# Developing an OntoLex Morphology Module

**Telco: Jan 20, 2021**

**Participants:**

Bettina

Christian C

Max

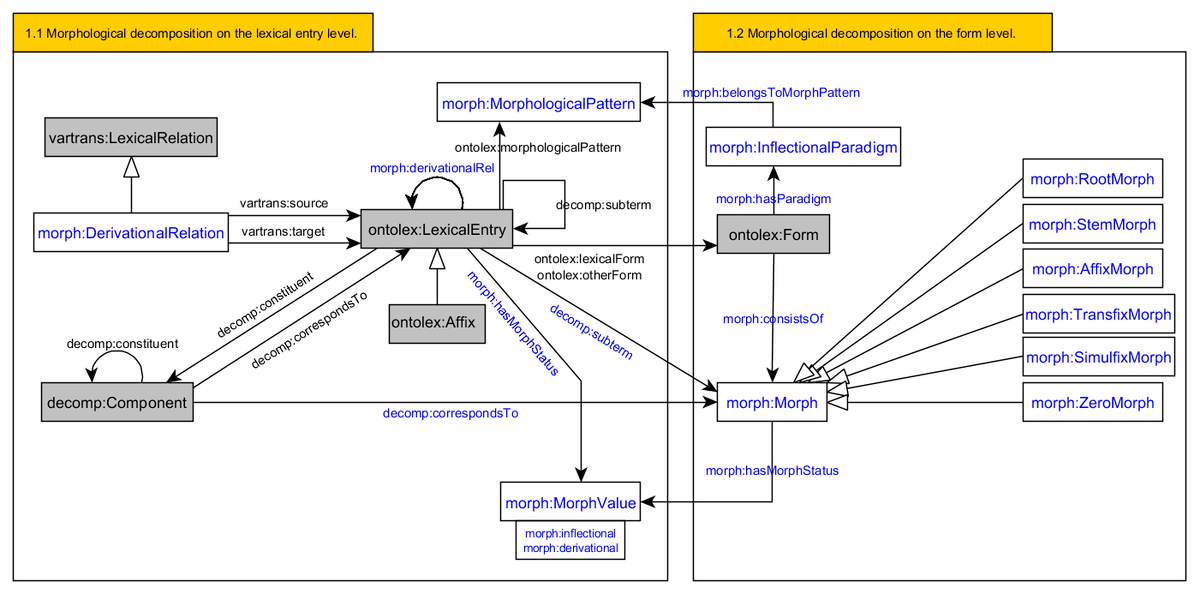
Matteo

Fahad Khan

Francesco

**Agenda**

1. **Status Quo**



Defined purpose and scope of the morphology module:

### **Purpose and scope**

The morphology module aims at fulfilling two modelling purposes:

**1.Stating elements that are involved in the decomposition of lexical entries and forms.**

1.1 Morphological decomposition on the lexical entry level.

scope: The kind of elements of which a lexical entry can consist should be as non-restrictive as possible. I.e. The decomposition of lexical entries encompass lexical entries, components, derivational affixes, inflectional affixes, stems, roots and zero morphs. However, a lexical entry can NEVER be composed of a form!

1.2 Morphological decomposition on the form level.

scope: Elements of which a form can consist include roots, stems, inflectional affixes and zero morphs. A form can NEVER be decomposed into lexical entries (including ontolex:Affix), components and forms.

Points 1.1 and 1.2 have been finished to a preliminary status according to the paper: <https://elex.link/elex2019/wp-content/uploads/2019/09/eLex_2019_33.pdf>

*Do we need to discuss all classes and properties again?*

*Shall I create a questionnaire to find out which class and property definitions need to be reevaluated?*

**2. Enabling the representation of building patterns that are involved in the formation of lexical entries and forms.**

2.1 Representation of decompositional building patterns for lexical entries.

2.2 Representation of decompositional building patterns for forms.

Point 2.2 was discussed with different examples but is not finished.

Point 2.1 still needs to be modeled.

Current status?: <https://docs.google.com/document/d/1nzcxiyBfKe5rRWrur-gmOzqFaginfydDPR7uFlPtjHo/edit?usp=sharing>

1. **Future Workflow**

Define new Workflow for finishing module development

1. Finish modeling proposal that addresses all modeling needs
   1. Finish collection and evaluation of modeling needs
   2. Finish modeling of purpose 2.1 and 2.2 according to modeling needs
2. Apply test datasets for every vocabulary element
3. Refine model
4. Repeat step 2 and 3 until each vocabulary element is used at least for modeling 2 resources
5. Create specification document
6. **Editing of the specification document**

<https://github.com/ontolex/morph/blob/master/index.html>

Who will contribute to the editing of the specification document?

Shall we start writing already? - slowly with Google doc for collaborative editing, later move to markup and git document - only a few people (to be decided) transfer the Google doc to its final git file.

[https://ontolex.github.io/morph/](https://meet.google.com/linkredirect?authuser=0&dest=https%3A%2F%2Fontolex.github.io%2Fmorph%2F)

1. **Remaining modeling needs**

Decide on N 1 - N20 here: <https://www.w3.org/community/ontolex/wiki/Morphology#Purpose_and_scope>

Collect more with regard to purpose 2.1 and 2.2?

**Tasks for next telco:**

* All:Collect more modeling needs with regard to purpose 2.1 and 2.2

Modelling needs to be added to be Wiki:

4 ways to represent infected word-forms

1. rules with transformations(Max)
2. Inflection Tables with full inflected word form(Max)
3. set of affixes/word endings (Max)
4. Assimilation rules as additional rules on top of 1.

* All: Collect test data here<https://drive.google.com/drive/folders/1hNOj5qm34J7uGJrsudI4eDuQAMCFFfzv?usp=sharing>
* Bettina: Prepare vocabulary description document\*<https://docs.google.com/document/d/1r6P9pTibsIy1iug1ptRtI-h9oVoA9N2wwDJJVaXt56s/edit?usp=sharing>

*\*edit this document to work on vocabulary and address these questions: Do we need to discuss all classes and properties again?*

*Shall I create a questionnaire to find out which class and property definitions need to be reevaluated?*

* Bettina: Update modelling needs and decision results in Wiki
* Max: prepare latest diagram of vocabulary for purpose 2.2 with respective vocabulary definitions
* Bettina: Prepare modelling need google doc to share
* Bettina: Prepare specification as Google doc to share with everyone who likes to be author
* Fahad + CC + Bettina: Bettina to write invitation mail to be sent out to NexusLinguarum to participate and share language resources
* Bettina: add morphological language types that are supposed to be covered with the morph module to the definition of the scope in the Wiki (most frequent morphological phenomena but also try to aim at representing languages which are morphologically very complex, e.g. CCs sample data for an eskimo language)

**Mail to Nexus Linguarum**

Dear Nexus Linguarum researchers,

we are reaching out to you to invite you to participate in the development of the new OntoLex-lemon Morphology Module.

After a longer break, the regular work on the module has been taken up again with the first telco last Wednesday. This module pursues to enable the representation of morphological language data by addressing computational linguistic as well as language descriptive needs. Moreover, it aims at allowing for the automatic generation of inflectional paradigms and word-forms of lexemes and the generation of lexemes that result from word-formation processes.

A preliminary report of the current modeling status can be consulted here:

<https://elex.link/elex2019/wp-content/uploads/2019/09/eLex_2019_33.pdf>

Right now we are working on the modeling draft for the automatic generation of inflected word-forms and derived lexemes.

Since there is considerable overlap between the OntoLex-lemon community and your effort we are hoping to work together in building a Morphology Module that fulfills the modeling needs current ontologies and vocabularies lack to express.

There are several ways of a collaborative exchange and participation in the module development:

1. You are welcome to join our bi-weekly telcos on Wednesday at 1-2 pm CET (the next telco will be on February, 3rd): <https://meet.google.com/rrs-garw-oxu?hs=122&authuser=0>
2. We would also like to encourage you to share language data with us that addresses a specific modeling need you would like to be representable with the Morphology Module or that you like to provide as test data.
3. Further ways for collaboration/participation?!

You can enter your representation needs into this document: <https://docs.google.com/document/d/1z1HjQ0_WmdQwdRoktSFPGewOkzcHesdJiz0O4CFX1Q8/edit?usp=sharing>.

For the contribution of test data please use the folder: <https://drive.google.com/drive/folders/1hNOj5qm34J7uGJrsudI4eDuQAMCFFfzv?usp=sharing>.

We believe that a collaboration between our research efforts will lead to more accurate language data as Linked Data and are looking forward to hearing back from you soon.

Kind regards,

Christian, Fahad and Bettina