Morph telco 2022-02-23, 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)

Matteo Pellegrini (MP)

Max Ionov (MI)

Katerina Gkirtzou (KG)

Penny Labropoulou (PL)

Fahad Khan (FK)

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

[**0. Orga**](#_ltn48ch8cnns) **1**

[**1. Module draft**](#_iu5bilovzl4h) **3**

[**2 Evaluation of inflectional data modeling**](#_h3gli2fvks3y) **4**

[2.1 POSTPONED: vartrans:orthVariant (Greek, Penny)](#_z07otlr5e986) 4

[2.2 POSTPONED: Orthographical variants in Old English (Fahad)](#_4mr3o3oi5cf7) 5

[2.3 Latin inflectional data example by Matteo](#_nuozmk7g1l68) 6

[2.4 morph:stemType (Greek/Latin)](#_vj6s4tgj246t) 7

[2.5 POSTPONED: Other data sets](#_dmwdwvjzy3w9) 7

[**3. Other todos**](#_imja6y9svso) **7**

[3.1 OntoLex TermElement vs. Morph subclasses](#_qb6wo82rlngi) 7

[3.2 crafting/collecting definitions](#_n0e2ll1nl5iu) 7

[3.3 LDL submission ? [=> next week]](#_4u89668ejc7q) 8

# 0. Orga

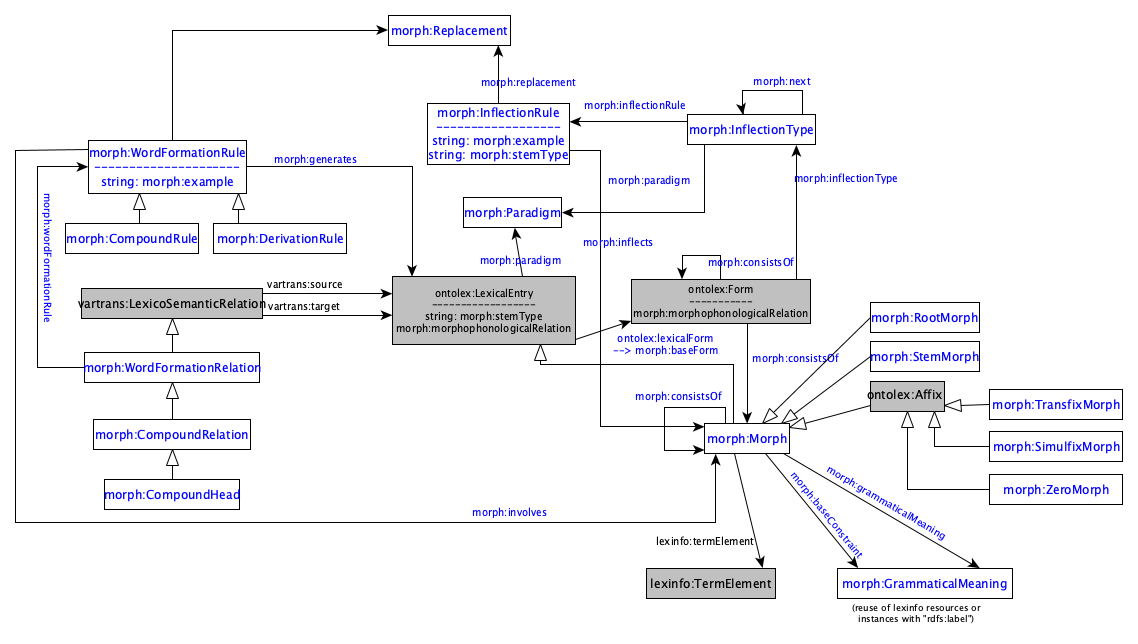
* consolidated wiki and documentation
  + diagram sources and minutes now mirrorred under <https://github.com/ontolex/morph>
* draft definitions under <https://github.com/ontolex/morph/blob/master/draft.md>, (to be) continuously updated
  + restructured into logical sections
  + completed (according to diagram), but no additional definitions created

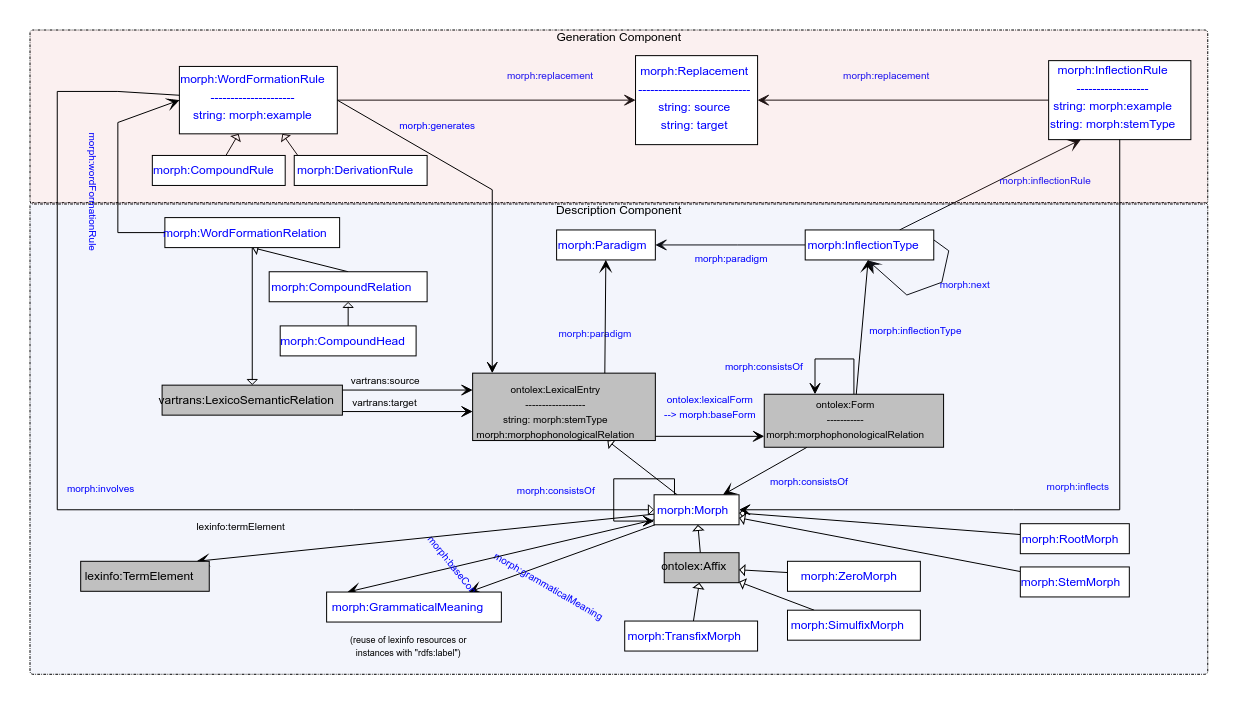
**TODO** (@all): Add definitions that are not there or outdated

# 

# 1. Module draft

**draft 4.11 (old layout & reformatted)**

****

****

**Model draft 4.11 updates & questions**

* changed display of datatype properties (as in UML)
* round edges replaced by squared edges (in order to work with online editor)
* morph:replacement points to a class Replacement which takes source and target datatype properties (as modelled in Thierry’s paper)
* Question: shouldn’t LexicalEntry (except as superclass of morph:Morph) be an ontolex:Word
  + TODO: think about this before publishing the final model
  + MI: I think there was a discussion about this in the very beginning (about affixes). To check
* Question: definition of morph:morphophonologicalRelation? (@Penny?)
  + drop it
* the arrow for involves is wrong. must be \*filled\*

**Model draft 4.12 updates:**

* grammaticalMeaning also pointing from ontolex:Form to GrammaticalMeaning (per request by Matteo and Penny) – mostly for filtering convenience
* stemType moving from LexicalEntry to ontolex:Form

# 2 Evaluation of inflectional data modeling

## 2.1 POSTPONED: vartrans:orthVariant (Greek, Penny)

New examples with vartrans:orthVariant as a vartrans:category (found it more interesting to instead of adding it just as a subproperty). What is still not clear though is how to define that the two (or more) orthographic variants share the same senses and syntactic behaviour.

lexis\_data:augo\_26232 a ontolex:LexicalEntry ;

rdfs:label "αυγό"@el ;

ontolex:canonicalForm lexis\_data:augo\_form

lexinfo:partOfSpeech lexinfo:commonNoun ;

morph:paradigm lexis\_data:vouno ;

ontolex:sense lexis\_data:augo\_sense\_USem1074 ;

synsem:synBehavior lexis\_data:augo\_SUNo25013 .

lexis\_data:avgo\_34067 a ontolex:LexicalEntry ;

rdfs:label "αβγό"@el ;

ontolex:canonicalForm lexis\_data:avgo\_form

lexinfo:partOfSpeech lexinfo:commonNoun ;

morph:paradigm lexis\_data:vouno ;

~~ontolex:sense lexis\_data:augo\_sense\_USem1074 ;~~

~~synsem:synBehavior lexis\_data:augo\_SUNo25013~~ .

lexis\_data:avgo\_form ontolex:writtenRep "αβγό"@el .

lexis\_data:augo\_form ontolex:writtenRep "αυγό"@el .

lexis\_data:avgo\_orthVariants a vartrans:LexicalRelation ;

vartrans:source lexis\_data:avgo ;

vartrans:target lexis\_data:augo ;

vartrans:category lexis:orthVariant .

Questions:

* termbase instead of lexis for orthVariant?
* any relation with `morph:baseForm` ?

Option with lexicog module works but multiplies the lexical entry data, e.g. senses have to be created for all lexicographic components.

* consensus

→ to be developed as part of the vartrans module but applied in order to connect two ontolex:LexicalEntry resources that only differ in their orthographic representation but share the same senses: a vartrans:orthVariant subproperty of vartrans:lexicalRel that entails that the same senses apply to both variants (restriction has to be formulated)

* possibilities (last time):

1. create separate entries and link them with a subproperty of vartrans:lexicalRel (your\_namespace:orthVariant sub vartrans:lexicalRel) between two ontolex:LexicalEntries that are orthographic variants - the shared data of both would have to be repeated for both entries then or it has to be stated only on one lexical entry resource (duplicate all senses)
2. alternatively, leave senses of the variant empty, define in the semantics of orthVariant that this entails that the same senses apply to both variants)
3. use owl:sameAs and share senses (a bit messy, because semantically this means they are the same thing)

## 2.2 POSTPONED: Orthographical variants in Old English (Fahad)

* was TODO@Fahad: sample data

## 2.3 **Latin inflectional data example by Matteo**

<https://drive.google.com/file/d/1v2M-LbCdrdaPl2LVuPKzOfU-fQjv_5_1/view?usp=sharing>

was this complete?

Representation of inflectional information of orthographic variants:

class: lila:Flexeme (subclass of ontolex:LexicalEntry) are linked to lila:Lemma

*lexemes → lexical units with a unique meaning*

*▶ flexemes → lexical units with a unique form (i.e., a unique inflectional paradigm)*

object property between lila:Lemma: lila:lemmaVariant (subclass of ontolex:Form)

Matteo continues his presentation on Latin inflection and raises further questions.

bundles of features

* grammaticalMeaning can be either a set or a single feature
* need property between form and grammatical meaning (reification for bundles of lexinfo properties)

multiple ontolex:Forms for the same feature bundle?

* yes, unless purely orthographical
* but what is purely orthographical is a matter of debate
* if the underlying resource requires a differentiation, one can also create separate forms for orthographic variants
  + OntoLex: “Different forms are used to express different morphological forms of the entry. They should not be used to represent ortographical variants, which should be represented as different representations of the same form. For example, for the lexical entry *color*, we would have two different representations of the same form, one for the British English written representation *colour* and one for the American English written representation *color*. Both representations have the same pronunciation and the same meaning, so they are two different lexicographic variants of the same lexical entry”

=> if it is not sure whether they have the same pronunciation or meaning, a researcher/resource who wants/has to distinguish it can (this is not MUST anyway, but SHOULD, so we can deviate under justified circumstances)

* different base forms for generation
  + cf. stemType (unsolved yet)
  + MI: can be illustrated/documented with morph:example (if generation with regular expressions is too hard)
  + Penny: my example is computational, not lexicographical, needed there

## 2.4 morph:stemType (Greek/Latin)

(from modern Greek inflection)

=> morph:stemType (identifier for different stems, also needed for Latin)

question: why at inflection rule and at ontolex:LexicalEntry ?

cf. Question (CC, 2022-02-23, offline): wouldn't it make more sense to model stemType at a (base) form rather than at the level of lexical entries?

next week: tbc whether this works for Latin/Matteo

Added slides for tomorrow's discussion: <https://docs.google.com/presentation/d/1DI50ytMI1xcDjxpPQhAT1t6-jWbr5zCX/edit#slide=id.p1>

- Greek inflection: slides 1-3

- orthographic variants: slide 4

## 2.5 POSTPONED: Other data sets

define discussion order for

* FST grammars (German, Christian+Max, <https://github.com/acoli-repo/acoli-morph/tree/main/morphisto>)
* Old English (Fahad)
* sample data on GDrive

# 3. Other todos

## 3.1 OntoLex TermElement vs. Morph subclasses

* <https://www.w3.org/community/ontolex/wiki/Morphology#Fixed_set_of_morph:Morph_classes:_Telco_12.05.2021>
* <https://github.com/ontolex/lexinfo/issues/21>
* question: extend lexinfo (and abandon morph subclasses) or morph subclasses (and replicate lexinfo) ?
  + the current modelling with some subtypes of morph(eme)s modelled as subclasses and others as TermElements is inconsistent

## 3.2 crafting/collecting definitions

* <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 password)

## 3.3 LDL submission ? [=> next week]