

DSpace da Universidade de Santiago de Compostela

<http://dspace.usc.es/>

Instituto da Lingua Galega

Rosario Álvarez (2016): “*O Tesouro do léxico patrimonial galego e portugués. Concepción, metodoloxía, obxectivos e utilidades*”. *40º Aniversário do Centro de Linguística da Universidade do Porto. Colóquio Comemorativo*. Faculdade de Letras. Porto, 24-25 de novembro de 2016.



You are free to copy, distribute and transmit the work under the following conditions:

- **Attribution** — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
- **Non commercial** — You may not use this work for commercial purposes.

O Tesouro do léxico patrimonial galego e portugués

Concepción, metodoloxía, obxectivos e utilidades

Rosario Álvarez

Instituto da Lingua Galega. Universidade de Santiago de Compostela

rosario.alvarez@usc.es

Lema

Variante

Procurar Lema...



Pesquisa avançada ▾

Sensível a diacríticos

O projeto *Tesouro do léxico patrimonial galego e portugués* é uma base de dados lexical que permite o acesso rápido e fácil à informação contida em obras sobre o léxico dialectal do galego, do português europeu e do português do Brasil.

O *Tesouro* permite consultar facilmente toda a informação completa contida em obras lexicográficas e dialectais que nem sempre se encontram acessíveis aos investigadores. A informação sobre as fontes originais é fornecida na sua totalidade e encontra-se organizada a partir de variantes, lemas, classificação semântica e localização geográfica. A informação textual complementa-se com uma representação cartográfica da distribuição espacial das formas.

O projecto do *Tesouro* teve início em 2009, promovido pelo Instituto da Lingua Galega (USC). Actualmente, colaboram nele 20 universidades da Galiza, do Brasil e de Portugal.

As pesquisas na base podem realizar-se a partir das formas lexicais (variantes ou lemas), da localização (Brasil, Galiza e Portugal) ou do campo semântico ao qual estão associadas.

O *Tesouro* continua a enriquecer-se com a participação de todos os colaboradores. Desde 2014, a base de dados do *Tesouro* está acessível de maneira livre e gratuita.



Canastro



Dados gerais

Obras introduzidas: 144 (Brasil 26, Galicia 62, Portugal 55)

Número total de registos: 198649

Número total de lemas: 52721

Número de variantes: 85484

Forma de citação

[\[txt\]](#) [\[bib\]](#) [\[ris\]](#) [\[rdf\]](#)

Álvarez, Rosario (coord.): *Tesouro do léxico patrimonial galego e portugués*. Santiago de Compostela: Instituto da Lingua Galega. <<http://ilg.usc.es/Tesouro>> [Consultado: <data>]

▶ *Tesouro*

- Un gran banco de datos de léxico con localización xeográfica
 - Fontes: traballos previos
 - Múltiples autores, variados enfoques, éditos / inéditos, etc.
 - Condicións básicas: recollida de campo e localización xeográfica (máx. concello).

▶ *do léxico patrimonial*

- Inventario do léxico patrimonial, con toda a información que consta nas fontes.

▶ *galego e portugués*

- Territorio galego (Galicia + territorio limítrofe galegófono)
- Portugal + enclaves limítrofes polo leste / Brasil

▶ Cal é o obxectivo? Que ofrece?

- Inventario de fontes: busca por autor / territorio
- Acceso a todas as formas lexicais documentadas: lema / variante.
- Información lexicográfica ou doutro teor, para cada voz. Localización bibliográfica e xeográfica.
- Lematización, clasificación gramatical e clasificación semántica.
- Cartografía: distribución espacial de lemas e variantes.
- Descarga de arquivos con información lexicográfica e mapas.

Metodoloxía

▶ **Coñecemento** e localización da obra

- Ficha de **inventario**: ficha, autor, data, tipoloxía, territorio, citación, ubicación...
- O **inventario** ofrécese como un servizo bibliográfico para os investigadores

▶ **Traslado** dos datos de cada obra seleccionada

- Nunha **táboa** propia. Propósito de principio: manter a literalidade da fonte
- **Homoxeneización**: número de campos, características de cada un, convencións

1	2	3	4	5	6	7	8	10	11	12	13
variante	fonética	categoría	definición	exemplos	comentarios manipulador	páxina	sección	referencia bibliográfica	código xeográfico	imaxes	remisións

▶ **Tratamento** lingüístico de cada táboa. Criterios unificados

9	16	17	6
clasificación semántica	lematización	clase e categoría	comentarios manipulador

▶ **Incorporación** á base de datos co **corpus** común. Acceso aos usuarios

Metodoloxía: criterios de lematización

- ▶ **Manual** a disposición dos membros do equipo. Blog de traballo
 - Criterios, procedementos, exemplos. Elaboración colaborativa e permanente
 - Versións galega e portuguesa, on-line
- ▶ **Criterios básicos**
 - Dous **lemarios**: galego / portugués. Desideratum: relacionalos na fase seguinte
 - Recorrer ás formas **padrón** en galego e en portugués (*dereito / direito*). Actuar
 - cando hai máis dunha forma padrón (p. ex., gal. *coio* e *croio*, port. *redemunho* e *redemoinho*);
 - cando falta o padrón (p. ex., gal. *baldema*, *baldela* ‘póla esmoucada da parra’).
 - **Unificar** baixo o mesmo lema variantes gráficas e fonéticas. **Non unificar** variantes con afixos diferentes (salvo excepcións acordadas).
 - Establecemento dun **novo lema** a partir de variantes
 - Proximidade á forma “esperada” por etimoloxía
 - Sanción polo uso maioritario ou convencional fóra dos dicionarios (internet, escrita...)
 - Documentación na tradición escrita ou lexicográfica
 - Existencia na outra lingua (ou linguas veciñas)
 - Procura de congruencia: familia de palabras, procedementos similares, etc.
 - Hai un protocolo específico sobre **lemas simples e complexos**.

Sensible a diacríticos

Localización (2)

Brasil

Portugal

Campo semántico (3)

0.0 - Non clasificable

6.1 - Animais silvestres (denominacións)

7.1 - Ser humano (aspectos físicos, psíquicos e condutuais)

Lemas (2)

minhoca (pt-eu)

minhoca (pt-br)

Variantes (6)

minhoca (pt-br)

minhoca (pt-eu)

minoca (pt-eu)

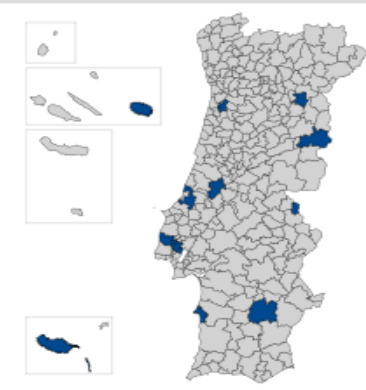
menhoca (pt-eu)

mioca (pt-eu)

mióca (pt-eu)

Resultados atopados (17)

menhoca [mi'ɲokɐ] s.f. Minhoca. Baptista 1967:333. LEMA: minhoca sf.	PT-EU	6.1
menhoca [mi'ɲokɐ] s.f. Minhoca. Braga 1971:284. Vid. mioca . LEMA: minhoca sf.	PT-EU	6.1
menhoca [mɲokɐ] s.f. Minhoca. Delgado 1970:344. LEMA: minhoca sf.	PT-EU	6.1
menhoca [mɲokɐ] s.f. Minhoca. Delgado 1970:375. LEMA: minhoca sf.	PT-EU	6.1
minhoca CarvalhoS 1974. Vid. minhoca-da-água . LEMA: minhoca sf.	PT-EU	6.1
minhoca sf Verme anelídeo. Etimología. Micælis derivou de mioca, lat. <i>minare</i> . Sampaio cita como corrente, <i>mi-nhoca</i> ou <i>minhoga</i> , o que é extraído, arrancado ou tirado. Cordeiro da Mata apud Imaguire (1999, p. 365) afirma que o vocábulo é quimbundo <i>mu</i> , prefixo + <i>nhoka</i> , cobra, em que o u átono pretónico se mudou em i. Lino 2000:89. LEMA: minhoca sf.	PT-BR	6.1
minhoca s.f. (Menos frecuente.) MaiaB 1965:144. Vid. bichas da terra . LEMA: minhoca sf.	PT-EU	6.1
minhoca [mi'ɲokɐ] s.f. Órgão genital da creança. Do africano munhoca. A.N. {a palabra ainda não vem registada com este sentido no Dicionário de Cândido Figueiredo}. OliveiraR 1948:119. Vid. gaita . LEMA: minhoca sf.	PT-EU	7.1
minhoca [mi'ɲokɐ] s.f. Var. mioca . SilvaM 1972:318. Vid. mioca . LEMA: minhoca sf.	PT-EU	6.1
minoca s.f. Minhoca. Nunes 1965:138. LEMA: minhoca sf.	PT-EU	6.1
mioca [mj'okɐ] s.f. Minhoca. Sobre a etimologia desta palabra V. A.N. Dic. s.v. Observa-se nesta palabra o desaparecimento da palatal <i>nh</i> (= ñ), portanto, o fenómeno de síncope. Barros 1940. LEMA: minhoca sf.	PT-EU	6.1
mioca [mi'okɐ] s.f. Minhoca. Braga 1971:284. Vid. menhoca . LEMA: minhoca sf.	PT-EU	6.1
mioca s.f. Verme anelídeo, da familia dos lumbricídeos, próprios dos lugares húmidos. (cf. Ericeira, pág. 340). Caldeira 1960:368. LEMA: minhoca sf.	PT-EU	6.1
mioca s.f. Minhoca; verme da terra. Costa 1961:261. LEMA: minhoca sf.	PT-EU	6.1
mioca [mj'okɐ] s.f. Minhoca. SilvaM 1972:318. Vid. minhoca . LEMA: minhoca sf.	PT-EU	6.1
mióca s. f. Minhoca. Alves 1993:222. LEMA: minhoca sf.	PT-EU	6.1
mióca Alves 1993:222. Vid. mióca de mari , mióca de terra . LEMA: minhoca sf.	PT-EU	0.0



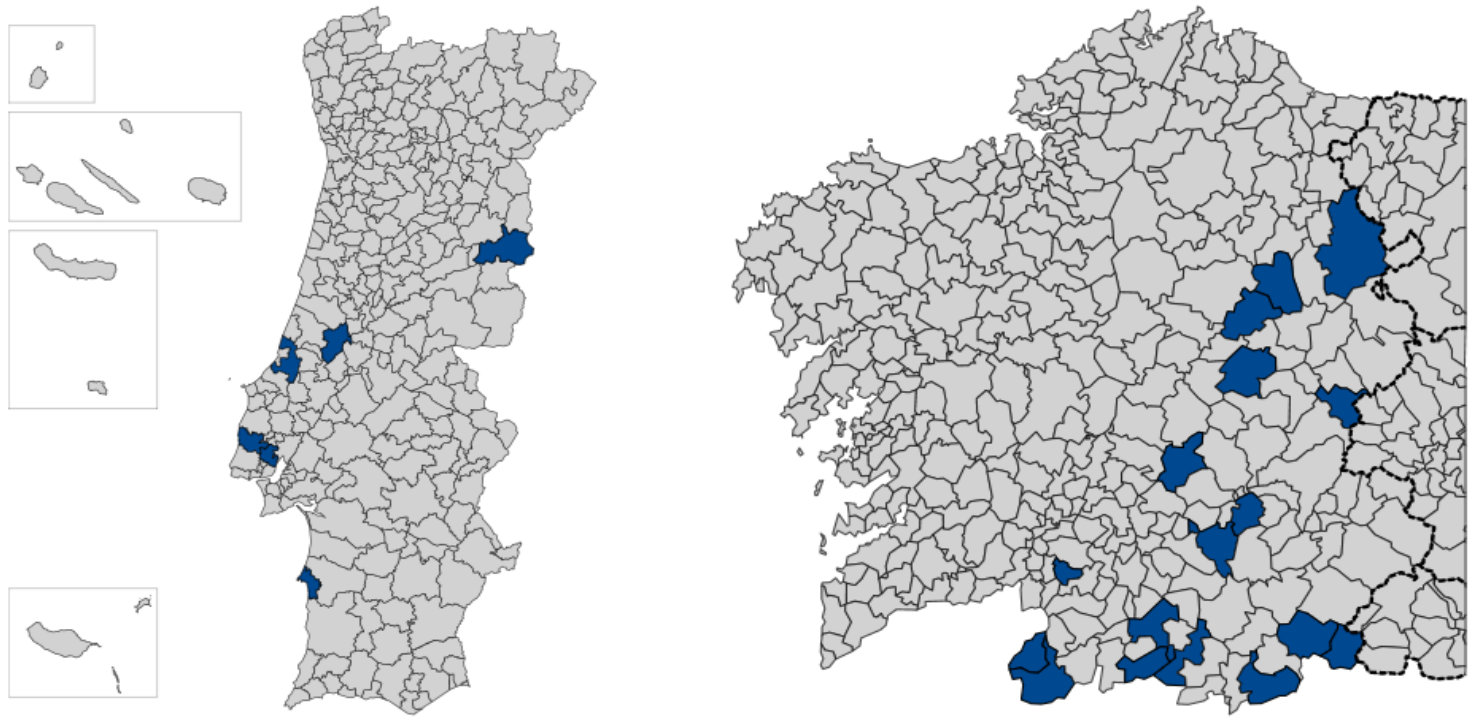
Lema Variante Busca avanzada ▾

Sensible a diacríticos

- Localización (2)**
- Galicia
 - Portugal
- Campo semántico (2)**
- 0.0 - Non clasificable
 - 6.1 - Animais silvestres (denominación)
- Lemas (2)**
- minhoca (pt-eu)
 - miñoca (gl)
- Variantes (3)**
- mioca (pt-eu)
 - mioca (gl)
 - mióca (pt-eu)

Resultados atopados (22)

mioca	Lombriz de terra. Acosta 1970:42. LEMA: miñoca sf.	GL	6.1
mioca	Miñoca, en cast. "lombriz de terra". Álvarez 1974:148. LEMA: miñoca sf.	GL	6.1
mioca [mj'oke]	s.f. Minhoca. Sobre a etimoloxía desta palabra V. A.N. Dic. s.v Observa-se nesta palabra o desaparecemento da palatal <i>nh</i> (= ñ), portanto, o fenómeno de síncope. Barros 1940. LEMA: minhoca sf.	PT-EU	6.1
mioca [mi'oke]	s.f. Minhoca. Braga 1971:284. Vid. menhoca . LEMA: minhoca sf.	PT-EU	6.1
mioca	s.f. Verme anelídeo, da familia dos lumbrídeos, propios dos lugares húmidos. (cf. Ericeira, pág. 340). Caldeira 1960:368. LEMA: minhoca sf.	PT-EU	6.1
mioca mi[ɔ]ca	s. f. Bicho qu' anda por baixo da terra úmeda, e sirbe de cebo pr' os truíteiros. Castro 1986:64. LEMA: miñoca sf.	GL	6.1
mioca	s.f. Minhoca; verme da terra. Costa 1961:261. LEMA: minhoca sf.	PT-EU	6.1
mioca mi[ɔ]ca	Lombriz de terra. Díaz 1971:26. LEMA: miñoca sf.	GL	6.1
mioca	Gusano en general. Gil 1969:14. LEMA: miñoca sf.	GL	6.1
mioca	Lombriz de terra. Gil 1969:14. LEMA: miñoca sf.	GL	6.1



Lema

Pódese apagar un territorio

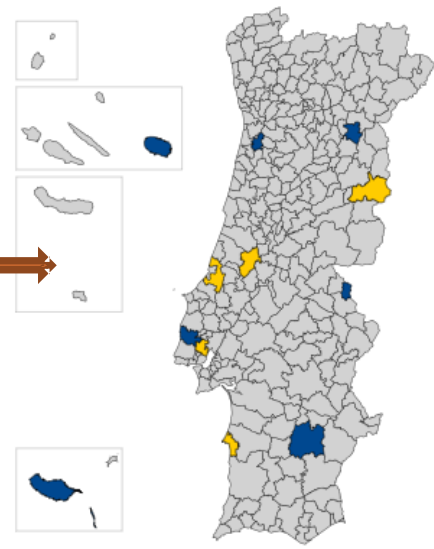
- Localización (2)
- Brasil
 - Portugal
- Campo semântico (3)
- 0.0 - Não classificável
 - 6.1 - Animais bravios (denominações)
 - 7.1 - Ser humano (aspectos físicos, psíquicos e comportamentais)
- Lemas (2)
- minhoca (pt-eu)
 - minhoca (pt-br)
- Variantes (6)
- minhoca (pt-br)
 - minhoca (pt-eu)
 - minoca (pt-eu)
 - menhoca (pt-eu)
 - mioca (pt-eu)
 - mióca (pt-eu)

Resultados encontrados (17)

minhoca [mɨ'ɲɔkɐ] s.f. Minhoca. Delgado 1970:344. LEMA: minhoca sf.	PT-EU 6.1
menhoca [mɨ'ɲɔkɐ] s.f. Minhoca. Delgado 1970:375. LEMA: minhoca sf.	PT-EU 6.1
minhoca CarvalhoS 1974. Vid. minhoca-da-água . LEMA: minhoca sf.	PT-EU 6.1
minhoca s.f. (Menos frecuente.) MaiaB 1965:144. Vid. bichas da terra . LEMA: minhoca sf.	PT-EU 6.1
minhoca [mi'noke] s.f. Órgão genital da creança. Do africano munhoca. A.N. (a palabra ainda não vem registada com este sentido no Dicionário de Cândido Figueiredo). OliveiraR 1948:119. Vid. gaita . LEMA: minhoca sf.	PT-EU 7.1
minhoca [mɨ'ɲɔkɐ] s.f. Var. mioca . SilvaM 1972:318. Vid. mioca . LEMA: minhoca sf.	PT-EU 6.1
minoca s.f. Minhoca. Nunes 1965:138. LEMA: minhoca sf.	PT-EU 6.1
mioca [mɨ'ɲɔkɐ] s.f. Minhoca. Sôbre a etimologia desta palabra V. A.N. Dic. s.v. Observa-se nesta palabra o desaparecimento da palatal nh (= ñ), portanto, o fenómeno de síncope. Barros 1940. LEMA: minhoca sf.	PT-EU 6.1
mioca [mi'ɲɔkɐ] s.f. Minhoca. Braga 1971:284. Vid. menhoca . LEMA: minhoca sf.	PT-EU 6.1
mioca s.f. Verme anelídeo, da família dos lumbricídeos, próprios dos lugares húmidos. (cf. Ericeira, pág. 340). Caldeira 1960:368. LEMA: minhoca sf.	PT-EU 6.1
mioca s.f. Minhoca; verme da terra. Costa 1961:261. LEMA: minhoca sf.	PT-EU 6.1
mioca [mɨ'ɲɔkɐ] s.f. Minhoca. SilvaM 1972:318. Vid. minhoca . LEMA: minhoca sf.	PT-EU 6.1
mióca s. f. Minhoca. Alves 1993:222. LEMA: minhoca sf.	PT-EU 6.1
mióca Alves 1993:222. Vid. mióca de mari , mióca de terra . LEMA: minhoca sf.	PT-EU 0.0



Selecciona unha variante: ilumina atestacións no corpus / localizacións no mapa



Metodología: criterios de clasificación gramatical

- ▶ Os criterios son comúns
 - Toma de decisión convencional sobre determinados aspectos problemáticos. Exemplos
 - *carpinteiro* **s**
 - *idiota* **a, s**
- ▶ Relación de abreviaturas convencionais unificada

Metodoloxía: criterios de clasificación semántica

- ▶ Manual a disposición, on-line. Blog de traballo
 - Criterios, procedementos, exemplos...
 - Versión galega e portuguesa
- ▶ Uso de claves numéricas, de acordo cunha táboa convencional
- ▶ Tres niveis de clasificación, de especificación sucesiva

6. ANIMAIS NÃO-DOMÉSTICOS

6.1. Denominações de animais: mamíferos, peixes, aves, anfíbios, insectos...

- ✓ Denominações genéricas: *animal, bicho, pássaro, peixe...*
- ✓ Denominações de animais: *corço, furão, lebre, lobo, raposa, morcego...; borboleta, vespa; aranha; águia, andorinha, cotovia, corvo, mocho; badejo, barbo, peixe-galo, truta; caracol, caranguejo, lesma; lagarto, rã, víbora...*
- ✓ Variedades e crias: *lobato, perdigão...*

6.2. Léxico relacionado com os animais não-domésticos (partes, aspecto, tocas, costumes...)

- ✓ Partes do corpo: *bico, concha, ferrão, garra...*
- ✓ A habitação: *cama, covil, lura, ninho, toca...*
- ✓ Acções: *aninhar, chocar, cantar, nadar, saltar, voar...*
- ✓ Propriedades e características (adjectivos e substantivos): *bravo, toural...*

6.3. A caça. Armadilhas, armas e léxico relacionado

- ✓ Armadilhas: *armadilha, esparrela, laço, lousa, ratoeira...*
- ✓ Armamento e munição: *escopeta, espingarda,...*
- ✓ Equipamento: *cartucheira,...*
- ✓ Acções: *apontar, caçar, disparar...*
- ✓ Outros: *montaria...*

6.0. Outro léxico dos animais não-domésticos de difícil classificação nas secções anteriores

Ilumina atestacións no corpus

- Galicia
- Campo semántico (4)**
- 0.0 - Non clasificable
- 6.1 - Animais silvestres (denominacións)
- 7.1 - Ser humano (aspectos físicos, psicolóxicos e condutuais)
- 8.6.1 - Artes de pesca
- Lemas (1)**
- miñoica (gl)
- Variantes (9)**
- miñoica (gl)
- miñoica, -ña (gl)
- miñoica de río (gl)
- miñoica da pedra (gl)
- cebo de miñoica da barra (gl)
- mioca (gl)
- miñoicas (gl)

Resultados atopados (51)

miñoica s. f. Nombre que reciben varios anélidos policetos del género *Nereis*. {Cangas} GonzálezGM 2002:544. LEMA: **miñoica** sf.

miñoica **miñ**[ˈɔ]ca Caralla. López 1969:60. LEMA: **miñoica** sf.

miñoica **miñ**[ˈɔ]ca Verme anélido, longo e de corpo mol e dividido en argolas de carne quitinosa, que viven en terra húmeda López **miñoica** sf.

miñoica **miñ**[ˈɔ]ca López 1969:44. Vid. **andar á miñoica**. LEMA: **miñoica** sf.

... ser humano na súa dimensión individual: aspectos físicos, psicolóxicos e condutuais. O corpo: anatomía e fisioloxía. Características físicas. Enfermidades, eivas e tratamento. As características psicolóxicas e a conduta. A actividade física, movementos e accións (comprende as accións non incluíbles noutros apartados, aínda que non se especifique que son realizadas por humanos).

... que vive nos terreos húmidos. Louredo 2012. LEMA: **miñoica** sf.

Monteagudo 1998. LEMA: **miñoica** sf.

... ca LEMA: **miñoica** sf.

Apaga clasificadores: só 6.1 [animais não-domésticos (denominações)]

TESOURO DO LÉXICO PATRIMONIAL GALEGO E PORTUGUÉS

Lema Variante Busca avanzada Sensible a diacríticos

Localización (1)

- Galicia

Campo semántico (4)

- 0.0 - Non clasificable
- 6.1 - Animais silvestres (denominacións)
- 7.1 - Ser humano (aspectos físicos, psicolóxicos e condutuais)
- 8.6.1 - Artes de pesca

Lemas (1)

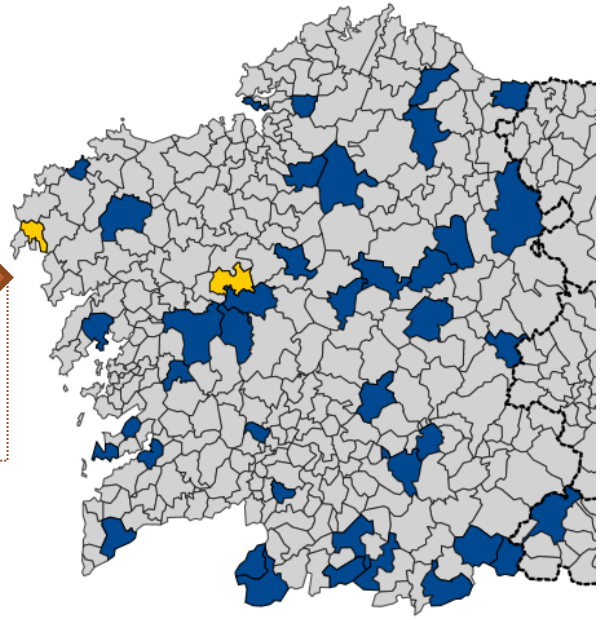
- miñoica (gl)

Variantes (9)

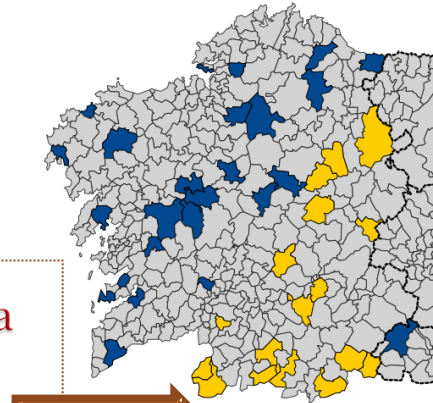
- miñoica (gl)
- miñoica, -ña (gl)
- miñoica de río (gl)
- miñoica da pedra (gl)
- cebo de miñoica da barra (gl)
- mioca (gl)
- miñoicas (gl)

Desprega contido do campo

Ilumina localizacións no mapa



Ilumina no mapa localizacións dunha variante



Busca por lema MIÑOICA

Busca avanzada

- ▶ Selección:
 - Galicia / Portugal / Brasil
 - Clasificadores semánticos
 - Categoría gramatical
- ▶ Obtense
 - Listaxe por lemas, coas variantes de cada un e localización no mapa. Acceso ao banco de datos.
 - Listaxe por variantes, coa información completa en cada atestación. Tamén o lema. Localización no mapa e acceso ao banco de datos.

TESOURO DO LÉXICO PATRIMONIAL GALEGO E PORTUGUÉS

Lema Variante

Categorías... Busca simple ▲

Resultados atopados (1177)

abegoiro <i>sm</i> Variantes Portugal: abigoilo .	PT-EU
abelha brava <i>loc sf</i> Variantes Portugal: avelha brava .	PT-EU
abelhão <i>sm</i> Variantes Portugal: abelhão , abilhão , avelhão .	PT-EU
abesardo <i>sm</i> Variantes Portugal: abesardo .	PT-EU
abetarda <i>sf</i> Variantes Portugal: abatarda , batarda .	PT-EU
abibe <i>sm</i> Variantes Portugal: abibe .	PT-EU
abigoeiro <i>sm</i> Variantes Portugal: abigoeiro .	PT-EU
aboa <i>sf</i> Variantes Portugal: aboa , aboinha .	PT-EU
aboeirinha <i>sf</i> Variantes Portugal: aboeirinha .	PT-EU
aboiinha <i>sf</i> Variantes Portugal: aboiinha .	PT-EU
abrecu <i>sm</i> Variantes Portugal: abre-cu .	PT-EU
abre-cu <i>sm</i> Variantes Portugal: abre-cús .	PT-EU
abrótea <i>sf</i> Variantes Portugal: abroita , abróita , abróta , abrótea .	PT-EU
abrótega <i>sf</i> Variantes Portugal: abrótica , abrótiqa .	PT-EU
abrótega pequena <i>loc sf</i> Variantes Portugal: abrótica pequenina .	PT-EU

Estado actual

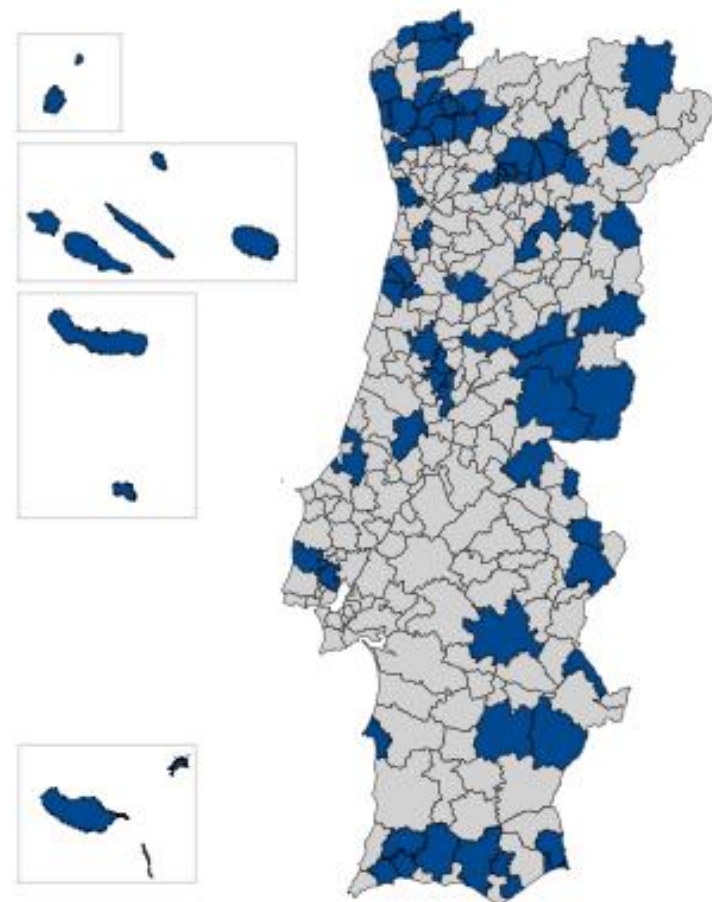
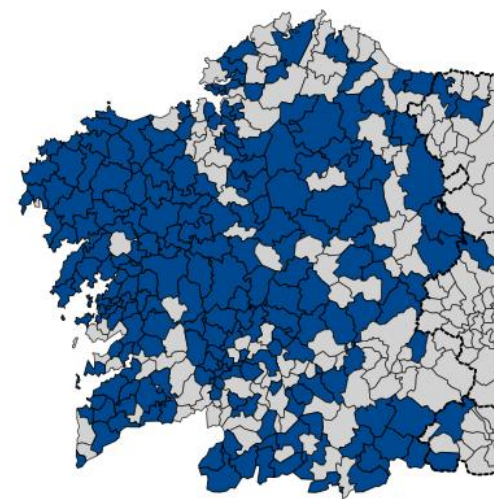
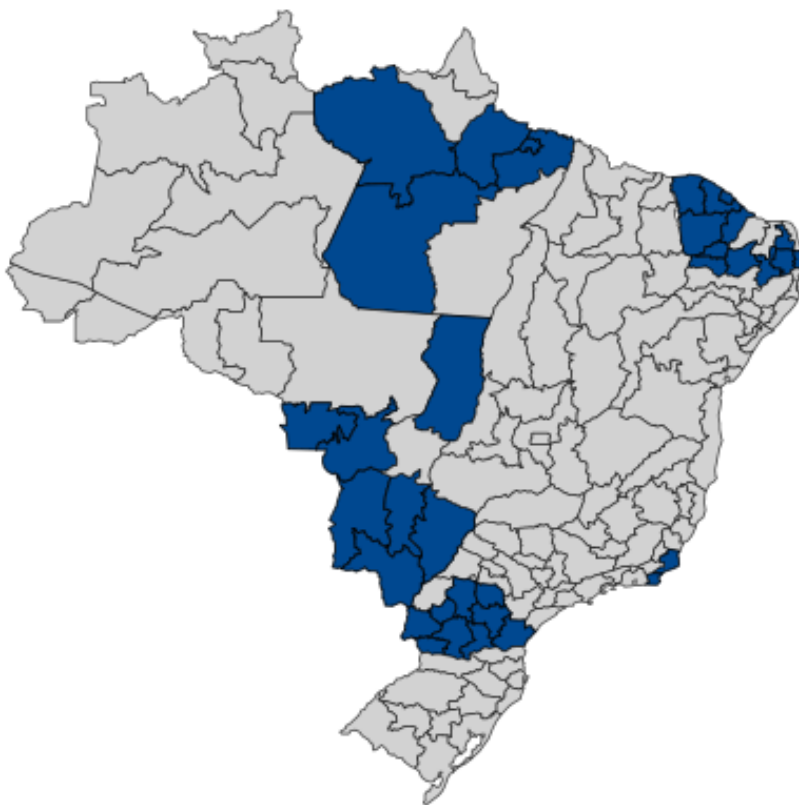
Datos xerais

Obras introducidas: 144 (Brasil 26, Galicia 62, Portugal 55)

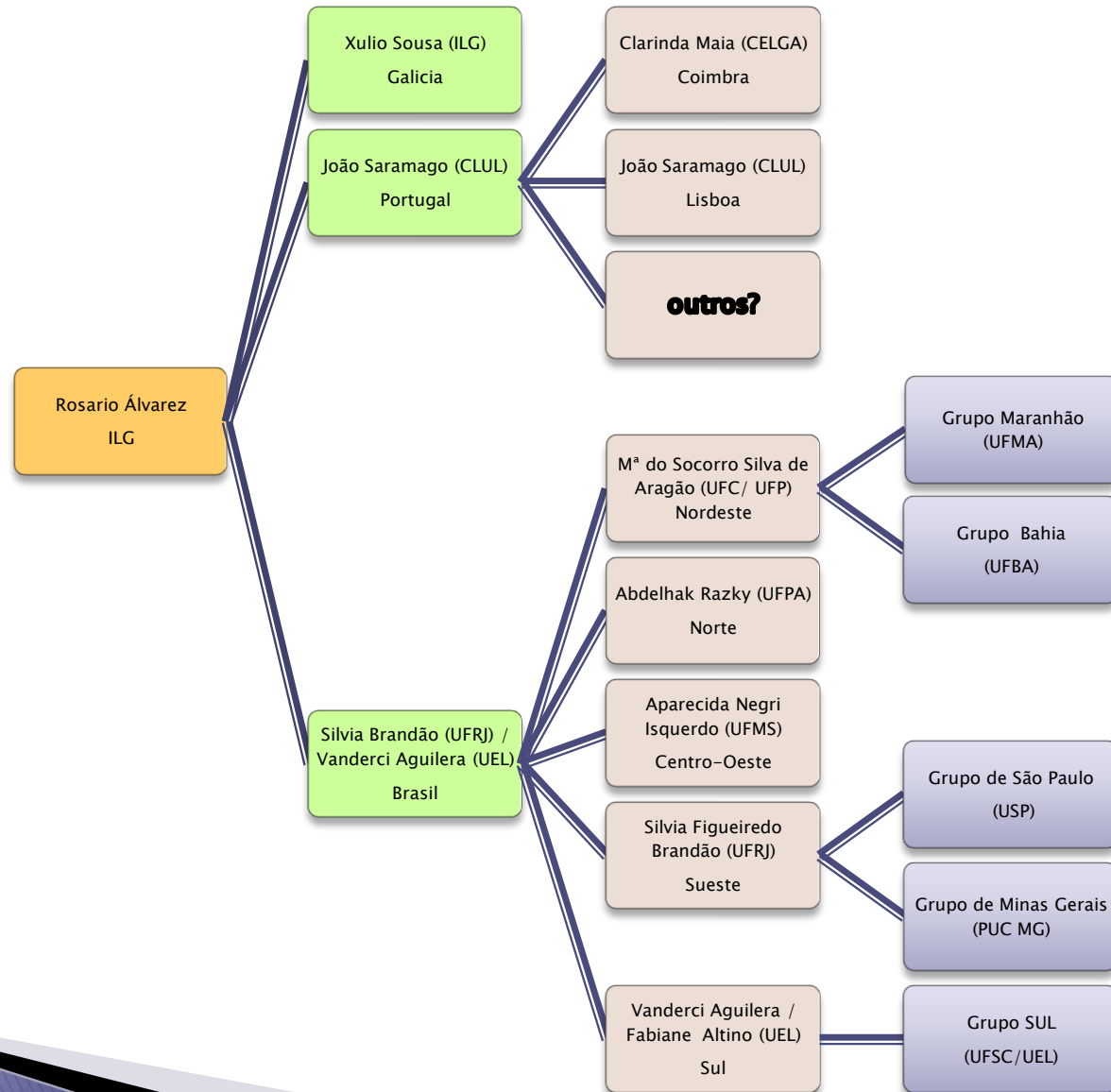
Número total de rexistros: 198849

Número de lemas: 52721

Número de variantes: 85484



Equipo



Seguimos a traballar...

- ▶ O equipo funciona en rede: estamos abertos a colaboracións de diverso tipo e á integración de persoas e equipos
- ▶ Seguimos tratando obras e aumentando o corpus
- ▶ Temos novos retos:
 - Vinculación dos lemarios galego e portugués
 - Refinar a clasificación e tratamento semánticos
 - Vinculación de xeosinónimos: xeosinónimo de referencia
 - Maior automatización no tratamento dos datos

- ▶ Accesible en <http://ilg.usc.gal/Tesouro/>
- ▶ Ao voso dispor en rosario.alvarez@usc.es

Instituto da Lingua Galega – Universidade de Santiago de Compostela
Praza da Universidade 4
15782 Santiago de Compostela

Graciñas
Obrigadísima