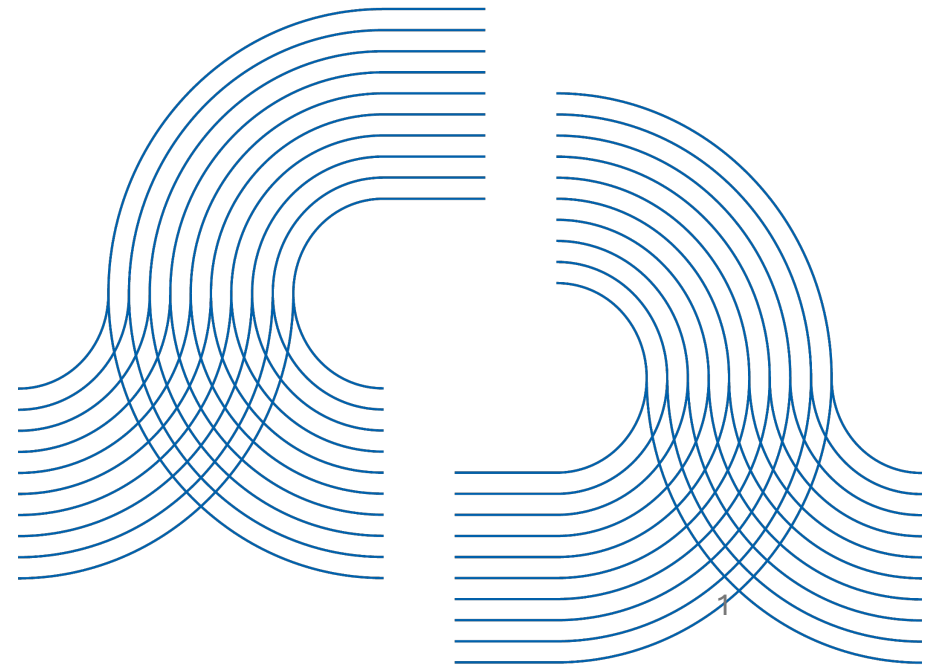


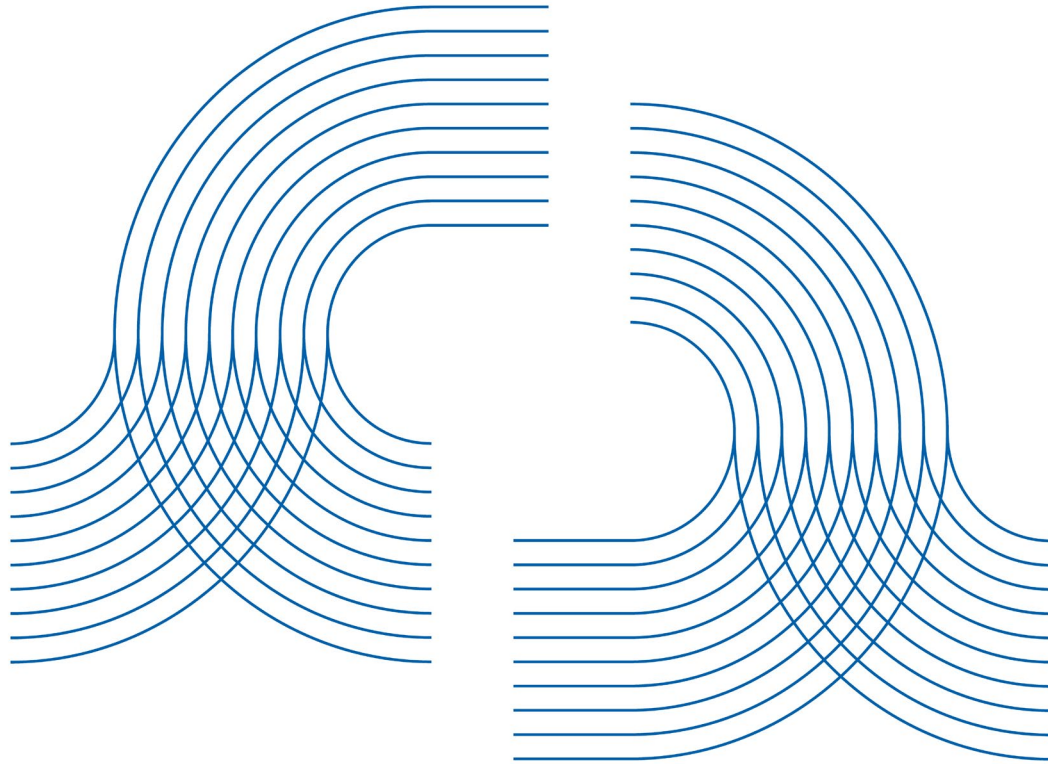
Proxectos e recursos do **Instituto de investigación LINGUA** da Universidade de Vigo

Javier Pérez Guerra
jperez@uvigo.gal

II Xeira CLARIAH-GAL, 9 maio 2025



Instituto de investigación LINGUA (iLingua)



Instituto de investigación
Lingua
Universidade de Vigo

- creado a finais de 2024
- persoal investigador de:
 - linguas (galego, español, inglés, francés, alemán, portugués, lingua de signos) (e literaturas)
 - tradución e interpretación
 - lingüística
 - comunicación
 - socioloxía
- 74 investigadores/as vinculados/as, 6 non permanentes, 15 en formación

Instituto de investigación LINGUA (iLingua)

- Liñas:
 - Lingüística interna e variación lingüística
 - Sociedade, ideoloxía e política lingüística
 - Linguas, cognición e educación
 - Tradución e interpretación
 - Lingua e literatura
 - Lingua e accesibilidade: medios e lingua de signos
 - Humanidades dixitais

Instituto de investigación LINGUA (iLingua)



Universidade de Vigo

SECRETARÍA

Instituto de Investigación LINGUA (iLingua)

ES / EN

iLingua ▾ Membros ▾ Investigación Transferencia Blog Contacto

Humanidades dixitais

La Disputación de Stephanie Malher

— Dans quatre jours seulement. Je suis encore là pendant quatre jours. Lundi, quand je l'ai vue, Stephanie disait avoir un rendez-vous qui allait lui apporter les éléments manquants à son dossier...

— Laisse l'affaire à l'un de tes collègues, me suggéra-t-elle.

— Hors de question ! Derek, cette fille m'a assuré qu'en 1994...

Il ne me laissa pas terminer ma phrase :

— On a bondé l'enquête, Jesse ! C'est du passé ! Qu'est-ce qui te prend tout d'un coup ? Pourquoi veux-tu à tout prix te replonger là-dedans ? Tu as vraiment envie de revivre tout ça ?

Je regrettais son manque de soutien.

— Alors, tu ne veux pas venir à Orphée avec moi ?

— Non, Jesse. Désolé. Je crois que tu mérites complètement.

C'est donc seul que je me rendis à Orphée, vingt ans après y avoir mis fin pour de bon à ma vie...

Blog

<https://ilingua.uvigo.gal/>

Universidade de Vigo

O iLingua cos posgraos da FFT

O iLingua no diario Atlántico

De linguas, tradución e cousas xurídicas

4

Bases de datos

ADESSE

Base de datos de verbos, alternancias de diátesis y esquemas sintáctico-semánticos del español
(coordinador: José María García-Miguel)

- <https://adesse.uvigo.es>
- base de datos de verbos e construcións verbais do español:
 - completa caracterización sintáctico-semántica por verbo
 - alternancias de diátese
 - frecuencias relativas a relacións semánticas semellantes

DAR I

Hacer tener, proporcionar, ofrecer

[-] Clasificación semántica y potencial valencial

Tipo de proceso:	Transferencia	
Argumentos:		Frecuencia
A0 DNTE	DONANTE	402 (89.9 %)
A1 POS-F	POSEEDOR-FINAL	357 (79.9 %)
A2 POS	POSESIÓN	403 (90.2 %)
A3 Concepto	Concepto	20 (4.5 %)

Perfil combinatorio >>

[-] Realización

Datos buscados

DAR I - Hacer tener, proporcionar, ofrecer

- Activa

act **A0:DNTE** **A2:POS** **A1:POS-F** **A3:Concepto** >
= **SUJ** = **ODIR** = **a OIND** = **por OBL**

Hay 4 ejemplos de 1 entradas verbales

Página: 1/1 Elementos por página: 20

Verbo	(ess)	Texto
DAR-I1	S-0 D-2 I-1 O-3	Intentar mentir una vez más, recurrir a identificaciones ficticias y dar al lector zeta por erre -en una palabra, admitir la verdad de los telegrafistas
DAR-I1	S-0 D-2 I-1 O-3	» --que me, me, me parece que nos daban trescientas pesetas por tarde ¿no?--.
DAR-I1	S-0 D-2 I-1 O-3	Nos daban por un dólar dólar yo no sé qué cantidad de eslotis, que es la moneda polaca, pero un montón.
DAR-I1	S-0 D-2 I-1 O-3	bonos y, por un esloti, por la peseta te dan muchos más bonos que eslotis, o sea, que tú sales ganando al cambio.

Recursos sobre corpus para español e LSE do
Grupo GRADES da Universidade de Vigo

ADESSE

ADESSE é unha base de datos de palabras e construcións verbais de español, ofrece a análise morfolóxica e sintáctica das palabras, así como a análise de palabras e frases completas. A análise morfolóxica inclúe a análise da flexión, a análise da derivación e a análise da composición. A análise sintáctica inclúe a análise da estrutura da frase, a análise da función das palabras e a análise da relación entre as palabras.

iSignos

iSignos é unha plataforma para traballar con signos. Permite crear e xestionar signos, así como a análise de signos. A análise de signos inclúe a análise da estrutura do signo, a análise da función do signo e a análise da relación entre os signos.

Trebank de LSE

Trebank de LSE é unha plataforma para traballar con signos. Permite crear e xestionar signos, así como a análise de signos. A análise de signos inclúe a análise da estrutura do signo, a análise da función do signo e a análise da relación entre os signos.

ECEG

Eighteenth-Century English Grammars database
(coordinadora: Nuria Yáñez)

- <https://eceg.iatext.ulpgc.es>
- base de datos compilada no contexto da estandarización escrita e oral do inglés no século XVIII
- inclúe metadatos bibliográficos e biográficos sobre 323 gramáticas do inglés e 275 autores/as

[Enable advanced mode](#)

1730?], 1731

No. 383 - A new method of teaching or acquiring the English tongue. London: R. Baldwin.

Author

Written by **Stirling, John**, male

Born in unknown, unknown, unknown

Occupation

Education schoolmaster**Religion** vicar, chaplain (domestic chaplain)

Biography

Stirling, John (d.1777).

'M.A., and Master of St. Andrew's School, Holbourn' (1735: title page). [Holbourn ~ Holborn].

'Vicar of Gaddesden, Hertfordshire, and Chaplain to his Grace the Duke of Gordon' (title page 1740, 2nd edn).

'Little is known about Stirling's life except that he held a MA, served as chaplain to the Duke of Gordon, and wrote a large number of extremely rare' (Alston). These books include English titles such as A short View of English Grammar (1735), A Course of Theology (1736), The Private Tutor to British Youth (1778), and Lily's Pedagogical Admonition (1787); but his most important publication was A System of rhetoric (1733), an elementary school text that listed and defined 94 rhetorical figures.' (Mora)

Author

Occupation

Biography

Contents

Type of work and division

Subsidiary content

Target audience

Imprint

Editions

References

Comments



ECEP

The Eighteenth-Century English Phonology database
(coordinadora: Nuria Yáñez e equipo do Digital Humanities Institute de Sheffield)

- <https://www.dhi.ac.uk/ecep/>
- base de metadatos de máis de 21.000 palabras con información fonolóxica (IPA) extraída de 11 dicionarios de pronuncia do inglés publicados na segunda metade do século XVIII

About the Author

Name	Life Dates	Gender	Class
Sheridan, Thomas	1719?-1788	male	unknown

Place of Birth	Occupation	Occupation Description
unknown_IS, unknown_IS, IrelandSouth	Education, Miscellaneous	lexicographer, grammarian, elocutionist, author, writer of education, actor

Biography

"Actor and orthoepist, was probably born in early 1719, though no records of birth or baptism have been found. His father, Dr Thomas Sheridan (1687–1738), was by then running a highly regarded school at his house in Capel Street, Dublin, but there is a possibility that Thomas, his third son, was born in Quilca, co. Cavan, perhaps in the childhood home of his mother, Elizabeth MacFadden. Jonathan Swift, a close friend of Dr Sheridan, was Thomas's godfather and, by Sheridan's own account, a beneficent influence on him from childhood."

Author Works

Show 10 entries

Filter: Start typing to filter...

Title

A general c
is, to estab
is prefixed
printed [by
Poultry; an

Showing 1 to

Lexical Search

Lexical Categories	Lexical SubSets	Example Words	IPA	IPA Variant	Attitudes
CHOICE	BATH_c	absorb			

Clear the Search

II XEIRA CLARIAH-GAL, 9 maio 2025, Santiago de Compostela

Recursos sobre corpus para a língua inglesa

Elena Seoane, Nuria Yéñez Bouza, Javier Pérez Guerra e grupo de investigación LVTC

Recursos sobre a ecoloxía lingüística en Xibraltar

Recursos para o estudo do inglés como L2

Recursos para o estudo da historia da lingua inglesa

Clear the Search

Corpus: Lingua de signos española

iSignos

Interfaz de datos de la Lengua de Signos Española
(coordinadores: Carmen Cabeza e José María García-Miguel)

- <https://isignos.uvigo.es>
- recurso de consulta de datos lingüísticos da lingua de signos española
- 30 gravacións en vídeo de signantes +
glosas das mans +
tradución ao español e ao inglés

iSignos

	Identificador	Género	Tema	Signante	Rango de edad	Sexo	Duración	Año
▶	ET-MJC	Entrevista	Vida sorda	MJC	35-54	Mujer	00:05:07	2015



x1 x0,5 x0.25

- ▶ T: Tú... [Sonríe.]
1 D: INDX.PRO:2sg
I: INDX
- ▶ T: [Señala al interlocutor y se ríe.]
2 D:
I: INDX
- ▶ T: [Se ríe.]
3 D: G:no-mirar-emisor
I: G:no-mirar-emisor
- ▶ T: [Se ríe.]
4 D: G.ADAPT
I:
- ▶ T: Esa cara no...
5 D: INDX G:no-pasa-nada CARA NO
I:
- ▶ T: Nada. nada.

iSignos

Contexto del ejemplo 1 del listado anterior.

Identificador: PS-NXG

Signante: NXG

Rango de edad: 35-54

Género: Narración

Mano: Derecha

Tema: Pear Film

Sexo: Hombre



▶ NXG: ¡Hola!

#1 D: HOLA

I:

▶ NXG: Voy a hablar de (la historia) de las peras,

#2 D: TEMA PERA4(Xc) DL:PERA

I: TEMA

▶ NXG: voy a contarla.

#3 D: AHORA EXPLICA

I: EXPLICA

Recursos sobre corpus para español e LSE do
Grupo GRADES da Universidade de Vigo
Xosé M. García-Rojas, Galindo, Perea del Carmen Cabeza Prieto
Instituto Lingüístico, Universidade de Vigo



UD_Spanish_Sign_Language-LSE

Universal Dependencies treebank da Lingua de Signos Española
(coordinador: José M^a García-Miguel)

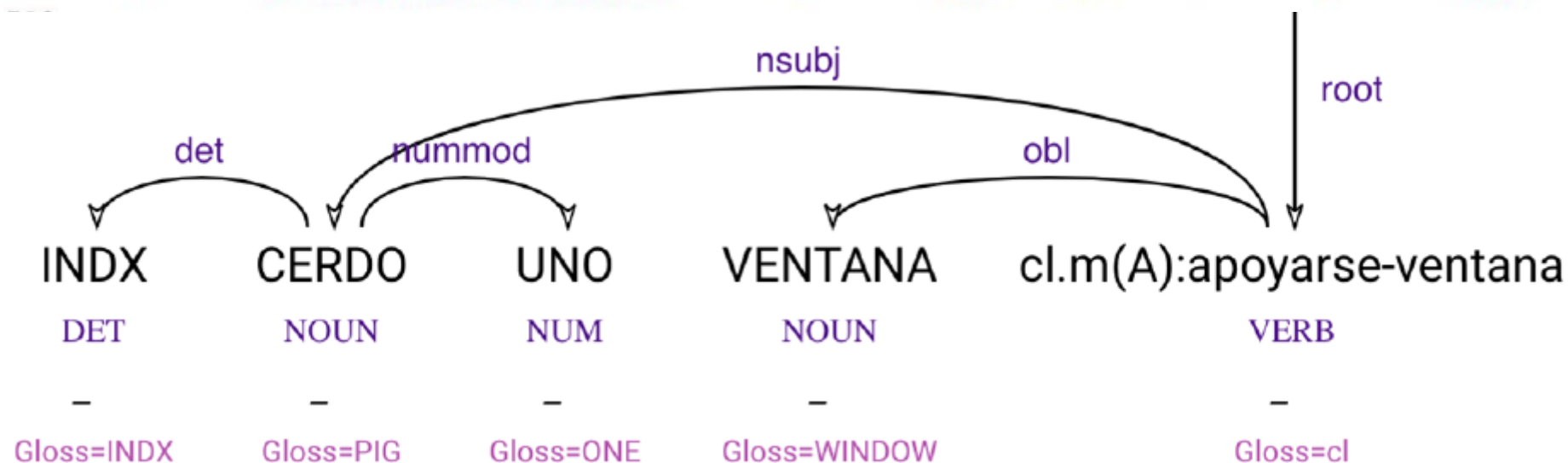
- https://universaldependencies.org/treebanks/ssp_lse/index.html,
https://github.com/UniversalDependencies/UD_Spanish_Sign_Language-LSE
- sistema de anotación gramatical (formato ELAN .eaf):
‘Annotations for LSE-RADIS corpus’
<https://zenodo.org/records/10670864>
- 500 oracións analizadas mediante gramática de dependencias

UD_Spanish_Sign_Language-LSE

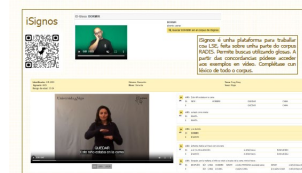
```

506 # sent_id = HC-JRV#057
507 # text = INDX CERDO UNO VENTANA cl.m(A):apoyarse-ventana
508 # gloss_en = INDX PIG ONE WINDOW cl
509 # text_es = Uno de los cerdos mira (aterrorizado) por la ventana.
510 # text_en = One of the (terrified) pigs look out the window.
511 1      INDX      _      DET      Det      _      2      det      _      Gloss=INDX
512 2      CERDO    _      NOUN    N        _      5      nsubj   _      Gloss=PIG
513 3      UNO      _      NUM     Num     _      2      nummod  _      Gloss=ONE
514 4      VENTANA  _      NOUN    N        _      5      obl      _      Gloss=WINDOW
515 5      cl.m(A):apoyarse-ventana _      VERB    V.D     _      0      root     _      Gloss=cl

```



Recursos sobre corpus para español e LSE do
Grupo GRADES da Universidade de Vigo



Corpus: Lingua inglesa e variedades

ICE-GBR

Compoñente xibaltareño do International Corpus of English
(coordinadora: Elena Seoane)

- International Corpus of English: consorcio internacional formado por 26 grupos de investigación con sede na Universidade de Zürich
- 1 millón de palabras, 20 rexistros de inglés oral e escrito (dende conversas espontáneas ata debates parlamentarios)
- parte escrita: <https://view0.webs.uvigo.es/ice-gibraltar>

ICE INTERNATIONAL CORPUS OF ENGLISH

SPOKEN (300)	Dialogues (180)	Private (100)	Face-to-face conversations (90) Phonecalls (10)
		Public (80)	Classroom Lessons (20) Broadcast Discussions (20) Broadcast Interviews (10) Parliamentary Debates (10) Legal cross-examinations (10) Business Transactions (10)
	Monologues (120)	Unscripted (70)	Spontaneous commentaries (20) Unscripted Speeches (30) Demonstrations (10) Legal Presentations (10)
		Scripted (50)	Broadcast News (20) Broadcast Talks (20) Non-broadcast Talks (10)

ICE-GBR GIBRALTAR COMPONENT

SPOKEN & WRITTEN DATA

REGISTER DISTINCTIONS

**COMPARABLE WITH
OTHER ENGLISHES**

WRITTEN (200)	Non-printed (50)	Student Writing (20)	Student Essays (10) Exam Scripts (10)
		Letters (30)	Social Letters (15) Business Letters (15)
	Printed (150)	Academic writing (40)	Humanities (10) Social Sciences (10) Natural Sciences (10) Technology (10)
		Popular writing (40)	Humanities (10) Social Sciences (10) Natural Sciences (10) Technology (10)
		Reportage (20)	Press news reports (20)
		Instructional writing (20)	Administrat Skills/hobb
		Persuasive writing (10)	Press edito
		Creative writing (20)	Novels & s

II XEIRA CLARIAH-GAL, 9 maio 2025, Santiago de Compostela
Recursos sobre corpus para a lingua inglesa
Elena Seoane, Nuria Yéñez Bouza, Javier Pérez Guerra* e grupo de investigación LVTC

Grupo de investigación
Linguistic Variation and Social Language
Universidade do Vigo
Instituto de Investigación Lingüística
(Universidade do Vigo)

Recursos sobre a ecoloxía lingüística en Xibraltar

ICE-GBR
GIBRALTAR
COMPONENT

SCIDE
CORPUS OF
SPOKEN
GIBRALTARIAN
ENGLISH

GibPress
CORPUS OF
GIBRALTARIAN
PRESS

Recursos para o estudo do inglés como L2

VICOLSE

DATA

TOOLS

Recursos para o estudo da historia da lingua inglesa

20

20

20

SGibE (en construcción)

Corpus of Spoken Gibraltar English
(coordinadora: Elena Seoane)

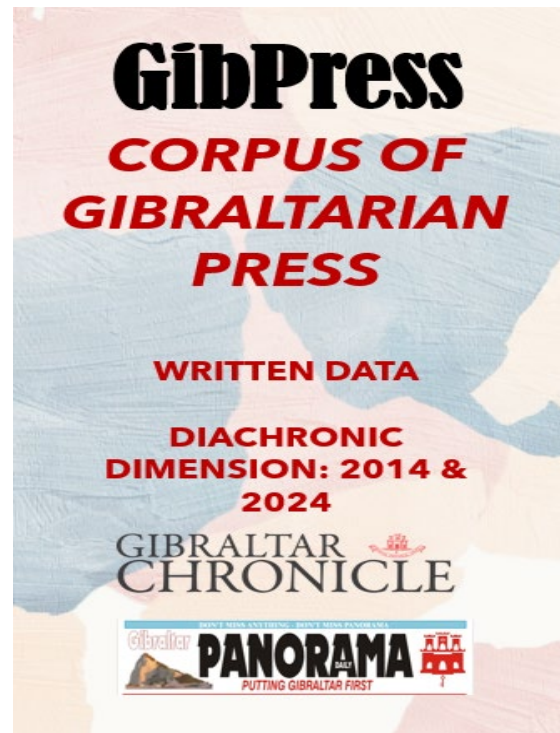
- Corpus oral de inglés de Xibraltar, espontáneo de falantes xibaltareños do podcast *Street Talk*



GibPress (en construcción)

Corpus of Gibraltar Press
(coordinadora: Elena Seoane)

- corpus de prensa xibraltareña: reportaxes e editoriais dos sitios web do *Gibraltar Chronicle* e *Panorama*, anos 2014 e 2024
- sistema automatizado para recuperar e archivar textos



ARCHER

A Representative Corpus of Historical English Registers
(coordinadora: Nuria Yáñez)

- <https://www.projects.alc.manchester.ac.uk/archer>
- corpus histórico de rexistros textuais en lingua inglesa dispoñible na interfaz CQPweb, con anotación morfolóxica e semántica
- 3,3 millóns de palabras, 4 séculos, 12 rexistros de carácter formal e informal nas variedades de inglés británico e americano



The APU Writing and Reading Corpus

(coordinadora: Nuria Yáñez e a Universidade de Liverpool)

- <https://datacat.liverpool.ac.uk/2456>
- combina a edición dixital (imaxes orixinais) e corpus lingüístico (transcricións en versión diplomática e normalizada)
- anotación morfolóxica e semántica
- corpus elaborado no marco da lingüística educativa na época de Margaret Thatcher, con materiais do ano 1979 e 1988 escritos por estudantes de nivel *Year 6* e *Year 11* en Reino Unido
- 172.000 palabras

The APU Writing and Reading Corpus

The screenshot displays the APU Writing and Reading Corpus interface. At the top, there is a search bar and a navigation menu with buttons for 'TEI Tagset', 'POS Tagset', 'Semantic Tagset', and 'Frequency lists'. Below this, a text input field shows the filename 'At home their is a rule I have to'. The main content area is divided into two panels. The left panel shows the 'Original spelling' view, displaying the text: 'At home their is a rule I have to obey, I must go to bed at 9:00 at the latest. At home I start to get changed at 8:00 and do our packlunches which takes about 5 minutes. Then I can go and sit down. But if I feel tierd I can go to bed before if I wanted too but when 9:00 comes I must, must go to bed.' The right panel shows the 'POS tagging' view, displaying the same text with various tags and symbols, such as 'At_Z5 home_H4/H1c there_Z5 is_A3+ a_Z5 rule_G2.1 |_Z8mf have_S6+[i1.2.1 to_S6+[i1.2.2 obey_S7.1- ,_PUNC |_Z8mf must_S6+ go_B1[i2.3.1 to_B1[i2.3.2 bed_B1[i2.3.3 at_Z5 9:00_N1 at_Z5 the_Z5 latest_T3-- At_Z5 home_H4/H1c |_Z8mf start_T2+ to_Z5 get_A9+ changed_A2.1+ at_Z5 8:00_N1 and_Z5 do_A1.1.1 our_Z8 pack_O2 lunches_F1 which_Z8 takes_A9+ about_A13.4 5_N1 minutes_T1.3 ._PUNC Then_N4 |_Z8mf can_A7+ go_M1 and_Z5 sit_M8[i3.2.1 down_M8[i3.2.2 ._PUNC But_Z5 if_Z7 |_Z8mf feel_X2.1 tired_B1 |_Z8mf can_A7+ go_B1[i4.3.1 to_B1[i4.3.2 bed_B1[i4.3.3 before_Z5 if_Z7 |_Z8mf wanted_X7+ to_Z5 but_Z5 when_Z5 9:00_N1 comes_M1 |_Z8mf must_S6+ ,_PUNC must_S6+ go_B1[i5.3.1 to_B1[i5.3.2 bed_B1[i5.3.3 . PUNC'. The 'POS tagging' button is highlighted with a red box. On the far right, there is a vertical sidebar with labels 'MC', 's_VVZ', and 'I'.

Search

TEI Tagset POS Tagset Semantic Tagset Frequency lists

Original spelling POS tagging Semantic tagging Metadata description

Filename

At home their is a rule I have to

TEI Tagset POS Tagset Semantic Tagset Frequency lists

Original spelling POS tagging Semantic tagging Metadata description

At home their is a rule I have to obey, I must go to bed at 9:00 at the latest

At home I start to get changed at 8:00 and do our packlunches which takes about 5 minutes. Then I can go and sit down. But if I feel tierd I can go to bed before if I wanted too but when 9:00 comes I must, must go to bed.

At_Z5 home_H4/H1c there_Z5 is_A3+ a_Z5 rule_G2.1 |_Z8mf have_S6+[i1.2.1 to_S6+[i1.2.2 obey_S7.1- ,_PUNC |_Z8mf must_S6+ go_B1[i2.3.1 to_B1[i2.3.2 bed_B1[i2.3.3 at_Z5 9:00_N1 at_Z5 the_Z5 latest_T3-- At_Z5 home_H4/H1c |_Z8mf start_T2+ to_Z5 get_A9+ changed_A2.1+ at_Z5 8:00_N1 and_Z5 do_A1.1.1 our_Z8 pack_O2 lunches_F1 which_Z8 takes_A9+ about_A13.4 5_N1 minutes_T1.3 ._PUNC Then_N4 |_Z8mf can_A7+ go_M1 and_Z5 sit_M8[i3.2.1 down_M8[i3.2.2 ._PUNC But_Z5 if_Z7 |_Z8mf feel_X2.1 tired_B1 |_Z8mf can_A7+ go_B1[i4.3.1 to_B1[i4.3.2 bed_B1[i4.3.3 before_Z5 if_Z7 |_Z8mf wanted_X7+ to_Z5 but_Z5 when_Z5 9:00_N1 comes_M1 |_Z8mf must_S6+ ,_PUNC must_S6+ go_B1[i5.3.1 to_B1[i5.3.2 bed_B1[i5.3.3 . PUNC

MC

s_VVZ

I

VICOLSE

Vigo Corpus of Learner Spoken English
(coordinadores: Beatriz Tizón, Javier Pérez)

- aprox. 100.000 palabras de inglés oral narrativo e argumentativo, producido por estudantes universitarios de inglés como lingua estranxeira na UVigo
- información sociolóxico-cultural das persoas participantes
- plataforma (en construción) facilita acceso aos audios e ás transcricións, seguindo as convencións de LINDSEI

VICOLSE

<stdnt 01>

[Task 1]

Little Red Riding Hood *ah* continued ahead (tch) and *ah* through the forest and then he arrived to a little house a wooden house (tch) which was her grandma's one *ahm* once she was there she knocked on the door (tch) and *ahm* her grandmother's voice said *eh* come in . she went in and *mm* she (..) kissed her grandmother and then she said *eh* here you have a gift some things I've bought for you like honey and *ah* cakes and whatever *ah* the grandmother was very happy and *ah* yeah *ah* the problem was that she was in bed cause she was quite ill and *ah* Little Red Riding Hood *ah* (..) started to ask her why was she so so ill in bed (tch) *ahm* (..) then the the grandmother said *aom* I'm very ill because *m* I have a cough (tch) and the the little *child child *started saying *mm* yeah I *?? I think I'm starting to notice some kind of strange features on on your face like for instance you're your eyes they they *really look bigger

VIGO CORPUS OF LEARNER SPOKEN ENGLISH

VICOLSE

The Vigo Corpus of Learner Spoken English, VICOLSE, is a collection of real samples of spoken English produced by Galician students of English as a foreign language at the University of Vigo.

VIGO CORPUS OF LEARNER SPOKEN ENGLISH

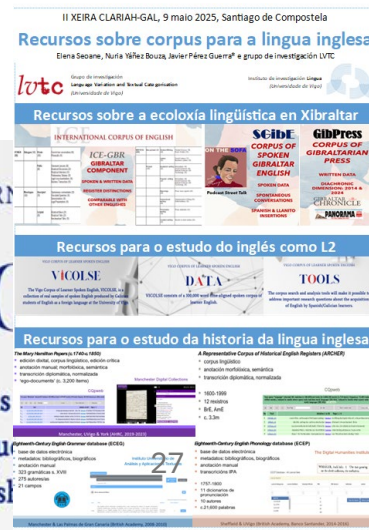
DATA

VICOLSE consists of a 100,000 word time-aligned spoken corpus of learner English.

VIGO CORPUS OF LEARNER SPOKEN ENGLISH

TOOL

The corpus search and analysis address important research questions of English by Spanish



Plataformas e portais de consulta

RILG

Recursos Integrados da Lingua Galega

(coordinado polo grupo TALG, Alexandre Fernández)

- <https://ilg.usc.gal/rilg>
- portal de recursos textuais e léxicos de tecnoloxía lingüística da lingua galega -- dicionarios, corpus de lingua galega e corpus paralelos de tradución

RILG



Portal RILG



Inicio



Información



Referencias



Recursos ▾



· Consultar todo o RILG

Insira un termo

Vai

Seleccionar: ☒ RILG

Procurar: ☒ palabra ☐ lema

Todas as categorías ▾

Todas as categorías

Adxectivo

Numeral

Adverbio

Determinante

Nome

Verbo

Pronome

Demostrativo

Posesivo

· Ou seleccionar un recurso

Dicionarios galegos

Corpus de lingua galega

Corpus paralelos de tradución



INSTITUTO DA LINGUA GALEGA (ILG)



NEOTECA

(coordinado polo grupo TALG, Alexandre Fernández)

- <https://ilg.usc.gal/neoteca>
- banco de datos sobre neoloxía léxica do galego
- neoloxismos documentados na prensa galega dende 1997
- información contextual, datación, fonte, categoría gramatical e procedemento de formación

NEOTECA



Neoteca



Inicio



Información



Referencias



Recursos ▾



· Consultar a Neoteca

Vai

Adverbio ▾

Todos os tipos ▾

Todos os tipos

Abreviación

Composición culta

Composición híbrida

Composición léxica

Composición sintagmática

Parasíntese

Prefixación

Préstamo

Simple

RERCOR v3.1.2 2018-2025

Recursos sobre enfermedades raras

(coordinadoras: Elena Sánchez e Tamara Varela)

- <http://www.rercor.org>
- portal de recursos lingüísticos multilingües do ámbito biomédico sobre un amplo grupo de enfermedades raras
- base: corpus de máis de 8 millóns de palabras, comparables e paralelos en francés, inglés e español:
 - MYOCOR 2.0 (Corpus de enfermedades neuromusculares)
 - EMCOR (Corpus de enfermedades metabólicas)
 - ERCOR (Corpus de guías de práctica clínica sobre ER)
 - ORPHACOR (Corpus de medicamentos orfos)
 - PRODAPCOR (Corpus de produtos de apoio)
 - TERAPCOR (Corpus de terapia ocupacional)

RERCOR v3.1.2 2018-2025

Busca

Resultado

1	TERAPCOR / en / 2019	 mental health care emphasizes that recovery is more complex than the absence of symptoms, proposing that recovery from mental illness occurs when one can live a "hopeful, satisfying and contributing life, even with ongoing limitations" ()
2	ORPHACOR / en / 2019	Consultation with a physician is recommended during illness as blood phenylalanine levels may increase.
3	ORPHACOR / en / 2019	 in the blood (Hyperbilirubinemia), abnormal results of blood test (Hypomagnesaemia and hyponatraemia), influenza like illness , chills, bone marrow cancer (myelodysplastic syndrome), blood transfusion.
4	ORPHACOR / en / 2019	 you only. Do not pass it to others. It may harm them, even if their signs of illness are the same as yours.
5	ORPHACOR / en / 2019	 only. Do not pass it on to others. It may harm them, even if their signs of illness are the same as your child's.
6	TERAPCOR / en / 2004	You know, Just because you've got an illness doesn't mean to say we're confined to, say, sitting and reading a book.
7	ORPHACOR / en / 2019	 only. Do not pass it on to others. It may harm them, even if their signs of illness are the same as yours.
8	ORPHACOR / en / 2019	- Flu like symptoms, such as high temperature, aches and chills (influenza-like illness)
9	ORPHACOR / en / 2019	- psychotic behaviour (mental illness with hallucinations, difficulty in thinking clearly and understanding reality, anxiety), initial difficulty in sleeping
10	ORPHACOR / en / 2018	Use in intercurrent illness
11	ORPHACOR / en / 2019	 required respiratory intervention >6 hours/day continuously for =7 days, all of whom initiated ventilatory support during an acute reversible illness
12	TERAPCOR / en / 2017	- Disease and illness
13	ORPHACOR / en / 2019	 only. Do not pass it on to others. It may harm them, even if their signs of illness are the same as yours.
14	TERAPCOR / en / 2017	- work with people of all ages with a variety of conditions caused by injury or illness , psychological or emotional difficulties, developmental delay or the effects of aging
15	ORPHACOR / en / 2019	 only. Do not pass it on to others. It may harm them, even if their signs of illness are the same as yours.

Palabras ortográficas

illness

Software

NER Buddy

Software de análise automática, semiautomática e asistida de subtítulos intralingüísticos en directo con IA
(coordinadores: Pablo Romero, Luis Alonso, grupo GALMA)

- [*subtitulado intralingüístico en directo*: subtítulos na mesma lingua ca o orixinal]
- Modelo NER: sistema complexo para a avaliación da calidade de subtítulos en directo máis utilizado no mundo. Factores:
 - impacto comunicativo
 - lexibilidade



NER Buddy

Software de análise automática, semiautomática e assistida de subtítulos intralingüísticos em directo com IA
(coordinadores: Pablo Romero, Luis Alonso, grupo GALMA)

- Empregado e testado polo Parlamento Europeo, Netflix e cadeas públicas e empresas internacionais, e testado tanto con modelos comerciais (GPT) como con modelos *open-source* (Openchat, Llama, Gemma, Mixtral)



NETFLIX

NER Buddy

NER accuracy: 98.74% Reduction rate: 0.42% [Show details](#)

transcript.txt

1	00:00:00	Water is a collective good .	
2	00:00:04	It should therefore also be our greatest source of cooperation .	
3	00:00:10	It's in this spirit that the world is meeting at the United Nations in New	
4	00:00:15	In the 20th century , signed and in fact	
5	00:00:24	more cooperation over	
6	00:00:28	But with climate press water becomes an inc	
7	00:00:36	On the path we are o will suffer from acute	
8	00:00:42	by 2030 .	
9	00:00:45	Europe is facing the v	
10	00:00:49	Freshwater res	
11		And millions are globa	

Drag and drop each word on the corresponding error.

Add Add Add Add Add

1 - Serious

Recognition Edition

Until this summer , for example , we are implementing a water reuse for agricultural

2 - Standard

Recognition Edition

Until this summer , for example , we are implementing a water reuse regulation for agricultural

3 - Minor

regulation Recognition Edition

NER accuracy		Reduction rate	
98.85%		0.42%	

Recognition errors		Edition errors		Unevaluated	
Unevaluated	0	Score	0	Unevaluated	0
Correct	2	Score	0	Correct	0
Minor	12	Score	3	Minor	0
Standard	1	Score	0.5	Standard	0
Serious	2	Score	2	Serious	0
Total	17	Score	5.5	Total	0

NER Buddy
Software asistido por IA
para a avaliación da calidade de subtítulos

Pablo Romero Fresco e Luis Alonso Radigalupe

1. Mais da metade da poboación mundial utiliza subtítulos intralingüísticos.

2. O modelo NER, creado por Pablo Romero-Fresco (lingua - UVigo), é o método de avaliación da calidade do subtítulo máis utilizado no mundo.

3. NER Buddy é o único software que aplica automaticamente o modelo NER coa axuda de IA.

38

Demro do NER Buddy, que está sendo testado por:

NETFLIX

Proxectos e recursos do **Instituto de investigación LINGUA** da Universidade de Vigo



Javier Pérez Guerra
jperez@uvigo.gal

II Xeira CLARIAH-GAL, 9 maio 2025

