Morph telco 2022-03-23, 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)

Max Ionov (MI)

Katerina Gkirtzou (KG)

Fahad Khan (FK)

Matteo Pellegrini (MP)

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 stemType**](#_c0gfniblyea4) **2**

[**2 LDL submission**](#_5q4h03dxajtd) **3**

[**3 Evaluation of inflectional data modeling**](#_letz9qce1sim) **3**

[3.1 morph:stemType (Greek/Latin)](#_vj6s4tgj246t) 3

[3.2 vartrans:orthVariant (Greek/Penny; Old English/Fahad)](#_z07otlr5e986) 3

[3.3 POSTPONED: Other data](#_dmwdwvjzy3w9) 5

[**4. Other todos**](#_imja6y9svso) **5**

[3.1 OntoLex TermElement vs. Morph subclasses](#_qb6wo82rlngi) 5

[3.2 crafting/collecting definitions](#_n0e2ll1nl5iu) 5

# 0. Module draft

**draft 4.12**

****

**Model draft 4.13 updates:**

* Rename morph:stemType property in the morph:Form (not in the InflectionRule).

# 

# 1 stemType

* last time: different base forms for generation
  + stemType for “co-indexing” inflection rules and baseForms
  + Penny: my example is computational, not lexicographical, needed there
  + slides: <https://docs.google.com/presentation/d/1DI50ytMI1xcDjxpPQhAT1t6-jWbr5zCX/edit#slide=id.p1>

- Greek inflection: slides 1-3

- orthographic variants: slide 4

* summary: stemType not a transparent term for a form (what is a stem type for a form?), Matteo: formType ? could be something else than a morphological stem!
* rename in both classes?
  + agreed
* Matteo: formIndex ?
* Matteo’s examples
  + the term "stem type" is pretty much in line with the traditional literature. "The Verb Stem has three forms , the Present Stem , the Perfect Stem , and the Supine Stem" (Chase 1882)
  + Matteo: lauo vs. lauare is two different base forms, but both present (“pres1” and “pres2”)
  + CC: are these at initial or generated forms?
  + Matteo: both
  + Max: potentially problematic if we generate inflected forms from generated forms
  + CC: morph:baseForm can be used to mark non-generated forms
  + suggestion: baseType (to resemble baseForm)
    - baseForm (in the diagram, but not deeply discussed yet => German FST)
      * marks a form as being the basis of generation
      * doesn’t have to be a real word
    - if multiple base forms exist, these can be distinguished by baseType
      * same as in inflection rule
      * baseType can be propagated to generated forms
      * all inflected forms should be generated from either a canonical form or a base form that has the same baseType (if a baseType is provided)
  + FK: any relation with semitic roots?
  + CC: could be baseForms (if the underlying dictionary encodes it as such, they could also be canonical forms)
  + 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
* FK: adopt Flexeme into model?
* Matteo: possibly too LiLa-specific.
  + can be added if another, LiLa-independent dataset requires that

# 2 LDL submission

suggested last week, could document modelling progress

Deadline: April 8th

Willing to contribute:

* Christian
* Matteo
* Penny
* Max
* Katerina
* ?Fahad (if needed for an example)
  + Potential examples: [wiht - Wiktionary](https://en.wiktionary.org/wiki/wiht#Old_English), [cuman - Wiktionary](https://en.wiktionary.org/wiki/cuman#Old_English)

TOOD@Max: send a Doodle poll to discuss the plan and responsibilities

* idea: diff to 2019 and examples

MOVE ALL THE FOLLOWING TO NEXT TIME

# 3 Evaluation of inflectional data modeling

## 3.2 vartrans:orthVariant (Greek/Penny; Old English/Fahad)

OE sample data: …

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:

* morph namespace 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)

## 3.3 POSTPONED: Other data

define discussion order for

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

# 4. 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 username)

## 