

DSpace da Universidade de Santiago de Compostela

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

Instituto da Lingua Galega

María Álvarez de la Granja (2015): “A edición dos materiais no léxico dispoñible do galego”.
Gallæcia – III Congreso Internacional de Linguística Histórica. Universidade de Santiago de Compostela, 27-30 xullo 2015.



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.



GALLÆCIA - III CONGRESSO INTERNACIONAL DE LINGUÍSTICA HISTÓRICA
Homenagem aos Professores Ramón Lorenzo e Antón Santamarina

A edición dos materiais no léxico dispoñible do galego



María Álvarez de la Granja

Instituto da Lingua Galega - Universidade de Santiago de Compostela

maria.alvarez.delagranja@usc.es

01. o corpo humano

| | |
|---------------------|--------------|
| 1. pés | 25. cerebro |
| 2. braças | 26. neuronas |
| 3. cabeça | 27. sangue |
| 4. mãos | 28. boca |
| 5. coração | 29. olhos |
| 6. pulmón | 30. tímpano |
| 7. tronco | 31. oído |
| 8. estómago | 32. orelha |
| 9. pâncreas | 33. pelo |
| 10. vesícula | 34. pele |
| 11. ovários | 35. suor |
| 12. vagina | 36. uréteres |
| 13. escroto | 37. |
| 14. veias | 38. |
| 15. artérias | 39. |
| 16. válvulas | 40. |
| 17. esterno | 41. |
| 18. clavícula | 42. |
| 19. omóplato | 43. |
| 20. fémur | 44. |
| 21. fígado | 45. |
| 22. vesícula biliar | 46. |
| 23. intestino | 47. |
| 24. ano | 48. |

Tratamento das respostas obtidas

**026 01 01 pés brazos cabeza
mans corazón pulmón tronco
estómago páncreas vexiga ovarios
vaxina escroto veas arterias válvulas
esternón clavícula omóplato fémur
fígado vesícula biliar intestino ano
cerebro neuronas sangue boca ollos
tímpano oído orella pelo pel súor
unllas**

Dúas tarefas fundamentais no proceso de edición

- **Supresión dalgunhas voces**
- **Lematización de todas as respostas conservadas**

Eliminación de respostas

- **De maneira xeral, conservamos as palabras e expresións ofrecidas polos informantes, pero en todos os centros de interese levamos a cabo a eliminación dunhas poucas respostas.**

Criterios para a eliminación de respostas

- Formas repetidas na mesma listaxe.
- Formas que están moi lonxe ou fóra do ámbito temático solicitado ou que teñen un carácter demasiado xenérico dentro deste:
 - fotosíntese* no CI 'Os animais'
 - mirar* en 'Os traballos do campo e do xardín'
- Formas que constitúen series das que botan man os informantes para ir cubrindo as celas do cuestionario con pouco esforzo:
 - Fifa 00, Fifa 01, Fifa 02, Fifa 03, Fifa 04* no CI 'Xogos e distraccións'

Criterios para a eliminación de respostas

- Formas nas que é palpable o desexo do informante de transgredir a tarefa encomendada:

algúns só decoran —despois de profesor— no CI ‘A escola: mobles e materiais’

mar-celo, mar-mota, mar-tín no CI ‘O mar’

- Formas que conteñen predicacións:

calma os problemas de cada un no CI ‘O mar’

A lematización

Co obxectivo de

- 1** Unificar variantes (sobre todo, aínda que non só, fónicas, morfolóxicas, ortográficas e flexivas).
- 2** Ofrecer, na medida do posible, formas normalizadas que faciliten a busca do usuario en Dispogal

vinculamos a un *lema* (noutros léxicos dispoñibles, *vocábulo*) todas e cada unha das palabras conservadas nas listaxes.

Algúns criterios de lematización

- As formas en plural lematízanse en singular, agás que estean lexicalizadas:

lobos → *lobo*

- Os adxectivos en feminino lematízanse en masculino:

bonita → *bonito*

- As formas con erros ortográficos lematízanse coa ortografía correcta

avella → *abella*

- As formas con abreviaturas lematízanse coa forma plena:

sist nervioso → *sistema nervioso*

Agrupación de respuestas

- Deste modo agrupamos respuestas:

CI 'O campo'

lobo 47

lobo 19

lobos 27

lovos 1

Agrupación de respostas

Tamén agrupamos

- **Masculinos e femininos de seres sexuados baixo unha forma que representa ambas as posibilidades:**
 - médico* e *médica* → *médico-a*
 - polbeira* e *polbeiro* → *polbeira-o*
- **Sintagmas complexos e as formas que resultan da abreviación destes baixo unha forma que representa ambas as posibilidades:**
 - vaqueiros* e *pantalóns vaqueiros* → *(pantalóns) vaqueiros*
- **Acurtamentos e formas plenas baixo unha forma que representa ambas as posibilidades:**
 - tele* e *televisión* baixo *tele(visión)*

Agrupación de respostas

- Deste modo temos, por exemplo, no CI 'A roupa':

(pantalóns) vaqueiros 168

| | |
|---------------------|----|
| baqueiro | 4 |
| baqueiros | 9 |
| baquero | 1 |
| baqueros | 2 |
| p vaqueiro | 1 |
| pantalón baqueiro | 1 |
| pantalón vaqueiro | 2 |
| pantalón vaquero | 4 |
| pantalón vaquero | 1 |
| pantalón vaquero | 7 |
| pantalón vaqueros | 1 |
| pantalóns vaqueiro | 1 |
| pantalóns vaqueiros | 7 |
| pantalóns vaqueros | 1 |
| vaqueiro | 29 |
| vaqueiros | 64 |
| vaquero | 11 |
| vaqueros | 22 |

Dispogal

| | | | | | | |
|----|------------------------------|---------|---------|-----|--------|---------|
| 19 | suxeitador | 0,18774 | 33,625% | 269 | 1,610% | 48,782% |
| 20 | cazadora | 0,18562 | 29,750% | 238 | 1,425% | 50,207% |
| 21 | vestido | 0,18513 | 30,875% | 247 | 1,478% | 51,685% |
| 22 | saia | 0,17675 | 22,750% | 182 | 1,089% | 52,774% |
| 23 | calzoncillos | 0,15297 | 24,625% | 197 | 1,179% | 53,953% |
| 24 | (pantalóns) vaqueiros | 0,15158 | 21,000% | 168 | 1,006% | 54,959% |
| 25 | guante | 0,15067 | 28,125% | 225 | 1,347% | 56,306% |
| 26 | chándal | 0,14030 | 24,125% | 193 | 1,155% | 57,461% |

Apéndice 1. Listaxe de lemas e variantes por CI

| | |
|-----------------------------------|------------|
| (pantalóns) piratas | 67 |
| pantalon pirata..... | 2 |
| pantalón pirata..... | 12 |
| pantalones piratas..... | 3 |
| pantalóns pirata | 1 |
| pantalons piratas..... | 3 |
| pantalóns piratas..... | 4 |
| pantalos pirata..... | 1 |
| pirata..... | 12 |
| piratas | 28 |
| pirátas | 1 |
| (pantalóns) vaqueiros..... | 168 |
| baqueiro | 4 |
| baqueiros..... | 9 |
| baquero | 1 |
| baqueros..... | 2 |
| p vaqueiro | 1 |
| pantalón baqueiro..... | 1 |
| pantalon vaqueiro..... | 2 |
| pantalón vaqueiro..... | 4 |
| pantalon vaquero | 1 |
| pantalón vaquero | 7 |
| pantalón vaqueros | 1 |
| pantalóns vaqueiro | 1 |
| pantalóns vaqueiros..... | 7 |
| pantalóns vaqueros..... | 1 |
| vaqueiro | 29 |
| vaqueiros..... | 64 |

Tratamento das respostas obtidas

026 01 01 pés/pé brazos/brazo
cabeza mans/man corazón pulmón
tronco estómago páncreas vexiga
ovarios/ovario vaxina escroto veas/vea
arterias/arteria válvulas/válvula
esternón/esterno clavícula
omóplato/omoplata fémur fígado
vesícula biliar intestino ano cerebro
neuronas/neurona sangue boca
ollos/ollo tímpano oído orella pelo pel
súor unllas/uña

Tratamento das respostas obtidas

026 01 01 pés/pé brazos/brazo
cabeza mans/man corazón pulmón
tronco estómago páncreas vexiga
ovarios/ovario vaxina escroto veas/vea
arterias/arteria válvulas/válvula
esternón/esterno clavícula
omóplato/omoplata fémur fígado
vesícula biliar intestino ano cerebro
neuronas/neurona sangue boca
ollos/ollo tímpano oído orella pelo pel
súor unllas/uña

O tratamento das variantes galegas

CI 'O campo'



legón
ligón
logón

ensilar
silar



Lematización conxunta das variantes fónicas e morfolóxicas de orixe galega

O lema será a forma estándar do galego co mesmo significado, designación e raíz.



legón
legón
ligón
logón

ensilar
ensilar
silar



Lematización separada dos sinónimos galegos

Teñen lemas diferentes:

- *parrulo e pato*
- *oculista e oftalmólogo*
- *xabaril e porco bravo*

porco bravo

porco bravo
porcos bravos



xabaril

| | |
|----------|-----------|
| ghabalís | xabaril |
| jabali | xabariles |
| jabalí | xabarin |
| jabalies | xabarín |
| jabalíes | xabarins |
| jabalís | xabaríns |
| xabalí | xabaris |
| xabalín | xavaril |
| xabalins | xavaríns |
| xabalis | xavarís |
| xabalís | |

O tratamento das voces do castelán



maceira
maceiras
maceiro
maciñeira
maciñeiras
macira
manzano
mazeiros
mazaira
mazanceiros
mazeiro

aboelos
abós
abuelo
abuelos
avó
avoa
avós
avo



O tratamento das vozes do castelán



morcéago
morcego
morcélago
murcélago
murciélago



iglesia
igresas
igreja
igrexas
igrexia

O tratamento das voces do castelán

maceira

maceira
maceiras
maceiro
maciñeira
maciñeiras
macira
manzano
mazeiros
mazaira
mazanceiros
mazeiro



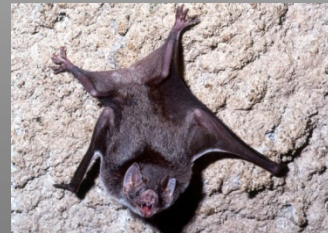
avó,-a

aboelos
abós
abuelo
abuelos
avó
avoa
avós
avo



morcego

morcéago
morcego
morcélago
murcélago
murciélago



igrexia

iglesia
igresas
igrexia
igrexas
igrexia



O tratamento das voces do castelán

Teñen lemas diferentes:

- *silla e cadeira*
- *perro e can*
- *cerdo e porco*

cerdo

cerdo

cerdos

porco

porco

porcos





GALLÆCIA - III CONGRESSO INTERNACIONAL DE LINGUÍSTICA HISTÓRICA
Homenagem aos Professores Ramón Lorenzo e Antón Santamarina

Muito obrigada!
Moitas grazas!



GALLÆCIA III CILH

GALLÆCIA - III CONGRESSO INTERNACIONAL DE LINGÜÍSTICA HISTÓRICA
Homenagem aos Professores Ramón Lorenzo e Antón Santamarina

Muito obrigada!
Moitas grazas!



María Álvarez de la Granja

Instituto da Lingua Galega- Universidade de Santiago de Compostela

maria.alvarez.delagranja@usc.es