Morph telco 2023-11-15, 13:00 CET

**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)

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)

Ranka Stanković (RS)

Gilles Sérasset (GS)

Mike Rosner (MR)

Table of Contents

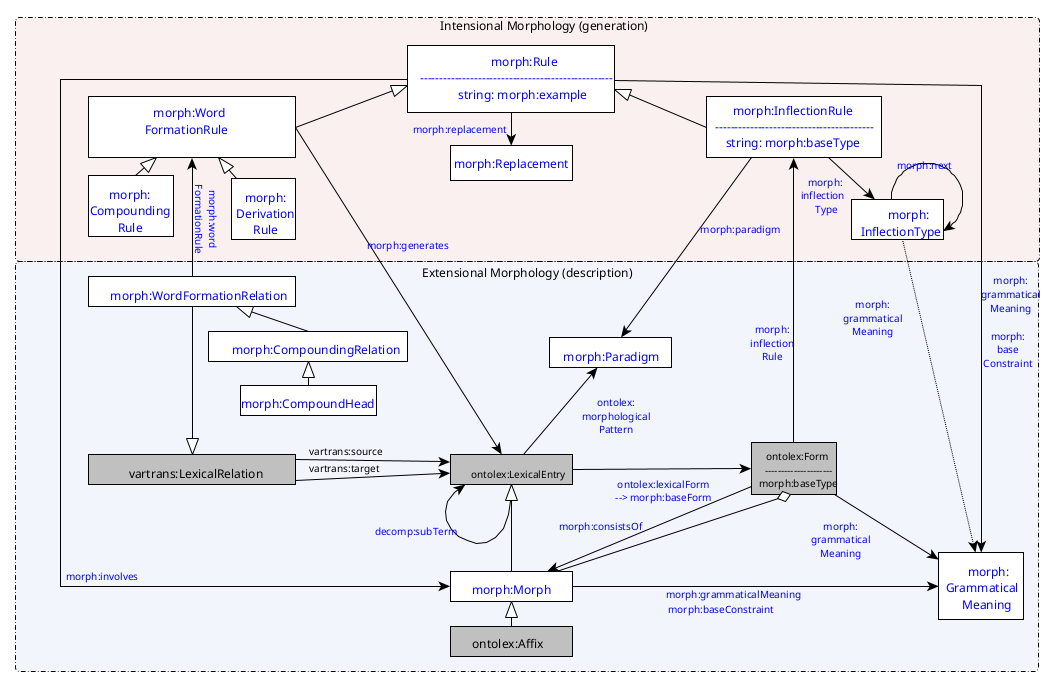
[0. Module draft (4.18)](#_v4addnpz2vtu)

[1. CC: Inuktikut data and problems with LexicalEntries](#_8pfoi21gwxpp)

[2. Grammatical meaning connection for InflectionType](#_crkrxhcgs4bw)

[3. Next call](#_gs9uncz12oeu)

# 0. Module draft (4.18)



* Removed ??? from morph:grammaticalMeaning. We can support this property for morph:InflectonType (we might discuss this today if we have time)
* Made the round arrow pretty :)

# 1. CC: Inuktikut data and problems with LexicalEntries

<https://github.com/acoli-repo/morph-addenda/tree/master/data/polysynthetic>

The problem: ambiguity in word formation — <https://github.com/acoli-repo/morph-addenda/tree/master/data/polysynthetic#5-encode-ambiguity-in-derivation>

Discussion:

* Example atausiulugu
  + This is an inflected word, but the inflection part is unproblematic. The word formation part is challenging, and that part would also be relevant for a dictionary.
* Can a morph:Morph have more than one form?
  + CC: inuktitut morph. Segmentation provides canonical form of affix along with current string realization. How to model that? (proposal: multiple forms of the same morph)
  + If yes, this could be used to represent (some cases of) allomorphy
  + MI: original consensus to not require morph:Morphs to be morphemes, morpheme handling within Mmoon
* Allomorphy
  + MI: one morph:Morph per form, relate allomorphs indirectly by repurposing grammatical meaning (i.e., one Morpheme=one grammatical meaning)
  + CC: redundancy and information loss, incompatible with otherForms for other lexical entries.
  + MI: if morphophonological, then just provide rules instead of segments
  + CC: we cannot require everyone who wants to give a morphological segmentation to write a morphological analyzer
* CC: Instead of forms being a sequence of morphs, model them as a sequence of forms (example below), otherwise we lose information about the realization of a morph(eme) at a particular position
  + If allomorphs are modelled as alternative forms of the same underlying morph(eme), this allows us to provide information about the specific form variant of a morph used as a particular position in a particular word without having to write full-fledged assimilation rules
  + This means to allows one ontolex:Form to be a rdfs:Seq of ontolex:Forms (rather morph:Morphs). No changes to morph:contains
  + MI: leads to problems with downward compatibility
  + CC: if we don’t arrive at a consensus here, we could skip the rdfs:Seq part of forms, because this is nothing defined in morph namespace anyway
* MI: Next meeting supposed to be the last regular morph meeting
  + FK: really? MI: There may be irregular meetings

**POST-CALL ADDENDA:** Excerpt from <https://github.com/acoli-repo/morph-addenda/tree/master/data/polysynthetic> (apologies for not being able to find them on the spot)

Original analysis:

{atausi:atausiq/1n}{u:u/1nv}{lugu:lugu/tv-part-1s-3s-fut} {one}{existence; is}{part. future: while I \...him/her/it}

As IGT (with morpheme boundaries added, these are indirectly expressed from the fourth row with grammaticalMeaning/baseConstraints)

| atausiulugu | | |
| --- | --- | --- |
| atausi- | -u- | -lugu |
| atausiq- | -u- | -lugu |
| 1n | 1nv | tv-part-1s-3s-fut |
| one | existence; Is | part. future: while I \...him/her/it |

Original dictionary:

{atausi:atausiq/1n} {one}

{u:u/1nv} {existence; is}

{lugu:lugu/tv-part-1s-3s-fut} {part. future: while I \...him/her/it}

In OntoLex-Morph:

# CURRENT Seq with morphs

# FAIL 1: we loose information about the actual forms

:atausiulugu\_le a ontolex:Word;

ontolex:canonicalForm :atausiulugu\_atausiulugu\_f.

:atausiulugu\_atausiulugu\_f a ontolex:Form;

ontolex:writtenRep "atausiulugu"@iu-Latn.

**:atausiulugu\_atausiulugu\_f** a rdfs:Seq;

**rdf:\_1 :atausiq\_1\_le;**

**rdf:\_2 :u\_1\_le;**

**rdf:\_3 :lugu\_tv\_part\_le.**

# REVISION Seq with forms

# FAIL 1: SOLVED

formation rules, see data)

:atausiulugu\_le a ontolex:Word;

ontolex:canonicalForm :atausiulugu\_atausiulugu\_f.

:atausiulugu\_atausiulugu\_f a ontolex:Form;

ontolex:writtenRep "atausiulugu"@iu-Latn.

**:atausiulugu\_atausiulugu\_f** a rdfs:Seq;

**rdf:\_1 :atausiq\_atausi\_f;**

**rdf:\_2 :u\_u\_f;**

**rdf:\_3 :lugu\_lugu\_f.**

Affix/base dictionary:

**:atausiq\_1\_le** a lexinfo:RootMorph, ontolex:LexicalEntry;

ontolex:canonicalForm :atausiq\_atausiq\_f;

**ontolex:otherForm :atausiq\_atausi\_f**;

ontolex:sense :atausiq\_1n;

lexinfo:partOfSpeech lexinfo:noun. # SAME: "type: nominal root"

:atausiq\_atausiq\_f a ontolex:Form;

ontolex:writtenRep "atausiq"@iu-Latn, "ᐊᑕᐅᓯᖅ"@iu-Cans.

**:atausiq\_atausi\_f** a ontolex:Form;

ontolex:writtenRep "atausi"@iu-Latn, "ᐊᑕᐅᓯ"@iu-Cans.

:atausiq\_1n a ontolex:LexicalSense;

skos:definition "one"@en; # SAME

ontolex:concept :number\_quantity. # NEW

# this shows allomorphy for a root morph, it’s the same for affixes, though

# {u:u/1nv} {existence; is}

**:u\_1\_le** a ontolex:Affix;

**ontolex:canonicalForm :u\_u\_f**;

# we have more than one u-form, so the ids have to be more specific,

# as these forms differ in their phonological context

ontolex:sense :u\_1nv;

morph:grammaticalMeaning :verb; # ../1n\*v\*

morph:baseConstraint :noun. # ../1\*n\*v

:u\_1nv a ontolex:LexicalSense;

skos:definition "existence; is"@en.

**:u\_u\_f** a ontolex:Form;

ontolex:writtenRep "u".

# -lugu is somewhat ambiguous:

# {lugu:lugu/tv-part-1d-3s-fut} {part. future: while we (two) \...him/her/it}

# {lugu:lugu/tv-part-1p-3s-fut} {part. future: while we (many) \...him/her/it}

# {lugu:lugu/tv-part-1s-3s-fut} {part. future: while I \...him/her/it}

# {lugu:lugu/tv-part-2d-3s-fut} {part. future: while you (two) \...him/her/it}

# {lugu:lugu/tv-part-2p-3s-fut} {part. future: while you (many) \...him/her/it}

# {lugu:lugu/tv-part-2s-3s-fut} {part. future: while you \...him/her/it}

# {lugu:lugu/tv-part-4d-3s-fut} {part. future: while they (two) \...him/her/it}

# {lugu:lugu/tv-part-4p-3s-fut} {part. future: while they (many) \...him/her/it}

# {lugu:lugu/tv-part-4s-3s-fut} {part. future: while he/she/it \...him/her/it}

**:lugu\_tv\_le** a ontolex:Affix;

**ontolex:canonicalForm :lugu\_lugu\_f**;

ontolex:sense :lugu\_tv\_part\_fut;

ontolex:baseConstraint :verb; # we're doing verbal inflection here, so we can attach to a verbal base, only

ontolex:grammaticalMeaning

:tv\_1d\_3s, :tv\_1p\_3s, :tv\_1s\_3s, # these are alternative meanings

:tv\_2d\_3s, :tv\_2p\_3s, :tv\_2s\_3s,

:tv\_4d\_3s, :tv\_4p\_3s, :tv\_4s\_3s.

# Note: In this way, we cannot disambiguate forms for their different grammatical meanings

# If that would be intended, we would need to create one lexical entry per feature combination.

lexinfo:mood :verbal\_participle;

lexinfo:tense :future.

# Note: Inuktitut does not inflect for grammatical tense, but only for mood. Some moods have future readings, though.

:lugu\_tv\_part\_fut a ontolex:LexicalSense;

skos:definition "part. future: while s.o. does something to s.t. (object)"@en.

**:lugu\_lugu\_f** a ontolex:Form;

ontolex:writtenRep "lugu"@iu-Latn.

Note on incorporating verbs like :u\_1\_le

* Morphologically, these behave like affixes, but semantically, they are lexical verbs. These really need to be lexical entries.

Note on allomorphy:

* Except for the stem *atausi-*, all morphemes in this example happen to take the canonical form, but they have other allomorphic variants. By allowing ontolex:otherForm for the these variants, adding a new variant requires two triples (ontolex:otherForm and ontolex:writtenRep – the others can be inferred). Creating a separate lexical entry for *atausi* means that it and its sense have to be completely duplicated. Also, there is no direct link between the allomorphic variants. Same for -*ut-* and its variant -*u-*, for *-uq-* and its variant *-u-*, for *-liq-* and its variants *-siq-, -si-* and *-li-*, etc. (depending on following morpheme, final consonants can be assimilated. But this is described with the following morpheme, see <https://github.com/acoli-repo/morph-addenda/blob/master/data/polysynthetic/atausiulugu.morphs.ttl>).
* As for how that would be presented in a dictionary, see <https://www.inuktitutcomputing.ca/DataBase/index.php?lang=en&c=DefinitionDeSuffixe&m=liq%2F2nv> or <https://uqausiit.ca/sites/default/files/2020-04/Affix-Dictionary-V21.pdf>.

**Post-call addendum**

* As requested by Max, a comparison of the modelling with forms being rdfs:Seq of forms resp. Morphs under <https://github.com/acoli-repo/morph-addenda/tree/master/data/polysynthetic>
  + directory one-morph-with-multiple-forms/
    - incl. linking of contextual variant with canonical form (via otherForm)
    - 260 triples (all data)
    - Morph dictionary (atausiulugu.morphs.ttl): 202 triples
  + directory every-form-one-morph/
    - Information loss: no linking between different form variants yet
    - 312 triples (all data, +20%)
    - Morph dictionary (atausiulugu.morphs.ttl): 254 triples (+25%)

# 2. Grammatical meaning connection for InflectionType

What is InflectionType right now? <https://github.com/ontolex/morph/issues/11> Option 2?

# 3. Next call

**29.11.2023**

Agenda:

* Confirm that all the datasets are compatible
* Go through all the GH issues
* A quick look at the companion vocabulary (for regex rules and sound classes)