

DÚAS FERRAMENTAS PARA O ESTUDO DA VARIACIÓN EN GALEGO: DICIONARIO DE FRECUENCIAS E FLEXIONADOR

EVA DOMÍNGUEZ (EDOMIN@CIRP.GAL / EVAMARIA.DOMINGUEZ@USC.GAL)

VÍTOR MÍGUEZ (VITOR.MIGUEZ@USC.GAL)

NESTA EXPOSICIÓN IMOS VER...

Diccionario de frecuencias

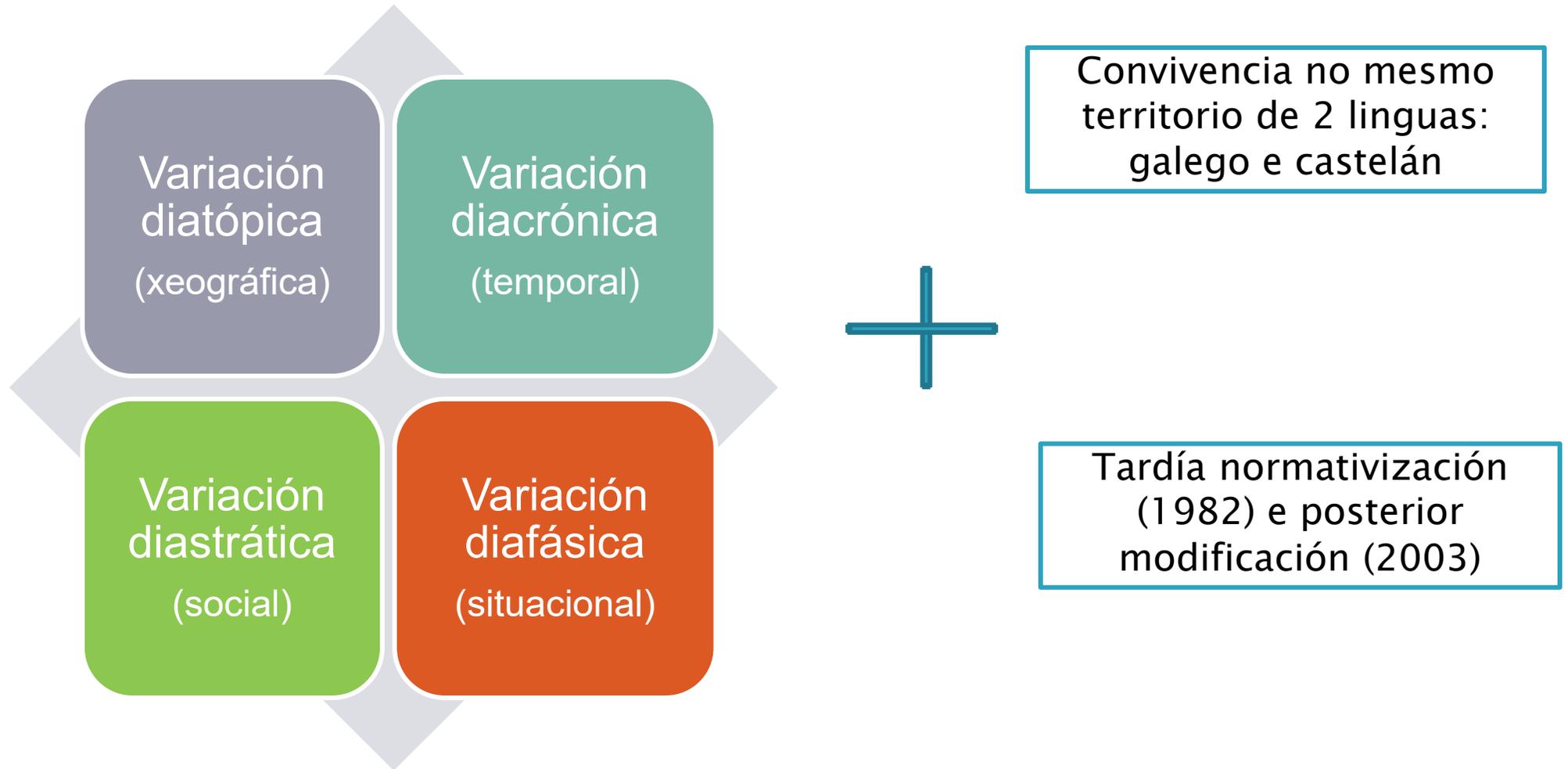
- Un precedente
- Localización da ferramenta e descrición xeral
- Estudo de caso: denominación pagá e cristiá dos días da semana

Flexionador

- Breve panorámica sobre conxugadores e flexionadores para o galego
- Por que un máis?
- O flexionador de XIADA: principais características
- Demo do seu funcionamento
- Algúns desaxustes e opcións de mellora

Conclusións

O CALDO DE CULTIVO DA VARIACIÓN E O CAMBIO LINGÜÍSTICO



CALADOIROS: CORPUS DE GALEGO CONTEMPORÁNEO

Corpus	Acceso	Período	Tamaño	Rexistro
AGO	http://ilg.usc.es/ago	1974-2010	2000 h de gravacións ?? palabras	Oral (1000 h etnotextos + 100 h relatorios, presentacións de libros, programas de radio etc.) 
CORGA	http://corpus.cirp.gal/corga	1975-2021	45.665.649 palabras	Escrito (Oral 1,5%: programas de radio, conferencia) 
CORILGA	http://ilg.usc.es/corilga	1965-2020	?? palabras	Oral 
CORTEGAL	https://ilg.usc.gal/cortegal	2017	1000 textos (sobre 200-250 palabras cada un)	Escrito (textos argumentativos estudantes ABAU)
CTAG	https://ilg.usc.gal/ctg	[2000-2009]	17.984.638 palabras	Escrito (textos especializados de 6 ámbitos temáticos)
TILG	https://ilg.usc.es/TILG	1612-2013	30.924.082 palabras	Escrito (Oral 4,7%: poesía trad. oral, conferencia etc.)
PARLAMINT-ES-GA (noSketchEngine)	https://www.clarin.si/ske/#dashboard?corpname=parlamint41_es_ga	2015-2022	17.511.047 palabras	Oral (debates parlamentarios)

Seguindo a derrota do CORGA, contan tamén con dicionario de frecuencias dinámico:

- [ESLORA – Corpus para el estudio del español oral](#)
- [CAES – Corpus de aprendices de español](#)

CORGA 4.1 <[HTTP://CORPUS.CIRP.GAL/CORGA](http://corpus.cirp.gal/corga)>

❑ **Características xerais:** corpus documental aberto, de acceso libre, de balde, codificado en XML, abrangue cronoloxicamente dende 1975-, representativo do galego actual (=> modelos de lingua), etiquetado automaticamente.

❑ **Composición** (45.660.649 palabras / 54.737.277 elementos gramaticais):

758 xornais

348 revistas

220 novelas

196 ensaios

144 coleccións de relato curto

162 obras de teatro

20 guións

203 entradas de blog

69 horas de audio transcritas aliñadas (689.000 palabras)

❑ **Ferramenta de recuperación** (consultas por contigüidade ou por proximidade)

- Obtención dos datos: procuras por formas ortográficas e/ou información gramatical (elementos gramaticais, etiquetas, lemas e hiperlemas).
- Filtros: orixe, bloque, xénero, tipo de documento, data, área temática, medio, obra, autor, interlocutor, sexo do autor, sexo do intelocutor, titulares, encabezamentos, alongamentos, solapamentos etc.
- Visualización dos resultados: concordancias, frecuencias, coaparicións e **inventario**.
- **Diccionario de frecuencias estático**, que ten en conta o índice de dispersión, de maneira que a información que proporciona é moito máis relevante que a simple frecuencia absoluta ou relativa do global.

O DICCIONARIO DE FRECUENCIAS

VOCABULARIOS BÁSICOS PARA ESTUDANTES

Routledge
Taylor & Francis Group

Frequency dictionary

3171 Product Results

Frequency dictionary

Showing: Keywords: Frequency dictio... Clear All

Relevance 12 per page

PERSIAN
A Frequency Dictionary of Persian: Core Vocabulary for Learners
Corey Miller, Karineh Aghajanian-St...

GERMAN
A Frequency Dictionary of German: Core Vocabulary for Learners
Erwin Tischner, Jupp Mörning

FRENCH
A Frequency Dictionary of French: Core Vocabulary for Learners
Deryle Lonsdale, Yvon Le Bras

ARABIC
A Frequency Dictionary of Arabic: Core Vocabulary for Learners
Tim Buckwalter, Dikwongh Parkinson

CZECH
A Frequency Dictionary of Czech: Core Vocabulary for Learners
František Čermák, Michal Kriváček

KOREAN
A Frequency Dictionary of Korean: Core Vocabulary for Learners
Sun-Hae Lee, Seok-Bae Jang, Sang-K...

TURKISH
A Frequency Dictionary of Turkish: Core Vocabulary for Learners
Yegim Aksan, Mustafa Aksan, Ümit...

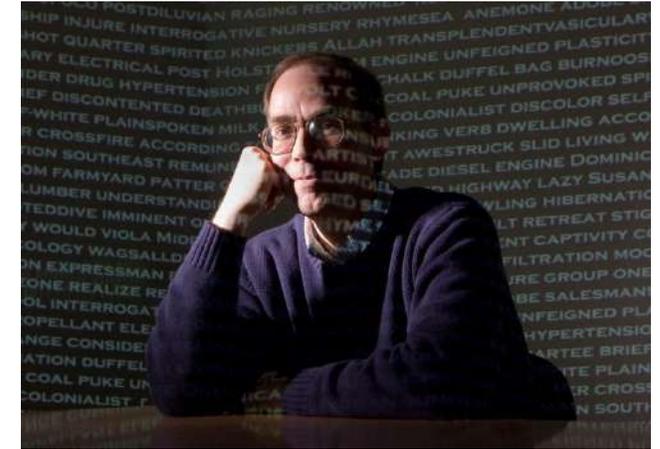
JAPANESE
A Frequency Dictionary of Japanese: Core Vocabulary for Learners
Yukio Tono, Makoto Yamazaki, Kikui...

DUTCH
A Frequency Dictionary of Dutch: Core Vocabulary for Learners
Fransje van Marrewijk, J. van Marrewijk

SPANISH
A Frequency Dictionary of Spanish: Core Vocabulary for Learners
Mark Davies

RUSSIAN
A Frequency Dictionary of Russian: Core Vocabulary for Learners
Mark Davies

MANDARIN CHINESE
A Frequency Dictionary of Mandarin Chinese: Core Vocabulary for Learners
Mark Davies



Mark Davies

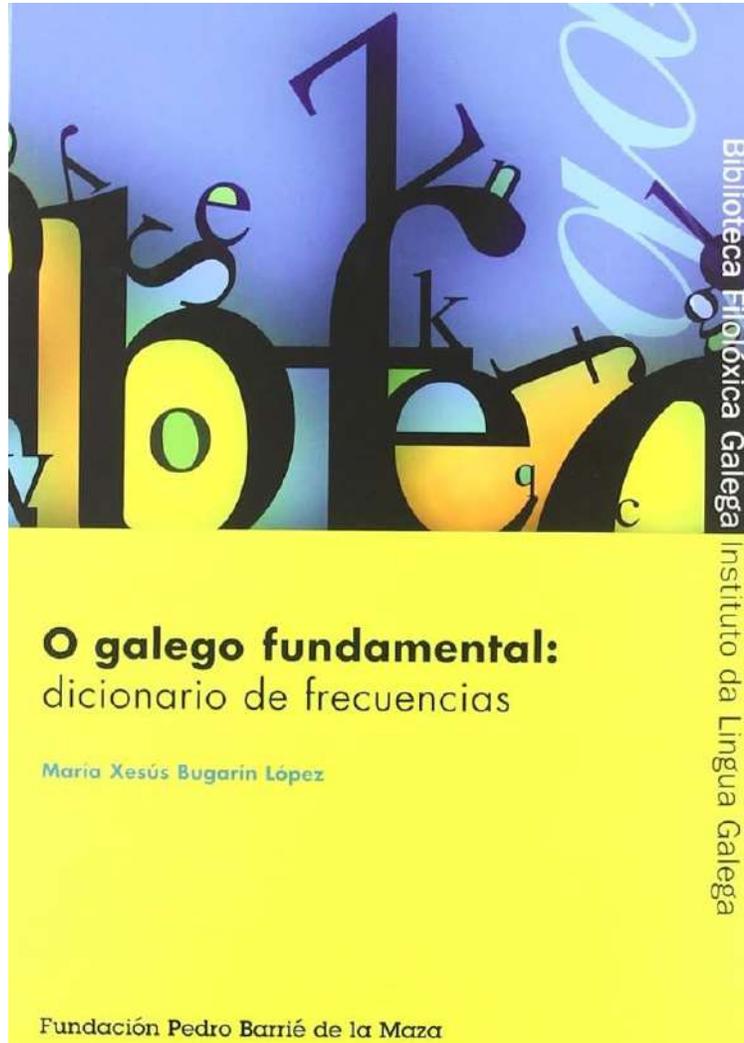
Davies, M. / Hayward Davies, K. (2018): *A Frequency Dictionary of Spanish: Core Vocabulary for Learners. Second edition: revised and expanded.* Routledge

Davies, M. / Raposo Preto-Bay, M. (2007): *A Frequency Dictionary of Portuguese: Core Vocabulary for Learners.* Routledge.

Davies, M. (2006): *A Frequency Dictionary of Spanish: Core Vocabulary for Learners.* Routledge



O PRECEDENTE GALEGO



2004 - Tese de doutoramento

2007 - Publicación

- **Frecuencias dos elementos léxicos a partir dun corpus de 2.064.847 palabras tiradas de textos e gravacións orais datados entre 1975 e 1999.**
 - Listaxe de lemas (32.529) ordenados por frecuencia decrecente, coa indicación da porcentaxe de aparición. Na publicación uns 17.200 (cálculo aproximado) cunha frecuencia de 3 ou superior.
 - Listaxe de lemas ordenados alfabeticamente, coas formas morfolóxicas que lle pertencen (87.273 distintas), acompañadas da clase de palabra e mais da frecuencia absoluta.

Orde	Lema	Ocorrencias	Frecuencia
1	o ¹	129.771	6,28
2	e	77.624	3,76
3	de	77.220	3,74
4	do	63.735	3,09
5	que ¹	45.200	2,19
6	ser ²	40.150	1,94
7	un	38.985	1,89
8	que ²	36.533	1,77
9	no	32.311	1,56
10	se	30.488	1,48
11	non	30.026	1,45
12	ó	26.402	1,28
13	a ²	24.503	1,19
14	en	19.339	0,94
15	ter	17.199	0,83

- Hai 10.117 lemas cunha frecuencia de 1 (hápax): 31,1%.
- Os 53 primeiros lemas representan un 49,98%.
- Os 1000 lemas máis frecuentes cobren o 80,32% do corpus.

Lema	Categoría	Forma	Ocorrencias
CARA			
	loc.adv.	cara, de	6
	loc.prep.	cara a	512
	loc.prep.	cara a, de	31
	loc.prep.	cara de	1
	loc.prep.	cara para	9
	prep.	cara	304
	s.f.	cara	415
	s.f.	caras	52
	s.f.	cariña	15
			1345

PEXAS DOS DICIONARIOS DE FRECUENCIAS IMPRESOS

- ❑ O formato impreso impón unha extensión limitada, que bate cos principios expostos por Zipf.
- ❑ Limitacións no rexistro (escrito, oral ou ambos)
- ❑ Limitacións na clasificación temática (xeral, deportes, economía, medicina...)
- ❑ Limitacións na clasificación tipolóxica (prensa, novela, ensaio, conferencia...)
- ❑ Limitacións no rango temporal
- [...]
- ❑ Pouco ecolóxico



George K. Zipf

Un pequeno número de palabras son utilizadas con moita frecuencia, mentres que un gran número de palabras son pouco empregadas

Busca

Corpus

Etiquetado automaticamente

Tipo

Palab. ortográficas

Palab. ortográficas

Elem. gramaticais

Palab. ortográficas próx.

Elem. gramaticais próx.

Resultado

Tipo de resultado

Inventario

Ordenación

Total

Coincidencia

Tamaño de páxina

50

Agrupación

Unidade

Filtros

Orixe

Calquera

Bloque

Calquera

Xénero

Calquera

Subtipo

Calquera

Medio

Calquera

Sección

Calquera

Dende

Calquera

Ata

Calquera

Área temática

Calquera

Subárea

Calquera

Sexo-autor

Calquera

Sexo-interlocutor

Calquera

Autor

Calquera

Obra

Calquera

Documento

Calquera

Buscar en

Todo

Palab. ortográficas

Cinco palabras máximo

Volver

Limpar

Buscar

PEQUENA CALA SOBRE A DENOMINACIÓN PAGÁ E CRISTIÁ DOS DÍAS DA SEMANA

PAGÁ

luns

martes

mércores

xoves

venres

CRISTIÁ

segunda feira

terza ~ terceira feira

corta ~ cuarta feira

quinta feira

sexta feira

<https://portaldaspalabras.gal/video/os-nomes-dos-dias-da-semana>

"PODES ATOPAR TERZA FEIRA OU CORTA FEIRA NUN DICIONARIO PERO XA MOI POUQUÍÑA XENTE USA AQUÍ ESAS FORMAS"



<https://www.tiktok.com/@digochoeu/video/7307720551181618464?lang=es>

Busca

Corpus

Etiquetado automaticamente

Tipo

Elem. gramaticais

Sensibilidade

- Acentos
- Maiúsculas

Resultado

Tipo de resultado

Inventario

Ordenación

Total Coincidencia

Tamaño de páxina

50

Agrupación

Elemento gramatical

Filtros

Orixe

Calquera

Bloque

Calquera

Xénero

Calquera

Subtipo

Calquera

Medio

Calquera

Sección

Calquera

Dende

Calquera

Ata

Calquera

Área temática

Calquera

Subárea

Calquera

Sexo-autor

Calquera

Sexo-interlocutor

Calquera

Autor

Calquera

Obra

Calquera

Documento

Calquera

Buscar en

Todo

Elem. gramaticais

Elemento gramatical

Etiqueta

Lema

Hiperlema

Unidade



[Volver](#) [Limpar](#) [Buscar](#)

A BUSCA: LUNS|MARTES|MÉRCORES|XOVES|VENRES|SEGUNDA FEIRA|TER*A FEIRA|C*RTA FEIRA|QUINTA FEIRA|SEXTA FEIRA

Elem. gramaticais

ajc*rtafeira|quinta feira|sexta feira | Etiqueta  Lema Hiperlema Unidade 

Descargar  Volver Limpar Buscar

Resultados 1 a 11 de 11(12.530)

Selección de columnas

Valor absoluto, Valor relativo (por millón) 	Lustro 	Área temática 
Medio 	Orixe 	Bloque 
Subtipo 	Xénero 	Sexo-autor 
Sexo-interlocutor 		

<

		Total (54.737.277 / 57.693)
1	venres	3046 56 / 2206
2	luns	2745 50 / 2015
3	xoves	2525 46 / 1899
4	martes	2100 38 / 1626
5	mércores	1941 35 / 1561
6	terza feira	53 1 / 33
7	segunda feira	37 1 / 31
8	cuarta feira	28 1 / 26
9	sexta feira	28 1 / 24
10	quinta feira	26 0 / 25
11	corta feira	1 0 / 1

Resultados 1 a 11 de 11(12.530)

Selección de columnas

Valor absoluto, Valor relativo (por millón) ▾

Medio ▾

Subtipo ▾

Sexo-interlocutor ▾

Lustro (10) ▾

Orixe ▾

Xénero ▾

Área temática

Bloque

Sexo-autor

		Lustro					
		Total (54.737.277 / 57.693)	1975-1979 (978.095 / 117)	1980-1984 (1.669.899 / 143)	1985-1989 (2.019.335 / 120)	1990-1994 (5.414.080 / 386)	1995-1999 (9.829.561 / 12.422)
1	venres	3046 56 / 2206	3 3 / 2	9 5 / 8	16 8 / 8	104 19 / 44	676 69 / 495
2	luns	2745 50 / 2015	12 12 / 6	19 11 / 11	20 10 / 7	68 13 / 38	623 63 / 502
3	xoves	2525 46 / 1899	4 4 / 4	4 2 / 3	12 6 / 8	108 20 / 50	676 69 / 506
4	martes	2100 38 / 1626	6 6 / 5	12 7 / 8	9 4 / 6	57 11 / 36	505 51 / 398
5	mércores	1941 35 / 1561	2 2 / 2	1 1 / 1	9 4 / 4	41 8 / 24	489 50 / 392
6	terza feira	53 1 / 33	0 0 / 0	0 0 / 0	0 0 / 0	0 0 / 0	0 0 / 0
7	segunda feira	37 1 / 31	0 0 / 0	0 0 / 0	0 0 / 0	0 0 / 0	1 0 / 1
8	cuarta feira	28 1 / 26	0 0 / 0	0 0 / 0	0 0 / 0	1 0 / 1	1 0 / 1
9	sexta feira	28 1 / 24	0 0 / 0	0 0 / 0	0 0 / 0	0 0 / 0	0 0 / 0
10	quinta feira	26 0 / 25	0 0 / 0	0 0 / 0	0 0 / 0	1 0 / 1	0 0 / 0
11	corta feira	1 0 / 1	0 0 / 0	0 0 / 0	0 0 / 0	0 0 / 0	0 0 / 0

Concentración de ocurrencias nos últimos lustros representados no corpus que evidencia un intento de recuperación das formas con denominación cristiá

Resultados 1 a 11 de 11(12.530)

Selección de columnas

Valor absoluto, Valor relativo (por millón) ▾	Lustro (10) ▾	Área temática ▾
Medio ▾	Orixe ▾	Bloque ▾
Subtipo ▾	Xénero ▾	Sexo-autor ▾
Sexo-interlocutor ▾		

9 (2.019.335 / 120) 1990-1994 (5.414.080 / 386) 1995-1999 (9.829.561 / 12.422) 2000-2004 (7.341.493 / 9479) 2005-2009 (9.536.629 / 13.877) 2010-2014 (10.565.828 / 14.936) 2015-2019 (6.963.086 / 6055) 2020-2024 (419.271 / 158)

	9 (2.019.335 / 120)	1990-1994 (5.414.080 / 386)	1995-1999 (9.829.561 / 12.422)	2000-2004 (7.341.493 / 9479)	2005-2009 (9.536.629 / 13.877)	2010-2014 (10.565.828 / 14.936)	2015-2019 (6.963.086 / 6055)	2020-2024 (419.271 / 158)
1 venres	104 19 / 44	676 69 / 495	559 76 / 419	478 50 / 377	760 72 / 567	433 62 / 282	8 19 / 4	
2 luns	68 13 / 38	623 63 / 502	593 81 / 443	502 53 / 372	535 51 / 397	357 51 / 232	16 38 / 7	
3 xoves	108 20 / 50	676 69 / 506	468 64 / 382	404 42 / 326	468 44 / 360	377 54 / 257	4 10 / 3	
4 martes	57 11 / 36	505 51 / 398	469 64 / 372	346 36 / 287	401 38 / 314	288 41 / 195	7 17 / 5	
5 mércores / 4	41 8 / 24	489 50 / 392	406 55 / 331	306 32 / 253	366 35 / 316	316 45 / 236	5 12 / 2	
6 terza feira	0 0 / 0	0 0 / 0	1 0 / 1	2 0 / 2	13 1 / 10	33 5 / 17	4 10 / 3	
7 segunda feira	0 0 / 0	1 0 / 1	0 0 / 0	3 0 / 2	11 1 / 10	15 2 / 13	7 17 / 5	
8 cuarta feira	1 0 / 1	1 0 / 1	0 0 / 0	0 0 / 0	13 1 / 11	9 1 / 9	4 10 / 4	
9 sexta feira	0 0 / 0	0 0 / 0	1 0 / 1	0 0 / 0	13 1 / 10	14 2 / 13	0 0 / 0	
10 quinta feira	1 0 / 1	0 0 / 0	1 0 / 1	0 0 / 0	10 1 / 10	9 1 / 8	5 12 / 5	
11 corta feira	0 0 / 0	0 0 / 0	0 0 / 0	0 0 / 0	0 0 / 0	1 0 / 1	0 0 / 0	

Resultados 1 a 11 de 11(12.530)

DISTRIBUCIÓN POR XÉNERO

Elem. gramaticais

Resultados 1 a 11 de 11(12.530)

Selección de columnas

		Xénero					
Total (54.737.277 / 57.693)		Dramático (2.009.843 / 213)	Ensaístico (10.477.936 / 735)	Narrativo (18.108.691 / 2397)	Xornalístico (23.448.769 / 54.209)	Non aplica (692.038 / 139)	
1	venres	3046 56 / 2206	17 8 / 10	60 6 / 24	440 24 / 173	2480 106 / 1969	49 71 / 30
2	luns	2745 50 / 2015	16 8 / 14	70 7 / 34	426 24 / 163	2185 93 / 1779	48 69 / 25
3	xoves	2525 46 / 1899	30 15 / 17	126 12 / 58	291 16 / 142	2038 87 / 1661	40 58 / 21
4	martes	2100 38 / 1626	21 10 / 12	42 4 / 23	226 12 / 122	1749 75 / 1437	62 90 / 32
5	mércores	1941 35 / 1561	11 5 / 5	27 3 / 18	165 9 / 92	1700 72 / 1426	38 55 / 20
6	terza feira	53 1 / 33	1 0 / 1	2 0 / 1	12 1 / 3	38 2 / 28	0 0 / 0
7	segunda feira	37 1 / 31	0 0 / 0	5 0 / 3	1 0 / 1	31 1 / 27	0 0 / 0
8	cuarta feira	28 1 / 26	0 0 / 0	1 0 / 1	1 0 / 1	26 1 / 24	0 0 / 0
9	sexta feira	28 1 / 24	0 0 / 0	3 0 / 1	1 0 / 1	24 1 / 22	0 0 / 0
10	quinta feira	26 0 / 25	0 0 / 0	0 0 / 0	2 0 / 2	24 1 / 23	0 0 / 0
11	corta feira	1 0 / 1	0 0 / 0	0 0 / 0	1 0 / 1	0 0 / 0	0 0 / 0

CONCORDANCIAS DE TERZA FEIRA

Elem. gramaticais

terza feira Etiqueta Lema Hiperlema Unidade +

Descargar ▾ Volver Limpar Buscar

Resultados 1 a 38 de 38 Cabeceira

Documento: Sermos Galiza / SG2012-11-30/8 Editorial: Sermos Galiza S.A. Orixe: Escrita Subtipo: Xornal Sección no texto: Corpo	Autor: Rodil, Olalla Data: 2012 Bloque: Non ficción Áreas temáticas: Economía e política / Outros	Sexo-autor: Muller Medio: Xornal Xénero: Xornalístico Sección: Actualidade
1 2012 / Xornal	en o discurso de investidura pronunciado a pasada terza feira por lo xa presidente de a Xunta de Galiza, avanzaba que durante esta lexislatura o s	
Documento: Sermos Galiza / SG2012-11-30/8 Editorial: Sermos Galiza S.A. Orixe: Escrita Subtipo: Xornal Sección no texto: Corpo	Autor: Rodil, Olalla Data: 2012 Bloque: Non ficción Áreas temáticas: Economía e política / Outros	Sexo-autor: Muller Medio: Xornal Xénero: Xornalístico Sección: Actualidade
2 2012 / Xornal	de a oposición, que en este período de sesións contarán con unha forza máis, xa amosaban a pasada terza feira , após o discurso de investidura pronunciado por Alberto Núñez Feijóo, o seu rexeita	
Documento: Sermos Galiza / SG2012-11-02/22 Editorial: Sermos Galiza S.A. Orixe: Escrita Subtipo: Xornal Sección no texto: Corpo	Autor: Vázquez, Etti Data: 2012 Bloque: Non ficción Áreas temáticas: Economía e política	Sexo-autor: Muller Medio: Xornal Xénero: Xornalístico Sección: Cultura sociedade
3 2012 / Xornal	"chamaron a familia para unha reunión en a vindeira terza feira ", conta María, integrante de o colectivo cidadán.	
Documento: Sermos Galiza / SG2012-09-14/4 Editorial: Sermos Galiza S.A. Orixe: Escrita Subtipo: Xornal Sección no texto: Corpo	Autor: Partal, Vicent Data: 2012 Bloque: Non ficción Áreas temáticas: Economía e política / Ciencias sociais / Outros	Sexo-autor: Home Medio: Xornal Xénero: Xornalístico Sección: Opinión
4 2012 / Xornal	"escoita, España", a respectábel oda de Joan Maragall apareceu esta terza feira en a portada de un xornal de Barcelona.	
Documento: Sermos Galiza / SG2014-11-27/23 Editorial: Sermos Galiza S.A. Orixe: Escrita Subtipo: Xornal Sección no texto: Corpo	Autor: Lorenzo, X Data: 2014 Bloque: Non ficción Áreas temáticas: Economía e política	Sexo-autor: Descoñecido Medio: Xornal Xénero: Xornalístico Sección: Cultura sociedade
5 2014 / Xornal	convocou unha protesta en a porta de o centro a a que lle querían dar continuidade días despois, a terza feira pasada.	

O FLEXIONADOR DE XIADA

PRECEDENTES: CONXUGADORES

CilenisConxugador™

Só verbos coñecidos.

Conxugación do verbo: *andar*

Presente do Indicativo

Eu **ando**
Ti **andas**
El **anda**
Nós **andamos**
Vós **andades**
Eles **andan**

Pretérito Perfecto do Indicativo

Eu **andei**
Ti **andaches**
El **andou**
Nós **andamos**
Vós **andastes**
Eles **andaron**

Pretérito Imperfecto do Indicativo

Eu **andaba**
Ti **andabas**
El **andaba**
Nós **andabamos**
Vós **andabades**
Eles **andaban**

Pret. Pluscuamperfecto do Indicativo

Eu **andara**
Ti **andaras**
El **andara**
Nós **andáramos**
Vós **andárades**
Eles **andaran**

Futuro do Indicativo

Eu **andarei**
Ti **andaras**
El **andará**
Nós **andaremos**
Vós **andaredes**
Eles **andarán**

Condicional

Eu **andaría**
Ti **andaría**
El **andaría**
Nós **andaríamos**
Vós **andaríades**
Eles **andarían**

Presente do Subxuntivo

Eu **ande**
Ti **andes**
El **ande**
Nós **andemos**
Vós **andedes**
Eles **anden**

Pretérito Imperfecto do Subxuntivo

Eu **andase**
Ti **andases**
El **andase**
Nós **andásemos**
Vós **andáseades**
Eles **andasen**

Futuro do Subxuntivo

Eu **andar**
Ti **andares**
El **andar**
Nós **andáramos**
Vós **andáreades**
Eles **andarán**

Infinitivo Conxugado

Eu **andar**
Ti **andares**
El **andar**
Nós **andáremos**
Vós **andáredes**
Eles **andaren**

Imperativo Afirmativo

Ti **anda**
El **ande**
Nós **andemos**
Vós **andade**
Eles **anden**

Imperativo Negativo

Ti **andes**
El **ande**
Nós **andemos**
Vós **andedes**
Eles **anden**

Formas Nominais

Infinitivo **andar**
Xerundio **andando**
Participio **andado**

Acerca de L@verca@



Edicións Xerais de Galicia S.A.
Doutor Marañón, 12. 36211 Vigo (ES)
Tfno. +34 986 214888 / 986 214880



Laverca 1.0 © Copyright 2002
Reservados todos los derechos

Autores :

- Manuel González González
- Carmen García Mateo
- Eduardo Rodríguez Banga
- Elisa Fernández Rei

Programación :

- Xavier Fernández Salgado
 - Camilo Giráldez Ruibal
 - Francisco J. Méndez Pazó
- Avaliación e Correccións:**
- Begoña González Rei
 - Luis Xuncal Pereira
 - María Vallejo González

Nota: Para a elaboración deste programa utilizáronse algúns módulos do conversor texto-fala Cotovia desenvolvido no proxecto "Síntese de fala" do "Centro Ramón Piñeiro para a Investigación en Humanidades".

L@VERC@ - [Resultado de conxugar: andar]

Ficheiro Edición Texto Opcións Ventá Axuda

Modo	Tempo	Persoa	Conxugación
Presente		eu	ando
		ti	andas
		el/ela	anda
		nós	andamos
		vós	andades
Imperfecto		eu	andaba
		ti	andabas
		el/ela	andaba
		nós	andabamos
		vós	andabades
Perfecto		eu	andei
		ti	andaches
		el/ela	andou
		nós	andamos
		vós	andastes
Pluscuamperfecto		eu	andara
		ti	andaras
		el/ela	andara
		nós	andáramos
		vós	andárades
Futuro		eu	andarei
		ti	andará
		el/ela	andará
		nós	andaremos
		vós	andaredes

Ler Imprimir Parar Axuda

Analizador

Analiza Palavra

Analiza Arquivo

Conxugador

andar

Opcións Conxugador

Conxugar

Conxuga Arquivo

REAL ACADEMIA GALEGA

BIBLIOTECA HEMEROTECA ARQUIVO PUBLICACIONES

A LINGUA A INSTITUCIÓN LETRAS GALEGAS COMUNICACIÓN DICCIONARIOS

INICIO | DICCIONARIO

Diccionario

DICCIONARIO ACTUALIZACIÓNS CRÉDITOS

RUSCA SIMPLE SINÓNIMOS BUSCA AVANZADA

andar

andar **Conxugar**

verbo intransitivo

1 Trasládarse dun lugar a outro dando pasos.
Volveu á casa **andando**. O neno está aprendendo a **andar**.
= SINÓNIMO **caminar**

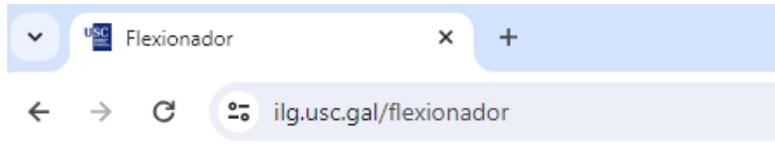
Conxugación do verbo: andar

Indicativo	Subxuntivo	Imperativo
Presente	Presente	Presente
ando	ande	-
andas	andes	anda
anda	ande	-
andamos	andemos	-
andades	andedes	andade
andan	anden	-
Pretérito imperfecto	Pretérito imperfecto	Infinitivo Conx.
andaba	andase	andar
andabas	andases	andares
andaba	andase	andar
andabamos	andásemos	andamos
andabades	andáseades	andades
andaban	andasen	andaren
Presente Perfecto	Futuro	Xerundio
andei	andarei	andando
andaches	andaras	-
andou	andará	-
andamos	andaremos	-
andastes	andaredes	-
andaron	andarán	-
Condicional	Participio	Participio
andaría	andando	andado
andaría	andares	andada
andaría	andar	andados
andaríamos	andamos	andadas
andaríades	andades	-
andarían	andaren	-



Conshuga (2009)
Grupo ProLNat da USC

PRECEDENTES: FLEXIONADORES



Flexionador

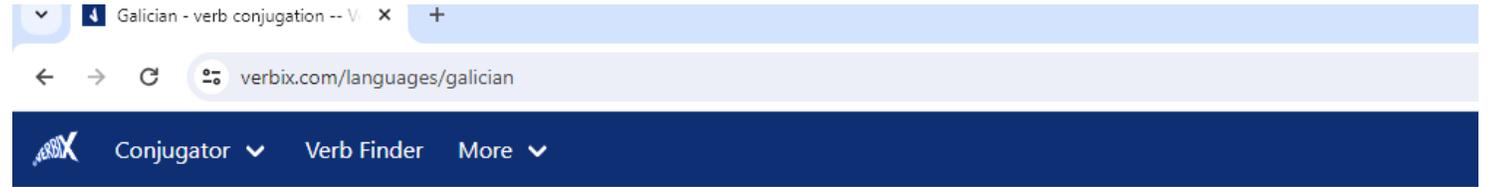
Lema:

Idioma:

Salida:

```
reitor (M,S,N)
reitora (F,S,N)
reitoras (F,P,N)
reitores (M,P,N)
reitor (M,S,A)
reitora (F,S,A)
reitoras (F,P,A)
reitores (M,P,A)
```

*(Xénero, Número, PoS)



Galician verb conjugation

Galician is a regional language spoken by 4 million native speakers in North Western Spain. It's a [Romance](#) language that evolved from the [Latin](#) language.

Conjugate a Galician Verb

Fill in the infinitive. Don't use any capital letters!

Some of the Galician characters don't exist in the English alphabet. If you can't input them here - try the following:

To Get	Enter	Sample Verbs
í	i/	oi/r equals oír
ñ	n~	pon~er equals poñer
ü	u:	liqu:idar equals liqüidar

See also:

- [Galician Nouns](#)
- [Galician Adjectives](#)

Recently Conjugated Verbs

excomungar, cribar, perforar, orixinar, galardoar, arensar, sobresaltar, parolar, coligar, ambalar, recolocar, rechazar, acougar, deter, pelar, antiquar, leccionar, bocexar, desencolerizar, engarameñar, enganchar, enfracuecer, fundir, viciar, achantar, esvaecer, frequentar, incoar, suicidar, relampadejar,

VERBIX: UNHA MOSTRA



Galician: andar Galician verb 'andar' conjugated

[Cite this page](#) | [Conjugate another Galician verb](#)

SHARE

Nominal Forms

Infinitivo: andar
Participio: andado
Xerundio: andando

Cognates

Indicative

Present

eu ando
ti andas
el anda
nós andamos
vós andades
eles andan

Imperfecto

eu andaba
ti andabas
el andaba
nós andabamos
vós andabades
eles andaban

Futuro

eu andarei
ti andarás
el andará
nós andaremos
vós andaredes
eles andarán

Pretérito perfecto

eu andei
ti andache
el andou
nós andamos
vós andastes
eles andaron

Subxuntivo

Presente (subxuntivo I)

eu ande
ti andes
el ande
nós andemos
vós andedes
eles anden

Imperfecto (subxuntivo II)

eu andase
ti andases
el andase
nós andasemos
vós andasedes
eles andasen



EN RESUMIDAS CONTAS, POR QUE UN NOVO FLEXIONADOR?

- ❑ Só hai un integral para galego, que coñezamos, e presenta erros.
- ❑ Posuímos unha base de datos léxica con información sobre a normatividade dos lemas e das formas, que ten en conta norma e uso.
- ❑ Cremos que se pode aproveitar para dispoñibilizar unha nova ferramenta que resulte de utilidade:
 - na aprendizaxe do galego, tanto en L1 como en L2.
 - no procesamento da linguaxe natural [os correctores ortográficos e os sistemas de tradución automática precisan contar coas formas flexionadas]
 - a lexicógrafos e tradutores, ao poder dispoñer de exemplos en contextos reais de uso.

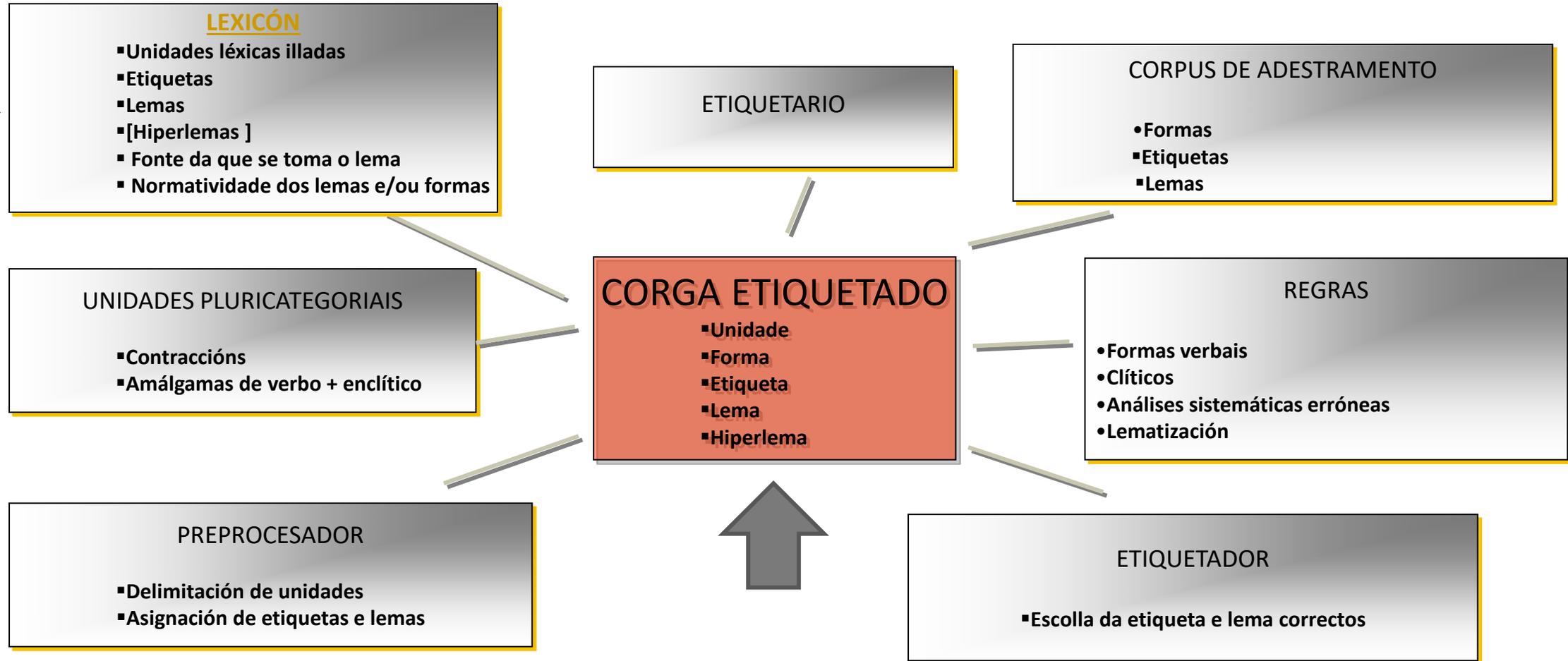
A CONEXIÓN CO CORGA

O FLEXIONADOR DE XIADA

Ferramenta para a anotación morfosintáctica automática do galego

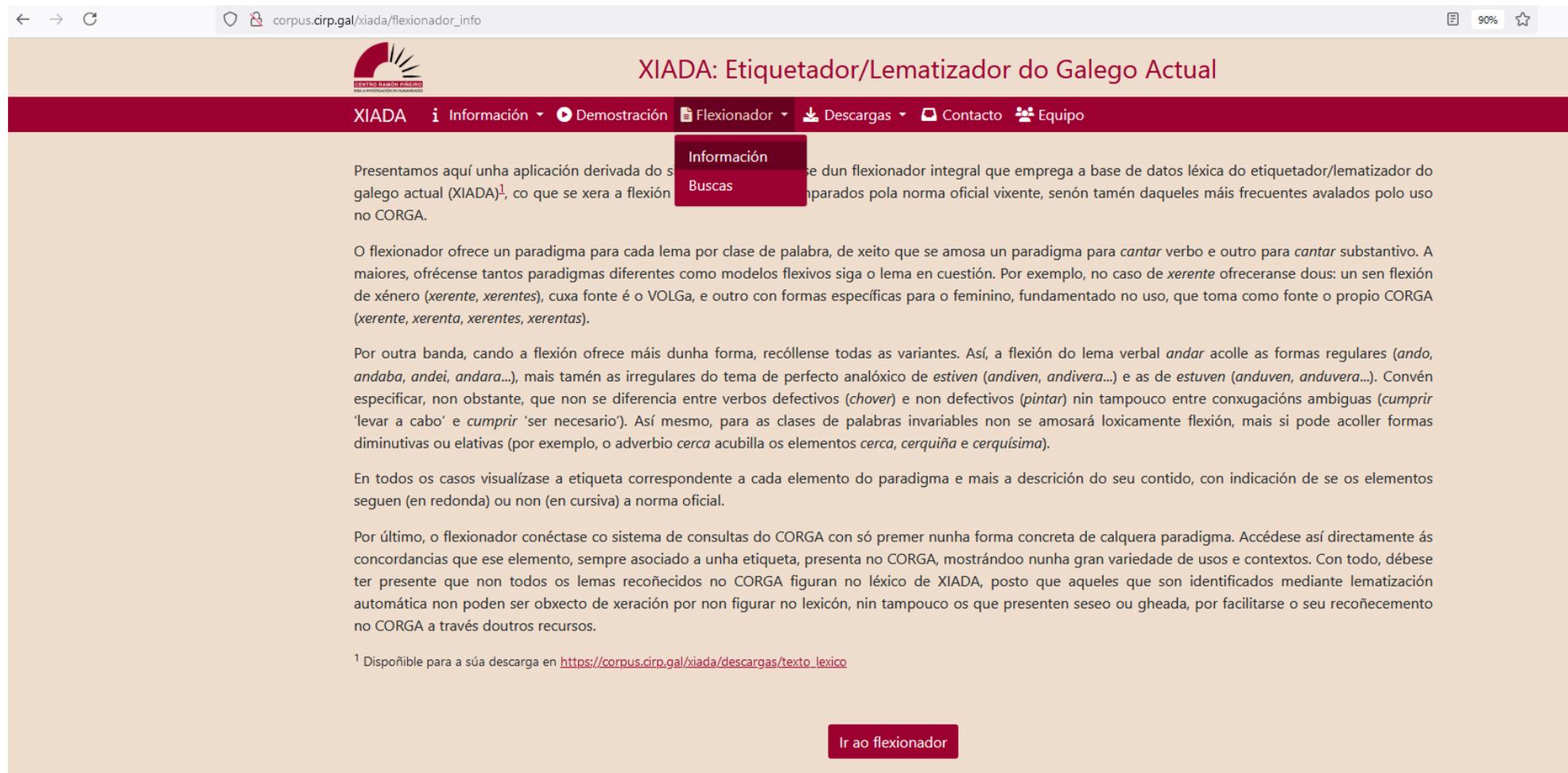
XIADA

Obxectivo: etiquetaxe e lematización automáticas do CORGA



Etiquetaxe e desambiguación automáticas en galego: o sistema XIADA

ACCESO AO FLEXIONADOR DE XIADA: [HTTP://CORPUS.CIRP.GAL/XIADA](http://corpus.cirp.gal/xiada)



The screenshot shows a web browser window with the address bar displaying corpus.cirp.gal/xiada/flexionador_info. The page title is "XIADA: Etiketador/Lematizador do Galego Actual". The navigation menu includes "XIADA", "Información", "Demostración", "Flexionador", "Descargas", "Contacto", and "Equipo". The "Flexionador" menu item is highlighted, and a dropdown menu is visible with "Información" and "Buscas" options. The main content area contains several paragraphs of text describing the application and its features.

Presentamos aquí unha aplicación derivada do sistema de un flexionador integral que emprega a base de datos léxica do etiquetador/lematizador do galego actual (XIADA)¹, co que se xera a flexión comparados pola norma oficial vixente, senón tamén daqueles máis frecuentes avalados polo uso no CORGA.

O flexionador ofrece un paradigma para cada lema por clase de palabra, de xeito que se amosa un paradigma para *cantar* verbo e outro para *cantar* substantivo. A maiores, ofrécese tantos paradigmas diferentes como modelos flexivos siga o lema en cuestión. Por exemplo, no caso de *xerente* ofreceranse dous: un sen flexión de xénero (*xerente, xerentes*), cuxa fonte é o VOLGa, e outro con formas específicas para o feminino, fundamentado no uso, que toma como fonte o propio CORGA (*xerente, xerenta, xerentes, xerentas*).

Por outra banda, cando a flexión ofrece máis dunha forma, recóllense todas as variantes. Así, a flexión do lema verbal *andar* acolle as formas regulares (*ando, andaba, andei, andara...*), mais tamén as irregulares do tema de perfecto analóxico de *estiven* (*andiven, andivera...*) e as de *estuvén* (*anduvén, anduvera...*). Convén especificar, non obstante, que non se diferencia entre verbos defectivos (*chover*) e non defectivos (*pintar*) nin tampouco entre conxugacións ambiguas (*cumprir* 'levar a cabo' e *cumprir* 'ser necesario'). Así mesmo, para as clases de palabras invariables non se amosará loxicamente flexión, mais si pode acoller formas diminutivas ou elativas (por exemplo, o adverbio *cerca* acubilla os elementos *cerca, cerquiña* e *cerquísima*).

En todos os casos visualízase a etiqueta correspondente a cada elemento do paradigma e mais a descrición do seu contido, con indicación de se os elementos seguen (en redonda) ou non (en cursiva) a norma oficial.

Por último, o flexionador conéctase co sistema de consultas do CORGA con só premer nunha forma concreta de calquera paradigma. Accédese así directamente ás concordancias que ese elemento, sempre asociado a unha etiqueta, presenta no CORGA, mostrándoo nunha gran variedade de usos e contextos. Con todo, débese ter presente que non todos os lemas recoñecidos no CORGA figuran no léxico de XIADA, posto que aqueles que son identificados mediante lematización automática non poden ser obxecto de xeración por non figurar no lexicón, nin tampouco os que presenten seseo ou gheada, por facilitarse o seu recoñecemento no CORGA a través doutros recursos.

¹ Dispoñible para a súa descarga en https://corpus.cirp.gal/xiada/descargas/texto_lexico

[Ir ao flexionador](#)

FLEXIÓN DE CIRCULAR

 **XIADA: Etiquetador/Lematizador do Galego Actual**

XIADA Información Demostración Flexionador Descargas Contacto Equipo

Escriba na caixa de texto o lema que queira flexionar:

Lema: Flexionar

Verbo

Indicativo

Presente <ul style="list-style-type: none">• circulo• circulas• circula• circulamos• circulades• circulan	Copretérito <ul style="list-style-type: none">• circulaba• circulabas• circulaba• circulabamos <i>circulábamos</i>• circulabades <i>circulábades</i>• circulaban	Antepretérito <ul style="list-style-type: none">• circulara• circularas• circulara• circularamos <i>circuláramos</i>• circularades <i>circulárades</i>• circularan
Futuro <ul style="list-style-type: none">• circularéi <i>circularéi</i>• circularás• circulará• circularemos• circularedes• circularán	Pretérito <ul style="list-style-type: none">• circulei <i>circuléi</i>• circulaches <i>circulache</i>• circulouse <i>circulóu</i>• circulamos• circulastes• circularon	Pospretérito <ul style="list-style-type: none">• circularía• circularías• circularía• circularíamos <i>circularíamos</i>• circulariades <i>circularíades</i>• circularían

Non normativo - Vii10p:
Verbo, Copretérito,
Indicativo, Primeira, Non
aplica, Plural

Scfs: Substantivo, Común,
Feminino, Singular **VO** volga_2004

circular Scfs
circulares Scfp

Adxectivo volga_2004

circular A0as
circulares A0ap

- 3 clases de palabras, por tanto 3 paradigmas distintos.
- Fonte da que se toma o lema: VOLGa
- Etiqueta que caracteriza no lexicón cada elemento.
- Desenvolvemento da descrición da etiqueta.
- Indicación sobre a normatividade dos elementos do paradigma:
 - Formato da letra:
 - redonda = normativo
 - itálica = non normativo
 - Indicación expresa de que o elemento non segue a norma, xunto coa súa etiqueta e descrición desta.
- Ligazón aos resultados en modo concordancias da busca de cada elemento do paradigma no CORGA.

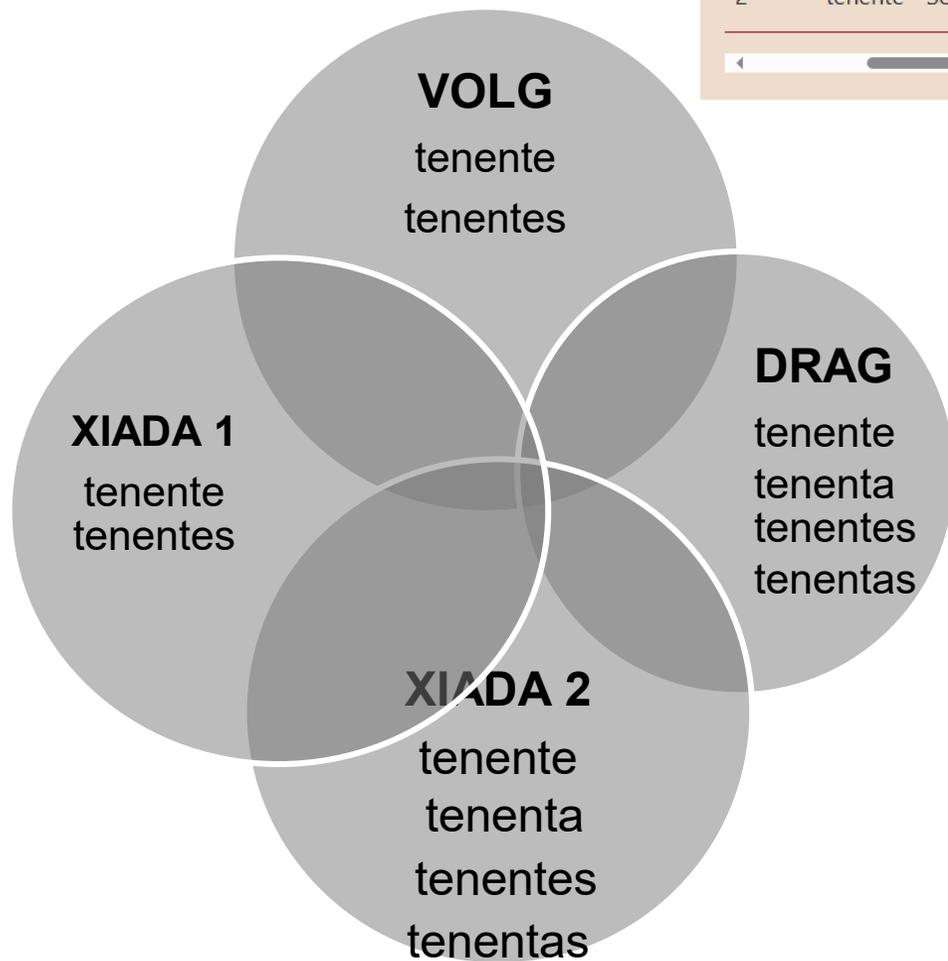
FLEXIÓN DE TENENTE

Lema cuxa flexión orixina dous paradigmas diferentes, documentado ademais en dúas fontes distintas.



The screenshot shows the website 'XIADA: Etiquetador/Lematizador do Galego Actual'. The navigation bar includes links for 'Información', 'Demostración', 'Flexionador', 'Descargas', 'Contacto', and 'Equipo'. The main content area has a text input field with 'Lema: tenente' and a 'Flexionar' button. Below this, two sections are shown: 'Substantivo drag_2012' and 'Substantivo volga_2004'. The first section lists various inflected forms of 'tenente' (e.g., tenente Scms, tenenta Scfs) and includes a tooltip for 'Non normativo - Scxp: Substantivo, Común, Non binario, Plural'. The second section lists forms for 'volga_2004' (e.g., tenente Scas, tenentes Scap).

Raíz	Subetiqueta	Grupo	Lema	Hiperlema	Lema recomendado	Normativa lema 2003	Categoría lema	Centros de interese	Lexicón	Sublexicón	Fonte	Válida
1	tenent	Sc	G13	tenente	tenente	✓	S		Xeral	principal	drag_2012	✓
2	tenente	Sca	G29	tenente	tenente	✓	S		Xeral	principal	volga_2004	✓



A **tenente** comprendeu o estraño que tería sido actuar daquel xeito e tamén dubidou do seu propio argumento.

A **tenente** puxo a súa man libre sobre o cadril e deu a volta.

A **tenente** suspirou.

O **tenente** de alcalde de vigo, Santi Domínguez, inaugurou c

O **tenente** coronel destacou o resultado "moi positivo" dos la

"Sen dúbida foi o peor momento da misión", díxolles o **tenente** coronel aos xornalistas, antes de asegurar que a

Logo de aterrar o voo comercial ás 07.20 horas, o **tenente** coronel Louro, xefe da expedición, quixo recordar

A **tenente** de alcalde responsable da área de Urbanismo do

(Convulsa, nos brazos da **tenenta** de aia, chorando e gritando.)

A **tenenta** de navío Agartha Posse II pasara a noite no cuarto infantil.

FLEXIÓN DE ALGÚN

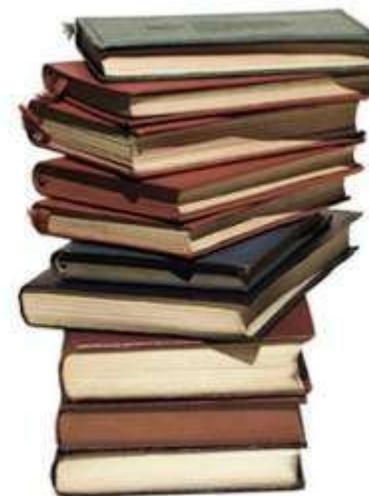
Desaxustes entre o flexionador e o CORGA:

- Etiquetas con hipervalores
- Lematización automática (*teletraballar, alghún, apareser...*)

The screenshot shows the web interface for 'XIADA: Etiketador/Lematizador do Galego Actual'. The header includes the logo of 'CENTRO RAMÓN PIRELLO' and the title. A navigation bar contains links for 'XIADA', 'Información', 'Demostración', 'Flexionador', 'Descargas', 'Contacto', and 'Equipo'. The main content area has a text input field with the word 'algún' and a 'Flexionar' button. Below the input, the word 'Indefinido' is shown with a red underline, followed by the text 'volga_2004'. A list of flexionated forms is displayed: 'algún lams', 'algunha lafs', 'algúns lamp', 'algunhos lamp', 'algús lamp', and 'algunhas lafp'.

FONTES DOS LEMAS

- Corpus de Referencia do Galego Actual [4.1] - <http://corpus.cirp.gal/corga>
- Dicionario da Real Academia Galega - <http://www.realacademiagalega.org/dicionario>
- Dicionario de galego - DIGALEGO - <https://digalego.xunta.gal>
- Dicionario de dicionarios - <https://ilq.usc.gal/ddd>
- *TERGAL: Banco de termos galegos recomendados* [1.5.2] - <https://www.cirp.gal/tergal>
- Carballeira Anllo, X. M. (coord.): *Gran dicionario Xerais da lingua*. Vigo, Xerais, 2009.
- García, C. e M. González González (dirs.): *Diccionario da Real Academia Galega*. A Coruña, Galaxia, 1997.
- García, C. e M. González González (dirs.): *Diccionario castelán-galego da Real Academia Galega*. A Coruña, Fundación Pedro Barrié de la Maza, 2004.
- González González, M. e A. Santamarina Fernández (coords.): *Vocabulario ortográfico da lingua galega (VOLGa)*. Santiago de Compostela / A Coruña, Real Academia Galega / Instituto da Lingua Galega, 2004.
- González González, B.: *Ortografía da lingua galega*. A Coruña, Galinova Editorial, 2004.
- Ledo Cabido, Bieito (dir.): *Diccionario de galego*. Vigo, Ir Indo, 2004.
- Pena, Xosé Antonio (dir.): *Gran dicionario século 21 da lingua galega*. Vigo, Galaxia, 2006.
- Real Academia Galega e Instituto da Lingua Galega: *Normas ortográficas e morfolóxicas do idioma galego*. Santiago de Compostela / A Coruña, Real Academia Galega / Instituto da Lingua Galega, 2003.



MELLORAS DESEXABLES

- **Incluír as unidades multipalabra.**
- **Ofrecer no paradigma flexivo a información de que elementos están presentes no CORGA, de xeito que se poida contrastar visualmente xa o inventario (a flexión íntegra do lema á que dá orixe a base de datos léxica) e o uso (fundamentado nas ocorrencias que un determinado elemento teña no corpus de referencia da lingua galega).**



CONCLUSIÓNS I

- ❑ Os avances en lingüística de corpus e lingüística computacional permitíronnos desenvolver dous recursos máis para facilitar o estudo da lingua galega: un dicionario de frecuencias e un flexionador.
- ❑ Os avances en lingüística computacional simplifican e optimizan a análise lingüística.
- ❑ Describimos como a modalidade *Inventario* no xeito de amosar os resultados dunha procura se converte nun dicionario de frecuencias dinámico.
- ❑ Demostramos a importancia deste dicionario para o estudo da variación e o cambio lingüístico en expresións concretas, fronte ao uso máis habitual de coñecer a macroestrutura estatística do galego.
- ❑ Trátase dun dicionario dinámico, flexible e cunha gran potencialidade grazas á distribución segundo os variados parámetros que ofrece: lustro, área temática, medio, orixe, bloque, subtipo de documento, sexo do autor e sexo do interlocutor.
- ❑ Localízase no interior do CORGA, co cal se pode acudir ás concordancias, frecuencias completas, cabeceiras ou calquera outra información que se puidera precisar para entender e completar os datos.

CONCLUSIÓN II

Poñemos á disposición pública un novo flexionador para o galego que:

- ❑ cobre tanto a flexión verbal coma a nominal;
- ❑ proporciona a flexión de lemas normativos, mais tamén inclúe variantes dialectais que non seguen necesariamente a norma oficial e mesmo castelanismos ou hipercorreccións;
- ❑ diferencia a normatividade dos elementos do paradigma a través do tipo de letra: redonda = normativo, cursiva = non normativo;
- ❑ dá conta dos distintos modelos flexivos que seguen os diferentes lemas: *tenente/tenentes* e *tenente/tenenta/tenentes/tenentas*;
- ❑ conecta co CORGA para amosar os exemplos reais de uso que no corpus se rexistren do elemento asociado á etiqueta que se seleccione;
- ❑ actualiza o seu lemario cada vez que se incorporan entradas ao lexicón.

ALGUNHAS REFERENCIAS

- Bugarín López, M. X. 2007. *O galego fundamental: diccionario de frecuencias*. A Coruña: Fundación Pedro Barrié de la Maza.
- Cilenis Conxugador: <https://fegalaz.usc.es/~gamallo/php/conjugador/conjuga-web/>
- Conshuga: <https://gramatica.usc.es/pln/gl/tools/conjugador/conjugador.php>
- Corpus de Referencia do Galego Actual (CORGA): <http://corpus.cirp.gal/corga>
- Díaz Regueiro, Manuel. 1992. *Os verbos galegos*. Santiago de Compostela: Xunta de Galicia.
- Etiquetador/lematizador do galego actual (XIADA): <http://corpus.cirp.gal/XIADA>
- Flexionador ILG: <https://ilg.usc.gal/flexionador>
- Gamallo P., Garcia, M., González, I., Muñoz, M., Del Río, I. 2013. *Learning verb inflection using Cilenis conjugators*, *Eurocall Review*, 21(1): 12-19. <http://hdl.handle.net/10251/138315>
- González González, M. (dir.): *Diccionario da Real Academia Galega*. A Coruña: Real Academia Galega. (<https://academia.gal/diccionario>)
- González González, M. e A. Santamarina Fernández (coords.): *Vocabulario ortográfico da lingua galega (VOLGa)*. Santiago de Compostela / A Coruña, Real Academia Galega / Instituto da Lingua Galega, 2004.
- González González, M., C. García Mateo, E. Rodríguez Banga e E. Fernández Rei. 2002. *Diccionario de verbos galegos*. Laverca. Vigo: Edicións Xerais de Galicia.
- Verbix Verb Conjugator (<https://www.verbix.com/>)

<pausa/> foi todo señores <pausa/> **gracias** pola **atención** prestada <pausa/> pasen boa tarde

así rematamos <pausa/> **gracias** pola **atención** prestada <pausa/> pasen <pausa/> boa tarde

foi todo <pausa/> señores <pausa/> **gracias** pola **atención** prestada <pausa/> pasen boa tarde

<pausa/> as dúas e media <pausa/> **gracias** pola **atención** prestada <pausa/> pasen boa tarde

isa/> **son** as dúas e media <pausa/> **gracias** pola **atención** prestada <pausa/> pasen boa tarde

Moitas **grazas** pola **atención** _dixo bruscamente para rematar a conversa.

Moitas **grazas** pola súa **atención**.

Moitas **grazas** pola súa **atención**.

Preguntas?

O LEXICÓN DE XIADA: LEMA ANDAR

XIADA: Etiquetador/Lematizador do Galego Actual

[XIADA](#)
[Raíces](#)
[Locucións](#)
[Propios](#)
[Periféricas](#)
[Formas](#)
+ ▾

Raíces / Consultar

Consultar raíces

Lexicón

Sublexicón

Fonte

Centro de interese

Raíz

Normativa raíz 2003

Subetiqueta

Grupo

Lema

Hiperlema

Lema recomendado

Normativa lema 2003

	Raíz	Normativa raíz 2003	Subetiqueta	Grupo	Lema	Hiperlema	Lema recomendado	Normativa lema 2003	Categoría lema	Centros de interese	Lexicón	Sublexicón	Fonte
1	and	V	V1	andar	andar	✓	V	O corpo humano	Xeral	principal	volga_2004		
2	andar	Scm	G30	andar	andar	✓	S	O corpo humano	Xeral	principal	volga_2004		
3	andiv	✗	V	Vi23c	andar	andar	✓	V	O corpo humano	Xeral	principal	volga_2004	
4	anduv	✗	V	Vi23c	andar	andar	✓	V	O corpo humano	Xeral	principal	volga_2004	

GRUPOS DE DERIVACIÓN: PARTE DO V1

 **XIADA: Etiquetador/Lematizador do Galego Actual**

XIADA Raíces Locucións Propios Periféricas Formas + ▾

[Grupos](#) / Editar

Editar grupo

Nome Representante

[Modificar](#)

Desinencia	Subetiqueta	Siguiente grupo	Normativa 2003	
a	0m20s		✓	 
a	pi30s		✓	 
aba	ii20s		✓	 
abades	ii20p		✓	 
abamos	ii10p		✓	 
aban	ii30p		✓	 
abas	ii20s		✓	 
ache	ei20s		✗	 
aches	ei20s		✓	 
ad	0p0	G1	✓	 
ade	0m20p		✓	 