Morph telco 2022-04-19, 13:15 CEST

**Link:** [**https://meet.google.com/nsj-tbcy-yop**](https://meet.google.com/nsj-tbcy-yop) **[CHECK HERE FOR UPDATED LINK(S)]**

**Latest Definitions:** [**https://github.com/ontolex/morph/blob/master/draft.md**](https://github.com/ontolex/morph/blob/master/draft.md)

**Nexus:** [**https://nexuslinguarum.eu/the-action/join-us**](https://nexuslinguarum.eu/the-action/join-us)

**Participants [please add yourself]:**

Christian Chiarcos (CC) (excused for being 10 min late)

Max Ionov (MI)

Katerina Gkirtzou (KG)

Besim Kabashi (BK)

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)

Ranka Stanković (RS)

Gilles Sérasset (GS)

Mike Rosner (MR)

Table of Contents

[0. Module draft (4.17)](#_v4addnpz2vtu)

[1. Semitic Roots](#_f94a165qguj0)

[2. Morph ordering](#_uly9ufoswxxy)

[3. Open topics (from 2023-04-05)](#_mmf5sao0z9x2)

[POSTPONED](#_yw6tpr61n4hz)

[Next time](#_s2vgjo4z02ta)

# 0. Module draft (4.17)



# 1. Semitic Roots

* 2023-04-19
  + Suggestion (MI): Roots are morphs (lexinfo:RootMorph), canonical forms specified for a lexical entry consist of them.
* Example (with minor comments)

roots:k-t-b a lexinfo:RootMorph, ontolex:LexicalEntry ;

ontolex:evokes :k-t-b\_meaning;

rdfs:label "k-t-b" .

:k-t-b\_meaning a ontolex:LexicalConcept.

# not: LexicalSense

:kiteb a ontolex:Word ;

lexinfo:partOfSpeech lexinfo:verb ;

morph:morphologicalPattern <kiteb\_paradigm> ;

ontolex:canonicalForm <kiteb\_form> ;

morph:baseForm <kiteb\_form> .

<kiteb\_form> a ontolex:Form ;

morph:consistsOf roots:k-t-b ;

ontolex:writtenRep "kiteb"@mlt ;

ontolex:phoneticRep "/kɪtɛp/" .

* + Approved by participants
* **Tbc**: solved?
  + Consensus: yes
  + **TODO**: add as example to draft
* Do we need unordered along with numbered properties for consistsOf ?
  + Consensus: yes
* Can we abandon consistsOf in favour of rdfs:member
  + Consensus: no (because it’s more established, it might have special semantics [which is to be confirmed], and this double modelling has precedent in OntoLex-Decomp)
* No lexical senses for RootMorph
  + **TODO**: put into RootMorph definition that the set of lexical senses for RootMorph MUST be empty because root morph are not independent lexemes but used in different lexemes with different senses.
    - otherwise, the ontolex definition may lead people to use lexical sense here, because the root morph actually \*is\* a single lexical entry.
* RootMorph rdfs:subClassOf morph:Morph?
  + KG: Using the morph:consistsOf here, implies (by inference) that the roots:k-t-b is a morph:Morph. Do we want this inference? Meaning does linguistically make sense? If it is always true I would propose to declare lexinfo:RootMorph as a subclass of morph:Morph. I have a similar question/issue regarding lexinfo:StemMorph
  + CC: yes, this is the intuition, but in the current proposal, this is not done (<https://github.com/ontolex/lexinfo/pull/29/commits/c50438a57e33152586bfc38cef73ef4a1b5535f0>), because the Morph vocabulary is unpublished and we shouldn’t contaminate Lexinfo with premature stuff.

=> **TODO**: fix/update pull request *after Morph publication*

* lexinfo:etymologicalRoot
  + 2023-04-19, FK: lexinfo:etymologicalRoot: lexical entry -> lexical entry; “Morpheme that has a particular status with regards to the word's etymology.” can be used to link lexical entries (!) with a lexinfo:RootMorph
* **Tbc**: is that enough or do we need another relation?
  + Confirmed, yes

# 2. Morph ordering

One of the requirements.

* RDF, unordered: morph:consistsOf <morph1>, <morph2>
  + NB: for circumflexes, we either need to decompose or just to provide morph:consistsOf
* It already is in the diagram: diamond-symbol operator (that’s standing for rdfs:Seq)

:x morph:consistsOf :m1, :m2.

:x a rdfs:Seq; rdf:\_1 :m1; rdf:\_2 :m2.

* KG:
  + x is the ontolex:Form right? I thought that the rdfs:Seq implied from the diamond in the diagram would result in something like this, no?

:x a ontolex:Form ;

morph:consistsOf [

a rdfs:Seq ;

rdf:\_1 :y1;

rdf:\_2 :y2;

]

.

:y1 a morph:Morph .

:y2 a morph:Morph .

* That was also Matteo’s intuitive interpretation
* CC: modelling inspired by the usage of morph:Morph for morph(ological segment)s in interlinear glossed text
* This was not the intended modelling, indeed, but it has advantages
  + PRO: multiple segmentations
  + PRO: defining form as a seq of morph(eme)s may be counterintuitive to linguists
  + CONTRA: Could make the diagram quite bulky. Need to create another aggregator class
  + CONTRA: we cannot define RDFS constraints on the components of a container
  + CONTRA: consistsOf should then always point to Seqs
* To be further discussed
* Alternatively, we could use rdf:List (not discussed)

Turtle:

:x owl:sameAs ( :m1, :m2 ).

equivalent to:

:x a rdf:List;

rdf:first :m1;

rdf:rest [

a rdf:List;

rdf:first :m2;

rdf:rest rdf:nil ] .

rdf:List have a very nice Turtle syntax, but are much harder to query.

previously, we had a consensus *against* rdf:List, also following other OntoLex

modules (esp., decomp)

TBC: do we stay with rdfs:Seq/rdfs:Bag?

* **Tbc**: this is solved, right?
  + No, we might need to reconsider to follow the modelling along the lines of katerina’s proposal, because that allows for multiple different segmentations

# 3. Open topics (from 2023-04-05)

* morph:baseConstraint
  + **todo@kg**: more comprehensible definition needed
* MP: Naming “paradigm” may evoke incorrect associations, definition of paradigm in linguistics as a set of word forms

Note: we also need to rephrase the “paradigm” property, then

* + morphological pattern
  + Inflection type?
  + Inflection class?
    - MP: +1
    - GS: sounds like owl class
  + Inflection?

=> hard to get to a consensus, vote?

# POSTPONED

* MP: generates to lexical forms?
* Besim: Clitics, e.g., Macedonian
  + Max: idea was to treat them like separable prefixes in German, it’s ok to have a space in a written representation of a form
  + **TO-BE-DONE**: find example data of lexical entries in 5 languages
    - Ranka: Samples for Serbian (link below)
* Character / sound classes
  + Options: either within morph or within the scope of a new module?
    - If the latter:
      * together with signs/multimodality?
      * together with other transformation rules (e.g., for transcription)?
        + Could also include finite state terminology
        + Could also include diachrony
* Finite state terminology
  + Options: either within morph or within the scope of a new module?
    - If the latter:
      * Together with diachrony?
      * Together with sound classes?
* MWEs
  + 2023-04-05 Gilles: describe decomposition of MWEs in DBnary w. Thierry
* Non-concatenative morphology
  + 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> (Fahad)
* **TO-BE-DONE**@Katerina:
  + list challenges from LDL
  + Tbc: are they still challenges or not, if so, do they fall under morph?
* GitHub issues
  + <https://github.com/ontolex/morph/pull/16>
  + <https://github.com/ontolex/morph/issues>
* Wiki: requirements (see 2023-04-05)
* github/gdrive: data sets (see 2023-04-05)
* comments from draft
  + <https://github.com/ontolex/morph/blob/master/draft.md>

## Next time

Discussing problems with Old Irish in its representation in morph