Morph telco 2022-10-19, 13:00 CET

**Link:** [**https://meet.google.com/nsj-tbcy-yop**](https://meet.google.com/nsj-tbcy-yop) **[check here for link updates if it doesn’t work]**

**Latest Definitions:**  <https://github.com/ontolex/morph/blob/master/draft.md>

**Nexus:** https://nexuslinguarum.eu/the-action/join-us

**Participants [please add yourself]:**

Christian Chiarcos (CC)

Max Ionov (MI)

Katerina Gkirtzou (KG)

Fahad Khan (FK)

Khadija Ait ElFqih (KAE)

Matteo Pellegrini (MP)

Ciprian-Octavian Truică (CT)

Penny Labropoulou (PL)

Elena Simona Apostol (ESA)

Sina Ahmadi (SA)

Elena Benzoni (EB)

Petra Steiner (PS)

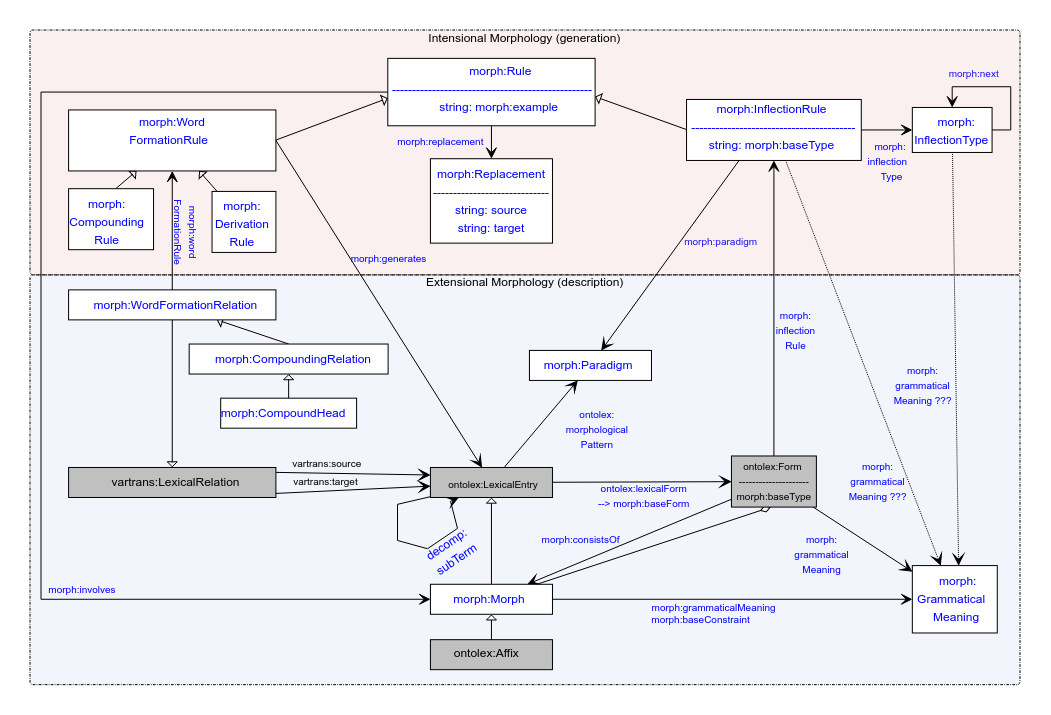
Theodorus Fransen (TF)

Thierry Declerck (DFKI) excused

Ranka Stanković

# 0. Module draft

**Draft 4.16**



No changes

# 1. Publications

## LLODream postproceedings

deadline 01.12, post-presentation publication

For Rasprave journal:

* Croatia’s linguistic journal
* All papers, including abstracts, summaries, notes and references should be not shorter than 24 000 and not longer than 30 000 characters with spaces.
* Manuscripts should be sent by e-mail to [rasprave@ihjj.hr](mailto:rasprave@ihjj.hr)
* Info from : <https://llodapproaches2022.mruni.eu/wp-content/uploads/2022/02/Rasprave_guidelines-for-authors_vilnius.pdf>

**DONE@MI**: ask CC if he wants to take charge on this

* Consensus: wait with an overview until model is stable

## MWE Chapter

* Initiated in FrAC
* Doodle poll: <https://doodle.com/meeting/participate/id/dyrZg8ze>
* Draft for chapter: <https://www.overleaf.com/8285444258rpfnbwgwbrdp>
* Template: <https://www.overleaf.com/latex/templates/langsci-skeleton-for-edited-volumes-2022-06/vvjbpxnrbnkk>, chapters/01.tex
* Initial abstract: <https://docs.google.com/document/d/1HLTsYScMiZE4Nb6b7mozN-SoKAIWtisdiKLa0OoXPuU/edit?usp=sharing>
* Slack channel: #frac-mwe-chapter
* authors: CC KG MI BK FK CT EA
  + We also asked Matteo to join
  + Anyone else?

## eLex 2023

<https://elex.link/elex2023/>

Submission deadline: 31.01.2023

Dates: 27.06-29.06

MI:

* Overview of the model
  + (Maybe just the dynamic part and changes since the last presentation)
* Use case: Semitic? [if we have a consensus on that]

# 2. Handling of Semitic languages

**TO-BE-DONE**: figure out how to include missing categories to LexInfo

Khadija: data prepared for Arabic (cf. last time):

* [Morph Module semitic.odt](https://drive.google.com/file/d/1NST-qdkxAw7am6F3vGAap5o4eQv0ulV1/view?usp=sharing)

Khadija: special-purpose modelling for Arabic?

* GS, CC: no, if that turns out to be necessary, it probably indicates a mistake in the current modelling
  + Fahad Khan\_ aren't hebrew and arabic similar morphologically? even the writing systems don't usually include vocalisations (of short vowels)
  + Khadija Ait ElFqih No, they are not similar. Morphologically
* Always possible to introduce more specific subclasses (in a separate namespace), but OntoLex concepts should be language-independent

Necessary features:

* lexinfo:POS extensions
  + Process:
    - Fork repo
    - Change data under <https://github.com/ontolex/lexinfo/tree/master/data> (in your repo)
    - Pull request
      * Cf. <https://github.com/ontolex/lexinfo/pulls>
    - Problems: mail John McCrae (or ask via Nexus Slack)
* Modelling diacritics
  + Diacritics change the meaning (and morphology) of the word (root?)
    - clarification: the symbols (diacritics) are vowels, right?
      * Vowels, no vowel or repeat consonant
    - and they are applied to the consonantal root, not to a regular word
      * Both, changes meaning
  + MI: if modelled as derivation or inflection, we need to specifiy a morph that is added to the lexical entry, but it would consist of a diacritic only, no written rep, thus. But this diacritic would probably not be modelled as a LexicalEntry (= Morph)
    - Khadija: might be a good idea to model it as lexical entry
    - MI: Concern: might not be in scope of lex entry definition
    - CC: Similar to Umlaut in German
  + CC: suggestion to model in rules (if Morph modelling is dis-preferred)
    - GS: positive experiences with linguists in developing an MT system
    - GS: possible problem: if you change a diacritic, you might have contextual dependencies, maybe better if the underlying representation is not a written representation, but something more abstract
    - GS: diacritics just anpother way to change a string (as a string operation, an infix is exactly the same operation as a diacritic)
      * For that, we need to split diacritics from base characters
      * Requirement: Unicode NFD normalization (<https://towardsdatascience.com/what-on-earth-is-unicode-normalization-56c005c55ad0>)
        + **DONE@CC (2022-10-24)**: ADDED AS A RECOMMENDATION INTO DOC for morph:Replacement
    - GS: Is that even different from Morph?
      * Khadija: change in semantics
      * CC: at the moment, no way of modelling this without morph:Morph, but that was one of the reasons for connecting morph:InflectionRUle with miorph\_:GrammatgicalMeaning
    - there seems to be a gap in the modelling: represent the meaning/function of a morphological rule without an explicit morph
      * Suggestion: use morph:grammaticalMeaning as a property of morph:Rule
        + Matteo Pellegrini In general I think it would be good to be able to express variation in meaning without having to rely on morph, So I am in favour of linkin morph:Rule to morph:GrammaticalMeaning
        + Gilles Sérasset +1
      * **DONE@CC (2022-10-24)**: update the model accordingly
      * **TODO@FK**: model a Hindi/Urdu example, e.g., <https://en.wiktionary.org/wiki/%E0%A4%96%E0%A5%81%E0%A4%B2%E0%A4%A8%E0%A4%BE#Hindi>
      * **TODO@Khadija**: Modelling examples for Arabic entries

# 3 Postponed

* Semitic: Circumfix
* Consolidating Definitions
* Thierry’s STSM

# 4. AoB

Next call in 2 weeks