

Carolina Julià Luna (Universidad Nacional de Educación a Distancia)
 Soraya Almansa Ibáñez (Universidad Nacional de Educación a Distancia)
 Borja Alonso Pascua (Universidad de Salamanca)
 Joseph García Rodríguez (Universidad Nacional de Educación a Distancia)
 Natalia Terrón Vinagre (Universitat Autònoma de Barcelona)

Assumpció Rost Bagudanch (Universitat Autònoma de Barcelona)
 Alba Agüete Cajiao (Universidad de Salamanca)
 Beatriz Blecua Falgueras (Universitat de Girona)
 César Gutiérrez (Wake Forest University)
 M.ª Jesús Machuca Ayuso (Universitat Autònoma de Barcelona)

CARACTERÍSTICAS DE CORPAT

- Es una base de datos MySQL 5.6 en código abierto.
- Recoge información de **987** puntos geográficos de España.
- Contiene datos de **9 atlas lingüísticos** regionales del español europeo.

1. Estudio, selección y recogida de datos para el análisis y la informatización en el corpus en línea.

2. Perfeccionamiento, ampliación y corrección de las funcionalidades de CORPAT.

3. Análisis, estudio y tratamiento de la información **fonética** en los mapas.

4. Análisis, estudio y tratamiento de la información **léxico-semántica** en los mapas.

OBJETIVOS GENERALES

MARCO CIENTÍFICO-ECONÓMICO



Corpus digital para la preservación y el estudio del patrimonio lingüístico del español (CORPAT-PEPLES)



Lengua oral y cambio lingüístico en los atlas españoles (CORPAT-LOCALES)

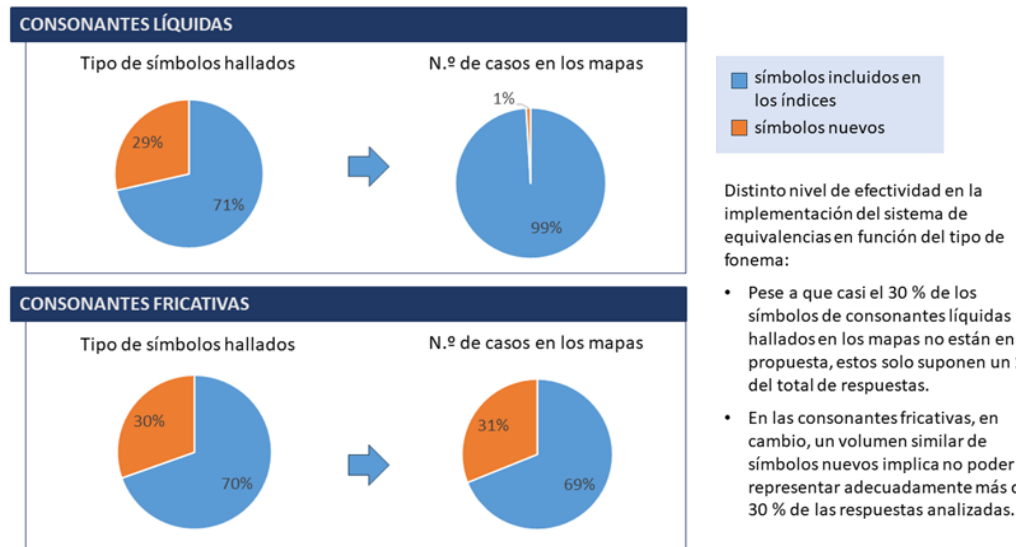
Julià (2021a, 2021b, 2023a)

FONÉTICA

La fonética constituye una parte fundamental en los atlas lingüísticos. Su digitalización permite establecer comparaciones entre distintas áreas y extraer datos fonéticos. Para ello es imprescindible contar con una transcripción unificada en AFI.

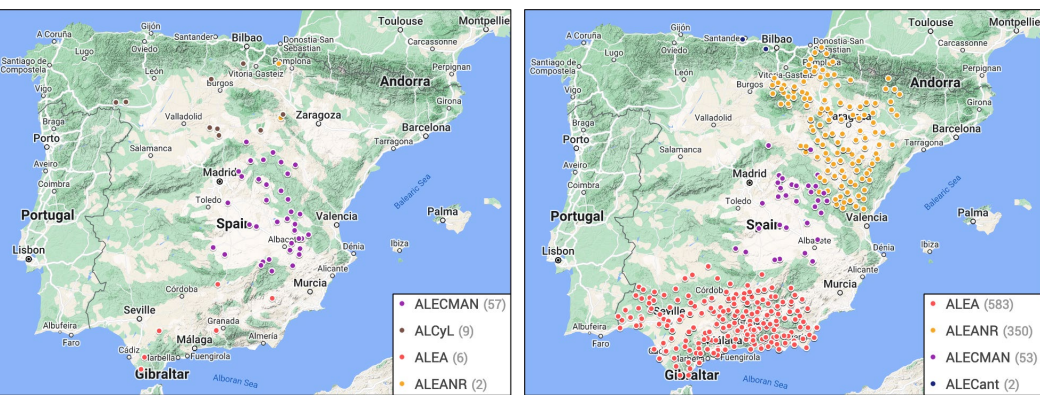
- Homogeneización de los símbolos fonéticos y propuesta de adaptación a AFI → mecanismos de validación.

Mecanismos para la validación de la transcripción fonética en AFI y ejemplos



Marrero y Rost (2024)

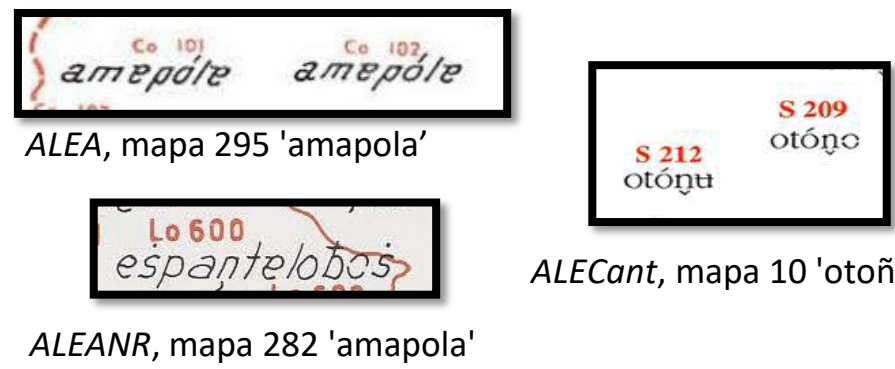
[aj̟̞̚] > [ej̟̞̚] (n=74) / [ej̟̞̚] > [aj̟̞̚] (n=988)



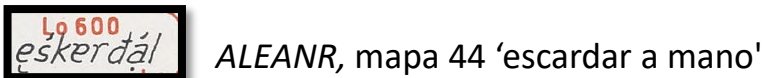
- Identificación de procesos fonéticos no descritos ampliamente en la bibliografía.

Agüete (en prensa), Blecua et al. (2024), Gutiérrez (2023a) y Marrero y Rost (2024)

Vocales "relajadas"



Lambdacismo en la Rioja Baja

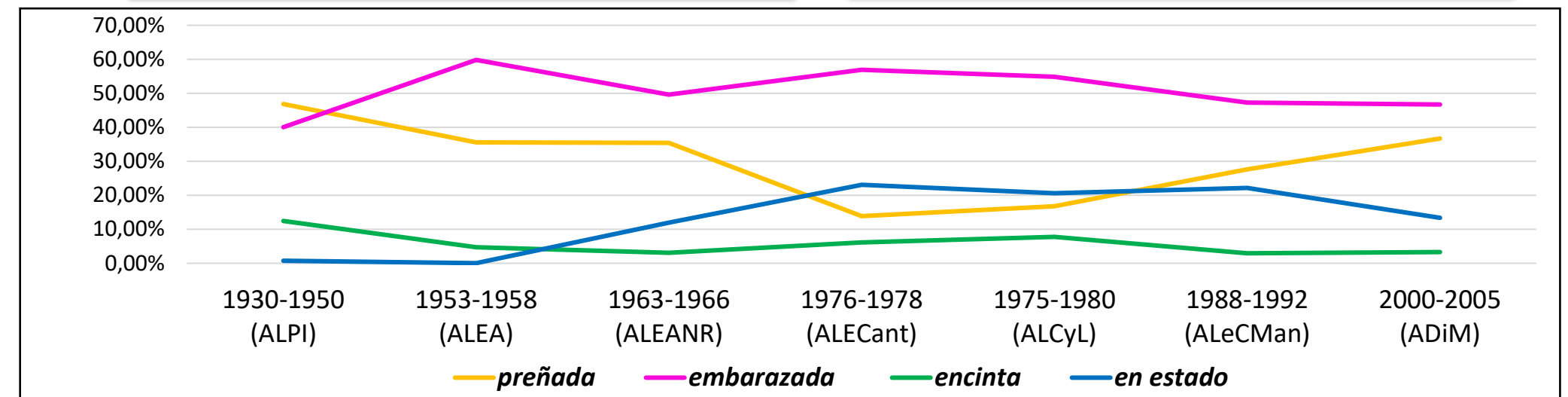
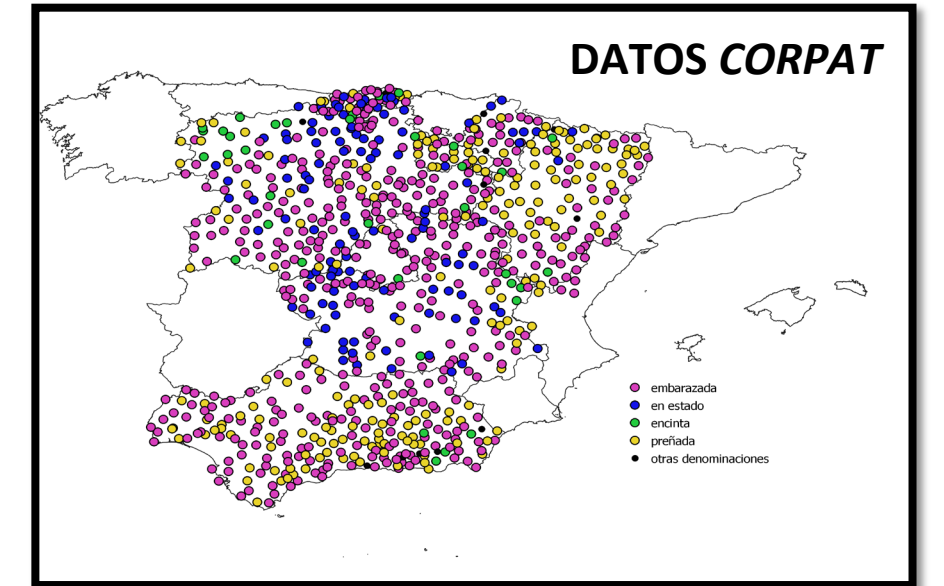
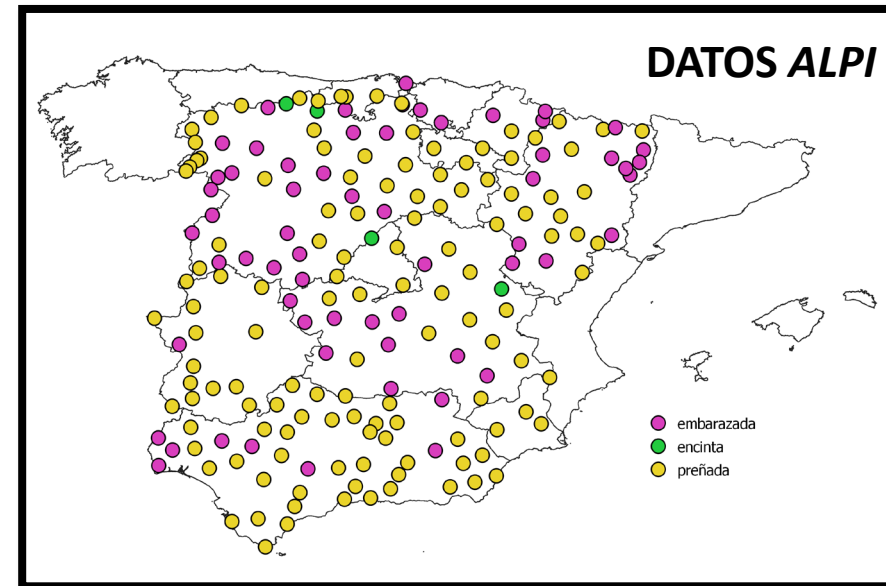


LÉXICO

El léxico constituye el área lingüística con mayor representación en los atlas regionales: supone entre el 75 % y el 100 % de los datos que contienen. La digitalización permite avanzar en su conocimiento desde diferentes puntos de vista.

- Procesos y motivaciones del cambio léxico-semántico en el espacio y en el tiempo.
- Mortandad y vitalidad léxicas.
- Innovaciones léxicas.
- Productividad de los procesos de formación de palabras.
- Distribución y evolución de las áreas léxicas.
- Cambio léxico en relación con la evolución de la sociedad.
- Procesos de hibridación y transferencia léxica.
- Etimología léxica.

Variación designativa del concepto 'estar encinta'



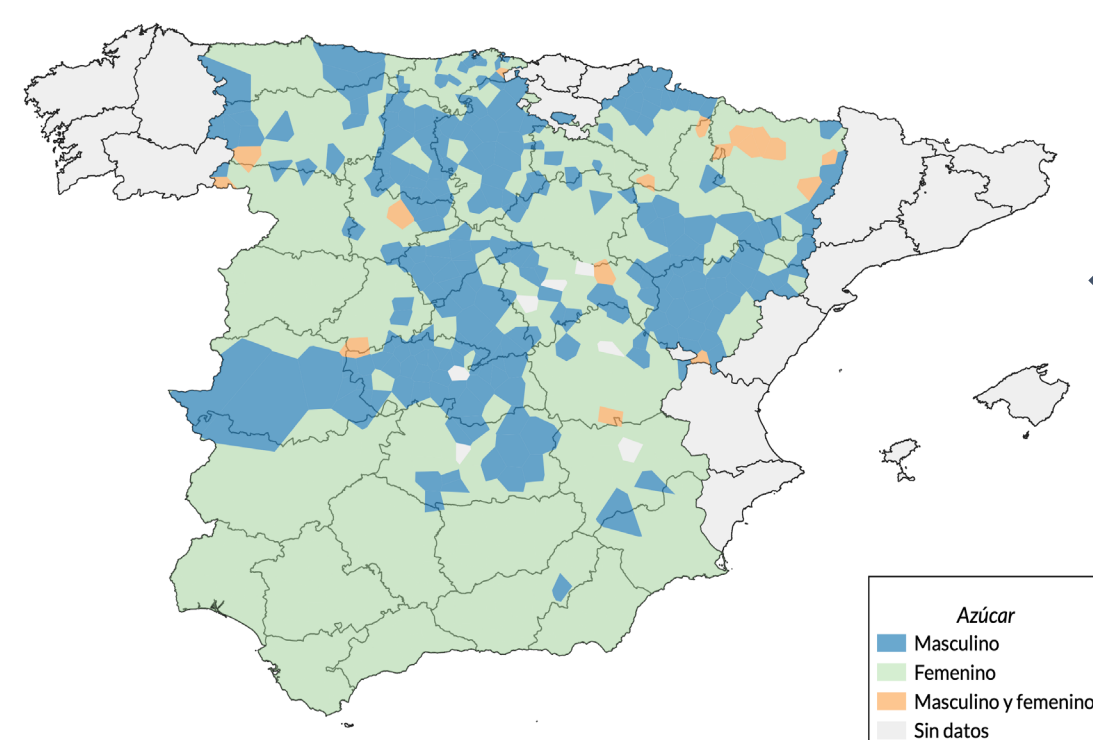
Gutiérrez (2023b), Julià (2022, 2023b, 2024, en prensa), Julià, Rost y Quijada (2022) y Julià y Rost (en prensa)

MORFOSINTAXIS

Al igual que ocurre con buena parte de los estudios dialectológicos del siglo XX, los aparatos morfosintácticos han ocupado un lugar secundario en los atlas lingüísticos del español.

- La morfología recibe más atención que la sintaxis, que suele estar infrarrepresentada.
- No está claro cuándo un fenómeno se asigna a la sintaxis o a la fraseología; a veces hay clasificaciones erróneas.
- Son pocos los fenómenos que aparecen en todos los atlas, lo que dificulta la detección de patrones.
- En ocasiones, la misma cuestión se rastrea con ítems distintos según el atlas y esto complica la interpretación geolectal.

Variación genérica del sustantivo azúcar a partir de CORPAT



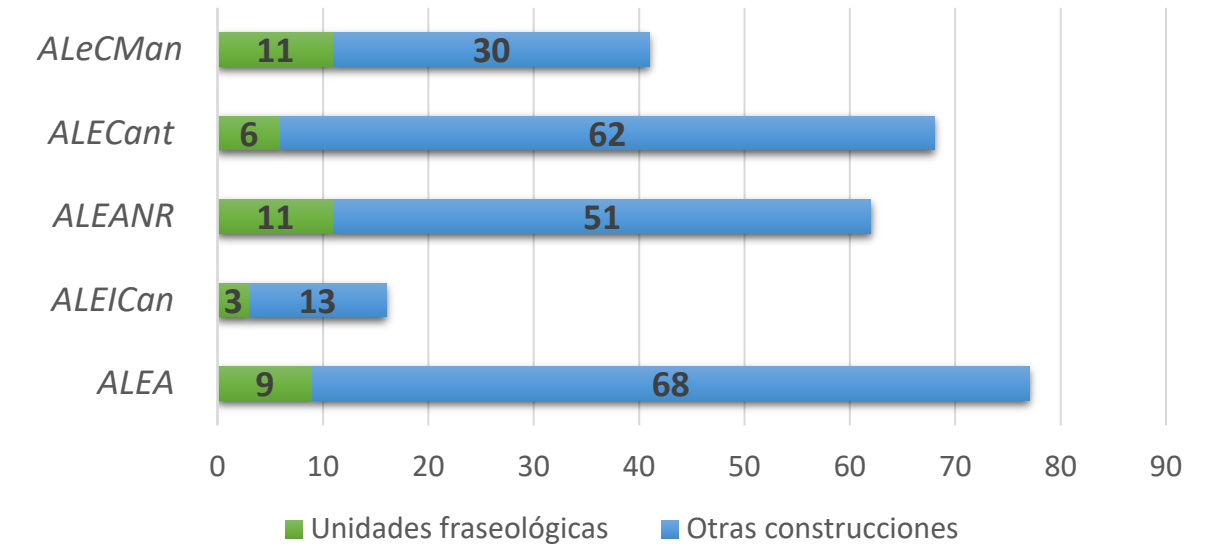
Alonso (2023a, 2023b, 2023c, 2024)

FRASEOLOGÍA

El proyecto no tiene como objetivo el estudio de la fraseología. Un análisis general revela su escasa representatividad y los problemas metodológicos que hay en la recolección de datos.

- 4 de los 7 atlas incluyen un apartado sobre sintaxis y fraseología.
- La fraseología representa el 1 % del total de los datos de los atlas.
- Es el área lingüística con un menor número de mapas en la geografía regional.
- Se observan inconsistencias en la clasificación del material fraseológico, sobre todo en los atlas publicados entre 1950-1990.

Apartado de sintaxis y fraseología en los atlas



Terrón (2023) y García y Julià (2023)

PRIMEROS RESULTADOS

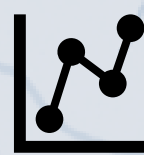
CORPAT contiene los siguientes datos (julio de 2024):

540 mapas

195 conceptos

64500 registros

POSIBLES FUTURAS LÍNEAS DE TRABAJO



Dialectometrización de los datos.



Cartografiado de la información por concepto.



Incorporación de los audios de algunos atlas.

BIBLIOGRAFÍA

