



Spanish Learner Language Oral Corpus

Funded by the UK
Economic and Social Re-
search Council (ESRC)
(RES-000-23-1609 and
RES-062-23-1075).



SPLLOC members



Maria J. Arche



Laura Dominguez



Ros Mitchell



Flo Myles



Nicole Tracy-Ventura

Contact us:

N.Tracy-Ventura@soton.ac.uk

- **SPLLOC** is a new corpus of oral L2 Spanish now freely available to the research community through the project website.
- The data have been collected from classroom learners in UK schools and universities using a series of specially designed elicitation tasks.
- Data come from **120 learners** (L1 English) at each of 3 levels: **beginners** (aged 13-15), **intermediate** (aged 17-18), and **advanced** students (fourth year undergraduates aged 20-23 majoring in Spanish).
- Thirty native speakers were also recorded undertaking the same tasks.
- The resulting database contains 461 digital soundfiles (learner recordings - 269,262 words) and 100 native speaker recordings (64,229 words), the corresponding transcripts and XML files also included. Some files also have an extra layer of tagging which identifies parts of speech.
- Transcripts of all recordings in **CHILDES** format.
- The database contains two sets of data: SPLLOC I and SPLLOC II

SPLLOC I:

- 40 hours of oral semi-spontaneous data
- Elicitation tasks used:
 - Impersonal Narrative (Loch Ness)
 - Pictured-based guided interview (personal narrative)
 - Paired-discussion task
 - Clitic production task
- 221 recordings from classroom learners from three age groups
- 50 native speaker recordings

SPLLOC II:

- 30 hours of oral semi-spontaneous data
- New tasks designed to elicit the **past tense** (descriptions available)
- Elicitation tasks used:
 - Free impersonal narrative
 - Controlled impersonal narrative
 - Guided interview (personal narrative)
 - Simultaneous actions task
- 240 recordings from classroom learners from three age groups
- 60 native speaker recordings