodologia rígida que, necessariamente, precisa considerar os seguintes elementos: i) a *rede de pontos* deve ser bem traçada para que abranja toda a região estudada; ii) os *informantes* são selecionados segundo o mesmo perfil; iii) o *questionário* deve ser o mesmo para toda a coleta de dados e, por fim, iv) o *inquiridor* deve ter preparação específica para a realização das entrevistas. A partir de um atlas linguístico podem ser estudados todos os níveis da língua: fonético, morfológico, lexical, pragmático, entre outros. Neste estudo, focaliza-se o nível lexical.

1.2. Pressupostos metodológicos

Como já mencionado, este trabalho analisa dados inéditos do Projeto Atlas Linguístico do Brasil que tem como objetivo mais amplo o registro do falar de habitantes de grandes, médios e pequenos centros urbanos, com vistas a traçar um panorama da realidade linguística brasileira. Para tanto, a equipe de pesquisadores realizou inquéritos com 1.100 informantes naturais de 250 localidades distribuídas ao longo do território nacional. A definição desses pontos linguísticos foi pautada em estudo sistemático acerca da realidade brasileira, considerando, dentre outros aspectos, os processos de povoamento e de desenvolvimento socioeconômico de cada área. Já a seleção dos informantes considerou o seguinte perfil: falante nascido na localidade pesquisada e com pais naturais da mesma região linguística; alfabetizado e ter cursado até o último ano do Ensino Fundamental nas localidades do interior. Nas capitais, além dos quatro informantes com escolaridade fundamental, foram entrevistados falantes com Curso Superior completo. Em cada localidade do interior, foram selecionados quatro informantes de duas faixas etárias: Faixa I (18-30 anos) e Faixa II (50-65 anos) e dos sexos masculino e feminino. Nas capitais, além desses quatro informantes, foram realizados inquéritos com mais quatro falantes

com escolaridade universitária, selecionados conforme as mesmas características em termos de idade, sexo e naturalidade[4]. Neste trabalho, como já anunciado, a amostra examinada pertence aos inquéritos efetuados nas localidades do interior de duas regiões brasileiras.

A realização deste estudo obedeceu às seguintes etapas: 1) levantamento dos dados e respectiva distribuição diatópica com indicação do grau de frequência por localidade/região por meio de gráficos; 2) verificação da dicionarização das variantes lexicais catalogadas por meio da consulta a dicionários produzidos em diferentes períodos da história da Língua Portuguesa; 3) análise semântica das variantes.

2. Análise dos dados

2.1. Distribuição geográfica das variantes léxicas

Ao todo, para nomear o conceito expresso na pergunta 092/QSL, foram registradas seis unidades lexicais e respectivas variantes que totalizaram 237 ocorrências. A Tabela 1 apresenta a produtividade dessas unidades lexicais, levando-se em conta todo o universo pesquisado.

Item lexical	Nº de ocorrências	Porcentagem
Vesgo	152	64,15%
Zarolho	54	22,80%
Estrábico	15	6,33%
Caolho	06	2,52%
Olho torto	05	2,10%
Birolho	05	2,10%
Total	**237**	**100%**

Tabela 1. Distribuição percentual das denominações para a 'pessoa que tem os olhos voltados para direções diferentes', em localidades do interior das regiões Norte e Sul do Brasil

Observa-se que, no conjunto geral do universo pesquisado, a disposição dos dados demonstra, em um primeiro momento, a maior incidência de *vesgo* em detrimento das demais, evidenciando, assim, que esse item lexical pertence ao léxico ativo dos falantes das regiões Norte e Sul do Brasil. Em seguida, na segunda posição

4 Cf. Metodologia do Projeto ALiB no site do projeto: https://alib.ufba.br/content/metodologia

em termos percentuais, situa-se a variante lexical *zarolho*, fornecida por grande parcela da população com alta frequência, embora *vesgo* configure-se como o mais produtivo. Na sequência, os itens léxicos documentados foram analisados segundo a região de origem, iniciando-se pela região Norte do Brasil.

2.1.1. Região Norte

No universo das localidades do interior da região Norte houve 75 ocorrências do conjunto de cinco itens léxicos catalogados. O Gráfico 1 contém a produtividade de cada uma dessas unidades léxicas no âmbito de seis[5] estados da região Norte que integram a rede de pontos do ALiB (Amazonas, Pará, Amapá, Acre, Rondônia e Tocantins).

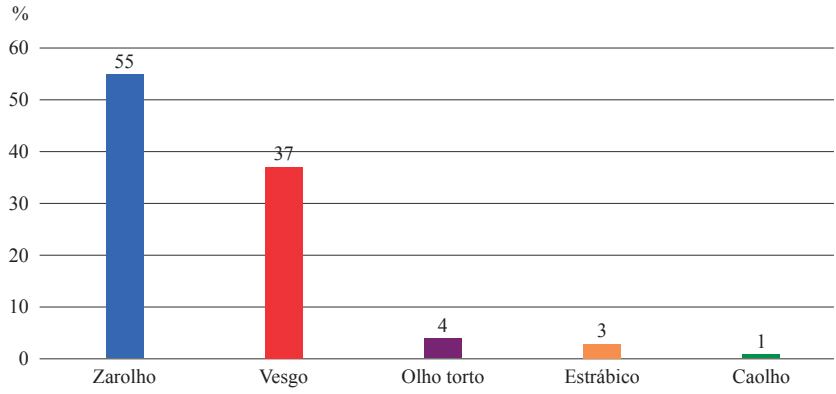


Gráfico 1. Produtividade das denominações para *estrábico* no interior da região Norte do Brasil

O Gráfico 1, que contém a produtividade das variantes no conjunto dos estados da região Norte, evidencia a polarização das variantes *zarolho* e *vesgo* no universo pesquisado, pois as três outras formas concentram baixos níveis de ocorrência. Já o Gráfico 2, também relativo à região Norte, mas com a distribuição dos dados conforme o estado federativo, fornece outro panorama em termos de distribuição da produtividade das variantes lexicais em exame. No estado do Amapá, a variante *vesgo* alçou o maior índice de ocorrência: 60% do conjunto de dados apurados no Oiapoque, única localidade no interior desse estado em oposição às variantes *zarolho* e *olho torto* que tiveram 20% de registros cada uma. Outras duas particularidades foram observadas: no estado do Pará houve empate

5 No estado de Roraima, apenas a capital Boa Vista integra a rede de pontos do ALiB, em razão da baixa densidade demográfica desse Estado da Federação.

técnico entre *zarolho* e *vesgo*, respectivamente, com 45% e 47% de registros, e no Acre onde o item lexical *zarolho* reinou absoluto com 100% de ocorrências em Cruzeiro do Sul. Dentre as demais unidades lexicais documentadas, foi apurada uma ocorrência única de *instalação trocada* em Benjamin Constant (AM), citada pelo informante masculino da segunda faixa etária. A unidade léxica *olho torto*, por sua vez, foi mencionada como resposta por três informantes: Oiapoque (AP): informante masculino, faixa etária I; Óbidos (PA): informante masculino, faixa etária II e Pedro Afonso (TO): informante feminina, faixa etária II, enquanto a forma *caolho* foi citada em Marabá (PA) pelo informante masculino de faixa etária I. No conjunto das localidades do interior da região Norte houve dois casos de ausência de respostas (não-resposta): a informante feminina da primeira faixa etária de São Gabriel da Cachoeira (AM) e a informante feminina da segunda faixa etária de Itaituba (PA) não souberam a resposta ou optaram por não responder a questão.

O Gráfico 2 traz a distribuição diatópica dessas variantes e suas respectivas produtividades, agora considerando individualmente os seis estados da região Norte em que foram documentadas.

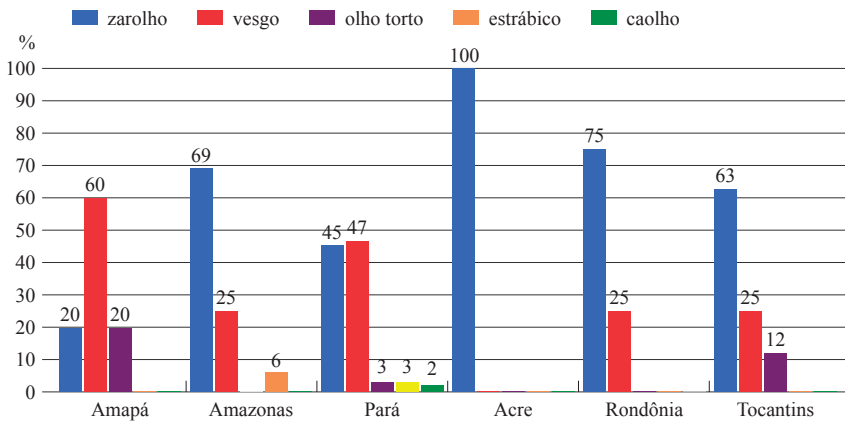
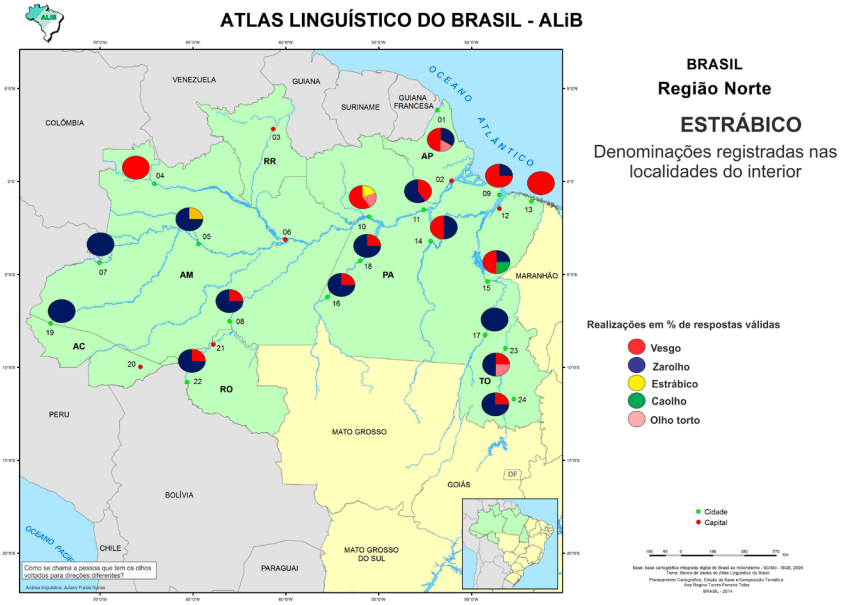


Gráfico 2. Distribuição diatópica das denominações para *estrábico*, em estados da região Norte do Brasil

A Carta Lexical 1, elaborada de forma experimental para este trabalho, contém o mapeamento das unidades léxicas em exame no conjunto das 18 localidades do interior da rede de pontos do Projeto ALiB da região Norte. Traz, portanto, uma visão mais particularizada em relação aos dados apresentados por meio dos gráficos 1 e 2.



Carta 1. Denominações para *estrábico* em localidades do interior da região Norte do Brasil

Observando-se agora a distribuição diatópica e a produtividade da amostra de dados em exame em cada uma das 18 localidades investigadas, nota-se que *vesgo* alçou 100% de produtividade em São Gabriel da Cachoeira (AM) e em Bragança (PA) e foi a forma mais produtiva no Oiapoque (AP), em Soure (PA) e em Óbidos (PA). Em contrapartida, o item lexical *zarolho* obteve 100% de frequência em Benjamin Constant (AM), em Cruzeiro do Sul (AC) e em Conceição do Araguaia (PA). Pode-se depreender, pela disposição dos dados no mapa, que a unidade lexical *vesgo* sobressai-se ao norte e o item léxico *zarolho* ao sul da região Norte do Brasil, com graus distintos de ocorrência. Há casos bem pontuais de realização de três itens lexicais: *estrábico* em Tefé (AM) e em Óbidos (PA); *caolho* somente em Marabá e *olho torto* em Pedro Afonso (TO) e em Óbidos (PA). Verifica-se ainda que em apenas três localidades houve o registro de três variantes distintas: Óbidos (PA): *vesgo*, *olho torto* e *estrábico*; Marabá (PA): *vesgo*, *zarolho* e *caolho*; Pedro Afonso (TO): *vesgo*, *zarolho* e *olho torto*. Esse fenômeno pode ser explicado pelas migrações internas tão frequentes nessa região do Brasil.

Na sequência, são discutidos os dados documentados na região Sul.

2.1.2. Região Sul

Nas 41 localidades do interior da região Sul foram registradas seis designações que somaram um total de 162 ocorrências. O Gráfico 3, a seguir, informa o percentual de ocorrência de cada uma dessas designações no conjunto dos três estados sulistas.

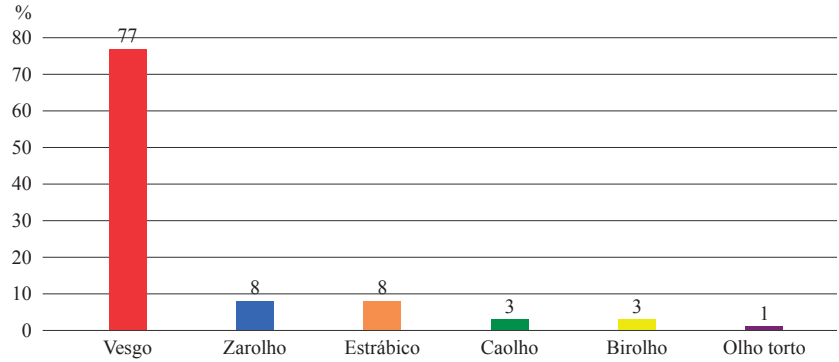


Gráfico 3. Distribuição diatópica das denominações para *estrábico* em estados da região Sul do Brasil

Ao contrário do observado com os dados da região Norte (Gráficos 1 e 2), nota-se, pelo informado no Gráfico 3, a absoluta predominância da unidade lexical *vesgo* entre os falantes sulistas, com 76,54% das respostas. Na sequência situam-se as variantes lexicais *zarolho* e *estrábico*, ambas com o mesmo índice de ocorrência (8,02%); *caolho* e *birolho*, também empatados com 3,08% e, por fim, *olho torto* com 1,23%. Nas localidades do interior do Rio Grande do Sul e em uma de Santa Catarina, houve a ocorrência de *birolho*, unidade não registrada na região Norte. No Gráfico 4, na sequência, são apresentados os dados mencionados, agora distribuídos segundo cada um dos três estados da região Sul.

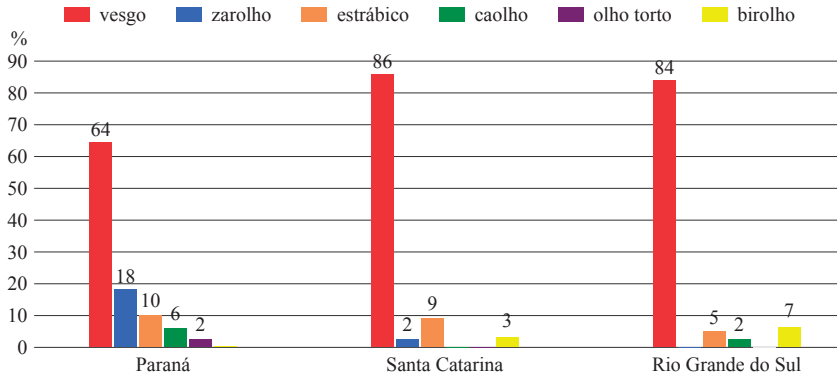
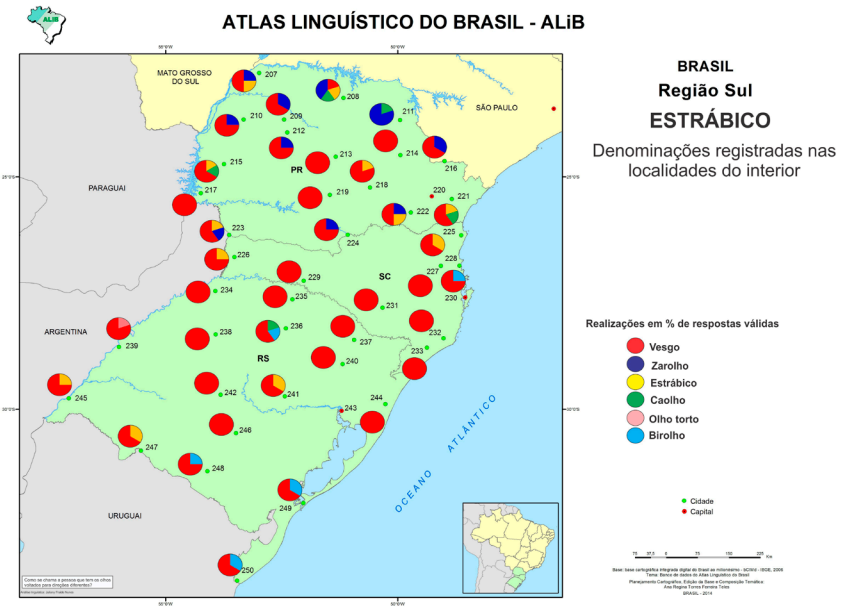


Gráfico 4. Distribuição diatópica das denominações para *estrábico* nos estados da região Sul do Brasil

Observando-se a distribuição percentual dos dados, verifica-se que a unidade lexical *vesgo* continua liderando a produtividade, mesmo com uma leve queda no Paraná em relação a Santa Catarina e ao Rio Grande do Sul. O item léxico *zarolho*, por sua vez, foi documentado em oito localidades do Paraná: Nova Londrina (informante feminina, faixa etária I), Londrina (informante feminina, faixa etária I), Tomazina (todos os informantes da localidade), Umuarama (informante feminina, faixa etária II), Terra Boa (informante masculino, faixa etária II), Campo Mourão (informante feminina, faixa etária I), Lapa e Barracão (informante masculino, faixa etária II) e em uma de Santa Catarina: Porto União (informante masculino, faixa etária I). Já o item lexical *estrábico* foi mencionado nos três estados, com maior grau de ocorrência no Paraná, onde foi citado em sete localidades, embora com baixo índice de ocorrência: Nova Londrina (informante feminina, faixa etária II), Londrina (informante feminina, faixa etária II), Barracão (informante masculino, faixa etária II), Toledo (informante feminina, faixa etária II), Imbituva (informante feminina, faixa etária II), Lapa (informante feminina, faixa etária II) e Morretes (informante masculino, faixa etária I). Em Santa Catarina foi apontado como resposta em duas cidades: São Miguel do Oeste (informante feminina, faixa etária II) e São Francisco do Sul (informante feminina, faixa etária I e II) e no Rio grande do Sul em três pontos da rede: Santa Cruz do Sul (informante feminina, faixa etária I), Uruguaiana (informante feminina, faixa etária II) e Santana do Livramento (informante feminina, faixa etária II). O designativo *caolho*, por seu turno, foi registrado em cinco localidades, quatro no estado do Paraná: Londrina (informante feminina, faixa etária I), Tomazina (informante feminina, faixa etária II), Toledo (informante masculino, faixa etária I), Morretes (informante masculino, faixa etária II) e em uma do Rio Grande do Sul: Passo Fundo (informante feminina, faixa etária II).

A unidade lexical *birolho* foi mencionada apenas na região Sul do Brasil, com cinco ocorrências, quatro delas no Rio Grande do Sul: Passo Fundo (informante masculino, faixa etária II), Bagé (informante masculino, faixa etária II), São José do Norte (informante feminina, faixa etária II) e Chuí (informante masculino, faixa etária II) e uma em Santa Catarina: Itajaí (informante masculino, faixa etária I). Nota-se, a partir do perfil dos informantes que citaram *birolho*, que, com exceção de Itajaí, todos são da segunda faixa etária, o que pode evidenciar um traço de conservadorismo em relação ao uso desse designativo em detrimento dos demais. Por fim, em Chuí (RS), ponto 250 da rede de pontos do ALiB, foi documentada uma ocorrência única da unidade lexical *bisca* (informante feminina, faixa etária II).

Também a Carta Lexical 2 contém a distribuição pontual de ocorrência das variantes em exame segundo a localidade onde foram documentadas, o que permite verificar como se comportam esses itens lexicais em cada uma das 41 localidades da rede de pontos do Projeto ALiB pertencentes à região Sul.



Carta 2. Denominações para *estrábico* no interior da região Sul do Brasil

Como já anteriormente assinalado e pelos dados da Carta 2, nota-se que o item lexical *zarolho* foi registrado apenas no Paraná e em uma localidade de Santa Catarina na fronteira com o Paraná. Logo, esse designativo foi mais produtivo ao norte do Paraná e, consequentemente, no norte da região Sul, enquanto no sul do estado predomina a forma *vesgo*, cuja área de dispersão estende-se para os outros dois estados sulistas (Santa Catarina e Rio Grande do Sul), confirmando, assim, a predominância de *vesgo* na norma lexical dos falantes do sul do Brasil. A unidade lexical *birolho*, por seu turno, merece destaque pela distribuição das ocorrências: quatro localidades (Passo Fundo, Bagé, São José do Norte e Chuí) próximas da fronteira com o Uruguai, o que denota a influência do contato linguístico dos sul-rio-grandenses com a língua espanhola[6]. Constatou-se que *birolho* é de origem galega e se configura como a denominação mais produtiva para a 'pessoa que tem os olhos voltados para direções diferentes', na Carta 120 do *Atlas Lingüístico Galego* (Álvarez *et alii*, 2005: 201). Nesse sentido, pode-se inferir que essa unidade

6 Esses dados confirmam dois dos aspectos que os atlas linguísticos podem revelar assinalados por Meya Llopart (1975): *regiões fronteiriças* e *influências de umas línguas sobre outras*. No caso em pauta, nota-se a influência entre línguas em contato nessa região brasileira que faz fronteira com a Argentina e com o Uruguai, países de língua espanhola.

léxica foi transplantada da Espanha para as colônias espanholas na América e, consequentemente, em virtude de situação de contato linguístico, foi assimilada pelos falantes brasileiros dessa faixa de fronteira.

Fato curioso foi observado em relação ao uso da unidade léxica *estrábico*, que foi apontado por alguns informantes como de uso por estratos específicos na sociedade. Na região Sul, em Adrianópolis (PR), o entrevistado masculino, da segunda faixa etária cita *vesgo* e *estrábico* e, ao ser questionado sobre qual é de uso mais comum, responde: "*vesgo*, *estrábico* nem sei como falar direito, só os ricos falam assim, já os peões de rua usam *vesgo*", enquanto na região Norte, em Altamira (PA), o informante masculino, da segunda faixa etária teceu o seguinte comentário sobre o emprego dessas duas formas lexicais: "Aqui chama *vesgo*, mas ele é *estrábico*, algo assim, não estou lembrado". Esses excertos mostram que a norma lexical também pode demarcar a classe social dos falantes.

A distribuição diatópica das variantes que nomeiam a 'pessoa que tem os olhos voltados para direções diferentes' aponta para distinções entre os falares nortista e sulista, como bem ilustra a distribuição diatópica e a produtividade de uso das unidades lexicais *zarolho* e *vesgo*, com a predominância massiva da primeira na região Norte e da segunda na região Sul. O Gráfico 5 traz informações que permitem o cotejo dos dados apurados nas duas regiões em exame e, consequentemente, a constatação da distribuição diatópica, da produtividade e do grau de dispersão das variantes com maior grau de ocorrência no âmbito dos dois espaços geográficos delimitados para este estudo.

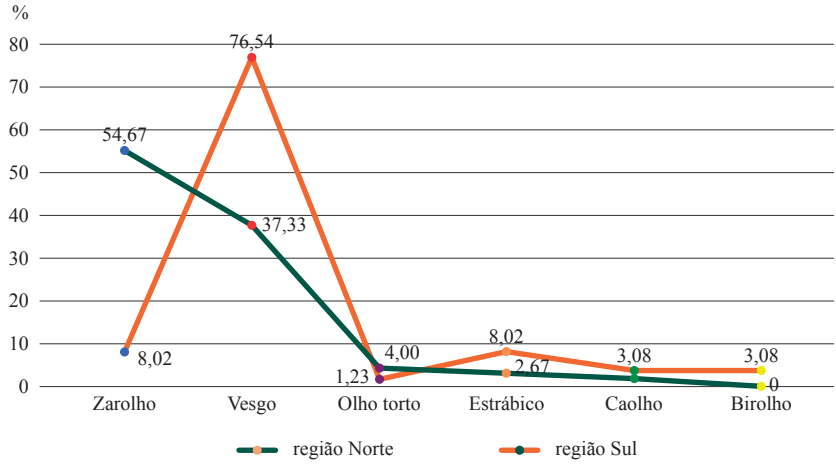


Gráfico 5. Grau de dispersão de denominações para *estrábico* nas regiões Norte e Sul do Brasil

2.2. Análise léxico-semântica

Para fundamentar a análise semântica das unidades lexicais documentadas, foram consultados sistematicamente quatro dicionários da Língua Portuguesa com vistas a verificar a questão da dicionarização das formas em análise: Bluteau (1712-1728); Silva (1813); Houaiss (2001); Ferreira (2004)[7].

As unidades lexicais *vesgo* e *zarolho* já aparecem dicionarizadas nos dicionários de língua portuguesa produzidos nos séculos XVIII e XIX, respectivamente, Bluteau (1712-1728) e Silva (1813). Bluteau (1712-1728) registra os verbetes *vesgo*, na acepção de 'o que mete hum olho por outro... *strabo*' e *zanolho,* como 'falto de hum olho, ou quem os tem atravessados... strabo'. Há também *strabismo,* 'palavra de medico'. O dicionário de Silva (1813), por sua vez, define *strabismo* como 's.m. Cirurg. Má posição do olho dentro da sua órbita'; *vesgo*, como 'adj. Que tem a vista torcida, metendo um olho pelo outro' e *zanòlho,* na acepção de 'zarolho'[8]. Percebe-se que essas acepções têm como ponto de partida os semas atribuídos a *strabo,* termo derivado do Latim que quer dizer 'que eh o vesgo, ou torto dos olhos' (Bluteau, 1712-1728).

Nos dicionários contemporâneos, Houaiss (2001) e Ferreira (2004), as definições dos itens léxicos *vesgo*, *caolho* e *zarolho* não são muito esclarecedoras. Em Ferreira (2004), no verbete *vesgo* só há remissiva para *estrábico*. Já *caolho*, em Houaiss (2001), é definido como 'que ou quem não tem um dos olhos; estrábico' e *zarolho* como 'que não tem um olho, ou é cego de um olho; que sofre de desvio de um ou ambos os olhos; estrábico, vesgo'. Ferreira (2004), por sua vez, define *zarolho* como 'cego de um olho' e remete o usuário para *estrábico*. Já no verbete *caolho*, o dicionarista faz remissiva para *estrábico* e *zarolho*. Esses dados retirados dos dicionários ratificam o registro de Isquerdo (2003: 172) de que "o usuário comum da língua confunde-se ao estabelecer a diferença entre *olho torto* e *caolho* [...], por isso atribui a eles as mesmas designações".

Uma possível etimologia atribuída às formas *caolho* e *zarolho* foi buscada em Cunha (1986), para quem *caolho* deriva do termo 'ka', originado da língua quimbundo do continente africano, que significa 'pequeno', com a palavra de origem latina 'olho'. O termo *caolho* não foi encontrado nos dicionários de Bluteau (1712-1728) e de Silva (1813), o que pode indicar a origem mais recente do termo, em comparação às demais variantes.

---

7 De acordo com a necessidade de complementação e/ou confirmação de informações foram consultadas as obras de Cunha (1986); Machado (1952); Cavero (1977); Corominas (1998); Real Academia Galega (2012).

8 O dicionário faz remissiva para *zarolho*, mas não inclui esse item lexical na nomenclatura da obra.

O item lexical *zarolho*, como indicam Bluteau (1712-1728) e Silva (1813), deriva do termo ZANOLHO que, segundo Machado (1952), procede da junção do item lexical ZANAGA, de origem obscura, com o termo de origem latina OLHO. Já em Houaiss (2001), *zarolho* é apresentado como derivado da junção do vocábulo ZANAGA, que advém de ZAYARCA, oriunda da língua Basca na acepção de atravessado, com o termo de origem latina OLHO.

Em relação à etimologia das unidades lexicais *vesgo* e *estrábico*, observa-se, a partir da consulta aos dicionários selecionados, que *estrábico* é derivado do termo STRABICO, proveniente da palavra STRABO, que tem origem do Latim e significa *vesgo* ou torto dos olhos (Bluteau 1712-1728). Já o termo *vesgo,* segundo Machado (1952), tem origem obscura. Houaiss (2001), todavia, registra que esse item lexical deriva do espanhol *bisgo*/*bizco*. Segundo Corominas (1998), *bizco* "deriva de ESTRÁBICO, 1240. De origen incierto, seguramente de una raíz de creación expresiva que designó primero la ida de 'guiñar el ojo' y luego bizcar'". O dicionário de Cavero (1977) também define *bizco* como 'estrábico, vesgo, zarolho, torto dos olhos, estrabão'. Essa informação trazida pelo dicionário espanhol explica a etimologia registrada em Houaiss (2001). Nesse sentido, quando o informante de Chuí (RS) emprega *bisca* para nomear o *vesgo*, vale-se dos conhecimentos de outra língua em contato com o português na fronteira seca do Brasil com o Uruguai, no caso o espanhol.

As unidades léxicas *birolho* e *mirolho* não constam nos dicionários de Língua Portuguesa consultados. Como já mencionado, *birollo* é uma unidade léxica de origem galega na acepção de 'Que ten a vista torta, cada ollo mirando nunha dirección' (Real Academia Galega 2012), nomeando, assim, a conceito em foco neste estudo. *mirolho* foi aqui considerada uma variante de *birollo*. Já a unidade léxica *olho torto* não está dicionarizada. Em relação a *olho torto*, entende-se que se trata de um termo genérico usado para designar o conceito em causa.

2.2.1. Tabus linguísticos

Outro aspecto que merece atenção em relação aos dados em exame é a relação entre as escolhas lexicais e os meios de substituição de palavras-tabu apontados por Guérios (1979). É sabido que o olho, parte do corpo humano, tradicionalmente carrega consigo muitos tabus arraigados na cultura popular. Segundo Guérios (1979), antigamente, durante a época das caçadas, era proibido pronunciar a palavra *olho*, em virtude da crença de que isso atrairia mau-olhado e, assim, provocaria um período de pouca caça. Desse modo, no decurso da história esse termo foi sendo substituído por outros, com carga semântica menos ofensiva. Por essa razão, no estudo do vocabulário relativo às partes do corpo humano, em especial ao olho, a questão dos tabus linguísticos não pode ser desconsiderada. Neste trabalho, as variantes lexicais documentadas para nomear 'a pessoa que tem os olhos voltados

para direções diferentes' foram examinadas, nessa perspectiva dos tabus linguísticos, segundo os meios de substituição propostos por Guérios (1979) para palavras-tabu, como o mencionado.

Dentre o conjunto de variantes estudadas, o item lexical *estrábico* não se enquadrou nos meios de substituição, pois é considerado o termo técnico para nomear o conceito em causa. Já como designativos disfêmicos, assim classificados por se configurarem como uma "expressão grosseira ou desagradavelmente direta, em vez de outra, indireta ou neutra" (Ferreira, 2004), foram classificadas as denominações *vesgo*, *zarolho* e *caolho*. O item *olho torto*, por sua vez, configura-se como uma expressão genérica, enquanto os itens lexicais *birolho* e *bisca* não foram classificados dentre os meios de substituição, pois a possível origem dessas duas unidades lexicais remete a outra base linguística, razão pela qual, por ora, elas não puderam ser classificadas de acordo com os meios de substituição de palavras-tabus (Guérios, 1979).

3. Considerações finais

O léxico de uma língua reflete os aspectos constituintes de uma sociedade, como a cultura, a história e a ideologia. Por essa razão, os estudos de dados lexicais se tornam uma rica fonte de análise, não só de aspectos linguísticos, como também de traços culturais e históricos do espaço geográfico de uso do recorte lexical em estudo.

Nessa vertente, este trabalho analisou dados lexicais oriundos de pesquisas dialetais que fornecem, sobretudo, uma amostra do falar de cada uma das regiões pesquisadas. Assim, no que diz respeito às regiões Norte e Sul do Brasil, o estudo demonstrou a presença de características próprias de cada um desses espaços geográficos, ao mesmo tempo em que evidenciam semelhanças e individualidades.

O conjunto das respostas fornecidas para designar 'a pessoa que tem os olhos voltados para direções diferentes' (92/QSL/ALiB) resultou em oito variantes. A análise dos dados aqui proposta demonstrou que, enquanto *zarolho* alçou alta produtividade nas localidades da região Norte, *vesgo* representou a maioria das respostas dos informantes da região Sul. O item lexical *estrábico*, por sua vez, considerado o termo técnico para nomear o referente em questão, foi documentado nas duas regiões investigadas, com maior nível de ocorrência na região Sul[9]. As unidades léxicas analisadas, no nível semântico, de acordo com seu uso e sua origem, podem ser classificadas em três categorias: formas cultas – *estrábico*; formas patrimoniais simples – *vesgo*; formas patrimoniais compostas ou derivadas – *caolho*, *zarolho*, *birolho*, *olho torto*.

9 Os estudos de Nunes e Isquerdo (2012) e Nunes (2013) apontaram conclusões similares com dados da região Centro-Oeste do Brasil.

O estudo evidencia também a influência dos contatos linguísticos entre o português e o espanhol na região Sul do Brasil, com o uso de *birolho* em quatro localidades do estado do Rio Grande do Sul situadas na fronteira com o Uruguai e da ocorrência única de *bisca* em Chuí (RS), ponto 250 da rede de pontos do ALiB, também na fronteira com o Uruguai.

Os dados também demonstraram casos de tabus linguísticos materializados na seleção lexical dos grupos investigados como o uso de unidades léxicas disfêmicas apontadas: *vesgo*, *zarolho* e *caolho*. O produto deste estudo mostrou, enfim, aspectos da visão de mundo dos habitantes das localidades pesquisadas, uma vez que apontou como os tabus linguísticos estão arraigados na cultura de uma comunidade, independentemente de fatores sócio-histórico-geográficos.

Os resultados aqui apresentados acrescentam dados ao estudo de Aguilera e Silva (2015), por exemplo, que analisaram unidades lexicais que nomeiam partes do *olho* tendo como *corpus* dois atlas linguísticos, o do Paraná (1994) e o do galego (2005). Em especial, para a pergunta 'a pessoa que tem os olhos voltados para direções diferentes', esse estudo aponta para a alta produtividade de *vesgo* e a pouca frequência de *birolho* no Paraná, aventando a hipótese de um item em vias de desuso nesse estado sulista, tese também aceita neste trabalho, pois os dados do ALiB não mais registram o item lexical *birolho* no Paraná.

Este estudo também dialoga com os resultados alcançados por Juliá Luna (2007; 2011) em pesquisa sobre designações de partes do *olho* em diferentes atlas linguísticos românicos, especificamente as seguintes partes: *ceja*, *párpado*, *pestaña*, *pupila* e *esclerótica*.

Em suma, este trabalho e os estudos aqui mencionados ratificam a importância de investigações acerca das denominações de partes do *corpo humano* que, por sua vez, representam uma rica fonte de análise, nas dimensões diatópica e semântica, com destaque para a valorização dos atlas linguísticos de diferentes domínios como fontes de dados lexicais, já atestada, dentre outros, pelos estudos apontados neste trabalho.

Referências bibliográficas

Aguilera, Vanderci de Andrade, & Silva, Hélen Cristina da (2015). "Convergências lexicais: a área dos olhos no *Atlas Lingüístico Galego* e no *Atlas Lingüístico do Paraná*", *Estudos de Lingüística Galega*, *7*, 5-18. DOI: http://dx.doi.org/10.15304/elg.7.2333 (Consulta em 17.01.2017).

Álvarez, Rosario *et alii* (2005). *Atlas Lingüístico Galego*. *Volume V: léxico. O ser humano (I)*. A Coruña: Fundación Pedro Barrié de la Maza / Instituto da Lingua Galega.

Bluteau, Raphael (1712-1728): *Vocabulario portuguez & latino: aulico, anatomico, architectonico...* Coimbra: Collegio das Artes da Companhia de Jesus, 8 vols. http://dicionarios.bbm.usp.br/pt-br/dicionario/edicao/1. (Consulta em 17.01.2017).

Cardoso, Suzana Alice (2010). *Geolinguística: tradição e modernidade*. São Paulo: Parábola Editorial.

Cavero, David Ortega (1977). *Diccionario Português-Español / Español-Português*. Barcelona: Sopena.

Comitê Nacional do Projeto ALiB (2001). *Atlas linguístico do Brasil*: *questionário*. Londrina: EDUEL.

Corominas, Joan (1998). *Breve dicionário etimológico de la lengua castellana.* Madrid: Gredos.

Coseriu, Eugenio (1979). *Teoria da linguagem e linguística geral* (trad. de Agostinho Dias Carreiro). Rio de Janeiro: Presença; São Paulo: Edusp.

Coseriu, Eugenio (1982). *O homem e sua linguagem*. São Paulo: Universidade de São Paulo.

Cunha, Antônio Geraldo (1986). *Dicionário etimológico Nova Fronteira da Língua Portuguesa.* Rio de Janeiro: Editora Nova Fronteira.

Ferreira, Aurélio Buarque de Holanda (2004). *Novo Dicionário Aurélio da Língua Portuguesa*, Versão 5.0. Curitiba: Editora Positivo.

Guérios, Rosário Farâni Mansur (1979). *Tabus Linguísticos* (2.ª edição). São Paulo: Editora Nacional.

Hjelmslev, Louis (1991). *Ensaios linguísticos*. São Paulo: Perspectiva.

Houaiss, Antônio (2001). *Dicionário eletrônico Houaiss da Língua Portuguesa*. Rio de Janeiro: Objetiva.

Isquerdo, Aparecida Negri (2003). "Léxico em tempo e espaço: a questão dos regionalismos". Em Marin, Jérri Roberto, & Vasconcelos, Cláudio Alves de (orgs.), *História, região e identidades*, 165-181. Campo Grande – Mato Grosso do Sul: Editora da Universidade Federal de Mato Grosso do Sul.

Isquerdo, Aparecida Negri, & Teles, Ana Regina Torres Ferreira (2014). "A rede de pontos". Em Cardoso, Suzana Alice Marcelino *et alii*, *Atlas linguístico do Brasil*, vol. 1, 37-77. Londrina: EDUEL.

Julià Luna, Carolina (2007). *Léxico y variación: las denominaciones de las partes del ojo*. Bellaterra: Universitat Autònoma de Barcelona. Trabajo de investigación. <http://www.recercat.cat/bitstream/handle/2072/4360/Treball%20de%20recerca.pdf?sequence=1> (Consulta em 17.01.2017).

Julià Luna, Carolina (2011). "Procedimientos de creación léxica en las designaciones iberorrománicas del *párpado* (I)", *Revista de Filología Románica, 28*, 49-68. DOI: http://dx.doi.org/10.5209/rev_RFRM.2011.v28.37219 (Consulta em 17.01.2017).

Machado, José Pedro (1952). *Dicionário Etimológico da Língua Portuguesa*, vol. 5 (Q-Z) (1.ª edição). Lisboa: Livros Horizonte.

Malinowski, Bronislaw (1948). *Una teoría científica de la cultura*. Buenos Aires: Editorial Sudamericana.

Matoré, Georges (1953). *La méthode en lexicologie*. Paris: Marcel Didier.

Meya Llopart, Montserrat (1975). "Interpretación del mapa 1204 del ALEA «orzuelo»", *Revista Dialectología y tradiciones populares, 31 (1-4)*, 103-110.

Mouton, Pilar García (1990). "El estudio del léxico en los mapas lingüísticos". Em Moreno Fernández, Francisco (comp.), *Estudios sobre variación lingüística*, 25-75. Alcalá de Henares: Universidade de Alcalá de Henares.

Nunes, Juliany Fraide, & Isquerdo, Aparecida Negri (2012). "Variantes para *vesgo* nas capitais das regiões Centro-Oeste e Norte: contribuições do ALiB". Em Isquerdo, Aparecida Negri, Altino, Fabiane Cristina, & Aguilera, Vanderci de Andrade, *Atlas Linguístico do Brasil: Descrevendo a língua, formando jovens pesquisadores...* Londrina-PR: EDUEL.

Nunes, Juliany Fraide (2013). "Vocabulário do corpo humano: um estudo das designações para *vesgo*". Em XXI Seminário do CELLIP, *Anais do XXI Seminário do Centro de Estudos Linguísticos e Literários do Paraná*, vol 1. Universidade Estadual do Paraná / Instituto Federal do Paraná / Universidade Federal do Paraná, 1581-1594.

Real Academia Galega (2012). *Dicionario da Real Academia Galega*. A Coruña: Real Academia Galega. http://academia.gal/dicionario (Consulta em 17.01.2017).

Sapir, Edward (1969). *Linguística como ciência*. Rio de Janeiro: Editora Livraria Acadêmica.

Silva, Antonio de Moraes (1813). *Diccionario da lingua portugueza*. Lisboa: Typographia Lacerdina, 2 vols. http://dicionarios.bbm.usp.br/pt-br/dicionario/edicao/2 (Consulta em 17.01.2017).

Novos caminhos na morfologia lexical do Português: perspetiva histórica[1]

Maria José Carvalho
Faculdade de Letras da Universidade de Coimbra (CELGA-ILTEC)

Resumo:

Como tem sido reconhecido, os estudos de morfologia lexical do Português pouco têm beneficiado de uma investigação histórica em documentação cronologicamente seriada, particularmente nas etapas mais recuadas da língua. Com este contributo, pretende-se fornecer dados (bem como algumas cronologias) que possam contribuir para os estudos de morfologia lexical do Português assentes numa conceção de língua mais ampla e abrangente, que integre a diacronia na descrição das suas atuais variedades. Com base num *corpus* documental medieval dos fundos do mosteiro de Santa Maria de Alcobaça (IAN/TT-Lisboa), descreve-se a evolução e a produtividade de alguns sufixos, nomeadamente: *-airo/-eiro*, *-deira/-doiro* > *-douro*, *-inho*, *-nte*, *-agem*, *-al*, *-ific-*, *-ondo*, *-oso*, *-ura*, *-issimo*, *-ad-* e *-(d)iça*. Apresentam-se, igualmente, algumas pistas para futuras investigações, particularmente as que dizem respeito à interface morfologia lexical/semântica ou, ainda, morfologia lexical/dialetologia, cognição e sociolinguística.

Palavras-chave:

Morfologia lexical histórica, morfologia lexical e cognição, derivação sufixal, história dos sufixos no Português, lexicalização e formação de palavras.

Sumário:

1. Introdução. 2. Análise do *corpus*. 2.1. O sufixo *-airo/-eiro* no léxico dos ofícios/profissões "medievais". 2.2. O sufixo *-deira/-doiro* > *-douro.* 2.3. Lexicalização do sufixo *-inho.* 2.4. O sufixo *-nte.* 2.5. O sufixo *-agem.* 2.6. O sufixo *-al* nas expressões formulares. 2.7. O sufixo *-ific-*. 2.8. O sufixo *-ondo.* 2.9. O sufixo *-oso.* 2.10. O sufixo *-(d)ura.* 2.11. O sufixo *-íssimo* na linguagem religiosa. 2.12. O sufixo *-ad-*. 2.13. O sufixo *-(d)iça.* 3. Conclusões e pistas para investigações futuras.

1 Este trabalho constitui uma versão refundida de parte de um capítulo ("Operadores sufixais") da tese de doutoramento da autora (Carvalho, 2006: 530-538). Resultou da comunicação "Subsídios para a morfologia derivacional: alguns produtos lexicais sufixados no Português medieval", apresentada a *Gallæcia – III Congresso Internacional de Linguística Histórica. Homenagem aos Professores Ramón Lorenzo e Antón Santamarina (Facultade de Filoloxía, Universidade de Santiago de Compostela, 27-30 de xullo de 2015).*

New paths in Portuguese lexical morphology: a historical perspective

Abstract:

It has been recognised that Portuguese lexical morphology has benefitted very little from the historical study of chronologically arranged documents, particularly in the earliest stages of the language. The present article aims to provide data, together with some chronologies, as a contribution towards a lexical morphology of Portuguese based on a wider and more inclusive idea of language, which includes the diachronic in the description of its present varieties. Based on a medieval corpus of documents from the Monastery of Santa Maria de Alcobaça (IAN/TT-Lisboa), the evolution and productivity of some suffixes are described, especially: -airo/-eiro, -deira/-doiro > -douro, -inho, -nte, -agem, -al, -ific-, -ondo, -oso, -ura, -issimo, -ad- and -(d)iça. *Some suggestions for future research are also presented, particularly in the interface of lexical morphology/semantics and in lexical morphology/dialectology, cognition, and sociolinguistics.*

Key words:

Historical lexical morphology, lexical morphology and cognition, suffixal derivation, history of Portuguese suffixes, lexicalization and word formation.

Contents:

1. Introduction. 2. Analysis of the corpus. 2.1. Suffix -airo/-eiro *in the lexicon of 'medieval' trades/professions. 2.2. Suffix* -deira/-doiro > -douro. *2.3. Lexicalization of suffix* -inho. *2.4. Suffix* -nte. *2.5. Suffix* -agem. *2.6. Suffix* -al *in formulaic expressions. 2.7. Suffix* -ific-. *2.8. Suffix* -ondo. *2.9. Suffix* -oso. *2.10. Suffix* -(d)ura. *2.11. Suffix* -íssimo *in religious discourse. 2.12. Suffix* -ad-. *2.13. Suffix* -(d)iça *3. Conclusions and suggestions for future research.*

1. Introdução

Como tem sido reconhecido, o impacto dos resultados da investigação histórica nos estudos de morfologia lexical do Português ainda não se efetivou completamente, ao contrário do que seria expectável[2]. Iremos centrar-nos, neste estudo, na análise morfo-semântica dos produtos lexicais mais comuns no português medieval, muito particularmente nos aspetos que apresentam variação e que, por esse motivo, divergem ou apresentam semelhanças relativamente ao português atual. Consideramos, de facto, que é esse percurso que nos pode fornecer a chave para entender as motivações e mecanismos da mudança, bem como para interpretar a diversidade atual, embora estejamos conscientes das dificuldades metodológicas que envolvem o estabelecimento da cronologia dos afixos[3]. Com este trabalho, pretende-se fornecer alguns dados, cronologicamente datados, que possam contribuir para uma Morfologia lexical do Português assente numa conceção de língua mais ampla e abrangente, que integre a diacronia, a semântica, a cognição, a dialetologia e a sociolinguística na sua descrição. Conforme já salientou Mário Eduardo Viaro (2010: 179):

> Não levar em conta a história, equivaleria à postura de um zoólogo que pretenda descrever a fisiologia dos animais e que ignore dados da evolução das espécies, imaginando *a priori* que o apêndice cecal humano, por exemplo, por não ter supostamente nenhuma função, seja algo que não precisa ser considerado nessa descrição.

Com base num *corpus* medieval, descreve-se a evolução e a produtividade de alguns sufixos, nomeadamente: *-airo/-eiro*, *-deira/-doiro* > *-douro*, *-inho*, *-nte*, *-agem*, *-al*, *-ific-*, *-ondo*, *-oso*, *-(d)ura*, *-issimo*, *-ad-* e *-(d)iça*[4]. Apresentam-se, igualmente, algumas pistas para futuras investigações, particularmente as que dizem respeito à interface morfologia lexical/semântica, derivação/dialetologia e sociolinguística ou, ainda, derivação e cognição[5].

---

2 O estudo do léxico obriga, indiscutivelmente, a uma visão em que o sincrónico e o diacrónico se entrelaçam. A verdade é que, tal como já referiu Mário Viaro, "chega-se à conclusão que os *corpora* são insuficientes: o dicionário apresenta palavras demais e os textos, palavras de menos. A neologia, fonte de produtividade, evidentemente pode ter uma abonação num dicionário, mas arduamente figura em *corpora* de textos" (Viaro, 2006: 4, pdf.).

3 As dificuldades metodológicas que este tipo de pesquisa encerra foram já equacionadas a propósito de alguns prefixos (Carvalho, 2016: 48, n. 6). O nosso objetivo neste trabalho é, contudo, pôr em relevo as divergências e convergências diacrónicas e dialetais no espaço da língua (galego)-portuguesa.

4 Os sufixos mais produtivos no português medieval revelaram-se, no entanto, *-mento, -ção* e *-ria.* Sobre a sua evolução no português antigo, consulte-se, a esse propósito, Carvalho (2015), artigo elaborado tendo como base o mesmo *corpus* (Carvalho, 2006: 33-287).

5 A importância da diacronia e da semântica foi também salientada por Freitas e Areán-García que, reconhecendo "o valor dos estudos históricos e etimológicos, e verificando-se as datações", concluíram que "não se pode prescindir dos estudos das mais várias áreas, como mínimo: Morfologia Sincrônica,

O *corpus* que serviu de base a esta pesquisa é constituído por 153 documentos notariais originais, por nós transcritos, oriundos dos fundos do mosteiro cisterciense de Santa Maria de Alcobaça, um importante centro na cultura portuguesa medieval. Integra uma coleção de documentos compreendidos entre 1289 e 1565, que fazem parte da coleção *Mosteiro de Alcobaça, 1ª e 2ª incorporações* (IAN/TT), redigidos não apenas no mosteiro, mas também nas áreas periféricas sob sua jurisdição, os chamados "coutos" (Carvalho, 2017).

2. Análise do *corpus*[6]

2.1. O sufixo *-airo/-eiro* no léxico dos ofícios/profissões "medievais"

O sufixo latino *-ARIU, -A, de alta produtividade (adjunto a bases substantivas ou adjetivas), teve como continuador histórico *-eiro*, depois de se registar a atração da semivogal para junto da vogal da sílaba tónica e a assimilação exercida por esta última. Em alguns casos, não se registou a assimilação da vogal da sílaba tónica, daí resultando *-airo*, como em *ffamiliayro* (1307 Alp 13), *ffamyliayro* (1307 Alp 13). Trata-se de um agentivo[7] com pouca vitalidade na área geográfica em estudo, tendo sido substituído por *familiar* 'comensal', 'caseiro': "Johãm Martjz *famjliar*" (1467 Mai 117) ou, no idioleto de outro tabelião, "Johãm Martjz *famaliar*" (1479 MA 124). Pelo contrário, tal como hoje, a língua arcaica conheceu um grande número de produtos agentivos parafraseáveis por "aquele que exerce atividade relacionada com Nome-base"[8], em que a vogal do operador sufixal sofreu assimilação. Em consequência das transformações profundas sofridas pela sociedade, a natureza destes produtos agentivos é, naturalmente, substancialmente diferente da dos agentivos observáveis no português atual. Centrando-se a vida social da Idade Média à volta dos mosteiros, e constituindo a riqueza fundiária a coluna vertebral da economia na sociedade feudal, é natural que os agentivos mais frequentes nesta coleção documental sejam *celeireiro* (43 v.) 'guarda ou administrador do celeiro; o monge que cuida de tudo o que pertence ao sustento

Linguística de *Corpus*, Filologia Românica, Historiografia Linguística, Sociolinguística e Dialetologia, Lexicologia, Toponímia e Terminologia, Estilística, Semântica, dentre outras" (2010: 117).

6 Os exemplos extraídos do *corpus* fazem-se acompanhar de ano, local de redação (em sigla ou acrónimo) e número, dentro da mesma coleção. A terminologia usada ao longo deste estudo histórico-descritivo é a de Rio-Torto (1998).

7 Para a definição de agentivo e sua complexificação, veja-se Viaro (2012: 3, pdf).

8 Tal como refere Boullón Agrelo, em certos casos, é difícil distinguir se estamos perante profissões realmente exercidas ou de verdadeiras alcunhas (1999: 74), embora concordemos com a autora ao considerar que "nos casos en que seguen a estructura *prenome* + *patronímico* + *profesión* o adxectivo profesional designa claramente o oficio" (1999: 75).

dos que vivem nos mosteiros', *porteiro* (22 v.), *rendeiro* (15 v.), *pitanceiro* (7 v.) 'distribuidor dos rendimentos de um convento', *emfermeiro*[9] (5 v.) e *ferreiro* (5 v.). Por outro lado, tendo em conta a estratificação social medieval, um ato notarial era, muitas vezes, solicitado e/ou testemunhado por *caualeiros* (14 v.) e *escudeiros* (21 v.).

Há, no entanto, a menção a nomes de ofícios generalizados na época[10], alguns eventualmente hoje ameaçados de extinção a longo prazo, em virtude de poderem vir a desaparecer as atividades manuais desenvolvidas pelos seus agentes e, portanto, os seus referentes:

> *ffferreiro* (1326 MA 19; 1388 MA 58; 1485 MA 128), *ffferreyro* (1345 MA 33), *fereyro* (1434 SC 91); *teſſoureiro* (1490 MA 131), *teſoureiro* (1336 Alj 26), *teſoureyro* (1321 Alc 17), *theſoureyro* (1300 Alj 8); *carpenteiro* (1429 MA 88), *carpinteiro* (1565 Alc 153); *çapateiro* (1440 MA 96; 1448 Alj 103), *çapateyro* (1307 Alp 13), *capateyro* (1526 Ped 145); *cozinheiro* (1477 MA 121); *oleiro* (1528 MA 147); *pedreiro* (1565 Alc 153), *pedreyro* (1399 MA 65; 1526 ped 145); *molneiro*[11] (1305 Alp 11), *moleiro* (1478 MA 123), etc.

Há ainda que ter em conta que a mudança da paisagem física, humana e social, bem como o modo de vida das populações e o progresso tecnológico, têm inequivocamente um forte impacto na mudança lexical, nomeadamente sob a forma de extinção ou de substituição de alguns produtos derivados por outros[12].

Embora tipologias deste género contenham alguma avaliação subjetiva, uma vez que envolvem o *saber sobre* o grau de mudança da sociedade e a consciencialização que dela têm os seus habitantes, podemos considerar como formas agentivas extintas ou em vias de extinção (pelo menos, em espaço lusitano) as seguintes, presentes no *corpus* em apreço (e com o significado que aí contêm). O português europeu atual conserva ainda reminiscências dialetais de algumas dessas formas:

---

9 Este produto lexical apresenta as seguintes formas gráficas e/ou fonéticas: *emfermeiro* (1416 MA 75 e 1529 MA 148), *emffremeiro* (1522 MA 144), *jmfermeiro* (1528 MA 147) e *jnfermeiro* (1527 MA 146).

10 Encontrámos também o nome *foreiros* (1459 MA 111), apresentando o mesmo processo de formação, embora não correspondendo propriamente a um ofício ou profissão. O caráter agentivo deste nome está associado a transitoriedade, ao contrário da generalidade das profissões que têm, normalmente, um caráter permanente. Estamos conscientes de que "o que é um significado agentivo […] é nebuloso em um recorte sincrónico qualquer" (Viaro, 2012: 9, pdf.).

11 Deverá tratar-se de um moçarabismo de caráter fonético (< *MOLINARIU).

12 Lembremos, a título de exemplo, a forma *carniceiro*, que referiremos mais adiante. De facto, no Português do Brasil existe a forma *açougueiro*, correspondendo à forma *talhante* no Português europeu.

> *ataqueiro* 'fabricante ou vendedor de atacas' (1495 MA 134); *camareiro* (1425 MA 84; 1482 MA 125), *camareyro* (1399 MA 65); *carniçeyro* (1346 SC 34), *carniceyro* (1362 MA 44), *carnjçeiro* (1460 MA 112); *caruoeiro* (1416 MA 78); *cauouq[u]eiro* 'cavador, indivíduo que trabalha em minas ou pedreiras' (1447 Alj 101); *cũqueyro*[13] (1298 Alc 6); *corryeiro* 'o que faz ou vende objetos de couro' (1521 ped 143); *forneiro* (1351 Alv 37); *lleyteyro* (1541 SM 152); *mãpoſteiros*[14] (1430 Cós 89); *obreiros*[15] (1372 MA 47); *odreiro* (1528 MA 147, 2 v.) 'que fabrica odres', *picheleiro*[16] (1482 MA 125); *pregoeiro* (1396 Ped 62), *pregoeyro* (1343 Cós 32; 1346 SC 34; 1434 SC 91); *quinteyro* 'homem encarregado do trabalho de uma quinta' (1345 MA 33); *ſtrabeyro*[17] (1399 MA 65); *tenoeiro* 'tanoeiro; o que faz ou conserta tonéis ou outras vasilhas semelhantes' (por dissimilação) (1428 MA 87); *tonoeiro*[18] 'tanoeiro' (por assimilação) (1436 Alf 93); *tjmtoreiro* (1521 Ped 143), *tyntoreiro* (1509 ped 140); *uaqeyro* 'que guarda/pastoreia gado bovino' (1434 SC 91), *uaqueyro* (1334 Alf 25), *vaqueyro* (1536 SC 150), *vaqueyro* (1536 SC 151), etc.

De facto, na sua maioria, são formas dicionarizadas e potencialmente existentes, embora muitas vezes careçam de informações de natureza dialetal, sociolinguística e cronológica precisas. De acordo com António de Morais e Silva (1992: s. v.), o nome de profissão *cunqueiro* existe ainda no Minho com o significado de 'louceiro', constituindo, portanto, um regionalismo. Em alguns casos, como em *camareiro*, regista-se uma transferência do seu significado originário ('fidalgo que servia na câmara real', 'dignidade da corte pontifícia', etc.) para outros contextos, em virtude das mudanças na estrutura social (hoje, 'pessoa que em hotéis, navios, comboios, etc., arruma os aposentos e está ao serviço dos hóspedes'). Em outros, como em *carniceiro*, alterou-se a sua classe gramatical, e o produto adquiriu outros significados que não são o seu significado primário. Como refere Mário Viaro:

> *Carniceiro* é, no sentido básico da palavra, aquele que descarna o animal ou que vende sua carne, isto é, o açougueiro […]. Mas também é aquele

13 Trata-se de "Pedro Pay(z) *cũqueyro*". O produto lexical *cũqueyro* deriva de *cunca*, 'tigela'.

14 Deverá tratar-se de 'malaposteiros' (por deturpação popular), tendo-se formado, neste caso, um agentivo a partir de uma palavra composta.

15 Trata-se de um agentivo com caráter de transitoriedade e não de permanência, ou seja, não era propriamente um nome de profissão. No português do Brasil, *obreiro* está, muitas vezes (eventualmente metaforicamente), associado à linguagem da Igreja.

16 Trata-se de "Lujs Afom*ſſo picheleiro*". Para além de 'fabricante de pichéis' , também é o nome dado no Norte do país ao canalizador ou soldador.

17 O *estrabo* é o'excremento animal', pelo que tudo leva a crer que o *estrabeiro* fosse a pessoa que tinha a seu cargo a limpeza dos estábulos.

18 A estrutura é, neste caso, Nome + patronímico + profissão + topónimo. Trata-se de "Afom*ſo* Anes *tonoeiro* da Çella".

> que come carne, referindo-se com frequência a um animal […] O uso de *carniceiro* como "que é sanguinário" é uma metonímia criada a partir do sentido composicional primário. Muitas vezes, sentidos derivados obscurecem o sentido composicional ou, quando muito, torna pouco clara a relação entre o significado do sufixo e o significado da base (2012: 6, pdf.).

Mais duvidosa é a forma *landeyro*[19], que, tal como na generalidade dos restantes casos, surge aposta a nome próprio + apelido: "P*edr*o Do*minguiz* landeyro" (1345 MA 33). O contexto leva-nos a acreditar que se trata de um nome de ofício ('o que apanha as landes') ou, eventualmente, 'o proprietário de florestas de landes' mas a verdade é que nos dicionários do português contemporâneo esta forma é apresentada como um adjetivo cujo significado é 'árvore que produz landes' ou ainda 'carvalho ou outras árvores de bolota, quando são muito produtivas'[20].

O mesmo se passa com os nomes próprios sob a forma de alcunhas que, embora contenham em si a noção de agentivo se tornam difíceis de parafrasear: o *Vijntanejro* (1436 Alf 93, 2 v.), por exemplo, cuja palavra-base deverá ter sido *vinte* (confronte-se com *quarenteiro*, que apresentaremos a seguir). Alcunhas de cariz agentiva construídas com o sufixo *-eiro* foram, certamente, bastante frequentes no português medieval, tendo sido, em alguns casos, cristalizadas em topónimos: *Caſal do Figeneiro* (1453 MA 107), *Caſſal do Gayteyro* (1536 SC 150) são alguns exemplos, embora desconheçamos as suas motivações. Como refere Viaro, "encontrar paráfrases que sirvam como uma definição sucinta, funcional e verdadeira do sufixo é um grande desafio. Valendo-nos da nossa intuição, com muita frequência, esquecemo-nos que as palavras não foram criadas hoje, mas em sincronias pretéritas" (Viaro, 2012: 3, pdf.).

Há, por fim, um conjunto de agentivos que existiram na sociedade medieval, baseada na riqueza fundiária e no sistema feudal a ela ligada, mas que já não têm vitalidade atualmente, porque desapareceram os seus referentes:

> *beeſteiros* 'soldado armado de besta e que com ela peleja' (1421 Evo 80, 6 v.); *mateiros* (1391 MA 59), *mateyro mor* (1541 SM 152, 2 v.), *mateyro pequeno* (1541 SM 152, 2 v.), *mateyros* (1346 SC 34; 1541 SM 152); *peſſoeiro*[21] (1375 MA 48), *piſſueiro* 'pisoeiro; pisador de uvas' (1478 MA 123); *quarent[ei]-*

19 Remetemos novamente o leitor para a definição de agentividade, em Viaro (2012: 3, pdf).

20 Quer num caso, quer noutro, torna-se importante atentar na complexa noção de agentividade, tal como já sublinhou Viaro (2012: 3): "Podemos dizer que uma árvore é um agentivo como um ser humano? Tudo depende da nossa postura enquanto pesquisador da língua. Um cientista da linguagem de viés dedutivo não terá dúvida em agrupar o *-eiro* de *cozinheiro* com o *-eira* de *pereira*, pois os verbos empregados nas interpretações ('trabalhar em' e 'produzir') não são muito distintos semanticamente".

21 'Cabeça de um prazo ou casal, cabecel, que recebe as rendas e porções dos seus consortes, para as entregar por junto ao senhorio'.

ro (1479 MA 124); *rraçoeiro* (1405 MA 70), *rraçoeir[os]* (1472 TC 120), *raçoeyros* (1491 Alj 132) 'que distribuem as rações (nas igrejas, por exemplo)'; *rrelegueiros* (1391 MA 59), *rrelegueyro* 'o homem que cobra as rendas dos senhorios' (1467 Mai 117); *ſeſmeyros*[22] (1321 Alc 17, 2 v.); *ſiſeyro* 'cobrador de sisas' (1469 Cel 118); *ſobradeiro*[23] (1465 MA 116).

O sufixo -*eiro* atualmente não goza, na língua portuguesa, da mesma produtividade de séculos anteriores, como já notou Viaro, pois foi suplantado, sobretudo a partir do século XVIII, por -*ista*, o formador de nomes de profissões por excelência na língua atual (2010: 176). Diferente evolução teve o galego onde este sufixo, não só em nomes de profissões e agentivos[24] mas também em outros substantivos denominais expressando variadas noções (coletivos, abundanciais, árvores, continentes, locativos…) é altamente produtivo (González Seoane, 2012: 278). O mesmo acontece, aliás, com adjetivos denominais (relacionais[25], caracterizadores disposicionais…), deverbais caracterizadores (*falar – falangueiro*) e com adjetivos deadjetivais de valor predominantemente aproximativo (*grande – grandeiro*).

Valores locativos[26] do antigo sufixo *-ARIU parecem ter-se cruzado com o de outro sufixo, *-ORIU, tendo-se posteriormente cristalizado em formas que se converteram em toponímicas, como nos exemplos *Refeitorejro* (1455 MA 108) e *Reffectorejro* (1456 MA 109) 'Refeitório'. Igualmente cristalizada em topónimo é *Lagoeiras* (1529 MA 148), um locativo que designa o 'local de grande quantidade de Nome-base', ou, ainda, 'conjunto de N-base'.

2.2. O sufixo *-deira/-doiro,-a > -douro,-a*

O sufixo -*deiro* terá resultado do cruzamento do sufixo *-TO (do particípio latino) com o sufixo *-ARIU, observando-se a sonorização da consoante surda, a atração

22 Almoxarifes que faziam a inspeção das sesmarias, 'datas das terras, casais, pardieiros em ruína e desaproveitados e que os senhorios, depois de avisados, não fazem aproveitar e valer'.

23 Não vem identificada a pessoa. Este nome de ofício vem precedido da estrutura "pagada + SP (circunstancial de lugar) + SP (complemento indireto)": "pagada dentro no d*i*cto nolſo m*oeſtei*ro ao *ſobradeiro*".

24 Na Galiza, é comum o uso de agentivos em -*eiro* como alcunhas apostos a nomes próprios + apelidos.

25 No "léxico disponível do galego", por exemplo, o lema *(pantalóns) vaqueiros* ocorre 168 vezes (Álvarez de la Granja, 2015).

26 Como refere Viaro, "infinitas taxonomias podem emergir, se não forem norteadas e restringidas pelo conhecimento histórico de algumas transformações, fartamente documentadas […]. Soluções rígidas com viés determinista ou teleológico dão azo a excessivas posturas *ad hoc* e, nem sempre se compatibilizam com os dados, distanciando-se da real solução dedutiva desejada pela Linguística, como já o foi por outras ciências" (Viaro, 2012: 10, pdf.).

da vogal para junto da sílaba tónica e, finalmente, a evolução do ditongo (*-dairo* > *-deiro*). A forma feminina *-deira* tem a sua origem no acrescentamento da desinência nominal de género feminino ao sufixo *-deiro*. Rio-Torto & Renca, a propósito dos nomes em *-deira* no português, referem a crescente tendência para a afectação de *-eira* a "denominações de realidades artesanais e não teconologizadas" [sic] (Rio-Torto, & Renca, 2007: 103). Ora, a verdade é que não se trata verdadeiramente de uma tendência crescente mas apenas da sobrevivência da situação morfológica mais antiga, a do século XIV[27], que vai ficando progressivamente submersa, à medida que as sociedades progridem. Em finais daquele século, encontramos o adjetivo deverbal parafraseável por 'que há-de/deverá Vb', apresentando esse mesmo sufixo: *ualedeyra* (1388 MA 57; 1399 MA 65, 2 v.) / *ualledeira* (1388 MA 57), o que faz prever a preferência trecentista pelo resultado do sufixo *-(T)ARIA/-U (> *-deira*), mas a partir do século XV preferiu-se *-doiro*, procedente de *-(T)ORIU[28]. Apresentam-se a seguir os adjectivos colhidos nesta colecção, designativos de qualidade ou estado:

> *[cõ]p[r]idoyros* (1434 SC 91), *[cõ]p[r]idoiro* (1422 MA 82), *[con]p[r]idoiras* (1440 MA 96), *[con]p[r]idoyrro* (1500 MA 136), *conp[r]idoyros* (1437 Ped 94; 1453 MA 107), *conp[r]idoyrro* (1507 MA 139), *cõp[r]idoiro* (1469 Cel 118); *poſſoydoyra* (1300 Alj 8); *valedoira* (1447 Alj 101), *valedoyra* (1484 MA 127); *vymd[oy]ra* (1502 MA 137), *vymd[oy]ro* (1502 MA 137), *vymdoyrra* (1500 MA 136) e *vindoiro* (1522 MA 144).

No segundo quartel do século XVI, o ditongo *oi* do sufixo é substituído por *ou*[29]: *valedoura* (1536 SC 150) e *valedouro* (1536 SC 150). Este sufixo pode igualmente criar substantivos locativos, como nos exemplos seguintes: *log[r]adoyras*[30] (1495 MA 134), *logradoiros* (1521 Ped 143) e *log[r]adoyrros* (1502 MA 137), à semelhança, aliás, do que é possível verificar ainda no galego: *obradoiro*, *lavadoiro*[31]. É apenas nos documentos do segundo quartel do século XVI que se regista a substituição de *oi* por *ou*: *logradouros* (1529 MA 148; 1532 Tur 149) e *saydouros* (1541 Sal 152).

---

27 A ausência de historicidade pode, na nossa opinião, conduzir a um erro metodológico em morfologia derivacional, que é o de levar a uma distorção de perspetiva, falseadora da realidade.

28 De acordo com Ferreiro (1997: 167), os adjetivos formados com o sufixo *-(d)oiro* "continúan semanticamente as formas latinas do participio de futuro en *-turus*".

29 Sobre a evolução da variação *oi* ~ *ou* em Português, veja-se o que se diz em Carvalho (2011).

30 *Logradouro* designa o 'terreno contíguo a uma habitação; terreno público ou pastagem para os gados de uma povoação' e, mais recentemente e, pelo menos em algumas regiões, 'garagem'.

31 Consulte-se https://gl.wikipedia.org/wiki/Sufixo.

2.3. Lexicalização do sufixo *-inho*[32]

Quanto ao sufixo *-inho*, se aceitarmos como válida a informação dada por José Pedro Machado, para quem a forma *afoucinhado* (proveniente de *foucinha*, diminutivo de *fouce*), na sua variante sem prefixo, já se documenta no século XIII (Machado, 1995: s. v. *afoucinhado*)[33], é possível que a emergência deste sufixo em época tão remota[34] possa estar ligada à necessidade de evitar a homonímia com *fouce*, uma vez que a distinção semântica entre *foice* e *foicinha* sobrevive ainda dialetalmente. Com grande vitalidade, ao longo de toda a coleção, *afoucinhado* é um adjetivo relacional parafraseável por "que evoca, que tem x propriedades de Nb"[35] ('diz-se do frango que tem as penas da cauda crescidas, viradas como foice'): [capões] *afoucinhados* (1375 MA 48), *afoucjnhados* (1416 MA 78), *afouçinhad[oσ]* (1479 MA 124), *affouçinhad[os]* (1489 MA 130).

Este produto derivado foi, portanto, gerado a partir de uma base nominal em *-inha* cuja lexicalização se processou, eventualmente, para evitar a homonímia. Por outro lado, e tendo em conta que o operador da lexicalização foi de tipo morfológico (o sufixo *-inho*), este caso aponta para a necessidade de relacionar gramaticalização com lexicalização, aspecto já salientado por Juan Carlos Moreno Cabrera:

> a strong tendency exists for grammaticalization processes to feed lexicalization processes, and that there is a close interaction between both procedures. This interaction is crucial for having a better understanding of the evolutionary dynamics of lexical and grammatical elements (1998: 223).

Para além destes exemplos, outros existem na nossa colecção, de outra natureza. As abonações arroladas deixam supor que um tipo de base em que o sufixo *-inho* foi mais produtivo é aquele que diz respeito aos antropónimos (sejam eles nomes próprios, apelidos ou alcunhas), cujo Nb é, por vezes, um nome de animal. Tal como sublinhou já Boullón Agrelo (1999: 72), do ponto de vista da história da língua, é este o terreno onde mais claramente podemos ver o interesse da onomástica, particularmente para a história do léxico. Ainda que seja impossível

32 Na maior parte do território galego, o sufixo é *-iñ-o* e, na zona mais oriental (nomeadamente no falar galego do ocidente das Astúrias), *-ín*. Veja-se Alonso Núñez (2000: 162). Também Boullón Agrelo considera *-ín* como solução dialectal alternativa, documentando-se dialectalmente em falas do leste de Lugo e no galego das Astúrias e do Bierzo (2007: 30).

33 O contexto recolhido pelo autor é o seguinte: "…e hũu par de Boos Capoes *ffouçynhados* Reçebondos", no *Arquivo Histórico Português*, III, 23.

34 Referindo-se aos problemas que a diacronia enfrenta, exprime-se deste modo Mário Viaro: "Uma data, contudo, raramente é segura: não se pode ignorar o fato de que a palavra surja na fala antes (não raro, séculos antes) de aparecer testemunhada na escrita, mas, com certeza, uma data exclui a ilusão de que ela é antiga ou recente, baseada na intuição" (Viaro, 2006: 8, pdf.).

35 *Afoucinhados* aparece como qualificativo de "capões" ("capões *afoucinhados*").

avaliar o sentido conotativo de muitas formas, dada a inexistência de contexto, trata-se de importantes testemunhos do léxico histórico, uma vez que, pelo seu caráter datado e localizado, as abonações encontradas fornecem-nos um valioso contributo para a cronologia da língua e para uma abordagem dialetal numa perspetiva histórica.

Tal como refere a autora (1999: 88), é difícil averiguar quando o sufixo *-inho* (que, no latim, e ainda na Idade Média, tinha um carácter patronímico, já que aparecia em nomes de crianças formados a partir dos cognomes dos pais) começou a adquirir um sentido diminutivo ou afectivo. O primeiro exemplo de *-inho* remonta, no nosso corpus, a 1377 (*Johãnjnho*) e data de um documento escrito em Alvorninha, o couto mais periférico do mosteiro de Alcobaça:

> *Cauadinho* [Vááſque An*e*s, d*i*cto *Cauadinho*[36]] (1391 MA 59); *Caualinho* [Johã *Caualinho*] (1391 MA 59); *F[er]moſynho* [f*r*ey Johã *Fermoſynho*] (1429 MA 88); *Giũq[ue]irinho*[37] (1391 MA 59); *Johãnjnho* [*Johãnjnho* da Mouta] (1377 Alv 50); *Pasarynho* [J*o*ão R*odr*yguyz *Pasarynho*] (1541 Sal 152); *Perrachjnho* [a cuba do *Perrachjnho*] (1414 Alv 76, 2 v.); *Ratjnho* [*Ratjnho Mouro*] (1423 MA 83); *ſenhorinha* [Afonſo] (1460 MA 112).

Os sufixos *-inho* em nomes próprios são, nos nossos documentos, relativamente tardios, e não deverão pertencer todos à mesma camada linguística. De facto, face a uma matéria que só recentemente começou a ser objeto de atenção, torna-se necessário, como bem notou Boullón Agrelo a propósito do hipocorístico (2007), não só problematizar a delimitação conceptual de nomes com sufixo diminutivo ou afetivo (uma vez que alguns destes produtos não seriam formas "afetivas" dos nomes pessoais) mas igualmente reconhecer as dificuldades metodológicas para o seu estudo. Assim, por exemplo, nomes como *ſenhorinha e Fermoſynho* (este último, apelido de frade), apresentam *-inha*, *-inho* não como sufixo mas por evolução fonética, ou seja, *-inha* e *-inha* não terão sido, nestes casos, verdadeiramente diminutivos mas as formas derivadas de SENIORINA[38] e FORMOSINU, respetivamente. A verdade é que, mesmo nesses casos, parece ter havido uma reabilitação tardia, pois só aparecem a partir do século XV.

O que poderá ter acontecido é que, só a partir de certa altura, as alcunhas deixariam de se circunscrever à intimidade familiar ou a situações informais para se refletirem

36 Boullón Agrelo (2007: 16) refere o aparecimento, na baixa Idade Média, da forma romance *dito*, que vai seguida da alcunha, com uso frequente de sufixos diminutivos, habitualmente a partir de bases léxicas (neste caso, trata-se do adjetivo *cavado*).

37 De realçar a acumulação de dois sufixos: *-eiro* e *-inho* (*junqueirinho* < *junqueiro* < *junco*).

38 Nome próprio apresentado por Boullón Agrelo (1999: 413): *Seniorina* [F, NOM], fem. de *Seniorinus* (10.921 TSobrado I 125).

igualmente nos registos dos atos jurídicos, devido eventualmente à crescente necessidade de expressividade.

Fora do âmbito da onomástica, os exemplos com sufixo *-inho* são ainda mais tardios, o que parece falar a favor de algum tipo de estigmatização pela norma culta portuguesa desse tipo de elemento formativo, pelo menos na linguagem técnica dos notários: *peredinha* (1484 MA 126) e *cafinhas* (1527 MA 146, 4 v.).

Atentemos na explicação dada por José Joaquim Nunes para a origem do sufixo:

> Dando ao sufixo -īnu, donde este evolucionou, sentido que originariamente não tivera, a língua popular a ele principalmente recorreu para a formação de nomes –substantivos ou adjectivos– indicadores de objectos de dimensões limitadas ou de qualidade em grau diminuto (1989: 379-380).

Ora, se cruzarmos esta informação com as abonações aduzidas, e se nos lembrarmos do hipocorístico citado *Perrachjnho* (de *parracho* 'homem atarracado, baixo', por dissimilação vocálica), seremos levados a supor que a atribuição destes nomes às entidades em causa assentou num processo de natureza metafórica, numa série de "viajes de un dominio conceptual a otro, una transferencia de un significado más concreto a otro más abstracto que no aparece en el contexto" (Tornel Sala, 2000: 122). Podemos, eventualmente, ser levados a supor que alguns dos referentes seria de pequena estatura, atendendo ao que é dito sobre o sufixo por José Joaquim Nunes. A questão que se coloca para o leitor de hoje é saber até que ponto *-inho*/*-inha*, pospostos a nomes de animais (alguns deles de grande porte, como *cavalo*) e aplicados a humanos transportam informação depreciativa na sua significação, em casos como *Caualinho, Passarynho* e *Ratjnho*[39]. De facto, a escolha deste operador afixal, aposta a bases deste tipo, torna opaca a relação entre o significante/significado e o referente, uma vez que está em causa um signo morfológico que foi transposto para outra situação de significação[40].

Assim, definindo, como Elizabeth Traugott, o processo de subjectivização como o processo pragmático-semântico onde "meanings become increasingly based in the speaker's subjective belief, state/atttitude toward the proposition" (1995: 31), e relacionando a autora a subjectivização com o processo de gramaticalização,

39 Boullón Agrelo (2007: 17) apresenta também uma alcunha interessante com um nome de animal: *Johan Patiño* (a. 1454 DocDevOurense 2, 61).

40 Será o apelido *Ratjnho* a designação, já no primeiro quartel do século XV, dos habitantes da Beira Interior que se deslocavam para o Centro-litoral do País para cultivar a terra dessas zonas? Por outro lado, surpreende que o outro nome que o segue seja um nome designativo de um habitante de outra zona populacional: *Mouro* ("Ratjnho Mouro"). Ora, de acordo com Boullón Agrelo, o adjetivo *mouro* designa certas características físicas 'de pele ou cabelo escuro' (1999: 68). Normalmente, contrapõe-se a condição de *mouro* 'judeu' à de *galego*, 'cristão velho' (1999: 231).

na medida em que se trata de "a gradient phenomena, whereby forms and constructions that at first express primarily concrete, lexical, and objective meanings come through repeated use in local syntactic contexts to serve increasingly abstract, pragmatic, interpersonal and speaker-based functions" (1995: 32), propomos o inverso para os casos de lexicalização analisados, ao nível da antroponímia. Ou seja, de um fenómeno que deverá ter sido morfológico (e, portanto, gramatical), na sua origem, uma vez que consistiu na adjunção de um constituinte sufixal a uma base, passou-se, gradualmente, através de certas associações mentais e cognitivas, a um processo de lexicalização onomástica[41]. Concluindo, analisando os casos citados, é lícito concluir com Moreno Cabrera que "exactly the same hierarchy that some scholars have proposed for explaining how grammaticalization takes place can be succesfully applied to the opposed and in some sense contradictory process of lexicalization" (1998: 224).

2.4. O sufixo *-nte*

O operador sufixal *-nte*, que no latim servia para formar o particípio presente, gera, ao longo da história da língua, agentivos deverbais parafraseáveis por "que V", na sua grande maioria inaceitáveis no português atual. Em alguns casos, este constituinte sufixal foi substituído por -(*d)or* (*guardante* → *guardador*; *cumprinte* → *cumpridor*); em outros, este tipo de significação deixou de existir, como em *estante*, *consirante*, *estipulante*, tendo sido absorvida pela contida em formas de gerúndio do verbo em causa. Apresentamos, a seguir, as ocorrências registadas com *-nte* sufixal:

> *aceitante* (1490 MA 131); *[cõ]prĩt[e]* (1438 Ped 95), *[cõ]p[r]int[e]* (1434 SC 91); *cõſíjrantef* (1304 Alc 9); *correnteσ* (1479 MA 124); *eſtant[e]* (1422 MA 81; 1484 MA 126), *eſtantes* (1489 MA 130), *eſtamte* (1529 MA 148); *g[u] a[r]dant[e]* (1434 SC 91), *gua[r]dant[e]* (1438 Ped 95); *jſtipulant[e]* (1490 MA 131), *jſtipolant[e]* (1490 MA 131); *moent[e]σ* (1479 MA 124); *poent[e]* (1490 MA 131).

Consideremos os contextos apresentados abaixo. Nos exemplos de 1 a 5, os produtos assim formados podem ser substituídos por formas de gerúndio. Saliente-se que no exemplo (3) a forma de gerúndio *eſtando* é usada com adverbial locativo ("hi") e a forma de particípio presente *eſtant[e]* (nos exemplos 4 e 5) com a de adverbial temporal ("ora" ou "ao t*en*po de ora"):

> 1) nos, ffr*e*y P*edr*o abbade e o conuẽto de Alcobaça, *cõſíjrantef* a bõdade e a ſciẽcia e a aiuda q*ue* recebemos (1304 Alc 9)

41 Essa lexicalização onomástica (que envolve, neste caso, uma mudança semântica de âmbito avaliativo/emotivo) é o resultado de um processo cognitivo de des-referencialização, cuja transferência de significados assenta na semelhança, podendo, por isso, ser comparável à metáfora.

2) p*er*ant*e* el pareceo F*er*nam G*õçal*l*uez* de Porto de Moos e ora *eſtant[e]* em Aljubarrota (1422 MA 81)
3) e out*r*oſy *eſtando* hi Johãm de Figeiredo, *eſtant[e]* ao p*r*eſent*e* no d*i*cto m*o*ſt*eir*o ẽ p*r*eſença de mỹ, t*abali*am (1484 MA 126)
4) tod*os eſtantes* ao t*en*po de ora cõm o dito s*e*nn*h*or ẽm o dito moſt*eyr*o (1489 MA 130)
5) e derã d'emp*ra*zamẽto em fatiosym deste dia p*er*a todo sempre ao bacharell Amdre Lopez, *eſtamte* ora neste most*ei*ro (1529 MA 148)

Consideremos, agora, os seguintes exemplos:

6) q*ue* pagaſſe aa out*r*a par*te [cõ]prĩte* e *guardant[e]* dez mj*l*l rr*eae*s brãcos de pena (1434 SC 91)
7) q*ue* pagaſe a out*r*a p*ar*te *[cõ]print[e]* e *guardante* bjnt*e* mjll rr*eae*s brãcos de pena (1438 Ped 95)
8) e oσ d*ic*toσ moynhoσ *moenteσ* e *correnteσ* (1479 MA 124)
9) prometendo o d*i*to dom abb*a*de e ſop*r*ior e ſan*chr*iſtãoo e monjes e *con*vento ao d*i*to Pero Lobo ao *poente* e *jſtipulante* p*er* ſolene jſtipulaçam (1490 MA 131)
10) ao d*i*to Pero Lobo *aceitante* e *jſtipolante* p*er* ſolene eſtipulaçam (1490 MA 131)

Observando os exemplos acima, em (6) e (7) ("aa outra parte *[cõ]prĩte* e *guardant[e]*" / "a outra parte *[cõ]prĩt[e]* e *guardante*", de 1434 e 1438, respetivamente) e em (9) e (10) ("ao *poente* e *jſtipulante*" / "*aceitante* e *jſtipolante*", de 1490) o sufixo *-nte* forma verdadeiros nomes agentivos deverbais; no exemplo (8) ("moynhoσ *moenteσ* e *correnteσ*", de 1479), forma adjetivos deverbais. Perante a escassez de dados, o que se pode dizer é que o funcionamento 'verbal' do particípio presente (com o sentido do gerúndio atual) parece registar-se ao longo de todo o período medieval (1304-1529), ao passo que os nomes e adjetivos deverbais começam a registar-se no segundo quartel do século XV. Esta constatação poderá conduzir a pensar que se desenvolveu um processo de nominalização (eventualmente, um subtipo de lexicalização) que conviveu com os usos verbais do particípio presente. Assim, a tensão ou rivalidade entre o que é mais gramatical ou mais lexical parece ter sido uma constante na história da língua, para tornar mais rigorosos e objetivos os conceitos que envolviam uma escritura, nomeadamente os relacionados com os seus intervenientes bem como com as suas coordenadas espácio-temporais[42].

42 De acordo com Oliveira (2014: 69), "o particípio latino foi adequado ao paradigma verbal, mas começaram a ser iniciadas nominalizações com o particípio presente ainda nessa língua, estabelecidas definitivamente no português. A sua forma singular *-nte(m)* e plural *-ntes* são provenientes, como as palavras da língua portuguesa em geral, do acusativo latino, e passaram ao português pelo latim culto, devido ao próprio desenvolvimento de sua forma, que, em sua essência, é o acusativo com a queda final do *-m*. No entanto, apesar do particípio presente não estar mais presente no português, as formas verbo-nominais latinas deixaram suas marcas no sistema morfossintático, servindo de ponto

O produto lexical *poente*, com o sentido gramatical de agentivo e traço [+ animado] extinguiu-se totalmente no português atual. Em alguns contextos, sofreu, por metonímia, um processo de lexicalização (cf. o local do *sol poente* → o *poente*), processo esse que foi (tal como os processos de gramaticalização) unidirecional, irreversível. Como é sabido, no francês o particípio presente conservou-se (pelo menos na escrita) como parte da conjugação verbal, por influência latina. No italiano são comuns os nomes em *-nte* (*cantante* e não *cantor*, por exemplo).

2.5. O sufixo *-agem*[43]

De acordo com José Joaquim Nunes,

> do sufixo -TĪCU, que no latim formava adjectivos, nasceu o português *-ádego* que, na antiga língua, junto, como naquela, a substantivos, produziu igualmente adjectivos, que passaram à classe daquele, no sentido de impostos, cargos, sendo depois substituído pelo francês *-agem*, de idêntica proveniência, o qual exprime impostos, aglomeração e acção (Nunes, 1989: 372).

A mais antiga abonação encontrada é o item lexical *linhagem*, de origem provençal (< *linhatge*), que era, no português arcaico, de género masculino. O documento 1291 Alc 2 apresenta duas ocorrências cujos contextos não deixam lugar a ambiguidades: "do uoſo *linagẽ*" e "de uoſo *li[n]agẽ*". As acepções de 'impostos' e 'acção' encontram-se presentes em *portagees* (1321 Alc 17) e *ujagẽes* (1443 Alf 99), mas não se encontra ainda o sentido atual de 'ato ou estado' em *ancorajẽns* (1460 MA 113), que, no *corpus* em análise, é um "nomen quantitatis", parafraseável por 'conjunto de Nb'. De facto, é interessante verificar que *ancoragem* atualmente não tem o sentido concreto que lhe era afetado no século XV ('conjunto de âncoras'), designando antes a operação de V-base ou o processo. Vejam-se os seguintes exemplos, ambos no mesmo documento do *corpus*:

---

de partida para inovações ou conservando, em grau maior ou menor, suas primitivas feições. Por fim, vemos que a passagem do particípio presente a nome se deu definitivamente nas línguas românicas, com graus diversos e algumas exceções". Por outro lado, Fiéis, Lobo e Martins, ao referirem-se ao particípio presente no português antigo, reconhecem a necessidade de "estudar detalhadamente as outras construções em que o particípio presente verbal ocorre, nos dados do português antigo, que não de adjunção adverbial, de forma a percebermos melhor a estrutura dessas construções. Possivelmente, esse estudo ajudará a explicar a motivação para o desenvolvimento do particípio presente verbal no português e quando se pode considerar que ele deixou de existir" (Fiéis, Lobo, & Martins, 2016: 437).

43 Sobre a produtividade deste sufixo (assim como a de *-ádego*) em época medieval, veja-se Caetano (2014: 55-58). Segundo Gonçalves, "a acepção de conjunto, apesar de presente em poucas palavras, se estabeleceu fortemente desde sua origem e está em vigor até os dias atuais. A partir do séc. XIII, o sufixo passa a apresentar várias acepções, as quais todas se repetem no séc. XIV, além da criação de outra subclasse, a REL. TAX" (Gonçalves, s.d.: 13).

1) q*ue* rrepartiſem os mareant*e*s e as carauelas cada q*ue con*pr*i*ſe po*r* ſ*er*ujço d'el Rej e q*ue* outroſſy podya e deuya aue*r* as *ancorajẽns* dos nauyos q*ue* ao p*or*to da d*i*cta villa chegaſem e ancoraſem (1460 MA 113)
2) e outroſſy viſtas as ca*r*tas duas p*er* q*ue* era mãdado q*ue* o almjrant*e* rrepa*r*tiſe as barcas e os omẽns do mar e ouueſe as *ancorajẽns*, as quaes ſſe entendyam [sic] nas teras d'el Rej (1460 MA 113).

No contexto seguinte, surge o produto *apeyragẽ*, com o significado de 'comida para os animais', tal como a palavra-base *apeiro*[44]. A forma derivada é mais tardia, o que denota já, eventualmente, a consciência da associação do sufixo ao sentido de 'grande quantidade' ou 'conjunto de N-base':

3) e posa colher madeira p*er*a o dito casall e *apeyragẽ* e cousas necesarias p*er*a o dito casall (1529 MA 148).

Foi, ainda, encontrado o produto *promajes*, muito provavelmente de origem provençal. Poderá ser parafraseável por "árvores que dão Nb", sendo *proma* o Nb[45]:

4) e ponhã ẽ cad'hũm anno no d*i*c*t*o pum*ar* ſejs eyx*er*tos de bõaas *promajes* (1500 MA 136).

Assim, o sufixo *-agem* poderá, eventualmente, ser um sinónimo de *-eira* em algumas bases substantivas, cujo resultado final é 'nome de árvore de fruto'.

2.6. O sufixo *-al* nas expressões formulares

Inserido em expressões formulares, e, portanto, tipicamente do tecnoleto jurídico, a partir de meados do século XV, começa a surgir um novo produto com o operador sufixal *-al* (< *-ALE) (parafraseável por "em relação com Nb"), a partir de bases substantivas, reforçando, pleonasticamente, uma sequência de adjetivos com essa terminação. De facto, até ao primeiro quartel do século XV, esse tipo de expressões continha apenas o adjectivo *corporal*:

44 Não encontrámos o produto *apeiragem* (ou seu étimo) nos dicionários etimológicos do Português e nos da língua latina consultados. António de Morais e Silva refere, contudo, esta forma com o mesmo significado de *apeiro* (Silva, 1992: s. v. *apeiragem* 'recheio de abegoaria' e *apeiro* 'o mesmo que *apeiragem*'). Em outros dicionários surge com o significado de 'conjunto de apeiros, peças próprias para jungir os bois ao arado', 'animais de trabalho', 'apetrechos de abegoaria' ou, ainda, 'comida para o gado'. No nosso *corpus*, a base *apeiro* pode ter como significado 'peças próprias para jungir os bois ao arado', mas também 'comida para o gado'. No contexto seguinte, apenas a primeira ocorrência não oferece dúvidas quanto ao seu significado: "e q*ue* a madeira p*er*a adubío de ſuas caſas e p*er*a *apeiro* de ſuas lau*r*as aiam p*er* aluara do d*i*cto Senhor dom Abbade. E talhãdo ell*e*s mais da d*i*cta rrama ou madeira p*er*a caſas e p*er*a *apeiros* ſem mãdado" (1391 MA 59).

45 Recordamos que, de acordo com Meyer-Lübke, as formas *praoma* e *prõma* (< prūna) existem num dialeto franco-provençal sobrevivente na Suíça e na região da Sabóia, respectivamente (1935, s. v. "prūna".)

1) meterõ en *corp[or]al* poſſiſſõ (1315 Alj 15)
2) e tomẽ a poſſe e *corporal* poſſiſſom (1388 MA 57)
3) e tomem a poſſe e *co[r]poral* poſſiſſom (1399 MA 65)
4) em poſe e *corporal* poſiſam (1422 MA 81)
5) em poſe e *corporal* poſyſam (1467 Mai 117)

A partir do 3º quartel do século XV os exemplos do *corpus* evidenciam, normalmente, uma sequência de três adjectivos, sendo um deles *real* e o outro *autoal* (de *auto*):

6) de poſe *rreal corporall* e *autoall* poſyſãm (1462 Mai 114)
7) ẽ poſe *rreall autoall corporall* poſyſã (1509 Ped 140)
8) poſe *rreall* e *autoall corporall* poſſyſam (1536 SC 151)
9) "de poſe *corporall rreall autoall* e ẽ *corporall* poſjſãm" (1532 Tur 149)

Trata-se, na nossa opinião, de um neologismo erudito criado na linguagem jurídica para reforçar, pela aliteração de *-al*, a anulação de qualquer argumento por parte do empossado no sentido de inviabilizar o contrato por qualquer vício de forma. O elevado número de significantes criados na formalização de um ato para a expressão de realidades de natureza jurídica, cuja função discursiva é meramente reiterativa, faz parte da mentalidade do redator de um ato medieval, denunciando "un souci de réalisme, un besoin de précision notariale" (Zimmermann, 1989-1990: 317). Como afirma o autor atrás citado, "les rédacteurs d'actes usent d'un language redondant et diffus, pratiquent à profusion l'énumération, la glose, le *pléonasme*, mais aussi –involontairement– *l'allitération*, l'antithèse, voire le lapsus" (1989-1990: 309). Por outro lado, a cronologia do surgimento de *autoall*, como terceiro adjetivo de uma sequência enumerativa, confirmam a constatação do mesmo autor: "la renaissance du droit romain, l'influence des formulaires italiens favorisent la recherche d'une plus grande conceptualisation juridique, hâtent l'évolution des actes notariés vers une plus grande concision" (1989-1990: 334).

2.7. O sufixo *-ific-*

O exemplo sufixal deste tipo mais expressivo do nosso *corpus* é o que constrói produtos verbais que explicitam "uma mutação de um estado inicial em um estado final", consistindo numa "afectação/transferência de Nb, traduzível pela paráfrase 'afectar Nb', 'prover de Nb'" (Rio-Torto, 1998: 120). Referimo-nos a um verbo derivado por sufixação, já existente no latim[46], apresentando uma configuração do tipo: Xb-*ific-a-r*. O sufixo poderá exibir diversos alomorfes: *-efeg-* (porventura o mais antigo), *-iuig-*, *-ifig-* (com alógrafo *-iffig-*), como a seguir se apresenta:

46 Veja-se, por exemplo, o doc. 20 de *Um formulário monástico português medieval*, editado por Saúl António Gomes (1999: 168): "plantetis et fructificetis".

frutefegedeſ (1291 Alc 2), *frotificarõ* (1304 Alc 10), *frutiuiguedeſ* (1291 Alc 3), *ffrotiuigaſſe* (1315 Alj 15), *frutifiguedeſ* (1317 Alc 16), *frutifig[ue] des* (1324 Alc 18), *frutiffiga[r]des* (1324 Alc 18), *froytiuigar* (1337 Alc 27), *froytiuigardes* (1337 Alc 27), *frotifiguedes* (1375 MA 48).

Excecionalmente, a configuração do derivado é Xb-*a*-*r*: *froytedes* (1321 Alc 17), de *froytar*. Importa salientar que a semântica de *frutificar* no português medieval era mais restrita do que a do português atual. Assim, no português actual este produto verbal tem como significado 'dar frutos', enquanto que na época medieval significava 'plantar árvores de fruto' ("meter a fruyto", como também se encontra documentado em 1297 Alc 5). No português atual este produto conheceu, por isso, uma extensão polissémica.

2.8. O sufixo *-ondo*

No português medieval, o sufixo *-ondo* produziu adjectivos deverbais parafraseáveis por 'que está em condições de se Vb':

rrecebõdas (1495 MA 134), *rrecebõdo* (1500 MA 136), *rreçebond[os]* (1489 MA 130), *rreçebond[oσ]* (1479 MA 124; 1485 MA 128), *rrecebondos* (1383 Alj 53; 1408 MA 71), *rreçebonda* (1450 Alv 104; 1453 MA 107; 1482 MA 125; 1484 MA 126), *rrecebonda* (1397 MA 63), *rrecebondas* (1451 MA 105; 1452 MA 106), *rreçebondas* (1453 MA 107; 1460 MA 112; 1502 MA 137; 1507 MA 139) e *rreçebondo* (1426 MA 85; 1477 MA 121; 1484 MA 126), etc.

Este adjetivo (inexistente no português atual), e cujo sufixo apenas se acrescenta à base *receber*, qualifica os nomes *capões*, *frangões*, *galinhas*, *ovos*, mas também (eventualmente por generalização analógica da expressão) *reais de prata*:

1) do*us* (*ca*)pões e do*us* frangaãos e duas duzias d'ouos po*r* dja de Natal, bõos e *rrecebondos* (1383 Alj 53)
2) Out*r*oſſy nos daredes de foro da d*i*cta noſſa vinha em cada hũu a*n*no por dia de Sam Miguel de Setẽbro hũa galinha bõa e *rrecebonda* (1397 MA 63)
3) Out*r*oſſy nos dared*e*s mais em cada hũu a*n*no, como d*i*cto h*e*, sete reaaes de p*r*ata de moeda de Caſtela bõos e *rrecebondos* de juſto peſo de rrenda da d*i*c*t*a noſſa loia por dia de Paſcoa da Surreiçõ (1408 MA 71).

2.9. O sufixo *-oso*

Também foi possível detetar, no português medieval, um outro produto adjetival formado com o sufixo *-oso* (<*-oso) a partir de uma base verbal, parafraseável por 'que V', em que V significa 'ser bastante', 'suficiente'. Embora bastante produtivo

no português medieval, como sufixo caracterizador (apresentando, normalmente, a noção de 'abundância de'), este sufixo não existe atualmente, no Português, aposto a esta base: *auõdoſo* (1379 Alc 51; 1405 MA 70), *auõdoſſo* (1485 MA 128), *auõdoſſa* (1383 Alj 53), *avondoſa* (1405 MA 70) e *auõdoſos* (1328 Alj 21).

No nosso *corpus*, normalmente, este adjetivo vem associado a outro (com conteúdo semântico muito positivo) e carateriza, normalmente, o nome *procurador*: "procurador lidimo *auõdoſo*", "procurador *auõdoſſo* perfecto".

Extinto no português atual, este adjetivo tem, curiosamente, grande vitalidade no galego[47].

2.10. O sufixo *-(d)ura*

Com o operador *-ura* existiu o "nomen qualitatis" deadjectival *falſura* (1329 Evo 22), nome abstrato completamente estranho ao português atual, que substituiu esse sufixo (em certos produtos) por *-dade*: *falsidade*, por exemplo.

Já o sufixo *-(d)ura* (< *-(T)URA) serviu para formar nomes de ação, parafraseáveis por 'ação de V-b' ou 'resultado de V-b', neste último caso podendo identificar-se igualmente com a noção de 'conjunto' (*pregadura*, por exemplo):

> *eſ[cri]ptura* (1500 MA 136); *f[e]c[t]ura* (1410 MA 73), *feitura* (1402 MA 67; 1460 MA 112; 1478 MA 122; 1478 MA 123; 1490 MA 131; 1495 MA 134; 1509 Ped 140; 1521 Ped 143; 1528 MA 147); *fejt[ur]a* (1397 MA 64), *feytura* (1433 Ped 90; 1453 MA 107; 1485 MA 128; 1507 MA 131; 1529 MA 148), *ffeyturra* (1489 MA 130); *p[re]gadura* (1363 MA 45; 1388 MA 58; 1399 MA 66; 1413 MA 75; 1423 MA 83; 1452 MA 106; 1453 MA 107; 1465 MA 116; 1478 MA 123; 1495 MA 134; 1507 MA 131), *pregadura* (1362 MA 43; 1386 MA 56; 1453 MA 107; 1479 MA 124), *p[re]guadura* (1416 MA 78); *semeadura* (1429 MA 88; 1541 Sal 152), *ſemeadura* (1372 MA 47; 1429 MA 88; 1502 MA 137), *ſemeadurra* (1372 MA 47), *ſſemeadura* (1416 MA 78).

Excecionalmente, encontra-se *dereituras* (1459 MA 111), com o significado de 'conjunto de N-b', em que N é *dereito* 'direito'.

2.11. O sufixo *-íssimo* na linguagem religiosa

Um outro sufixo presente no nosso *corpus*, a partir do último quartel do século XV, é o que expressa a ocorrência de uma propriedade em grau superlativo.

47 Veja-se o *Dicionario de dicionarios* (s. v. *abondoso*). Segundo Lorenzo (1968: s. v. *abondoso*), este adjetivo documenta-se nas *Cantigas de Santa Maria* (*avondosos*) e, ainda, em Gil Vicente (*abondosa*). Segundo o mesmo autor (1977: s. v. *auondo*), *avondo* (advérbio) é muito frequente na produção cronística galega e ainda sobrevive no Alentejo, no Algarve e no galego atual.

Na generalidade dos casos, aplica-se a nomes de dignidade ou títulos honoríficos[48]. Apresentamos a seguir os exemplos recolhidos:

> *rreuerẽdiſſîmo* [*rreuerẽdiſſîmo* S*enh*or dom Jorje, p*er* mer*ç*ee de D*euσ* ca*r*deall ẽ Roma] (1478 MA 122), *Reuerẽdiſymo* [*Reuerẽdiſymo* em *Chriſt*o padre S*enh*or dom Jorje] (1478 MA 123), *rreuerendiſſîmo* [*rreuerendiſſîmo* em *Chriſt*o padre S*enh*or dom Jorje] (1479 MA 124), *rreuerẽdiſymo* [*rreuerẽdiſymo* ẽ *Chriſt*o padre S*enh*or dom abb*a*de de Çiſtell] (1484 MA 126), *Reuerẽdyſſymo* [*Reuerẽdyſſymo* cardeall em a vylla de Porto de Moos] (1491 Alj 133); *Jluſtriſîmo* [*Jluſtriſîmo* p*r*incipe e ſ*enh*or o cardeall Jnffante de Purtugall] (1526 Ped 145).

Os exemplos colhidos neste *corpus* confirmam, de facto, o que a propósito desta cronologia afirmam Hernán Urrutia Cárdenas e Manuela Álvarez Álvarez:

> El uso de tales superlativos se incrementó durante el Renacimiento, por influencia culta, para decaer en siglos posteriores. Es sobre todo a partir de finales del siglo pasado cuando su empleo se hace abundante" (Urrutia Cárdenas, & Álvarez Álvarez, 1988: 82).

A mesma ideia se encontra em Augusto Epiphanio da Silva Dias, para quem o emprego "dos superlativos em -*issimo*, como superlativos exclusivos, é latinismo raro" (Dias, 1970: 172).

2.12. O sufixo *-ad-*[49]

O sufixo -*ad*- produziu, no português medieval, interessantes unidades lexicais. Uma delas, inexistente no português atual, é *cubadas* (1343 AM 31), um nome denominal que, a avaliar pelo contexto, poderá ser sinónimo do N-base ou apresentar um conteúdo semântico correspondente a 'conjunto', 'grande quantidade':

> 1) "jt*en* mãdo q*ue* os vi[n]hos q*ue* ora téémos en *cubadas* q*ue* os vendan e pagen aqe*l*lo q*ue* deuem*os*".

Em outros casos, com base de género feminino, o operador sufixal -*ad*- produziu uma mudança de género na forma derivada sem, contudo, ter havido qualquer alteração categorial ou de significado. A avaliar pela única ocorrência do *corpus*, já tardia, -*ado* não parece ter acrescentado qualquer valor ao nome base, constituindo, portanto, um sinónimo: *baleado* significa, assim, 'baleia' e não 'conjunto de baleias', no nosso *corpus*:

48 Constitui excepção a forma *antjq[u]ifyma* (1515 SM 141).

49 Para uma análise exaustiva deste setor derivacional que não envolve alteração categorial entre base e derivado, veja-se Rio-Torto (1987: 105-107).

2) "jumto donde ſayo hũu *baleado*, no termo de ſam Martjnho" (1515 SM 141).

De facto, a forma *balea*, com o mesmo referente, ocorre seis vezes ao longo do documento.

2.13. O sufixo *-(d)iça*

Com o sufixo *-(d)iça* (< *-(T)ICIA) formou-se, no português antigo, sobre uma base verbal, o adjectivo *llauradyças* ("ate hir dar nas teras *llauradyças*", 1541 Sal 152, 3 v.), que poderá ser parafraseável por 'estar em condições de Vb' (= 'arável'). Embora não integrando o léxico da linguagem comum[50], este produto tem ainda alguma vitalidade dialectal.

3. Conclusões e pistas para investigações futuras

Tentámos, ao longo deste trabalho, fornecer alguns dados que possam contribuir para integrar a dimensão histórica das línguas nos estudos de morfologia lexical. De facto, negligenciar as exceções ou os sufixos pouco produtivos no português contemporâneo (e que, em muitos casos, sobrevivem dialetalmente ou em outros espaços que o português partilhou com outras línguas, nomeadamente o galego) é escamotear uma dimensão importante das línguas que é a sua *historicidade*. Como já sublinhou Mário Viaro, "não é preciso insistir mais sobre quão indesejáveis são essas soluções *ad hoc* que tentam excluir dados diacrônicos, rejeitando significados que não são freqüentes segundo a usualidade" (Viaro, 2006: 9, pdf.).

Seria interessante analisar o percurso histórico-espacial do sufixo *-eiro*, gerador de agentivos, integrando-o na conjuntura sociocultural em que vigorou, bem como pesquisar as motivações da direcionalidade da mudança no espaço português e galego. Uma tipologia deste agentivo (ligado a ofícios ou profissões), elaborada de acordo com critérios de natureza sociolinguística e baseada na metodologia da Dialetologia percetual, tornaria mais rigorosas as informações de tipo lexicográfico, e evidenciaria a perceção e atitude valorativa dos falantes relativamente à sua vitalidade, o que poderia trazer informação relevante do ponto de vista sociológico. Seria igualmente útil uma pesquisa que averiguasse as motivações históricas para os atuais usos metonímicos e metafóricos deste tipo de agentivo.

50 Têm atualmente bastante vitalidade adjectivos com outro tipo de base, como *estaladiço* e *irritadiço*, cuja paráfrase ('que Vb facilmente') mantém afinidades com *llauradyça,* embora não contendo a noção de 'telicidade'.

A substituição de alguns sufixos por outros sinónimos está, inquestionavelmente, relacionada com toda a estrutura e reorganização do léxico (muitas vezes, associado às combinações fonemáticas)[51] ao longo da história da língua (cf. *-deira* > *-doiro* > *-douro*). A análise de outros requer, por outro lado, não só uma problematização em torno das dificuldades metodológicas do seu estudo como um enquadramento teórico que integre a análise das motivações dos processos de lexicalização (o sufixo *-inho*, por exemplo), que são, naturalmente, de natureza histórica e cognitiva. A análise das ocorrências dos produtos em *-nte* exigiria um tipo de pesquisa sobre os usos do particípio presente ao longo do período medieval que desse conta da emergência de eventuais processos de lexicalização a este fenómeno associados bem como das suas motivações e/ou consequências sintático-semânticas. Por outro lado, a variação entre o agentivo com sentido de gerúndio e os outros tipos de agentivo em *-nte* (não comutáveis com esse valor verbal) poderá estar associado à semântica da palavra-base.

Em alguns casos, há sufixos que são atualmente mais usados em contextos de grande informalidade (até humor) e cujo uso está, portanto, condicionado por fatores de ordem pragmática (*-deira, -doiro*, *-diço*, por exemplo)[52]. Alguns sufixos, como *-al*, permitiram criar neologismos no tecnoleto jurídico (sobretudo a partir de meados do século XV), surgindo associados ao renascimento do Direito romano e à busca de uma maior conceptualização jurídica. Outros ainda (como *-issimo*) foram particularmente produtivos em outras fases da língua, em variedades específicas, nomeadamente na linguagem religiosa, onde foi usado profusamente nos nomes de dignidade ou títulos honoríficos. Seria, igualmente, interessante, investigar, em outros *corpora*, como e quando se efetuou a extensão polissémica de alguns sufixos que geravam, no português medieval, produtos lexicais associados a um significado da esfera do concreto e que vieram, com o decorrer do tempo, a conferir a esses produtos um significado mais abstrato (*-ific-*, *-agem*, por exemplo).

Em suma, a produtividade sufixal bem como a transformação semântica dos sentidos vinculados aos sufixos requer, de facto, "pressupostos claros de semântica histórica", tal como já assinalou Viaro (2010: 180). Alargar este tipo de análise a um conjunto significativo de produtos lexicais em *corpora* diversificados constitui um desafio para a Morfologia, pois só um estudo de semântica sócio-histórica, aliado à cognição e à perceção poderá esclarecer a trajetória dos sufixos em certos produtos e justificar a sua vitalidade no (galego)-português, quer na(s) sua(s) modalidade(s) normativa(s), quer nas suas variedades diastrásticas e dialetais.

51 Consulte-se o artigo de Carvalho (2011: 353-364).

52 Em alguns casos, o uso desses sufixos está associado a questões de género, mais concretamente a uma certa visão e mundividência social (*trabalhadeira* mas não **trabalhadeiro*; *casadoira* mas, pouco aceitável, **casadoiro*).

Referências bibliográficas

Alonso Núñez, Aquilino S. (2000). "Os sufixos nominais diminutivos do galego actual", *Verba, 27*, 133-174.

Álvarez de la Granja, María (2015). "A edición dos materiais no léxico dispoñible do galego". Comunicação apresentada em *Gallaecia – III Congresso Internacional de Linguística Histórica. Universidade de Santiago de Compostela, 27-30 xullo 2015*. http://dspace.usc.es/bitstream/10347/13558/1/MariaAlvarezdelaGranja_Gallaecia_2015.pdf.

Boullón Agrelo, Ana Isabel (1999). *Antroponimia medieval galega (ss. VIII-XII)*. Tübingen: Max Niemeyer Verlag.

Boullón Agrelo, Ana Isabel (2007). "Farruco, Paco, Fran: datos históricos e evolución dos hipocorísticos en Galicia", *Revista Galega de Filoloxía, 8*, 11-56. http://ruc.udc.es/dspace/bitstream/handle/2183/7579/RGF_8_art_1.pdf?sequence=1&isAllowed=y.

Caetano, Maria do Céu (2014). "Herança sufixal: os sufixos *-ático*, *-ádig-*, (*-ádeg-*) e *-agem*", *Revista Galega de Filoloxía, 15*, 53-68. http://illa.udc.es/rgf/pdf/RGF_15_enteiro.pdf (Consultado em 7.12.2015).

Carvalho, Maria José (2011). "From etymology to cognitive sociolinguistics. A case-study from European Portuguese". Em Silva, Augusto Soares da, Torres, Amadeu, & Gonçalves, Miguel (orgs.), *Línguas pluricêntricas. Variação linguística e dimensões sociocognitivas/Pluricentric languages. Linguistic variation and sociocognitive dimensions*, 353-364. Braga: Universidade Católica Portuguesa.

Carvalho, Maria José (2015). "Linguagem e Justiça: Polissemia, *desambiguidade*, e produtividade sufixal nos textos jurídicos, ao longo dos tempos", *Biblos, 3ª série, 1*, 367-394. DOI: http://dx.doi.org/10.14195/0870-4112_3-1_15.

Carvalho, Maria José (2016). "Prefixes in the history of Portuguese: a cognitive and semantic approach", *Estudos de Lingüística Galega, 8,* 45-67. DOI http://dx.doi.org/10.15304/elg.8.2608.

Carvalho, Maria José (2017). *Documentação medieval do mosteiro de Santa Maria de Alcobaça. Edição, com introdução e notas de aparato crítico.* Anexo 74 de *Verba*. Santiago de Compostela: Universidade de Santiago de Compostela. DOI: http://dx.doi.org/10.15304/9788416533817.

Dias, Augusto Epiphánio da Silva (1970). *Syntaxe histórica portuguesa* (5.ª edição). Lisboa: Livraria Clássica Editora.

Dicionário da Língua Portuguesa. Acordo ortográfico: o antes *e o* depois. Edição revista e atualizada. Porto: Porto Editora, 2011.

Dicionario de dicionarios = Antón Santamarina (ed.). *Dicionario de dicionarios da lingua galega*. http://sli.uvigo.es/ddd/ (Consultado em 4.12.2015).

Dicionário Priberam da Língua Portuguesa (DPLP). Disponível em: https://www.priberam.pt/DLPO/

Ferreiro, Manuel (1997). *Gramática histórica galega, II. Lexicoloxía*. Santiago de Compostela: Laiovento.

Fiéis, Alexandra, Lobo, Maria, & Martins, Cláudia (2016). "O particípio presente no português antigo", *Revista da Associação Portuguesa de Linguística, 1*, 417-440. http://ojs.letras.up.pt/index.php/APL/article/view/1598/1421.

Freitas, Érica Santos Soares de, & Areán-García, Nilsa (2010). "Morfologia histórica do Português", *Revista Philologus*, 16 (47), 109-118. http://www.usp.br/gmhp/publ/FreAre1.pdf (Consultado em 7.12.2015).

Gomes, Saúl António (1999). *Um formulário monástico português medieval: o manuscrito alcobacense 47 da BNL*, *Hvmanitas, 51*, 141-184.

González Seoane, Ernesto X. (2012). "O sufixo -*eiro* en galego e no portugués brasileiro. Propostas para un estudo contrastivo". Em Oliveira, Maria Lúcia da Cunha Victorio de, & Rauber, André Luiz (orgs.), *Castilho – Anais de resumos do II Congresso Internacional de Linguística Histórica. Homenagem a Ataliba Teixeira de Castilho. Universidade de S. Paulo, 7-10 de fevereiro de 2012.* São Paulo: USP, http://cilh.fflch.usp.br/sites/cilh.fflch.usp.br/files/cad_new.pdf.

Infopédia. Dicionário Porto Editora (Disponível em: https://www.infopedia.pt/dicionarios/lingua-portuguesa).

Lorenzo, Ramón (1968). *Sobre a cronologia do vocabulário Galego-Português (Anotações ao 'Dicionário etimológico' de José Pedro Machado)*. Vigo: Galaxia.

Lorenzo, Ramón (1977). *La traducción gallega de la Crónica General y de la Crónica de Castillla.* Edición crítica anotada, con introducción, índice onomástico y glosario, vol. II (Glosario). Ourense: Instituto de Estudios Orensanos "Padre Feijoo".

Machado, José Pedro (1995). *Dicionário etimológico da língua portuguesa* (7.ª edição), 5 vols. Lisboa: Livros Horizonte.

Meyer-Lübke, Wilhelm (1935). *Romanisches Etymologisches Wörterbuch* (3.ª edição). Heidelberg: Carl Winter Universitätsverlag.

Moreno Cabrera, Juan Carlos (1998). "On the relationship between grammaticalization and lexicalization". Em Ramat, Anna Giacalone, & Hopper, Paul J. (eds.), *The limits of grammaticalization*, 211-227. Amsterdam / Philadelphia: John Benjamins. DOI: 10.1075/tsl.37.10mor.

Nunes, José Joaquim (1989[9]). *Compêndio de gramática histórica portuguesa (Fonética e Morfologia)*. Lisboa: Clássica Editora.

Oliveira, Anielle Aparecida Gomes Gonçalves Jacometti de (2014). *Os sufixos agentivos* -nte *e* -(/t/s)or *no português: um estudo semântico-diacrónico.* Tese apresentada como requisito parcial para obtenção do grau de Doutor. Universidade de São Paulo. www.teses.usp.br/…/2014_AnielleAparecidaGomesGoncalvesJacomettiDeO. (Consultada em 18.05.2015).

Oliveira, Anielle Aparecida Gomes Gonçalves Jacometti de (s..d.). *Análise diacrônica da semântica do sufixo* -agem *em palavras formadas no século XII, XIII e XIV*. http://www.usp.br/gmhp/publ/GonA1.pdf (Consultado em 7.12.2015).

Rio-Torto, Graça (1987). "Morfologia das palavras construídas em *-ad(-a)*", *Biblos, 63*, 97-178.

Rio-Torto, Graça (1998). *Morfologia derivacional. Teoria e aplicação ao Português*. Porto: Porto Editora.

Rio-Torto, Graça, & Renca, Nuno (2007). "Nomes em *-deira* no português", *Filologia e Linguística Portuguesa, 9*, 103-116. http://www.revistas.usp.br/flp/article/view/59775/62884 (Consultado em 7.12.2015).

Silva, António de Morais (1992). *Novo dicionário compacto da língua portuguesa*. Edição compacta do texto fundamental do *Grande dicionário da língua portuguesa*, *segundo a 10ª Edição Revista, muito aumentada e actualizada conforme as regras do acordo ortográfico luso-brasileiro de 10 de Agosto de 1945, por Augusto Moreno, Cardoso Júnior e José Pedro Machado, aliviada de etimologia, formas verbais e abonações, constituindo o maior repertório de vocábulos da língua falada e escrita em Portugal e no Brasil* (7.ª edição), 5 vols. Mem Martins: Editorial Confluência.

Tornel Sala, José Luís (2000). "Gramaticalización y cognición: pautas definitorias e interrelación", *Letras de Deusto, 30 (88)*, 111-140.

Traugott, Elizabeth C. (1995). *Subjectification in grammaticalisation*. Em Stein, Dieter, & Wright, Susan (eds.), *Subjectivity and subjectivisation. Linguistic perspectives*, 31-54. Cambridge: Cambridge University Press. DOI: http://dx.doi.org/10.1017/CBO9780511554469.003.

Urrutia Cárdenas, Hernán, & Álvarez Álvarez, Manuela (1988[2]). *Esquema de morfosintaxis histórica del español*. Bilbao: Universidad de Deusto, 1983.

Viaro, Mário Eduardo (2006). "Problemas de morfologia e semântica histórica do sufixo *-eiro*", *Estudos Lingüísticos, 35*, 1443-1452. http://www.usp.br/gmhp/publ/Via30.pdf (Consultado em 1.10.2016).

Viaro, Mário Eduardo (2010). "Sobre a inclusão do elemento diacrônico na teoria morfológica: uma abordagem epistemológica", *Estudos de Lingüística Galega, 2*, 173-190. DOI 10.3309/1989-578X-10-8.

Viaro, Mário Eduardo (2012). "A formação do significado agentivo de *-eiro*". Em Cestero Mancera, Ana M., Molina Martos, Isabel, & Paredes García, Florentino (eds.), *Actas del XVI Congreso Internacional de la Alfal (Alcalá de Henares, 6-9 de junio de 2011)*, 2671-2679. Alcalá de Henares: Universidad de Alcalá. www.usp.br/gmhp/publ/ViaA5.pdf (Consultado em 1.10.2016).

Viterbo, Fr. J. de S. R. (1962-1966). *Elucidário das palavras, termos e frases que em Portugal antigamente se usaram e que hoje regularmente se ignoram,* 2 vols. Edição crítica baseada nos manuscritos e originais de Viterbo, por Mário Fiúza (Porto e Lisboa: Livraria Civilização).

Zimmermann, Michel (1989-1990). "Glose, tautologie ou inventaire? L'énumération descriptive dans la documentation catalane du Xème au XIIème siècle", *Cahiers de Linguistique Hispanique Médiévale, 14-15*, 309-338.

O vocabulário dos tecidos no banco de textos do *Dicionário Histórico do Português do Brasil – séculos XVI, XVII e XVIII*: um recorte lexicográfico

Clotilde de Almeida Azevedo Murakawa
Faculdade de Ciências e Letras da Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP). Câmpus de Araraquara

Resumo:

A lexicografia contemporânea, em especial a brasileira, vem se apoiando nos fundamentos propostos pela linguística de corpus para a construção de bancos de dados ou bases textuais como um recurso imprescindível para a elaboração de dicionários quer gerais quer de especialidades. Foi com base nesses fundamentos que o banco de dados do *Dicionário Histórico do Português do Brasil – séculos XVI, XVII e XVIII* (DHPB) foi concebido e organizado. Uma base textual constituída de documentos da mais variada natureza e gênero dos três séculos do período colonial brasileiro. Usando de ferramentas computacionais do Programa Philologic, fez-se uma busca de unidades lexicais referentes a tecidos e panos com o objetivo de organizar um pequeno vocabulário, em que tais unidades estão contextualizadas em documentos dos séculos XVI, XVII, XVIII e começo do XIX. Além de se obter o conteúdo semântico dessas unidades à época, obteve-se também a informação histórica e cultural que elas veiculam. Este texto apresenta essas unidades especializadas, a definição lexicográfica que a elas é dada nos dicionários de língua portuguesa dos séculos XVIII e XIX e o relacionamento delas com a época em que estão documentadas.

Palavras-chave:

Banco de dados, lexicografia histórica, definição lexicográfica, tecido, pano, história colonial brasileira.

Sumário:

1. Introdução. 2. Metodologia. 3. Análise dos dados. 4. Conclusão.

Clothes vocabulary in the database of the Dicionário Histórico do Português do Brasil – séculos XVI, XVII e XVIII: *a lexicographical clipping*

Abstract:

Contemporary lexicography, especially Brazilian, is being based on the foundations proposed by the Linguistic corpus for the construction of a database or textual basis as an essencial resource for the preparation of the general or specific dictionaries. The Dicionário Histórico do Português do Brasil – séculos XVI, XVII e XVIII *(DHPB) was conceived and*

organised upon these foundations. A textual base comprising documents of a varied nature and genre from three centuries of the Brazilian colonial period. Using computacional tools from the Philologic Program, a search was conducted of lexical units relating to fabrics and clothes in order to construct a small vocabulary, where these units are contextualised in documents from the sixteenth, seventeenth, eighteenth and beginning of the nineteenth centuries. In addition to obtaining semantic content from the period, historical and cultural information conveyed was also recorded. This study shows which these specialised units are, lexicographical definitions shown in the Portuguese dictionaries of the eighteenth and nineteenth centuries, and their relationship with the time in which they were documented.

Key words:

Database, historical lexicography, lexicographical definition, fabrics, clothes, Brazilian colonial history.

Contents:

1. Introduction. 2. Methodology. 3. Data analysis. 4. Conclusion.

1. Introdução

A lexicografia contemporânea tem se apoiado, a todo o momento, na linguística de corpus para a construção de dicionários gerais ou especiais da língua. Segundo Sardinha (2004: 3):

> A Linguística de Corpus ocupa-se da coleta e da exploração de *corpora*, ou conjuntos de dados linguísticos textuais coletados criteriosamente, com o propósito de servirem para a pesquisa de uma língua ou variedade linguística. Como tal, dedica-se à exploração da linguagem por meio de evidências empíricas, extraídas por computador.

Biderman (2001) já afirmava com segurança que uma base de textos bem construída era ferramenta imprescindível para a elaboração de dicionários gerais. Sua afirmação pode ser ampliada também para os diversos tipos de dicionários sejam eles de especialidades, históricos ou etimológicos, etc. Dizia ela:

> Para uma seleção criteriosa e científica, a estratégia correta é o recurso a uma grande base textual, um enorme corpus de dados linguísticos, de discursos realmente realizados –língua escrita e oral– para daí extrair a nomenclatura desejada com base em critérios léxico-estatísticos (Biderman 2001: 34).

Partindo do princípio de que o banco de textos, ou banco de dados, ou ainda corpus lexicográfico, conforme denomina Porto-Dapena (2002), é peça fundamental para a construção de um dicionário, ou como diz Sardinha "um artefato produzido para a pesquisa" (2004: 17), foi construído para o projeto *Dicionário Histórico do Português do Brasil – séculos XVI, XVI e XVIII* um banco de dados especial, onde estão inseridos textos de três séculos e de alguns anos mais produzidos no período do Brasil Colonial, ou seja, de 1500 a 1808, duas datas marcantes na história do Brasil. Começando com a Carta de Pero Vaz de Caminha relatando o "descobrimento da terra brasileira" e terminando com documentos escritos até 1808, ano da vinda da Família Real Portuguesa para o Brasil, este banco de textos foi organizado a partir do objetivo do projeto do *Dicionário Histórico*, ou seja, a elaboração de uma obra de referência que pudesse reunir um repertório lexical de aproximadamente 10.000 verbetes, representativo de uma época em que o português europeu trazido pelos colonizadores sofreu, no país, modificações e acréscimos.

Ao término do projeto, cumprindo prazo estabelecido pelo CNPq, agência financiadora da pesquisa, pôde-se elaborar um dicionário com 10.470 verbetes, a partir desse banco construído. Nele foram inseridas obras de natureza diversa, de todo o tipo e gênero, escritas por portugueses sobre o Brasil ou ainda por aqueles que vieram para o Brasil e aí se fixaram. Desta feita, pode-se encontrar depois da Carta de Caminha obras dos missionários viajantes, na sua maioria jesuítas que vieram em missão catequética e no país permaneceram; diários de navegação,

como o de Pero Lopes de Sousa, irmão de Martim Afonso de Sousa; cartas de sesmarias; roteiros descritivos da flora e fauna brasileiras; descrições geográficas; cartas e sermões do Pe.Vieira, pregados no Brasil e de outros oradores sacros que tiveram sua correspondência reunida em obras esparsas; obras e documentos que tratam do Estado do Grão Pará, durante a era pombalina; cartas comerciais trocadas entre comerciantes da colônia com outros de Portugal; obras sobre a nobiliarquia paulistana; atos de câmaras municipais; anais de câmaras de diversos municípios brasileiros; documentos cartoriais; autos de devassas feitos durante a Inconfidência Mineira; processos; inventários; testamentos; alvarás; posturas; bandos; atos de doações de terras, casas e terrenos; cartas de ofício; patentes; cartas dos governadores gerais; provisões; documentos forenses; estatutos de sociedades; constituições dos bispados do Brasil; regimentos militares; obras sobre medicina, farmácia, agricultura, mineração, além da produção literária do barroco e arcadismo.

Com 7.492.472 ocorrências (23.858 páginas escaneadas), este banco deu sustentação para a redação dos 10.470 verbetes, e muito mais teria sido redigido se mais tempo houvesse.

Como é do conhecimento da maioria dos pesquisadores que se dedicam às ciências da lexicologia, lexicografia e terminologia, o projeto do DHPB foi concebido pela Profª. Maria Tereza Biderman, no âmbito do programa Institutos do Milênio do CNPq. Com o falecimento da Profª. Biderman, em maio de 2008, coube à autora deste texto a tarefa de, pelo menos, atingir os propósitos do projeto de elaborar os 10.000 verbetes calculados por Biderman. Pode-se, então, perguntar se o DHPB está completo. E responde-se que não. A partir desse total, muito mais poderá ser acrescentado ao dicionário histórico, pois o banco de dados permite ao lexicógrafo aperfeiçoá-lo, porque nem um terço dele foi utilizado. A par desse banco, há ainda outro com mais 2.049.249 ocorrências, resultante de 8.500 páginas escaneadas. Os dois bancos totalizam, assim, 9.541.721 ocorrências.

Dentre as inúmeras obras que integram o banco de textos, uma delas permitiu reunir duas dezenas de unidades lexicais referentes a tecidos ou panos. É a obra monumental intitulada *Negócios coloniais*, em cinco volumes, organizada pelo Prof. Luís Lisanti e financiada pelo Ministério da Fazenda, com publicação pelo mesmo ministério, em 1973. A obra contém cartas comerciais inéditas, num total de 1792, trocadas entre Francisco Pinheiro, comerciante em Lisboa e seus correspondentes no Brasil, África e praças europeias, na primeira metade do século XVIII.

As cartas que compõem a quase totalidade dos cinco volumes permitem visualizar o intenso comércio de mercadorias variadas entre os representantes no Brasil e Francisco Pinheiro, em Lisboa. A maior parte dessas cartas é relativa ao Brasil e mostra a atividade comercial da época.

Sobre Francisco Pinheiro há que se informar que foi um rico comerciante português com comércio intenso com o Brasil –diga-se Bahia, Rio de Janeiro, Mato Grosso, São Paulo, Ceará, Pernambuco, Minas Gerais– e com países da Europa como França, Holanda, Inglaterra e Itália, e com alguns países da África, como Angola e Costa da Mina.

Ao falecer em 1749, Francisco Pinheiro, ao dividir sua fortuna, "instituiu a alma por herdeira da maior parte dela e erigiu a Congregação de Nossa Senhora da Doutrina por testamentária" (Lisanti 1973: XXXIX). Entretanto, em 1750, uma bula papal e um decreto real, ordenaram que essa parte da herança fosse entregue ao Hospital Real de Todos os Santos, em Lisboa, antigo Hospital de São José, depositário de todos os documentos que comprovam as atividades comerciais de Francisco Pinheiro. Como afirma Lisanti (1973: XXXIX): "São milhares de documentos que servem a história econômica e social do Brasil e Portugal".

Foram objeto de estudo de Lisanti as 1.792 cartas e desse total algumas dezenas foram selecionadas para entrarem no banco de dados do DHPB.

2. Metodologia

A partir de um glossário inserido na obra *Negócios coloniais*, selecionou-se vinte e duas unidades lexicais referentes a tecidos ou panos, com o objetivo de confirmar sua presença ou não no banco de dados do *Dicionário Histórico*. Com esta pequena nomenclatura, começou a pesquisa no banco de textos utilizando os vários recursos do programa *Philologic*, que deu suporte informático ao DHPB. São procedimentos de busca que levam à obtenção de uma variedade de dados linguísticos. Este programa apresenta dois tipos de busca: pelo contexto e por similaridades. A primeira busca leva diretamente ao contexto e através do KWIC chega-se também aos contextos maiores e ao total de ocorrências da forma procurada; já a segunda busca, por similaridades, é de grande valia, pois leva às variantes linguísticas. O programa permite ainda a procura por sintagmas ou por excerto de frase, auxiliando na ampliação da pesquisa.

Tomando os sintagmas *vestido de*..., *pano de*..., *tecido de*..., ou utilizando ainda os referentes a peças do vestuário como *camisa de*..., *calção de..., capa de..., saia de..., capote de..., sobretudo de..., casaca de*..., ou *saiote de*..., sintagmas que foram escolhidos ao acaso, conseguiu-se chegar a vários tipos de textos onde outras unidades lexicais relativas a tecidos foram coletadas. Foram textos como data de sesmarias, inventários, testamentos, confirmação de estatutos religiosos e militares, cartas comerciais, autos de devassa, autos de prisão, descrição de festas, ou ainda livros de despesas e gastos de igrejas e sacristias; além de outros onde estão descritas viagens, costumes dos índios, etc.

Ressalte-se ainda que as unidades no banco de dados apresentaram variantes, e cada uma em contextos e épocas diferentes. Assim, em *cetim*, tem-se a grafia com <s> *setim*; e para *algodão* tem-se as variantes *algodaõ, algudam* e *algudão*.

Procedendo à busca no banco de textos ora por similaridades, ora por contextos, chegou-se ao total de 126 unidades lexicais simples ou sintagmáticas, incluídas neste total as variantes. Numa segunda etapa, reunindo as variantes às formas canônicas, de acordo com o *Vocabulário Ortográfico da Língua Portuguesa* (VOLP – 2009), obteve-se o total de 83 unidades, que formam a nomenclatura objeto deste texto.

O quadro a seguir relaciona as unidades, confirmado seu registro no VOLP, acompanhadas da datação do documento onde estão inseridas, e as variantes também com a respectiva datação. Vale lembrar que apenas uma única unidade (*osteda*) não está registrada no VOLP; entretanto as edições de Morais confirmam sua existência.

Unidades lexicais	Datação	Variantes
1. Algodão	1549	Algodaõ (1591), Algudam (1625), Algudão (1704)
2. Aniagem	1727	Aniage (1769), Niage (1727)
3. Baeta	1583	
4. Baetão	1798	
5. Barbarisco	1654	Berberisco (1654)
6. Barrigana	1717	Barregana (1727), Baregana (1725)
7. Berne	1784	
8. Bertangil	1654	
9. Bertanhão	1764	
10. Bocachim	1725	
11. Bretanha	1757	Bertanha (1725)
12. Bretanha de Hamburgo	1791 (2 oc.)	
13. Brim	1735	
14. Burel	1583	
15. Calamânia	1769	Calamanha (1764)
16. Cambraeta	1727	Cambraetta (1727)
17. Cambraia	1726	Cambraya (1801), Cambrai (1750), Cambray (1757)
18. Camelão	1726	
19. Cassa	1802	
20. Cassemira	1803	

Unidades lexicais	Datação	Variantes
21. Seda	1583	Ceda (1624)
22. Setim	1583	Cetim (séc. XVIII)
23. Chamalote	1587	Chamalotte (1699), Chonbalote (1655), Xamalote (1756)
24. Chita	1757	Xita (1764)
25. Crê	1749	
26. Crepe	1743	
27. Damasco	1561	
28. Damasquilho	1654	
29. Droguete	1753	Dorogete (1725), Durguette (1727), Druguete (1727), Druguette (1727)
30. Duqueza	1727	
31. Durante	1797	
32. Escumilha	1801	
33. Esguião	1750	
34. Esperregão	1742	
35. Espiguilha	1771	
36. Estamenha	1709	
37. Estofo	1762	
38. Estopa	1587	
39. Estopinha	1726	
40. Fustão	1791	
41. Gala	1767	
42. Ganga	1792	Canga (1798)
43. Gingão	1654	Guingão (1771)
44. Holandilha	1645	Olandilha (1725)
45. Hollanda	1767	
46. Lã	1587	Lãa (1680), Lam (1607), Lãm (1797)
47. Lemiste	1802	Llemiste (1725)
48. Linhagem	1709	Linhage (1756)
49. Linho	1500	
50. Lona	1587	
51. Marroquim	1725	
52. Milaneza	1654	
53. Mosselina	1798	Mossellina (1798)
54. Osteda	1750	

Unidades lexicais	Datação	Variantes
55. Panico	1727	
56. Panículo	1756	
57. Panno de aguardente	séc. XVIII	
58. Panno de breu	1587	
59. Panno de estopa	1711	
60. Panno de munção	1730	
61. Panno de Serafina	1764	
62. Panno-rei	séc. XVIII	
63. Pano de algodão	1637	
64. Pano de linho	1618	
65. Pano de serviço	1618	
66. Picaró	1725	
67. Pinhoela	1654	
68. Primavera	1764	
69. Raxetta	1654	Raxeta (1709)
70. Retroz	1727	Retrós (1750)
71. Ruão	1726	Roão (séc. XVIII), Ruam (1728)
72. Saeta	1727	
73. Serafina	1654	Sarafina (1727)
74. Sarge	1627	Sarja (1676)
75. Sargeta	1655	
76. Serapilheira	1714	
77. Sufulie	1727	Sufuli (1771)
78. Tafetá	1583	Tafeta (1624), Thafetá (1712)
79. Troçal	1725	
80. Veludo	1561	Velludo (1587)
81. Veludilho	1779	
82. Xita de paquete	1798	
83. Xita paquete	1798	

Quadro 1. Unidades referentes aos tecidos com as variantes (Banco de dados do DHPB)

Do total obtido, passou-se à consulta nos dicionários, em especial o *Vocabulario Portuguez e Latino* de Raphael Bluteau (1712/1728) e nas várias edições do *Diccionario da Lingua Portugueza* de Antonio de Morais Silva (a primeira de 1789, a segunda de 1813; a terceira de 1823; a sexta de 1858; a nona de finais do

século XIX e a décima de 1949, todas impressas em Portugal), não esquecendo ainda o *Grande Diccionario Portuguez ou Thesouro da Lingua Portugueza* de Frei Domingos Vieira (1871/1874).

A razão da consulta aos dicionários foi para saber a definição registrada para cada tecido, informação sobre sua descrição e procedência e do que era feito, se de *algodão, linho, seda, lã ou cetim,* tecidos básicos que entravam na composição de vários outros. Além da cor e textura, as definições lexicográficas apresentam o local de fabricação e o país de sua procedência. É pela definição que se obtem dados sobre a comercialização de um tecido. O tecido *ruão*, por exemplo, e as variantes *roão* e *ruam*, indicam a procedência do pano, e sua definição em Silva é: "panno de linho tosado, e talvez tinto que serve para forros de vestidos" (1789: II, 359). E Bluteau em seu *Vocabulario* define: "Ruaõ também se chama certo gênero de lenço[1], que se faz na Cidade, & contorno de Ruão" (1720 : VII, 391).

Um outro exemplo é o de *bretanha* (com a variante *bertanha*) definida em Bluteau como: "Panno de linho que nos vem de Bretanha (França)" (1712: II, 689). A nona edição do dicionário de Morais acrescenta que este nome *bretanha* passou a ser dado aos tecidos de igual qualidade, manufaturados em outras terras: *bretanha de França*, *de Irlanda* (1712: I, 360). Tal informação justifica a presença na nomenclatura organizada de *bretanha de Hamburgo*, registrada em uma certidão de autos de sequestro, de 1791.

Na impossibilidade de se apresentar todas as unidades em seus contextos e a definição lexicográfica registrada nos dicionários consultados, foram selecionados alguns exemplos para ilustrar o que se pretende afirmar: a importância de bancos de textos construídos com objetivos precisos para um determinado fim e a possibilidade que os bancos oferecem para inúmeras outras pesquisas linguísticas, quer de natureza léxica, quer morfológica, sintática ou semântica. No caso específico deste texto, o foco é a construção de um vocabulário especial referente a tecidos, e pelas definições lexicográficas registradas nos dicionários, conhecer sua definição, comercialização e destacar alguns aspectos da época.

3. Análise dos dados

As unidades lexicais *tecido* e *pano* foram consideradas como ordenadoras de todo o conjunto, ou seja, arquilexemas da nomenclatura construída. Como mencionado anteriormente, o *algodão*, o *linho*, a *seda*, a *lã* e o *cetim* aparecem como fios que entram na trama dos mais variados tecidos. Apenas a unidade *linho* está na Carta de Caminha, de 1500.

1 **Lenço**, s .m. "Toda a tela de linho, e algodão" (Silva, 1813: II, 214).

Do total de 126 unidades, consideradas as variantes, quinze estão em documentos do século XVI, vinte e uma no século XVII, sete no século XIX e todo o restante 86, estão no século XVIII, indicando o que Lisanti mencionou em seu livro sobre o intenso comércio de mercadorias neste século e nelas incluídos os tecidos.

É interessante observar que o banco de dados registra inúmeros sintagmas formados com a unidade *pano / panno*: *panno de estopa*, *panno de breu*, *panno de aguardente*, *pano de munção* (*monção*), *pano de serafina*, *pano de algodão, pano de linho e pano de serviço*. Destas combinatórias depreende-se que *panno de breu*, *panno de aguardente*, dado o contexto onde estão empregados, significam 'tecido embebido em uma substância' (*panno de breu* aparece em texto que trata de navegação e instrumentos náuticos, e *panno de aguardente,* em texto sobre medicamento curativo para feridas). Com relação a *pano de munção*, deduz-se que era tecido usado para proteger pessoas ou objetos durante o período de tempo chuvoso. As duas ocorrências estão documentadas em duas cartas remetidas de Minas Gerais para Lisboa, solicitando este tipo de tecido.

A fim de ilustrar a análise proposta, foram selecionadas, dentro das 83 unidades do vocabulário, algumas unidades de cada século com seu contexto, conforme registro no banco de dados do DHPB, e a definição lexicográfica de cada uma, já que a definição é ponto fundamental para conhecimento do tecido e da época.

Do século XVI, selecionou-se *burel*, *chamalote*, *tafetá* e *veludo*[2].

> **Burel**, s.f. Panno grosseiro de lã, de que andão vestidos os Capuchos; e que antigamente se trazia por luto (Silva, 1789: I, 202).
>
> > «Piratininga é villa da invocação da conversão de São Paulo; está do mar pelo sertão dentro doze leguas; é terra muito sadia, ha nella grandes frios e geadas e boas calmas, é cheia de velhos mais que centenários, porque em quatro juntos e vivos se acharam quinhentos annos. Vestem-se de **burel**, e pellotes pardos e azues, de pertinas compridas, como antigamente se vestiam. Vão aos domingos á igreja com roupões ou berneos de cacheira sem capa. A villa está situada em bom sitio ao longo de um rio caudal».
> >
> > Padre Fernão Cardim (1980) [1583], *Informação da missão do P. Christovão Gouvêa ás partes do Brasil, anno de 83, ou Narrativa epistolar de uma viagem e missão jesuítica* [A00_0751 p. 173].
>
> **Chamalote**, s.f. Tecido de lã ou pello de camelo.§ Fazenda encorpada de seda, ou de seda e lã, cujo tecido apresenta varias ondulações (Silva, 18...: I, 448).

2 As definições foram extraídas dos dicionários mencionados nas referências e os contextos estão transcritos conforme registro no banco de dados do DHPB, indicando o número do arquivo e página da obra entre colchetes.

«Jucuriaçu é uma arvore que se dá em terras fracas [...]. Ha uma casca de côr parda, com aguas pretas, e outra vermelhaça, com aguas tambem pretas, umas e outras da feição do **chamalote**; e umas e outras tem o cheiro suavíssimo, e na casa onde se lavra sahe o cheiro por toda a rua, e os seus cavacos no fogo cheiram muito bem; a qual madeira é muito estimada em toda a parte pelo cheiro e formosura».
Gabriel Soares de Sousa (1938) [1587], *Das arvores means com differentes propriedades, dos cipós e folhas uteis* (Parte segunda, Título 9) [A00_0185 p. 253].

Tafetá, s.m. Tecido de seda muito lustroso, em fios lizos e rectilineos; emprega-se em cortinas, forros, etc. (Silva, 18...: II, 866).

«Martim Rodrigues hum vistido de berberisco calsão Roupeta E capa a Roupeta forrado o corpo de bertangil E as abas de **tafeta** preto tudo novo en sua avaliasão de des mil rs. »
Maria da Silva [1654], *Inventário e testamento de Martim Rodrigues Tenório* (1654) [A00_0757 p. 275].

Veludo, s.m. Tecido de seda, ou de algodão com pello muito macio; e é lizo, lavrado, etc. (Silva, 18...: II, 976).

«E acabado o offertorio se assentou elle em huma cadeira no degrao do altar com a mitra de brocado na cabeça, e assi elle como os dous, scilicet, diacono e subdiacono, revestidos de vestimenta e dalmaticas de **veludo** verde e sabastros de brocado muyto rico, que foy da capella d'El-Rey, afora outros 4, que estavão ao redor delle [...]».
P. Leonardo do Vale (1956) [1561], *Carta do P. Leonardo do Vale por comissão do P. Luís da Grã ao P. Diego Laynes* (Baía 23 de setembro 1561) [a00_0055 p. 61].

Das quatro unidades lexicais descritas, três delas ainda estão na língua portuguesa, tendo caído em desuso apenas *burel*.

Para representar o século XVII, servem de exemplo *barbarisco*, *damasquilho*, *milaneza* e *raxetta*. A unidade *raxeta*, em forma variante, entretanto, está em documento de 1709, e *raxetta*, abaixo contextualizada, em 1654. Seguem os contextos e as definições:

Barbarisco, s.m. Certo tecido de lã, vindo da Barbaria (região) e relacionado na pauta dos mercadores de retalho, junto aos Estatutos approvados em *Alv. de 16 de dezemb., de 1757* (Silva, 18...: I, 318).

«E oitenta rs lhe derão hua espingarda de tres palmos en tres mil E duzentos rs lhe derão o vistido de **barbarisco** calsão E Roupeta en sua avaliasão de des mil rs [...]».
Maria da Silva [1654], *Inventário e testamento de Martim Rodrigues Tenório* (1654) [A00_0757, p. 287].

Damasquilho, s.f. Panno de seda a modo de Damasco, mas mais leve (Bluteau, 1713: III, 7).

«Roupetta e jubão de pano de algodão listrado com huas mangas de **damasquilho** verde em sua avaliação por dous mil e oito sentos Reis foi avaliado hum calção e Roupetta de algodão de gingão [...]».
Gaspar Dias Peres [1654], *Inventário e testamento de Gaspar Dias Peres* (1654). [A00_0755, p. 18].

Milaneza, s.f. Certo panno tecido em Milão (Silva, 1789: II, 82).

Esta definição se encontra mais bem descrita em Vieira (1873: IV, 239): "Especie de panno cujo fundo é um fio coberto por dous fios de seda, dos quaes um menos comprido que o outro, fórma sobre o fio um pequeno relevo a distancias iguaes".
«[...] e os mandou apregoar, [...] fis este termo eu ignaccio gomes velles escrivão dos orfãos que o escrevi foi Rematado o vestido de **milaneza** Roxo em Miguel nunes camacho por tres mil, e duzentos Reis, fiado por seis mezes, e deu por seu fiador, e principal pagador, a domingos Barboza e por não aver quem mais desse o dito juis lho mandou Rematar de que fis este termo em que asinarão com, o ditto juis e eu ignaccio gomes [...]».
Gaspar Dias Peres [1654], *Inventário e testamento de Gaspar Dias Peres* (1654). [A00_0755 p. 18].

Raxetta, s.f. Diminutivo de raxa. Especie de raxa menos grossa. [**Raxa**, Antigo panno grosseiro de menos valor. Hoje, tecido de lã com largas riscas, que serve para cobertores de cama] (Vieira, 1874: V, 91).

«[...] foi avaliado hum vestido de **Raxetta** calção e Roupetta e jubão de pano de algodão listrado [...]».
Gaspar Dias Peres [1654], *Inventário e testamento de Gaspar Dias Peres* (1654). [A00_0755 p. 18].
«Ordenamos que nenhum Religioſo de qualquer qualidade que ſeja, ſaya nem vá ás Communidades em habitos, ou tunicas pardas de **raxeta**, ou eſtamenha, ou outra qualquer cor, que naõ ſeja cinzenta ſobpena de tres diſciplinas pela primeyra vez, & a ſegunda ſeja caſtigado ao arbitrio do Irmão Miniſtro Provincial».
Desconhecido [1709], *Breve da Confirmaçam destes Estatvtos* [A00_2499 p. 131].

As unidades selecionadas não mais são usadas no português e nem sempre registradas nos dicionários portugueses. Mas o VOLP as registra.

Finalmente, buscando os contextos e as definições da maior parte das unidades do século XVIII que compõem este pequeno vocabulário, foram selecionadas as seguintes: *calamânia*, *olandilha*, *serafina*, *sufolié*, *xita paquete* ou *xita de paquete*.

Calamânia, s.f. Tecido de lã listrado às vezes misturava-se na tecitura a seda ou pelo de cabra; existiam várias cores e qualidade; do francês "calmande" (Lisanti, 1973: I, LVI).

Apesar dos dicionários consultados não registrarem a unidade lexical *calamânia*, sua definição foi encontrada no glossário elaborado por Lisanti na obra *Negócios coloniais* (1973: LVI). A décima edição de Silva, entretanto, registra a unidade e diz o seguinte: "Tecido pouco apreciado e que se usou no século XVIII" (Silva, 1949: II, 746); o que não se pode considerar como uma definição lexicográfica. O contexto, a seguir, é da variante:

Calamanha

«Dous Frontaes de brim dos altares colateraes muito velhos Hum dito de Damasco branco, já muito velho Outro dito de **calamanha** prata já muito velho Outro dito de xita com ramos encarnados e franjas verdes. Huá cazulla de primeira com ramos vermelhos e campo branco com seo manipulo e Estola».

Jeronymo Alvares de Carvalho (1997) [1764], *Inventario dos bens pertencentes a Igreja Parochial desta Villa de Cyntra feita ao 1 de Janeiro de 1764* [A00_2504 p. 91].

Olandilha, s.f. Panno de linho grosso engommado, ou encerado, de fazer entretelas dos vestidos (Silva, 1813: II, 362).

«[...] os coais 2 enbrulhos são de **olandilha** amarella e a bara leva para huma das pontas hum pedaço da mesma **olandilha** cozida com a marqua a margem com o nº 3 e os dois vam com a mesma marca de n[1] e 2, q. D[s] permita por q[m] he levar tudo a salvam[to] a prezença de VM., e destes fazer lhe sem mil cruzados [...]».

Fran.[co] da Cruz (1973) [1727], *Cartas remetidas para Lisboa-Minas Gerais* [A00_0442 p. 314].

Serafina, s.f. Tecido de lã delgada para forros, cortinas, &.c. (Silva, 1813: II, 575).

«Uma medida de quartilho de cobre e outra de meio do mesmo. Nove varas encarnadas das insignias da mesma camara. Um panno de **serafina** verde para mesa. Um prato de jacarandá com tinteiro, areeiro e caixa de obreias do dito páo».

Victorino Soares Barbosa (1872) [1764], *Registro dos Autos da Creação da Villa de Montemór-o-Novo do Grão-Pará* [A00_1208 p. 194].

Sufulie, Certo estofo de algodão. *Pauta dos Merc. de Retalhos, de 13 de dezembro de 1757.* (Silva, 18...: II, 848).

«Perparou-se com ũa tanga de **sufuli** e baeta vermelha; ao pescoço várias miçangas, pente na testa, chapéu na cabeça, de que ficou muito alegre, e foi dizer aos seus, tanto que saiu da barraca, que estava muito bonita, o que se lhe percebeu por ser quase na língua da terra».

> Afonso Botelho de S. Paio e Sousa (1962) [1771], *Relação do primeiro encontro que tivemos com os índios do sertão do Tibagi, nos campos de Guarapuava, aos 16, e aos 17 de Dezembro de 1771* [A00_2297 p. 225].

Com relação às unidades *xita de paquete* ou *xita paquete*, esclarece-se que o sintagma não foi encontrado em nenhum dicionário, como sub-entrada, o que levou à suposição de alguns significados. Sabe-se que a *chita* era um tecido de algodão estampado com cores variadas. Com relação à *paquete*, além do significado de "navio a vapor de grande marcha", foi encontrado o sintagma *papel paquete* em Silva, nona edição (II, 474), significando "papel muito fino para cartas e contas, destinadas ao correio, a fim de se pagar menor porte". Mas também pode- se pensar que tal tipo de *xita* vinha transportada por paquete, daí seu nome. Nenhuma das duas possibilidades, entretanto, é satisfatória ou pode ser confirmada.

> «[...] sim o cabello, que he preto crescido por igual, orelhas pequenas rosto redondo testa curta, sobrancelhas pretas, olhos da mesma cor, naris afilado boca pequena labios grossos, com bastante barba: está vestido com camiza de pano de linho, calças de **xita paquete**, e meias brancas com xinellins. E nesta forma o deixei ficar nas ditas cadeas debaixo de chave entregue ao carcereiro dellas Bento Joze de Freitas, [...]».
> Desconhecido (1931) [1798], *A Inconfidencia da Bahia em 1798. Devassas e sequestros (conclusão)* [a00_226 3 p. 160].

Do início do século XIX, período ainda compreendido no recorte do tempo do DHPB, foram selecionadas como exemplo as unidades *cassa*, *cassemira* e *lemiste*. *Cassa* não apresenta variante no banco de dados; entretanto os dicionários indicam a grafia *caça* também usada.

> **Caça** ou **cassa,** s.f. (do arab. *Caça*, vestir, cobrir). Fazenda de algodão fina para roupas, cortinas, etc. §it. Feita de lã, chamada *cassa lã*. V. Musselina. (Silva, 1858: I, 369).
>
> «[...] muitas vezes cahem e se lhes veem os peitos, sem que esses máos criticos se lembrem de que estão debaixo da zona torrida, onde o grande frio corresponde ao que ahi sentimos em Mayo. Igualmente as nottão de andarem em suas cazas muitas vezes descalças, e de ordinario sem meyas, com camizas de **cassa** finissima e cambraya transparente; sem que deem attenção, como disse, ao clima em que se achão [...]».
> Luiz dos Santos Vilhena (1921) [1802], *Carta primeira* [A00_0404 p. 47].

Nos dicionários pesquisados não foi encontrado o tecido *cassemira*; a definição a seguir está no *Dicionário da Língua Portuguesa* de Antenor Nascentes (1961-1967), na forma *casemira*:

Casemira, s.f. Tecido fino de lã para roupas (Do ingl. *Kerseymere*, fazenda pura de Kersey, com influência do antropônimo *Casimiro*) (Nascentes, 1961: I, 405).

> «Entre estes as Pastas de Argila-branca, e amarella p.ª saber, q̃. alguns Inglezes a tem mandado hir p.ª alimparem, ou embranquecer a **Cassemira** branca, tingir a amarella, e Camursas, pela superioridade, q̃. lhe achaõ relativamente a q̃. vem de Inglaterra. Dellas há aqui vastas, e altas Minas, que a sua moça só preciza do beneficio da lavagem [...]».
> Antonio Joze da Franca e Horta (1990) [1803], Para o M.mo Ex.mo Snr. Anadia. Março 27 [A00_1079 p. 37].

Lemiste está em documento de 1802, e uma variante *llemiste* de 1725, mas já registrada na segunda edição de Silva (1813: II, 214) com a seguinte definição: "Panno fino de lã preto". Esta definição pode ser completada com o que está registrado na nona edição do mesmo Silva: "Panno de lã preto, o mais perfeito e fino dos de Segovia; e hoje de outras partes" (vol. II 253).

> «As pessas com que se ornão são de excessivo valor e quando a funcção o permitte aparessem com suas mulatas e pretas vestidas com ricas sayas de setim, becas de **lemiste** finissimo, e camizas de cambraya, ou cassa, bordadas de forma tal que vale o lavôr tres ou quatro vezes mais que a pessa [...]».
> Luiz dos Santos Vilhena (1921) [1802], *Carta primeira* [A00_0404 p. 47].

Finalmente, há que se registrar que uma única unidade não foi encontrada no VOLP (2009), muito embora esteja registrada já em Silva (1813). *Osteda* é definida: "s.f. Estofo antigo de França. *Ord. Af. 4.55. ou de Ostende?*" (1813: II, 374). Em Vieira encontra-se: "Estofo antigo de Ostende, cidade na Flandres occidental" (1873: IV, 593).

> «Uma campainha para tocar à Missa. Uma vestimenta de **osteda** com estóla, manipulo, almofada para o Missal, e bolsa de corporais com três véus de calices de nobrêsa, branco, vermelho e rôxo, tudo com guarnições de retrós côr de ouro. Um frontal da dita Osteda com as mesmas guarnições e franjas de retrós côr de ouro, para servir [...]».
> Francisco Xavier de Mendonça Furtado (1963) [1750], *Tratado de 1759: Material destinado a atos religiosos* [A00_0333 p. 381].

4. Conclusão

Este texto buscou dar evidência a dois aspectos da pesquisa. O primeiro refere-se à importância do banco de dados do DHPB, construído especialmente para a elaboração do dicionário. A partir dele inúmeras outras pesquisas poderão ser realizadas, pois a sua composição oferece material não só para o estudo do léxico, mas também para os demais níveis linguísticos da língua portuguesa nos séculos XVI, XVII, XVIII e início do XIX.

Um segundo aspecto refere-se ao vocabulário que foi construído a partir dos textos do mesmo banco. As 87 unidades que compõem este vocabulário estão, na sua maioria, no século XVIII e com uma curiosidade: a maior parte caiu em desuso, à exceção de *cambraia, casemira, chita, crepe, fustão e musseline,* que ainda são conhecidas e encontradas no comércio de tecidos. Acredita-se que com o desaparecimento do referente, também as palavras saíram de uso e permanecem hoje "guardadas" nos dicionários de Língua Portuguesa e também no VOLP. Através das definições lexicográficas pôde-se conhecer o tipo de tecido, sua textura, cor e procedência. Era a Europa, no século XVIII, a grande fornecedora dos tecidos e tal fato se encontra documentado nos textos.

Com relação ao século XVI, em que um número pequeno de unidades lexicais foi registrado, é o século que documenta o *linho, algodão, lã, seda* e *cetim,* que entraram na trama de vários tecidos. Em compensação, muitas outras unidades ainda permanecem em uso na língua portuguesa: *chamalote, damasco, estopa, lona, tafetá* e *veludo*.

Finalmente, do século XVII, nenhuma unidade permaneceu em uso; apenas os dicionários testemunham a sua existência e registram sua definição.

Referências bibliográficas

Biderman, Maria Tereza Camargo (2001). "Os dicionários na contemporaneidade: arquitetura, métodos e técnicas". En Oliveira, Ana Maria Pinto Pires de, Isquerdo, Aparecida Negri, & Alvez, Ieda Maria (eds.), *As ciências do léxico: lexicologia, lexicografia, terminologia* (2.ª edição), 129-142. Campo Grande: Humanitas / Editora UFMS,.

Bluteau, Raphael (1712-1728). *Vocabulario portuguez e latino*, 10 vols. Coimbra: Collegio das Artes da Companhia de Jesus; Lisboa: Joseph Antonio da Silva.

Lisanti, Luís (1973). *Negócios coloniais. Uma correspondência comercial do século XVIII*, 5 vols. Brasília: Ministério da Fazenda da República Federativa do Brasil; São Paulo: Visão Editorial.

Nascentes, Antenor (1961-1967). *Dicionário da Língua Portuguesa*, 4 vols. Rio de Janeiro: Imprensa Nacional.

Porto Dapena, José-Álvaro (2002). *Manual de técnica lexicográfica*. Madrid: Arco/ Libros.

Sardinha, Tony Berber (2004). *Linguística de corpus*. Barueri: Editora Manole.

Silva, António de Morais (1789). *Diccionario da lingua portugueza*, 2 vols. Lisboa: Officina de Simão Thaddeo Ferreira; Lisboa: Typographia Lacerdina, 1813 (2.ª edição); Lisboa: Typographia de M. P. Lacerda, 1823 (3.ª edição); Lisboa: Typographia de Antonio José da Rocha, 1858 (6.ª edição); Lisboa: Editora Empreza Litteraria Fluminense de Santos, Vieira & Commandita, s.d. (9.ª edição).

Silva, António de Morais (1949-1959). *Grande dicionário da língua portuguesa* (10.ª edição), 12 vols. Organização de Augusto Moreno, Cardoso Junior e José Pedro Machado. Lisboa: Editorial Confluência.

Vieira, Domingos (1871-1874). *Grande Diccionario Portuguez ou Thesouro da Lingua Portugueza*, 5 vols. Porto: Ernesto Chardron e Bartholomeu H. de Moraes Editores.

VOLP = *Vocabulário Ortográfico da Língua Portuguesa* (5.ª edição). Rio de Janeiro: Editora Global, 2009.

Arabismos no galego?

Juan A. Thomas
Utica College

Resumo:

O obxectivo deste estudo é elaborar unha lista de empréstimos léxicos exclusivos do galego para botar luz sobre posibles situacións de contacto lingüístico no pasado entre galegofalantes e falantes do árabe andalusí ou do romandalusí. A presenza de arabismos directos axudará a identificar situacións de contacto únicas entre o galego e as linguas faladas en Al-Andalus. A partir do *Diccionario de arabismos y voces afines en iberorromance* de Corriente (1999), atópanse dezasete entradas exclusivas do galego. Só sete son arabismos directos. Os demais son romandalusismos, evolucións dentro do galego e un berberismo. Discútense a etimoloxía, a relación coas demais linguas peninsulares, especialmente o castelán e o portugués, a actualidade de uso e a evolución semántica de cada vocábulo. A comparación da etimoloxía dos empréstimos cos seus campos semánticos e con certos eventos socio-históricos axuda a descubrir unha situación de contacto entre falantes de árabe e falantes de galego que resultou na incorporación destes vocábulos no galego.

Palabras chave:

Sociolingüística, linguas en contacto, empréstimos léxicos, arabismos, galego.

Sumario:

1. Introdución. 2. Metodoloxía. 3. Resultados. 3.1. Evolucións dentro do galego e vocábulos do latín. 3.2 Romandalusismos. 3.3. Arabismos. 4. Discusión. 5. Conclusión.

Arabisms in Galician?

Abstract:

The objective of this study is to prepare a list of exclusive loanwords in Galician in order to shed light on possible situations of linguistic contact in the past among speakers of Galician and speakers of Andalusian Arabic or Andalusian Romance. The presence of direct Arabisms will help to identify unique situations of contact between Galician and the languages spoken in Al-Andalus. Starting from Corriente's Diccionario de arabismos y voces afines en iberorromance *(1999), seventeen entries, unique to Galician, are found. Only seven are direct Arabisms. The rest are loanwords from Andalusian Romance, evolutions within Galician and one Berberism. Each word's etymology, relation to other peninsular languages, especially Castilian Spanish and Portuguese, current use and semantic evolution are*

discussed. The comparison of the loanwords' etymology with their semantic fields and with some socio-historical events helps to describe a situation of contact among Arabic speakers and Galician speakers which lead to the incorporation of this vocabulary in Galician.

Key words:

Sociolinguistics, languages in contact, loanwords, Arabisms, Galician language.

Contents:

1. Introduction. 2. Methodology. 3. Results. 3.1. Evolutions within Galician and words from Latin 3.2 Words from Andalusian Romance. 3.3. Arabisms. 4. Discussion. 5. Conclusion.

1. Introdución

O contacto entre linguas é un dos factores capaces de producir o cambio lingüístico, chegando a efectuar grandes modificacións estruturais nunha lingua. O contacto lingüístico realízase por medio de diversos procesos, por exemplo, o empréstimo, o cambio de código, o bilingüismo, e a aprendizaxe de segundas linguas. Os resultados do contacto lingüístico poden variar de situación en situación, pero inclúen a integración de novos vocábulos ou incluso a creación dunha lingua crioula. O empréstimo léxico é, quizais, un dos primeiros sinais dunha situación de contacto lingüístico e aínda que o proceso de por si parece incapaz de efectuar cambios lingüísticos, pode propiciar a incorporación de novos fonemas ou incluso desinencias verbais. Thomason e Kaufman (1988: 37) definen o empréstimo en xeral como "the incorporation of foreign features into a group's native language by speakers of that language". O empréstimo léxico é a adaptación fonolóxica, semántica e morfosintáctica dun vocábulo dunha lingua orixinal a outra lingua receptora e, máis importante, é asequible aos falantes monolingües da lingua receptora. Esta última característica diferencia o empréstimo léxico do *nonce borrowing* (empréstimo espontáneo; Poplack, Sankoff, & Miller, 1988:25). Cando un vocábulo pasa directamente dunha lingua orixinal, A, a unha lingua receptora, B, sen que medie unha terceira lingua, C, chámase un empréstimo directo. Se interfire unha terceira lingua, non se pode considerar un empréstimo da lingua A, xa que pasou polo proceso de adaptación fonolóxica, morfosintáctica e semántica segundo as regras da lingua C.

Kiesler (2003) compara os arabismos no portugués e no castelán: 22 arabismos exclusivamente portugueses fronte a 17 exclusivamente casteláns e un 1,13% de arabismos no vocabulario fundamental do portugués contra un 0,92% no castelán. Portugal ten a quinta parte da superficie de España pero hai 500 topónimos árabes en Portugal e 2000 en España. No vocabulario de Machado (1991) figuran uns mil vocábulos de orixe árabe no portugués. O castelán ten un número maior de empréstimos léxicos exclusivos árabes que o catalán, e, polo tanto, Kiesler (2003: 264) deduce que o oeste da Península Ibérica coñeceu unha arabización máis forte que o leste. Que pasa no noroeste peninsular e no galego? Thomas e Sayahi (2005) elaboraron unha lista duns catrocentos vocábulos galegos que teñen étimos árabes. A partir desta lista, escolléronse doce palabras ausentes no castelán para describir algúns procesos de adaptación ao galego sen que medie o castelán. O estudo indicou un alto grao de variabilidade tanto nos campos semánticos coma nos procesos de adaptación. O obxectivo do presente estudo é identificar outro grupo de arabismos galegos ausentes tanto no castelán coma no portugués.

O castelán, como lingua alta na condición diglósica galega secular, exerceu grande influencia sobre o galego, e dada a escasa presenza de árabes en terras da Galiza actual, crese que os arabismos en galego poden provir do castelán. Por tal motivo

resulta de interese estudar os arabismos exclusivos do galego, xa que poderían dar pistas sobre o contacto entre o árabe, ou unha variedade del, e o galego. Así e todo, cómpre considerar o portugués neste estudo. O portugués e o galego teñen unha relación xenealóxica estreita, e, por iso, comparten unha masa léxica común. Suponse que todos os arabismos galegos figuran no portugués (Piel, 1965: 1267). Aínda que o obxectivo do estudo citado arriba contemplou unicamente considerar a influencia do castelán (Thomas, & Sayahi 2005), seis dos arabismos galegos discutidos non se rexistran tampouco no portugués. Este feito é interesante porque o portugués chegou a forxar a súa propia identidade distinta do galego precisamente co contacto lingüístico coas falas mozárabes do sur e, polo tanto, o portugués conta con numerosos arabismos. Illar o galego dos arabismos en portugués axuda a descubrir situacións de contacto lingüístico posteriores ás que se produciron con árabes antes de que as falas galegas meridionais pasasen a formar o portugués. Así, os arabismos exclusivos do galego merecen unha análise semántica, histórica e etimolóxica para poder entender mellor a súa vía de entrada no galego, falado nun territorio onde historicamente houbo menos presenza de árabefalantes que noutras zonas da Península Ibérica.

O presente estudo aprovéitase tamén dunha descrición proporcionada por Frías Conde (2000) sobre o contacto entre o árabe clásico, o árabe andalusí, o romandalusí e o galego. É fundamental entender algo sobre a variedade do árabe falado na Península Ibérica. Sinalar o árabe como fonte do arabismo pode ser demasiado simplista ás veces (Corriente, 1996: 5). Aínda que o árabe clásico se usaba na Península Ibérica, o seu uso, igual que nos países arabófonos, limitábase ao ámbito relixioso ou literario. Non hai que descartar o árabe clásico como fonte de arabismos, xa que a tradución de textos árabes pode introducir arabismos no galego. Porén, estes adoitan ser cultismos e moitas veces non pasan por todos os procesos de adaptación como as voces patrimoniais, é dicir, esas voces transformadas polas regras fonolóxicas durante o uso diario a través dos séculos. O árabe falado na Península Ibérica, ou máis concretamente no Al-Andalus, o nome empregado polos árabes, é denominado aquí como árabe andalusí. Por outra banda, neses territorios ocupados polos árabes, falábase tamén unha variedade de romance, coñecida aquí polo nome romandalusí, fortemente influenciada polo árabe andalusí.

Como toda situación de contacto linguístico é produto dunha situación de contacto social, cómpre resumir a historia da presenza musulmá no territorio que hoxe comprende Galiza. Non temos datos certos sobre a incursión dos musulmáns no Noroeste cando someteron os pobos da Península Ibérica no século VIII (Pallares, & Portela, 1996: 165-167). Existen noticias posteriores, escritas no século XIII, sobre Musa ibn Nusayr, un dos xefes militares das tropas de árabes e bérberes. Dise que este xeneral chegou ata Lugo, dende onde volveu a Damasco. Noutra compilación de tradicións, datada no século XI (Pallares, & Portela, 1996:166), descríbese a primeira sublevación de bérberes e o seu subseguinte abandono dos territorios entre

os ríos Miño e Douro a mediados do século VIII, polo tanto, o máis probable é que os bérberes estivesen alí durante catro décadas. A presenza da lingua árabe posterior límitase ás incursións esporádicas, á entrega de escravos sarracenos a igrexas e mosteiros galegos e á chegada de mozárabes cristiáns do sur da península, buscando refuxio no norte. A expedición de Almanzor a Santiago en 977 é moi coñecida polo saqueo da cidade. Cunha (*apud* Frías Conde, 2000: 167) presenta unha serie de eventos dende 911 ata 1220, que documentan a entrega de prisioneiros musulmáns escravizados a institucións relixiosas galegas.

2. Metodoloxía

Recóllense os arabismos exclusivos galegos e os seus étimos do *Diccionario de arabismos y voces afines en iberorromance* de Federico Corriente (1999). Ademais das entradas únicas, recóllense tamén vocábulos galegos que teñen reflexos nas demais linguas penisulares se mostran diferenzas na súa forma fonolóxica, categoría léxica ou acepción. Empréganse os étimos de Corriente para afirmar a presenza dun determinado vocábulo en romandalusí ou árabe andalusí, aínda que eses étimos non estean documentados en todos os casos. Corriente (1999) non identifica a fonte das entradas galegas, pero parece que veñen do *Diccionario normativo galego-castelán* (Monteagudo Romero, & García Cancela, 1988) do que as recolleu para un traballo anterior (Corriente 1998). Empregouse o dicionario en liña, *Dicionario de dicionarios* (Santamarina, 2006-2013), para comprobar as acepcións en galego das palabras analizadas aquí, o *Dicionário Priberam da Língua Portuguesa* para o portugués e o *Diccionario de la Lengua Española* de la Real Academia para o castelán. A abreviatura *ra*. significa romandalusí, *lt*. o latín, *and*. o árabe andalusí, e *ac*. o árabe clásico. A táboa 1 presenta a correspondencia entre algúns símbolos e os sons que representan.

Símbolo	Valor fonético
ḥ	consoante fricativa, farínxea, xorda
x	consoante fricativa, velar, xorda
š	consoante sibilante, fricativa, postalveolar, xorda
ḍ	consoante oclusiva, dental, sonora, velarizada, enfática
ṭ	consoante oclusiva, dental, xorda, velarizada, enfática
q	consoante oclusiva, uvular, xorda
h	consoante fricativa, glotal, xorda
ʔ	consoante oclusiva, glotal, xorda
č	consoante africada, postalveolar, xorda

Táboa 1. Símbolos utilizados nas transcricións

3. Resultados

As táboas 2, 3 e 4 resumen as etimoloxías e os campos semánticos dos vocábulos estudados aquí. Todas as entradas están recollidas en Corriente (1999), e, de acordo con este autor, son arabismos. Os vocábulos recollidos na táboa 2 son evolucións dentro do galego, os da táboa 3 son romandalusismos e os da táboa 4 son os auténticos arabismos. Nove vocábulos teñen unha orixe híbrida. Algúns son de orixe románica ou incluso latina, entraron no árabe andalusí ou romandalusí, sufrindo adaptacións ou extensións semánticas, pero aínda así, salienta a forma románica da palabra. Outros teñen raíces claramente árabes, pero os seus significados evolucionaron no galego.

3.1. Evolucións dentro do galego e vocábulos do latín

Os vocábulos que figuran na táboa 2 non son arabismos, senón evolucións dentro do galego. Dous deles veñen directamente do latín. O *alcroque*, *digitalis purpurea,* é unha planta con usos médicos, tamén coñecida coma dedaleira ou dixital. A orixe de *alcroque* é romandalusí e consiste no artigo definido *al* mais a raíz latina *croccus* 'azafrán'. Corriente (1999: 145) tamén documenta *alcroco* en castelán, pero cunha diferenza semántica. *Alcroco* en castelán e o seu étimo romandalusí significan *azafrán*. Así, *alcroque* pasa por unha evolución semántica no galego.

A forma fonolóxica do adxectivo *desaxacado* 'desaliñado' consiste nun prefixo e sufixo romances e o radical árabe andalusí *šákl* 'arranxo'. Corriente non rexistra o verbo *axacar* 'arranxar, compor', pero está en Frías Conde (2002: 75), que propón unha derivación a partir da raíz árabe {ʃky}. *Desaxacado* está clasificado aquí como evolución dentro do galego polos seus afixos románicos.

Chuchamel é unha planta orixinaria de Europa, e non só dos territorios de Al-Andalus, senón de toda a Europa occidental, incluso fóra da Península Ibérica. Losada, Castro e Niño (1992) recollen tres nomes botánicos para *chuchamel*:[1] *anchusa officinalis, cerinthe major* e *lonicera periclymenum*. Franco Grande (1972) explica que é "unha planta cuxo talo chúpase e ten un certo sabor que apracenta". De todos os xeitos, *chuchamel* é outro vocábulo cunha forma moi románica, e incluso o mesmo Corriente (1999: 287) recoñece que a raíz pode provir do latín *sūgere.* Corriente indica que o romandalusí é a fonte que deu o verbo *chúch* ao árabe andalusí, pero non ofrece ningunha documentación que testifique unha orixe romandalusí do vocábulo galego. Polo tanto resulta difícil consideralo arabismo e, de feito, é máis probable que sexa unha evolución dentro do mesmo galego. O castelán tamén ten unha variante, *chupamieles.*

1 http://sli.uvigo.es/DdD/ddd_pescuda.php?pescuda=chuchamel&tipo_busca=lema (Consultado en 06.11.2015).

Alfóndega (substantivo) é un peixe, *trigla hirundo*, considerado arabismo (Corriente, 1999: 163). Ten orixe no árabe andalusí *xuṭṭáfa*, que vén do árabe clásico *xuṭṭāfah* 'andoriña'. Con todo, esta proposta require unha derivación complicada, a partir do reflexo hipotético **alfondefa*. Corriente tamén propón a extensión metafórica de *alfóndega* 'pousada' para significar o peixe que se serve na pousada, o que suxire que é unha evolución dentro do galego. O castelán antigo si ten a voz *alfónde/iga* 'pousada' e o castelán actual ten *alhóndiga*, que é un lugar para a venda e a compra do gran.

A partir da raíz árabe *{hdr}*, co sentido de 'falar demasiado e sen sentido', Corriente (1993: 81) propón a raíz **ḥaṭr* para o romandalusí, derivada doutra proposta **ḥṭr* para o árabe andalusí. Esta derivación require unha ultracorrección, para substituír os fonemas frecuentes /h/ e /d/ polos velarizados e menos frecuentes /ḥ/ e /ṭ/. Aínda que esta raíz non está documentada, explica algunhas formas fonolóxicas. Con todo, é difícil ver a relación semántica entre todos os derivados que reúne tal raíz. Corriente propón esta raíz **ḥaṭr/ *ḥṭr* como étimo para *androlla* e *(al)droga*. A palabra *androlla,* que designa un tipo de embutido, estaría formada pola raíz *ḥaṭr* mais o sufixo diminutivo *-ol*. *Androlla* non existe nin en castelán nin en portugués, pero aquel ten (*h)adrolla* que significa 'embuste'. Corriente (1999: 339) atribúe o galego *trola* á aférese da primeira sílaba de *(h)adrolla*. O problema coa derivación a partir de *ḥaṭr* + *-ol* é a orixe do *n* de *androlla*. A raíz **ḥṭr* non é o único étimo posible. O *Pequeno diccionario Xerais da lingua* (Navaza Blanco, 1997: 47) propón o latín *inductilia* como étimo de *androlla*. Aínda que esta proposta non explica a orixe do 'r', non só resolve o problema da orixe do 'n', senón que tamén é o étimo do francés *andouille* 'un tipo de embutido feito dos intestinos do porco' (*Nouveau petit Larousse*, Gillon *et alii*, 1968: 44), unha palabra que coincide semanticamente con *androlla*.

Ámboa (substantivo) é un recipiente grande para viño ou outros líquidos. O seu étimo andalusí *anbúba* derívase do árabe clásico *unbūbah* (Corriente 1999: 212). Por outra banda, eses étimos árabes non son os únicos posibles. O *Pequeno diccionario Xerais da lingua* (Navaza Blanco, 1997: 41) propón o latín **ambula* para *ámboa*. *Âmbula* en portugués é un vaso para gardar os santos óleos. O latín *ampulla* parece o étimo máis axeitado tanto para *âmbula* coma para *ámboa.*

	Campo semántico	Étimos (Corriente / étimo alternativo)	Castelán	Portugués
Do latín				
ámboa	vasillas	< and *anbúba* < ac *unbūbah*/ < lt. *ampulla*	tinaja	jarro; pote
androlla	alimentos	< ra. **ḥaṭr* < and. *ḥṭr*/ < lt. *inductilia*	cierto embutido	tipo de enchido

	Campo semántico	Étimos (Corriente / étimo alternativo)	Castelán	Portugués
Evolucións dentro do galego				
alcroque	plantas	< ra. *al+ qoroqo* < lt. *croccus*	digital; dedalera	digitalis; dedaleira
alfóndega	peixes	< and. *xuṭṭáfa* < ac. *xuṭṭāfah*/ ac. *funduq*	escacho; bejel	cabra-cabaço
chuchamel	plantas	< ra. *čúč* < lt. *sūgere*	chupamieles	língua-de-vaca
desaxacado	defectos	< and. *šákl*	desaliñado	desleixado; descuidado

Táboa 2. Vocábulos exclusivos do galego no *Diccionario de arabismos* de Corriente: do latín e evolucións dentro do galego

3.2. Romandalusismos

Os romandalusismos figuran na táboa 3. *Tarabito* e *alfiestra* eran orixinalmente palabras latinas prestadas ao árabe andalusí e adaptadas. Segundo Corriente (1999: 453), o galego (adxectivo) *tarabito* 'estrábico' é unha adaptación do árabe andalusí *talábri,* que procede do látin *terebra* 'foret, vrille, tarière' (Bornecque, & Cauët 1938: 491), é dicir, 'unha broca'. Resulta difícil ver a relación semántica entre *terebra* e *talábri*. Simonet (1888: 522) recolle *talabrí/ talábri* no seu *Glosario de voces ibéricas y latinas usadas entre los mozarábes*, un título que suxire que o vocábulo é romandalusí. Simonet propón o étimo latino *strabus*, cuxa forma coincide con *talábri* e *tarabito* mellor que con *terebra*. Con todo, segundo as pescudas no *Dicionario de dicionarios*, unicamente o *Glosario de voces galegas de hoxe* (García González, 1985) recolle *tarabito*[2]. Como *tarabito* ten un reflexo irmán no romandalusí é posible que proveña desa lingua.

	Campo semántico	Étimos (Corriente / étimo alternativo)	Castelán	Portugués
aldroga	doenzas	*al* + ra. **ḥaṭr+ óqa*	habón, roncha	pápula
alfiestra	costura	< ra. *al* + latín *fenestra*	ojo de la aguja	buraco da agulha
tarabito	doenzas	< and. *ṭalábri*, lt. *terebra* / < lt. *strabus*	estrábico	estrábico

Táboa 3. Vocábulos exclusivos do galego no *Diccionario de arabismos* de Corriente: romandalusismos

Alfiestra (substantivo) é o ollo dunha agulla. Corriente (1999: 162) propón unha orixe romandalusí a partir do latín *fenestra* 'fiestra' mais o artigo definido. *Alfiestra*

2 http://sli.uvigo.es/DdD/ddd_pescuda.php?pescuda=tarabito&tipo_busca=lema (Consultado en 06.11. 2015).

pasa por un proceso de adaptación moi parecido ao de *alcroque*, pero a diferenza deste, non ten ningún vocábulo irmán no castelán.

O único elemento romandalusí visible en *alporca* (substantivo) é o artigo *al*, pero segundo Corriente (1999: 206), a raíz románica *porca* é un calco semántico do árabe *xinzīrah* 'porca'. *Alporca* é a escrófula, unha doenza infecciosa que produce a inchazón dos ganglios linfáticos sobre todo no pescozo. O mesmo nome *escrófula* derívase do latín *scrofa* mais o sufixo diminutivo *-ula* para dar *scrofula* 'femia do porco' (Coromines, 2008: 225). Aínda que Corriente (1999: 206) indica a presenza de *alporca* unicamente no galego, este termo non é exclusivo desta lingua, xa que o *Priberam* o recolle coa mesma acepción, *escrófula*,[3] e polo tanto non figura nas táboas. Con todo, non parece seguro que *alporca* sexa arabismo, ben pode proceder do mesmo latín.

Corriente tamén cita a raíz **ḥaṭr/ *ḥṭr* para *(al)droga* 'unha erupción cutánea', unha acepción díficil de relacionar con **ḥṭr*. Como *alfiestra, alporca* e *alcroque*, o artigo *al* salienta tamén no vocábulo *aldroga*, pero aquí hai que afirmar unha orixe romandalusí do mesmo, xa que se fose árabe, habería asimilación coronal da consoante *d*.

3.3. Arabismos

	Campo semántico	Étimos (Corriente / étimo alternativo)	Castelán	Portugués
Arabismos				
alcouve	plantas	< and. *al-ḥáwḍ*; ac. *ḥawḍ*	semillero, almáciga	sementeira, berçário
alferga	costura	< and. *ḥílqat alxiyáta*	dedal	dedal
brico	vasillas	< and. *ibríq* < ac. *ibrīq*	vaso de barro	vaso de barro
cáceres	restrición	< and. *kásr* < ac. *kasr* /< árabe *qassara*	muescas para las coyundas del yugo	entalhes para os paus do jugo
ceibe	restrición	< and. *sáy(i)bah* < *sáy(i)b* < ac. *sāʔib*	libre / suelto	libre / solto; o acto de ceivar; de soltar
nezcre	costura	< and. *názq*	lanzadera	lançadeira
ñácaras	defectos	< ac. *nakirah*	homúnculo	homúnculo
Berberismo				
túzaro	defectos	< ber. *Tūzar*	insociable, travieso	insociável, travesso

Táboa 4. Vocábulos exclusivos do galego no *Diccionario de arabismos* de Corriente: arabismos e berberismo

3 http://www.priberam.pt/dlpo/alporca (Consultado en 05.11.2015).

Preséntase unha serie de vocábulos na táboa 4 que teñen formas que mostran máis influencia árabe que románica. Corriente non indica étimos romandalusís para os vocábulos discutidos nesta sección. *Alcouve* (substantivo), 'sementeiro', provén do árabe andalusí *alháwḍ* no que tamén se pode ver o árabe clásico *ḥawḍ* 'balsa' (Corriente, 1999: 144). *Alcouve* non se rexistra no español moderno, pero o reflexo *alholde* figura no navarro; é unha medida.

Alferga (substantivo), 'dedal', é arabismo exclusivo do galego (Baldinger, 1971: 63) que se deriva do árabe andalusí *ḥilqat alxiyáta* 'anel de costura'. Tamén do campo semántico da costura, *nezcre* (substantivo) é 'lanzadeira', é dicir, unha peza dun tear ou dunha máquina de coser na que se enlea o fío. O seu étimo árabe andalusí é *názq* 'lanza', que dá o infinitivo manchego *añasquear* 'arar superficialmente'.

Brico (substantivo) é un vaso de barro que pode proceder do árabe andalusí *ibrique*, que á súa vez se deriva do persa. *Ibrique* úsase no árabe moderno; é o cacharro no que se prepara o café turco ou a auga para facer a ablución.

Ceibe (adxectivo) existe en portugués na forma *ceive*, e aínda que as dúas voces comparten o mesmo campo semántico, o portugués *ceive*[4] é substantivo que significa 'o acto de ceivar, de soltar as águas', pero no galego, *ceibe* e *ceibo* son sinónimos que significan 'libre' e 'solto'. Derívanse do árabe andalusí *sáy(i)b*, que vén do árabe clásico *sāʔib* 'solto, libre'. Inclúese *ceibe* na lista a causa da diferenza na categoría léxica entre o galego e o portugués, pero o máis probable é que esta voz formase parte da masa léxica compartida polo galego e o portugués. Corriente descarta o étimo latín *celibe* porque o árabe andalusí *sáyib* significa tanto 'libre' coma 'solto', igual que o galego *ceibe*. A proposta etimolóxica do *Pequeno diccionario Xerais da lingua* (Navaza Blanco, 1997: 156), *ceibar* < o latín *sipare* 'derramar, tirar, esparcir' parece razoable, pero o étimo árabe *sáy(i)b* coincide mellor semanticamente co derivado galego.

Cáceres (substantivo) (Corriente, 1999: 267) é arabismo exclusivo do galego que provén do árabe andalusí *kásr* 'corte'. Os *cáceres* son as entradas das cangalleiras (paus) do xugo. O verbo árabe *qassara*, 'acurtar ou limitar o movemento dun animal', parece outro étimo posible (Thomas, & Sayahi, 2005:10). *Cáceres* non se atopa no portugués nin no castelán, a excepción do nome da cidade *Cáceres*.

Ñácaras (substantivo), 'homúnculo', aseméllalle ao árabe *nakirah* 'descoñecido'. Este vocábulo non está documentado no árabe andalusí, e de feito, figura como o único vocábulo neste estudo cun étimo do árabe clásico (Corriente, 1999: 407). O *Diccionario de la lengua española* documenta *ñácara* no español de Honduras

4 http://www.priberam.pt/dlpo/ceive (Consultado en 05.11.2015).

coa acepción, 'chaga na pel dunha persoa',[5] e cita o seu étimo como variante de *chácara*, polo tanto cónsiderase dunha orixe diferente do galego *ñácaras*.

O vocábulo *túzaro* é un adxectivo que significa insociable ou un substantivo que significa unha persoa insociable ou bruta. É único neste estudo xa que é probable que veña do topónimo bérber *Tūzar* (Corriente, 1999: 462). *Túzaro* non está nin no portugués nin no castelán. As palabras nas columnas *Castelán* e *Portugués* da táboa 4 tentan definilo nesas linguas; non son os vocábulos correspondentes.

4. Discusión

Un dos perigos inherentes de todo traballo léxico é a vixencia dos vocábulos. Os dezasete vocábulos estudados aquí están incorporados no *Dicionario de arabismos* de Corriente e no *Dicionario de dicionarios*. Con todo, isto non quere dicir que se empreguen con regularidade na fala cotiá, e *tarabito, alcroque, desaxacado, ñácaras, brico* e *cáceres* nin sequera figuran no *Dicionario da Real Academia Galega* (Real Academia Galega, 2014). No caso de *alfiestra*, pode ser que se use a expresión *ollo dunha agulla* con máis frecuencia. Segundo as pescudas realizadas no *Dicionario de dicionarios*, os seguintes vocábulos están recollidos en polo menos nove dicionarios: *alferga, alfiestra, ámboa, alcroque*, e *túzaro*. Aínda que *ceibe* está rexistrado en só catro obras, o seu irmán, *ceibo*, atópase en dez dicionarios. Con todo, o número de dicionarios que recolle un determinado vocábulo tampouco proba a súa vitalidade, porque, entre outras razóns, algúns dos dicionarios citados no *Dicionario de Dicionarios* son revisións de obras anteriores. Un segundo perigo inherente de todo traballo léxico é afirmar un determinado étimo como definitivo, polo tanto, tratouse de ofrecer alternativas neste traballo. Malia estas advertencias, e ante unha posible lista de palabras que usa pouca xente e unhas propostas etimolóxicas non definitivas, este estudo pode dicir algo sobre o contacto lingüístico entre o galego e o árabe andalusí e o romandalusí, porque estes vocábulos son coma fósiles que dan testemuño dunha situación de contacto entre linguas no pasado.

Optouse por escoller vocábulos exclusivos do galego para chegar a ver indirectamente o contacto entre galegofalantes e arabefalantes sen a interferencia do castelán e do portugués, porque estas linguas poden emprestarlle arabismos ao galego. Á hora de normativizar o galego, tomouse o portugués como referente e escolléronse solucións portuguesas, que, nalgúns casos, efectivamente eran arabismos, por exemplo, *garfo*, en vez do castelanismo *tenedor*. Aínda que se supuña que o conxunto de arabismos galegos estaba no portugués, este estudo deixa ver uns poucos vocábulos cuxa traxectoria é distinta. A ausencia destes dezasete vocábulos no castelán e portugués

5 http://dle.rae.es/?id=QjzCJlE (Consultado en 06.11.2016).

actuais non quere dicir que non formasen parte do léxico desas linguas no pasado, pero suxire a posibilidade. A exclusividade no léxico tampouco é o único indicador dun arabismo único e directo. Este estudo empregou estes criterios: a independencia semántica, a categoría léxica e a forma fonolóxica, sobre todo se unha das linguas substituíu un fonema por outro. Este estudio, por exemplo, non recolleu o vocábulo galego *acea* porque está no castelán (*aceña*) e no portugués (*azenha*) co mesmo significado que *acea*. A diferenza radica na ausencia do fonema /ɲ/. É case seguro que se producira debido ao contacto entre o galego e o árabe (Mariño Paz, 1999: 73), pero non é exclusivo do galego. Por outra banda, algúns vocábulos recollidos aquí figuran no castelán e no portugués, pero con acepcións, formas fonolóxicas ou categorías léxicas moi diferentes das dos vocábulos galegos. Un reflexo de *ceibe* (adxectivo) emprégase no portugués (*ceive*, substantivo) aínda que é doutra categoría léxica que o seu irmán galego. Reflexos de *alcroque*, *alfóndega* e *chuchamel* encóntranse no castelán actual. Inclúense neste estudo porque as formas fonolóxicas de *alcroque* e *chuchamel* son distintas das castelás *alcroco* e *chupamiel*. Tamén, *alcroque* e *alcroco* son plantas diferentes. *Alcroque* pasou por unha evolución semántica, porque o seu étimo significa *azafrán* no romandalusí. Rexístrase *alfóndega* porque o significado do castelán *alfónd(e/i)ga/ alhóndiga* non se refire ao peixe. Este estudo conclúe que *alcroque* e *chuchamel* son evolucións dentro do galego, igual que *alfónd(e/i)ga*, aínda que a forma fonolóxica desta ten raíz árabe.

Unha clasificación semántica pode botar luz sobre o proceso de transmisión dos arabismos ao galego e así, o contacto entre a lingua fonte e a lingua receptora, neste caso, o galego. Represéntanse seis categorías semánticas: a costura, *alferga, alfiestra* e *nezcre*; as plantas, *alcouve, alcroque* e *cuchamel*; as doenzas ou defectos, *aldroga, desaxacado, ñácaras, tarabito,* e *túzaro*; as vasillas, *ámboa* e *brico*; e a restrición de movemento ou da liberdade dun animal ou persoa, *cáceres* e *ceibe*. Os vocábulos que quedan, *alfóndega* e *androlla*, corresponden a un peixe e a un tipo de embutido, respectivamente.

Precísase tamén una clasificación das linguas de orixe dos elementos que compoñen cada palabra. Algunhas posúen unha raíz latina que entrou no árabe andalusí, pasou ao romandalusí, e despois, entrou no galego, e por tanto, o único elemento árabe visible é o artigo *al* xunto cunha raíz latina que adquiriu un novo significado durante esta traxectoria. Ás veces o artigo *al* non é árabe senón romandalusí. Os vocábulos *alcroque, aldroga, alfiestra, chuchamel, desaxacado* e *tarabito* mostran un alto grao de romanización, e menos arabización, e non entran no galego procedentes do árabe, tampouco do árabe andalusí, senón que algúns veñen do romandalusí e outros son solucións galegas. Non son arabismos directos. O latín *ampulla* é un posible étimo para *ámboa*, e *inductilia* para *androlla*. A forma fonolóxica de *alfóndega* pode provir do árabe andalusí ou do romandalusí, pero o galego transformou moito o seu significado, por iso non é romandalusismo e tampouco arabismo. Os demais

mostran máis arabización, e as súas formas lingüísticas proveñen do árabe andalusí: *alcouve, alferga, brico, cáceres, ceibe, nezcre*. *Ñácaras* non está documentado no andalusí, senón no árabe clásico. Como a forma de *ñácaras* mantén unha relación moi fiel co étimo árabe, resulta probable que sexa un arabismo directo. *Túzaro* é máis ben un berberismo.

Non se pode descartar a presenza de falantes de árabe na Galiza actual e entre os ríos Miño e Douro, pero o certo é que tampouco se pode contemplar un substrato apreciable de arabefalantes como o que se pode identificar nos territorios máis meridionais ao avanzar o galego-portugués cara ao sur durante a reconquista. Deste xeito, é máis probable que os arabismos entrasen no galego mediante o contacto con mozárabes durante a súa migración para o norte. Os mozárabes falaban unha variedade románica, chamada aquí romandalusí, en contacto co árabe durante séculos, e por iso, contaba con arabismos que os seus falantes transmitían ás demais linguas peninsulares. Os empréstimos máis románicos indicados anteriormente ao mellor entraron no galego a través do romandalusí, xa que están documentados nesa lingua. Os demais, é dicir, os que mostran máis arabización, non están documentados no romandalusí. A falta de documentación non proba a súa ausencia neste, pero como existe documentación de posibles étimos desas voces no árabe andalusí, non é imposible imaxinar a súa entrada no galego mediante contacto con mozárabes bilingües en árabe e romandalusí. Ao comparar os campos semánticos coa clasificación etimolóxica, vemos que dous vocábulos que representan ideas asociadas coa limitación da liberdade, *cáceres* e *ceibe*, e as voces relacionadas con defectos persoais, *ñácaras* e *túzaro*, forman parte da lista dos étimos máis fieis á súa lingua orixinal, o árabe. Como se sabe, relaciónase a presenza de árabes na Galiza medieval con incursións militares coma a de Almanzor en 977 (Pallares, & Portela, 1996: 182-183) e coa entrega de escravos andalusís (Cunha, apud Frías Conde, 2000: 167). Así, é posible que os arabismos entrasen durante estes eventos violentos. Por outra banda, as palabras do campo da costura *alferga* e *nezcre*, e das vasillas, *brico*, poderían explicarse a partir de contacto cultural cos andalusís. Do mesmo xeito, é verosímil a incorporación de *alcouve* porque os andalusís prestaron tanto a palabra coma a práctica de criar plantas desa maneira. Estes últimos casos son exemplos da entrada dun novo elemento cultural, por exemplo, unha vasilla ou unha nova práctica agrícola, acompañado do nome ao que se refire na lingua orixinal.

5. Conclusión

Do conxunto dos arabismos recollidos en Corriente (1999), figuran uns dezasete en galego sen irmáns nas demais linguas peninsulares. Así e todo, a maioría destes vocábulos vén do romandalusí ou resulta dunha evolución dentro do galego. Destes dezasete, unicamente oito poden proceder directamente do árabe ou dunha lingua

bérber, xa que non están documentados no romandalusí. Polo tanto, non se pode descartar situacións de contacto social entre falantes do galego e do árabe andalusí e dunha lingua bérber. A comparación da etimoloxía de cada empréstimo co seu campo semántico achega unha explicación da entrada destes vocábulos no galego en función duns eventos socio-históricos documentados.

Referencias bibliográficas

Augé, Gillon *et alii* (1968). *Nouveau petit Larousse*. París: Librairie Larousse.

Baldinger, Kurt (1958/1971[2]). *La formación de los dominios lingüísticos en la Península Ibérica*. Madrid: Gredos. Orixinal: *Die Herausbildung der Sprachräume auf der Pyrenäenhalbinsel. Querschnitt durch die neueste Forschung und Versuch einer Synthese*. Berlin: Akademie-Verlag, 1958.

Bornecque, Henri, & Cauët, Fernand (1938). *Le dictionnaire Latin-Français du baccalauréat*. Paris: Librairie Classique Eugène Belin.

Coromines, Joan (2008). *Breve diccionario etimológico de la lenga castellana.* Madrid: Gredos.

Corriente, Federico (1993). "Reflejos iberromances del andalusí {ḥṭr}", *Al-Andalus Magreb I*, 77-87.

Corriente, Federico (1996). "Novedades en el estudio de los arabismos en iberorromance", *Revista española de lingüística, 26:1*, 1-13.

Corriente, Federico (1998). "Arabismos y otras voces de origen semítico o medio-oriental en las hablas aragonesas y en gallego", *Romance Philology, 51*, 421-474.

Corriente, Federico (1999). *Diccionario de arabismos y voces afines en iberorromance*. Madrid: Gredos.

Dicionário Priberam da Língua Portuguesa (2008-2013) (edición electrónica). http://www.priberam.pt/dlpo (Consultado en 05.11.2015).

Franco Grande (1972[2]). *Diccionario galego-castelán*. Vigo: Galaxia. Consultado a través de Santamarina, Antón (coord.) (2006-2013).

Frías Conde, Xavier (2000). "O elemento árabe en galego (I)", *Revista galega de filoloxía, 1*, 157-171.

Frías Conde, Xavier (2002). "O elemento árabe en galego (II)", *Revista galega de filoloxía, 3*, 65-80.

García González, Constantino (1985). "Glosario de voces galegas de hoxe". Anexo 27 de *Verba*. Santiago de Compostela: Universidade. Consultado a través de Santamarina, Antón (coord.) (2006-2013).

Kiesler, Reinhard (2003). "¿Hay más arabismos en español o en portugués?". En Sánchez Miret, Fernando (ed.), *Actas del XXIII Congreso Internacional de Linguística y Filología Románica*, vol. 3, 263-266. Berlin: De Gruyter.

Losada, E., Castro. J., & Niño, E. (1992). *Nomenclatura vernácula da flora vascular galega*. Xunta de Galicia. Consultado a través de Santamarina, Antón (coord.) (2006-2013).

Machado, José Pedro (1991). *Vocabulário português de origem árabe*. Lisboa: Editorial Notícias.

Mariño Paz, Ramón (1999[2]). *Historia da lingua galega*. Santiago de Compostela: Sotelo Blanco Edicións.

Monteagudo Romero, Henrique, & García Cancela, Xermán (1988). *Diccionario normativo galego-castelán*. Vigo: Galaxia.

Navaza Blanco, Gonzalo (coord.) (1997). *Pequeno diccionario Xerais da lingua*. Vigo: Edicións Xerais de Galicia.

Pallares Méndez, María del Carmen, & Portela Silva, Ermelindo (1996). "Idade media". En Vázquez Varela, Xosé Manuel *et alii*, *Nova historia de Galicia*, 147-234. Oleiros: Editorial Tambre.

Piel, José M. (1965). "Caractères généraux et sources du lexique galicien". En *Actes du Xème Congrès International de Linguistique et Philologie Romanes (Strasbourg, 1962)*, vol. 3, 1261-1267. Paris: Klincksieck.

Poplack, Shana, Sankoff, David, & Miller, Chris (1988). "The social correlates and linguistic processes of lexical borrowing and assimilation", *Linguistic*s*, 26*, 47-104. DOI: http://dx.doi.org/10.1515/ling.1988.26.1.47.

Real Academia Galega (2014). *Dicionario da Real Academia Galega*. A Coruña: Real Academica Galega (edición electrónica). http://www.realacademiagalega.org/dicionario (Consultado en 05.11.2016).

Real Academia Española (2014). *Diccionario de la lengua española*. Real Academia Española (edición electrónica). http://www.rae.es (Consultado en 05.11.2016).

Santamarina, Antón (coord.) (2006-2013): *Dicionario de dicionarios: Corpus lexicográfico da lingua galega*. Seminario de lingüística Informática, Grupo TALG / Instituto da Lingua Galega) (edición electrónica). http://sli.uvigo.es/DdD/ (Consultado en 06.11.2015).

Simonet, Francisco Javier (1888). *Glosario de voces ibéricas y latinas usadas entre los mozarábes*. Madrid: Establecimiento tipográfico de Fortamets.

Thomas, Juan A., & Sayahi, Lotfi (2005). "Arabisms in Galician", *Cahiers linguistiques d'Ottawa, 33*, 23-42.

Thomason, Sarah G., & Kaufmann, Terrence (1988). *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.

Notas sobre toponímia portuguesa medieval

Maria Alice Fernandes
Universidade do Algarve

Esperança Cardeira
Universidade de Lisboa

Resumo:

Usando como fonte o *Reportório Toponímico de Portugal Continental* (1967) e o programa de georreferenciação toponímica *NOTVS* (Silva 2002) como ferramenta de trabalho, propomo-nos traçar um esboço da paisagem linguística medieval, através da identificação das áreas de implantação toponímica das línguas faladas durante a Reconquista no futuro território português (nomeadamente o árabe e o romance andalusis[1] e o galego-português) e da relação entre essa implantação e a geografia humana e ocupação do espaço.

Palavras-chave:

Toponímia, galego-português, Reconquista e repovoamento.

Sumário:

1. Introdução. 2. Tipologia da toponímia portuguesa. 3. Reconquista e implantação toponímica. 4. Toponímia dos assentamentos rurais. 4.1. Quintã/s. 4.1.1. Quintães. 4.1.2. Quintão/s. 4.1.3. Quinta/s. 4.1.4. Quintela/s. 4.1.5. Quintinha/s. 4.2. Casal/ais – Casares. 4.2.1. Casarelho/e. 4.2.2. Casarinhos – Casainho/s – Casalinho/s – Casaizinhos. 4.2.3. Casalito – Casaleco. 4.3. Aldeia/s. 4.3.1. Aldeinha – Aldeola - Aldeota. 4.4. Corte/s. 4.4.1. Cortelha – Cortelho/s. 4.4.2. Cortinha – Cortinhos. 4.5. Caria/s – Alcaria. 4.5.1. Alcarela – Alcariota. 5. Topónimos com origem em nomes de árvores: (MALA) MATTIANA e derivados. 6. Conclusões.

Notes on Portuguese medieval toponymy

Abstract:

In this paper, we will draw an outline of the medieval linguistic landscape by identifying the toponymic implantation of Galician-Portuguese, Arabic and Mozarabic languages in

1 Nos dois primeiros séculos do domínio árabe da Hispânia, os dialetos românicos e os árabes coexistiram numa situação de bilinguismo, razão pela qual a denominação de *romance moçárabe* (que restringe o seu uso aos cristãos que viviam no *Alandalús*) é aqui substituída pela de *romance andalusi* à semelhança de *romance andalusí* ou *romandalusí* (Corriente 2003).

Portuguese territory and relating those toponyms with human geography and the occupation of space. Source: Reportório Toponímico de Portugal Continental (1967).

Key words:

Toponymy, Galician-Portuguese, 'Reconquista' and repopulation.

Contents:

1. Introduction. 2. Typology of the Portuguese toponyms. 3. Toponymic implantation and the 'Reconquista'. 4. Toponymy of rural settlements. 4.1. Quintã/s. 4.1.1. Quintães. 4.1.2. Quintão/s. 4.1.3. Quinta/s. 4.1.4. Quintela/s. 4.1.5. Quintinha/s. 4.2. Casal/ais – Casares. 4.2.1. Casarelho/e. 4.2.2. Casarinhos – Casainho/s – Casalinho/s – Casaizinhos. 4.2.3. Casalito – Casaleco. 4.3. Aldeia/s. 4.3.1. Aldeinha – Aldeola – Aldeota. 4.4. Corte/s. 4.4.1. Cortelha – Cortelho/s. 4.4.2. Cortinha – Cortinhos. 4.5. Caria/s – Alcaria. 4.5.1. Alcarela – Alcariota. 5. Toponyms originated from tree names: (MALA) MATTIANA and derived forms. 6. Closing remarks.

1. Introdução

Os estudos sobre toponímia tiveram, em Portugal, um significativo desenvolvimento na primeira metade do século XX: as revistas que se dedicavam à História ou à Filologia ofereciam, quase sempre, um artigo sobre toponímia de José Leite de Vasconcelos, Joaquim da Silveira, José Joaquim Nunes, Joseph-Maria Piel, Cunha Serra e outros autores que se dedicaram à investigação sobre o onomástico. Uma parte significativa do trabalho de Piel consiste no estudo de topónimos e antropotopónimos, estudo que se destinaria, nas palavras do autor, "a fazer parte de um trabalho de conjunto sobre a toponímia de Portugal e do NO da Península" (Piel, 1945: 305-6). Esse trabalho, que seguiria a linha de Menéndez Pidal, usando a toponímia como ferramenta para a reconstituição de estádios de língua, nunca chegou a ser completado. Uma observação superficial do conjunto de estudos sobre toponímia portuguesa publicados no século passado aponta para uma acentuada diminuição a partir dos anos 70 e, mais grave ainda, uma quase ausência nos primeiros 15 anos do século XXI. Impõe-se o retomar do interesse pela toponímia portuguesa: possuímos um conjunto significativo de trabalhos realizados, que são, em geral, monografias, cada uma constituindo uma pequena unidade, desarticuladas entre si. O primeiro passo seria reunir todos estes trabalhos numa mesma base de dados e criar um *corpus* toponímico com base em documentação medieval portuguesa, por forma a proceder à sua análise, cartografia e interpretação.

Uma análise que nos parece particularmente interessante é a que relaciona a toponímia com a geografia humana e a ocupação histórica do espaço, a que, por isso, damos continuidade (Fernandes, & Cardeira, 2013). Esta abordagem, de natureza sociolinguística, permite-nos uma interpretação das mudanças linguísticas atestadas pelos topónimos que a unicamente estrutural não logra atingir, por não conseguir explicar os condicionamentos sociais da fixação toponímica. A perspetiva aqui considerada parte da transplantação para sul do galego-português, na esteira da Reconquista cristã do *Garbe Alandalús*[2] e tem em conta a correlação entre variáveis linguísticas e sociais. As linguísticas consistem nas mudanças induzidas pelo contacto dialetal entre colonos de diferentes proveniências regionais e, dependendo da etapa da conquista, também com falantes do romance andalusi; as variáveis sociais constam dos lugares e espaços ocupados com a colonização interna astur-leonesa e galego-portuguesa. A análise da toponímia medieval de que nos vamos ocupar tem, assim, em conta a correlação entre as soluções históricas atestadas pelos topónimos e a sua distribuição geográfica nos espaços históricos dos territórios políticos portugueses.

2 Adotamos a expressão *Garbe Alandalús*, que, segundo Ángeles Vicente (2006), testemunha a pronúncia coeva do domínio islâmico.

O objeto de estudo consiste na toponímia do povoamento rural, nomeadamente na fixada com base nos nomes andalusis (árabes e românicos) e galego-portugueses de propriedades com habitação e de povoados. Servimo-nos, como fonte, do *Reportório Toponímico de Portugal Continental* (1967) dos Serviços Cartográficos do Exército e, como ferramenta de trabalho para a cartografia dos dados, do *NOTVS*, um programa informático de análise e georreferenciação toponímica desenvolvido pelo geógrafo histórico Luís Fraga da Silva com a colaboração de Maria Alice Fernandes (Silva, 2002). Significa isto que a análise que iremos apresentar se baseia nos topónimos medievos remanescentes, as atestações toponímicas coevas servindo-nos apenas como fator validador. Por isso, não podemos dizer que se trata de um ensaio sobre toponímia medieval, mas tão só de notas sobre o assunto, à espera de um levantamento documental exaustivo que as enquadre devidamente. Todavia, e seguindo as leis da paisagem histórica formuladas pelo medievalista Jordi Bolòs (2010), a sobrevivência de topónimos medievais até à atualidade é um testemunho da continuidade do seu uso referencial pelos falantes e, acrescentamos nós, da sua continuidade ocupacional no caso dos lugares com origem em assentamentos rurais com *habitat*, nomeadamente propriedades e povoações.

2. Tipologia da toponímia portuguesa

No estudo da toponímia de Portugal podem considerar-se fundamentalmente os seguintes tipos de nomes de lugar, consoante a sua origem[3]:

1 – os nomes pré-romanos, como *Braga* ou os hidrónimos *Ave*, *Lima*, *Minho*;

2 – os nomes comuns, de origem parcial ou totalmente latina ou romance, que aludem a aspetos físicos ou sociais do território, como *Porto* ou *Bustelo*;

3 – os nomes de 'possessores' da primeira fase da Reconquista, na sua maioria de origem germânica, do tipo *Guimarães* ou *Gondomar*;

4 – os topónimos de origem árabe, relacionados com aspetos da paisagem, com o tipo de povoamento e com as atividades económicas. Estão neste caso formas como *Albufeira*, *Alcaria*, *Alcântara* e *Alferce*. A estes topónimos devem juntar-se os devidos aos assentamentos clânicos, de origem árabe, berbere ou românica, construídos com *Ben-* no singular ou *Bena-* em variação com *Bela-* no plural, que se tornaram topónimos: *Benafim*, *Benagil*, *Benfarras*, *Bencatel*, (1291 Benaçalema >) *Bela Salema* (Fernandes, 2001).

5 – os topónimos românicos andalusis, como *Messines* e *Moncarapacho*, ou os híbridos, como *Alfontes*, *Alportel*, *Cacela*, *Loulé*, *Monchique* e outros (Fernandes,

3 Outras classificações são possíveis, nomeadamente aquelas que se baseiam na distribuição dos topónimos por campos semânticos (fauna, flora, caraterísticas físicas do território, etc.).

2007). Embora não seja seguro que a Reconquista tenha encontrado variedades do romance andalusi com vitalidade nos séculos XII-XIII, alguns traços comuns aos dialetos centro-meridionais só parecem poder ser explicados tendo em conta uma influência românica local, o que favorece a hipótese da persistência de populações falantes daquele romance (moçárabes, berberes e árabes).

Partindo de uma perspetiva sociolinguística histórica, os topónimos aqui tratados enquadram-se nos seguintes períodos:

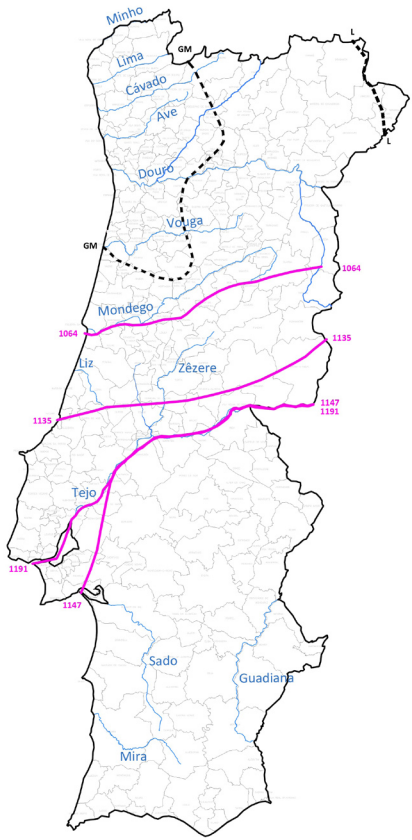
1. O período andalusi, que compreende o árabe e o romance andalusis, em situação de bilinguismo, (topónimos de origem árabe, moçárabe e híbridos resultantes do contacto linguístico);
2. O galego-português (a transplantação da toponímia desde a Galécia Magna até ao Algarve pode contribuir para o esclarecimento da difusão geográfica do galego-português a sul do Vale do Vouga e para a caraterização das formas de ocupação e organização do território);
3. O período português propriamente dito, representado na toponímia fixada a partir do início do português médio (século XV), e que permite verificar a criação de novas formas e a substituição das antigas.

3. Reconquista e implantação toponímica

A conquista cristã deu-se por fases (com avanços e recuos) ao longo do espaço e do tempo. No mapa 1 (Áreas linguísticas e extremas da Reconquista), as linhas a tracejado indicam a área pertencente à Galécia Magna (o domínio primitivo do galego-português) e a área leonesa, o restante espaço correspondendo aos territórios de colonização linguística galego-portuguesa; as linhas a rosa representam as extremas territoriais relativas a cada fase da Reconquista (Barroca, 1999; Marques, & Dias, 2003): 987 diz respeito ao reposicionamento da fronteira no Douro, na sequência das campanhas militares de Almançor; 1064 representa a conquista definitiva de Coimbra por Fernando I de Castela e Leão, que posiciona a fronteira no rio Mondego; 1135 a construção do castelo de Leiria por Afonso Henriques, que assegura a colonização a sul de Coimbra e as campanhas militares que levam à conquista de Lisboa e de Santarém em 1147; 1191 o contra-ataque almóada que faz recuar o domínio português novamente para o Tejo, no reinado de Sancho I, quando parte do Alentejo e do Algarve estava já conquistada; 1249, a tomada de Faro, com a qual termina o processo de conquista.

A esse processo corresponde um movimento de colonização que irá transplantar o galego-português e com ele as denominações das propriedades rurais senhorial e camponesa *quintãa* e *casal*, próprias do noroeste hispânico e, bem assim, do arabismo andalusi *aldea*, na aceção de pequeno povoado rural, que também se irão fixar toponimicamente nos territórios conquistados. Essas formas irão encontrar-se

com os topónimos andalusis *Corte* e *Alcaria* e com os antropotopónimos clânicos referidos.



Mapa 1. Áreas linguísticas e extremas da Reconquista

Não temos dados coevos do período andalusi sobre a ocupação do espaço rural. A cultura islâmica era de natureza essencialmente urbana e mercantil e os geógrafos e biógrafos árabes fornecem escassas informações sobre a ruralidade. Por outro lado, as informações documentais de que dispomos para o *Alandalús* não se podem aplicar ao *Garbe* sem o risco de distorções, em virtude de a sua localização geográfica periférica ter contribuído para formas de aculturação particulares em certas regiões, condicionadas pelo isolamento político, preservação da identidade hispânica ou inexistência de ulemas que transmitissem os ensinamentos do Corão e com ele as tradições culturais muçulmanas. São, assim, as áreas de implantação toponímica e

o pouco que sabemos sobre a organização sociopolítica coeva que proporcionam informações sobre a ocupação e a manipulação sociais do espaço físico durante o período andalusi. Para além disso, a expansão territorial cristã aumentou a rarefação populacional nas regiões fronteiriças muçulmanas e provocou movimentos migratórios que levaram ao despovoamento das zonas mais críticas, sobretudo a partir de meados do século XII. Nesse sentido, a distribuição da toponímia andalusi remanescente não pode deixar de ter em conta a reorganização social do espaço levada a cabo pelas monarquias cristãs. Não queremos com isto dizer que não possa ter havido outras reconstruções do espaço, mas o facto de os topónimos fixados no período andalusi terem sobrevivido até hoje é um indício da ocupação medieval dos lugares por eles nomeados. Assim sendo, começaremos a nossa análise pela toponímia galego-portuguesa difundida com a conquista cristã e as formas por ela adotadas no português propriamente dito, para podermos verificar quais as zonas de preservação da toponímia andalusi e a sua relação com a implantação da galego-portuguesa.

4. Toponímia dos assentamentos rurais

4.1. Quintã/s (90-93)[4]

Este topónimo tem como etimologia o lat. QUINTANA, originalmente um termo militar que designava a via que separava o quinto do sexto manípulo nos acampamentos romanos. No século VI, o termo é abonado nas *Etimologia*s de Isidoro de Sevilha com o significado de 'via carroçável'. Depois, na documentação latina do século X (Lapesa, & García, 2004), *quintana* designa a 'rua ou praça defronte da casa senhorial'[5], a própria 'casa senhorial' e a 'propriedade senhorial com casa de habitação e dependências agrícolas' (Fernandes, & Cardeira, 2013), num processo de extensão semântica. A razão desta evolução deve-se à morfologia da *pars urbana* da *villa*, cercada por um muro que se desenvolvia quadrangularmente em redor do pátio que lhe servia de entrada, as paredes das habitações (senhorial e servis) integrando frequentemente o próprio muro. É como identificação das honras inquiridas (*domus habitationis* e *dominicus*), que o nome comum e o topónimo ocorrem nas *Inquirições* do século XIII: *"*se em esta freguesia há casa de cavaleiro ou de dona que se defenda por honra – eis a pergunta sistemática [nas inquirições de 1288], que, no caso afirmativo, tem esta resposta: «há aí uma quintã», etc.*"* (Almeida Fernandes, 1978: 249). O termo era ainda usado no século XX, para designar o *eido* (< lat. ADITU) ou o *eixido* (<lat. EXITU), ou seja, a entrada ou saída da propriedade correspondentes à referida praça, no dialeto minhoto e como denominação da pocilga ou da estrumeira no beirão (Fernandes, & Cardeira, 2013).

4 Os números entre parênteses curvos representam as instâncias toponímicas.

5 Na época medieval, *praça* designava uma rua carroçável, como o era o pátio ou entrada das quintãs, que servia de mercado e de ponto de encontro para as exações fiscais.

O uso de *quintana* como designação da propriedade rural é comum aos romances setentrionais hispânicos, à exceção do castelhano. Na Galécia Magna meridional, correspondente ao Entre Douro e Minho, ocupou o lugar deixado vazio pela denominação romana *villa*, com a qual covariou durante a Alta Idade Média. A sua área focal pode ter sido as Astúrias, onde designou «una casería completa, com vivenda señorial, horro, cuadra, corripos, corripas, güertos, güertas…» (Concepción Suárez, 2007), ou seja, o domínio com habitação senhorial. Este nome estará, assim, associado à expansão territorial astur-leonesa, iniciada no Noroeste sob a forma de presúrias, isto é, de ocupação das terras em nome régio. Esta modalidade expansiva foi possível em virtude de a região a norte do Douro ter constituído um espaço marginal no *Alandalús* e ter sido destituída dos seus quadros civis e religiosos por Afonso I, aquando das suas campanhas militares no século VIII (Barroca, 2003). Dada a relevância da casa paçã no reconhecimento dos domínios senhoriais então criados, é natural que a denominação passasse a ser usada para identificar os próprios domínios. Na segunda metade do século XI, o termo é já transplantado para o Entre Douro e Mondego com os dois significados, com os quais se fixou também toponimicamente.

A forma galego-portuguesa evoluída de QUINTANA é *quintãa* (atestada desde o século XII: 1199 *quintáá*, 1220 *quintaa*, 1261 *quintãa*. Cf. Houaiss, & Salles, 2002, s.v. *quinta*), que resulta da síncope do -N- intervocálico específica daquele romance. Esta forma dará origem à portuguesa comum *quintã*, mediante a assimilação da nasalidade pela vogal átona final e a crase das vogais em hiato, e às dialetais *quintães* (plural de *quintã*) e *quintão* (v. adiante).

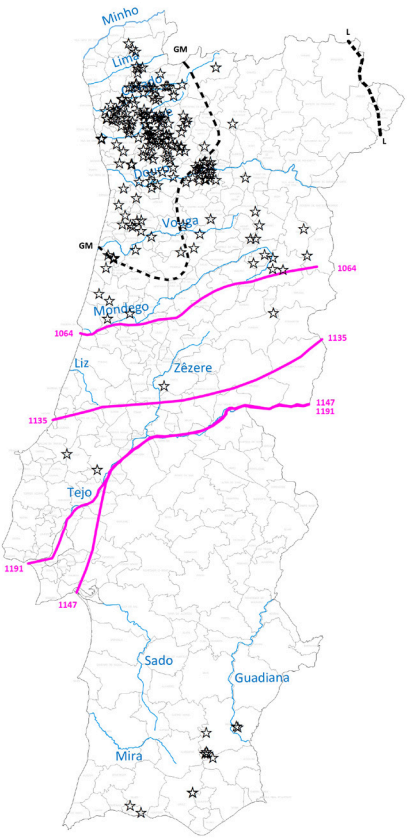
A área focal destas variantes na toponímia portuguesa é o Entre Douro e Minho, a área meridional da Galécia Magna. A variabilidade semântica e formal de *quintã* no Entre Douro e Minho reflete a manutenção linguística própria de sociedades antigas, com comunidades ligadas à terra por inúmeras gerações, caraterizadas pelas famílias alargadas e redes sociais complexas e densas (Milroy, 1992)[6].

Como referimos, a fixação toponímica de *Quintãa* inicia-se ainda na primeira fase da ocupação territorial asturiana (s. IX a XI), que chegou ao Mondego no reinado de Afonso III (866-910), com as presúrias do Porto em 868, de Chaves em 872 e de Coimbra em 878. Na sequência destas presúrias, dar-se-iam também as de Montemor-o-Velho, Viseu e Lamego. A fronteira cristã ocidental ficaria, então, posicionada, a sul, na margem esquerda do Mondego e, a leste, no vale do Côa. Os seus limites meridionais foram, no entanto, repostos no vale do Douro e, na zona costeira, ao longo

6 A rede social designa o conjunto dos falantes que interagem regularmente entre si. A densidade e complexidade da rede social advêm do tipo e do número de relações entre os falantes: quando as relações são fortes, as redes são densas; quando os falantes se conhecem uns aos outros nos seu diferentes papeis sociais (familiares e profissionais), as redes são complexas. Este tipo de redes é um dos fatores da manutenção das normas linguísticas.

do vale do Vouga em 987 pelos exércitos de Almançor, prefeito de Hisham II (Barroca, 1999; 2003). A ocupação cristã, que se concentrou, então, na zona costeira do antigo *conventus bracarensis*, procurou restaurar a organização eclesiástico-administrativa definida pelo reino suevo no século VI (Quiroga, & Lovelle, 1997).

A difusão geográfica entre o Douro e o Mondego corresponde já ao movimento expansionista senhorial nas Beiras Alta e Litoral, após a sua conquista por Fernando, o Magno, na segunda metade do século XI (Mattoso 1988). Com efeito, a geografia dos territórios cristãos a sul do Douro estender-se-ia nesta época a sul até à margem esquerda do Mondego, à zona de Soure e Penela, e a leste entre os rios Távora e Coa. Esta variante é depois difundida pelas conquista e colonização portuguesa até ao Algarve, o escasso número de instâncias toponímicas devendo-se, por um lado, à menor implantação senhorial a sul do Mondego e, por outro, à substituição de *quintã* pela variante comum *quinta*, que também se observou toponimicamente (V. adiante).



Mapa 2. Quintã/s - Quintães – Quintão/s

4.1.1. Quintães (24)

O plural *quintãs* conheceu uma variante mais tardia, resultante da confusão com os nomes da terceira declinação em -ANE, provavelmente devido ao processo de convergência no ditongo *õu* das terminações nasais *-ão* (-ANU), *-ã* (-ANE) e *-õ* (-ONE) observado no Minho e Douro Litoral[7]. Essa confusão está abonada pelo gramático quinhentista Fernão de Oliveira (1981) [1536]. Não encontramos testemunho toponímico de transplantação para fora da área original desta variante minhota.

7 O ditongo *-õu* consiste numa evolução da terminação *-õ* (< -ONE) e a convergência de *-ã* e *-ão* nesse ditongo explica-se pela sua frequência lexical, em muito superior à das outras terminações nasais. No Sul, a convergência em *-ão* pode ter-se ficado a dever ao processo de elaboração da língua.

4.1.2. Quintão/s (56/2)

Esta variante dialetal de *quintã*, também fixada toponimicamente, provém, segundo os autores que se ocuparam desta solução histórica, da variante alto-minhota e galega *quintáŋ,* por semivocalização da consoante nasal velarizada (Maia, 1981) ou da variante portuguesa *quintã*, por ditongação da vogal tónica em posição final absoluta (Martins, 1985). A solução histórica ditongada da terminação latina –ĀNA era, ainda no século passado, um traço dialetal minhoto, que, difundido com a colonização interna, veio a caraterizar outras variedades diatópicas (Cardeira, & Fernandes, 2008; Fernandes, & Cardeira, 2013). Devido ao seu estatuto dialetal, esta variante não tem registo escrito, sendo confundida com o aumentativo de *quinta*. Nos territórios de colonização linguística em que também se observa a ditongação daquela terminação latina nos nomes comuns, a preservação toponímica de *Quintão* é um argumento a favor da difusão deste traço dialetal a partir do noroeste, explicável pelo contacto com falantes com aquela proveniência.

4.1.3. Quinta/s (9307/192)

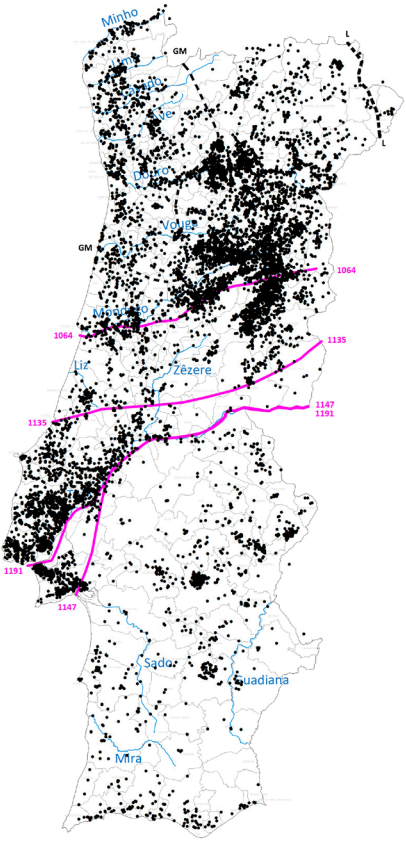
Trata-se de uma variante baixo-medieval de *quintã*, cuja evolução histórica não se encontra ainda esclarecida. Apesar de na toponímia do domínio leonês se atestarem as formas *Quintana* (< lat. QUINTANA) e *Quinta* (< lat. QUINTA), no caso português não há registo de *quinta* antes de finais do século XV, como se esperaria encontrar se este nome proviesse diretamente do latim QUINTA.

A solução histórica *quinta* é atribuída na literatura à dissimilação da ressonância nasal e retração do acento por Leite de Vasconcellos e José Joaquim Nunes, com possível intermediação da forma *quintá* para Williams (*apud* Rodrigues Lapa, 1932). Mas a verdade é que esta forma apenas se observa em galego, não havendo testemunho da sua ocorrência na variante portuguesa. Rodrigues Lapa prefere ver nesta evolução um fenómeno de analogia com o numeral *quinta* na aceção testamentária ou de fração em géneros ou numerário paga pelo rendeiro ao proprietário rural. O autor refere, em abono da sua tese, a mesma solução histórica em *ventana* > *ventã e campana* > *campã*, que teriam sido substituídas por *venta* e *campa* por analogia com os deverbais de *ventar* 'soprar com as ventãas' e *acampar* 'tocar a campãa', este último atestado na *Crónica de D. João I*.

A verdade é que estas variantes são tardias e meridionais. *Quinta* e *venta* atestam-se apenas no século XV (Fernandes, & Cardeira, 2013; Houaiss, & Salles, 2002), tendo sido adotadas na língua comum. De *campa* não encontrámos nenhuma atestação com o significado primitivo, o que não é de admirar se considerarmos que a forma erudita *sino*, que a substituiria, assim como o uso de *campã* (> *campa*), na aceção de 'laje sepulcral' (Houaiss, & Salles, 2002) se registam desde finais do século XIV. Da sua meridionalidade dá conta, para além da diferença relativamente às formas primárias

evoluídas diretamente da galego-portuguesa *quintãa*, a conservação e a vitalidade da formação de nomes deverbais nos dialetos portugueses do Sul, nomeadamente no alentejano e algarvio, uma caraterística da oralidade medieval.

Mas qual a razão da substituição das formas *quintã*, *campã* e *ventã* por *quinta*, *campa* e *venta* nesses dialetos? A resposta parece-nos ter de ser procurada nas mudanças linguísticas induzidas por contacto dialetal de longa duração provocadas pela conquista do Garbe. Com efeito, a transplantação do galego-português porá em contacto colonos com diferentes proveniências regionais e autóctones falantes do romance andalusi. Desenraizados, e, por isso, não sujeitos às imposições normativas das suas comunidades de pertença, os colonos acomodar-se-ão linguisticamente uns aos outros, o que levará à mistura dialetal e dará origem à formação de novas variedades dialetais meridionais galego-portuguesas por koineização. Este processo de mudança compreende, prototipicamente três gerações. Na primeira, para além da mistura de traços surgem, geralmente, as primeiras mudanças linguísticas por nivelação, i.e., por seleção dos traços mais frequentemente usados ou dos mais salientes. Na segunda, a dos filhos dos primeiros colonos, a solidariedade e a mobilidade social das novas comunidades em que se integram ou que procuram formar favorecem um aumento da variabilidade linguística, o que ativa a seleção e a realocação (i.e., a refuncionalização estilística ou social) de traços dialetais da mistura, bem como a criação de formas não existentes nas variedades originais em contacto. Na sua maioria devidos à aquisição imperfeita de dialetos segundos, estes interdialetalismos podem, na terceira geração, adquirir ou surgir em função da necessidade de identificação por contraste, pois é nesta etapa que ocorre a focalização dialetal que dá origem à nova variedade.



Mapa 3. Quinta/s

De entre essas mudanças, a simplificação, a nivelação e a realocação

podem explicar a adoção das inovações deverbais em detrimento das variantes arcaicas. A formação dos deverbais pode, com efeito, corresponder a uma inovação simplificadora, no sentido em que tem como resultado a regularização morfofonológica, pois de atemáticos os nomes adquirem o tema em *–a*, coincidente, ademais, com o morfema de género feminino. A competição entre as formas nasaladas e desnasaladas poderia, por sua vez, ter induzido a nivelação, a escolha das desnasaladas podendo ter sido condicionada pelo seu uso por um maior número de falantes ou pela sua saliência. A terceira mudança consistiria na realocação da variante sobrevivente à nivelação, mas cuja função/significado foi alterada. No caso de *quinta* e *venta* estaríamos, então, em presença de variantes niveladas e, no de *campa*, de uma variante realocada, cujo significado foi alterado para não competir com o de *sino*.

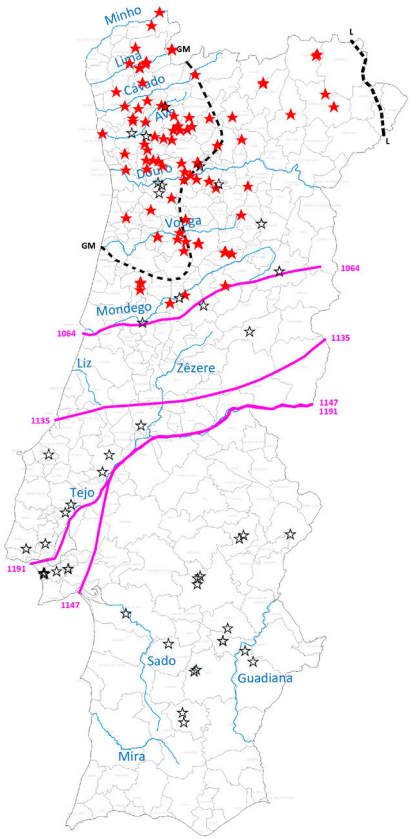
Durante o século XVI, a forma desnasalada *quinta* substitui a antiga *quintã*, que se conserva, essencialmente, na área setentrional. Fora dessa área, os topónimos com a terminação nasal devem-se à colonização interna galego-portuguesa que acompanhou ou se seguiu à Reconquista. A generalização das formas desnasaladas na língua comum terá de ser explicada, contudo, através da sua adoção pelo português médio (século XV). Efetivamente, esta variedade diacrónica escrita não só tem como base dialetal o português do sul, como consistiu na construção de um símbolo identitário próprio da nova dinastia reinante – a Casa de Avis, que se distanciou política e linguisticamente da velha nobreza nortenha. Neste sentido, a opção por estas inovações lexicais meridionais pode ter sido consciente e visado o distanciamento face às variantes setentrionais, portuguesa e galega (Cardeira, 2005).

4.1.4. Quintela/s (88/4)

Com base na forma latina QUINTANA criou-se o diminutivo QUINTANELLA (século X *quintanela*; XIII *quintaela*, *quinteela,* Machado, 1993, s.v. *Quintela*), na origem do topónimo galego-português *Quintela*, que ocupa uma área semelhante à de *Quintã*, estendendo-se ainda a Trás os Montes. O facto de o sufixo diminutivo ser ainda o latino-vulgar –ĔLLU (> gal.-port. *-elo*) é uma prova da toponimização de *quintãa* e de *quintaela* ainda durante a primeira fase da expansão territorial cristã, uma vez que esse sufixo entrará em variação com o galego-português *-ĩo* (< lat. –ĪNU), na origem do atual *–inho*, durante o século XI e será por ele suplantado no XII (Catalán, 1960). Merece referência, por isso, o topónimo **Quintelinha (1)**, junto ao rio Ave: trata-se de uma forma em que o sufixo diminutivo *-inha* se acrescenta a *quintela* num momento em que já não há memória de o sufixo *-ela* ser, também ele, diminutivo.

4.1.5. Quintinha/s (40/5)

Embora não se possa excluir a possibilidade de algumas instâncias toponímicas de *Quintinha* terem origem no latim QUINTANINA, a maioria há de ser já um diminutivo de *quinta*, formado com o sufixo *-inha*.



Mapa 4. Quintela/s – Quintinha/s

4.2. Casal/ais (4378/623) – Casares (12)

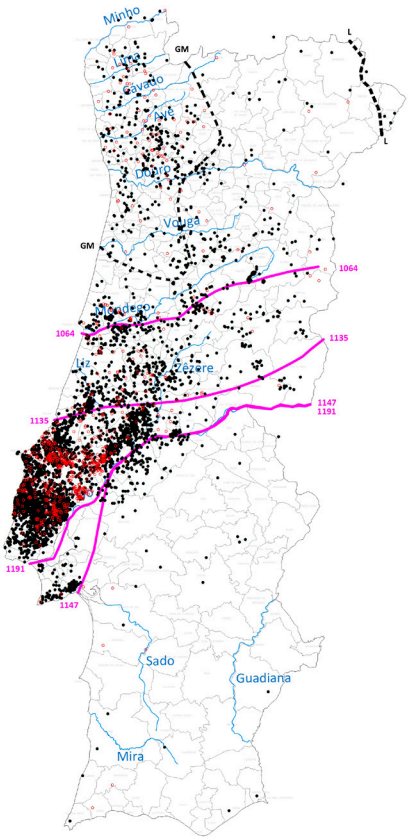
Proveniente do lat. CASALE 'relativo, pertencente à casa', o adjetivo *casalis* qualificou inicialmente o terreno pertencente à casa – o [*ager*] *casalis*, de acordo com a expressão *uinea casalis* abonada pelos gromáticos (i.e., agrimensores romanos), em que se especifica o tipo de cultivo do terreno (Vasconcellos, 1980 [1936]: 271). Quando no mesmo tipo de fontes encontramos a forma substantiva dele derivada, ela já não designa apenas o terreno, mas o conjunto do terreno e da casa, ou seja, a pequena exploração agrária (Chouquer, & Favory, 2002). O topónimo é bastante frequente e apresenta as variantes *-al* e *-ar* continuadoras das variantes latino-vulgares -ALE e -ARE do sufixo relacional. A variante em *–al* regista-se toponimicamente no singular

e no plural, a sua produtividade devendo-se ao facto de ter sido a adotada na língua comum. A variante em *-ar* é minoritária.

Casare e *casale* atestam-se na documentação latina do Norte em 865 e 870, respetivamente (Lapesa, & García, 2004). *Casal* aparece como topónimo tanto em textos latinos como em árabes andalusis do século XII: *Caxálex, Cachálex, El Caxár*, etc. (Simonet, 1967 [1888]). No entanto, a sua fixação toponímica é anterior, como o comprova a antiguidade de *Casarelho*, formado com o sufixo diminutivo latino clássico –ĪCULU.

A distribuição geográfica de topónimos com origem no nome comum *casal* (e na sua variante *casar*) verificou-se em toda a Hispania, o que explica a sua fixação toponímica em todo o território. O uso como nome comum não parece ter sido frequente no *Garbe*, pois não encontramos topónimos com essa origem atribuíveis ao período andalusi propriamente dito. Podemos, assim, supor que a sua difusão se tenha devido à transplantação do galego-português. Isso explicaria a razão por que a variante *casar* (na forma simples ou modificada) não se observa nos dialetos portugueses meridionais, excluída por nivelação em favor de *casal*, que foi a adotada na língua comum.

A distribuição geográfica das variantes toponímicas comprova esta interpretação. Com efeito, os formados com o sufixo –ARE registam-se apenas na forma plural e restringem-se aos concelhos do extremo norte confinantes com a Galiza. Os provenientes do lat. CASALINU, com síncope galego-portuguesa do -N-, não ultrapassam o paralelo de Aveiro. O mesmo é testemunhado pela cronologia das atestações documentais: na primeira metade do século XII, regista-se com frequência na documentação dos mosteiros de Grijó e de Pedroso (c. de Vila Nova de Gaia), já faz parte do vocabulário dos escrivães de Viseu, mas ainda é pouco conhecido dos de



Mapa 5. Casal – Casais

Coimbra; no decurso da segunda metade do mesmo século, o termo *casal* alcança as práticas notariais das Beiras. A sua progressão pára, no entanto, no Mondego e a Estremadura ignora-o quase por completo durante a época medieval (Durand, 1982).

A área de maior concentração de *Casal*, no litoral entre o Mondego e o Tejo, é recente, assim como também a de *Casalinho*, construção diminutiva com base em *casal* (+*inho*), que se documenta desde 1079 (Machado, 1993, s.v. *Casalinho*). A produtividade atual do nome na formação de topónimos não é tão grande como a de *quinta*, uma vez que modernamente o seu significado principal passou a ser o de 'casa de campo'. O topónimo no plural referencia um agregado de casais que forma um lugarejo (Vasconcellos, 1980 [1936]). Muitos dos topónimos *Casal* e *Casalinho* são fruto da colonização agrícola levada a cabo pelo Estado Novo, através da Junta de Colonização Interna criada em 1936.

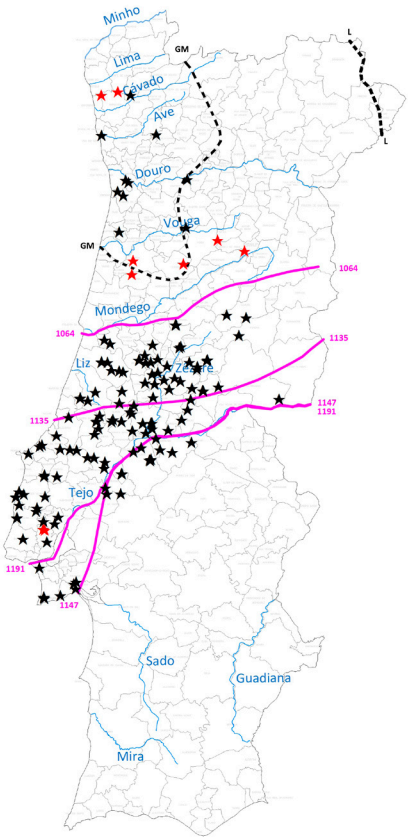
4.2.1. Casarelho/e (1/1)

Formada com o sufixo latino clássico -ĭculu (> *-eclu* > *-elho*), este diminutivo tem apenas duas ocorrências toponímicas, ambas em concelhos transmontanos (Montalegre e Vinhais), o que corrobora a ideia de *casar* ter sido uma variante minoritária nos dialetos da Galécia meridional.

4.2.2. Casarinhos (1) - Casainho/s (6/3) - Casalinho/s (120/11) – Casaizinhos (1)

Maior produtividade apresenta o sufixo latino relacional –īnu, que foi adotado em galego-português como diminutivo em substituição do latino-vulgar –ĕllu. É essa a razão que explica a sua aplicação às bases latinas casare-casale e a forma dele evoluída *–inho* às portuguesas *casal e casais.* Daí as formas divergentes *casainho/s* e *casalinho/s* e também *casaizinhos*, um diminutivo do plural *casais*. Com origem no latim casalinu, o primeiro destes diminutivos é galego-português, conforme atestado pela síncope do -n- intervocálico (1258 *casayno*, *casaino*, *casainhos*, Machado, 1993, s.v. *Casainho*); o segundo é português e consiste na junção do sufixo modificador *-inho* ao radical *casal*; o terceiro consiste num diminutivo da forma plural *casais*.

A distribuição geográfica de *casainho* corresponde praticamente à dos limites meridionais da área original galego-portuguesa, a sua expansão a leste devendo-se à colonização interna. O topónimo ter-se-ia, assim, fixado, antes da conquista definitiva de Coimbra em 1064, por Fernando o Magno. Conjuntamente com os dados documentais sobre a progressão do uso de casal com a Reconquista, esta distribuição pode ser usada como um marco da cronologia da síncope galego-portuguesa do –n- latino intervocálico.



Mapa 6. Casainho/s – Casalinho/s

4.2.3. Casalito (3) – Casaleco (1)

Casal conhece ainda os topónimos diminutivos com o sufixo *-ito e -eco,* que são minoritários.

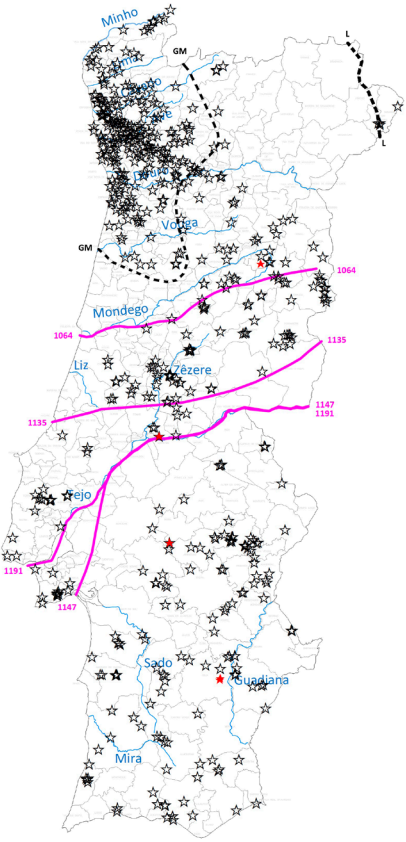
4.3. Aldeia/s (539/17)

Proveniente do árabe andaluz *aḍḍáyʿa* (Corriente, 2003, s.v. *aldea*) 'domínio senhorial', *aldea* é um arabismo incorporado no léxico iberorromânico desde a primeira fase da Reconquista, por intermédio dos moçárabes que imigraram para o Norte. Uma vez que *quintã* designava a 'casa senhorial' e *casal* a 'exploração rural

camponesa', *aldeia* passa a designar no noroeste a 'povoação rural', que pode integrar vários casais e quintãs.O nome comum covaria com *logo*, *lugar, povo, freguesia* e *sítio* como denominação de lugar habitado (Vasconcellos, 1980 [1936]). Na língua comum adquiriria o significado de 'povoação rural menor que a vila' e conheceria a epêntese de iode para resolução do hiato vocálico que hoje carateriza a forma *aldeia*. Não há evidências da sua fixação toponímica no período andalusi. Os topónimos com origem neste arabismo concentram-se no Entre Douro e Minho, observando-se ainda na Beira Alta, na Estremadura, em parte do Alentejo e no Algarve, mas não ocorrem nem em Trás os Montes nem na margem esquerda do Tejo.

4.3.1. Aldeinha (2) – Aldeola (1) - Aldeota (2)

Os diminutivos de *Aldeia* são poucos, em virtude de este nome ter passado a designar um povoado rural menor. O diminutivo de *aldeia*, como aliás de *quintã*, tem, na generalidade um valor afetivo, exceção feita para o sufixo *-ota* que adquiriu um sentido depreciativo.



Mapa 7. Aldeia/s – Alde-inha/ola/ota

4.4. Corte/s (428/115)

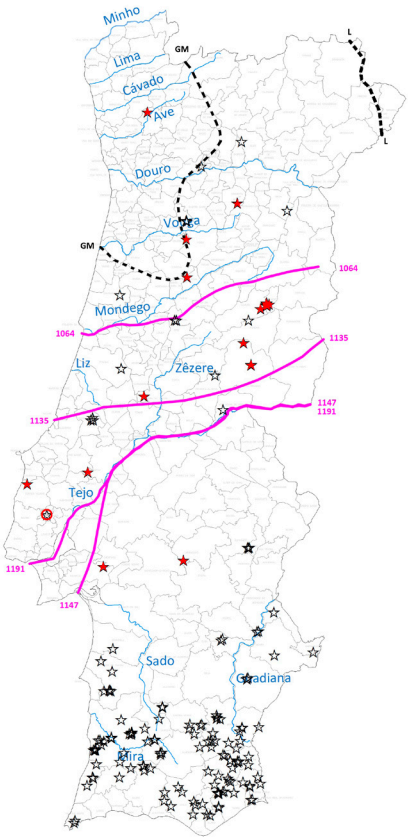
O nome comum tem como étimo remoto o latim COHORS, COHORTIS, um termo militar que designou uma divisão ou parte do acampamento romano e as tropas nela acantonadas e, especificamente, a décima parte da legião romana. O étimo próximo é a forma acusativa do baixo latim CORS, CORTIS, um sinónimo de *ātrĭum* (Torrinha, 1945), cujo significado evoluiria para propriedade rural com casa de habitação.

O átrio, que designava originalmente um pátio interior, tornou-se a parte mais importante da casa romana, uma sala situada logo depois da entrada e do vestíbulo, que se caraterizava pela sua forma quadrada (Torrinha, 1945). O termo, sob a forma galego-portuguesa

adro, seria aplicado ao terreiro descoberto em frente e à volta das igrejas, onde se reuniam os fiéis, depois da missa, e os homens-bons ou vereadores, quando o concelho não dispunha de paços próprios. Tal como ĀTRĬU, *corte* estendeu a sua significação à casa e, depois, como aconteceria com *quintana*, à propriedade por ela encabeçada. Paralelamente desenvolveu também o significado de recinto cercado e de curral, estábulo e pocilga.

O nome *corte* é panromânico, ocorrendo nos domínios gálico, itálico e ibérico, com o significado de curral, estábulo e pocilga; em italiano, castelhano, basco e andalusi tem também o significado primitivo de propriedade agrícola com casa de habitação (Simonet, 1967 [1888]), a sua primeira atestação sendo de 831 (Lapesa, & García, 2004). É com este último significado que o nome se fixou toponimicamente tanto na Andaluzia como nas Beiras, Alentejo e Algarve. A sobrevivência destes nomes de lugar na toponímia portuguesa, na maioria habitados (62%), é um indiciador da continuidade do seu povoamento durante o período andalusi e o galego-português.

A distribuição geográfica destes topónimos é complexa e vale a pena ser descrita em função não só da sua presença como da sua ausência. Começamos por esta última. Conforme se pode observar no mapa correspondente, este topónimo não ocorre na zona costeira do antigo *conventus bracarensis* correspondente ao eixo geográfico Braga-Porto, que funcionou como o eixo vertebrador do território galaico meridional durante o reino suevo e o astur-leonês (então acrescentado pela fundação de Guimarães). Também não ocorre entre o Vouga e o Mondego, sendo raro no litoral estremenho, na Beira Baixa e no Alto Alentejo. Ou seja, a sua área preferencial, são os limites orientais da Galécia Magna portuguesa, as zonas limítrofes de Trás os Montes e, nos territórios meridionais, para além



Mapa 8. Corte/s – Cortelha/s – Cortinhas

de uma ilhota ribatejana, a Beira Alta, o Baixo Alentejo e o Algarve. O que esta distribuição nos revela é que a fixação toponímica de *Corte* é andalusi e se encontra em distribuição complementar com *Quintã* e suas variantes, as sobreposições devendo-se à colonização interna.

4.4.1. Cortelha (14/16) – Cortelho/s (4/7)

Este nome conhece os diminutivos toponímicos *Cortelha/s* e as variantes masculinas *Cortelho/s*, formados com o sufixo latino -ĭculu (> *eclu* > gal.-port. *elho*). O nome feminino designa o recinto onde se cria ou recolhe o gado, sendo sinónimo de curral e malhada (Moraes Silva, 1890). O masculino *cortelho* referencia o lugar cerrado onde ficam as crias das cabras e ovelhas, para permitir a recolha do leite para o fabrico de queijos (Figueiredo, 1899). Toponimicamente, a forma feminina fixou-se preferencialmente na área meridional e a masculina no Alto Minho e na Beira Litoral.

4.4.2. Cortinha (24/38) – Cortinhos (2/1)

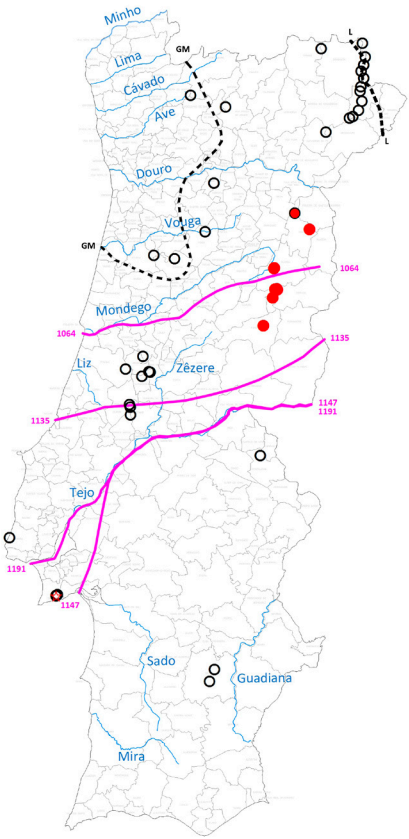
Do baixo latim cortina, foi originalmente um sinónimo de *cortale* 'portal de fora, átrio da corte'. Pensamos tratar-se de um adjetivo relacional que se substantivou quando *corte* passou a denominar a casa. Depois adquiriria também o significado diminutivo de horta murada, «minor cortis, seu rustica area, quae muris cingitur» (*apud* Simonet, 1967 [1888]), e courela lavradia (Figueiredo, 1899). O nome comum foi adotado em castelhano e no galego-português. A área toponímica portuguesa da forma feminina é setentrional, a masculina ocorrendo na região beirã.

4.5. Caria/s (18/2) – Alcaria (112/27)

Do árabe andalusi *alqaríyyah* (Corriente, 2003, s.v. *alcaria*), que designava, à época da reconquista cristã, as aldeias (i.e., os pequenos povoados rurais) da área meridional (século IX *karia*; XI *caria*; XIII *alcaria*). Por contraponto às denominações antropotoponímicas com *Ben-* e *Bena ~ Bela*, relativas aos assentamentos rurais clânicos, as alcarias dependiam das cidades islâmicas. Era, então, nas alcarias que se concentravam as populações autóctones, maioritariamente românicas ou delas descendentes. A forma toponímica *Caria* é Beirã e *Alcaria* meridional. Note-se que o facto de se terem conservado topónimos com esta denominação é um indicador do seu povoamento aquando da Reconquista. A concentração desses topónimos no Baixo Alentejo e no Algarve não se deve apenas à maior densidade populacional nesta região durante o domínio muçulmano, mas também ao refúgio de populações com o avanço da Reconquista cristã.

4.5.1. Alcarela (1) e Alcariota (2)

A forma híbrida *Alcarela*, junto ao Tejo, atesta a influência românica andalusi com o diminutivo românico *-ela* (< lat. ĕlla). *Alcariota* é provavelmente um diminutivo já da colonização interna, pois o sufixo é usado com um valor depreciativo, devido à pequena dimensão dos povoados respetivos.



Mapa 9. Caria/s – Alcaria/s

5. Topónimos com origem em nomes de árvores de fruta: (MALA) MATTIANA e derivados

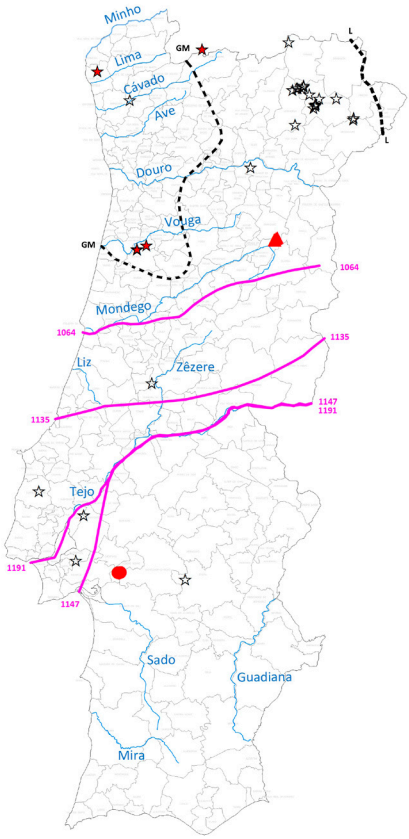
Passando, agora, para outro campo semântico, vejamos os topónimos que têm origem nos nomes comuns românicos que designam 'maçã' e 'pomar de maçãs'. O nome comum não coletivo encontra-se nos topónimos *Maçã,* do lat. (MALA) MATTIANA, atestado no século XIII (*maçãa,* vd. Houaiss, & Salles, 2002, s.v. *maçã*) ou *Maçãs*, atestado, como hidrónimo, em 1187, na forma *Maçãas* (Machado, 1995, s.v. *maçã*) juntamente com o diminutivo *Maçainha.* Uma vez que, na construção portuguesa de diminutivos, o sufixo *-inho* se associa a radicais, enquanto *-zinho* se junta a palavras, é provável que *-zinho* tenha sido, inicialmente, apenas uma ampliação de *-inho*, indispensável no caso das palavras atemáticas. A ser assim, *maçainha* será uma forma mais antiga do que a atual *maçãzinha*. Mas pode ser ainda mais antiga, se derivar diretamente do latim *MATIANINA.

Para 'pomar de maçãs' os topónimos (e nomes comuns) mais frequentes são *Macieira, Maceira, Macedo* e *Maçanedo. Macieira é* originalmente o nome da árvore, *Macedo* do coletivo (conjunto de macieiras). *Macedo* e *Maçanedo* formam-se com o sufixo latino -ĒTU (< *MATTIANETU) e documentam-se desde o século XIII (*maçanedo*, *maçaedo*, *mazanedo*, *mazaedo*, Machado, 1993, s.v. *Macedo*). A forma com síncope de -N- é a galego-portuguesa, a com a sua conservação a românica andalusi (mapa 10).

Se integrarmos a distribuição dos topónimos de tipo *Maçan-* no mapa das estremas da Conquista, verificamos que, à excepção de uma forma fronteiriça (que pode justificar-se por contacto de línguas) estes topónimos se registam em plena área andalusi, opondo-se à forma galego-portuguesa *Macedo*. No entanto, ambas as formas são raras no território português, tendo sido substituidas por *Maceira* e *Macieira*, usadas com um sentido coletivo. *Maçal*, sinónimo de *macedo*, não vingou como nome comum, designando hoje apenas uma povoação na Beira Alta. Dada a presença toponímica de *Manzanal* na Extremadura espanhola, é provável que o nome comum *maçal* (com síncope galego-portuguesa de -n-) represente uma forma resultante do contacto entre colonos galego-portugueses e leoneses. Este topónimo pode, assim, ser considerado um testemunho da proveniência regional dos colonos, evidenciando a interdialetalização induzida pelo contacto dialetal.

Para *Maceira* e *Macieira* podemos considerar dois étimos: a construção romance *maçã* + *eira*, com as variantes *maceira* e *macieira*, que resultam de diferentes resoluções do hiato (crase ou semivocalização) da forma intermediária *maceeira* (**maçãeira* > **maçaeira* > *maceeira* > *maceira* ~ *macieira*) ou a origem direta no étimo latino MATTIANARIA. Seja qual for o étimo, a evolução seria a mesma. As atestações são antigas: 974 *mazanaria*; 1066 *mazenaria*, 1072 *mazaneira*, 1077 *mazaria,* 1079 *macaneira,* 1080 *macenaria;* 1141 *maceira*; 1128 *maceejra*, 1162 *macaeira*; 1220 *maçeeyra*, *mazaeira*, *mazeira*, *mazeeira* (Machado, 1993, s.v. *Maceira* e *Macieira*).

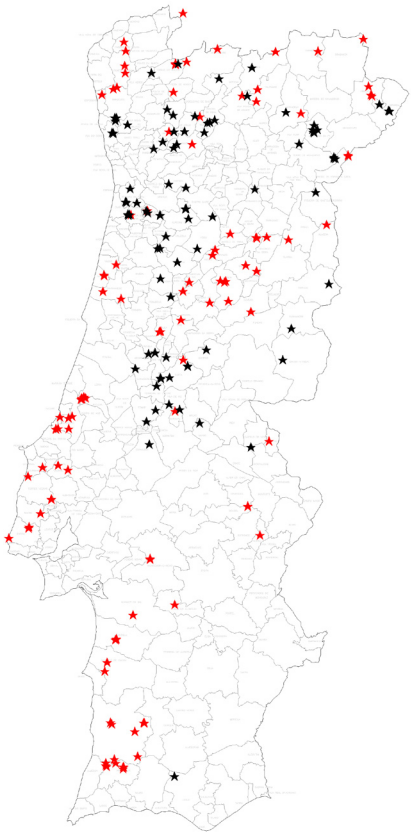
A distribuição de *Maceira* por todo o território mostra a antiguidade desta forma, levada para sul pela Reconquista. Trata-se de um topónimo claramente galego-português, com síncope de -N- latino intervocálico. A forma mais frequente na Galiza, *Maceira*, a par de outras que também revelam crase, como *Maceda* ou



Mapa 10. ☆ Macedo - ★ Macido - ● Maçanedo - ▲ Maçal

Macedo, continua no Alto Minho e difunde-se de norte para sul, acompanhando os movimentos da Reconquista e Repovoamento e chegando ao extremo meridional. No Algarve, nos séculos XIII e XIV, após um período de promulgação de forais e de criação de novos concelhos, a Coroa portuguesa limitou a sua intervenção às áreas fiscal e militar e à administração dos núcleos urbanos. As zonas rurais mantiveram, por esse motivo, as suas estruturas orgânicas, de que os vestígios dialetais minhotos ainda dão testemunho.

Ao lado da solução craseada *Maceira,* desenvolveu-se a variante com semivocalização *Macieira*, com atestações a partir do século XII: *maceejra*; XIII *mazeeira*; XIV *maçeeyras*. A sua concentração a norte do Tejo mas a sul do Cávado, parece indicar que esta forma poderá ser mais recente. De facto as atestações que permitem a interpretação de uma solução tritongada são mais recentes. Mas encontramos *Macieira* principalmente na região de Braga-Guimarães, que corresponde à área de uma primeira fase de reconquista, ainda na época das presúrias asturianas (finais do século IX, início do X), e as duas variantes encontram-se igualmente fixadas toponimicamente na Galiza (Cf. *Nomenclátor*). Parece, portanto, que as duas variantes se desenvolveram paralelamente. *Macieira* ocorre também na região de Coimbra, reconquistada e repovoada em fase posterior, por Afonso Henriques e Sancho I e pelos Templários. Uma hipótese que se pode colocar é que a variante tritongada fosse a preferida pelos possessores que se fixaram no Baixo Minho e Douro Litoral e aí criaram os seus senhorios. Não são, no entanto, esses senhores que vão colonizar os novos territórios reconquistados mas uma massa popular que pouco contacto teria com as variantes dialetais das classes mais cultas e que leva até ao sul, em vagas sucessivas, as variantes populares galego-portuguesas. Por outro lado, a



Mapa 11. Maceira/s – Macieira/s

variante do Baixo Minho e Douro Litoral pode ter acompanhado as conquistas de Afonso Henriques, fixando-se na região de Coimbra.

Certo é que a variante *Macieira* se concentra numa linha oeste, mais urbanizada. A leste e a sul (especialmente a sul do Tejo), as correntes migratórias, originárias, em vagas sucessivas, das regiões imediatamente a norte transportam traços linguísticos para sul: numa primeira fase, os colonos partem da Galiza e do Alto Minho e descem até à linha do Mondego; numa segunda fase descem do Mondego até ao Tejo e, finalmente, serão os colonos da região entre o Mondego e o Tejo que irão povoar o sul, embora sempre com o contributo de populações minhotas e galegas. As mudanças e variantes linguísticas são, assim, cumulativas, o que permite encontrar as mais antigas formas galego-portuguesas, como *maceira* ou *quintã* até ao extremo meridional do território.

6. Conclusões

A atual estruturação toponímica em Portugal continental atesta a presença sucessiva de vários grupos étnicos e linguísticos – a um norte-noroeste românico (com topónimos como *Quintã* ou *Casal*) opõe-se um centro-sul arábico (*Alcaria*) e românico (*Corte*) – e revela sucessivas camadas linguísticas que se vão sobrepondo (*Quintã-Quintela* / *Quinta-Quintinha*). As formas mais recentes, contudo, não eliminam as arcaicas, que se mantêm nas 'terras antigas' ou em 'ilhas' conservadoras, criando uma cartografia complexa que deve ser analisada à luz da história, da ocupação do espaço e da geografia humana. No estado atual dos conhecimentos, uma perspetiva estritamente linguística não parece suficiente para desvendar a história dos topónimos, mesmo que esclareça a sua etimologia. Disciplinas como a geografia física e humana, a dialetologia, a lexicologia e a sociolinguística, bem como a cartografia, podem revelar-se aliadas preciosas do estudo etimológico.

Uma análise da implantação toponímica, considerando a geografia humana e a ocupação do espaço, permite verificar a forma como o movimento da Reconquista cristã e o consequente processo de repovoamento transplantaram, de norte para sul, o dialeto primário da área original galego-portuguesa mas também revela um movimento de sul para norte, atestado pela implantação, a norte, de topónimos de origem árabe e moçárabe.

O território português parece ter uma relativa estabilidade demográfica desde a época pré-romana. As correntes migratórias ter-se-ão feito sentir, principalmente, nas zonas de mais intensa urbanização; as grandes regiões entre as cidades, por outro lado, conservaram grupos agro-pastoris de antigo assentamento. Assim, as variantes dialetais das classes mais cultas impõem-se nas áreas mais densamente urbanizadas, enquanto as regiões rurais conservam formas arcaizantes.

Fontes históricas e contemporâneas

Associação Nacional de Municípios Portugueses: http://www.anmp.pt/index.php/municipios

Barros, Mª de Fátima, Boiça, Joaquim Ferreira, & Gabriel, Celeste (1996). *As Comendas de Mértola e Alcaria Ruiva. As visitações e o Tombos da Ordem de Santiago 1482-1607*. Mértola: Campo Arqueológico de Mértola.

Dias, João José Alves (1998-1999). *Chancelarias Portuguesas. D. Duarte*, 2 vols. Lisboa: Centro de Estudos Históricos da Universidade Nova de Lisboa.

Dias, João José Alves (2002). *Cortes Portuguesas. Reinado de D. Manuel I (Cortes de 1498)*. Lisboa: Centro de Estudos Históricos da Unviersidade Nova de Lisboa.

Gonçalves, Iria (2002). *Tombos da ordem de Cristo. Comendas a sul do Tejo* Lisboa: Centro de Estudos Históricos da Universidade Nova de Lisboa.

Herculano, Alexandre (1977) [1856]. *Portugaliae Monumenta Historica: a saeculo octavo post Christum usque ad quintumdecimum...* Olisipone: Academiae Scientiarum Olisiponensis). Disponível em Biblioteca Nacional Digital: http://purl.pt/12270

Martins, Ana Maria (1998). *Documentos Portugueses do Noroeste e da Região de Lisboa. Da produção primitiva ao século XVI.* Lisboa: INCM.

Nomenclátor de Galicia: http://www.xunta.gal./nomenclator

Reportório Toponímico de Portugal. 03-Continente. Escala 1/25000, 3 vols. Lisboa: Ministério do Exército. Serviços Cartográficos, 1967.

Silva, Luís Fraga da (2002). NOTVS - *Exploração do Corpus Territoria do Sul [Computer software]*. Tavira: Campo Arqueológico de Tavira.

Referências bibliográficas

Almeida Fernandes, Armando de (1978). "A Nobreza na Época Vimarano-Portugalense", *Revista de Guimarães, 88*, 249-355.

Barroca, Mario Jorge (1999). "Fortificações e povoamento no Norte de Portugal (séc. IX a XI)", *Portugalia, Nova Série, vol. XXV*, 181-203.

Barroca, Mario Jorge (2003). "Da Reconquista a D. Dinis. Organização territorial e recrutamento militar". Em Barata, Manuel Themudo, & Teixeira, Nuno Severiano (eds.), *Nova História Militar de Portugal*, 69-94. Casais de Mem Martins: Círculo de Leitores.

Bolòs Masclans, Jordi (2010). "Cartografía, toponimia e historia medieval". Em X. Sousa Fernández (ed.), *Toponimia e Cartografía,* 39-71. Santiago de Compostela: Consello de Cultura Galega / Instituto de Lingua Galega.

Cardeira, Esperança (2005). *Entre o Português Antigo e o Português Clássico.* Lisboa: Imprensa Nacional-Casa da Moeda.

Cardeira, Esperança, & Fernandes, Maria Alice (2008). "O português medieval: koinização e elaboração", *Veredas*, *9*, 155-176.

Catalán, Diego (1960). "Hacia un atlas toponimico del diminutivo -INU en la toponimia hispano-románica", Separata do *Boletim de Filologia, XVII*.

Chouquer, Gérard, & Favory, François (2002). *L'Arpentage Romain.* Paris: Éditions Érrance.

Concepción Suárez, Julio (2007). *Diccionario Etimológico de Toponimia Asturiana.* Oviedo: KRK Ediciones.

Corriente, Federico (2003). *Diccionario de arabismos y voces afines en iberorromance* (2.ª edição). Madrid: Gredos.

Durand, Robert (1982). *Les Campagnes Portugaises entre Douro et Tage aux XIIe et XIIIe Siècles.* Paris: Fundação Calouste Gulbenkian – Centro Cultural Português.

Fernandes, Maria Alice (2001). "Toponímia algarvia de origem árabe". Em Coutinho, Valdemar (ed.), *Dinâmica Defensiva da Costa do Algarve. Do período islâmico ao séc. XVIII*, 21-33. Portimão: Instituto de Cultura Ibero-Atlântica.

Fernandes, Maria Alice (2007). "A toponímia do percurso". Em Khawli Abdallah, Silva, Luís Fraga da, & Fernandes, Maria Alice, *A viagem de Ibn Ammâr de São Brás a Silves. Estudo histórico, geográfico e toponímico.* São Brás de Alportel: CM.

Fernandes, Maria Alice, & Cardeira, Esperança (2013). "Toponímia do povoamento rural: os continuadores do latim QUINTANA e seus derivados". Em Galvão, Vânia *et alii* (eds.), *Língua Portuguesa: ultrapassando fronteiras, unindo culturas. Anais do IV Simpósio mundial de estudos de língua portuguesa,* 1913-1921. Goiânia: FUNAPE. http://www.simelp.letras.ufg.br/anais/simposio_44.pdf

Figueiredo, Cândido de (1899). *Novo Dicionário da Língua Portuguesa,* 2 vols. (3.ª edição). Lisboa: Imprensa Portugal-Brasil.

Houaiss, Antônio, & Salles Villar, Mauro de (2002). *Dicionário Eletrônico Houaiss da Língua Portuguesa*. CD-ROM, versão 1.0.5. Rio de Janeiro: Objetiva.

Lapesa, Rafael, & García, Constantino (2004). *Léxico hispánico primitivo (siglos VIII al XII).* Em Seco, Manuel (ed.), Madrid: Fundación Ramon Menéndez Pidal - Real Academia Española - Espasa Calpe.

López Carreira, Anselmo (1997). "Idade Media". Em *Historia Xeral de Galicia,* 93-204. Vigo: A Nosa Terra.

Machado, José Pedro (1993). *Dicionário Onomástico Etimológico da Língua Portuguesa,* 3 vols. (2. ª edição). Lisboa: Editorial Confluência.

Machado, José Pedro (1995). *Dicionário Etimológico da Língua Portuguesa*, 5 vols. (7.ª edição). Lisboa: Livros Horizonte.

Maia, Clarinda de Azevedo (1981). "Geografia dialectal e história do português. Resultados da terminação latina -ANA", *Biblos*, *57*, 73-96.

Marques, António Henrique R. de Oliveira, & Dias, João José Alves (2003). *Atlas Histórico de Portugal e do Ultramar Português.* Lisboa: Centro de Estudos Históricos.

Martins, Ana Maria (1985). "A evolução das vogais nasais finais [ã], [õ], [ẽ] no português". Em Pereira, Cilene da Cunha, & Pereira, Paulo Roberto Dias (eds.), *Miscelânea de Estudos Lingüísticos, Filológicos e Literários In Memoriam Celso Cunha*, 617-646. Rio de Janeiro: Editora Nova Fronteira.

Mattoso, José (1988). *Identificação de um País. Ensaio sobre as Origens de Portugal 1096-1325*, *Vol. II - Composição* (3.ª edição). Lisboa: Editorial Estampa.

Milroy, James (1992). *Linguistic Variation and Change: on the historical sociolinguistics of English.* Oxford: Basil Blackwell.

Moraes Silva, António de (1890). *Diccionario da Lingua Portugueza,* 2 vols. (8.ª edição). Rio de Janeiro: Empreza Litteraria Fluminense.

Oliveira, Fernão de (1981) [1536]. *Gramática da Linguagem Portuguesa*. Edição fac-similada. Lisboa: Biblioteca Nacional.

Piel, Joseph-M. (1945). "As águas na toponímia galego-portuguesa", *Boletim de Filologia, VIII*, 305-342.

Piel, Joseph-M. ([1960b] 1989). "Toponímia Germânica da Península Ibérica". Em Piel, Joseph-M., *Estudos de Linguística Histórica Galego-Portuguesa,* 149-171. Lisboa: IN-CM.

Piel, Joseph-M. (1989). *Estudos de Linguística Histórica Galego-Portuguesa.* Lisboa: Imprensa Nacional - Casa da Moeda.

Quiroga, Jorge L., & Lovelle, Mónica R. (1997). "Un modelo de análisis del poblamiento rural en el valle del Duero (siglos VIII-X) a partir de un espacio macro-regional: las tierras galaico-portuguesas", *Anuario de Estudios Medievales, 27*, 687-748.

Rodrigues Lapa, Manuel (1932). "Transposição e dissimilação nasal", *Boletim de Filologia, 1*, 163-165.

Simonet, Francisco Javier (1967) [1888]. *Glosario de voces ibéricas y latinas usadas entre los mozárabes.* Amsterdam: Oriental Press.

Torrinha, Francisco (1945). *Dicionário Latino-Português* (3ª edição). Porto: Edições Maranus.

Vasconcellos, José Leite de (1980) [1936]: *Etnografia Portuguesa, II.* Lisboa: Imprensa Nacional - Casa da Moeda.

Vicente, Ángeles (2006). *El proceso de arabización de Alandalús: un caso medieval de interacción de lenguas.* Zaragoza: Instituto de Estudios Islámicos y del Oriente Próximo.

Hagiotoponímia no Brasil: resultados de pesquisa

Ana Paula Mendes Alves de Carvalho
Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais, Campus Ouro Branco

Maria Cândida Trindade Costa de Seabra
Universidade Federal de Minas Gerais, Faculdade de Letras

Resumo:

Partindo do pressuposto de que os nomes de lugar formados por nomes de santos –os hagiotopônimos– constituem um grupo especial de topônimos em que se percebe a comunhão de aspectos psicológicos do ser humano com a geografia e a paisagem, este trabalho tem por objetivo apresentar resultados da pesquisa de Carvalho (2014), em que se realizou um estudo descritivo do léxico hagiotoponímico brasileiro, focalizando os 853 municípios do estado de Minas Gerais. Adotando os pressupostos teórico-metodológicos da ciência onomástica –Dauzat (1926) e Dick (1990a, 1990b, 2004 e 2006)–, este estudo vinculou-se ao Projeto ATEMIG –Atlas Toponímico do Estado de Minas Gerais–, projeto em desenvolvimento, desde 2005, na Faculdade de Letras da Universidade Federal de Minas Gerais. Em sua primeira etapa, o ATEMIG fez o levantamento e a classificação toponímica de todos os acidentes físicos e humanos do estado, documentados em cartas geográficas – fontes do IBGE, com escalas que variam de 1: 50.000 a 1: 250.000, perfazendo até o presente um total de 85.391 topônimos. Desse total de dados, os que constituem nosso corpus contemporâneo, totalizam 5.649 hagiotopônimos. Além disso, através da consulta a mapas dos séculos XVIII e XIX, construiu-se um segundo corpus de dados históricos, composto de 647 topônimos. A análise comparativa desses corpora permitiu perceber como o processo designativo a partir de hagiotopônimos se deu diacronicamente no estado. Verificou-se, ainda, que, nos dados históricos, a quantidade de topônimos relativos às invocações marianas é, em números percentuais, quase duas vezes maior em relação aos dados contemporâneos.

Palavras-chave:

Léxico, Cultura, religiosidade, toponímia.

Sumário:

1. Primeiras palavras. 2. Fundamentação teórica: religiosidade e toponímia. 2.1. O culto aos santos: breve histórico. 2.2. A devoção no Brasil: uma herança portuguesa. 2.3. Hagiotoponímia: nomes de santos no espaço geográfico. 3. Procedimentos metodológicos. 4. Análise dos dados. 4.1. Dados históricos. 4.2. Dados contemporâneos. 4.3. Discussão dos resultados. 5. Considerações finais.

Hagiotoponymy in Brazil: research results

Abstract:

Assuming that place names formed by names of saints (hagiotoponimic) are a special group of place names in which the communion of psychological aspects of human beings with the geography and landscape may be perceived, this paper aims to present results of Carvalho's research (2014), in which she performed a descriptive study of Brazilian hagiotoponimic lexicon, focusing on the 853 municipalities in the state of Minas Gerais. Adopting the theoretical and methodological assumptions of onomastic science - Dauzat (1926) and Dick (1990a, 1990b, 2004 and 2006) -, this study was linked to the ATEMIG Project - Toponymic Atlas of Minas Gerais State, a project in development since 2005 at Faculdade de Letras da Universidade Federal de Minas Gerais. In its first stage, the ATEMIG undertook the survey and toponymic classification of the entire State´s physical and human accidents documented on geographical maps - IBGE sources, with scales ranging from 1: 50.000 to 1: 250,000, with a present total of 85,391 place names. From this total, the data constituting our present corpus are represented by 5649 hagiotoponyms. Moreover, by consulting the maps of the eighteenth and nineteenth centuries, a second corpus of historical data, made up of 647 place names, was developed. The comparative analysis of these corpora has provided an insight into how the designation process from the hagiotoponimic occurred diachronically in the State. It was also verifdied through historical data obtained that the number of place names relating to Marian invocations is, in percentage figures, nearly twice as high in comparison to contemporary data.

Key words:

Lexicon, culture, religion, toponymy.

Contents:

1. First words. 2. Theoretical basis: religiosity and place names. 2.1. The cult of the saints: brief history. 2.2. Devotion in Brazil: a Portuguese heritage. 2.3. Hagiotoponimic: holy names in geographic space. 3. Methodological procedures. 4. Data analysis. 4.1. Historical data. 4.2. Contemporary data. 4.3. Discussion of results. 5. Final considerations.

1. Primeiras palavras

Dar aos lugares o nome de um santo é um costume muito antigo. De acordo com Diéguez González (2001: 65), remonta à segunda metade do século VI, quando as igrejas catedrais e paroquiais não tinham patrono, mas as igrejas particulares, basílicas e oratórios já eram normalmente erigidas em honra de um santo. Assim, o santo passava a ser o símbolo da igreja e a lhe dar o nome, que posteriormente era atribuído também às terras e à freguesia que iam se organizando nos arredores do templo religioso.

No Brasil, esse costume, que teve início na antiguidade cristã, veio, ao longo dos séculos, sendo passado de geração para geração e pode ser observado ainda nos dias atuais, "sendo possível até mesmo verificar-se a preferência regional pelo emprego sistemático de determinados santos e santas" (Dick, 2010: 187)

De acordo com Barbosa (1995), os portugueses que vieram povoar o Brasil eram originários de várias províncias de Portugal: Minho, Beira-Alta, Estremadura, Alentejo, e outros. Havia gente do Norte, do Sul, do Centro, entretanto, um sentimento comum a todos unia: muita religiosidade. Cada família recém-chegada ao Brasil tinha seu oratório, com a imagem do santo ou da santa, a quem dirigia suas preces.

Foi esse espírito de religiosidade que os portugueses, de um modo geral, transplantaram para os primeiros núcleos povoados em terras brasileiras, o que também ocorreu, no final do século XVII e início do século XVIII, na região onde se deu o início do povoamento de Minas Gerais, uma vez que, de acordo com Lima Júnior (1978: 87), também as primeiras povoações mineiras surgiram em torno de capelas. Nas palavras do autor:

> Mal começavam a prosperar, sem demora, surgia uma capelinha de taipa, em cujo altar se firmava a estátua que reproduzia o padroeiro da vila ou aldeia distante em Portugal. Reuniam-se vizinhos e iniciava-se o culto, pelas grossas mãos postas dessa geração de heroicos fundadores de Minas, que eles regaram de suor e de sangue (...) À beira dos grandes caminhos, nas proximidades dos locais onde se encontravam as minerações, ou nas quintas dos grandes sesmeiros, por todo o território imenso dessas extensas Minas Gerais, ficaram os marcos do espírito cristão e católico, plantados pelas gerações do século XVIII.

Assim, em território mineiro, o costume de atribuir aos lugares um nome religioso também veio, ao longo dos séculos, sendo perpetuado pelas gerações e pode ser observado, com nitidez, na contemporaneidade, conforme mostrou o trabalho de doutorado de Carvalho (2014). Intitulada *Hagiotoponímia em Minas Gerais*[1], a

1 Pesquisa de doutorado desenvolvida na Faculdade de Letras da UFMG no período de 2010 a 2014, sob a orientação da Profª Drª Maria Cândida Trindade Costa de Seabra, disponível em http://www.bibliotecadigital.ufmg.br.

pesquisa constituiu-se de um estudo descritivo –linguístico e cultural– da toponímia mineira, focalizando nomes de santos, nomes de santas e nomes de invocações à Virgem Maria que designam acidentes físicos e humanos nos 853 municípios do Estado.

Vinculado ao Projeto ATEMIG[2] –Atlas Toponímico do Estado de Minas Gerais–, o estudo contou com um Banco de dados de 85.391 topônimos, dentre os quais 5649 hagiotopônimos foram analisados quantitativa e qualitativamente para que se verificasse a preferência regional pelo emprego sistemático dessas denominações em cada uma das doze mesorregiões mineiras. Além desses topônimos, que constituíram o *corpus* de dados contemporâneos, através da consulta a mapas dos séculos XVIII e XIX, construiu-se um segundo *corpus* de dados históricos, composto de 647 topônimos. Assim, por meio da análise comparativa, percebeu-se como o processo designativo a partir de hagiotopônimos se deu diacronicamente no Estado. Desse modo, ao apresentar resultados da pesquisa de Carvalho, este estudo trata da hagiotoponímia no Brasil, focalizando o léxico toponímico do Estado de Minas Gerais.

2. Fundamentação teórica: religiosidade e toponímia

Os estudos toponímicos são de grande relevância para o conhecimento de aspectos históricos e socioculturais de um povo, uma vez que permitem a identificação de fatos linguísticos, de ideologias e crenças presentes no ato denominativo e, posteriormente, na sua permanência ou não em uma comunidade.

Destaca-se, por exemplo, a influência da religiosidade na motivação para a escolha dos nomes dados aos lugares ao longo dos tempos. No léxico toponímico brasileiro, isso pode ser percebido desde o primeiro contato do colonizador europeu com as terras recém-descobertas, o que se justifica, sobretudo, pelas circunstâncias que levaram ao descobrimento da terra e pelo momento histórico representado pelo quinhentismo português, em que a difusão dos preceitos da Igreja Católica representava uma preocupação máxima.

2.1. O culto aos santos: breve histórico

De acordo com a fenomenologia religiosa moderna, a santidade é um conceito central em qualquer religião. No Islamismo, por exemplo, as tumbas dos santos são

2 Projeto em desenvolvimento, desde 2005, na Faculdade de Letras da Universidade Federal de Minas. Gerais Em sua primeira etapa, o ATEMIG fez o levantamento e a classificação toponímica de todos os acidentes físicos e humanos dos 853 municípios do estado, documentados em cartas geográficas –fontes do IBGE, com escalas que variam de 1: 50.000 a 1: 250.000.

visitadas por suplicantes à procura de alguma bênção. O budismo reconhece como santos aqueles que, através de numerosos nascimentos, atingiram a iluminação.

O culto cristão aos santos, por sua vez, data dos primeiros séculos da Antiguidade, quando vários cristãos foram perseguidos até a morte pelo testemunho de fé, os mártires. Após a morte desses cristãos, seus companheiros passavam a venerar seus túmulos e suas relíquias (restos mortais e objetos pessoais) a que podiam ser atribuídas propriedades milagrosas. Acreditava-se que, por causa do martírio, o santo compartilhava dos sofrimentos de Jesus e ganhava acesso imediato à presença de Deus.

No Ocidente, terminadas as perseguições por parte do Império Romano, passa-se do culto aos mártires para o culto aos bispos. Já no Oriente, em que as Igrejas se destacam pela veneração dos santos, essa evolução ocorre de forma distinta, tendo depois profundas repercussões no cristianismo, de uma forma geral.

Apoiando-nos em Oliveira (1975: 4), no entanto, podemos dizer que a concepção popular de santo é muito mais abrangente do que se possa imaginar, uma vez que inclui, além dos santos canonizados pela Igreja, todas as denominações locais e titulares da Virgem Maria, de Jesus, bem como os santos locais e familiares. O que significa, em outros termos, que o culto a um determinado santo é histórico. Isto porque sua representação nos permite perceber a maneira pela qual uma determinada comunidade vive sua relação com a realidade social, definindo estratégias de convivência a partir da necessidade de resolução de seus problemas, contexto em que se investe de poder um grupo de especialistas, os santos, capazes de restaurar a ordem daquilo que é interpretado como caótico por intermédio do milagre.

Convém destacar ainda que, assim como o culto aos santos, a devoção a Maria, Mãe de Jesus Cristo, é um componente essencial do culto cristão desde a Antiguidade, sobretudo depois do Concílio de Éfeso (435), quando se reconheceu a Maria o título de 'Mãe de Deus'.

De acordo com Dias (1987), a Idade Média (séculos v a xv) fez que a devoção à Virgem Maria invadisse toda a piedade cristã, desde a liturgia às artes e literatura. Desdobraram-se as festas do calendário litúrgico e da denominação "Santa Maria" surgiram múltiplas invocações a Nossa Senhora, fomentando, assim, a construção de catedrais, igrejas e capelas a ela dedicadas.

Desse modo, desde os primeiros tempos da Igreja, a Virgem Maria, os apóstolos e os mártires foram considerados santos e venerados pela comunidade de fiéis, marcando a história do cristianismo em diversas sociedades.

2.2. A devoção no Brasil: uma tradição de origem portuguesa

A devoção fiel ao culto e à tradição da Igreja Católica sempre foi uma constante na nação ibérica. Esse costume se estendeu, à época dos descobrimentos e das conquistas ultramarinas, a todas as costas continentais e insulares aonde chegaram as embarcações lusas, como é o caso do Brasil.

A respeito dessa influência, Chaves (1957: 178), explica que os portugueses não podiam estabelecer um hiato histórico e de ação entre o que era nacional em Portugal e o que tinha de ser igualmente nacional para além do mar. Se as instituições metropolitanas foram o padrão da colonização e administração dos territórios ultramarinos e constituíam a força criadora dessa mesma atividade, era de esperar que toda a vida espiritual dos navegantes e colonizadores fosse transferida para as terras ultramarinas. Vale ressaltar aqui que a religiosidade cristã portuguesa não se limitou ao século XVI, pelo contrário, foi se fortalecendo ainda mais no decorrer dos séculos.

Assim, a religiosidade lusitana encontra, no novo continente, o clima fecundo ao seu expansionismo. Na carta de Pero Vaz de Caminha à Coroa Portuguesa, por exemplo, é possível perceber topônimos relacionados ao costume dominante entre os navegadores de nomear os acidentes geográficos de acordo com as inscrições do calendário religioso romano, como se verifica em Monte Pascoal (ou Pascal), já que a chegada à nova terra se deu na época da Páscoa.

Ao estudar a trilha das bandeiras na região em que se deu o início do povoamento de Minas Gerais, por exemplo, Megale (2002) constata que a nomeação dos lugares foi fortemente influenciada pela devoção religiosa. Nas palavras da autora:

> A toponímia na trilha das bandeiras documenta a presença desses religiosos: se o comando era de carmelitas, os núcleos habitacionais que surgiam perpetuavam a lembrança do orago de Nossa Senhora do Carmo; se o comando era de franciscanos, no de São Francisco, e assim com outras ordens da mesma maneira, os oragos marcam sua passagem. Lá onde o povoado é novo, o nome religioso o inaugura, se havia uma designação indígena o novo nome religioso a substitui. (Megale, 2000: 22)

Desse modo, devido à herança portuguesa, enraizada ancestralmente à religião cristã, a devoção à Nossa Senhora e aos santos, principalmente àqueles ligados ao Cristo e à Mãe de Deus, como São José, Santa Ana, São Joaquim, São João Batista e São Pedro, sempre mereceu carinho especial do povo brasileiro, o que pode ser percebido, dentre outras formas, pela observação das marcas deixadas no espaço geográfico, isto é, na toponímia.

2.3. Hagiotoponímia: nomes de santos no espaço geográfico

Os nomes de lugar formados por nomes de santos constituem um grupo especial de topônimos em que se percebe com clareza a comunhão de aspectos psicológicos do ser humano com a geografia e a paisagem. Conforme (Diéguez González, 2001: 63),

> São muitas as localidades que se honram de ser chamadas com o seu patrono e os oragos estão vivos no imaginário popular até ao ponto de existir com eles um trato igualitário e fraternal: o santo está na igreja mas é um vizinho como outro qualquer a quem pedir favores e a quem se convida a participar das alegrias e das tristezas. Quando o santo dá nome à freguesia será por ter sido capaz de ganhar o apreço dos seus covizinhos, salvando-os das calamidades e acompanhando-os nos bons e maus momentos de suas vidas.

Surgem aí os hagiotopônimos, vocábulo constituído dos elementos gregos hagios (santo), topos (lugar) e onoma (nome). "O esquema linguístico é sempre o mesmo: um hagiônimo (nome de santo) serve de determinante a um nome comum que exprime a noção de núcleo populacional: lugar, vila, aldeia, etc**."** (Diéguez González, 2001: 65).

No âmbito dos estudos toponímicos, além da pesquisa de Carvalho (2014) sobre a hagitoponímia mineira, vários estudos têm mencionado a influência da devoção aos santos e à Virgem Maria na nomeação dos acidentes geográficos. Alguns deles referem-se especificamente a essa temática, dentre os quais citamos[3]:

- na Polônia, "Geografical Names Deriving from Saints'Names (Patrocinia) in Poland", de Czopek-Kopciuch (2011);
- na Húngria, "Patrociny Settlement Names in the Carpathian Basin", de Tóth (2011);
- na República Tcheca e na Eslováquia, "Patrociny Settlement Names in the Czech Republic and Slovakia", de Stepán (2011);
- na Alemanha, "Patrociny Settlement Names in Eastern Germany", de Hengst (2011), "Patrociny Settlement Names in North-Western Territories of Germany", de Casemir (2011), "Sankt Blasien –Sammarei– Helena: Place Names with Sankt in Bavaria and Baden-Wurttemberg", de Buncher (2011);
- na França, "La Religion dans la Toponymie" de Lejeune (2002) e os estudos homônimos "Patrociny Settlement Names in France", de Billy (2011) e Taverdet e Gendron (2011);

3 Os estudos de 2011 encontram-se em Tóth (2011).

- na Itália, "Patrociny Settlement Names in Italy", de Marcato (2011);
- na Espanha, "Influencia de la devoción popular en la Onomástica", de Ayerbe Iríbar (1982); "Hagiotoponímia a la Segarra", de Turrul i Rubinat (1985), "Alguns nomes de santos, do latim ao galego-português", de Diéguez González (2001); "Hagiotoponyms in Catalonia (Spain)", de Tort-Donada (2011) e "Contributions to Spanish Hagiotoponyms" de Sastre (2011);
- em Portugal, "Os nomes dos santos tradicionais hispânicos na toponímia peninsular" de Piel (1950), "As influências religiosas na formação da Antroponímia e da Toponímia de Portugal" de Chaves (1957); e "Hierotoponímia portuguesa: os nomes de Nossa Senhora", de Carvalhinhos (2005) e,
- no Brasil, "A hierotoponímia no Brasil", de Dick (1990b).

Tantos pesquisadores se dedicarem a esse objeto de estudo evidencia o fato de que, venerados, desde os primeiros tempos da Igreja, como modelos de vida cristã e invocados como intercessores diante de Deus, os santos e, também, as diversas invocações marianas marcaram a história do cristianismo e, ainda hoje, são representativos em diversas esferas da sociedade:

> Encontramo-los na piedade popular, nas devoções e no culto litúrgico, nos nomes de batismo das pessoas e nos nomes de lugares, nas tradições dos povos e no folclore, nas lendas e nos provérbios, na arte e na literatura. [Fazem parte] do patrimônio religioso da humanidade, assim como do seu patrimônio cultural (Daix, 2000: 11).

3. Procedimentos metodológicos

Orientado pelos pressupostos teóricos dos estudos do léxico que se fundamentam na inter-relação língua, cultura e sociedade (Sapir, 1961; Matoré, 1953) e focalizando os mariotopônimos, este trabalho constitui um recorte da pesquisa de Carvalho (2014), em que se observou, por meio de uma análise diacrônica, como ocorreu o processo designativo a partir de hagiotopônimos em Minas Gerais.

Para tanto, adotando-se o conceito de cultura de Duranti (2000) e os princípios da ciência onomástica –Dauzat (1926) e Dick (1990a, 1990b, 2004 e 2006)–, a pesquisa contou com dois *corpora*: o primeiro, que foi obtido a partir da consulta ao Banco de dados do Projeto ATEMIG, composto de 5649 hagiotopônimos; e o segundo, que foi organizado por meio da consulta a mapas dos séculos XVIII e XIX que integram o trabalho Cartografia das Minas Gerais: da capitania à província, organizado por Costa et al (2002), constituído de 647 hagiotopônimos.

De acordo com Dick (1990a), os hagiotopônimos constituem uma subdivisão dos hierotopônimos, taxe toponímica referente aos nomes sagrados de diferentes crenças. Essa terminologia difere, em certa medida, daquela usada na Espanha por Luis López Santos que define hagiotopônimo "como todo vocábulo do léxico religioso convertido em topônimo, isto é, vinculado ao geográfico e convertido em nome de lugar[4]" (López Santos, 1960: 580 *apud* Dick, 1990a: 311).

A autora reconhece a divergência e esclarece que isso não invalida a classificação brasileira:

> trata-se, tão somente, de uma inversão de conceitos: o que, para um, é visto sob uma perspectiva particularizante, para o outro é percebido sob um caráter global e abrangente, o que, em absoluto, não desnatura a substância de conteúdo dos elementos pesquisados (Dick, 1990a: 311).

Retomando a explicação de Dick (1990b: 33) para a classificação toponímica dos nomes religiosos, observa-se que além se referir aos 1) aos nomes sagrados de diferentes crenças, 2) às efemérides religiosas, 3) às associações religiosas, 4) aos locais de cultos, os hierotopônimos podem apresentar, ainda, duas subdivisões: 5) hagiotopônimos - topônimos relativos aos santos e santas do hagiológio romano e 6) mitotopônimos: topônimos relativos às entidades mitológicas. Desse modo, em relação às outras, os hierotopônimos constituem uma taxe mais complexa por reunir, dentro do campo semântico dos nomes sagrados, as seis categorias distintas de topônimos relacionadas.

Neste estudo, adotou-se a terminologia de Dick, isto é, consideramos hagiotopônimos apenas os nomes de lugar que se referem aos santos e santas do hagiológio romano e não qualquer vocábulo do léxico religioso. Entretanto, diferentemente da autora e pautando-nos na literatura concernente à origem e à história do culto aos santos, incluímos nessa subtaxe as invocações marianas, visto que, conforme mostramos na seção anterior, os diversos títulos de Nossa Senhora surgem, na Idade Média, a partir do hagiônimo Santa Maria (Dias, 1987: 228).

Vale dizer que as denominações toponímicas relativas à Nossa Senhora vêm sendo classificadas, nos estudos toponímicos brasileiros, de forma diversa, ora com hagiotopônimos –Castiglioni (2008), Zamariano (2010) e Carvalho (2012)– ora pelo hiperônimo, isto é, como hierotopônimos –Seabra (2004), Carvalhinhos (2005), Ananias (2013)– assim como foi proposto por Dick.

---

4 "(…) es **hagiotopónimo** todo vocablo del léxico religioso convertido em topónimo, es decir, vinculado a lo geográfico y convertido en nome de lugar" (López Santos, 1960: 580 *apud* Dick, 1990a: 311) Tradução nossa.

Dessa maneira, como esses designativos não possuem a estrutura linguística dos hagiotopônimos –qualificativo (são/santo(a)) + antropônimo (prenome)–, seguindo a proposta de desdobramento da subtaxe relativa aos nomes de santos de Carvalho (2014), os topônimos relativos aos nomes de Nossa Senhora são classificados como mariotopônimos, conforme representamos na figura seguir:

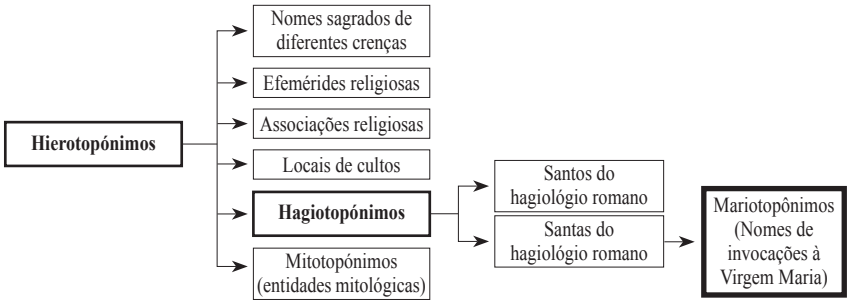


Figura 1. A taxe hierotopônimos e suas subdivisões. Fonte: Carvalho (2014: 89)

4. Análise dos dados

4.1. Dados históricos

O *corpus* de dados históricos, conforme dissemos anteriormente, foi constituído a partir da consulta a mapas dos séculos XVIII e XIX que integram o trabalho Cartografia das Minas Gerais: da capitania à província, organizado por Costa et al (2002).[5]

Dos 31 mapas consultados, apenas dois não estavam legíveis, impossibilitando, assim, a quantificação dos topônimos neles contidos.

Desse modo, nos 29 mapas legíveis, quantificaram-se, então, um total de 647 hagiotopônimos, dos quais 439 são ocorrências de nomes de santos, correspondendo a 67,8% dos dados, e 170, de nomes de santas, o que equivale a 26,3% do total de ocorrências, conforme pode ser observado no gráfico 13. Desses dados, apenas o topônimo S. Januário, que apresenta ocorrência única, não aparece no corpus de dados contemporâneos.

5 Trata-se de uma obra sobre a cartografia brasileira do período colonial e imperial composta de 31 pranchas cartográficas de Minas Gerais, sendo 25 do período colonial, das quais 6 referem-se a mapas regionais, 7 referem-se especificamente à área da Demarcação Diamantina, 9 tratam de Divisões Administrativas e 4 sobre as Capitanias, mas de modo amplo. As demais referem-se à Província.

As 38 ocorrências restantes, que representam 5,9% dos dados, referem-se aos mariotopônimos. Dentre elas, 13 aparecem apenas nos mapas históricos, isto é, não constam do banco de dados do ATEMIG, quais sejam Snra. das Candeyas (2), Sra do Livramento (1), N. S. da Lapa (1), N.S. das Correntes (1), N.S. de Nazareth (2), N.S. do Bom Sucesso (3) e N.S. do Desterro (3).

4.2. Dados contemporâneos

Por sua vez, o *corpus* de dados contemporâneos foi formado a partir da consulta ao banco de dados do Projeto ATEMIG, que se constitui, por ora, de um total de 85391 topônimos coletados em mapas contemporâneos de Minas Gerais.

Desse total de dados, focalizaram-se os nomes de lugar de natureza religiosa relativos aos santos e santas do hagiológio romano, os hagiotopônimos, que representam 5649 ocorrências dos topônimos que integram o Banco de dados.

Representando apenas 6,5% do total de topônimos do Estado, esses topônimos de origem religiosa parecem não constituir a taxe predominante em Minas Gerais, entretanto, no cômputo geral dos dados do ATEMIG, em que se observa, para cada mesorregião[6], o número de ocorrências de cada uma das 27 taxes propostas por Dick (1990b), verifica-se que, dentre as taxionomias de natureza antropocultural, os hagiotopônimos aparecem com destaque em todas elas, figurando entre o 1º e o 5º lugar.

A mesorregião Zona da Mata, por exemplo, sobressai em relação às demais, já que os nomes dessa taxe aparecem em 1º lugar, justificando, assim, o fato de maioria de nossos dados estarem situados nessa mesorregião.

Quanto às demais mesorregiões, observa-se que as denominações hagiotoponímicas aparecem em 2º lugar, no Vale do Mucuri, no Vale do Rio Doce, no Noroeste de Minas e no Sul e Sudoeste de Minas; em 3º lugar, no Jequitinhonha e Triângulo Mineiro e Alto Paranaíba; em 4º lugar, nas mesorregiões Central Mineira e Norte de Minas; e em 5º lugar, nas mesorregiões Campo das Vertentes, Metropolitana de Belo Horizonte e Oeste de Minas. A figura, a seguir, apresenta a distribuição, em números percentuais, dessas denominações toponímicas nas doze mesorregiões mineiras:

6 O levantamento e análise de dados foram realizados de acordo a divisão proposta pelo IBGE, que recorta o estado em doze mesorregiões, quais sejam: Campo das Vertentes (1), Central Mineira (2), Jequitinhonha (3), Metropolitana de Belo Horizonte (4), Noroeste de Minas (5), Norte de Minas (6), Oeste de Minas (7), Sul e Sudoeste de Minas (8), Triângulo Mineiro e Alto Paranaíba (9), Vale do Mucuri (10), Vale do Rio Doce (11) e Zona da Mata (12).

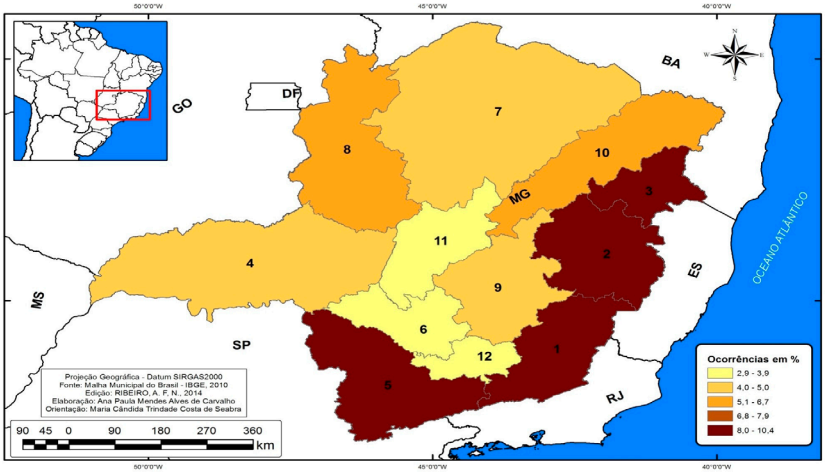


Figura 2. Distribuição dos hagiotopônimos nas mesorregiões mineiras em números percentuais.
Fonte: Carvalho (2014: 540)

Diante dos resultados apresentados, percebe-se que a porção leste e o sul do estado são as regiões que apresentam maior densidade hagiotoponímica, justamente a região por onde, direta ou indiretamente, inicia o povoamento mineiro. Acreditamos, em virtude disso, que a investigação dessa taxe, em território mineiro, tende a revelar valiosas informações no que concerne à organização histórica e sociocultural das mesorregiões mineiras.

Desse modo, após serem identificados, esses hagiotopônimos foram analisados a partir da observação de três categorias onomásticas, quais sejam: os nomes de *santos*, os nomes de *santas* e os nomes de invocações da *Virgem Maria* a que chamamos de mariotopônimos.

Constatou-se, então, que os nomes de *santos* são predominantes em território mineiro, como se pode verificar: de um total de 5649 hagiotopônimos, 3801 são dessa categoria onomástica, perfazendo um total de 67,3% do total de nomes, ao passo que os nomes de santas são representados por 1658 ocorrências ou 29,3% dos dados.

Os nomes relativos a invocações da Virgem Maria, por sua vez, figuram, no léxico toponímico do Estado, com um número bem mais reduzido, visto que foram quantificadas apenas 190 ocorrências de mariotopônimos, o que representa 3,4% do total de hagiotopônimos analisados.

Diferentemente dos dados históricos, os dados contemporâneos desta pesquisa foram analisados quanto à classificação dos acidentes geográficos, isto é, quantificaram-se

os acidentes físicos e os acidentes humanos a que se referem os hagiotopônimos sob análise, para cada categoria onomástica considerada, o que nos permitiu fazer as afirmações seguintes.

De uma forma geral, há, no léxico toponímico mineiro atual, a predominância de hagiotopônimos referentes a acidentes humanos, visto que dos 5649 nomes quantificados, 3563 são relativos a cidades, vilas, localidades, fazendas, dentre outros, perfazendo 63% do total de topônimos analisados. Os acidentes físicos –rios, córregos, serras etc.–, por sua vez, são representados por 2086 nomes, correspondendo, assim, a 37% dos dados.

No que se refere à classificação dos acidentes geográficos de acordo com as categorias onomásticas sob análise, verifica-se que, mantendo-se a tendência observada de forma geral, os acidentes humanos são predominantes em todas elas, o que, segundo Ananias (2013: 182), corrobora a tendência de predomínio de taxes antropoculturais, como é o caso dos hagiotopônimos, entre os acidentes humanos.

Ressalta-se que, para os mariotopônimos, o índice de ocorrências de denominações referentes a esse tipo de acidente é ainda mais significativo, pois corresponde a 94% dos dados, ou seja, dos 190 nomes referentes às invocações a Virgem Maria presentes em território mineiro, há apenas 6% –ou 12 ocorrências– de acidentes físicos.

Considerando os nomes de santos, observa-se que, do total de 3801 dados, 2268 são referentes a acidentes humanos, o que corresponde a 60% dos topônimos; já os acidentes físicos são designados por 1533 nomes dessa categoria toponímica, ou 40% dos dados. Os nomes de santas, em sua maioria, também se referem a acidentes humanos, visto que, dos 1658 topônimos quantificados, 1117 ocorrências (67%) são referentes a acidentes dessa natureza, ao passo que 541 nomes dessa categoria referem-se a acidentes físicos, o que equivale a 37% do total dessas denominações.

Dentre os acidentes humanos, merece destaque o número de fazendas. Verifica-se, pois, que a tendência geral nomes de santos > nomes de santas > maritopônimos observada nos dados se inverte:

	nomes de santos	<	nomes de santas	<	mariotopônimos
Número de fazendas =	65%	<	78%	<	85%

Essa inversão pode explicada pelo fato de que, em território mineiro, as fazendas são preferencialmente designadas por hagiotopônimos femininos, dentre eles,

a maioria refere-se a nomes de santas não reconhecidas oficialmente pela Igreja. Considerando a contribuição de Lima (1997), podemos aventar a hipótese de que tais nomes sejam hagiotopônimos aparentes atribuídos a esses acidentes pelo denominador para homenagear as mulheres que tinham como mais próximas, isto é, suas esposas, mães, filhas, avós etc.

Em relação aos acidentes físicos, que ocorrem em número bem menor para todas as categorias onomásticas analisadas, convém ressaltar que, dentre eles, seguidos dos ribeirões, os córregos figuram como o acidente dessa natureza que mais recebeu denominações hagiotoponímicas no estado, representando 59,4% dos nomes de santos, 66% dos nomes de santas e 42% dos mariotopônimos.

4.3. Discussão dos resultados

Da análise comparativa dos *corpora*, verificou-se que tanto contemporânea quanto diacronicamente, em território mineiro, há a seguinte tendência hagiotoponímica: nomes de santos > nomes de santas > mariotopônimos; sendo que, nos dados contemporâneos, a quantidade de nomes de santos (3801 ocorrências) é 2,3 vezes maior do que a quantidade de nomes de santas (1658 ocorrências) e 20 vezes maior do que o número de mariotopônimos (190 ocorrências). Seguindo essa mesma tendência, nos dados históricos, o número de ocorrências de nomes de santos (439) é 2,6 vezes maior do que as ocorrências de nomes de santas (170 ocorrências) e 12 vezes maior que o número de mariotopônimos (190 ocorrências). Em números percentuais, esse raciocínio pode ser representado da seguinte forma:

	nomes de santos	>	nomes de santas	>	mariotopônimos
Dados contemporâneos =	67,8%	<	26,3%	<	3,4%
Dados históricos =	67,3%	<	29,3%	<	5,9%

Diante do exposto, verifica-se que, em relação às demais categorias toponímicas analisadas, os nomes de santos se destacam tanto nos dados contemporâneos quanto nos dados históricos, e os hagiônimos Santo Antônio, São José, São João, São Domingos, São Sebastião, São Francisco e São Miguel apresentam-se, nessa ordem, com realce nos dois *corpora*.

Apesar de não aparecerem nos dados históricos, são muito recorrentes nos dados contemporâneos –figurando em todas ou em pelo menos em onze das doze mesorregiões– São Pedro (o 4º mais recorrente), São Bento, São Geraldo, São Lourenço, São Joaquim e São Vicente, dentre os quais São Geraldo se destaca, diferentemente dos demais que se sobressaem na Zona da Mata, por apresentar maior número de ocorrências no Vale do Rio Doce.

Com realce apenas nos dados históricos, São Gonçalo apresenta apenas 33 ocorrências nos dados contemporâneos, concentrando-se nas mesorregiões Norte, Noroeste de Minas e Metropolitana de Belo Horizonte. Apesar de apresentar, em relação aos mais frequentes, poucas ocorrências nos dois *corpora*, merecem destaque também, nos dados contemporâneos, Santo Inácio e São Tiago/Santiago. O primeiro porque a maioria de suas ocorrências está concentrada no Triângulo Mineiro e Alto Paranaíba, e o segundo porque, diferentemente da maioria, apresenta maior número de ocorrências nas mesorregiões centrais.

No que concerne ao emprego dos nomes de santas, destacam-se, tanto nos dados históricos quanto nos dados contemporâneos, Santa Ana, Santa Rita, Santa Maria, Santa Bárbara, Santa Isabel e Santa Luzia, sendo que, com exceção de Santa Rita que não aparece nas mesorregiões Campo das Vertentes e Central Mineira, os outros hagiônimos aparecem em todas as mesorregiões mineiras. Convém ressaltar aqui a presença significativa de Santa Anna nos dados históricos e Santa Ana/Santana (e outras variações) nos dados contemporâneos, sobretudo nas mesmas regiões do estado por onde passaram as bandeiras. Também com ocorrências em todas as mesorregiões, destacam-se ainda, na contemporaneidade, Santa Rosa, Santa Clara e Santa Helena.

Por outro lado, assim como ocorre com os hagiônimos masculinos, há nomes de santas que, diferentemente dos demais que são predominantes na Zona da Mata, destacam-se mais em outras mesorregiões, o que parece caracterizar um culto regional. Dentre esses nomes, podemos citar: Santa Luzia e sua variante Santa Lúcia, com realce no Sul e Sudoeste de Minas; Santa Terezinha, no Vale do Rio Doce e no Vale do Mucuri; e Santa Marta, no Vale do Rio Doce.

Nos dados contemporâneos, há ainda muitos outros hagiônimos: há os que são reconhecidos pela Igreja como os que mencionados anteriormente, os que são reconhecidos pelo culto popular, como Santo Expedito e Santa Efigênia, por exemplo, e, ainda, aqueles que parecem ser reconhecidos apenas pelo culto individual do denominador, isto porque não se encontrou, nas fontes consultadas, registro de seu culto oficial ou popular. Destes últimos, os nomes de santas sobressaem, uma vez que somam 42 hagiônimos, quais sejam: Santa Alda, Santa Angélica, Santa Apolinária, Santa Augusta, Santa Branca, Santa Cândida, Santa Cláudia, Santa Constância, Santa Delfina, Santa Elisa, Santa Elvira, Santa Elsa, Santa Emília, Santa Esméria, Santa Frutuosa, Santa Gabriela, Santa Glória, Santa Guilhermina, Santa Heloísa, Santa Hilária, Santa Isaura, Santa Isildinha, Santa Júlia, Santa Julieta, Santa Justa, Santa Laura, Santa Leocárdia, Santa Leonora, Santa Mafalda, Santa Márcia, Santa Mariana, Santa Maura, Santa Milana, Santa Olga, Santa Olinda, Santa Olívia, Santa Regina, Santa Rufina, Santa Silvéria, Santa Virgilina, Santa Virgínia e Santa Vitória.

Quanto aos hagiotopônimos masculinos, foram encontrados 18 desses nomes, a saber: Santo Apolônio, Santo Augusto, Santo Elias, Santo Onofre, São Cândido, São Cornélio, São Dinis, São Dimas, São Eduardo, São Fernando, São Gil, São Gotardo, São Julião, São Manoel/São Manuel, São Marino, São Modesto, São Sabino e São Vitorino.

Passando aos mariotopônimos, nota-se que, em relação à presença de hagiotopônimos em Minas Gerais, a principal diferença entre os dois corpora está nos topônimos relativos às invocações marianas, que, em números percentuais, nos dados históricos, é quase 2 vezes maior em relação aos dados contemporâneos. Acreditamos que essa diferença possa ser explicada por dois motivos principais: primeiro relaciona-se diretamente com o fato de que, a partir da segunda metade do século XVIII, o número dessas ocorrências começa a cair, conforme pode ser observado a seguir:

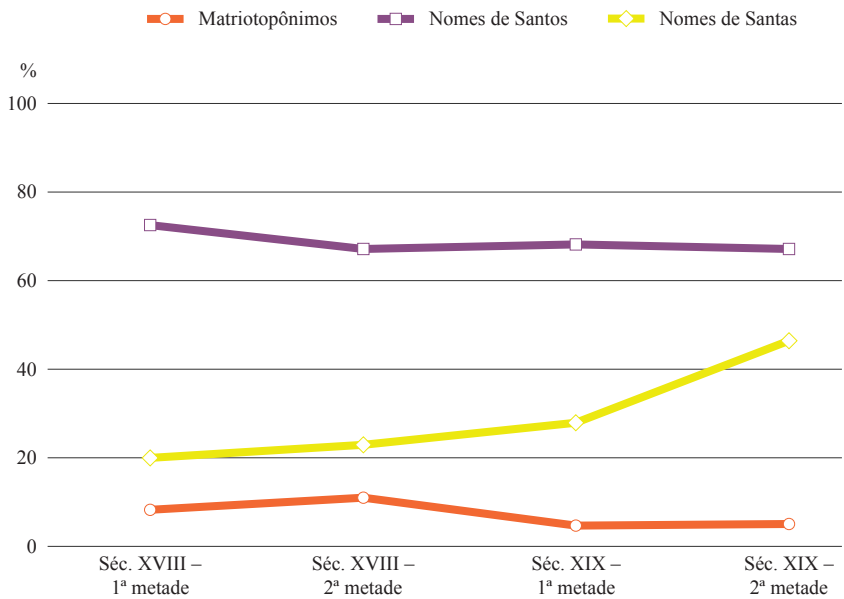


Gráfico 1. Evolução dos percentuais hagiotoponímicos ao longo dos séculos XVIII e XIX.
Fonte: Carvalho (2014: 644)

O segundo motivo refere-se à legislação toponímica –mais especificamente ao decreto-lei No 5.901 de 21 de outubro de 1943[7]– que, na primeira metade do

7 Esse decreto-lei dispõe sobre as normas nacionais para a revisão quinquenal da divisão administrativa e judiciária do país e pode ser consultado em http://www2.camara.leg.br/legin/fed/declei/1940-1949/decreto-lei-5901-21-outubro-1943-415891-publicacaooriginal-1-pe.html

século XX, recomenda, entre outras coisas, a não adoção de topônimos compostos de mais duas palavras. Assim, os municípios mineiros atualmente denominados Carmo do Cajuru, Conceição dos Ouros e Dores do Indaiá foram, anteriormente, chamados de Nossa Senhora do Carmo do Cajuru, Nossa Senhora da Conceição dos Ouros e Nossa Senhora das Dores do Indaiá, respectivamente.

Vale ressaltar, entretanto, que, muito anterior a essa lei federal, o denominador, mesmo que de forma inconsciente, já reduzia o sintagma toponímico, o que pode ser observado, por exemplo, na denominação do ribeirão, em torno do qual se deu o início do povoamento mineiro, que, mesmo sendo designado sob a invocação de Nossa Senhora do Carmo, recebeu apenas a forma específica 'do Carmo', constituindo, assim, um topônimo simples.

Nessa perspectiva, considerando os nomes de municípios, observa-se que, enquanto na atualidade apenas 5 municípios são denominados por mariotopônimos –Madre de Deus de Minas, Senhora de Oliveira, Senhora do Porto, Senhora dos Remédios e Virgem da Lapa–, aqueles que apresentam essa taxe como denominações anteriores somam 110 ocorrências. Observa-se que muitos desses se apresentam como reduções de invocações marianas, como os três municípios mencionados anteriormente.

No que concerne aos dados históricos, observa-se que, das 38 ocorrências de invocações marianas, 13 topônimos, ou 34%, aparecem apenas nos mapas consultados, isto é, não constam do banco de dados do ATEMIG, quais sejam Snra. das Candeyas (2), Sra do Livramento (1), N. S. da Lapa (1), N.S. das Correntes (1), N.S. de Nazareth (2), N.S. do Bom Sucesso (3) e N.S. do Desterro (3).

Convém destacar ainda que a diferença entre os dados históricos e os dados contemporâneos seria ainda maior, se desconsiderássemos destes últimos, os nomes mais recorrentes, quais sejam: Nossa Senhora Aparecida, Nossa Senhora das Graças e Nossa Senhora de Fátima. Estas –que não apareceram nos dados históricos– são invocações marianas recentes, sendo que, com 27 ocorrências, o título 'das Graças' data de 1830 e, com 11 ocorrências, o título 'de Fátima' data de 1917. A invocação a Nossa Senhora Aparecida, por sua vez, apresenta um número bem maior de ocorrências. São 75 topônimos em território mineiro. Esse título mariano, de ocorrência única nos dados históricos, apesar de não ser tão recente, uma vez que data de 1717, tem origem brasileira, no estado de São Paulo, o que poderia justificar, além de sua popularização tardia, os maiores índices desses nomes terem ocorrido nas mesorregiões Sul/Sudoeste de Minas e Triângulo Mineiro.

5. Considerações finais

Trazida pelos portugueses para o Brasil, no início do século XVI, a fé cristã e sua principal manifestação que é o culto aos santos foi ganhando extensão territorial, na medida em que se dava o povoamento das terras recém-descobertas. Isso porque, fincando cruzes e erguendo capelas para abrigar as diversas imagens de sua devoção, os desbravadores se apossavam do território, batizando-o, em sua língua e de acordo com a sua cultura.

É, pois, nesse contexto, que surgem os hagiotopônimos em território mineiro e que, ao longo dos séculos, mantiveram-se, sobretudo, em acidentes rurais, como as fazendas e os córregos, cuja motivação denominativa prende-se exclusivamente à subjetividade individual do denominador e, por serem referências em propriedades particulares, as alterações toponímicas são mais raras.

Por outro lado, os motivos para a escolha do nome de um munícipio –acidente urbano–, além de observar a legislação específica, pauta-se pela subjetividade coletiva e, muitas vezes, também por interesses políticos. Isso justifica o fato de muitas denominações hagiotoponímicas referentes aos municípios mineiros terem sido substituídas ao longo dos anos, principalmente no que concerne aos mariotopônimos. Entretanto, esses topônimos, pelas reduções referentes aos títulos que acompanham expressão Nossa Senhora, deixaram vestígios de sua origem religiosa em todo o território mineiro, o que pode ser observado pela quantidade considerável de denominações toponímicas atuais constituídas a partir dos determinantes 'Carmo', 'Conceição' e 'Dores', por exemplo.

Diante do exposto, concluímos, então, que o costume de dar aos lugares um nome relativo à devoção aos santos sempre foi uma constante nos diversos períodos da história mineira, desde o início do povoamento até os dias atuais, o que se revela como um verdadeiro patrimônio sócio-linguístico-cultural.

Referências bibliográficas

Ananias, Anna Carolina Chierotti dos Santos (2013). *Um passear pela toponímia das mesorregiões das mesorregiões de Toledo e Foz do Iguaçu*. Universidade Estadual de Londrina. Dissertação de mestrado inédita.

Ayerbe Iríbar, María Rosa (1982). "Influencia de la devoción popular en la onomástica", *Saitabi: Revista de la Facultat de Geografia i Història*, *32*, 5-16.

Barbosa, Waldemar de Almeida (1995). *Dicionário Histórico-Geográfico de Minas Gerais*. Belo Horizonte: Itatiaia.

Carvalhinhos, Patrícia de Jesus (2005). "Hierotoponímia Portuguesa. Os nomes de Nossa Senhora". Em *Livro de Resumos e Promagação do IX Congresso Nacional de Linguística e Filologia*, vol. IX, 75-76. Rio de Janeiro: Cifefil.

Carvalho, Ana Paula Mendes Alves de (2014). *Hagiotoponímia em Minas Gerais*. Universidade Federal de Minas Gerais. Tese de doutoramento inédita.

Carvalho, Francisco de Assis (2012). *Entre a palavra e o chão: memória toponímica da estrada real*. Universidade de São Paulo. Tese de doutoramento inédita.

Castiglioni, Ana Cláudia (2008). *Glossário de topónimos do bolsão sul-mato-grossense.* Universidade Federal do Mato Do Sul. Dissertação de mestrado inédita.

Chaves, Luís (1957). *Influências religiosas na formação da Toponímia e Antroponímia em Portugal.* Lisboa: Tip. Casa Portuguesa. http://bibliotecas.patrimoniocultural.pt/oarqueologo/OAP_S2_v3_1956/OAP_S2_v3_1956_150dpi_pdf/p177-210/p177-210.pdf (Consulta em 16.04.2014).

Costa, Antônio Gilberto *et alii* (2002). *Cartografia das Minas Gerais – da Capitania à Província*. Belo Horizonte: Editora da Universidade Federal de Minas Gerais.

Daix, Georges (2000). *Dicionário dos Santos: do calendário romano e dos beatos portugueses*. Lisboa: Terramar.

Dauzat, Albert (1926). *Les noms de lieux*. Paris: Delagrave.

Dias, Geraldo J. Coelho (1987). "A devoção do povo português a Nossa Senhora nos tempos modernos",*Revista da Faculdade de Letras. Historia, 4*, 227-256,. http://ler.letras.up.pt/uploads/ficheiros/2061.pdf.

Dick, Maria Vicentina de Paula do Amaral (1990a). *A Motivação Toponímica e a Realidade Brasileira.* São Paulo: Governo do Estado de São Paulo - Edições Arquivo do Estado.

Dick, Maria Vicentina de Paula do Amaral (1990b). *Toponímia e Antroponímia no Brasil. Coletânea de Estudos* (2.ª edição). São Paulo: Universidade de São Paulo.

Dick, Maria Vicentina de Paula do Amaral (2004) "Rede de conhecimento e campo lexical: hidrônimos e hidrotopônimos na onomástica brasileira". Em Negri Isquerdo, Aparecida, & Krieger, Maria das Graças, *As Ciências do Léxico: lexicologia, lexicografia, terminologia*, 121-130. Campo Grande, Mato Do Sul: Universidade Federal do Mato Do Sul.

Dick, Maria Vicentina de Paula do Amaral (2006). "Fundamentos Teóricos da Toponímia. Estudo de caso: o Projeto ATEMIG –Atlas Toponímico do estado de Minas Gerais (variante regional do Atlas Toponímico do Brasil)". Em Trindade Costa de Seabra, Maria Cândida (org.), *O léxico em estudo*, 91-117. Belo Horizonte: FALE / Universidade Federal de Minas Gerais.

Dick, Maria Vicentina de Paula do Amaral (2010) "Etnia e etinicidade. Um outro modo de nomear. Projetos ATESP/ATB." Em Negri Isquerdo, Aparecida, & Bocorny Finatto, Maria José, *As ciências do léxico: Lexicologia, Lexicografia, Terminologia.* Campo Grande (MS): Ed. Universidade Federal do Mato Do Sul; Porto Alegre: Editora da Universidade Federal do Rio Grande do Sul.

Diéguez González, Júlio (2001). "Alguns nomes de santos, do latim ao galego-português", *Cadernos Vianenses*, *30*, 63-79. http://gib.cm-viana-castelo.pt/documentos/20081013125552.pdf

Duranti, Alessandro (2000). *Antropologia Lingüística* (trad. de Pedro Tena). Madrid: Cambridge University Press.

Gaeta, Maria Aparecida J. V. (1999). "'Santos' que não são santos: estudos sobre a religiosidade popular brasileira", *Mimesis*, *1 (20)*, 57-76. http://www.usc.br/biblioteca/mimesis/mimesis_v20_n1_1999_art_05.pdf.

González Fernández, Rafael (2000). "El culto a los mártires y santos em la cultura cristiana. Origen, evolución y factores de su configuración", *Kalakorikos, 5*, 161-185. http://dialnet.unirioja.es/servlet/articulo?codigo=192201.

Lejeune, Sylvie (2002). "La religion dans la Toponymie" Em *Les Actes du FIG*. http://archives-fig-st-die.cndp.fr/actes/actes_2002/lejeune/article.htm.

Lima júnior, Augusto de (1978). *A capitania das Minas Gerais.* Belo Horizonte: Itatiaia.

Matoré, George (1953). *La méthode en lexicologie. Domaine Française*. Paris: Didier.

Megale, Nilza Botelho (2002). *Santos do povo brasileiro*. Petrópolis, RJ: Vozes.

Oliveira, Pedro Ribeiro de (1975). "Catolicismo popular e mudança social", *CEI Suplemento. Religiosidade Popular*, *12*, 3-11.

Piel, Joseph M. (1950). "Os nomes dos santos tradicionais hispânicos na toponímia peninsular". Separata de *Biblos*, *25, 26.* Coimbra.

Sapir, Edward (1961). *Linguística como ciência*. Rio de Janeiro: Acadêmica.

Seabra, Maria Cândida Trindade Costa de (2004). *A formação e a fixação da língua portuguesa em Minas Gerais: a toponímia da região do Carmo*. Universidade Federal de Minas Gerais. Tese de doutoramento inédita.

Tóth, Valéria (ed.) (2011). *Patrociny Settlment Names in Europe*. Debrecen-Helsinki. (Urálica Onomástica 8). University of Debrecen. http://mek.oszk.hu/10500/10518/10518.pdf (Consulta em 31.05.2014)..

Turrul i Rubinat, Albert (1985). "Hagiotoponímia a la Segarra", *Miscellània Ceverina-III (Ceverra)*, *3*, 11-28. http://www.raco.cat/index.php/MiscellaniaCerverina/article/viewFile/132161/182361

Zamariano, Márcia (2010). *Estudo Toponímico no espaço geográfico das mesorregiões paranaenses: Metropolitana de Curitiba, Centro-Oriental e Norte Pioneiro.* Universidade Estadual de Londrina. Tese de doutoramento inédita.

Zilles, Urbano (2007). "Lugar e veneração dos santos hoje", *Teocomunicação Porto Alegre, 158 (37)*, 491-507.