

# ***Dalin morphology***

## **1. BASIC INFORMATION**

### *1.1 Lexicon type*

Computational lexicon

### *1.2 Representation of the lexicon*

LMF

### *1.3 Character encoding*

The character encoding is in UTF-8

## **2. ADMINISTRATIVE INFORMATION**

### *2.1 Contact person (name, e-mail)*

Name: Markus Forsberg

E-mail: sb-info@svenska.gu.se

### *2.2 Copyright statement and information on IPR*

Open source: CC-BY-SA 3.0, LGPL 3.0

## **3. TECHNICAL INFORMATION**

### *3.1 Data structure of an entry*

Structured according to the LMF specification

(<http://www.lexicalmarkupframework.org/>).

### *3.2 Lexicon size*

62 797 lexical entries, 730 571 word forms

## **4. CONTENT INFORMATION**

### *4.1 The natural language(s) of the lexicon*

Swedish

### *4.2 Entry Type*

LexicalEntry in LMF.

### *4.3 Attributes*

The resource is a contains morphological information for the Dalin lexicon.

### *4.4 Coverage of the lexicon*

The resource enrich a large part of the dictionary with full morphological information, and use the baseform for the rest of the entries.

### *4.5 Intended application of the lexicon*

The intended use is language technology applications.

### *4.6 Reliability (automatically/manually constructed)*

The morphological description is done manually, but the assignment of inflectional information to entries in the Dalin lexicon has been done automatically. The quality needs improvement.

## **8 RELEVANT REFERENCES AND OTHER INFORMATION**

Lars Borin, Markus Forsberg 2011. [A diachronic computational lexical resource for 800 years of Swedish](#) *Language technology for cultural heritage*, Springer. 41-61