3 The Core Specification

The DMLex Core is for monolingual lexical resources, where headwords, definitions, examples etc. are all in one and the same language.

3.1 LexicographicResource object type

A data set which can be viewed and used by humans as a dictionary and - simultaneously - ingested, processed and understood by software agents as a machine-readable database.

**Warning**

The correct name of this data type in DMLex is *lexicographic resource*, not *lexical*.

**Properties:**

- **language** (optional, IETF language code)
  - The language of headwords, definitions, examples.

  **Note**

  DMLex is based on the assumption that all headwords in a single lexicographic resource are in the same language, and that definitions and examples, if any occur in the resource, are in that language too. The *language* property informs potential users of the resource which language that is.

- **transcriptionScheme** (optional, IETF language code, eg. en-fonipa for English IPA)
  - The scheme (e.g. IPA) in which the *transcription* property of *Pronunciation* objects is given.

  **Note**

  DMLex is based on the assumption that, if you do use any pronunciation transcriptions in your (monolingual) lexicographic resource, they all follow the same scheme. The *transcriptionScheme* tells potential users of the lexicographic resource which scheme that is.

**Children:**

- **Entry** (one or more)

*Example 1. Example of a lexicographic resource*

```xml
LexicographicResource: language="en"
  Entry: headword="aardvark"
    Sense: ...
    Sense: ...
  Entry: headword="abacus"
    Sense: ...
    Sense: ...
  Entry: headword="abandon"
    Sense: ...
    Sense: ...
```

3.2 Entry object type

A part of a lexicographic resource which contains information related to exactly one headword.

Child of:
LexicographicResource

Properties:

- headword (non-empty string)
  - The headword can be a single word, a multi-word expression, or any expression in the source language which is being described by the entry in the lexicographic resource.

- homographNumber (number, optional)
  
  **Note**

  Entries do not have an explicit listing order. An application can imply a listing order from a combination of the headword and the homograph number.

Children:

- PartOfSpeech (zero or more)
- Usage (zero or more)
- Pronunciation (zero or more)
- InflectedForm (zero or more)
- Sense (zero or more)

**Example 2. Example of a lexicographic entry**

Entry: headword="bank", homographNumber=1
PartOfSpeech: ...
Pronunciation: ...
Sense: ...
Sense: ...

3.3 Sense object type

A part of an entry which groups together information relating to one of the (possibly multiple) meanings (or meaning potentials) of the entry’s headword.

**Note**

An entry is a container for formal properties of the headword such as orthography, morphology, syntax and pronunciation. A sense is a container for statements about the headword’s semantics.

Child of:

- Entry

Properties:

- listingOrder
  - Can be implicit from the serialization.

- indicator (optional, non-empty string)
  - A short statement that indicates the meaning of a sense and permits its differentiation from other senses in the entry.

  **Note**
Indicators are sometimes used in dictionaries as "mini-definitions" instead of or in addition to regular definitions. They are short, one-word or two-word paraphrases of the sense. Their purpose is to allow the (human) user to find the desired sense quickly.

- **definition** (optional, non-empty string)
  - A long statement that describes and or explains the meaning of a sense.

**Note**

In DMLex, the term *definition* encompasses not only formal sciency definitions, but also less formal explanations.

Children:
- **Usage** (zero or more)
- **Example** (zero or more)

**Example 3. Example of an entry with two senses, each having both an indicator and a definition**

<table>
<thead>
<tr>
<th>Entry</th>
<th>headword=&quot;bank&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sense</td>
<td>indicator=&quot;financial institution&quot;, definition = &quot;an institution which...&quot;</td>
</tr>
<tr>
<td></td>
<td>Example: ...</td>
</tr>
<tr>
<td>Sense</td>
<td>indicator=&quot;riverside&quot;, definition = &quot;an earthen mound which...&quot;</td>
</tr>
<tr>
<td></td>
<td>Example: ...</td>
</tr>
</tbody>
</table>

### 3.4 PartOfSpeech object type

Any of the word classes to which a lexical item may be assigned, e.g. noun, verb, adjective, etc.

Child of:
- **Entry**

Properties:
- **value** (non-empty string)
  - Can be constrained and mapped to external authorities by the DMLex Controlled Vocabularies Module.
- **listingOrder**
  - Can be implicit from the serialization.

**Note**

If you want to model other grammatical properties of the headword besides part of speech, such as gender (of nouns) or aspect (of verbs), the recommended way to do that in DMLex is to conflate them to the part-of-speech label, for example "masculine noun" and "feminine noun", or "perfective verb" and "imperfective verb".

**Example 4. Example of a part of speech object**

<table>
<thead>
<tr>
<th>Entry</th>
<th>headword=&quot;aardvark&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>PartOfSpeech</td>
<td>value=&quot;noun&quot;</td>
</tr>
</tbody>
</table>

### 3.5 Usage object type

An indication of some restriction on the use of the lexical item such time (old-fashioned, neologism), region (dialect), register (formal, colloquial), domain (medicine, politics) or grammar (singular-only).
Child of:

- Entry
- Sense
- Pronunciation
- InflectedForm

Properties:

- value (non-empty string)
  - Can be constrained and mapped to external authorities by the DMLex Controlled Vocabularies Module.
- listingOrder
  - Can be implicit from the serialization.

Example 5. Example of a usage label

Entry: headword="bitch"
  Sense: definition="a female dog"
  Sense: definition="a woman who is difficult to get along with"
  Usage: value="slang"

3.6 Pronunciation object type

Information about the pronunciation of its parent.

Child of:

- Entry
- InflectedForm

Properties:

- At least one of:
  - transcription (non-empty string)
    
    **Note**
    
    The scheme of the transcription (e.g. IPA) can be communicated to users of the lexicographic resource through the transcriptionScheme property of the LexicographicResource object.
  - recording (string: name or URL of a sound file)
- listingOrder
  - Can be implicit from the serialization.

Children:

- Usage (zero or more)

Example 6. Example of pronunciation given as transcription

Entry: headword="aardvark"
  Pronunciation: transcription="a:rdva:rk"

Example 7. Example of pronunciation given as a sound file
3.7 InflectedForm object type

An inflected headword is a form of the inflectional paradigm of its parent.

Child of:
- Entry

Properties:
- label (non-empty string) e.g. 'plural'
  - Can be constrained and mapped to external authorities by the DMLex Controlled Vocabularies Module.
- value (non-empty string)
- listingOrder
  - Can be implicit from the serialization.

Children:
- Usage (zero or more)
- Pronunciation (zero or more)

3.8 Example object type

A sentence or other text extract which illustrates the headword being used in the sense.

Child of:
- Sense

Properties:
- text (non-empty string)
- listingOrder
  - Can be implicit from the serialization.

3.9 SubsenseRelation relation type
Represents the fact that one sense (the subordinate sense) should be treated as a subsense of another sense (the subordinate). Both senses belong to the same entry.

Participants:
- the superordinate Sense (exactly one)
- the subordinate Sense (exactly one)

3.10 SubentryRelation relation type

Represents the fact that one entry (= the subordinate entry) should be treated as a subentry inside the sense (= the superordinate sense) of another entry.

Participants:
- the superordinate Sense (exactly one)
- the subordinate Entry (exactly one)

Properties:
- listingOrder
  - Can be implicit from the serialization.

4 Bilingual Module

Extends DMLex Core to support the encoding of bilingual lexicographic resources.

4.1 Extensions to LexicographicResource object type

Additional properties:
- translationLanguage (optional, IETF language code)
- translationTranscriptionScheme (optional, reference to some external authority - which?)
  - The scheme (e.g. IPA) in which the transcription property of TranslationPronunciation objects is given.

4.2 Extensions to Sense object type

Additional children:
- HeadwordTranslation (zero or more)

4.3 HeadwordTranslation object type

The translation equivalent of the headword in one of its senses.

Child of:
- Sense

Properties:
- text (non-empty string)
  - Can be a single word, a multi-word expression, or indeed any expression in the target language.
4.4 TranslationPartOfSpeech object type

Any of the word classes to which the translation may be assigned, e.g. noun, verb, adjective, etc.

Child of:
- HeadwordTranslation

Properties:
- value (non-empty string)
  - Can be constrained by the DMLex Controlled Vocabularies Module.
- listingOrder
  - Can be implicit from the serialization.

4.5 TranslationUsage object type

An indication of some restriction on the use of its parent. The restriction can be pragmatic (time, region, register), semantic (domain, semantic type) or formal ('no plural').

Child of:
- HeadwordTranslation
- TranslationPronunciation
- TranslationInflectedForm

Properties:
- value (non-empty string)
  - Can be constrained by the DMLex Controlled Vocabularies Module.
  - Its type (eg. whether register, temporal, geographic etc) can be specified by the DMLex Controlled Vocabularies Module.
- listingOrder
  - Can be implicit from the serialization.

4.6 TranslationPronunciation object type

Information about the pronunciation of its parent.

Child of:
- HeadwordTranslation
• TranslationInflectedForm

Properties (at least one):

• transcription (non-empty string)
• recording (string: name or URL of a sound file)
• listingOrder
  ○ Can be implicit from the serialization.

Children:

• TranslationUsage (zero or more)

4.7 TranslationInflectedForm object type

A form of the inflectional paradigm of its parent.

Child of:

• HeadwordTranslation

Properties:

• label (non-empty string) e.g. 'plural'
  ○ Can be constrained by the DMLex Controlled Vocabularies Module.
• value (non-empty string)
• listingOrder
  ○ Can be implicit from the serialization.

Children:

• TranslationUsage (zero or more)
• TranslationPronunciation (zero or more)

4.8 HeadwordTranslation object type

The translation equivalent of the headword in one of its senses.

Child of:

• Sense

Properties:

• text (non-empty string)
  ○ Can be a single word, a multi-word expression, or indeed any expression in the target language.
• listingOrder
  ○ Can be implicit from the serialization.

Children:

• TranslationPartOfSpeech (zero or more)
• TranslationUsage (zero or more)
4.9 Extensions to Example object type

Additional children:
- ExampleTranslation (zero or more)

4.10 ExampleTranslation object type

The translation of an example.

Child of:
- Example

Properties:
- text (non-empty string)
- listingOrder
  - Can be implicit from the serialization.

5 Internal Linking Module

Used to construct clickable (or otherwise navigable) hyperlinks between objects within the same LexicographicResource.

5.1 EntrySetRelation relation type

Represents an unordered set of two or more entries. Recommended for encoding variants such for example masculine and feminine counterparts or spelling variants.

Properties:
- label (non-empty string) eg. "synonyms"
  - Can be constrained and mapped to external authorities by the DMLex Controlled Vocabularies Module.

Participants:
- Entry (two or more, all in the same LexicographicResource)

5.2 SenseSetRelation relation type

Represents an unordered set of two or more senses (typically - but not necessarily - belonging to two or more different entries). Recommended for encoding synonyms, near synonyms etc.

Properties:
- label (non-empty string) eg. "synonyms"
  - Can be constrained and mapped to external authorities by the DMLex Controlled Vocabularies Module.

Participants:
- Sense (two or more, all in the same LexicographicResource)
5.3 SensePairRelation relation type

Represents an undirected pair of two senses (typically - but not necessarily - belonging to two different entries). Recommended for representing pairs of antonyms, opposites, converses and so on.

Properties:
- label (non-empty string) eg. "antonyms"
  - Can be constrained and mapped to external authorities by the DMLex Controlled Vocabularies Module.

Participants:
- Sense (exactly two, both in the same LexicographicResource)

5.4 SenseTupleRelation relation type

Represents a directed pair of two senses (typically - but not necessarily - belonging to two different entries). Recommended for representing pairs of hypernym/hyponym, broader/narrower, holonym/meronym.

Properties:
- label (non-empty string) eg. "hypernym/hyponym"
  - Can be constrained and mapped to external authorities by the DMLex Controlled Vocabularies Module.

Participants:
- the first Sense of the pair
- the second Sense of the pair

6 External Linking Module

Used to construct links between objects from different lexicographic resources.

6.1 SameRelation relation type

Represents the fact that two or more senses from two or more lexicographic resources are semantically the same.

Participants:
- Sense (two or more, no two of which belong to the same LexicographicResource)

6.2 SubsumptionRelation relation type

Represents the fact that one sense from one lexicographic resource is semantically wholly contained inside another sense from another lexicographic resource.

Participants:
- the broader Sense (exactly one)
- the narrower Sense (exactly one)

The two senses must not be within the LexicographicResource.
6.3 OverlapRelation relation type

Represents the fact that two or more senses from two or more lexicographic resources overlap semantically, without either one being semantically wholly contained inside the other.

Participants:

- Sense (two or more, no two of which belong to the same LexicographicResource)

7 Inline Markup Module

7.1 PlaceholderMarker marker type

Marks up a substring inside a headword (or inside a headword translation) which is not part of the expression itself but stands for things that can take its place, or constitutes some kind of meta-notation. Examples:

- beat [sb.] up
- continue [your] studies

Markup of:

- headword property of Entry
- text property of HeadwordTranslation

7.2 HeadwordMarker marker type

Marks up a substring inside an example (or inside an example translation) which corresponds to the headword (or to a translation of the headword).

Markup of:

- text property of Example
- text property of ExampleTranslation

8 Controlled Vocabularies Module

This module makes it possible to describe constraints on the values of certain plain-text properties of objects defined in DMLex Core and in DMLex Bilingual Module.

8.1 Extensions to LexicographicResource object type

Additional properties:

- labelLanguage (IETF language code)
  - The language of the display values of labels.

Additional children:

- PartOfSpeechLabel (zero or more)
- TranslationPartOfSpeechLabel (zero or more)
- UsageLabel (zero or more)
- EntrySetLabel (zero or more)
• SenseSetLabel (zero or more)
• SensePairLabel (zero or more)
• SenseTupleLabel (zero or more)
• TranslationUsageLabel (zero or more)
• InflectedFormLabel (zero or more)
• TranslationInflectedFormLabel (zero or more)

8.2 PartOfSpeechLabel object type

A PartOfSpeechLabel represents one of several allowed values for the value property of PartOfSpeech objects.

Properties:
- value (non-empty string)
- displayValue (optional)

Children:
- LabelMapping (zero or more)

Example 11. Example: restricting part-of-speech values

```xml
LexicographicResource: labelLanguage="en"
  PartOfSpeechLabel: value="n", displayValue="noun"
  PartOfSpeechLabel: value="v", displayValue="verb"
  PartOfSpeechLabel: value="adj", displayValue="adjective"
  Entry: headword="aardvark"
    PartOfSpeech: value="n"
  Entry: headword="speak"
    PartOfSpeech: value="v"
  Entry: headword="small"
    PartOfSpeech: value="adj"
```

Example 12. Example: mapping part-of-speech values to external inventories

```xml
PartOfSpeechLabel: value="n", displayValue="noun"
  LabelMapping: sameAs="http://www.lexinfo.net/ontology/3.0/lexinfo#noun"
```

Example 13. Example: mapping complex part-of-speech values to external inventories

```xml
PartOfSpeechLabel: value="Nma", displayValue="masculine animate noun"
  LabelMapping: sameAs="http://www.lexinfo.net/ontology/3.0/lexinfo#noun"
  LabelMapping: sameAs="http://www.lexinfo.net/ontology/3.0/lexinfo#masculine"
  LabelMapping: sameAs="http://www.lexinfo.net/ontology/3.0/lexinfo#animate"
```

8.3 TranslationPartOfSpeechLabel object type

A TranslationPartOfSpeechLabel represents one of several allowed values for the value property of TranslationPartOfSpeech objects.

Properties:
- value (non-empty string)
- displayValue (optional)

Children:
- LabelMapping (zero or more)
8.4 UsageLabel object type

A UsageLabel represents one of several allowed values for the value property of Label objects.

Properties:

- value (non-empty string)
- displayValue (optional)

Children:

- LabelMapping (zero or more)

Example 14. Example: restricting the values of usage labels

LexicographicResource: labelLanguage="en"
UsageLabel: value="sl", displayValue="slang"
UsageLabel: value="vul", displayValue="vulgar"

Example 15. Example: mapping the values of usage labels to external ontologies

LexicographicResource: labelLanguage="en"
UsageLabel: value="sl", displayValue="slang"
LabelMapping: sameAs="http://www.lexinfo.net/ontology/3.0/lexinfo#slangRegister"
UsageLabel: value="vul", displayValue="vulgar"
LabelMapping: sameAs="http://www.lexinfo.net/ontology/3.0/lexinfo#vulgarRegister"

8.5 EntrySetLabel object type

Represents one of several allowed values for the label property of EntrySet objects.

Properties:

- value (non-empty string)
- displayValue (optional)

Children:

- LabelMapping (zero or more)

8.6 SenseSetLabel object type

Represents one of several allowed values for the label property of SenseSet objects.

Properties:

- value (non-empty string)
- displayValue (optional)

Children:

- LabelMapping (zero or more)

8.7 SensePairLabel object type

Represents one of several allowed values for the label property of SensePair objects.

Properties:

- value (non-empty string)
- displayValue (optional)
Children:
- LabelMapping (zero or more)

### 8.8 SenseTupleLabel object type

Represents one of several allowed values for the label property of SensePair objects.

Properties:
- value (non-empty string)
- displayValue (optional)

Children:
- LabelMapping (zero or more)

### 8.9 TranslationUsageLabel object type

A TranslationUsageLabel represents one of several allowed values for the value property of TranslationLabel objects.

Properties:
- value (non-empty string)
- displayValue (optional)

Children:
- LabelMapping (zero or more)

### 8.10 InflectedFormLabel object type

An InflectedFormLabel represents one of several allowed values for the label property of InflectedForm objects.

Properties:
- value (non-empty string)
- displayValue (optional)

Children:
- LabelMapping (zero or more)

*Example 16. Example: restricting the labels of inflected forms*

```plaintext
LexicographicResource: language="ga", labelLanguage="en"
PartOfSpeechLabel: value="nf", displayValue="feminine noun"
InflectedFormLabel: value="gs", displayValue="genitive singular"
InflectedFormLabel: value="pl", displayValue="plural"
Entry headword="deirfiúr"
    PartOfSpeech: value="nf"
    InflectedForm: label="gs", value="deirféar"
    InflectedForm: label="pl", value="deirfiúracha"
```

*Example 17. Example: mapping the labels of inflected forms to external inventories*

```plaintext
InflectedFormLabel: value="gs", displayValue="genitive singular"
LabelMapping: sameAs="http://www.lexinfo.net/ontology/3.0/lexinfo#singular"
InflectedFormLabel: value="pl", displayValue="plural"
LabelMapping: sameAs="http://www.lexinfo.net/ontology/3.0/lexinfo#plural"
```
8.11 **TranslationInflectedFormLabel object type**

A TranslationInflectedFormLabel represents one of several allowed values for the label property of TranslationInflectedForm objects.

Properties:

- value (non-empty string)
- displayValue (optional)

Children:

- LabelMapping (zero or more)

8.12 **LabelMapping object type**

Represents the fact that an item in the controlled vocabulary is equivalent to item provided by an external authority.

Parents:

- PartOfSpeechLabel
- TranslationPartOfSpeechLabel
- UsageLabel
- EntrySetLabel
- SenseSetLabel
- SensePairLabel
- SenseTupleLabel
- TranslationUsageLabel
- InflectedFormLabel
- TranslationInflectedFormLabel

Properties:

- sameAs (URI)