

README file

Dataset title: **“Bibliografía esencial sobre sociedad, lengua e identidad en contexto migratorio, con enfoque en la comunidad hispana”**

Dataset author and contact for information: **Lisa Gozzi** (Università di Bologna, Dipartimento di Lingue, Letterature e Culture Moderne – LILEC), <https://orcid.org/0009-0004-3161-1100>, lisa.gozzi3@unibo.it

Dataset License: this dataset is distributed under license **Creative Commons CC BY – Attribution 4.0**.

Publication year: **2025**

Dataset content

The dataset consists of the following documents:

- 10 files containing tabled text data in .xlsx and .csv format
 - “1_Sociolinguistica_urbana.csv”
 - “1_Sociolinguistica_urbana.xlsx”
 - “2_Dinamicas_migratorias.csv”
 - “2_Dinamicas_migratorias.xlsx”
 - “3_Migraciones_latinas_mundo.csv”
 - “3_Migraciones_latinas_mundo.xlsx”
 - “4_Migraciones_latinas_italia.csv”
 - “4_Migraciones_latinas_italia.xlsx”
 - “5_Recursos.csv”
 - “5_Recursos.xlsx”
- 1 file in Spanish containing the introduction to the dataset in .pdf format
 - “Bibliografia_migraciones_social_lengua_espanol_introduccion.pdf”
- 1 README file in Italian in .pdf format
 - “README_it_Bibliografia_migraciones_social_lengua_espanol.pdf”
- 1 README file in English in .pdf format
 - “README_eng_Bibliografia_migraciones_social_lengua_espanol.pdf”

Dataset documentation

Abstract

The following dataset collects and organizes bibliographical data in three different languages (Italian, Spanish, English) related to sociolinguistics in migration, focusing especially on Spanish-speaking communities' mobility, as it belongs to the workflow of the post-doc project titled *Spagnolo e italiano in contatto: repertori linguistici, dinamiche migratorie e mobilità internazionale a Bologna* ("Spanish and Italian in contact: linguistic repertoires, migration dynamics and international mobility in Bologna" – LILEC Department, Università di Bologna).

The corpus structure follows a path starting from a broader perspective on urban sociolinguistics and focusing next on the studies related to migration dynamics. Afterwards, it takes more into consideration the community researched through the project, examining the Hispanic group mobility firstly on global scale and secondly on the Italian territory. As far as themes are concerned, the dataset considers specific dynamics regarding linguistic contact phenomena, identity redefinition and communicative practices in plurilingual urban environments.

The bibliographical references were collected through different online academic databases, such as Google Scholar, Dialnet, Scielo, Redalyc and e_Buah. Following the databases search and collection of relevant references, these were classified into four thematic categories previously mentioned, that is 1) urban sociolinguistics, 2) migration dynamics, 3) global Hispanic migration and 4) Hispanic migration in Italy. Finally, a fifth file is included, where references to corpora and databases available online dedicated to linguistic contact in migration contexts are presented.

Files' features

Files "**1_Sociolinguistica_urbana.csv/.xlsx**", "**2_Dinamicas_migratorias.csv/.xlsx**", "**3_Migraciones_latinas_mundo.csv/.xlsx**" and "**4_Migraciones_latinas_italia.csv/.xlsx**" are organized into 15 columns, in Spanish, to report the following pieces of information:

1. Author (*Autor*).
2. Year (*Año*).
3. Title (*Título*).
4. Contribution type (*Tipo de contribución*) – monograph, curatorship, chapter in book, article, thesis/dissertation, conference proceedings.
5. Book authors or editors (*Autores/editores del volumen*) – compiled where necessary (chapter in book, conference proceedings).
6. Book title (*Título del volumen*) – compiled where necessary (chapter in book, conference proceedings).
7. City (*Ciudad*) – compiled where necessary (monograph, coauthorship, chapter in book, conference proceedings, thesis/dissertation).
8. Publisher (*Editorial*) – compiled where necessary (monograph, coauthorship, chapter in book, conference proceedings, thesis/dissertation).
9. Journal (*Revista*) – compiled where necessary (article).

10. Volume and/or issue (*Volumen/número*) – compiled where necessary (article) and only if it was specified in the bibliographical reference considered. In those cases where only of the two values considered was present, the dataset documents just that one.
11. Pages (*Páginas*) – compiled where necessary (chapter in book, conference proceedings, article).
12. Language (*Lengua*) – Italian (IT), Spanish (ES), English (EN).
13. Available online (*Disponibilidad en línea*).
14. DOI – if present.
15. Link (*Enlace*) – if available online.

The file “**5_Recursos.csv/.xlsx**” is organized into 3 columns, in Spanish, to present the following pieces of information:

1. Database or resource’s name (*Nombre del recurso*)
2. Link (*Enlace*)
3. Resource’s related publications (*Publicaciones relacionadas*) – these are included only in the case of specific databases, where they represent a way to access the database itself, as it was verified during the compilation of this dataset.

Finally, the .pdf file “**Bibliografia_migraciones_social_lengua_espanol_introduccion.pdf**”, in Spanish, constitutes an introduction to the dataset. It includes a general presentation about the total amount of references documented in the corpus, as well as the number of bibliographical references per thematic group. In addition, it provides a graphic representation of the references included according to the publication year.

.csv file specifics

In the case of those files in .csv format, the field delimiter is the semicolon (“;”) and the encoding character is UTF-8.