Que é?

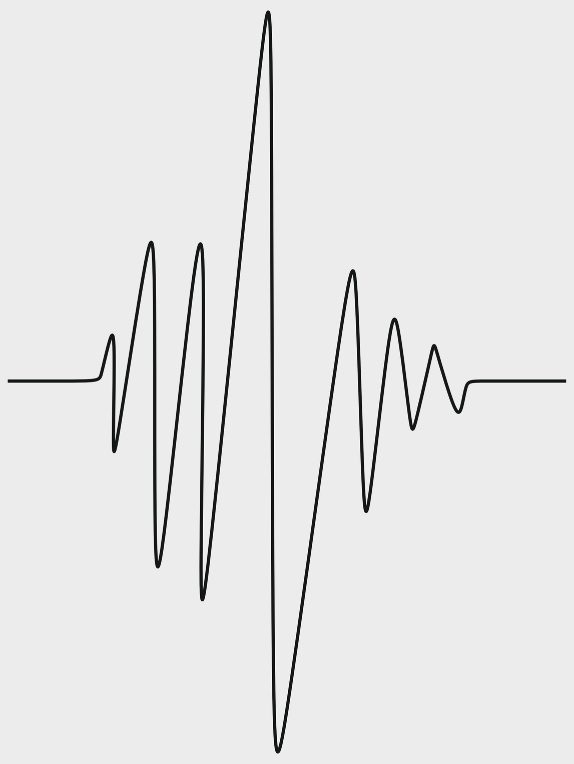
- CORILGA é un corpus oral con transcricións ortográficas aliñadas co arquivo sonoro, nunha correlación temporal texto-voz.
- As transcricións están lematizadas e anotadas en diferentes niveis (fonético, morfolóxico, textual e temático).
- Esta aplicación permite realizar buscas combinadas de maneira funcional e inmediata, sen restricións de uso.
- O corpus, de acceso libre, conta actualmente con 60 horas de gravacións transcritas en fase de revisión.

CORILGA

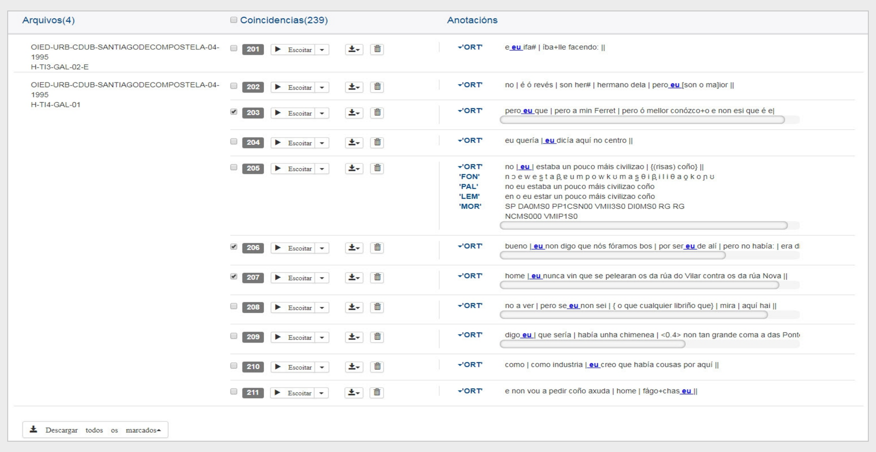
Corpus Oral Informatizado da Lingua Galega

Obxectivos

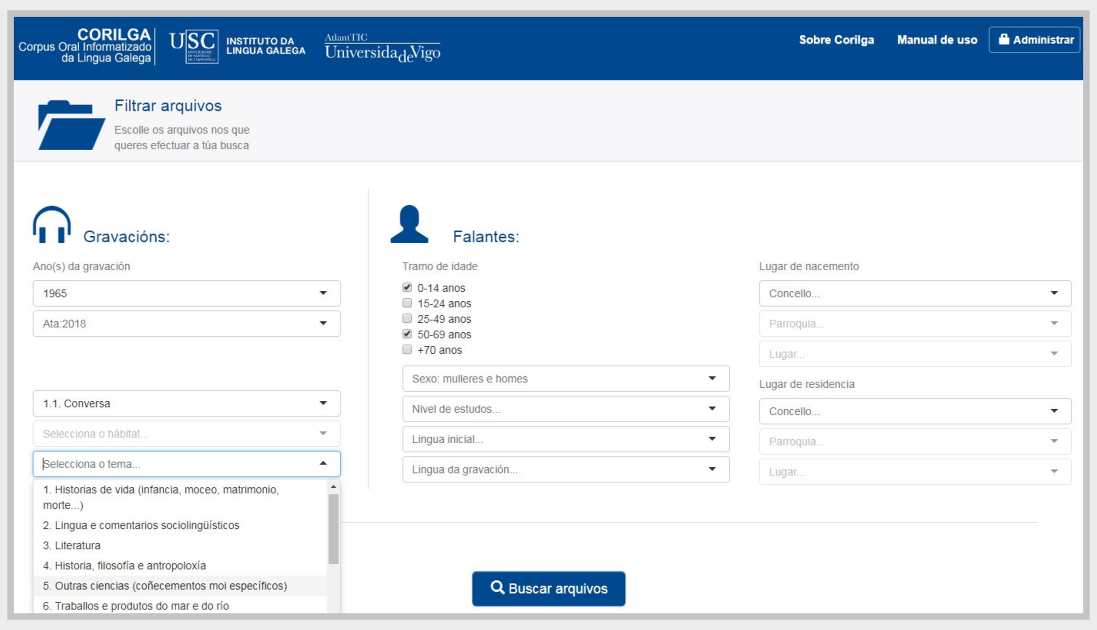
- Dispor dun corpus de gravación e das correspondentes transcricións que permitan estudar a oralidade da lingua galega contemporánea, desde a década de 1960 até a actualidade.
- Reflectir a variación diafásica -con atención a diferentes rexistros e xéneros textuais-, diastrática, diacrónica e diatópica.
- Facilitar a análise do cambio lingüístico en tempo real e en tempo aparente.
- Favorecer a elaboración de estudos humanísticos interdisciplinares, nomeadamente no ámbito lingüístico (análise do discurso, pragmática, morfoloxía, fonética, dialectoloxía etc.).
- Colaborar coa creación e desenvolvemento de tecnoloxías da fala.



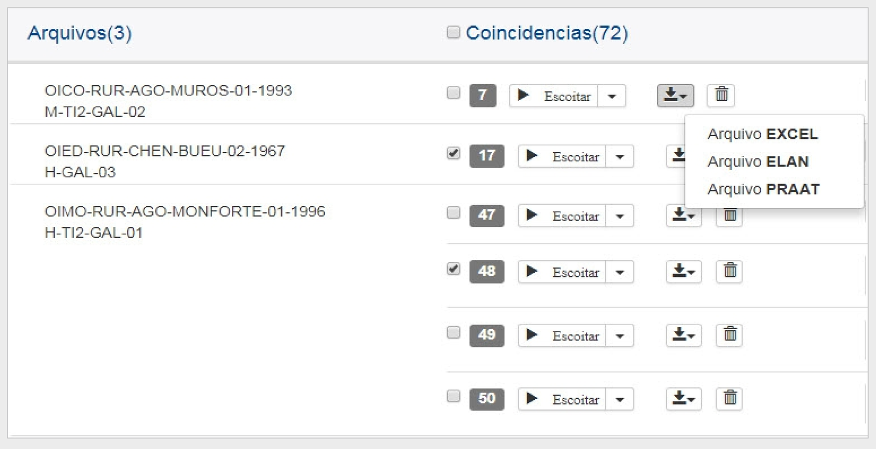
Imaxe 2. Patróns de busca ortográfica, morfolóxica, fonética e lemática.



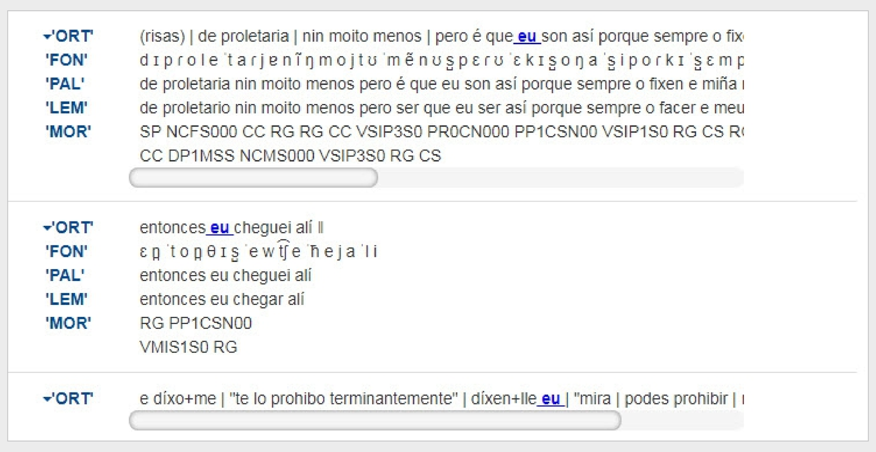
Imaxe 3. Exemplo de resultados da busca.



Imaxe 1. Parámetros principais de busca: data da gravación, tipo de texto e características do falante.



Imaxe 5. Datos da gravación e falante e opcións de escoita e descarga.



Imaxe 4. Detalle de resultados expandidos.

Daniel Amarelo Montero[1]
José Manuel Dopazo Entenza[1]
Tamara Rial Montes[1]
Reyes Rodríguez Rodríguez[1]

Xosé Luís Regueira[1]
Carmen García Mateo[2]
Elisa Fernández Rei[1]
Eduardo Rodríguez Banga[2]
Antonio Cardenal López[2]

[1] Instituto da Lingua Galega - Universidade de Santiago de Compostela
[2] Universidade de Vigo

UNIÓN EUROPEA
Fondo Europeo de Desarrollo Regional