



Mass adding lexemes in Basque language

User:Theklan

Adding by hand: not an option

(inanimate noun) declension of <i>etxe</i> [hide ▲]			
	indefinite	singular	plural
absolute	etxe	etxea	etxeak
ergative	etxek	etxeak	etxeek
dative	etxeri	etxeari	etxeei
genitive	etxeren	etxearen	etxeen
comitative	etxerekin	etxearekin	etxeekin
causative	etxerengatik	etxearengatik	etxeengatik
benefactive	etxerentzat	etxearentzat	etxeentzat
instrumental	etxez	etxeaz	etxeez
inessive	etxetan	etxean	etxeetan
locative	etxetako	etxeko	etxeetako
allative	etxetara	etxera	etxeetara
terminative	etxetaraino	etxeraino	etxeetaraino
directive	etxetarantz	etxerantz	etxeetarantz
destinative	etxetarako	etxerako	etxeetarako
ablative	etxetatik	etxetik	etxeetatik
partitive	etxerik	-	-
prolative	etxetzat	-	-

- Words have declension (actully potposition)
- Traditional approach: using rules

Partnership with Elhuyar

- Elhuyar Foundation has a great amount of language-technology related tools.
- They edit the “Students Dictionary”
- 2019: uploading most common sustantives from this dictionary (10.000), with all forms and senses.

Current statistics

- <https://tools.wmflabs.org/ordia/language/>

What's next?

- Adding more lexemes but...

What's next?

- What can we do with this?
 - How do we get automatic translations?
 - Is it possible to transclude at Wiktionary?
- Find meaning of a word at Wikipedia? Or even external?
- Can we track (somehow) geographical usage of word variants?
 - How to encode pronunciation variants?