Morph telco 2022-04-20, 13:00 CET

**Link:** [**https://meet.google.com/nsj-tbcy-yop**](https://meet.google.com/nsj-tbcy-yop)

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

**Participants [please add yourself]:**

Christian Chiarcos (CC) [christian.chiarcos@gmail.com](mailto:christian.chiarcos@gmail.com)

Max Ionov (MI)

Katerina Gkirtzou (KG)

Fahad Khan (FK)

Matteo Pellegrini (MP)

Ciprian-Octavian Truică (CT)

Penny Labropoulou (PL)

Elena Simona Apostol (ESA)

Agenda (please add, *but* do not edit table of contents directly, but add sections below and then update here):

[**0. Module draft**](#_iu5bilovzl4h) **1**

[**1 follow-up LDL submission**](#_5q4h03dxajtd) **2**

[1.1 crafting/collecting definitions](#_n0e2ll1nl5iu) 2

[**1.2 Discussion points from paper**](#_letz9qce1sim) **3**

[1.3 Other observations](#_9il2am7fjrme) 3

[**2. German FST**](#_hnsslo6efssf) **4**

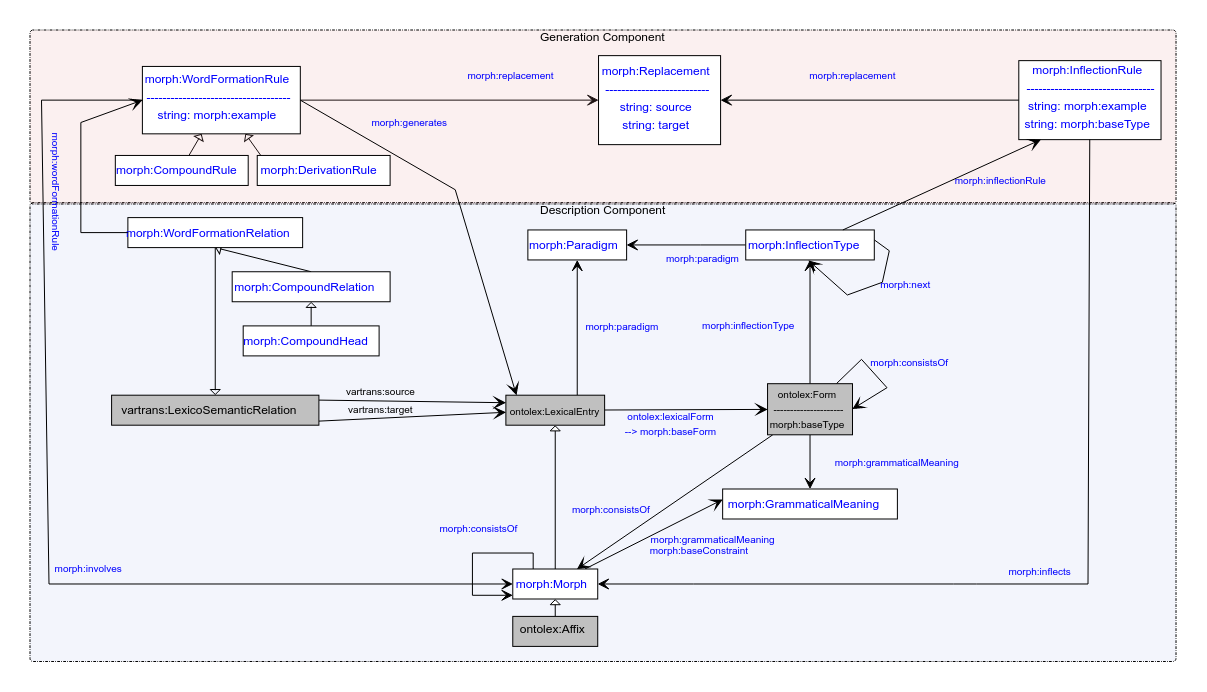
[**3 vartrans:orthVariant (Old English/Fahad)**](#_btgor4akt3q) **4**

[**4 Discussion order for other data**](#_jwteu367w2rp) **4**

[**5. AOB**](#_4u89668ejc7q) **4**

# 0. Module draft

**draft 4.13 (updated for LREC paper)**

****

**Model draft 4.13 updates:**

* Renamed morph:stemType property in the morph:Form (not in the InflectionRule) => morph:baseType
* dropped Morph subclasses (to be moved to LexInfo)

**Model draft 4.14 updates:**

* see below

# Definition consolidaiton

# 1 follow-up LDL submission

documents modelling progress, 8 pages

submitted Aug 14, revised Aug 18 (Fahad to be added in metadata, Max to be removed from metadata)

overleaf: <https://www.overleaf.com/4868363189kczjzdndgxwc> (folder submission/)

## other publication plans

* COLING?
  + may 17
  + could be an altewrnative venue for the LDL submission (if rejeted or revoked) or a novel paper (if we have novel content)
    - decision postponed
  + idea for novel paper: word formation in OntoLex-Lemon
    - not original content, but more like a survey and documentation of best practices?
    - can be helpful to consolidate/revise word formation part of the module
    - possible input from LiLa
  + paper@COLING:
    - FK: update of OntoLex, incl. FrAC, Morph, MModality
* LLODREAM? (Matteo, Marco) <https://easychair.org/cfp/llodream2022>
  + 500 words
  + deadline june 15th
  + ?publication
* later journal paper

## 1.1 crafting/collecting definitions

* FK: definitions ready by May 17 (before COLING deadline)
* <https://github.com/ontolex/morph/blob/master/draft.md>
* you can contribute suggestions by creating issues (<https://github.com/ontolex/morph/issues>), via pull requests, or by direct editing (share your GitHub username)
* definition refinement
  + procedure:
    - open an issue
    - pull request + close the issue
  + TODO@CC: update definition of inflects in GitHub draft
  + Matteo: read definitions until next telco
  + Penny: read definitions => issue

## 1.2 walk-through LDL paper & discussion:

* + define morph subclasses in LexInfo rather than OntoLex-Morph, cf. <https://github.com/ontolex/lexinfo/issues/21> (needs update: not as TermElement)
    - John: can be defined as being equivalent
      * lexinfo:Prefix subclassOf [ lexinfo:termElement lexinfo:prefix ]
  + morph:inflects => morph:involves (these things are largely parallel)
    - tentative consensus
  + morph:paradigm (between lexical entry and paradigm) => ontolex:morphologicalPattern
    - we never used ontolex:morphologicalPattern. Suggestion: use that in place of the paradigm property \*between LexicalEntry and Paradigm\* (not between paradigm and inflection type/rule) [tentative consensus]
    - John: that was the original intention
  + vartrans:LexicoSemanticRelation => vartrans:LexicalRelation
    - why is WordFormationRelation not a LexicalRelation but a LexicoSemanticRelation?
      * it used to be the former
    - this may just have been by error => revise
  + CompoundXY => CompoundingXY?
    - * proposed by Matteo, should describe the process, esp. for CompoundingRule
      * CompoundHead should stay
      * CompoundingRelation instead of CompoundRelation
    - CompoundingRelation?
      * **pro renaming: 1**
      * contra renaming: 0
    - CompoundingHead?
      * pro: 0
      * **contra: 1**
    - CompoundingRule?
      * **pro renaming: 2**
      * contra renaming: 0
  + in the report, we must describe the relation between decomp and CompoundRelation
    - CC: I’m not sure decomp:Component is well-suited for compound analysis, because it doesn’t relate to lexicosemantic relations.
    - decomp:subTerm definitely is well-suoited for compound analysis. morph:CompoundRelation could be \*motivated\* as an alternative reification of decomp:subTerm. and the reason for reifying is to link it to a word formation rule (and, via the rule, with a morpheme).
    - the current modelling of decomp is oriented towards an analysis of synsem (semantic) roles within a compound. in morphology, we normally don’t have that, what we have, instead, are relations between lexemes and morphemes.
    - in particular there is no added value if modelling morph as decomp:Component, because it’s correspondsTo points to other lexical entries (incl. morph, but this can be done with decomp:subterm, too) or \*syntactic\* arguments (inapplicable to morphology).

=> define CompoundingRelation as reification of decomp:subterm

* + cuts in LEXIS and LILA sections to be double-checked by Penny & Katerina, resp., Matteo & Marco
  + German FST: use inflection type to represent finite states and their transitions
    - FST grammars (German, Christian, <https://github.com/acoli-repo/acoli-morph/tree/main/morphisto>)
      * works: we can model and retrieve replacement sequences
      * terminologically, this is a bit of a stretch => discuss inflection type
  + inflection type => separate discussion
  + variation in inflection:
    - Matteo: would like to see some grouping of (sub-)lexical entries with paradigm (“Flexeme”)
      * flexeme may be too specific, and too little used
      * suggestion: model the grouping by lexicog, have both the overarching lexical entry and the flexemes as separate lexical entries, no vocabulary extension needed, but a usage note in the report
    - Penny: would like to trace contracted and non-contracted versions of the same paradigm
      * can be partially modelled by means of “markers”, i.e., lexinfo usage properties, instead, tbc. whether lexinfo needs to be extended for that
  + metadata properties for LexInfo (hypothetical/unattested form, etc.) => tentative consensus, but details to be discussed

## 1.2 Postponed

* revise/abandon inflection type? (needs to be checked on data => discuss German FST data)
  + Max: no, this is a complete revision of generation model
  + we don’t seem to have a common understanding of what InflectionType is meant to be
    - Max: hard to explain
  + to be discussed next time
* <https://link.springer.com/chapter/10.1007/978-3-030-98876-0_34>: MMoOn modelling for Greek morphology. Worth a comparison?
* consistsOf
  + currently
    - ambiguous: form -> form, form -> morph, morph -> morph
    - doesn’t represent order
  + suggestion:
    - replace by aggregation (rdfs:Seg or rdfs:Bag)
    - define a form as an aggregate of morphs
    - no such relation between form -> form and morph -> morph
    - instead of morph -> morph: because morph is a lexical entry, we go from morph -> form, and then use “consistsOf” equivalent at form level
    - instead of form -> form: use form -> morph and assign the “sub-forms” as lexicalForm to the morph
  + has the original “consistsOf” ever been used by anyone before?

# 4 other data

define discussion order for

* sample data on GDrive
  + where is the link?
* Fahad: OE sample data: coman/quoman example, cf. <https://en.wiktionary.org/wiki/cuman#Old_English>
  + issues with dialects (reference dialect vs. other dialects) and diachrony (phonological processes); treatment of syncopation, suppletion, fusion of different roots [=> variants?] ?
* semitic roots
  + last time:
    - FK: need a Semitic example, from the same consonant cluster, we can generate different POSes, <https://en.wikipedia.org/wiki/K-T-B>, <https://en.wiktionary.org/wiki/%D9%83_%D8%AA_%D8%A8>
    - OntoLex limitations: one POS per lexical entry
    - Max: example from Arabic dictionary (<https://en.wikipedia.org/wiki/Dictionary_of_Modern_Written_Arabic>): organized by roots, but root is not made explicit
    - postponed until we have a Semitic speaker
    - Ilan?
      * but first, check Bettina’s conversion of KDictionaries’ Hebrew dict

# 5. AOB

next call in two weeks