User Documentation of

Cornetto LMF

Lexical Resource for Dutch

Isa Maks, Hennie van der Vliet, Attila Görög, Piek Vossen
VU University Amsterdam

May 2013

CLARIN project CLARIN-NL-11-020
Deliverable D9

Project CORNETTO-LMF-RDF

Table of Contents

1 Introduction

Cornetto is a lexical resource for Dutch (www.tst.nl: Cornetto) including a set of lexical entries and a set of synsets related to each other. The present version (May, 2013) includes:

- Synsets
- Lexical entries: word –sense combinations
- Sense Relations
- Sentiment Information
- References to English WordNet
- References to SUMO
- Refernces to wordNet Domains

The lexicon is modeled according to ISO standard Lexicon Markup Framework (LMF) (LMF: ISO 24613: 2008) a metamodel for the description of lexical resources. The data categories are selected from the ISOcat datacategory registry (www.isocat.org) which is a central registry for concepts relevant in linguistics and in the domain of language resources[1].

This document is structured as follows: Section 2 gives a short introduction to the Cornetto Resource; Sections 3 – 8 describe the most important linguistic concepts used in Cornetto; Section 9 gives the complete list of Cornetto's data categories and definitions and Section 10 presents the Cornetto-LMF data model and the design for the XML data file.

If you use the Cornetto Resource, then please cite one the following papers:

Vossen, P., I.Maks, R. Segers and H. van der Vliet (2008). *Integrating Lexical Units, Synsets, and Ontology in the Cornetto Database*. In Proceedings of LREC-2008, Marrakech, Morocco.

Vossen, P., I.Maks, R.Segers, H.van der Vliet, M.F. Moens, K.Hofmann, E.Tjong Kim Sang, M.de Rijke (2013) *Cornetto: a lexical semantic database for Dutch*, Chapter in: P. Spyns & J. Odijk (eds): Essential Speech and Language Technology for Dutch, Results by the STEVIN-programme, Publ. Springer series Theory and Applications of Natural Language Processing, ISBN 978-3-642-30909-0.

[1] The conversion of Cornetto XML, i.e.the original resource, into Cornetto LMF XML is carried within the CLARIN (www.clarin.eu) project Cornetto-LMF-RDF (2012-2013)

2 Description of the Cornetto Lexical Resource

2.1 Short overview of Cornetto

The Cornetto lexical resource for Dutch covers the most generic and central part of the language. Cornetto combines the structures of the Princeton Wordnet, some of the features from the FrameNet for English and the information on morphological, syntactic, semantic and combinatorial features of lexemes normally found in dictionaries. The Cornetto resource is compiled by combining and aligning two existing semantic resources for Dutch: the Dutch wordnet (DWN) (Vossen 1998) and the Referentie Bestand Nederlands (RBN) (Martin et al. 1999). Recently, the resource is revised and extended with sentiment values in the From Text to Political Positions project , and with semantic annotations in SONAR, CGN and texts from the Web in the DutchSemCor project. This documentation describes the release of Cornetto of May 2013 (version 2.1).

The Cornetto Lexical Resource consists of two large repositories of lexicon data: the lexical entry repository and the synset repository. A Lexical Entry (LE) is a word-meaning pair (i.e. a single meaning of a certain word form), for which morphological, syntactical, semantical and combinatorial information is given. As such, LEs are word senses in the lexical semantic tradition, containing the linguistic knowledge that is needed to properly use the word in a specific meaning in a language. Since the LEs follow a word-to-meaning view, the semantical and combinatorial information for each meaning clarify the differences across the meanings. LEs focus on the polysemy of words and typically follow an approach to represent condensed and generalised meanings from which more specific ones can be derived.

Each LE is aligned with a synset (set of synonyms) in the synset repository. As such, a synset can be seen as a set of LEs with the same meaning and every synset stands for a concept. The synsets in Cornetto are interconnected by different semantic relations such as hyponymy, antonymy and meronymy. The Cornetto Resource is aligned with the English Wordnet, from which domain information was imported. The domains represent clusters of concepts that are related by a shared area of interest, such as *sport*, *education* or *politics*.
The data structures are represented in Figure 1.

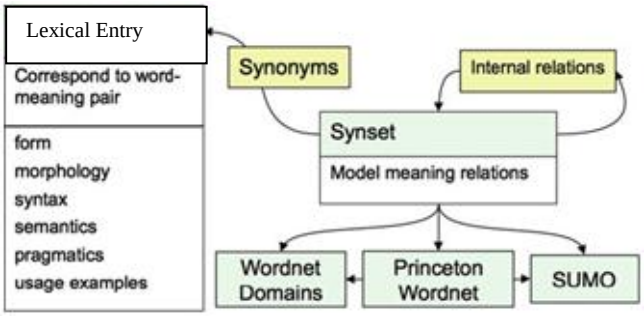

Figure 1: The Cornetto data structures

The definitions of LEs from the same synset should be semantically equivalent and the LEs of a single word form should belong to different synsets. The LEs of a single word form typically differ in terms of connotation, pragmatics, syntax *and* semantics but synonymous words in the same synset can be differentiated along connotation, pragmatics and syntax but not semantics. This structure of the resource makes it possible to combine the very detailed information on form and usage of a specific LE or a group of LEs with the semantic relations which are specified in the corresponding synset(s).

2.2 Statistics

The current version of Cornetto (version 2.1) includes the following items:

Lexical entries	127,334
Noun Lexical entries	85,368
Verb Lexical entries	16,502
Adjective Lexical entries	15,458
Multiword Lexical Entries	9,397
Lexical Entries with polarity labels	22,202
Sense group relations	11,812
Sense examples	80,512
Lexical collocations	19,166
Grammatical collocations	10,373
Synsets	70,497
Synset relations (internal relations between Dutch synsets)	91,734
Synset equivalent relations (between Dutch and English Princeton WordNet synsets)	84,031
References to SUMO (from synsets to SUMO)	69,610
References to WordNet domains (from synsets to WordNet domains)	93,165

Table 1: statistics of Lexical Resource Cornetto (version 2.1, May 2013)

2.3 Identifier Formats

Cornetto identifiers have a fixed format.

Lexical Entry id(entifier)

Format	Lemma _ part-of-speech_sequencenumber
Example	huis_n_2
Explanation	part-of-speech is an abbreviation of the part-of-speech used in Cornetto. Its values are n(noun), v(verb) or a(djective). lemma is the headword sequence number is the number of the sense of the lemma. Often, though not always , the lower the sequence number the more frequent or common the sense.

Synset id(entifier)

Format	language_versionnumber_xxx_part-of-speech
Example	nld-21-d_v-1541-v
Explanation	part-of-speech is an abbreviation of the part-of-speech used in Cornetto. Its values are n(noun), v(verb) or a(djective). language is the language described in the resource : Dutch (nld) xxx is a meaningless part of the id, It may contain alphabetic and numerical xxxx as well as hyphens (-) and uderschores (_). It is composed of identifiers of earlier versions of Cornetto but has no meaning in the current version. the version number refers to the version of Cornetto ; the latest version (May 2013) is 2.1 (21)

Other identifiers, such as sense example id, and sense id do not have a fixed format

3 Lexical Entry

The most important data categories of the lexical entry (LE) are described in this section. Examples illustrate what kind of information can be found and how this information is distributed among the different lexicons layers like syntax, morphology and semantics.
In Cornetto, a lexical entry (LE) is a unique pair of a word form and only one meaning. In case of polysemy, each combination of the word form and one of its meanings is described in a separate LE. A lexical entry contains information on the lemma and word form, the morphology, the syntax, the semantics and the pragmatics, as well as on its combinatorial behaviour. In this section we will give an overview and highlight some characteristic aspects of Cornetto's lexical entries. The complete set of data categories and their values can be found in section 9.

3.1 Lemma and Word Forms

A lexical entry (LE) consists of **lemma, form type, part-of-speech** (noun, verb or adjective) and information on **word forms**. For all parts-of-speech the same categories of lexical information can be found, but of course, details will be different for different parts of speech, as will become clear in this section. Whenever information is specific for a specific part of speech, this will be explicitly mentioned.
Form types can be contractions (*bus* from *autobusl*), acronyms (like *NATO*) or abbreviations (like *etc*.).
LEs may have (a set of) **word forms** with different attributes. Attributes of the word form are:

- one or more plural forms (for nouns and verbs)
- the comparative and superlative forms (for adjectives)
- past tense and past participle, if not regular formed (for verbs)
- related forms, like form variants and spelling variants

An example of a form variant is *droppel(drop)* as a variant of the word form *druppel(drop)*.

3.2 Morphology and morphosyntax

Morphological information on the LE can be found in the data category (**morphoType**) morphological type. The morphological type can for instance be a simplex, a compound, an idiomatical compound, a derivation, a word group or a zero-derivation (as for the noun *loop* which is derived from the verb *lopen* (*to walk*)). For adjectives, the morphological type is augmented by **comparison type** (regular or irregular) and **declinability**, for verbs by **separability**. The adjective *klinisch* (clinical) is a derivative and is declinable. An example of a separable phrasal verb is *doorwerken*, because of *ik werk door*.
The morphosyntactic information on nouns is on **pronominal and grammatical gender**. In Dutch, there are two definite articles (*de* and *het*) for three types of grammatical gender: *de* for masculine and feminine nouns and *het* for neuter nouns. Most speakers of Dutch are uncertain about the grammatical gender of *de*-words unless the referent has biological (male of female) gender. Pronominal gender, however, has pronouns for masculine (*hij, hem, zijn*), feminine (*zij, haar*) and neuter (*het, zijn*). As a result, speakers of Dutch are uncertain about the personal pronoun for referring to, for example, *de schaar (scissors)* which is of feminine grammatical gender; should this be *hij* or *zij*? Almost all speakers of Dutch[2] , however, use the personal pronoun *hij* to refer to *de*-words, like *schaar*, whether they are masculine or feminine.

[2] In Belgium, the distinction between the three genders is more often maintained. Words that are of feminine grammatical gender are usually still referred to with *zij*.

In Cornetto, the following values are used to indicate grammatical and pronominal gender of nouns.

name	**explanation**
m	grammatical and pronominal gender are m (masculine)
f	*hij/hem/zijn* or *zij/haar* (*schaar, originally female gender*) ; pronominal gender is m (masculine); grammatical gender is f (feminine)
mf	*zij/haar* (*dame* (lady)); biological, grammatical and pronominal gender are f (feminine)
n	*het/zijn* (*huis* (house)) ; grammatical and pronominal gender are n (neuter)
mfn	*hij/hem/zijn* or *zij/haar* or *het/zijn (kind* (child)) pronominal gender is m (masculine) or f (feminine) depending on the biological gender, or n (neuter) ; grammatical gender is n (neuter)
mn	*hij/hem/zijn* of *het/zijn (joch (lad*)) pronominal gender is m (masculine) ; grammatical gender is n (neuter)
fn	*het/zijn* of *zij/haar* (*meisje* (girl)) pronominal gender is feminine (f) ; grammatical gender is neuter (n)
m_f	*hij/hem/zijn* of *zij/haar* (b.v. *dokter* (*doctor*)), pronominal gender is m (masculine) or f (feminine) depending on the biological gender

Table 2: grammatical and pronominal gender values for nouns

The morphosyntax for verbs encodes for the auxiliary in the pefect tense, *hebben*, *zijn* or both, and for reflexivity (*zich vergissen*).

Adjectives are labeled for adverbial usage, i.e. whether they can be used in adverbial position. For instance, slecht (*bad, badly*) can be used both as an adjective, like in *een slechte dag* (*a bad day)* and as an adverb like in *hij gedraagt zich slecht* (*he behaves badly).* Adjectives are also encoded for possible restrictons on their use in attributive or predicative position. For example, *militair (*military) can occur in attributive position only, as in *militair vliegveld (military airport*) and not in predicative position (* *het vliegveld is militair* –* *the airport is military*). Other adjectives can be used in both positions.

3.3 Syntax, Syntactic behaviour and syntactic subcategorization

Syntactic behaviour encodes the characteristic combinatorical properties of the LE. Values for **complementation** of nouns are:

name	**explanation**
psmodnoun	a noun that can be followed by a postmodifying noun (*fles => een fles wijn(a bottle of wine*))
factive	a sentential complement with *dat* (*problem => het probleem dat we niet genoeg koffie hebben (the problem that we do not have enough coffee)).*
ofclause	*de vraag of …* (*the question whether*)
whclause	*het is de vraag waarom …*(*it is the question why*)
toinf	an infinitive clause introduced by *te*: *zijn bewering te zullen komen (his claim to come)*
omtoinf	like in **toinf**, but with optional *om*: *zijn poging (om) te fietsen (his attempt to ride a bicycle)*
fixprep	fixed preposition cannot be predicted just by its meaning: *een gevoel van angst (a feeling of fear)*
prep	prepositions that are not fixed, but clearly associated with the noun or adjective and frequently used

Table 3: complementation of nouns and adjectives

The syntactic behaviour of verbs is described by the **valency** i.e. the number of arguments the verb takes, and **transitivity** (transitive or intransitive). Explicit and detailed complementation patterns of the verbs are described in the **syntactic subcategorization frame**. The subcategorization frame matches the constituents a verb possibly takes as an argument (NP, PP, Subclause, AP) in combination with their function (like direct object, indirect object and prepositional object). The frame may be completed with an optional complementizer introducing the Subclause or an optional preposition introducing the PP. The verb *vitten (carp)*, for example, takes the constituent PP with the function of prepositional object, like *hij vit op haar* (*he carps at her*). The verb may be assigned multiple patterns. For example, one of the senses of the verb *aangeven (to report)* takes 2 patterns:

- a NP (constituent) as a direct object (function) e.g. *iets/iemand aangeven (advise something/somebody*) like in *Ik wil de diefstal aangeven (I want to report the theft)*
- a NP (constituent) as a direct object (function) in combination with an PP (constituent) as a prepositional object (function) e.g. *iets/iemand aangeven bij iemand (report somebody/something to somebody)* like in *Ik wil hem aangeven bij de politie (I want to report him to the police)*

The following fragment represents the syntactic subcategorization frame of *aangeven (report)* in XML format.

```
<SyntacticBehaviour valency="tri" transitivity="transitive">
    <SyntacticSubcategorisationFrame>
      <syntacticArgument constituent="np" function="directObject"/>
      <syntacticArgument constituent="pp" function="prepositionalObject"/>
    </SyntacticSubcategorisationFrame>
    <SyntacticSubcategorisationFrame>
      <syntacticArgument constituent="np" function="directObject"/>
    </SyntacticSubcategorisationFrame>
  </SyntacticBehaviour>
```

3.4 Sense and semantics

The datacatagories describing the lexical semantic properties are Sentiment (cf. section 6), Pragmatics (cf. section 3.6), Semantics (cf. section 3.4) and SenseExamples (cf. section 3.5).
The values for Semantics differ for each separate part-of-speech. The information for nouns is on countability, reference (proper name or common) and semantic type. Eleven semantic types are selected: human, nonhuman, time, place, artifact, dynamic, non-dynamic, substance, abstract, concrete, concrother. Polysemy can partially be described as a systematic shift in semantic type. As such the noun *fabriek (factory)* is categorized as artefact, and can shift to place and institution.

artefact:	*een fabriek bouwen*	*(to build a factory)*
place*:*	*in een afbriek werken*	*(to work at a factory)*
institution*:*	*de fabriek is gesloten*	*(the factory is closed)*

Adjectives are described with a similar set of semantic types and shifts and verbs are described with the types *action*, *process* and *state*.
In addition, verbs refer to a semantic feature set (**semanticFeatureSet**) which links the syntactic information on verb complementation to semantic features. The labels refer to a set of syntactic features (valency, transitive and ditransitive) and to a set of semantic features:

- Valency (the number of arguments of the verb)
- Transitive (the verb takes a direct object)
- Ditransitive (the verb takes a direct and an indirect object)

- Control (the subject of the verb is capable of acting with volition)
- Attributive (the verb expresses a relation of ownership)
- Spatial (the verb expresses a location or movement of (one of) the participants(s))
- Cognition (the verb demands emotional, perceptual or mental activitity)
- Dynamic (the verb expresses a non-static, changing situation)

Only the positive values are shown. For example, state1 means -Control, -Dynamic, -Attributive, -Spatial, -Cognitive, -Transitive, -Ditransitive. Likewise, process1 means –Control, +Dynamic, –Attributive, –Cognitive, –Spatial, –Transitive, -Ditransitive. The numbers at the last position of the label (e.g. action**1**) correspond to the valency of the verb (mono, di or tri).

The following table presents an overview of the labels with an explanation of the features and an example.

name	**features**	**example**
action1	Control Dynamic	*applaudisseren (applaud)*
action2	Control Dynamic Transitive	*afgieten (drain)*
action3	Control Dynamic Ditransitive	*combineren(combine)*
echprod2	Control Dynamic Attributive Transitive	*inkopen (buy in)*
echprod3	Control Dynamic Attributive Ditransitive	*afnemen (take up (goods))*
mvmt1	Control Dynamic Spatial	*duikelen (tumble)*
mvmt2	Control Dynamic Spatial Transitive	*patrouilleren (patrol)*
mvmt3	Control Dynamic Spatial Ditransitive	*deponeren (deposit)*
cognt1	Control Dynamic Cognitive	*huichelen (dissemble)*
cognt2	Control Dynamic Cognitive Transitive	*bagatelliseren (play down)*
cognt3	Control Dynamic Cognitive Ditransitive	*inpeperen (get even with s.o.)*
state1		*sneuvelen (fall in action)*
state2	Transitive	*leunen (lean)*
state3	Ditransitive	*noodzaken (compel)*
possess2	Attributive Transitive	*grossieren (collect)*
location1	Spatial	*schoolblijven (stay in after school)*
location2	Spatial Transitive	*stranden (run aground)*
location3	Control Spatial Ditransitive	*vasthouden (hold)*
stcognt1	Cognitive	*wanhopen (despair)*
stcognt2	Cognitive Transitive	*aanmatigen (assume)*
stcognt3	Cognitive Ditransitive	*ontlenen (derive)*
process1	Dynamic	*achteruitgaan (decline)*
process2	Dynamic Transitive	*schaden (damage)*
prmvmt1	Dynamic Spatial	*stuiteren (bounce)*
prmvmt2	Dynamic Spatial Transitive	*struikelen (stumble)*
procognt1	Dynamic Cognitive	*hallucineren (hallucinate)*

procognt2	Dynamic Cognitive Transitive	*pochen (boast)*
procognt3	Dynamic Cognitive Ditransitive	*inspireren (inspire)*

Table 4: semantic features of verbs

3.5 Sense examples

Cornetto contains a lot of information on the combinatoric properties of the LEs. We already mentioned the data categories for syntactic behaviour and the corresponding semantic feature set. In this section we will describe the **Sense Examples**, which also contain information on the use of the LE in context. Cornetto contains much more sense examples then commonly found in dictionaries and they are described systematically. The sense examples are built around five data categories: **canonicalForm**, **textualForm, Semantics_ex**, **Syntax_ex** and **Pragmatics**.

Each example is presented in canonical form and/or in textual form. Textual forms are fully inflected (parts of) sentences; canonical forms are representations in "dictionary" or basic mode with non-inflected verbs and singular nouns. If both a canonical and textual form are given, the latter is an illustration of the former in a broader context.

All examples are syntactically characterized as a Sentence, NP, VP, AP or PP. For example, the canonical form *iemand een advies geven(to give somebody an advise)* is a VP.

Moreover, the examples are categorized in different types (expressionType) according to their degree of semantic compositionality and syntactic fixedness. The expressionTypes range form "free combinations" which are fully compositional to more fixed combinations like grammatical and lexical collocations and pragmatic formulas. Other frequent subtypes are:

Grammatical collocations: fixed combinations consisting of a LE and a function word which is often a preposition. For example, *angst voor (fear of)* and *in antwoord op (in answer to).*

Lexical collocations: fixed combinations of frequently co-occurring lexical words like verbs, nouns and adjectives. For example, *een zwerm bijen (a swarm of bees), zwaar weer(heavy weather)* and *het anker lichten* (*lift the anchor).*

The **semantics of lexical collocations** are systematically characterized by a meaning collocator. This approach is broadly based on the work of Mel'čuk (1996). Some examples of the values are: *bonus* (a good x) like in *een wijs besluit (a wise decision), magnus* (an intensification of x) like in *een oorverdovend applaus (a deafening applause)* and *oper* for support verbs, such as *begaan (do, commit),* like in *een overtreding begaan (commit an offense).*

3.6 Pragmatics

Lexical entries (multiword and singleword units), synsets and sense examples have usage information. The values for pragmatics include information on:

- the domain the LE is used in (e.g. economics, for begroting (estimate), electricity, for condensator (condensator))
- chronology (old fashioned: alkoof (alcove), neologism: fluisterasfalt(silent asphalt))

- connotation (pejorative: gepeupel (mob), euphemistic: troostmeisje (consolation girl), offensive: geitenbreier (duffer, bore), jocular: zakjapanner(small calculator; lit. a Japanese person for in your pocket))
- geography (belg: ajuin, onion). The label geography refers to the regional occurance of a word; LEs labeld belg are more frequently used in Belgian Dutch than in the Dutch spoken in the Netherlands.
- register (formal: *zwerk (wrack)*, informal: *zwieper (wallop)*, vulgar: *gezeik (bullshit)*, *slang: bajes (can, cooler)*)

Domain labels can be found at two places in the Cornetto entry: the synsets have a reference to WordNet domains and the lexical entry also includes domain labels. Both sets of domain labels are based on Wordnet Domains Hierarchy. For more information, please refer to section 8.2.

4 Multiword Expressions

Lexical entries may be singleword expressions (i.e words such as *lopen* (*walk)* and *huis* (*house*)) or multiword expressions. Multiword expressions are expressions which consist of more than one word and have a meaning which cannot be derived from the individual parts of the expression. There are two types of multiword expressions (i.e. expressionType) in Cornetto:

- Idioms are considered as a group of words in a fixed order that have a particular meaning that is different from the meaning of each word understood on its own. Idiomatic expressions are semi-fixed expressions which means that word order and composition are invariable, while inflection, variation in reflexive form and determiner selection is possible.For example: *Eieren voor zijn geld kiezen* (literal translation: *choose eggs for your money*) which means "taking the easiest and safest way out".
- Proverbs are short sayings or sentences that usually contain words of wisdom, truth or morality. They are complete sentences which are syntactically completely fixed. For example: *Haastige spoed is zelden goed* (*haste is waste* ; literal translation: *hasty speed is seldom good*) which means "doing something fast does not mean it is done well"

The following XML fragment represents the idiom *eieren voor zijn geld kiezen* (expressionType=idiom) with a definition and an example *de regering lijkt nu eieren voor zijn geld te kiezen (the government now seems to do the rational thing)* that illustrates the use of the idiom in a broader context.

```
<LexicalEntry id="buigen-mwe-c_545469-1">
      <MultiwordExpression writtenForm="eieren voor zijn geld kiezen" expressionType="idiom"/>
      <Sense senseId="c_545469-1"
           synset="unknown_000"
           definition="de verstandige weg kiezen"> /* do the rational thing */
        <Pragmatics/>
        <SenseExamples>
          <SenseExample id="ex-c_545469-1">
            <textualForm textualform="de regering lijkt nu eieren voor zijn geld te kiezen"/>
          </SenseExample>
        </SenseExamples>
      </Sense>
  </LexicalEntry>
```

Multiword expressions should be linked to synsets. However, this is not yet realized in the current version of Cornetto. Therefore, the synset id is set to "unknown_000".

5 Sense groups

Sense groups are relations between the different senses of a polysemous word. In the DutchSemCor project, 4 sets of sense groups have been derived using automatic and semi-automatic methods. A sense sengroup is a set of meanings of a lemma that are semantically close and therefore difficult to discriminate both for humans and machines. Metonymy, specialization and generalization of meaning of words can lead to closely related meanings that are compatible and can apply simultaneously in a context. An example of metonymy is "academie" (academy) referring to the institution (sense 1) and the building (sense 2). A case of specialization/generalization is "behandeling" (treatment), referring to a medical treatment (sense 1) but also to treatment in general (sense 2). By distinguishing sense groups it is possible to apply WSD at different levels of precision and relevance. Sense groups are derived from properties of the Cornetto resource or from the annotation in DutchSemCor. The following types (**relationType**) of sense groups are distinguished

- **Co-hyponyms** : lexical entries that belong to synsets that have the same hypernym
- **Co-relations:** lexical entries that belong to two synsets that have a direct semantic relation (like a NEAR_SYNONYM, SUBEVENT, RESULT, PART relation) with each other
- **Co-synonyms:** lexical entries that belong to synsets that have more than one synonym in common, i.e. in addition to the lemma itself there is at least lemma that occurs in both synsets
- **Co-annotations:** lexical entries that were both assigned to a single token in the corpus by human annotators. In the DutchSemCor project, at particular stages of the annotation, annotators were instructed to assign more than one lexical unit to a token if the meanings were compatible and it was not clear from the context which of the meanings applies. These double-annotations were used to extract sense groups.

For more information on sense groups and on the DutchSemCor project, please refer to http://www2.let.vu.nl/oz/cltl/dutchsemcor/.

The following XML fragment shows the 3 lexical entries of the lemma *das (scarf, necktie, badger*). The first two senses (*scarf* and *necktie*) are conceptually closer to each other than to the third sense (*badger*). They form a sense group as they are related to each other by a co-hyponym relation.

```
  <LexicalEntry id="das-n-1" partOfSpeech="noun">       /* scarf  */
      <Lemma writtenForm="das"/>
          <Sense senseId="r_n-10435"
            synset="nld-21-d_n-10772-n"
            definition="om de hals voor de warmte">
          <SenseRelations>
            <SenseGroup relationType="co-hyponyms" targetSenseId="r_n-10436"/>
          </SenseRelations>
      ….
          </Sense>
    </LexicalEntry>
<LexicalEntry id="das-n-2" partOfSpeech="noun">      /* necktie */
      <Lemma writtenForm="das"/>
  …
   <Sense senseId="r_n-10436"
            synset="nld-21-d_n-14300-n"
            definition="stropdas">
          <SenseRelations>
            <SenseGroup relationType="co-hyponyms" targetSenseId="r_n-10435"/>
          </SenseRelations>
…
      </Sense>
    </LexicalEntry>
    <LexicalEntry id="das-n-3" partOfSpeech="noun">      /* badger */
 …
     <Sense senseId="r_n-10437"
            synset="nld-21-d_n-38852-n"
            definition="zoogdier met zwart-witte streep">
          <SenseRelations/>
…
      </Sense>
    </LexicalEntry>
```

6 Sentiment

The Cornetto senses are linked to automatically derived sentiment labels such as “positive” and “negative”. They express or evoke a positive or negative opinion, belief or attitude. Polarity is automatically determined and imported from an external resource “VuPolarityLexAutomatic” (Maks and Vossen, 2011). As the labels are automatically generated, there may be incorrect ones.

The following XML-fragment gives the first sense of the word *ramp* (*disaster*) with a negative poarity label.

```
<LexicalEntry id="ramp-n-1" partOfSpeech="noun">
<Lemma writtenForm="ramp"/>
<WordForms>
<WordForm writtenForm="ramp" grammaticalNumber="singular" article="de"/>
<WordForm writtenForm="rampen" grammaticalNumber="plural" article="de"/>
</WordForms>
<Morphology/>
<MorphoSyntax pronominalAndGrammaticalGender="mf"/>
        <Sense senseId="r_n-30452" synset="d_n-19123"  definition="ongeluk van grote    omvang of
        invloed">
<SenseRelations/>
<Sentiment polarity="negative" externalReference="VuPolarityLexAutomatic"/>
..
</Sense>
</LexicalEntry>
```

7 Synsets

The Cornetto synsets group Dutch words into sets of synonyms called synsets. Every synset contains a group of synonymous words; different senses of a word are in different synsets. Synsets are interlinked by means of conceptual-semantic and lexical relations which result in a network of meaningfully related words and concepts.
These relations include:

- hypernyms: Y is a hypernym of X if every X is a (kind of) Y (canine is a hypernym of dog)
- hyponyms: Y is a hyponym of X if every Y is a (kind of) X (dog is a hyponym of canine)
- antonyms: Y is an antonym of X if every Y is a opposite concept of X (clean is a antonym of dirty)
- role_agent: which can be found between *hond*:1 (*dog*) and *blaffen:*1 (*bark*) where *dog* has the role of agent of the verb *bark*.

A complete list of synset relations and their definitions can be found in section 9.9.

The following XML-fragment presents part of the relations of the synset nld-21-d_n-17260-n which contains words the word *hond (dog)* and its synonyms.

```
<Synset id="nld-21-d_n-17260-n">          / * dog, etc. */
 <SynsetRelations>
  <SynsetRelation
     target="nld-21-d_n-20460-n" relType="HAS_HYPERONYM">      /*  carnivore*/
  </SynsetRelation>
  <SynsetRelation
     target="nld-21-n_a-512728-a" relType="BE_IN_STATE">        /* rabid, hydrophobic */
  </SynsetRelation>
  <SynsetRelation
     target="nld-21-d_n-39722-n" relType="HAS_HOLO_MEMBER"> /* dog basket */
  </SynsetRelation>
  <SynsetRelation
     target="nld-21-d_n-22878-n" relType="HAS_MERO_PART">     /* dog shit */
  </SynsetRelation>
  <SynsetRelation
     target="nld-21-d_n-22882-n" relType="HAS_MERO_PART">      /* doglike devotion */
  </SynsetRelation>
  <SynsetRelation
     target="nld-21-d_n-12269-n" relType="ROLE">                    /* dog breeding */
  </SynsetRelation>
  <SynsetRelation
    target="nld-21-d_v-1541-v"  relType="ROLE_AGENT">            /* bark */
  </SynsetRelation>
  <SynsetRelation
     target="nld-21-d_v-7284-v"  relType="ROLE_PATIENT">        /* trim, clip */
  </SynsetRelation>
  <SynsetRelation
     target="nld-21-d_v-116-v"   relType="ROLE_PATIENT">          /* put a leash */
  </SynsetRelation>
 </SynsetRelations>
 <MonolingualExternalRefs>
 </Synset>
```

8 External References

8.1 Princeton WordNet

English or Princeton WordNet is a lexical database of English with a synset structure similar to the Cornetto synsets.
Synsets in Cornetto-LMF have been automatically mapped onto synsets of the Princeton WordNet. Many mappings are being manually validated and corrected or have been done so in previous projects. Relations between Dutch and English synsets are called Equivalence relations (EQ). For instance, the Dutch *slurpen* (*sip*) and the English *drink* illustrate a relation between two concepts in two different languages where concept A is a type of concept B and is called EQ_HAS_HYPERONYM.
Through these translation links with the Princeton WordNet, the Dutch Cornetto synsets can be linked to other resources which are mapped onto the Princeton WordNet, like, for example, WordNet Domains (cf. section 8), SUMO (cf. section 9.2) and WordNets of other languages. More information on the Princeton WordNet can be found here: http://wordnet.princeton.edu/.

8.2 Domains

Domain labels in Cornetto can be found at two places in the Cornetto entry: the synsets have a reference to WordNet domains and the lexical entry also includes domain labels (cf. section 3.6). Both domain labels are based on the Wordnet Domains Hierarchy which is a flat (two-level) hierarchy of basic domains suitable for NLP tasks such as text categorization and word sense disambiguation. For example, there is a category *doctrine* which includes categories like *archeology*, *linguistics* and *psychology*. The synset domain labels are automatically generated through the translation equivalent links with the Princeton Wordnet. They have been partly manually post-edited. The LE domain labels have been manually assigned. Another difference between the 2 sets of labels refers to the rules for assigning them. The LE are (usually) labeled only if they are considered as expert concepts or words , such as for example *diabetes (medicine).* Synsets, however, are also labeled if they refer to general language words. For instance, the synset which includes *huisdier* (*pet*) is labeled with the domain label *zoology*, whereas the LE *huisdier* is not. More information about the WordNetDomains hierarchy can be found here: http://wndomains.fbk.eu/.[3]

The following XML-fragment shows domain labeling in synsets:

```
<Synset id="nld-21-d_n-39784-n">    /* dog */
</MonolingualExternalRef>
<MonolingualExternalRef externalSystem="wordnet_domain" externalReference="zoology"/>
</MonolingualExternalRefs>
</Synset>
```

The following XML-fragment shows domain labeling in LEs:

```
<LexicalEntry id="diabetes-n-1" partOfSpeech="noun">
<Sense senseId="r_n-10792"  synset="nld-21-d_n-14325-n"  definition="suikerziekte">
<Semantics-noun reference="common" countability="uncount" semanticType="nondynamic"/>
<Pragmatics>
    <Domains domain="medicine"/>
</Pragmatics>
</Sense>
</LexicalEntry>
```

[3] accoustics , administration , aeronautics , agriculture , alimentation , anatomy , anthropology , archeology , architecture , art , artisanship , astrology , astronomy , astronautics , astronomie , biochemistry , biology , botany , building_industry , chemistry , cinema , commerce , computer_science , cycling , dance , diplomacy , doctrines , ecology , economy , electronics_electricity , empty , engineering , fashion , fishing , folklore , gastronomy , geography , geology , golf , herladry , history , housekeeping , hunting , hydraulics , industry , insurance , law , linguistics , literature , mathematics , media , medicine , merchant_navy , meteorology , metrology , military , money , music , mythology , pedagogy , pharmacy , philosophy , photography , physics , play , politics , psychology , publishing , railway , religion , school , science , sculpture , sexuality , soccer , sociology , sport , state , swimming , telecommunication , tennis , theatre , theology , transport , zoology

8.3 SUMO

Synsets in Cornetto-LMF contain labels from The Suggested Upper Merged Ontology (SUMO). SUMO and its domain ontologies form the largest formal public ontology in existence today. They are being used for research and applications in search, linguistics and reasoning.
SUMO labels have been generated automatically through the mapping of Dutch synsets onto English synsets using Equivalence relations (EQ). During the Cornetto project and other projects later on, many of these Ontology labels have been validated and corrected manually. See below an example of SUMO labels.

veroordeling:1 - het bepalen v.e. vonnis

SUMO labels: (+, , RegulatoryProcess) , (+, , Declaring) , (+, , ExpressingDisapproval) , (+, , TimeInterval)

For more information on SUMO and an extensive list of all ontological labels (including the hierarchy which is not available in Cornetto), please refer to the SUMO portal: http://www.ontologyportal.org/.

9 Cornetto data categories and ISOcat definitions

The datacategories used in Cornetto are selected from the ISOcat registry. The following table gives data category names (column1) , data category values (column 1 with indentation) and definitions (column 2). The categories are grouped into broader categories like Lexicon Information which correspond to the several lexicon layers and to the elements in the DTD structure. Some categories apply only for a particular part-of-speech. If this is the case, the particular part-of-speech is given between brackets immediately after the name of the datacategory.

An alphabetic overview of the data categories including ISOcat urls is given in Appendix I.

9.1 Lexicon Information

GlobalInformation	Global Information is a class representing administrative information and other general attributes. There is an aggregation relationship between the Lexical Resource class and the Global Information class in that the latter describes the administrative information and general attributes of the entire resource. The Global Information class does not allow subclasses.
LexicalResource	Lexical Resource is a class representing the entire resource. Lexical Resource occurs once and only once. The Lexical Resource instance is a container for one or more lexicons.
Lexicon	Lexicon" may also be used as a general term covering different types of lexical resources such as dictionaries, WordNets, lexicons, etc.
external Reference	Reference to a particular node of an external descriptive system.
external System	Reference to an external descriptive system.
MonolingualExternalRef	Monolingual External Ref is a class representing the relationship between a Sense or a Synset instance and an external system.

9.2 LexicalEntry

LexicalEntry	Container to represent the LMF Lexical Entry Class which represents a lexeme in a given language
formType	Indication whether the form is a full form or an abbreviated form
acronym	An abbreviation made up of the initial letters of the components of the full form of the designation or from syllables of the full form and pronounced syllabically.
full	full form (i.e. The form is not abbreviated)
abbreviation	indication that a word resulted by leaving out characters of a longer word or words, keeping the original meaning, while the resulting word is not considered a proper noun nor behaves as a common noun
contraction	A lexical unit formed by a shortening of a word, syllable, or word group by omission of a sound or letter.
Lemma	Base form a word or term that is used as the formal entry in a dictionary.
writtenForm	representation of the written string of a form
partOfSpeech	A category assigned to a word based on its grammatical and semantic properties.
verb	word that can be conjugated when combined with other words or groups of words in a sentence
adjective	word characterizing an independent entity like a noun or its equivalent, while it is not a word form of another part of speech
noun	word that can be combined with a demonstrative pronoun, while it is not a

	word form of another part of speech
WordForms	WordForm – Word Form is a Form subclass representing a form that a lexeme can take when used in a sentence or a phrase. So, Word Form class can manage simple lexemes, compounds and multi-word expressions.
grammaticalNumber (noun)	Grammatical category for the variation in form of nouns, pronouns, and any words agreeing with them, depending on how many persons or things are referred to
plural	Value that expresses more than one element
singular	Value that expresses more than one element
tense (verb)	Property referring to the way the grammar marks the time at which the action denoted by the verb took place.
pastParticiple	non-finite form of a verb expressing an event as completed
pastTense	Past tense is an absolute tense that refers to a time before the moment of utterance
RelatedForm	Related Form is a Form subclass representing a word form or a morph that can be related to the Lexical Entry in one of a variety of ways (e.g. Derivation, root). The Related Form can be typed. There is no assumption that the Related Form is associated wit
variantType	Refers to whether the variant form is different with regard to spelling (i.e. Spelling variant) or pronunciation (i.e. form variant)
spellingVariant	Variants with different written spelling but with the same pronunciation.
formVariant	Variants with different written spelling and pronunciation.

9.3 Morphology and MorphoSyntax

Morphology	The study of the structure and constituency of individual words.
adverbialUsage (adjective)	property of an adjective to be used also as an adverb (yes) or only as an adjective (no)
article (noun)	Definite article which is used for singular elements
auxiliary (verb)	relation between a subordinate verb and the main verb
comparisonType (adjective)	The type of rule (i.e. regular, irregular or mixed regular and irregular) which accounts for the forming of the comparative and superlative forms of adjectives
regular	Modification of adjectives to express relative degree by adding -er for the comparative and -st for the superlative form
mixed	Modification of adjectives to express relative degree using both regular and irregular forms such as in the case of 'achterbaks - achterbakser - meest' achterbaks.
irregular	Modification of adjectives to express relative degree using irregular forms such as 'goed - beter - best'.
degree (adjective)	property concerning comparison
superlative	value expressing (via affix or suppletion) the highest level of intensity of the named feature
comparative	value expressing (via affix, suppletion) a higher than basic level of intensity of the named feature, while it is not the highest level.
mode (verb)	One of a set of distinctive forms that are used to signal modality. Modality is a facet of illocutionary point or general intent of a speaker, or a speaker's degree of commitment to the expressed proposition's believability, obligatoriness, desirability or reality.
infinitive	Mood cited as unmarked or base form.
morphoType	morphological type of a word refers to the internal morphological structure of a word

compound	A lexical unit that combines two or sometimes more different words, frequently such that the sense of the new lexical unit is not clearly derivable from the combination of its parts.
phrasal	a verb that is always combined with a preposition-like element (particle)
zero-derivation	The creation of a word from an existing word of a different word class without any change in form.
derivation	A lexeme that is related to another lexeme by a rule of derivation.
x-compound	idiomatic compound
compderiv	A word that is morphologically related to other words by rules of both derivation and composition
wordgroup	a wordgroup is a lemma that consists of one or more content words separated by space of hyphens
position (adjective)	refers to the property of an adjective to be used in attributive or predicative position or in both positions
attrpred	An adjective that can be used in both attributive and predicative position.
attributive	An adjective that comes before a noun and not after a copula verb, like BE, SEEM, etc
predicative	indication that a word is realized at the position meant for the nonverbal part of a predicate or that of an adverb
pronominalAndGrammaticalGender (noun)	grammatical gender refers to the noun classification system that divides common nouns in classes like feminine, neuter and masculine; proniminal gender is the gender of the pronoun that refers to a noun; these types of gender are related but not identic
mn	pronominal gender is m (masculine) ; grammatical gender is n (neuter)
m_f	pronominal gender is m (masculine) or f (feminine) depending on the biological gender
fn	pronominal gender is feminine (f) ; grammatical gender is neuter (n)
f	grammatical and pronominal gender are f (feminine)
mfn	pronominal gender is m (masculine) or f (feminine) depending on the biological gender, or n (neuter) ; grammatical gender is n (neuter)
mf	pronominal gender is m (masculine); grammatical gender is f (feminine)
n	grammatical and pronominal gender are n (neuter)
m	grammatical and pronominal gender are m (masculine)
reflexivity (verb)	verb with an optional or obligatory reflexive pronoun
reflexive	voice for a construction where the semantic agent and patient are the same
optionalReflexive	reflexive voice is optional

separability (verb)	refers to the characteristics of a verb whether it consists (or not consists) of a lexical core and a separable particle; the phenomenon of separable verbs is frequently found in Dutch, German and Hungarian
unseparable	if a verb is unseparable , it does not consist of a lexical core and a separable particle but it appears always in one form
separable	A separable verb is a verb that is composed of a lexical core and a separable particle

9.4 Syntax and Syntactic Behaviour

Syntax	The study of grammatical relations between words and other units within a sentence' (Concise Oxford Dictionary of Linguistics). To be distinguished from morphology, which applies to units smaller than the word.

SubcategorizationFrame (verb)	Subcategorization Frame is a class representing one syntactic construction. A Subcategorization Frame instance is shared by all Lexical Entry instances that have the same syntactic behaviour in the same language. A Subcategorization Frame can inherit relationships and attributes from another more generic Subcategorization Frame by means of a reflexive link. Therefore, it is possible to integrate a hierarchical structure of Subcategorization Frame instances.
syntacticArgument (verb)	Syntactic Argument is a class representing an argument of a given Subcategorization Frame. A Syntactic Argument can be linked recursively to a Subcategorization Frame instance in order to describe deeply complex arguments. Syntactic Argument allows the connection with a semantic argument by means of a SynSemArgMap instance.
syntacticBehaviour	Syntactic Behaviour is a class representing one of the possible behaviours of a lexeme. The Syntactic Behaviour instance is attached to the Lexical Entry instance and optionally to the Sense instance. The presence in a given lexicon of one Syntactic Behaviour instance for a lexical entry means that this lexeme can have this behaviour in the language of the lexicon. Syntactic description is optional, so it is possible to describe morphology and semantics without any syntactic description. Lexical Entry, Syntactic Behaviour and Sense instances form a triangle representing Morphology, Syntax and Semantics. Detailed description of the syntactic behaviour of a lexical entry is defined by the Subcategorization Frame instance.
complement	a complement is a word group that functions as the complement of a verb, adjective or noun. Usually it is headed by a complementizer
fixprep	optional complement headed by a fixed preposition
psmodnoun	Post-modifying noun complement
oblprep	obligatory complement headed by a fixed preposition
quant	complement with optional specification of quantity
factive	complement headed by the Dutch factive complementizer "dat"
omtoinf	complement with "om te"-infinitive
whclause	complement headed by a WH-word like waarom (why), waar (where) , etc.
oblobj	complement with obligatory object 'het'
toinf	complement with "te"-infinitive
dancomp	complement headed by "dan" which is in Dutch the conjunction for introducing comparative clauses
complementizer	A complementizer is a connective which marks a complement clause [Crystal 1997: 75]
dat	the embedded sentence is headed by the subordinating conjunction "dat" (in Dutch)
WH	the embedded sentence is headed by a question word or wh-word like waarom (why), wat (what), wanneer (when), etc.
te	the embedded sentence is headed by the subordinating conjunction "te" (in Dutch, cf. "to" in English)
of	the embedded sentence is headed by the subordinating conjunction "of" (in Dutch; cf. "whether" in English)
hoe	the embedded sentence is headed by the subordinating conjunction "hoe" (in Dutch; cf. "how" in English)
omte	the embedded sentence is headed by the subordinating conjunction "om te" (in Dutch; cf. "to" in English)
constituent	A constituent is a word or a group of words that functions as a single unit within a hierarchical structure
np	phrase headed by a noun

ap	phrase headed by an adjective
vp	phrase headed by a verb
pp	phrase beginning by one or several prepositions and a complement such as a noun phrase
subclause	A clause which does not constitute a complete sentence in itself, but must be connected with or attached to an independent clause. [Pei and Gaynor 1980: 206]
pp	phrase beginning by one or several prepositions and a complement such as a noun phrase
function	The grammatical relationship of one constituent to another within a syntactic construction.
specifyingComplement	complement of a verb which is realized by NP, AP or PP, specifying a quantity or a quality
directObject	A noun, pronoun, or noun phrase whose referent receives the direct action of a verb.
objectComplement	A complement that is used to predicate a description of the direct object.
indirectObject	A noun, pronoun, or noun phrase indicating the recipient or beneficiary of the action of a verb and its direct object.
phraseType	A group of words forming a syntactic constituent with a single grammatical function.
pp	phrase beginning by one or several prepositions and a complement such as a noun phrase
vp	phrase headed by a verb
ap	phrase headed by an adjective
np	phrase headed by a noun
sentence	A sentence is a grammatical unit that is composed of one or more clauses
preposition	Adposition placed at the beginning of a noun phrase.
transitivity (verb)	A property of verbs that relates to whether a verb can take direct objects and how many such objects a verb can take.
intransitive	Refers to a verb that does not take a direct object; that is, to a verb that does not express an action which directly affects another person or thing
transitive	A verb which takes a direct object; that is, a verb that expresses an action which directly affects another person or thing.
valency (verb)	The number of arguments controlled by a verbal predicate
mono	a monovalent verb (also called intransitive) takes only one argument which is the subject
tri	a trivalent verb is a verb that takes 3 arguments. One of them is the subject, the other one is the direct object and the third one is an indirect object or other type of prepositional object.
di	a divalent verb takes two arguments. One of them is the subject and the other is the direct object or a (obligatory) prepositional object.

9.5 Sense

Semantics	The study of the meaning of linguistic structures.
Sense	One of zero to many meanings or concepts associated with a given head word in a lexical entry.
countability (noun)	Countability is motivated by the semantic distinction between object and substance reference.
mass	Designation of a term or word that is not countable and cannot generally be

	used with the indefinite article or in the plural.
uncount	An uncountable noun is a noun with the property that any quantity of it is treated as an undifferentiated unit, rather than as something with discrete subsets.
coll	A collective noun is the name of a number (or collection) of people or things taken together and spoken of as one whole.
count_uncount	Nouns that can be both countable or uncountable according to their different meanings.
pluraleTantum	Plurale Tantum is a noun that appears only in the plural form and does not have a singular variant for referring to a single object.
count	noun referring to one or several concepts that may be counted
definition	Representation of a concept by a descriptive statement which serves to differentiate it from related concepts.
reference (noun)	attribute dealing with the type of noun, like common and proper
proper	Noun that is the name of a specific individual, place, or object.
common	Noun that signifies a non-specific member of a group.
semanticFeatureSet (verb)	a set of semantic properties which describe - in a formal way - the meaning of a verb. The set consists of the following features : transitive, intransitive, ditransitive, dynamic, cognitive, control, state, spatial
mvmt3	refers to a verb concept with the following semantic features:ditransitive, dynamic, control, spatial
action2	refers to a verb concept with the following semantic features:transitive, dynamic, control
action3	refers to a verb concept with the following semantic features::ditransitive, dynamic, control
cognt1	refers to a verb concept with the following semantic features: intransitive, dynamic,control, cognitive
action1	refers to a verb concept with the following semantic features:intransitive, dynamic, control
cognt3	refers to a verb concept with the following semantic features:ditransitive, dynamic,control, cognitive
stcognt3	refers to a verb concept with the following semantic features:ditransitive, dynamic, control, cognitive
echprod3	refers to a verb concept with the following semantic features:ditransitive, dynamic, control, attributive
echprod2	refers to a verb concept with the following semantic features: transitive, dynamic, control, attributive
cognt2	refers to a verb concept with the following semantic features: transitive, dynamic,control, cognitive
process2	refers to a verb concept with the following semantic features:intransitive, dynamic
mvmt2	refers to a verb concept with the following semantic features:transitive, dynamic, control, spatial
mvmt1	refers to a verb concept with the following semantic features:intransitive, dynamic, control, spatial
prmvmt3	refers to a verb concept with the following semantic features:ditransitive, dynamic, spatial
process1	refers to a verb concept with the following semantic features:intransitive, dynamic
location2	refers to a verb concept with the following semantic features: transitive, spatial
possess2	refers to a verb concept with the following semantic features: transitive, attribu-

	tive
process3	refers to a verb concept with the following semantic features:ditransitive, dynamic
prcognt1	refers to a verb concept with the following semantic features:intransitive, dynamic, cognitive
location1	refers to a verb concept with the following semantic features:intransitive, state, spatial
state1	refers to a verb concept with the following semantic features: state, intransitive
prcognt2	refers to a verb concept with the following semantic features:transitive, dynamic, cognitive
state2	refers to a verb concept with the following semantic features: state, transitive
prcognt3	refers to a verb concept with the following semantic features:ditransitive, dynamic, cognitive
state3	refers to a verb concept with the following semantic features: state, ditransitive
prmvmt1	refers to a verb concept with the following semantic features:intransitive, dynamic, spatial
prmvmt2	refers to a verb concept with the following semantic features: transitive, dynamic, spatial
possess3	refers to a verb concept with the following semantic features: ditransitive, attributive
stcognt2	refers to a verb concept with the following semantic features: transitive, dynamic, control, cognitive
stcognt1	refers to a verb concept with the following semantic features:intransitive, dynamic, control, cognitive
semanticShifts	a semantic shift refers to a process of meaning extension following a -usually systematic- shift between semantic types
semanticType	a semantic Type is part of a set of values that categorizes concepts in a given domain according to their semantic characteristics
time	specifying a period or moment in time
abstract	Properties or qualities as distinguished from any particular embodiment of the properties/ qualities in a physical medium. Instances of Abstract can be said to exist in the same sense as mathematical objects such as sets and relations, but they cannot exi
colour	refers to colour
phyper	refers to physical or perceptive properties
place	a point or an area on the Earth's surface or elsewhere
measure	nonanimate and a unit of length, measure, distance, money etc.
nondynamic	abstract concept not involving change or internal stages
action	Action verbs are verbs that describe actions and things taking place rather than states. Unlike most stative verbs, an action verb can usually be used in the progressive aspect, which is used for actions that are in progress.
emomen	referring to emotional or mental properties
institute	groups and organizations that can act as an agent
human	animate and human being (cf. nonhuman)
artefact	refers to physical entities constructed by man
nonhuman	animate but not human (cf. human)
process	refers to phenomenon marked by changes through a series of states
dynamic	abstract concept involving change or internal stages

concrother	refers to concepts which are concrete, but not substance or artifact
animate	Perceived as alive.
concrete	A concrete object is an object which exists at a particular time or place. It is the opposite of abstract.
state	refers to situations or conditions that are static
substance	nonanimate and material substances, incl. natural kinds

9.6 Sense Examples

CanonicalForm	The canonical form is a representation of the example in a basic mode ("dictionary mode") with not inflected verbs and singular nouns
SenseExample	SenseExample – Sense Example is a class used to illustrate the particular meaning of a Sense instance. A Sense can have zero to many examples.
TextualForm	The textual form is a representation of an example as a fully inflected sentence; if it is combined with a canonical form, it is an illustration of the latter in a broader context.
combiWord	sense examples and multiword expressions are considered as combinations of a head word and one or more combination words (combiWords).
expressionType	expression types refer to a range of subtypes in accordance with the degree of semantic non-compositionality and syntactic fixedness of the multiword expression
grammaticalCollocation	a collocation (a combination of frequently co-occurrung words) of function words (e.g a preposition or determiner) and content words (i.e. a noun, verb, adjective or adverb). collocations are syntactically fixed and semantically transparent.
properName	a noun that is the name of a specific individual, place, or object
slogan	a memorable phrase often used in a political, commercial, religious, and or other context as a repetitive expression of an idea or purpose (en.wikipedia.org/wiki/slogan)
term	A verbal designation of a general concept in a specific subject field.
pragmaticFormula	a combination of words used repeatedly and always in the same form. They are "highly conventionalized prepatterned expressions whose occurrence is tied to more or less standardized communication situations" (Coulmas, F. (1981) Conversational Routine)
lexicalCollocation	a collocation (a combination of frequently co-occurrung words) of two or more content words , i.e. nouns , verbs and adjectives and adverbs; collocations are syntactically fixed and semantically transparent
freeCombination	a free word combination is a combination of words which is semantically completely transparent and syntactically not fixed.

9.7 SenseRelation

SenseRelation	Sense Relation is a class representing the oriented relationship between Senses instances
reltype	synset relation type refers to the type of relationship that exists between two synsets within one wordNet or between different wordNets.
SenseGroup	A sense group is a set of meanings of a word that are semantically close and therefore difficult to discriminate both for humans and machines (source: DutchSemCor-project).
senseGroupRelationType	sense group relation type refers refers to the type relation which exist between different senses of a word
co-synonyms	senses of one word that belong to synsets that have more than one synonym in common, i.e. in addition to the lemma itself there is at least lemma that occurs in both synsets.

co-relations	senses of one word that belong to synsets that have a direct semantic relation, e.g. NEAR_SYNONYM, SUBEVENT, RESULT, PART, etc.
co-annotations	senses of one word that were both assigned to a single token in the corpus during manual sense-tagging in the DutchSemCor-project.
co-hyponyms	senses of one word that belong to synsets that have the same hypernym

9.8 Sentiment

polarity	polarity refers to the property of a word in a particular sense to express attitude. Usually there are 3 values: positive , negative or neutral (no) attitude
negative	the property of a word (sense) to express a negative attitude
positive	the property of a word (sense) to express a positive attitude

9.9 Synset

synset	A synonym set; a set of words that are interchangeable in some context without changing the truth value of the preposition in which they are embedded.
synsetRelation	synsetRelation is a class representing the oriented relationship between Synset instances.
CO_INSTRUMENT_RESULT	A relation between two concepts where concept B is the result of an action carried out by the instrument expressed by concept A. (cf. CO_RESULT_INSTRUMENT)
CO_AGENT_INSTRUMENT	A relation between two concepts where concept B is the instrument used by concept A in a certain action. (cf. CO_INSTRUMENT_AGENT)
ROLE_AGENT	A relation between two concepts where concept A is typically the agent of the action expressed by concept B. (cf. INVOLVED _AGENT)
ROLE_DIRECTION	A relation between two concepts where concept A is typically the direction or location of the action or event expressed by concept B. (cf. INVOLVED _DIRECTION)
BE_IN_STATE	A relation between two concepts where concept A is qualified by concept B (cf. STATE_OF)
ROLE_INSTRUMENT	A relation between two concepts where concept A is the instrument necessary for the action or event expressed by concept B. (cf. INVOLVED _INSTRUMENT)
CAUSES	A relation between two concepts where concept B comes into existence as a result of concept A (cf. IS_CAUSED_BY)
ROLE	A relation between two concepts where concept A is typically involved in the action or event expressed by concept B. (cf. INVOLVED)
CO_AGENT_PATIENT	A relation between two concepts where concept B is the patient undergoing an action carried out by concept A. (cf. CO_PATIENT_AGENT)
CO_AGENT_RESULT	A relation between two concepts where concept B is the result of an action carried out by concept A. (cf. CO_RESULT_AGENT)
CO_INSTRUMENT_PATIENT	A relation between two concepts where concept B undergoes an action for which the instrument expressed by concept A is used. (cf. CO_PATIENT_INSTRUMENT)
CO_PATIENT_AGENT	A relation between two concepts where concept B undergoes an action carried out by concept A. (cf. CO_AGENT_PATIENT)
CO_PATIENT_INSTRUMENT	A relation between two concepts where concept A undergoes an action for which the instrument expressed by concept A is used. (cf. CO_INSTRUMENT_PATIENT)
CO_RESULT_AGENT	A relation between two concepts where concept A is the result of an action carried out by concept B. (cf. CO_AGENT_RESULT)
CO_RESULT_INSTRUMENT	A relation between two concepts where concept A is the result of an action for which the instrument expressed by concept B is used. (cf.

	CO_INSTRUMENT_RESULT)
CO_ROLE	A relation between two concepts where one concept undergoes an action in which the other concept is involved (bidirectional).
HAS_SUBEVENT	A relation between two concepts where concept B takes place during or as part of concept A, and whenever concept B takes place, concept A takes place. (cf. IS_SUBEVENT_OF)
CO_INSTRUMENT_AGENT	A relation between two concepts where concept A is the instrument used by concept B for a certain action. (cf. CO_AGENT_INSTRUMENT)
HAS_XPOS_HYPERONYM	A relation between two concepts of different part of speech where concept B is a type of concept A. (cf. HAS_XPOS_HYPONYM)
HAS_HYPERONYM	A relation between two concepts where concept A is a type of concept B. (cf. HAS_HYPONYM)
INVOLVED_LOCATION	A relation between two concepts where concept B is typically the location where the action or event expressed by concept A takes place. (cf. ROLE _LOCATION)
INVOLVED_INSTRUMENT	A relation between two concepts where concept B is typically the instrument necessary for the action or event expressed by concept A. (cf. ROLE _INSTRUMENT)
INVOLVED_DIRECTION	A relation between two concepts where concept B is typically the direction or location of the action or event expressed by concept A. (cf. ROLE _DIRECTION)
INVOLVED_AGENT	A relation between two concepts where concept B is typically the agent of the action expressed by concept A. (cf. ROLE_AGENT)
INVOLVED	A relation between two concepts where concept B is typically involved in the action or event expressed by concept A. (cf. ROLE)
INVOLVED_RESULT	A relation between two concepts where concept B comes into existence as a result of concept A. (cf. ROLE _RESULT)
IN_MANNER	A relation between two concepts where concept B qualifies the manner in which an action or event expressed by concept A takes place. (cf. MANNER_OF)
INVOLVED_SOURCE_DIRECTION	A relation between two concepts where concept B is the place from where the action or event expressed by concept A begins/starts/happens. (cf. ROLE _SOURCE_DIRECTION)
HAS_MERONYM	A relation between two concepts where concept B makes up a part of concept A. (cf. HAS_HOLONYM)
HAS_MERO_PART	A relation between two concepts where concept B is a component of concept A. (cf. HAS_HOLO_PART)
HAS_MERO_MEMBER	A relation between two concepts where concept B is a member/ element of concept A. (cf. HAS_HOLO_MEMBER)
HAS_MERO_MADEOF	A relation between two concepts where concept A is made of concept B. (cf. HAS_HOLO_MADEOF)
HAS_MERO_LOCATION	A relation between two concepts where concept A is a place located in concept B. (cf. HAS_HOLO_LOCATION)
HAS_HYPONYM	A relation between two concepts where concept B is a type of concept A. (cf. HAS_HYPERONYM)
HAS_MERO_PORTION	A relation between two concepts where concept A is an amount/piece/portion of concept B. (cf. HAS_HOLO_PORTION)
ROLE_LOCATION	A relation between two concepts where concept A is the location where the action or event expressed by concept B takes place. (cf. INVOLVED _LOCATION)
XPOS_NEAR_SYNONYM	A relation between two concepts of different part of speech where concept A and concept B are similar in meaning.
XPOS_NEAR_ANTONYM	A relation between two concepts of different part of speech where concept A has the opposite meaning to concept B.

XPOS_FUZZYNYM	A relation between two concepts of different part of speech where concept A is strongly associated with concept B but no proper relation has been defined between the two.
STATE_OF	A relation between two concepts where concept B is qualified by concept A. (cf. BE_IN_STATE)
ROLE_TARGET_DIRECTION	A relation between two concepts where concept A is the place where the action or event expressed by concept B leads to. (cf. INVOLVED_TARGET_DIRECTION)
ROLE_SOURCE_DIRECTION	A relation between two concepts where concept A is the place from where the action or event expressed by concept B begins/starts/happens. (cf. INVOLVED _SOURCE_DIRECTION)
INVOLVED_PATIENT	A relation between two concepts where concept B is typically the patient undergoing an action or event expressed by concept A. (cf. ROLE _PATIENT)
ROLE_PATIENT	A relation between two concepts where concept A is the patient undergoing an action or event expressed by concept B. (cf. INVOLVED _PATIENT)
HAS_XPOS_HYPONYM	A relation between two concepts of different part of speech where concept B is a type of concept A. (cf. HAS_XPOS_HYPERONYM)
NEAR_SYNONYM	A relation between two concepts where concept A and concept B are closely related in meaning but are not in the same synset.
NEAR_ANTONYM	A relation between two concepts where concept A has the opposite meaning to concept B.
MANNER_OF	A relation between two concepts where concept A qualifies the manner in which an action or event expressed by concept B takes place. (cf. IN_MANNER)
IS_SUBEVENT_OF	A relation between two concepts where concept A takes place during or as part of concept B, and whenever concept A takes place, concept B takes place. (cf. HAS_SUBEVENT)
IS_CAUSED_BY	A relation between two concepts where concept A comes into existence as a result of concept B. (cf. CAUSES)
INVOLVED_TARGET_DIRECTION	A relation between two concepts where concept B is the place where the action or event expressed by concept A leads to. (cf. ROLE _TARGET_DIRECTION)
ROLE_RESULT	A relation between two concepts where concept A comes into existence as a result of concept B. (cf. INVOLVED _RESULT)
EQ_HAS_MERONYM	A relation between two concepts in two different languages where concept B makes up a part of concept A. (cf. EQ_HAS_HOLONYM)
EQ_SYNONYM	A relation between two concepts in two different languages where concept A and concept B are translation equivalents. (cf. EQ_NEAR_SYNONYM)
EQ_ROLE	A relation between two concepts in two different languages where concept A is typically involved in the action or event expressed by concept B. (cf. EQ_INVOLVED)
EQ_NEAR_SYNONYM	A relation between two concepts in two different languages where concept A and concept B are similar in meaning and are possible translations of each other. (cf. EQ_SYNONYM)
EQ_IS_SUBEVENT_OF	A relation between two concepts in two different languages where concept A takes place during or as part of concept B, and whenever concept A takes place, concept B takes place. (cf. EQ_HAS_SUBEVENT)
HAS_HOLONYM	A relation between two concepts where concept A makes up a part of concept B. (cf. HAS_MERONYM)
EQ_IS_CAUSED_BY	A relation between two concepts in two different languages where concept B causes concept A. (cf. EQ_CAUSES)
EQ_UNSPECIFIED	A relation between two concepts in two different languages where concept A is strongly associated with concept B but no proper relation has been defined between the two.

EQ_HAS_SUBEVENT	A relation between two concepts in two different languages where concept B takes place during or as part of concept A, and whenever concept B takes place, concept A takes place. (cf. EQ_IS_SUBEVENT_OF)
EQ_IS_STATE_OF	A relation between two concepts in two different languages where concept B is qualified by concept A. (cf. EQ_BE_IN_STATE)
EQ_HAS_HYPONYM	A relation between two concepts in two languages where concept B is a type of concept A. (cf. EQ_HAS_HYPERONYM)
EQ_HAS_HYPERONYM	A relation between two concepts in two different languages where concept A is a type of concept B. (cf. EQ_HAS_HYPONYM)
EQ_HAS_HOLONYM	A relation between two concepts in two languages where concept A makes up a part of concept B. (cf. EQ_HAS_MERONYM)
EQ_CO_ROLE	A relation between two concepts in two languages where one concept undergoes an action in which the other concept is involved (bidirectional).
EQ_CAUSES	A relation between two concepts in two different languages where concept A causes concept B. (cf. EQ_IS_CAUSED_BY)
EQ_BE_IN_STATE	A relation between two concepts in two different languages where concept A is qualified by concept B (cf. EQ_IS_STATE_OF)
EQ_INVOLVED	A relation between two concepts in two different languages where concept B is typically involved in the action or event expressed by concept A. (cf. EQ_ROLE)
HAS_HOLO_LOCATION	A relation between two concepts where concept A is a place located in concept B. (cf. HAS_MERO_LOCATION)
HAS_HOLO_MADEOF	A relation between two concepts where concept B is made of concept A. (cf. HAS_MERO_MADEOF)
HAS_HOLO_PORTION	A relation between two concepts where concept B is an amount/piece/portion of concept A. (cf. HAS_MERO_PORTION)
HAS_HOLO_MEMBER	A relation between two concepts where concept A is a member/ element of concept B. (cf. HAS_MERO_MEMBER)
FUZZYNYM	A relation between two concepts where concept A is strongly associated with concept B but no proper relation has been defined between the two.
HAS_HOLO_PART	A relation between two concepts where concept A is a component of concept B. (cf. HAS_MERO_PART)

9.10 Pragmatics

Pragmatics	The study of the use of language in terms of the context in which it is spoken.
chronology	the use of a word is related to a certain period in time
neologism	A newly coined term.
oldfashioned	A term or lexeme that has fallen from fashion, but the meaning of which is readily recognizable.
connotation	A connotation is a commonly understood subjective cultural or emotional association that some word or phrase carries, in addition to the word's or phrase's explicit or literal meaning, which is its denotation.
pejorative	A word or grammatical form that connotes negativity and expresses a contempt or distaste.
euphemistic	Innocuous word, name, or phrase that replaces an offensive or suggestive one.
offensive	Register that expresses a situation that people should avoid because it is extremely offensive or embarrassing.
jocular	Fond of or characterized by joking; humorous or playful.
geography	Regional occurrence of a word
belg	used for words that are more frequently used in Belgian Dutch than in Dutch as used the Netherlands

register	the way in which something is said or written.
formal	A broad term for speech or writing marked by an impersonal, objective, and precise use of language.
informal	Language use characterized by spontaneous speech.
vulgar	Register of a term or text type that can be characterized as profane or socially unacceptable.
slang	An extremely informal register of a word, term, or text that is used in spoken and everyday language and less commonly in documents.
domain	area of activity, interest, or knowledge

9.11 Multiword expressions

expressionType	expression types refer to a range of subtypes in accordance with the degree of semantic non-compositionalit and syntactic fixedness of the multiword expression
proverb	a brief popular axiom or saying
idiom	a group of words in a fixed order that have a particular meaning that is different from the meaning of each word understood on its own

10 Cornetto-LMF

10.1 Design and model

The design of the Cornetto-LMF resource is in accordance with the ISO standard Lexical Markup Framework (LMF: ISO24613: 2008), a metamodel for lexical-semantic resources (Francopoulo et al., 2006, http://en.wikipedia.org/wiki/Lexical_Markup_Framework). It provides a common model for the creation and of use of lexical resources. LMF is composed of the following components:

- The core component which presents the basic hierarchy of information in a lexical entry.
- Extensions of the core component which describe additional components required for a specific lexical resource.

A component includes several classes (e.g. Sense class. Morphology class, etc.) and Classes have attributes. For instance, the Lexical Entry class has a part-of-speech attribute and an identifier attribute. For the design and implementation of the Cornetto Lexical Resource (cf. figure (2)) we used the following components and classes:

Core component: the root of the LMF core is a structural skeleton whose root is the Lexical Resource class. A Lexicon is a container for the words of a given language. Lexical Entry is a class that allows the connection between a form and a sense. Form is a class representing the way a word is spoken and/or written. A Form instance is associated with different Form Representation instances, for example, when the language has various ways of to express written forms. In Cornetto, a Lexical Entry instance is always linked to one Sense instance. Morphology and Morphosyntax are sub classes which describe the morphology and morphosyntactic characterisctics of the LE.

The **semantic component** is an extension of the core component for describing semantic information linked to the Sense class. The purpose is to describe one sense and its relations with other senses belonging to the same language.

The **syntactic component** is an extension of the core package describing syntactic information. It includes a syntactic behaviour class which is attached to the lexical entry. Detailed description of the syntactic behaviour of a lexical entry is defined by the Subcategorization Frame instance which is a class representing one syntactic construction.

The **WordNet LMF component** (Soria et al., 2009) is an extension that models WordNet lexical resources within the LMF model. The component describes two distinct repositories of Lexical Entries and Synsets, respectively. Links between the Lexical Entries and the synsets are defined as references from the Sense instances to the Synset instances.

In addition, we defined new classes for the description of **sentiment** and **pragmatics,** both of them optional subclasses of Sense.

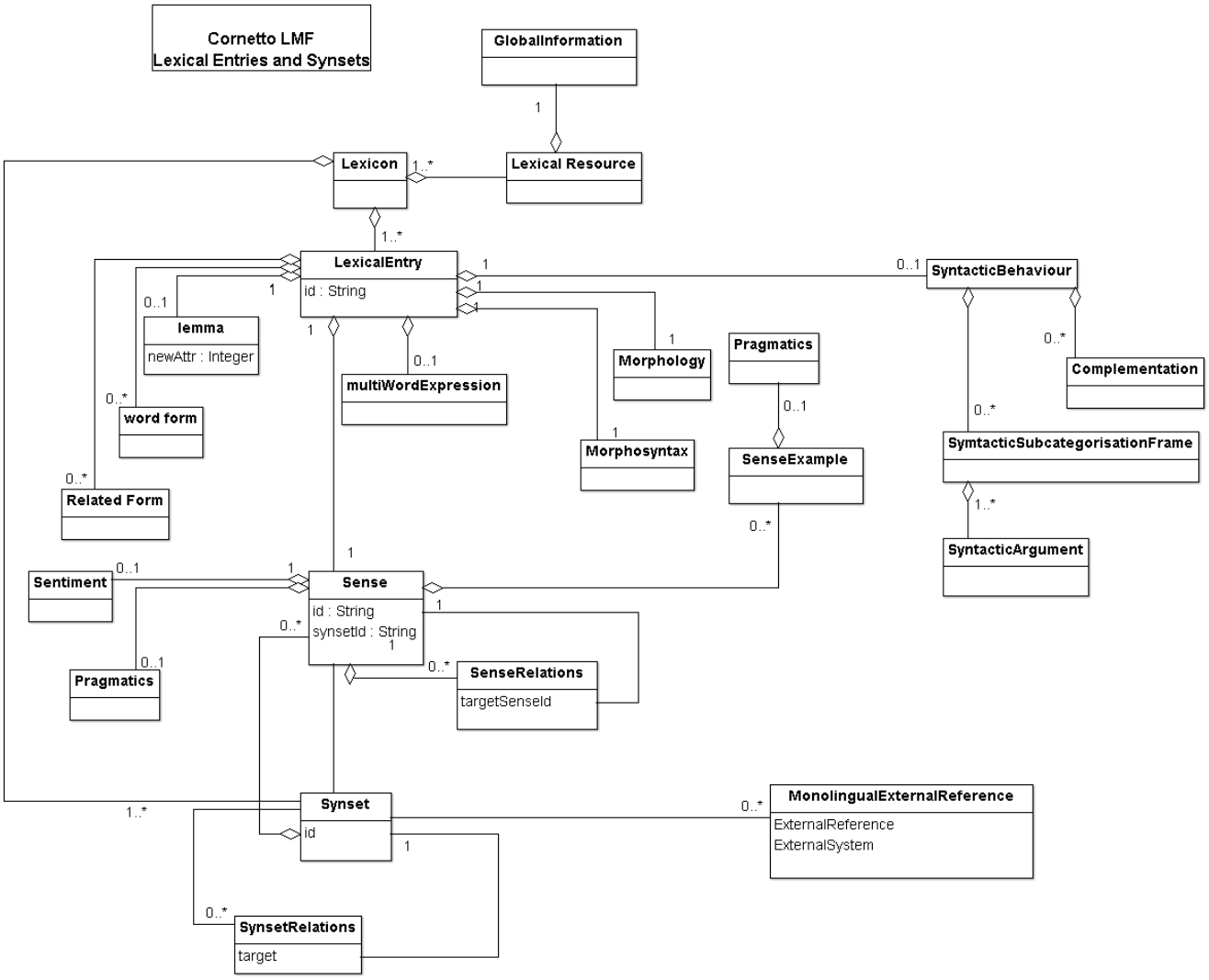

Figure (2) : schematic representation of the Cornetto-LMF model.

Although our model is in complete accordance with LMF's core model structure , we did not follow the common approach where each Lexical Entry groups different Senses of a word. In Cornetto, one Lexical Entry is linked to one Sense and each Sense of a polysemous word is linked to a unique Lexical Entry. This implies that all components, such as Morphology, MorphoSyntax and Syntax , unambiguously refer to one Lexical Entry-Sense pair.

10.2 Format : DTD

The data are represented in XML format which is conformant the following dtd. ELEMENTS correspond to LMF-components (cf. figure (2)).

```
<!-- -->
<!ELEMENT LexicalResource (GlobalInformation,Lexicon)>
<!ELEMENT GlobalInformation EMPTY>
<!ATTLIST    GlobalInformation label CDATA #REQUIRED>
<!ELEMENT Lexicon (LexicalEntry+, Synset*)>
<!ATTLIST    Lexicon label NMTOKEN  #REQUIRED >
<!ATTLIST    Lexicon language NMTOKEN #REQUIRED >
<!ATTLIST Lexicon languageCoding CDATA #REQUIRED >
<!ATTLIST Lexicon owner NMTOKEN #REQUIRED >
<!ELEMENT LexicalEntry (((Lemma, WordForms) | MultiwordExpression),  RelatedForms*, Morpholo-
gy*, MorphoSyntax*, SyntacticBehaviour*, Sense ) >
<!ATTLIST LexicalEntry id CDATA #REQUIRED >    <!-- id consists of lemma-partOfSpeech-sense
sequence number-->
<!ATTLIST LexicalEntry partOfSpeech (adverb | adjective | noun | verb | other ) #IMPLIED >
<!ATTLIST LexicalEntry formType (full | contraction | acronym | abbreviation) #IMPLIED>
<!ELEMENT Lemma EMPTY >
<!ATTLIST Lemma writtenForm CDATA #REQUIRED >
<!ATTLIST Lemma mode (infinitive) #IMPLIED > <!-- verb -->
<!-- -->
<!ELEMENT MultiwordExpression EMPTY >
<!ATTLIST MultiwordExpression writtenForm CDATA #REQUIRED >
<!ATTLIST MultiwordExpression expressionType (idiom | proverb) #IMPLIED >
<!-- -->
<!ELEMENT WordForms ( WordForm* ) >
<!ELEMENT WordForm EMPTY >
<!ATTLIST WordForm article CDATA #IMPLIED >
<!ATTLIST WordForm grammaticalNumber ( plural | singular ) #IMPLIED >  <!-- noun -->
<!ATTLIST WordForm comparison ( comparative | superlative ) #IMPLIED > <!-- adjective -->
<!ATTLIST WordForm tense ( pastTense | pastParticiple  ) #IMPLIED > <!-- verb -->
<!ATTLIST WordForm writtenForm CDATA #REQUIRED >
<!-- -->
<!ELEMENT  RelatedForms (RelatedForm+ ) >
<!ELEMENT RelatedForm EMPTY>
<!ATTLIST RelatedForm    writtenForm CDATA #REQUIRED>
<!ATTLIST RelatedForm    variantType  ( formVariant | spellingVariant) #REQUIRED>
<!-- -->
<!ELEMENT Morphology EMPTY >
<!ATTLIST Morphology morphoType ( compderiv | derivation | compound |  zero-derivation | x-
compound | wordgroup | phrasal ) #IMPLIED >
<!ATTLIST Morphology comparisonType (regular | irregular | mixed) #IMPLIED >  <!-- adjective -->
<!ATTLIST Morphology declinable (yes | no) #IMPLIED >  <!-- adjective -->
<!ATTLIST Morphology separability (  separable | inseparable ) #IMPLIED >          <!--verb-->
<!-- -->
 <!ELEMENT MorphoSyntax (auxiliaries*)  >
<!ATTLIST MorphoSyntax pronominalAndGrammaticalGender ( f | m | n | fn | m_f  | mfn | mn  | mf )
#IMPLIED >  <!--noun-->
<!-- -->
<!ATTLIST MorphoSyntax adverbialUsage ( yes | no) #IMPLIED >    <!--adjective-->
<!ATTLIST MorphoSyntax position ( attributive | predicative | attrpred ) #IMPLIED > <!--adjective-->
```

```
<!ELEMENT  auxiliaries EMPTY>   <!---verb-->
<!ATTLIST  auxiliaries auxiliary (hebben | zijn) #IMPLIED > <!---verb-->
<!ATTLIST MorphoSyntax reflexivity (  optionalReflexive | reflexive  ) #IMPLIED ><!---verb--> <!--
verb-->
<!--  -->
<!ELEMENT SyntacticBehaviour ( Complementation* , SyntacticSubcategorisationFrame* ) >
<!ATTLIST SyntacticBehaviour valency (mono | di | tri ) #IMPLIED>
<!ATTLIST SyntacticBehaviour transitivity (transitive | intransitive ) #IMPLIED>
<!--  -->
<!ELEMENT Complementation EMPTY >
<!ATTLIST Complementation complement ( datclause | ofclause | dancomp | whclause |  oblobj |
quant | oblprep |  factive | fixprep | omtoinf | prep | psmodnoun | toinf ) #IMPLIED >
<!ATTLIST Complementation preposition (wegens | te |  zonder | achter | ter...van | met...van |
in...van | met...tot | op...bij |  per | aangaande | langs | inzake | boven | door | binnen | via |
ten...van | uit | omtrent | om | aan | als | bij | in | jegens | met | naar | onder | onder...van | op |
over | rond | tegen | tegenover | tot | tussen | van | voor ) #IMPLIED >
<!ELEMENT SyntacticSubcategorisationFrame (syntacticArgument+ ) >
<!ELEMENT syntacticArgument EMPTY >
<!ATTLIST syntacticArgument constituent   (  nil | np | pp | s | ap) #IMPLIED >
<!ATTLIST syntacticArgument function   ( nil | specifyingComplement | objectComplement | direc-
tObject | indirectObject | prepositionalObject | specifyingObject) #IMPLIED >
<!ATTLIST syntacticArgument preposition   CDATA  #IMPLIED >
<!ATTLIST syntacticArgument complementizer   (  dat | omte | te| WH | hoe | of ) #IMPLIED >
<!--  -->
<!ELEMENT Sense ( Sentiment | Pragmatics | (Semantics-verb | Semantics-noun | Semantics-
adjective) | SenseExamples | SenseRelations | MorphoSyntax )* >
<!ATTLIST Sense senseId CDATA #REQUIRED >
<!ATTLIST Sense synset NMTOKEN #IMPLIED > <!--refers to synset Id-->
<!ATTLIST Sense definition CDATA #REQUIRED >
<!--  -->
<!ELEMENT Sentiment EMPTY >
<!ATTLIST Sentiment polarity ( negative | positive ) #REQUIRED >
<!ATTLIST Sentiment  externalReference NMTOKEN #REQUIRED >
<!--  -->
<!ELEMENT Pragmatics (Domains)* >
<!ELEMENT Domains   EMPTY >
<!ATTLIST Domains domain (accoustics | administration | aeronautics | agriculture | alimentation |
anatomy | anthropology | archeology | architecture | art | artisanship | astrology | astronomy | as-
tronautics | astronomie | biochemistry | biology | botany | building_industry | chemistry | cinema |
commerce | computer_science | cycling | dance | diplomacy | doctrines | ecology | economy | elec-
tronics_electricity | empty | engineering | fashion | fishing | folklore | gastronomy | geography | ge-
ology | golf | herladry | history | housekeeping | hunting | hydraulics | industry | insurance | law |
linguistics | literature | mathematics | media | medicine | merchant_navy | meteorology | metrology |
military | money | music | mythology | pedagogy | pharmacy | philosophy | photography | physics |
play | politics | psychology | publishing | railway | religion | school | science | sculpture | sexuality |
soccer | sociology | sport | state | swimming | telecommunication | tennis | theatre | theology |
transport | zoology ) #IMPLIED >
<!ATTLIST Pragmatics chronology (  neologism | oldfashioned ) #IMPLIED >
<!ATTLIST Pragmatics connotation ( euphemistic | jocular | offensive | pejorative ) #IMPLIED >
<!ATTLIST Pragmatics geography ( belg | ind | dialect) #IMPLIED >
<!ATTLIST Pragmatics style ( formal | informal | slang | vulgar | archaic) #IMPLIED >
<!--  -->
<!ELEMENT Semantics-noun (semanticShifts-noun)*>
<!ATTLIST Semantics-noun countability (count | uncount | count_uncount | coll | mass | plurtant)
#IMPLIED >
<!ATTLIST Semantics-noun reference (common|proper) #IMPLIED >
<!ATTLIST Semantics-noun semanticType (substance | abstract | animate | artefact | concrete | con-
crother | dynamic | human | institute | measure | nondynamic | nonhuman | place |  time )
#IMPLIED>
```

```
<!ATTLIST Semantics-noun semanticSubType CDATA #IMPLIED >
<!ELEMENT semanticShifts-noun EMPTY >
<!ATTLIST semanticShifts-noun semanticType  ( animate | abstract | artefact | concrete | concrother
| dynamic | human |  institute | measure | nondynamic | nonhuman |  place |  substance | time )
#IMPLIED>
 <!--  -->
<!ELEMENT Semantics-adjective (semanticShifts-adjective)*>
<!ATTLIST Semantics-adjective semanticType ( substance | place |  temp | stuff | colour | phyper |
emomen | abstract) #IMPLIED>
<!ELEMENT semanticShifts-adjective EMPTY >
<!ATTLIST semanticShifts-adjective semanticType  (  place | temp | stuff | colour | phyper | emomen
| abstract ) #IMPLIED>
<!--  -->
<!ELEMENT Semantics-verb (semanticTypes*)  >
<!ELEMENT semanticTypes  EMPTY >
<!ATTLIST semanticTypes semanticType (action | process | state)   #IMPLIED >
<!ATTLIST semanticTypes semanticFeatureSet ( stcognt1 | stcognt2 |stcognt3 |state1 | state2
|state3 |process1 | process2 |process3 |prmvmt1 | prmvmt2 | prmvmt3 | possess2 | possess3 | prc-
ognt1 | prcognt2 |prcognt3 |location1 | location2 | mvmt1 | mvmt2 | mvmt3 | echprod1 | echprod2 |
echprod3 | action1 | action2 | action3 | cognt1 | cognt2 | cognt3 )  #IMPLIED>
<!--  -->
<!ELEMENT SenseExamples ( SenseExample* ) >
<!--  -->
<!ELEMENT SenseExample ( canonicalForm | Pragmatics | Semantics_ex | Syntax_ex | textualForm
)* >
<!ATTLIST SenseExample id NMTOKEN #REQUIRED >
<!--  -->
<!ELEMENT canonicalForm EMPTY >
<!ATTLIST canonicalForm phrase  (s | np | vp  | ap | pp) #IMPLIED >
<!ATTLIST canonicalForm canonicalform CDATA #IMPLIED >
<!ATTLIST canonicalForm expressionType ( freeCombination | slogan | properName | term |  lexi-
calCollocation | grammaticalCollocation | pragmaticFormula ) #IMPLIED>  <!--enumeration-->
<!--  -->
<!ELEMENT Semantics_ex  (lex-collocator)* >
<!ATTLIST Semantics_ex gracol-complem CDATA #IMPLIED >
<!ATTLIST Semantics_ex gracol-gramword ( expletive | object | material | subject-matter | condition
| relation | source | cause | circumstance | direction | goal | location | manner | means | measure |
modality | origin | place | purpose | reason | restriction | result | time ) #IMPLIED >
<!ATTLIST Semantics_ex definition CDATA #IMPLIED >
<!--  -->
<!ELEMENT  lex-collocator EMPTY >
<!ATTLIST lex-collocator collocator ( action | antibonus | anticause | antiexist | antifunct | antimag-
nus | antimanif | antintense | antioper1 | antioper2 | antipossess | antireal | antiverus | bonus |
cause | causenotre | cont | decrease | degrade | dimin | enum | excess | exist | fin | funct | incept |
intens | liqui | magnus | malus | manif | measure | minusexist | modif | mult | oper | oper1 | oper2 |
process | support | place | plusexist | possess | qualifier | quantifier | real | sing | son | state | type |
upgrade | verus ) #IMPLIED >
<!ATTLIST Semantics_ex lexcol-speccollocator (appear | attend | caretake | consult | consume | con-
tainer | create | engage | evoke | give | hold | install | introduce | keep | lose | own | portion | prac-
tise | prepare | putoff | receive | refuse | reign | remove | solve | suffer | take | takeplace | turnon |
use | violate | wear)  #IMPLIED >
<!--  -->
<!ELEMENT  Syntax_ex (combiWord+ ) >
<!ELEMENT combiWord EMPTY >
<!ATTLIST combiWord partOfSpeech ( adjective | adverb | article | conjunction | noun | numeral|
preposition | pronoun | verb ) #IMPLIED >
<!ATTLIST combiWord lemma CDATA #IMPLIED >
<!--  -->
<!ELEMENT textualForm EMPTY >
```

```
<!ATTLIST textualForm textualform CDATA #IMPLIED >
<!ATTLIST textualForm phrase (s | np | vp | ap | pp) #IMPLIED >
<!-- -->
<!ELEMENT SenseRelations ( SenseGroup* ) >
<!ELEMENT SenseGroup EMPTY >
<!ATTLIST SenseGroup relationType ( co-annotation | co-hyponyms | co-synonyms | co-relations)
#REQUIRED >
<!ATTLIST SenseGroup targetSenseId NMTOKEN #REQUIRED >
<!-- -->
<!ELEMENT Synset ( Definition | MonolingualExternalRefs | SynsetRelations )* >
<!ATTLIST Synset baseConcept CDATA #IMPLIED >
<!ATTLIST Synset id ID #REQUIRED >
<!-- -->
<!ELEMENT Definition EMPTY >
<!ATTLIST Definition gloss CDATA #REQUIRED >
<!-- -->
<!ELEMENT MonolingualExternalRefs ( MonolingualExternalRef* ) >
<!-- -->
<!ELEMENT MonolingualExternalRef ( Meta? ) >
<!ATTLIST MonolingualExternalRef externalReference CDATA #REQUIRED >
<!ATTLIST MonolingualExternalRef externalSystem ( SUMO | pwn-20 | pwn-30 | wordnet_domain )
#REQUIRED >
<!ATTLIST MonolingualExternalRef relType CDATA  #IMPLIED >
<!-- -->
<!ELEMENT Meta EMPTY >
<!ATTLIST Meta author CDATA #IMPLIED >
<!ATTLIST Meta confidence NMTOKEN #REQUIRED >
<!ATTLIST Meta date NMTOKEN #IMPLIED >
<!ATTLIST Meta source CDATA #REQUIRED >
<!ATTLIST Meta status ( yes ) #IMPLIED >
<!-- -->
<!ELEMENT SynsetRelations ( SynsetRelation* ) >
<!-- -->
<!ELEMENT SynsetRelation ( Meta? ) >
<!ATTLIST SynsetRelation relType (XPOS_NEAR_SYNONYM | HAS_XPOS_HYPONYM |
ROLE_SOURCE_DIRECTION | BE_IN_STATE  | CAUSES  | CO_AGENT_INSTRUMENT  |
CO_AGENT_PATIENT  | CO_AGENT_RESULT  | CO_INSTRUMENT_AGENT  |
CO_INSTRUMENT_PATIENT  | CO_INSTRUMENT_RESULT  | CO_PATIENT_AGENT  |
CO_PATIENT_INSTRUMENT  | CO_PATIENT_RESULT  | CO_RESULT_AGENT  |
CO_RESULT_INSTRUMENT  | CO_RESULT_PATIENT  | CO_ROLE  | FUZZYNYM  |
HAS_HOLO_LOCATION  | HAS_HOLO_MADEOF  | HAS_HOLO_MEMBER  | HAS_HOLO_PART  |
HAS_HOLO_PORTION  | HAS_HOLONYM  | HAS_HYPERONYM  | HAS_HYPONYM  |
HAS_MERO_LOCATION  | HAS_MERO_MADEOF  | HAS_MERO_MEMBER  | HAS_MERO_PART  |
HAS_MERO_PORTION  | HAS_MERONYM  | HAS_SUBEVENT  | HAS_XPOS_HYPERONYM  | IN_MANNER
| INVOLVED  | INVOLVED_AGENT  | INVOLVED_DIRECTION  | INVOLVED_INSTRUMENT  |
INVOLVED_LOCATION  | INVOLVED_PATIENT  | INVOLVED_RESULT  |
INVOLVED_SOURCE_DIRECTION  | INVOLVED_TARGET_DIRECTION  | IS_CAUSED_BY  |
IS_SUBEVENT_OF  | MANNER_OF  | NEAR_ANTONYM  | NEAR_SYNONYM  | ROLE  | ROLE_AGENT  |
ROLE_DIRECTION  | ROLE_INSTRUMENT  | ROLE_LOCATION  | ROLE_PATIENT  | ROLE_RESULT  |
ROLE_TARGET_DIRECTION  | STATE_OF  | XPOS_FUZZYNYM  | XPOS_NEAR_ANTONYM  )  #IMPLIED
>
<!ATTLIST SynsetRelation target CDATA #REQUIRED >
```

11 References

Francopoulo G., Bel N., George M., Calzolari N., Monachini M., Pet M., Soria C. (2007) Lexical Markup Framework: ISO standard for semantic information in NLP lexicons. In: GLDV (Gesellschaft für linguistische Datenverarbeitung), Tubingen

Görög, A., Vossen, P. (2010) *Computer Assisted Semantic Annotation in the DutchSemCor Project*. In Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10). 2010, Malta, Valletta

Mel'čuk, I. A. (1996) *Lexical functions: a tool for the description of lexical relations in a lexicon*. In: Wanner, L. (ed.), Lexical functions in lexicography and natural language processing. Amsterdam: Benjamins

Maks, I., P. Vossen, R. Segers and H. van der Vliet (2008). *Encoding Adjectives in the Dutch semantic database Cornetto*. In Proceedings of LREC-2008, Marrakech, Morocco, 2008.

Maks, I. and P. Vossen (2011) *Different Approaches to Automatic Polarity Annotation at Synset Level.*In: Proceedings of the First International Workshop on Lexical Resources, WoLeR, Slovenia.

Martin, W., I.Maks, M. Groot and S. Bopp (2005) RBN (Referentie Bestand Nederlands) documentation , Internal Report VU Amsterdam, http://tst-centrale.org/ images/ stories/producten /documentatie/ rbn_documentatie_nl. pdf

Soria C., Monachini M., Vossen P. (2009) *Wordnet-LMF: fleshing out a standardized Format for Wordnet Interoperability*. In: proceedings of CHI-2009, Boston, USA.

Vliet, H. van der (2007) *The Referentiebestand Nederlands as a Mult-Purpose Lexical Database*' In: International Journal of Lexicography 20(3): 221-238

Vossen, P., Hofmann, K., de Rijke, M., Tjong Kim Sang, E., and Deschacht, K. (2007). *The Cornetto Database: Architecture and User-Scenarios*. In DIR 2007.

Vossen, P., I.Maks, R. Segers and H. van der Vliet (2008). *Integrating Lexical Units, Synsets, and Ontology in the Cornetto Database*. In Proceedings of LREC-2008, Marrakech, Morocco.

Vossen P., Görög, A., Izquierdo, R. (2012). *DutchSemCor: targeting the ideal sense-tagged corpus*. In: Proceedings of the Eighth conference on International Language Resources and Evaluation (LREC'12), Istanbul, Turkey.

Vossen P., Görög, A., Laan, F., Van Gompel, M., Izquierdo, R. , Van den Bosch, A. (2011). *DutchSemCor: building a semantically annotated corpus for Dutch*. In: Proceedings of Electronic Lexicography in the 21st century: New Applications for new users (eLEX2011), Bled, Slovenia.

Vossen, P., I.Maks, R.Segers, H.van der Vliet, M.F. Moens, K.Hofmann, E.Tjong Kim Sang, M.de Rijke (2013) *Cornetto: a lexical semantic database for Dutch*, Chapter in: P. Spyns & J. Odijk (eds): Essential Speech and Language Technology for Dutch, Results by the STEVIN-programme, Publ. Springer series Theory and Applications of Natural Language Processing, ISBN 978-3-642-30909-0.

Windhouwer,M., I. Schuurman & S. E. Wright (2013). *Collaboratively Defining Widely Accepted Linguistic Data Categories in the ISOcat Data Category Registry*. In: ESWC2013 Workshop on Semantic Web Collaborative Spaces (SWCS2013). Montpellier, France.

Appendix I: Cornetto data categories and ISOcat urls

The following list is a list of Cornetto data categories in alphabetic order. The list gives the Data category names (column 1), the ISOcat url (column2) amd a broader linguistic category (column3) which corresponds to the categorization of data categories in section 9.

abbreviation	http://www.isocat.org/datcat/DC-5036	LexicalEntry
abstract	*to be decided*	Sense and Semantics
acronym	*to be decided*	LexicalEntry
action	http://www.isocat.org/datcat/DC-5974	Sense and Semantics
action1	http://www.isocat.org/datcat/DC-5943	Sense and Semantics
action2	http://www.isocat.org/datcat/DC-5944	Sense and Semantics
action3	http://www.isocat.org/datcat/DC-5945	Sense and Semantics
adjective	http://www.isocat.org/datcat/DC-1230	LexicalEntry
adverb	http://www.isocat.org/datcat/DC-1232	Morphology and Morphosyntax
adverbialUsage	http://www.isocat.org/datcat/DC-5922	Morphology and Morphosyntax
animate	*to be decided*	Sense and Semantics
ap	http://www.isocat.org/datcat/DC-5767	Syntax
ap	http://www.isocat.org/datcat/DC-5767	Syntax
artefact	http://www.isocat.org/datcat/DC-5991	Sense and Semantics
article	http://www.isocat.org/datcat/DC-1892	Morphology and Morphosyntax
article	http://www.isocat.org/datcat/DC-3589	Morphology and Morphosyntax
attributive	http://www.isocat.org/datcat/DC-5242	Morphology and Morphosyntax
attrpred	http://www.isocat.org/datcat/DC-5791	Morphology and Morphosyntax
auxiliary	http://www.isocat.org/datcat/DC-2262	Morphology and Morphosyntax
BE_IN_STATE	http://www.isocat.org/datcat/DC-5878	synset
belg	http://www.isocat.org/datcat/DC-5936	Pragmatics
CanonicalForm	http://www.isocat.org/datcat/DC-6002	Examples
CAUSES	http://www.isocat.org/datcat/DC-5827	synset
chronology	http://www.isocat.org/datcat/DC-5921	Pragmatics
CO_AGENT_INSTRUMENT	http://www.isocat.org/datcat/DC-5879	synset
CO_AGENT_PATIENT	http://www.isocat.org/datcat/DC-5838	synset
CO_AGENT_RESULT	http://www.isocat.org/datcat/DC-5849	synset
CO_INSTRUMENT_AGENT	http://www.isocat.org/datcat/DC-5860	synset
CO_INSTRUMENT_PATIENT	http://www.isocat.org/datcat/DC-5864	synset
CO_INSTRUMENT_RESULT	http://www.isocat.org/datcat/DC-5876	synset
CO_PATIENT_AGENT	http://www.isocat.org/datcat/DC-5812	synset
CO_PATIENT_INSTRUMENT	http://www.isocat.org/datcat/DC-5806	synset
CO_RESULT_AGENT	http://www.isocat.org/datcat/DC-5807	synset
CO_RESULT_INSTRUMENT	http://www.isocat.org/datcat/DC-5808	synset
CO_ROLE	http://www.isocat.org/datcat/DC-5809	synset
co-annotations	http://www.isocat.org/datcat/DC-5912	SenseRelation
cognt1	http://www.isocat.org/datcat/DC-5946	Sense and Semantics
cognt2	http://www.isocat.org/datcat/DC-5947	Sense and Semantics
cognt3	http://www.isocat.org/datcat/DC-5952	Sense and Semantics
co-hyponyms	http://www.isocat.org/datcat/DC-5913	SenseRelation

colour	http://www.isocat.org/datcat/DC-5987	Sense and Semantics
combiWord	http://www.isocat.org/datcat/DC-6003	Examples
common	http://www.isocat.org/datcat/DC-1256	Sense and Semantics
comparative	http://www.isocat.org/datcat/DC-4924	Morphology and Morphosyntax
comparisonType	http://www.isocat.org/datcat/DC-5924	Morphology and Morphosyntax
compderiv	http://www.isocat.org/datcat/DC-5902	Morphology and Morphosyntax
complement	http://www.isocat.org/datcat/DC-5865	Syntax
complementizer	http://www.isocat.org/datcat/DC-5923	Syntax
compound	*to be decided*	Morphology and Morphosyntax
concrete	http://www.isocat.org/datcat/DC-5984	Sense and Semantics
concrother	http://www.isocat.org/datcat/DC-5976	Sense and Semantics
conjunction	http://www.isocat.org/datcat/DC-1260	Morphology and Morphosyntax
connotation	http://www.isocat.org/datcat/DC-5929	Pragmatics
constituent	http://www.isocat.org/datcat/DC-5757	Syntax
contraction	*to be decided*	LexicalEntry
co-relations	http://www.isocat.org/datcat/DC-5914	SenseRelation
co-synonyms	http://www.isocat.org/datcat/DC-5915	SenseRelation
count	*to be decided*	Sense and Semantics
count_uncount	*to be decided*	Sense and Semantics
countability	*to be decided*	Sense and Semantics
dancomp	http://www.isocat.org/datcat/DC-5867	Syntax
dat	http://www.isocat.org/datcat/DC-5891	Syntax
definition	http://www.isocat.org/datcat/DC-1972	Morphology and Morphosyntax
degree	http://www.isocat.org/datcat/DC-4920	Morphology and Morphosyntax
derivation	*to be decided*	Morphology and Morphosyntax
di	http://www.isocat.org/datcat/DC-5765	Syntax
directObject	http://www.isocat.org/datcat/DC-1274	Syntax
domain	http://www.isocat.org/datcat/DC-2212	domain
dynamic	http://www.isocat.org/datcat/DC-5977	Sense and Semantics
echprod2	http://www.isocat.org/datcat/DC-5941	Sense and Semantics
echprod3	http://www.isocat.org/datcat/DC-5970	Sense and Semantics
emomen	http://www.isocat.org/datcat/DC-5990	Sense and Semantics
EQ_BE_IN_STATE	http://www.isocat.org/datcat/DC-5811	synset
EQ_CAUSES	http://www.isocat.org/datcat/DC-5877	synset
EQ_CO_ROLE	http://www.isocat.org/datcat/DC-5810	synset
EQ_HAS_HOLONYM	http://www.isocat.org/datcat/DC-5805	synset
EQ_HAS_HYPERONYM	http://www.isocat.org/datcat/DC-5813	synset
EQ_HAS_HYPONYM	http://www.isocat.org/datcat/DC-5850	synset
EQ_HAS_MERONYM	http://www.isocat.org/datcat/DC-5851	synset
EQ_HAS_SUBEVENT	http://www.isocat.org/datcat/DC-5852	synset
EQ_INVOLVED	http://www.isocat.org/datcat/DC-5883	synset
EQ_IS_CAUSED_BY	http://www.isocat.org/datcat/DC-5884	synset
EQ_IS_STATE_OF	http://www.isocat.org/datcat/DC-5855	synset
EQ_IS_SUBEVENT_OF	http://www.isocat.org/datcat/DC-5856	synset
EQ_NEAR_SYNONYM	http://www.isocat.org/datcat/DC-5857	synset
EQ_ROLE	http://www.isocat.org/datcat/DC-5858	synset
EQ_SYNONYM	http://www.isocat.org/datcat/DC-5880	synset
EQ_UNSPECIFIED	http://www.isocat.org/datcat/DC-5859	synset
euphemistic	http://www.isocat.org/datcat/DC-5931	Pragmatics
expressionType	http://www.isocat.org/datcat/DC-5905	Examples

external Reference	http://www.isocat.org/datcat/DC-1975	external Reference
external System	http://www.isocat.org/datcat/DC-1974	external Reference
f	http://www.isocat.org/datcat/DC-5801	Morphology and Morphosyntax
factive	http://www.isocat.org/datcat/DC-5873	Syntax
fixprep	http://www.isocat.org/datcat/DC-5870	Syntax
fn	http://www.isocat.org/datcat/DC-5797	Morphology and Morphosyntax
formal	http://www.isocat.org/datcat/DC-1992	Pragmatics
formType	*to be decided*	LexicalEntry
formVariant	http://www.isocat.org/datcat/DC-5917	LexicalEntry
freeCombination	http://www.isocat.org/datcat/DC-5910	Examples
full	http://www.isocat.org/datcat/DC-5911	LexicalEntry
function	*to be decided*	Syntax
FUZZYNYM	http://www.isocat.org/datcat/DC-5832	synset
geography	http://www.isocat.org/datcat/DC-5935	Pragmatics
GlobalInformation	http://www.isocat.org/datcat/DC-5994	LexicalEntry
grammaticalCollocation	http://www.isocat.org/datcat/DC-5907	Examples
grammaticalNumber	http://www.isocat.org/datcat/DC-251	Morphology and Morphosyntax
HAS_HOLO_LOCATION	http://www.isocat.org/datcat/DC-5816	synset
HAS_HOLO_MADEOF	http://www.isocat.org/datcat/DC-5831	synset
HAS_HOLO_MEMBER	http://www.isocat.org/datcat/DC-5830	synset
HAS_HOLO_PART	http://www.isocat.org/datcat/DC-5817	synset
HAS_HOLO_PORTION	http://www.isocat.org/datcat/DC-5828	synset
HAS_HOLONYM	http://www.isocat.org/datcat/DC-5836	synset
HAS_HYPERONYM	http://www.isocat.org/datcat/DC-5824	synset
HAS_HYPONYM	http://www.isocat.org/datcat/DC-5823	synset
HAS_MERO_LOCATION	http://www.isocat.org/datcat/DC-5821	synset
HAS_MERO_MADEOF	http://www.isocat.org/datcat/DC-5820	synset
HAS_MERO_MEMBER	http://www.isocat.org/datcat/DC-5819	synset
HAS_MERO_PART	http://www.isocat.org/datcat/DC-5846	synset
HAS_MERO_PORTION	http://www.isocat.org/datcat/DC-5826	synset
HAS_MERONYM	http://www.isocat.org/datcat/DC-5854	synset
HAS_SUBEVENT	http://www.isocat.org/datcat/DC-5853	synset
HAS_XPOS_HYPERONYM	http://www.isocat.org/datcat/DC-5882	synset
HAS_XPOS_HYPONYM	http://www.isocat.org/datcat/DC-5848	synset
hoe	http://www.isocat.org/datcat/DC-5895	Syntax
human	http://www.isocat.org/datcat/DC-5978	Sense and Semantics
idiom	*to be decided*	mwe
IN_MANNER	http://www.isocat.org/datcat/DC-5837	synset
indirectObject	http://www.isocat.org/datcat/DC-1310	Syntax
infinitive	*to be decided*	Morphology and Morphosyntax
informal	*to be decided*	Pragmatics
institute	http://www.isocat.org/datcat/DC-5979	Sense and Semantics
intransitive	http://www.isocat.org/datcat/DC-1322	Syntax
INVOLVED	http://www.isocat.org/datcat/DC-5847	synset
INVOLVED_AGENT	http://www.isocat.org/datcat/DC-5885	synset
INVOLVED_DIRECTION	http://www.isocat.org/datcat/DC-5845	synset
INVOLVED_INSTRUMENT	http://www.isocat.org/datcat/DC-5844	synset
INVOLVED_LOCATION	http://www.isocat.org/datcat/DC-5843	synset
INVOLVED_PATIENT	http://www.isocat.org/datcat/DC-5842	synset
INVOLVED_RESULT	http://www.isocat.org/datcat/DC-5841	synset

INVOLVED_SOURCE_DIRECTION	http://www.isocat.org/datcat/DC-5840	synset
INVOLVED_TARGET_DIRECTION	http://www.isocat.org/datcat/DC-5815	synset
irregular	http://www.isocat.org/datcat/DC-5925	Morphology and Morphosyntax
IS_CAUSED_BY	http://www.isocat.org/datcat/DC-5839	synset
IS_SUBEVENT_OF	http://www.isocat.org/datcat/DC-5818	synset
jocular	http://www.isocat.org/datcat/DC-5932	Pragmatics
Lemma	http://www.isocat.org/datcat/DC-1324	Morphology and Morphosyntax
lexicalCollocation	http://www.isocat.org/datcat/DC-5906	Examples
LexicalEntry	http://www.isocat.org/datcat/DC-5505	LexicalEntry
LexicalResource	http://www.isocat.org/datcat/DC-5992	LexicalEntry
Lexicon	http://www.isocat.org/datcat/DC-4360	LexicalEntry
location1	http://www.isocat.org/datcat/DC-5965	Sense and Semantics
location2	http://www.isocat.org/datcat/DC-5966	Sense and Semantics
m	http://www.isocat.org/datcat/DC-5800	Morphology and Morphosyntax
m_f	http://www.isocat.org/datcat/DC-5799	Morphology and Morphosyntax
MANNER_OF	http://www.isocat.org/datcat/DC-5919	synset
mass	*to be decided*	Sense and Semantics
measure	http://www.isocat.org/datcat/DC-5980	Sense and Semantics
mf	http://www.isocat.org/datcat/DC-5795	Morphology and Morphosyntax
mfn	http://www.isocat.org/datcat/DC-5796	Morphology and Morphosyntax
mixed	http://www.isocat.org/datcat/DC-5927	Morphology and Morphosyntax
mn	http://www.isocat.org/datcat/DC-5798	Morphology and Morphosyntax
mono	http://www.isocat.org/datcat/DC-5764	Syntax
MonolingualExternalRef	http://www.isocat.org/datcat/DC-5886	external Reference
mood	http://www.isocat.org/datcat/DC-1427	Morphology and Morphosyntax
Morphology	http://www.isocat.org/datcat/DC-2638	Morphology and Morphosyntax
morphoType	http://www.isocat.org/datcat/DC-5887	Morphology and Morphosyntax
mvmt1	http://www.isocat.org/datcat/DC-5967	Sense and Semantics
mvmt2	http://www.isocat.org/datcat/DC-5968	Sense and Semantics
mvmt3	http://www.isocat.org/datcat/DC-5948	Sense and Semantics
n	http://www.isocat.org/datcat/DC-5802	Morphology and Morphosyntax
NEAR_ANTONYM	http://www.isocat.org/datcat/DC-5971	synset
NEAR_SYNONYM	http://www.isocat.org/datcat/DC-5972	synset
negative	http://www.isocat.org/datcat/DC-5540	sentiment
neologism	*to be decided*	Pragmatics
nondynamic	http://www.isocat.org/datcat/DC-5981	Sense and Semantics
nonhuman	http://www.isocat.org/datcat/DC-5982	Sense and Semantics
noun	http://www.isocat.org/datcat/DC-1333	LexicalEntry
np	http://www.isocat.org/datcat/DC-2256	Syntax
np	http://www.isocat.org/datcat/DC-2256	Syntax
numeral	http://www.isocat.org/datcat/DC-1334	Morphology and Morphosyntax
objectComplement	*to be decided*	Syntax
oblobj	http://www.isocat.org/datcat/DC-5871	Syntax
oblprep	http://www.isocat.org/datcat/DC-5874	Syntax
of	http://www.isocat.org/datcat/DC-5896	Syntax
offensive	*to be decided*	Pragmatics
oldfashioned	http://www.isocat.org/datcat/DC-5787	Pragmatics
omte	http://www.isocat.org/datcat/DC-5892	Syntax
omtoinf	http://www.isocat.org/datcat/DC-5866	Syntax
optionalReflexive	http://www.isocat.org/datcat/DC-5934	Morphology and Morphosyntax

partOfSpeech	http://www.isocat.org/datcat/DC-1345	LexicalEntry
pastParticiple	http://www.isocat.org/datcat/DC-4963	LexicalEntry
pastTense	*to be decided*	LexicalEntry
pejorative	http://www.isocat.org/datcat/DC-5933	Pragmatics
phrasal	http://www.isocat.org/datcat/DC-5901	Morphology and Morphosyntax
phraseType	http://www.isocat.org/datcat/DC-5769	Syntax
phyper	http://www.isocat.org/datcat/DC-5989	Sense and Semantics
place	http://www.isocat.org/datcat/DC-5988	Sense and Semantics
plural	http://www.isocat.org/datcat/DC-253	Morphology and Morphosyntax
pluraleTantum	*to be decided*	Sense and Semantics
polarity	http://www.isocat.org/datcat/DC-5537	sentiment
position	http://www.isocat.org/datcat/DC-5930	Morphology and Morphosyntax
positive	http://www.isocat.org/datcat/DC-5538	sentiment
possess2	http://www.isocat.org/datcat/DC-5949	Sense and Semantics
possess3	http://www.isocat.org/datcat/DC-5942	Sense and Semantics
pp	http://www.isocat.org/datcat/DC-5768	Syntax
pp	http://www.isocat.org/datcat/DC-5768	Syntax
pragmaticFormula	http://www.isocat.org/datcat/DC-5908	Examples
Pragmatics	http://www.isocat.org/datcat/DC-2643	Pragmatics
prcognt1	http://www.isocat.org/datcat/DC-5969	Sense and Semantics
prcognt2	http://www.isocat.org/datcat/DC-5964	Sense and Semantics
prcognt3	http://www.isocat.org/datcat/DC-5963	Sense and Semantics
predicative	http://www.isocat.org/datcat/DC-4943	Morphology and Morphosyntax
preposition	http://www.isocat.org/datcat/DC-1366	Syntax
preposition	http://www.isocat.org/datcat/DC-2757	Syntax
prmvmt1	http://www.isocat.org/datcat/DC-5953	Sense and Semantics
prmvmt2	http://www.isocat.org/datcat/DC-5951	Sense and Semantics
prmvmt3	http://www.isocat.org/datcat/DC-5950	Sense and Semantics
process	http://www.isocat.org/datcat/DC-5985	Sense and Semantics
process1	http://www.isocat.org/datcat/DC-5956	Sense and Semantics
process2	http://www.isocat.org/datcat/DC-5955	Sense and Semantics
process3	http://www.isocat.org/datcat/DC-5954	Sense and Semantics
pronominalAndGrammaticalGender	http://www.isocat.org/datcat/DC-5793	Morphology and Morphosyntax
pronoun	http://www.isocat.org/datcat/DC-1370	Morphology and Morphosyntax
proper	http://www.isocat.org/datcat/DC-1371	Sense and Semantics
properName	*to be decided*	Examples
proverb	*to be decided*	mwe
psmodnoun	http://www.isocat.org/datcat/DC-5869	Syntax
quant	http://www.isocat.org/datcat/DC-5872	Syntax
reference	http://www.isocat.org/datcat/DC-4908	Sense and Semantics
reflexive	http://www.isocat.org/datcat/DC-3842	Morphology and Morphosyntax
reflexivity	http://www.isocat.org/datcat/DC-5928	Morphology and Morphosyntax
register	http://www.isocat.org/datcat/DC-1988	Pragmatics
regular	http://www.isocat.org/datcat/DC-5926	Morphology and Morphosyntax
RelatedForm	http://www.isocat.org/datcat/DC-5998	LexicalEntry
reltype	http://www.isocat.org/datcat/DC-5938	SenseRelation
ROLE	http://www.isocat.org/datcat/DC-5940	synset
ROLE_AGENT	http://www.isocat.org/datcat/DC-5939	synset
ROLE_DIRECTION	http://www.isocat.org/datcat/DC-5863	synset
ROLE_INSTRUMENT	http://www.isocat.org/datcat/DC-5862	synset

ROLE_LOCATION	http://www.isocat.org/datcat/DC-5861	synset
ROLE_PATIENT	http://www.isocat.org/datcat/DC-5814	synset
ROLE_RESULT	http://www.isocat.org/datcat/DC-5835	synset
ROLE_SOURCE_DIRECTION	http://www.isocat.org/datcat/DC-5834	synset
ROLE_TARGET_DIRECTION	http://www.isocat.org/datcat/DC-5833	synset
semanticFeatureSet	http://www.isocat.org/datcat/DC-5937	Sense and Semantics
Semantics	http://www.isocat.org/datcat/DC-2645	Sense and Semantics
semanticShifts	http://www.isocat.org/datcat/DC-5997	Sense and Semantics
semanticType	http://www.isocat.org/datcat/DC-5973	Sense and Semantics
Sense	http://www.isocat.org/datcat/DC-464	Sense and Semantics
SenseExample	http://www.isocat.org/datcat/DC-5673	Examples
SenseGroup	http://www.isocat.org/datcat/DC-5996	SenseRelation
senseGroupRelationType	http://www.isocat.org/datcat/DC-5794	SenseRelation
SenseRelation	http://www.isocat.org/datcat/DC-5995	SenseRelation
sentence	http://www.isocat.org/datcat/DC-1386	Syntax
separability	http://www.isocat.org/datcat/DC-5792	Morphology and Morphosyntax
separable	http://www.isocat.org/datcat/DC-5803	Morphology and Morphosyntax
singular	http://www.isocat.org/datcat/DC-252	Morphology and Morphosyntax
slang	http://www.isocat.org/datcat/DC-1995	Pragmatics
slogan	http://www.isocat.org/datcat/DC-5909	Examples
specifyingComplement	*to be decided*	Syntax
spellingVariant	http://www.isocat.org/datcat/DC-5918	LexicalEntry
state	http://www.isocat.org/datcat/DC-5986	Sense and Semantics
STATE_OF	http://www.isocat.org/datcat/DC-5825	synset
state1	http://www.isocat.org/datcat/DC-5959	Sense and Semantics
state2	http://www.isocat.org/datcat/DC-5958	Sense and Semantics
state3	http://www.isocat.org/datcat/DC-5957	Sense and Semantics
stcognt1	http://www.isocat.org/datcat/DC-5962	Sense and Semantics
stcognt2	http://www.isocat.org/datcat/DC-5961	Sense and Semantics
stcognt3	http://www.isocat.org/datcat/DC-5960	Sense and Semantics
SubcategorizationFrame	http://www.isocat.org/datcat/DC-4620	Syntax
subordinateClause	*to be decided*	Syntax
substance	http://www.isocat.org/datcat/DC-5983	Sense and Semantics
superlative	http://www.isocat.org/datcat/DC-4925	Morphology and Morphosyntax
synset	http://www.isocat.org/datcat/DC-4613	synset
synsetRelation	http://www.isocat.org/datcat/DC-5993	synset
syntacticArgument	http://www.isocat.org/datcat/DC-6000	Syntax
syntacticBehaviour	http://www.isocat.org/datcat/DC-5999	Syntax
Syntax	http://www.isocat.org/datcat/DC-2647	Syntax
te	http://www.isocat.org/datcat/DC-5893	Syntax
tense	http://www.isocat.org/datcat/DC-1286	LexicalEntry
term	http://www.isocat.org/datcat/DC-5788	Examples
TextualForm	http://www.isocat.org/datcat/DC-6001	Example
time	http://www.isocat.org/datcat/DC-5975	Sense and Semantics
toinf	http://www.isocat.org/datcat/DC-5868	Syntax
transitive	http://www.isocat.org/datcat/DC-1405	Syntax
transitivity	http://www.isocat.org/datcat/DC-5266	Syntax
tri	http://www.isocat.org/datcat/DC-5766	Syntax
uncount	*to be decided*	Sense and Semantics
unseparable	http://www.isocat.org/datcat/DC-5804	Morphology and Morphosyntax

valency	http://www.isocat.org/datcat/DC-5763	Syntax
variantType	http://www.isocat.org/datcat/DC-5916	LexicalEntry
verb	http://www.isocat.org/datcat/DC-1424	LexicalEntry
vp	http://www.isocat.org/datcat/DC-2255	Syntax
vp	http://www.isocat.org/datcat/DC-2255	Syntax
vulgar	http://www.isocat.org/datcat/DC-1998	Pragmatics
WH	http://www.isocat.org/datcat/DC-5894	Syntax
whclause	http://www.isocat.org/datcat/DC-5875	Syntax
WordForms	http://www.isocat.org/datcat/DC-5700	LexicalEntry
wordgroup	http://www.isocat.org/datcat/DC-5888	Morphology and Morphosyntax
writtenForm	http://www.isocat.org/datcat/DC-1836	LexicalEntry
x-compound	http://www.isocat.org/datcat/DC-5889	Morphology and Morphosyntax
XPOS_FUZZYNYM	http://www.isocat.org/datcat/DC-5822	synset
XPOS_NEAR_ANTONYM	http://www.isocat.org/datcat/DC-5829	synset
XPOS_NEAR_SYNONYM	http://www.isocat.org/datcat/DC-5881	synset
zero-derivation	http://www.isocat.org/datcat/DC-5897	Morphology and Morphosyntax