

ECPC VUELVE SU MIRADA HACIA LOS REFUGIADOS

Miembros

IP: María Calzada

Equipo de investigación: M^a Jesús Blasco, Beatriz Méndez, Celia Rico, M^a Mar Sánchez.

Equipo de trabajo: José Aleixandre, Vaclav Brezina, Carla Fonte, Àngela Francès, Fernando Gabarrón, Andrew Hardie, Cristian Hernández, Paula Porcar, Elena Ríos

Metodología

1. Recopilación de archivos HTML
2. Gestión de archivos HTML
3. Traducción de HTML a XML
4. Validación del proceso de traducción
5. Alineación PE-EN y PE-ES



- Método semiautomático: *scripts* de Perl sobre el uso de expresiones regulares
- Criterios cualitativos y cuantitativos

XML

```
METADATA {structure,
{speaker,
{name, status,
country_of_origin,
contituency, gender, post,
date/place_of_birth, party
}text
{speech, language, omit,
```

Origen

TEC: estudios traductológicos de corpus – Mona Baker

<https://genealogiesofknowledge.net/translational-english-corpus-tec/>

ENPC: estudios de lingüística contrastiva encabezados por Stig Johansson y Knut Hofland

<http://www.helsinki.fi/varien/CoRD/corpora/ENPC/>

RE-CRI (20-24)

- Ampliación de los corpus actuales
- Incorporación de debates orales (frente a actas escritas) transcritos por el grupo de forma semiautomatizada
- Creación de herramientas de exploración textual

PARA EL ESTUDIO DE LA REPRESENTACIÓN DE LOS REFUGIADOS



ECPC (European Comparable and Parallel Corpora)

desde 2005

Equipo multidisciplinar, plurinacional y multilingüe

Proyecto en desarrollo actualmente: RE-CRI (20-24)

(Representaciones originales, traducidas e interpretadas de la(s) crisis de refugiados: triangulación metodológica desde el análisis del discurso basado en corpus PID2019-108866RB-100)

Deliverables

Corpus

- Congreso de los Diputados (CD) (hasta 2014): aprox. 38.000.000 palabras
- House of Commons (HC) (hasta 2014): aprox. 100.000.000 palabras
- Parlamento Europeo (EP) (2004-2011):
 - EP-ES: aprox. 28.200.000 palabras
 - EP-EN: aprox. 28.900.000 palabras

Herramientas

- ConcECPC
- Glossa_ECPC
- ECPC-Web (corpus con metadatos)

Objetivo

Estudiar las tendencias lingüísticas del discurso parlamentario en el Congreso de los Diputados (CD), la Cámara de los Comunes (HC) y el Parlamento Europeo (en sus versiones en inglés y español, con discursos originales y traducidos)

