

TODOs

1. Catalog the sources of data for each language and list them here:
https://docs.google.com/spreadsheets/d/1NtG7jOHGqj4JrJJHjVYPZQ1Olr25_TgS5LDSjKsvSTs/edit?usp=sharing
2. Figure out how to merge the sources of data
3. Unify the annotation schemes to be compatible family wide (provide samples of paradigms in the Google spreadsheet shared in the folder:
https://docs.google.com/spreadsheets/d/11uHXL1aS8SErTqs-Z5pbgPkm-67Y_aF7oA--sl4l3E/edit?usp=sharing)
For tags/annotation examples: see languages that are already in UniMorph (see the corr. subsection)
4. Come up with a unified clitic annotation scheme
5. (optionally) provide for lemmas:
 - a. English translations
 - b. IPA
 - c. Cognates
6. For Apertium data: please add yourself/others to the languages with no annotators

UniMorph schema (follows Leipzig notation):

<https://unimorph.github.io/doc/unimorph-schema.pdf>

Meetings summary:

<https://docs.google.com/document/d/10Y-INfyzRMOhgqOnVEK5Ry9SlzTEcaxL80A8S-dUXNI/edit?usp=sharing>

Aragonese (Annotator's name here)

Comments/notes/questions

Aromanian

Arpitan

Asturian

Catalan

Corsican

French

Friulian

Galician

Italian

Ladin

Latin

Middle French

Neapolitan

Norman

Occitan

Old French

Picard

Portuguese

Romanian

Romansh

Sardinian

Sicilian

Spanish

Venetian