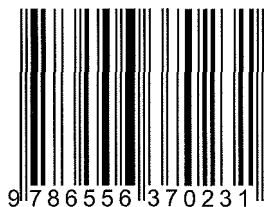


Com a obra *Estudos em Lexicografia – aspectos teóricos e práticos* pretendemos cooperar para as Ciências do Léxico, notadamente a Lexicografia, divulgando estudos lexicográficos que evidenciam a diversidade de olhares acerca do dicionário, oportunizando diferentes reflexões teóricas e práticas. Desse modo, esperamos que esta publicação alcance pesquisadores do léxico em distintas fases de sua formação acadêmica, bem como estudiosos de todas as demais áreas e interessados em geral no vocabulário de uma língua, que em muito se evidencia a partir de produções lexicográficas. Aos autores dos capítulos que compõem esta publicação e que gentilmente aceitaram nosso convite para participar desta coletânea, agradecemos pelas valiosíssimas contribuições.

Os Organizadores



DANIELA DE SOUZA SILVA COSTA
(ORGS.)

ESTUDOS EM LEXICOGRAFIA ASPECTOS TEÓRICOS E PRÁTICOS

gênero.
ciência. [Do lat. *scientia*.] S. f. 1. Conhecimento
tomar ciência. 2. Saber que se adquire pela
meditação; instrução, erudição, sabedoria. 3.
de conhecimentos socialmente adquiridos ou
dos, historicamente acumulados, dotados de
lidade e objetividade que permitem sua trans
estruturados com métodos, teorias e lingua
rias, que visam compreender e, poss., orient
eza e as atividades humanas. 4. Campo circ
entro da ciência (3), concernente a determin
u aspecto da natureza ou das atividades
mo, p. ex., a química, a sociologia, etc. 5. A
nhecimentos humanos considerados em co
progressos da ciência em nossos dias. 6. P
de intuitiva, sabedoria: “A ciência da a
elha e a minha, muito gente desconhece” (J
ão do Vale, da canção popular. Na *Árvore da*
ncia aplicada. Cien.



Pontes

Copyright © 2020 - Os organizadores, representante dos colaboradores
Coordenação Editorial: Pontes Editores
Editoração e Capa: Eckel Wayne
Revisão: Joana Moreira

CONSELHO EDITORIAL:

Angela B. Kleiman
(Unicamp – Campinas)
Clarissa Menezes Jordão
(UFPR – Curitiba)
Edleise Mendes
(UFBA – Salvador)
Eliana Merlin Deganutti de Barros
(UENP – Universidade Estadual do Norte do Paraná)
Eni Puccinelli Orlandi
(Unicamp – Campinas)
Glaís Sales Cordeiro
(Université de Genève - Suisse)
José Carlos Paes de Almeida Filho
(UnB – Brasília)
Maria Luisa Ortiz Alvarez
(UnB – Brasília)
Rogério Tilio
(UFRJ – Rio de Janeiro)
Suzete Silva
(UEL – Londrina)
Vera Lúcia Menezes de Oliveira e Paiva
(UFMG – Belo Horizonte)

PONTES EDITORES

Rua Francisco Otaviano, 789 - Jd. Chapadão
Campinas - SP - 13070-056
Fone 19 3252.6011
ponteseditores@ponteseditores.com.br
www.ponteseditores.com.br

SUMÁRIO

PREFÁCIO	7
LEXICOGRAFIA: A DICIONARIZAÇÃO DO LÉXICO.....	13
Maria da Graça Krieger	
O TESOURO DO LÉXICO PATRIMONIAL: PROPOSTA METODOLÓGICA PARA O ACERVO DAS VARIANTES XEOLECTAIS.....	33
Rosario Álvarez Blanco	
BREVE PANORAMA DA LEXICOGRAFIA DIALETAL E SUA APLICAÇÃO EM TERRAS BRASILEIRAS	57
Daniela de Souza Silva Costa	
DICIONÁRIO HISTÓRICO DO PORTUGUÊS DO BRASIL SÉCULOS XVI, XVII e XVIII: RELAÇÃO ENTRE LEXICOGRAFIA E HISTÓRIA.....	75
Clotilde de Almeida Azevedo Murakawa	
LEXICOGRAFIA DE CORPUS: CORPORA E CORPUS <i>QUERY SYSTEM</i>	97
Regiani Aparecida Santos Zacarias Camila Höfling	
DICIONÁRIO BÍLINGUE KAIWÁ-PORTUGUÊS: UM REGISTRO DO LÉXICO DA LÍNGUA KAIWÁ.....	117
Eliane Berendina Loman de Barros Aparecida Negri Isquerdo	
PARÂMETROS PARA A ORGANIZAÇÃO LEXICOGRÁFICA DE FORMAS HOMÔNIMAS HOMÓFONAS NÃO HOMÓGRAFAS DESTINADOS A DICIONÁRIOS PEDAGÓGICOS	137
Renato Rodrigues-Pereira	

O TESOURO DO LÉXICO PATRIMONIAL:
PROPOSTA METODOLÓXICA PARA O ACERVO DAS
VARIANTES XEOLECTAIS

Rosario Álvarez Blanco
Universidade de Santiago de Compostela

INTRODUCCIÓN

Novas metodoloxías e novas tecnoloxías – man con man, aquelas non posibles sen estoutras –, viñeron dar nova vida ás ciencias do léxico, con outras vertentes orixinais ou renovadas, ora nos estritos marcos das disciplinas clásicas ora en feliz interface con outras disciplinas. A informática posibilitou a creación e dispoñibilización universal de recursos (bancos de datos, corpus, repositorios, obras electrónicas...), de modo que os *investigadores dispomos de información accesible nunha cantidade e dunha calidade que non poderíamos imaxinar hai unhas poucas décadas*; por parte, permitiu soñar con novas maneiras de achegar e explotar os datos.

Atendendo ao convite dos organizadores, aos que moito agradezo a deferencia e a ocasión honrosa de facer parte da obra inaugural deste ambicioso proxecto editorial, vou presentar parte dos principios metodolóxicos que subxacen no *Tesouro do léxico patrimonial galego e portugués*, coordinado por min, no que participa unha ampla rede investigadores de centros universitarios de Brasil, Galicia e Portugal. O proxecto atende a catro aspectos importantes tanto para as ciencias do léxico coma para a dialectoloxía e a historia da lingua: a recolleita e sistematización da rica

e complexa variación formal, a clasificación semántica, a categorización gramatical e a distribución territorial. Chamámoslle *Tesouro* non só porque sexa un *thesaurus*, que tamén o é, senón porque o tipo de obras volcadas no seu banco de datos contén tal cantidade e tal variedade de xoias lingüísticas, ausentes da escrita e do repertorio de obras lexicográficas, que ningunha denominación podería acaer mellor ao seu conxunto. O TLPGP está accesible ao público, en aberto, desde febreiro de 2014 no portal do Instituto da Lingua Galega (<http://ilg.usc.gal/Tesouro/pt/>), mais segue en permanente construción¹. No mesmo lugar figuran todos os datos relevantes: características, obxectivos, equipo, corpus, forma de cita, titorial de axuda, etc. O texto que segue é descritivo, apoiado en casos particulares que serven de exemplificación para problemas e solucións; carece por iso dunha conclusión, compoñente que raramente debe faltar dun artigo científico. Dos problemas metodolóxicos que tivemos que enfrontar, escollemos o tratamento da variación formal (variantes: gráficas, fónicas, morfolóxicas) sistematizada baixo as correspondentes formas representativas ou de cita (lemas), o que obrigou a elaborar un rigoroso protocolo de lematización. Son os lemas os que finalmente permitirán estudar as áreas léxicas e a súa distribución territorial, mais tendo en conta que nunca descoidamos a rica variación que subxace a cada un deles.

O TESOURO DO LÉXICO PATRIMONIAL GALEGO E PORTUGUÉS (TLPGP)

Son innumerables os traballos de recolleita de léxico patrimonial que se deron e seguen a dar non amplo dominio lingüístico galego-portugués, sobre todo a partir da primeira metade do século xx, con diferentes crono-

1 Tivo diferentes axudas públicas específicas: AECID (2009), MICINN (España) FFI2009-12110 (2010-2012) e FCT (Portugal) PTDC/CLE-LIN/102650/2008 (2010-2012). Beneficiouse tamén das Axudas á consolidación e estruturación de unidades de investigación do SUG, Secretaría Xeral de Universidades, Consellería de Educación da Xunta de Galicia. Na actualidade beneficiase de dous proxectos relacionados aos que serve de apoio, *Léxico Patrimonial (Galicia-Portugal-Brasil)* (Convenio coa Secretaría Xeral de Política Lingüística da Xunta de Galicia / Axuda ao grupo de investigación FILGA [Xunta de Galicia – FEDER]) e *O patrimonio léxico da Gallaecia* (Programa INTERREG V-A España - Portugal [POCTEP] [0358_GEOAPARD_1_E-4]. PO FEDER Galicia 2014-2020 “Unha maneira de facer Europa”).

loxías segundo os países. En Galicia, que se incorpora máis tardiamente aos estudos lingüísticos², danse sobre todo a partir dos anos ‘60, foron especialmente intensos nas dúas décadas seguintes – programados desde a universidade –, e seguen a producirse por parte de diferentes entidades, públicas e privadas.

Os traballos ofrecen unha grande heteroxeneidade, desde todos os puntos de vista: promovidos como parte do cursus académico universitario (teses de doutoramento e mestrado, memorias de licenciatura, traballos fin de grao ou de máster, traballos dirixidos...) ou como contribución social á recuperación e preservación do patrimonio inmaterial (asociacións culturais, centros de ensino, amadores...), con e sen método científico apurado, cun obxectivo preferente lingüístico ou etnográfico, con formatos variados (vocabularios, descrições, láminas, mapas, bases de datos...), editados ou inéditos...

O TLPGP naceu para centralizar os resultados desa recolleita patrimonial heteroxénea e polifacética nun único banco de datos que sistematizase toda esa información, preservase os datos e permitise o acceso conxunto a eles. Para iso foi preciso actuar en dúas fronteas igualmente importantes, interrelacionadas:

a. Garantir a fidelidade ás fontes orixinais: na forma recollida localmente coa expresión gráfica dada polo autor (*variante*), no traslado rigoroso da información lingüística ou doutro teor colectada (*definición* en sentido amplo, imaxes, material etnográfico...), na *localización* xeográfica dos termos e na *referencia* remisiva á fonte orixinal. Unha relación de CORPUS permanentemente actualizada permite recuperar as referencias bibliográficas, de modo que o usuario sempre pode coñecer os datos básicos de cada fonte, sexa para valorar a súa autoridade sexa para acudir a ela á procura doutras informacións.

b. Establecer un nivel de uniformización que permita sistematizar os datos da base e facilite a súa recuperación. Por iso, foi preciso agrupar as diferentes

2 Para os interesados na historia sociolingüística da lingua galega (MONTEAGUDO, 2017, MARTÍÑO, 1998), as súas repercusións na produción científica nesta e desta lingua, c o auxe actual remitimos a Monteagudo/Álvarez (coord.) (2017). A base de datos *BILEGA* dá noticia puntual de toda a produción pasada e presente, cun rico sistema de buscas.

variantes baixo un mesmo LEMA³ (por caso, baixo o lema COADEIRA agrúpanse as variantes cognadas {*coadeira*, *cuadeira*, *coladeira*, *coladera*...}, seguindo un procedemento metodolóxico que procura comprender todas as eventualidades (§ 4). As localizacións xeográficas foron unificadas baixo o límite do *concello/municipio* en Galicia e Portugal, e da *mesorregião* no Brasil, usando códigos de referenciación xeográfica estandarizados; esta acción permite visualizar de forma automática a selección sobre o mapa dos tres territorios. A todos os rexistros foi atribuída a categoría gramatical, de acordo cuns criterios e un sistema de abreviaturas unificados. Todas as voces foron clasificadas semanticamente seguindo un sistema numérico xerarquizado que se rexe por unha serie de convencións acordadas desde o inicio polos participantes; entre outras moitas utilidades está a de poder discriminar facilmente entre voces homónimas ou entre acepcións dunha mesma voz (por caso, GATO 5.5 ['animal doméstico'], xunto a 3.4.1 ['parte do carro'], 4.1 ['flor do salgueiro'], 6.1 ['tipo de peixe'], 8.1.1 ['peza do teitume do hórreo'], 8.1.2 ['grampo'], 8.2 ['prendedor'], 8.4.1 ['oficio'], 8.4.2 ['ferramenta'], etc.).

O número de traballos que recollen léxico patrimonial con referencia xeográfica é tan grande e a tipoloxía tan variada que no TLPGP asumimos tamén como tarefa ir construíndo un *INVENTARIO* onde dar noticia de todos eles, con independencia de que formen ou non parte do noso corpus. O usuario obtén a información procurando por *territorio* (grosso modo, país ou países; no caso de Brasil engádese a rexión, p. ex. Brasil [Centro-Oeste]), por calquera das palabras da referencia bibliográfica ou, moi importante, pola localización xeográfica dos datos contidos nela a través dos topónimos (p. ex., Mato Grosso do Sul [Sudoeste (Ponta Porã)]).

O TLPGP naceu como un proxecto colaborativo aberto e así segue na actualidade, acrecentándose de forma continuada co volcado de novas obras; na pestana *EQUIPO* figura a nómina de participantes e universidades a quen pertencen, que segue renovándose. Os manuais e procedementos metodolóxicos están en acceso nun blog privado para o colaboradores e parceiros do proxecto, mais a disposición de todas as persoas e entidades que manifesten interese. Todos os investigadores e investigadoras

interesados están convidados a participar, coa guía e asesoramento da coordinadora ou doutras persoas do equipo con competencias específicas.

VARIANTES E LEMAS

Xunto coa variación natural existente na lingua, no noso labor de sistematización dos datos lingüísticos topamos coa diversidade derivada de tradicións grafemáticas diferentes, de convencións académicas dispares, de solucións metodolóxicas particulares, etc. De forma sucinta, no tratamento da variación detectada ao colectar e disponibilizar os datos léxicos, debemos lidar cos seguintes tipos de variación, con diferentes graos de transcendencia.

a. Variación gráfica

1. A diferenciación pode corresponder ás convencións de dous padróns diferentes, como é o caso de /ɲ/, representado por <nh> en portugués e <ñ> en galego, sen que iso supoña unha diferenza fónica real (port. *banho* = gal. *baño*). Engádase que a forma padrón muda no tempo, consonte os acordos que se van tomando para cada lingua; así, por caso, aquilo que nunha altura era gal. *abó* é hoxe grafado como *avó* (repárese en que só mudou a forma ortográfica, pois en galego, coma nos dialectos do N de Portugal, non existe /v/ e os grafemas <b, v> representan por igual /b/; cf. ÁLVAREZ 2018: 67-70).

Con frecuencia a variación grafemática contida nas fontes débese á vontade de marcar un trazo dialectal que implica fonemas e sons que non teñen prevista correspondencia inequívoca na forma gráfica padrón, coma gal. *ghalopada* ou *jalopada* (lema GALOPADA), pronunciados con «gheada» (isto é, co paso de /g/ a /h/ en amplas áreas de Galicia, cf. ALGa: III 176ss), ou coma port. *tchabe* (lema CHAVE), onde se marca a existencia do fonema africado /tʃ/ que os dialectos setentrionais portugueses comparten co galego (ÁLVAREZ 2018: 65-67). O mesmo vale para casos en que o autor ou autora quere marcar que o trazo dialectal consiste na indistinción de dous fonemas ou outro tipo de solución non

3 Usaremos como convención a *cursiva* cando nos refiramos a formas variantes e a *VERSALETA* cando nos queiramos referir ao lema que as agrupa.

admitida no padrón, coma a falta de oposición /b/:/v/ no portugués do norte (de onde *bolta*, *baca*, *binho*... correntes nos glosarios da rexión no canto das grafías estandarizadas <volta, vaca, vinho>).

Noutros casos a variación nos grafemas débese a convencións metodolóxicas adoptadas polos autores, coma o uso uniforme con ou sen <h> frecuente nos glosarios académicos galegos (p. ex., *binagre* por <vinagre> ou *erdeiro* e *papaostias* por <herdeiro, papahostias>), tendo en conta que na ortografía galega a distinción grafemática entre e <v> non ten repercusión nin guía na pronuncia (insistimos: sempre /b/) e que <h> inicial ou intervocálico é sempre mudo.

2. O uso de diacríticos e o emprego das maiúsculas tamén introducen variación gráfica, debida tanto ás diferentes tradicións normativas (port. *língua*, *lingueta*, *para-raios*, *erva-de-nossa-senhora* / gal. *lingua*, *lingüeta*, *pararraios*, *herba da Nosa Señora*) coma á heterografía dos autores (*herba de Santa María*, *herba de santa maría*, *herba de santa María*). Procuramos a lematización unificada na variante estandarizada, mais as normas ortográficas non son tan precisas como deberan, de modo que, ao mesmo tempo, optamos por un sistema de buscas que non teña en conta as letras capitais e por ofrecer ao usuario a opción “con/sen diacríticos”.

3. Máis difícil de resolver de forma uniforme e estandarizada é o grao de integración dos elementos dun composto, un problema que en portugués se agrava co uso do trazo: gal. ANO NOVO {*ano novo*, *anonovo*, *aninovo*} pero CABODANO {*cabodano*, *cabo d'ano*, *cabo de ano*} / port. ANO-NOVO {*ano-novo*, *ano novo*, *an'novο*} e LUA NOVA {*lua nova*, *luanova*}. Os dicionarios non sempre son concluintes sobre este aspecto das voces que tratan nin adoptan unha posición unánime (v. gr. Houaiss, CABO DE GUIA = CABO-GUIA; Infopédia, OLHO DE BOI ≠ OLHO-DE-BOI, ambos baixo OLHO-DE-BOI en Houaiss). Xa se comprende a extrema dificultade de tratamento das moitas voces do corpus popular non recollidas nas obras de referencia e que, de facto, presentan diferentes graos de integración gramatical: por caso, ¿port. *arco da santíssima trindade* debe ser lematizado como sm. ARCO-DA-SANTÍSSIMA-TRINDADE seguindo o modelo de *arco-da-velha*, *arco-iris*, *arco-da-chuva*, *arco-celeste*... ou como loc.

sm. ARCO DA SANTÍSSIMA TRINDADE, sen darmos por concluído o proceso de integración?

b. Variación fonética

Este tipo de variación non sempre foi anotada na forma variante⁴, e en caso afirmativo non sempre se fixo da mesma maneira. Mais, mesmo así, son moitísimos os fenómenos de fonética sincrónica e diacrónica recollidos nestas fontes, dando como resultado alteracións que se presentan como variantes formais do mesmo lema. Débese ter presente a frecuencia con que as monografías dialectais recollen diferenzas fonéticas que son propias de grupos de falantes espontáneos, máis ou menos iletrados, mais non caracterizan todo o seu segmento sociocultural nin moito menos se corresponden coa norma social da maior parte da comunidade que habita o territorio estudado. 1. Poden afectar ao vocalismo, con moita frecuencia ao átono como en {*minhoca*, *menhoca*}, {gal. *fociño*, *fuciño* / port. *focinho*, *fecinho*}, {*efectivo*, *afectivo*}, {*semana*, *somana*, *sömana*, *sumana*}, {*alameda*, *lameda*}, {*broa*, *boroa*} ou {*viagem*, *viage*, *viagi*, *viaja*}, por pórmos só exemplos dalgúns dos fenómenos máis recorrentes. Mais tamén afectan ao tónico, cunha alteración de timbre marcada polos autores con diacríticos, coma en {*esterco*, *estêrco* ['e], *estêrco* ['ε]} e {*ovo*, *òvo* ['ɔ], *ôvo* ['o]}, ou con grafemas diferentes, coma en {gal. *cuncha*, *concha* / port. *concha*, *cuontcha*}, {*dende*, *dinde*}, {*escoitar*, *escuitar*, *escutar*} e {*teiroa*, *teirúa*, *teiró*}.

Non é infrecuente comprobar que, a pesar de compartir as variantes, con diversos graos de incidencia, a tradición consolidou padróns diferentes nas dúas linguas, coma en gal. DEREITO {*dereito*, *direito*, *dreito*, *dereto*} ≠ port. DIREITO {*direito*, *dereito*, *dereto*, *dreto*} (listadas en orde de frecuencia).

2. Os fenómenos que afectan ao consonantismo son moitos máis e máis variados. Nuns casos trátase de coñecidos trazos descritos como

4 Non é infrecuente que as obras, sobre todo as de formato académico, ofrezan a carón unha transcrición fonética, total ou parcial, que nós trasladamos a un campo específico, unificada en AFI e usando a fonte Doulos SIL Compact. Neste apartado referímonos á información fonética que, adaptando a grafía convencional, se tenta representar na forma da variante.

dialectais nunha das linguas, como a xa mencionada «gheada» do galego, que deu como resultado o paso /g/ a /h/ en todas as posicións menos no grupo interior /ŋg/, onde oscila entre [ŋg, ŋk, ŋh], de onde {*gato, ghato, jato*} mais {*domingo, dominco, domingho, dominjo*}⁵; algúns, como a tamén mencionada conservación da africada xorda /tʃ/ nos dialectos do norte de Portugal e en todo o galego (de onde {*chorar, tchorar*}), axudan a comprender a intensidade do continuum xeolectal a ambos os lados da fronteira política.

Un lugar especial, polo interese na deriva galego-portuguesa, ocupan os fenómenos que teñen que ver coa nasalidade, ben por mantemento/perda do antigo trazo –coma en {*gando, gado*} ou {*miñoca, minhoca, menhoca, minoca, mioca, miocra, moca*}– ben por desenvolvemento dunha nasalidade non etimolóxica en determinadas condicións, coma en {gal. *nube* / port. *nuvem, nuve, nube, núvia, núbia, nuvre*} (lat. NUBE-).

Na mesma liña están fenómenos menos considerados pola tradición dialectolóxica, mais que poden dar lugar a diferenzas do padrón nas dúas linguas, como {*estalar, estralar*} para gal. ESTALAR e port. ESTALAR (ESTRALAR), gal. BIRLO {*birlo, bilro*} ≠ port. BILRO {*bilro, birlo, birle, birro*} ou gal. XORDO {*xordo, sordo*} ≠ port. SURDO {*surdo, sordo, xordo*}.

3. Son moitos os fenómenos aparentemente menos sistemáticos que apuntan a grandes áreas históricas, coma as solucións vocálicas en hiato, ditongo ou crase a partir de -ŌLA {*teiroa, teirúa, teiró, teiróá, teirúa*} fronte ás que desenvolveron unha consoante velar {*queiroga, queiruga*}; en non poucas ocasións testemuñan procesos históricos dos que non teríamos noticia se non fose pola variación detectada a través dos testemuños actuais.

Moitos máis son de difícil adscripción a un territorio concreto (asimilacións e disimilacións, metáteses, síncope, epénteses, metáteses...): gal. {*paliza, palliza*}, port. {*gusano, busano, buzano*}, port. {*minhoca,*

5 O son da gheada varía entre realizacións fricativas, aproximantes e aspiradas localizadas entre o veo do padal e a glote. Non é infrecuente que nestas obras, para representala de forma inequívoca, no canto de <g> se usase o grafema <j>, que representa /x/ en español. Téñase en conta que, ao se perder a antiga oposición /3:/j/ en favor de /j/ no galego moderno (ÁLVAREZ 2018, p. 58-64), o grafema <j> non ten valor propio na ortografía galega.

minoca}, {*canterla, canterna*}, {*madalena, madanela*}, {*arangaño, aranganho, angaranho*}, {*canterla, cantelra, cantrela, cantela*}, etc. Algunhas palabras cultas semellan resumir na súa variación toda a planopia de procesos adaptativos, como gal. ICTERICIA = port. ICTERICIA {*ictericia, intrizia, itrícia, itrize, tiricia, tirixa, tiriça, trizia, trisa* (port. [z]), *trisa* (gal. [s]), *triça, atiriz, ictriz, itriz, teriz, tiriz, triz*}.

4. Un caso especial é a variación do acento, por traslado da intensidade a outra síbaba, coma nos xa citados resultados de -ŌLA {*teiroa, teirúa, teiróá, teirúa*} ou en palabras soltas moi comúns: gal. {*cantiga, cántiga, cântiga*}; gal. {*paxaro, páxaro*} fronte a port. {*pássaro, pássero, pássuro, passo*}, gal. {*perito, périto*}, gal. {*lóstrego, lústrigo, lústrigo, lostrego, lustrigo*}.

c. Variación morfolóxica

1. É coñecido que boa parte da variación formal no léxico se debe á amplitude de procedementos de formación de palabras e a que nestes procesos morfolóxicos poden estar implicados morfemas derivativos diferentes (*barbela* / *barbadela*, *morrinhento* / *morrinhoso* / *morrinhão*, *xentuza* / *xentalla* / *xentallada*, *gentuça* / *gentança*...); ocorre mesmo que nos atopemos coa base da derivación⁶ e a palabra derivada compartindo o mesmo significado, como en *queixo* = *queixelo*, *ombro* = *ombreiro* ou *barba* = *barbela*. A identificación destas formacións como formas léxicas diferenciadas é, na maior parte dos casos, tarefa doada para o lingüista. Ora ben, non sempre o é discriminar cando estamos diante dun proceso fonético ou dun proceso morfolóxico, cuestión que, como se verá, é

6 Empregaremos a seguinte convención terminolóxica: *palabra base* é a forma representativa que serve de forma de cita para todo o paradigma, para todas as realizacións concretas que unha unidade léxica debe adoptar como *palabra de discurso*: *estar* serve de forma de cita ás palabras de discurso *estar, estou, estás, estabades, estaremos, estiven, estivese...* e *arco da vella* é a palabra base que representa as palabras de discurso *arco da vella* e *arcos da vella*. *Palabra gráfica* é a comprendida entre dous espazos en branco, de modo que podemos dicir que unha palabra base como *arco da vella* está constituída por tres palabras gráficas. Este uso de “palabra base” non obsta para que o termo “base” sexa utilizado, ademais, para aquela forma léxica da que parten outras, por derivación ou por composición (*queixo* → *queixelo, amparo* → *desamparo, sapo* e *concho* → *sapoconcho*) (ÁLVAREZ; FERREIRO 2002, p. 221-231 e 673-678).

determinante para a aplicación dos nosos criterios metodolóxicos: por caso, ante unha palabra como *centulo* cabe preguntarse se se trata dunha evolución metafonética a partir de *centolo* (correlato masculino de *centola*, con /'o/), se foi engadido un morfema *-ulo* doutra procedencia ou incluso se a similitude con *centolo* é casual.

A casuística é tan variada como se poida imaxinar. O inventario de variantes morfolóxicas tratadas no TLPGP pode documentar todos os tipos de formación de palabras en galego e portugués, ao longo do tempo; por parte, a orixe da variación pode non ser romance e provir dunha base etimolóxica diferenciada, coma en *acender* < ACCENDERE e *encender* < INCENDERE. Son aspectos que interesa ter en conta para o tratamento das variantes morfolóxicas á hora da lematización, pois só se reúnen baixo un mesmo lema as que sexan estritamente cognadas (§ 4). Por iso é fundamental determinar se se trata de sufixos ou prefixos diferentes (*tecedor* / *tecelán*, *cozedela* / *cozedura*, *retrasar* / *atrasar*) ou de variacións formais do mesmo formante (como DES- {*des-*, *es-*}, en DESCABAÇAR {*descabaçar*, *escabaçar*}), e distinguir estas á súa vez de variacións semellantes que se producen de forma analóxica (como en ESPERTO {*esperto*, *desperto*} < lat. vg. EXPERTUS, por EXPERRECTUS).

3. As variantes morfolóxicas, en canto que palabras de discurso, son lematizadas pola palabra base, seguindo as convencións xerais (infinitivo nos verbos, forma do masculino singular para os adxectivos, etc.) e algunhas convencións particulares. O criterio é fácil de comprender e doado de aplicar, pois, en efecto, na maioría dos casos, os substantivos son citados por unha forma singular, masculina ou feminina (sm *dedo*, sf *perna*), e son poucos e bastante ben identificables aqueles que se citan por unha forma de plural e establecen a concordancia sempre nese número; son substantivo maculino plural (sm pl) ou substantivo feminino plural (sf pl), coma os seguintes: *quitas*, *libras* ou *secundinas* ‘envolturas fetais’; port. *óculos* e gal. *lentes*; gal. *albizaras* (mais port. *alvissara+s*), gal. *cartos* ‘diñeiro’, port. *sarandalhas*.

Con todo, o asunto non está exento de problemas, pois non é doado establecer cando o uso consagrou xa a forma de plural como única opción

para unha acepción dada ou cando se trata dunha cuestión de simple maior frecuencia. Por caso, nas nosas fontes é frecuente que se anote gal. *augas* = port. *águas* como substantivo feminino plural co significado ‘vertente do tellado’; e certamente este é o uso máis frecuente, porque polo xeral as casas son a dúas ou a catro augas, pero unha ollada máis atenta aos datos revela que cada vertente é “unha auga” (e de feito compróbase que o tellado pode ser de só unha auga ou que a comunicación pode xirar, en singular, a propósito dunha soa). Máis difíciles de tratar son os casos en que hai unha especie de continuum entre o uso do singular e o uso do plural, que volvemos exemplificar con sf AUGA (significado xeral) e sf pl AUGAS (significado específico): ¿cal é o punto límite entre *tomar (a) auga* ‘beber H₂O, do tipo que sexa’ e *tomar as augas* ‘acudir á fonte apropiada ou a un establecemento balneario para beber auga mineromedicinal ou bañarse nela’, toda vez que sempre se implica unha porción ou tipo do substantivo non contable *auga*? Respecto dos substantivos que denominan obxectos formados por dúas pezas, escollemos a variante máis común como padrón en cada lingua: port. *calças* e *ceroulas*, gal. *cirolas* vs port. *calção*, gal. *calza*, *calzón* e *pantalón*. Son máis problemáticas as palabras descoñecidas, na forma ou no significado, porque non podemos discernir se a forma plural empregada denomina unha unidade formada por varias partes (o que xustificaría en orixe o número) ou ás varias partes da entidade descrita: por exemplo, ante a voz *abeiras* recollida como sf pl co significado de ‘beirado, beiral’, ¿debemos supoñer un lema en -AS sinónimo dos termos equivalentes dados (en singular en ambos os casos), ou un lema en -A, entendendo que o beiral ten tantas partes como augas ten o tellado e que cada unha delas é unha (a)beira? A falta de máis información optamos por confiar no criterio do autor, sempre que sexa congruente, mais a dúbida persiste e a atribución é revisable no futuro.

3. Un dos problemas con que nos atopamos adoito é determinar en que casos os morfemas avaliativos cumpren estritamente a súa función e en cales se converteron en formantes dunha nova unidade léxica. Non é infrecuente que, en lugar da base, se recolla a variante en diminutivo sen máis valor có afectivo, como parece ser en FELOSA {*felosa*, *folosa*, *fulosiña*}, mais tamén se pode dar que a formación avaliativa chegase a

lexicalizarse e deixase de ser sentida como tal (cfr. *joaninha*, *xoaniña*, a partir de *joana*, *xoana*). No mundo animal é frecuente que o diminutivo sexa empregado para significar ‘fillo de + base’ (*burro* → *burriño*, *burrinho*), a partir de ‘máis pequeno ca + base’, polo que resulta case imposible determinar en que casos a gramaticalización alcanzou o grao necesario para considerar xa unha palabra derivada (*BURRIÑO* ‘cría de burro ou burra’) e non unha realización avaliativa do lema *BURRO*; por iso nestes casos tomamos o acordo de lematizar pola forma base (‘denominación do animal’, sen ter en conta a idade), mesmo nos casos en que se produce unha acumulación de sufixos que aparentemente podería dilucidar a cuestión; por iso baixo *BÁCORO* van agrupadas tanto as variantes {*bácoro*, *bácora*, *bácaro*, *bácuro*, *bacoiro*, *bac’ro*, *bacro*, *bacre*} coma {*bacorinho*, *bacoriño*, *bacuriño*, *bòcrinho*, *bôcrinho*, *bacorín*, *bacórico*, *bacorito*} e *bacorinhozinho*⁷. Noutros, en fin, resulya evidente que o uso xa consagrou ambas as solucións como formas léxicas diferenciadas, con significados que disipan as dúbidas ao respecto, como en *CABALO* {*cabalo*, *cabalos* → *cabaliño*, *cabaliños*} fronte a gal. *CABALIÑOS* ‘carrusel’ ou a port. *CAVALINHO* ‘jogo de azar’.

4. Finalmente, e esta é a maior dificultade, está o problema das palabras compostas e das locucións, pois os autores ou ben non teñen criterios metodolóxicos precisos para diferencialas e tratalas adecuadamente ou ben actúan de forma moi dispar entre uns outros.

O labor de lematización implica recoñecer a existencia dunha locución aínda que na obra non estea grafada nin clasificada como tal na súa entrada. Así, por exemplo, un autor categoriza como adxectivos as variantes *couracho* e *coiracho* ‘Nu, despido, não vestido, descoberto’ e dá como variantes os adx. *encoiracho* e *incoiracho*; a reflexión e análise lévanos a concluír que non se trata en ningún caso dun adxectivo senón da locución adverbial (quizais tamén adxectiva) *em couracho* ‘em couro, nu’, ben testemuñada noutras fontes.

7 Non, en cambio, *bacorelo*, lema *BACORELO*, porque *-elo* (< lat. *-ellu*), aínda que deixou moitas pegadas no léxico e, sobre todo, na toponimia de Galicia e do norte de Portugal, non é un sufixo avaliativo produtivo de formas de diminutivo en galego e portugués. Para este e outros aspectos relacionados coa diacronía e diatopía dos morfos do diminutivo en galego (e portugués), remitimos a Álvarez (2016).

O labor de lematización implica tamén facer o camiño contrario, isto é, tratar como lemas independentes colocacións ou sintagmas doutro tipo que os autores confunden con locucións. Nestoutro exemplo, a autora dálle entrada como locución a gal. *sudar a piche* ‘suar moito’, que nós debemos xerbrar en v. *SUAR* (*moito*, *pouco*, *a fío*, *coma un carneiro*...) e na locución adverbial ou adxectiva *A PICHÓ* ‘a chorros, a cano aberto’, combinable con verbos como *beber*, *chorar*, *caer*... ou con substantivos como *viño*.

Non sempre é posible ter o mesmo grao de certeza sobre como proceder. Son moitas as expresións de carácter local recollidas neste tipo de fontes, para as que non hai máis información cá contida nelas (nin dicionarios, nin uso escrito adicional), polo que non é infrecuente ter que volver sobre delas cando aparecen novos datos.

A LEMATIZACIÓN

O proxecto TLPGP conta cun manual de lematización que fixa os criterios metodolóxicos xerais e contempla solucións convencionais unitarias para casos particulares. Nesta presentación, pola necesaria brevidade, exporemos en só seis grandes liñas os principios de actuación.

1. As variantes cognadas (gráficas, fonéticas, morfolóxicas) agrúpanse baixo un único lema. Isto obriga a un traballo de análise dos compoñentes e a unha certa pesquisa histórica, pois para formar parte do mesmo lema non abonda con compartir a raíz: cómpre ter en orixe os mesmos constituíntes estruturais e formais. Así, gal. {*maceira*, *maseira*, *mazaira*, *mazaieira*, *macira*} agrúpanse baixo *MACEIRA* porque todas proceden de *MATIAN-* + *-ARIA*, a mesma orixe de port. *MACIEIRA* {*maceira*, *macieira*, *maciêira*, *macieirê*, *maciêra*, *manciêra*, *mecieira*, *amacieira*}; é diferente o caso de *MAZANCEIRA* {*mazanceira*, *masanseira*, *mansanseira*}, correlato de port. *MAÇAZEIRA*, formados ao engadir ao nome romance da froita (gal. occid. *mazán* e port. *maçã*) o sufixo *-eira* ‘árbore de + base’, co incremento consonántico *-c/z-* a que a gramática de galego e portugués obriga tras nasal.

O mesmo criterio atangue aos préstamos, desde que sexan cognados. Así, o castelanismo *manzanilla* vai lematizado como variante de gal. MACELA.

2. Nos casos de homonimia non se discriminan os lemas con ningún dos marcadores convencionais habituais nas obras lexicográficas. Cada lema vai acompañado da categoría gramatical (*querer* sm / *querer* v) e do clasificador semántico (*banco* 8.1.2 [enxoval doméstico] / *banco* 6.2 [relacionado con animais non domésticos]), suficientes para discriminar nesta fase de tratamento dos datos.

3. Consonte o criterio de primar a variación léxica, lematízanse separadamente os heterónimos (CABRA / BODE, GODALLO, CASTRÓN, CABRÃO); os substantivos que presentan sufixos de xénero “especiais”, herdados doutros estadios lingüísticos, se callar non romances, e non produtivos (coma GALO / GALIÑA OU ABADÉ / ABADESA); e os substantivos ‘non animados’ en que o xénero serve para formar novas palabras (do tipo CESTO / CESTA). En cambio, unifícanse baixo a forma de cita ‘masc. sing.’ os substantivos con moción xenérica co valor ‘macho / femia’, do tipo GATO {*gato, gata*} ‘mamífero felino doméstico’, categorizado como “s” (substantivo: masculino ou feminino); non é aplicable a sf GATA ‘tipo de peixe predador’, nin a sm GATO ‘peixe de fondo’.

4. Ponse especial coidado en lematizar de forma independente os lemas simples (formados por unha soa palabra gráfica) e os lemas complexos (integran máis dunha palabra gráfica), estes sempre categorizados como locucións, de diversos tipos: s GATO, sm GATO, sf GATA fronte a loc sm GATO MONTÉS, loc adv A GATAS, loc v FACER A GATA, loc com COMA O CAN E O GATO...

5. O lema toma a forma padrón sempre que haxa algún tipo de ditame normativo ou formas consagradas polo uso a través dos dicionarios de referencia; por iso, como xa se dixo, o lema galego pode non coincidir co lema portugués. Cando hai máis dunha variante padrón, unifícanse a favor dunha delas, seguindo un protocolo establecido: así, aínda que en portugués son aceptadas tanto *couro* coma *coiro*, posto que os dicionarios se inclinan a favor da primeira, tomamos como lema COURO.

6. Con máis frecuencia da esperada, nestas fontes hai unha porcentaxe elevada de formas non dicionarizadas, polo que hai que tomar seguido decisións sobre a solución que debe figurar como lema. A pesquisa previa implica procurar nos dicionarios e en corpora voces da mesma familia, para indagar nas formas padrón adoptadas polas que, sen seren cognadas, poidan ofrecer algunha vía de solución; e buscar testemuños non só na propia lingua, senón na lingua alternativa (galego ou portugués), noutros romances peninsulares e aínda noutras linguas europeas. Cómpre coñecer, coa maior certeza posible, a etimoloxía, pois o criterio é escoller como lema a variante máis achegada á esperable por deriva normal a partir da formación etimolóxica: así, lematizamos *belairo* como VELARIO ‘alpendre’ por supoñerlle relación con lat. UELARIUM; perante a dúbida de se *isar* ‘encerrar o can’ pode ser unha variante fonética seseante (equivalente a *izar*), a documentación en leonés e outros dominios hispánicos que non presentan este fenómeno lévanos a lematizar como ISAR; ante as variantes *redocelo* e *rodocelo* ‘envoltura das tripas do porco’, ambas documentadas por igual no DdD e ausentes no corpus escrito (TILG), optamos por REDOCELO debido ao seu parentesco con *redeño* e ser a máis frecuente no uso informal la través dunha busca en Internet.

No proceso de lematización tense en conta, de maneira permanente, o léxico recollido e padronizado para a outra lingua. Ante unha variante gal. *troboxar* ‘tronar, toar’ lematizamos por TROBOXAR, tendo en conta port. *trovejar* e gal. torboada, torbón, trebón; a forma trasmontana *crua* ‘rigoroso’, ausente dos dicionarios e exemplificada no proverbio *Dos Santos ao Natal, inverno cru*, debe ser lematizada como CAROAL {*caroal, carual, cru*}, perfectamente documentada en galego⁸ (DRAG: s.v.: “adxectivo | Máis intenso ca outros en canto a algunha das súas calidades. | *Un inverno caroal é un inverno moi frío*” | formando parte ou non do mesmo refrán, moi coñecido).

8 Xa desde o século xviii, en frei Martín Sarmiento (1746-1755c): “en Lemos significa, cosa corriente, cotidiana, frecuente y continua y trivial. Item, al tiempo oportuno para obrar dicen: ahora é o tempo caroal. Catálogo VF 1745-1755” (DdD: s.v. CAROAL).

A VARIACIÓN LÉXICA: XEOSINÓNIMOS

Unha vez sistematizada e considerada toda a variación existente (variantes gráficas, fónicas e morfolóxicas contidas nas fontes), chegamos ao obxectivo último do TLPGP: proporcionar instrumentos para poder estudar a variación léxica, que é a que nos dan os lemas, tendo en conta todas as fases e diverxencias evolutivas testemuñadas para cada un.

Escollemos para exemplificar este aspecto o tratamento das moitas variantes de todo tipo recolleitas para denominar a *Mustela nivalis*, coñecida xeralmente como *doninha* en Portugal e como *donicela* (*doniña* ou *denociña*) en galego⁹.

1. Aplicando os criterios metodolóxicos do proxecto agrupamos as variantes gráficas, fonéticas e morfolóxicas baixo estes seis lemas básicos¹⁰ relacionados con *dona*.

Galego	Portugués
DONIÑA {doniña, doriña}	DONINHA {doninha, dóninha, dòninha, duni- nha, deninhê, dônina}
DENOCIÑA {denociña, denosiña, denuciña, dinusiña, denouciña, denonciña, denunciña, delonciña, delunciña, donexiña}	DONEZINHA {donezinha, dônezinha, dônózinha, dunuzinha, denozinha, denezinha, nodezinha}
DONICELA {donicela, dunicela, denicela, dini- cela, denocela, dinocela, dinucela, dinocela, denunciela, dernicela, gunicela}	
DONA DAS PAREDES {dona das paredes, dona de paredes, doniña de paredes}	

9 O DRAG remite de *doniña* para *donicela*, onde se di que ten como sinónimos *doniña*, *garri-diña*, *saltaparedes* e, como variante menos recomendable, *denociña*. Fronte á recomendación normativa, os datos aínda inéditos da pregunta 1445 do cuestionario do *Atlas Lingüístico Galego* (ALGa) mostran que a metade do territorio, ao NW dunha isoglosa que vai desde a desembocadura do Eo (na estrema NE) á do Miño (na estrema SW) teñen como voz patrimonial algunha das variantes que agrupamos baixo DENOCIÑA; o resto do territorio lingüístico do galego repártese entre DONICELA, a segunda en extensión, e DONIÑA, esta nunha ampla franxa meridional sobre a fronteira portuguesa. Nos dicionarios portugueses a forma é *doninha*, só coa mención de *comadrinha* como rexionalismo.

10 Reducibles a catro de considerarmos conxuntamente as dúas linguas. Lémbrese que se os lemas galego e portugués non son idénticos deben aparecer por separado. Así ocorre con DONIÑA e DONINHA, cunha diferenza puramente ortográfica, e con DENOCIÑA e DONEZINHA, que se agruparían baixo un único lema de se tratar de variantes da mesma lingua.

As formas de variación detectadas son recorrentes nos distintos grupos:

a. variación na vogal pretónica inicial, que pode manter [ɔ] primitivo (< lat. DŌMINA, gal. e port. mod. /'o/), sufrir procesos de elevación [o, u] e de palatalización [e, i], e probablemente de centralización en portugués;

b. variación na vogal pretónica medial, na serie posterior [o, u] ou na anterior [e, i];

c. variación na vogal átona final, en dialectos portugueses, *deninhê*;

d. cambio na oclusiva inicial /d/ > /g/ en contexto anterior a vogal velar (*gunicela*), nunha área que está nas proximidades da conca do río Sil;

e. cambio do fonema /n/ por outro alveolar no ataque da sílaba, /l/ (*delonciña*, *delunciña*) sobre todo no norte de Galicia (área mindoniense) e /t/ (*doriña*) moi esporádico;

f. variación entre <c> /θ/ e <s> /ʒ, s/ debido ao fenómeno do «seseo» en galego¹¹, todos correlatos parciais de /z/ en portugués padrón (equivalente aos antigos e dialectais /z/ : /ʒ/); realización palatalizada de /s/ ante /i/, *donexiña*;

g. despalatalización de /ɲ/ no morfema avaliativo, *dônina*;

h. formación dunha coda na sílaba interior ante /θ/, case sempre nasal (*denonciña*, *delonciña*, *denunciela*...), máis frecuente no norte de Galicia; só esporadicamente, semivocálica [ow], *denouciña*;

i. formación dunha coda líquida na sílaba pretónica inicial, *dernicela*;

j. no límite oriental, ditongo *-iela* non autóctono, que evidencia un préstamo de falas asturianas ou leonesas próximas;

k. metátese, coma en *nodezinha* (no mesmo lugar existen ambas).

11 En parte do occidente de Galicia non se xerou o fonema interdental /θ/ dos antigos /s/ : /z/ posteriores ao proceso de queda do elemento oclusivo dos fonemas africados dentoalveolares; a perda do trazo distintivo de sonoridade en galego reduciu a parella a /s/, que mantén a relación de oposición con /s/ (*cocer* /s/ : *coser* /ʒ/), mais a tendencia é á neutralización a favor dun deles, preferentemente do apicoalveolar (para o fenómeno, en conexión coas variación xeolectal do portugués, cf. ÁLVAREZ, 2018, p. 60-64). A inexistencia de /ɲ/ coñécese como «seseo».

Lematizamos as formas galegas do segundo grupo baixo DENOCIÑA porque esta é a variante mencionada no *DRAG*, se ben como opción léxica menos recomendable, a consagrada pola lexicografía e a máis documentada no uso; as correspondentes portuguesas foron lematizadas baixo DONEZINHA (e non baixo un verosímil DENOZINHA que se asemellaría máis á forma galega) porque esta foi a única variante que atopamos dicionarizada nesta lingua (raramente, en todo caso: “s. f. || (ant. e Beira) o mesmo que doninha. Cf. Livro de Esopo, p. 26, e Rev. Lus., XII, p. 313”, Aulete: s.v. DENOZINHA).

DONIÑA, DENOCIÑA/DONEZINHA e DONICELA non son reducibles a un único lema, porque, aínda aceptando que compartan a base de partida DOMINA > *dona* e en todos eles haxa un elemento diminutivo, non son estritamente cognados.

— *Doninha* é unha formación romance a partir de *dona*, seguindo o modelo de formación do diminutivo común a galego e portugués a partir dun substantivo grave rematado en vogal átona: *dona* → *don-iñ-a* / *don-inh-a*.

— *Donicela* representa un estadio lingüístico anterior, como pon en evidencia o uso de -C-ĒLLA (unha das formas do diminutivo latino) e non de -INA (forma adoptada polo bloque galego-portugués, mais que non tiña ese valor avaliativo en latín). Non hai obxeccións fundadas a que derive do diminutivo lat. vulg. *DOMNICELLA ← DOMINA (cf. fr. *demoiselle* [ant. *domnizelle*, *dameisele*, *demiselle*...], ing. *damsel*, oc. *donaisèla*, it. *donzella*...)¹². A diferenza de gal. *doncela* / port. *donzela*, da mesma orixe, non perdeu a vogal intertónica, como non o fixeron outras voces galegas en -icela que tamén parten dun diminutivo latino en -ICELLA: *gradicela* (~ *granicela*) de CRATEM, *ponticela* de PONTEM, *manicela* de MANUM. Trátase, pois, da evolución a partir dunha denominación de uso no latín galaico, aquel que daría lugar ao romance primitivo do que, andado o

tempo, se diferenciarían galego e portugués. A formación ten a mesma motivación cá xerada no propio romance e que se acabaría impoñendo no S de Galicia e en todo o territorio do portugués europeo, que resulta ser un calco da primitiva: *doniña/doninha* é dim. de *dona*, como *dom(i)nicella* o era de *domina*.

— *Denociña/donezinha* levantan máis interrogantes. A presenza da consoante interdental /θ/ en toda a área de non seseo disipa calquera dúbida sobre a posibilidade dunha formación avaliativa a partir de *donosa* –cunha motivación semellante a port. *bonitinha*, gal. *garridiña*, fr. *bellette*, etc.– (Corominas: s.v. DONAR), pois en tal caso a única opción evolutiva sería con /ʒ/, *DONOSIÑA. Cómpre orientar a pesquisa, unha vez máis, cara a lat. DOMINA e os seus descendentes, porque, ademais, esta motivación é a que diferencia o bloque occidental (incluído todo o dominio galego-portugués) do resto do dominio hispánico, como puxo de manifesto Menéndez Pidal (1926, p. 400) e corroboran outros datos xeolingüísticos posteriores (PATO, 2011; GARCÍA MOUTÓN, 1987); no mapa de Pato (2011, p. 48), confeccionado a partir dos datos españois do ALPI, apréciase ben a singularidade do occidente con DOMINA fronte a toda península (faltan os datos portugueses, mais todas as informacións apuntan a un predominio de *doninha*).

Cremos, por iso, que a base de todas as variantes que agrupamos baixo DENOCIÑA/DONEZINHA proceden dun antigo *doneziña*, a mesma formación de donicela mais incorporando o morfo do diminutivo que triunfou en romance, *donic-ela* > *donic-iña* ~ *donec-iña*. Non estraña a permanencia da forma máis apegada ao latín na área actual de *donicela*, pois o eixe oriental formado por Lugo (Lucus Augusti), Ourense e Braga (Braccara Augusta), cunha derivación cara ao leste ata Astorga (Asturica Augusta), estaba fortemente romanizado e detentou o poder político e cultural ata o momento en que o auxe de Compostela o desprazou ao eixe occidental. Esta é a forma que atopamos nas *Cantigas de Santa María*, nunha curiosa pasaxe en que, ademais de informar sobre a presenza da *doniña* no ámbito doméstico, o poeta salienta a adscrición xeolingüística desta variante léxica:

12 O incremento consonántico -c- entre a base e os morfos diminutivos latinos era opcional e afecta sobre todo á terceira, cuarta e quinta declinacións (STRODACH, 1933, p. 10); os testemuños románicos mostran que tamén DOMINA, da primeira, xerou un diminutivo deste tipo. As obxeccións de Corominas (s.v. DUEÑO) sobre a falta de ditongación en esp. *doncella* e o vocalismo pechado da mesma palabra en variedades occitanas e catalás, que o levan a propoñer -ĈILLA, non afectan nin ao galego nin ao portugués, sempre cun /ε/ que non pode proceder de -Ĭ- latino.

Este pesar foi por hũa | bestiola que muit' amava
el rei, que sigo tragia | e que mui ben criava,
a que chaman donezy□a | os galegos, e tirava
con ela aves das covas, | e de taes ome vee.
(CSM 354)

A motivación de DONA DAS PAREDES é clara, tanto pola base *dona* coma pola alusión a un dos hábitats preferidos, os ocos dos muros e paredes de pedra. Optamos por lematizar *doniña de paredes* baixo DONA DAS PAREDES por considerala unha forma con valor avaliativo, baseándonos na infrecuencia dentro dunha área máis extensa en que predomina o formante *dona* (SE do territorio galego, en continuidade coa área de *doniña*). Canto a *de/das*, optamos pola solución común na tradición lexicográfica galega, desde frei Martín Sarmiento (1762; *DdD*: s.v. DONA)¹³. Por tanto, entre DONIÑA, DENOCIÑA, DONICELA e DONA DAS PAREDES cobren complementariamente todo o dominio lingüístico galego.

2. No corpus hai outras denominacións que se recollen de forma esporádica e sen formaren áreas: gal. SALTAPAREDES; port. COMADRINHA¹⁴, NORINHA, BONITINHA, BICHINHA DAS PAREDES e FARRAPEIRA. No noso corpus falta *garridiña*, unha das opcións normativas recollidas no *DRAG*; nas obras lexicográficas recollidas no *DdD* figuran, ademais, *bonitiña*, *cazoleira*, *cazoleta* e *mustela*; no traballo de campo do *ALGa* (cuestión 1446, materiais inéditos), engádense aínda *tixola vella* e *rabo de tixola* (pexorativos, coma *cazoleira*) e *ratonera*. Persiste a noticia de que no pasado foi domesticada coa mesma función cazadora do gato doméstico, pero a práctica foi abandonada porque facía grandes destrozos na roupa (*farrapeira*) e na louza da casa (*cazoleira*); tamén ten sona de mal xenio, a punto de servir como prototipo da furia (“En Baños de Molgas, *parecía unha doniña de paredes aquela muller, cómo se puxo...*”, Rivas 1988, apud *DdD*: s.v. DONIÑA); e, sobre todo, crese que é moi venenosa, astuta e vingativa, polo que se debe evitar todo contacto ou proximidade. Di un dos varios refráns alusivos ao

13 O criterio de lematización, nestes casos, é non ter en conta diferenzas entre certos elementos gramaticais con pouca carga semántica, como a presenza / ausencia de artigos, opción entre determinantes, diferente número nalgún dos compoñentes da expresión...).

14 Xa en Bento Pereira, *Thesouro*, 1697 (apud *CLP*: s.v. COMADRINHA). Compárese con esp. *comadreja*.

seu carácter mortífero: *Se te morde a donicela, busca viño e busca vela, que mañá che darán terra*. Así que mellor evitala ou, en todo caso, congraciarse con ela: de aí o uso de formas cariñosas ou de eufemismos, coma se o propio nome fose tabú ou se quixese evitar que detecte que falan dela; noutros casos son recriminatorias, rozando o insulto.

Non está clara a fronteira entre as formas denominativas desta entidade (eufemísticas ou non) e as apelativas, estas destinadas a apostrofar cariñosamente o animal para captar a súa benevolencia ou a increpala, coma nos recitados ou ditos populares recollidos nestas pasaxes:

Á doniña de paredes: *Denoncela, garridiña, busca a túa parediña; non tes capa, nin mantela, e tampouco sabes dela*. Dinlle *garridiña, bonita*, por mor de amansala, que cando a ofenden coa aquela de *donicela, tarabela, cazola vella, cazola vella*, pode alritarse e pegar a correr tras dun con dente mortal (RIVAS, 2011, p. 43).

Da donicela din qu'agarra a un cos dentes, e non o solta mentras non orneen sete burros e toquen sete campanas de sete eirexas. Dinlle: — *Donicela, bonitiña, garridiña...* Mais cando non lle teñen medo, dinlle: — *Donicela torresmeira, borrallenta*. — *Doniscela calderona malo raio que te coma* (RISCO, 1933, p. 452).

Para enfadarla dicen los niños: “*mouriña, mouriña*”, y para contentarla: “*branquiña, branquiña*” (A. OTERO, 1977, apud *DdD*: s.v. DENICELA).

No “Cancioneiro da Terra Chá (Pol)” (1980), de Isaac Rielo Carballo, dise que para encirriquitala e adoecela dinlle: *Dunicela fea, fea feixe de palla, cu de caldeira, fai un burato na lareira*. (...) Se lle dis: *Donosiña, garridiña!* ou *Donicela bonitiña!*, balancéase fachendosa. E se alguén fala mal dela, seica o persigue sen descanso. (REIGOSA, consulta 18/04/2019).

Ben se percibe que a relación de apelativos cariñosos podería ser enorme (*ruliña, pombiña, meniña, guapiña, carrapucheiriña, feiticeiriña, pombiña...*), sen que iso significase que todas fosen formas denominativas de *doniña*.

3. Non hai rexistros patrimoniais de lat. *MUSTELLA* e posibles descendentes. Nin no noso corpus nin noutras fontes hai indicios que verifiquen a información, dubidosa e pouco precisa, de Menéndez Pidal engadida en nota á afirmación de que “*mustela* tiene algún uso en gallego”: remite ao dicionario de Cuveiro Piñol, indicando que non a localiza, e engade “me dicen que *mustela* se usa al Sur de Ourense, en Verín” (1926, p. 397). A entrada *mustela* en Cuveiro (1876) e Filgueira et al. (1926) debe entenderse como inclusión da forma culta, atribuída de forma excesiva á doniña en lugar de facelo a todo o xénero dos mustélidos. Canto ao sur de Ourense, os datos son contundentes a favor dun uniforme *doniña*, salvo no límite oriental, xa fóra da área de Verín, onde concorre con *dona das paredes*.

No caso do galego, esta realidade é indisociable do feito de que na maior parte do territorio se documenten descendentes dunha forma que debeu ter moito vizo no latín galaico, *dom(i)nicella*, ben de forma directa (*DONICELA* e variantes) ben coa adaptación ao novo sufixo avaliativo (*donoziña* > *DENOCIÑA* e variantes); e onda non, foi *domina* a forma latina prevalente, non a clásica *mustella*.

4. No estado actual do TLPGP só é posible obter cartografía por separado de cada un dos lemas formalmente diferentes: isto é, se o lema GL, PT-EU e PT-BR coincide (p. ex. *ÂNCORA* ~ *ÂNCORA*), visualízase sobre os mapas de Galicia, Portugal e Brasil; mais se non é así (por exemplo, gal. *DENOCIÑA* / port. *DONEZINHA*), a consulta por lema só dá os resultados e cartografía procurados.

A visualización dos distintos lemas en variación para un mesmo concepto – isto é, dos xeosinónimos – é un obxectivo previsto desde o inicio do proxecto, mais forma parte dunha segunda etapa. Neste momento dispomos dunha aplicación informática que facilita a vinculación dos xeosinónimos e estamos a procesar os datos con esta finalidade, non de forma universal senón a partir dunha relación preestablecida de conceptos dunha serie de campos semánticos (fauna salvaxe, animais domésticos, flora, partes do corpo, xogos tradicionais, vinicultura, panificación, ferramentas, construcións...). O obxectivo é ofrecer unha nova utilidade do TLPGP, coa distribución espacial das variantes léxicas, a partir dos lemas, en construción permanente e ampliación progresiva.

REFERÊNCIAS

- ÁLVAREZ, Rosario. Diacronía e diatopía. Achea á variación morfolóxica no diminutivo en galego. In: RODRÍGUEZ GUERRA, Alexandre (ed.). **Lingüística histórica e dialectoloxía**: coordenadas do cambio lingüístico. Vigo: Servizo de Publicacións da Universidade de Vigo, 2016, p. 101-119.
- ÁLVAREZ, Rosario. Galego e portugués. Afinidade e diverxencia. In: CARRILHO, Ana Rita et al. (ed.). **Ao encontro das línguas ibéricas**. Covilhã: LusoSofia Press/Universidade da Beira Interior, 2018, p. 51-80.
- ÁLVAREZ, Rosario; FERREIRO, Xosé Xove. **Gramática da lingua galega**. Vigo: Galaxia, 2002.
- AULETE. **Dicionário Aulete Digital**. Lexicon Editora Dixital. Disponível em: <http://www.aulete.com.br/>. Acesso em: xaneiro 2019.
- COROMINAS, Joan; PASCUAL, José A. **Diccionario crítico etimológico castellano e hispánico**. 6 vols. Madrid: Gredos, 1991.
- CUVEIRO PIÑOL, Juan. **Diccionario Gallego**. Barcelona: Establecimiento tipográfico de N. Ramírez, 1876.
- FRANCISCO, García Gondar (dir.). **BILEGA**. Bibliografía informatizada da lingua galega. Santiago de Compostela: CIRP. Disponível em: <http://www.cirp.gal/bdo/bil/bilega.html>. Acesso em: xaneiro 2019.
- GARCÍA, Constantino; SANTAMARINA, Antón (drs.). **ALGa**: Atlas Lingüístico Galego. Vol. III: Fonética. A Coruña: Instituto da Lingua Galega/Fundación Pedro Barrié de la Maza, 1999.
- GARCÍA MOUTÓN, Pilar. Motivación en nombres de animales. **Lingüística española actual**, IX, 1987, p. 189-197.
- GRUPO PORTO EDITORA. **Infopédia**. Dicionários Porto Editora. Disponível em: <http://www.infopedia.pt>. Acesso em: xaneiro 2019.
- HOUAISS, Antônio et al. **Dicionário Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva, 2001 (versão eletrônica).
- MARIÑO PAZ, Ramón. **Historia da lingua galega**. Santiago: Sotelo Blanco, 1998.
- MENÉNDEZ PIDAL, Ramón. **Orígenes del español**. Estado lingüístico de la Península Ibérica hasta el siglo XI. Madrid: Espasa-Calpe, 1926.
- METTMANN, Walter (ed.). **Cantigas de Santa Maria**. Vigo: Xerais [Reed. de Coimbra: Acta Universitatis Conimbrigensis], 1981 [1959-72].

MONTEAGUDO, Henrique. **Historia social da lingua galega**. Vigo: Galaxia, 2017.

MONTEAGUDO, Henrique; ÁLVAREZ, Rosario (coord.). **Labor Histórico 3 (1)**. Lingüística Galega. A contribución do Instituto da Lingua Galega. 2017. Disponível em: <http://dx.doi.org/10.24206/lh.v3i1>. Acesso em: xaneiro 2019.

PATO, Enrique. Sobre geografia léxica española: distribución y áreas léxicas de la mustela. **Dialectologia** 6, 2011, p. 45-53.

REAL ACADEMIA GALEGA. **Dicionario da Real Academia Galega**. A Coruña: RAG. Disponível em: <https://academia.gal/dicionario/>. Acesso em: xaneiro 2019.

REIGOSA, Antonio. **Galicie encantada**. Enciclopedia da fantasía popular de Galicia. 2005. Disponível em: <https://galiciaencantada.com/index.asp>. Acesso em: abr. 2019.

RISCO, Vicente; RODRÍGUEZ MARTÍNEZ, Amador. Folklore de Melide. In: OTERO PEDRAYO, R. *et al.* **Terra de Melide**. [Santiago de] Compostela: Seminario de Estudos Galegos, 1933, p. 425-535.

RIVAS QUINTAS, Elixio. **Xogos e enredos**. Ourense: ed. Autor, 2011.

SANTAMARINA, Antón (dir.): **Dicionario de dicionarios**. Santiago de Compostela: Instituto da Lingua Galega. Disponível em: <http://sli.uvigo.gal/DdD/index.html>. Acesso em: xaneiro 2019.

SANTAMARINA, Antón; GONZÁLEZ SEOANE, Ernesto; ÁLVAREZ DE LA GRANJA, María. **Tesouro informatizado da lingua galega (Versión 4.1)**. Santiago de Compostela: Instituto da Lingua Galega. Disponível em: <http://ilg.usc.es/TILG/>. Acesso em: xaneiro 2019.

STRODACH, George Kleppinger. **Latin Diminutives in – ELLO/A - AND – ILLO/A -**. A Study in Diminutive Formation. Linguistic Society of America. [= *Language*, 9.1, 1933, p. 7-98]. Disponível em: <http://www.jstor.org/stable/522000>. Acesso em: xaneiro 2019.

VALVERDE, Filguera *et al.* **Vocabulario popular Galego-Castelán**. [Santiago de] Compostela: Seminario de Estudos Galegos, 1926.

VERDELHO, Telmo (dir.): **Corpus Lexicográfico do Português**. Universidade de Aveiro/CLUL. Disponível em: <http://clp.dlc.ua.pt/Inicio.aspx>. Acesso em: xaneiro 2019.

BREVE PANORAMA DA LEXICOGRAFIA DIALETAL E SUA APLICAÇÃO EM TERRAS BRASILEIRAS¹

Daniela de Souza Silva Costa
Universidade Federal de Mato Grosso do Sul

INTRODUÇÃO

O dicionário é uma fonte relevante para o conhecimento das línguas, uma vez que descreve a norma linguística sob diversos vieses, a depender de sua tipologia. Nesse sentido, também contribui sobremaneira para o ensino de idiomas e para o conhecimento das variedades linguísticas, inerentes a todos os sistemas. Tendo em vista esse caráter das obras lexicográficas e sua importância, este texto tem como objetivo apresentar brevemente um panorama da Lexicografia, desde suas primeiras produções, centrando-se na Lexicografia Dialectal, um de seus ramos de estudo, com a apresentação de alguns trabalhos produzidos no Brasil com base nesse aporte teórico.

Isso porque a história da Lexicografia remonta à Antiguidade, de modo que, desde o surgimento dos primeiros dicionários, muito antes de ela se instituir como disciplina linguística, ainda no século XIV, via-se a importância de compreender as línguas, de fazer conhecida a forma como outras sociedades se comunicavam e também para promover a interação entre as nações. A preocupação maior dos filólogos e gramáticos responsáveis pelas obras, até então paralexográficas², era a correção de

1 O presente trabalho foi realizado com apoio da Fundação Universidade Federal de Mato Grosso do Sul – UFMS/MEC – Brasil – e desenvolvido no Laboratório de Códigos e Linguagens – UFMS/CPAQ.

2 Obras paralexográficas são aquelas cuja produção ainda não tinha como base preceitos da Lexicografia teórica.