Getting started with the Cornetto demo

Cornetto.inl.nl

Cornetto is a lexical resource for Dutch and includes both the data of RBN (Referentie bestand Nederlands) and DWN (Dutch WordNet). It includes a set of lexical entries (LEs) and a set of synsets related to each other. It can be seen as a combination of a dictionary and a thesaurus.

The data can be accessed in one of the following ways:

• Simple Search for Lexical Entries

Just type a word and look up the meaning of the lexical entries and their (morpho)-syntactic, semantic and combinatorial properties. If you scroll down you can see also their place in the synset hierarchy.

<u>Advanced Search for Lexical Entries</u>

Permits you to select specific lexical entries from the resource by choosing and combining its linguistic properties.

• <u>Synset visualization</u>

Look up a word and get it visualized in a wordnet structure

The descriptions of the different screens in this document are also available under the HELP button of each indivual search modus.

A wordnet groups lexical entries into sets of synonyms called synsets. It records relations between these synsets of which the most important are hypernym and hyponym relations. Hypernym relations point at broader concepts (e.g. *animal* is a hypernym of *horse*) and hyponym relations point at narrower concepts (e.g. *rose* is a hyponym of plant.)

A synset is a set of lexical entries that are synonyms of each other. Synsets are the basic units of a wordnet and they are linked to each other by all kinds of relations.

A **lexical entry** is a combination of a word and its meaning. So, the Dutch word *toren* in its meaning of a building (*tower*) is one lexical entry and *toren* in its meaning of chess piece (*rook*) in is another one.

TIP: if you want to know more about the linguistic concepts you find, then please move your cursor over the field names to find an explanation. For more details, go to the Cornetto User Documentation (pp. 8-30) (http://cornetto.inl.nl/help/D9-Cornetto-Documentation.pdf)

Simple Search for Lexical Entries

Simple Search	Advanced Search	Synset Visualisation	Help
Simple Search	Matching Lexical	Entries	
Word Form:	*gebak]	
Use Wild-cards:	⊻ (?)		
Part of Speech:	Noun \$		
Items per Page:	20 ‡		
Search			

Begin a new search by typing a word or phrase into the text field and then clicking "Search" or pressing Return. The form for Simple Search consists of an input field for text in which a word form can be entered. The user can specify the word class (Part of Speech - POS) for the results (noun, verb or adjective) and he can select the number of results that should be displayed per page. The Simple Search allows the use of wild-cards:

*	An asterisk will match any (sub)string
?	A question mark will match any single character
*vogel	This example will match all words that end with "vogel"
open*	This example will match all words that begin with "open"
?open	This example will match with words like "lopen", "kopen", "dopen", etc.

For example, the you want to search for words ending with 'gebak', by typing in '*gebak' and clicking on the checkbox 'Use Wild-cards'.

The result set is only visible if a query has been submitted. If the query results in an empty set, the user is informed about that and is advised to revise the search parameters. If the result set is not empty, a table is shown with one row per found lexical entry. Several characteristics are displayed per column. If multiple Examples are available per cell of the table, only the first value is shown with an indication of the number of remaining elements not shown.

S Cornetto Demo						
Simple Search	Advanced Search	Synset Search	Visualisation	Help		
Simple Search Matching Lexical Entries						
Matching Lexical I	Matching Lexical Entries for '*gebak'					
Download Selection	(Max. 500)				1 to 6 from 6	Previous Next
ID	Written Form	Semantics	PoS	Examples		Pragmatics Syntax
appelgebak-n-1	appelgebak	-	noun			
gebak-n-1	gebak	(zoete) gebakken spijzen	noun	hier: dagelijks vers gebak! (5 more)		gastronomy
gebak-n-2	gebak	-	noun			
schuimgebak-n-1	schuimgebak	-	noun			
vruchtengebak-n-1	vruchtengebak	-	noun			
zandgebak-n-1	zandgebak	-	noun			

Since the results are shown in pages, the maximum length of the table is the number of items previously selected in the form. There are also buttons that allow the user to page forward or backward. **Click on the left file tab to return to the form**.

Clicking on a row of the table shows the details of the lexical entry. All available information from the data is displayed in this part. The data fields have been grouped in order to make the presentation more compact and clearer. **Move your cursor to the fields to show the explanation of the terms in a tooltip**. The hierarchic information of the corresponding synset is also provided. The marked synset in the hierarchy is the current one. In this example 'clubhuis' is marked.



The nodes that are indented to the left are hypernyms and those indented to the right are hyponyms. When you click on a square box, next to the LE in the hierarchy, the related Synset with detailed synset information will display in a new file tab. When you click on the circle, you jump to the related visualization. Links to other lexical entries are also available. **Clicking on these will display the entry in a new file tab.** Click on the left file tab to return to the Form.

- TIP: if you want to know more about the linguistic concepts you find, then please move your cursor over the field names to find an explanation. For more details, go to the Cornetto User Documentation (pp. 8-30) (http://cornetto.inl.nl/help/D9-Cornetto-Documentation.pdf)
- TIP: There is also an option available to download the selected items. You can click on the button and then download the file. The data are available on the Cornetto-LMF xml format. You can find more about this format in the Cornetto User Documentation (pp. 36-38)
- TIP: The number of items that can be downloaded is limited to 500 items. If you want to have more data then please go to and download the complete resource http://wordpress.let.vupr.nl/odwn/data/

Advanced Search

TIP: If you want to know more about the linguistic data categories you can choose, please read the Cornetto User Documentation (pp. 8-30) (http://cornetto.inl.nl/help/D9-Cornetto-Documentation.pdf)

TIP: You can always start a fresh query choosing the tab 'advanced search' again and hit the button 'clear'



The Advanced Search form permits you to select specific lexical entries from the resource. All properties of the lexical entries can be used as selection criteria. Since in Advanced Search there are many options available, they are grouped in sections. The sections are distributed over three rows on the screen, and every section can be collapsed or closed by clicking on the circle.

Two methods are used for the user to enter options. If there is a closed set of values for a parameter, a select list is used which allows one selection. If the set of values for a parameter is open (any string), then a text field will be available in which the user can specify a fixed string or he can make use of wild cards. In the screenshots below a select list is shown left and a text field is shown right.

🔇 Word Forr	ns		Word Form	e.g. water, denken 🛛 ?
Comparison	– Any – Comparative Superlative	× T	Multi-word Expression	Wild-cards (* or ?) allow ?
Tense	– Any – Past Participle Past Tense	* *		
Grammatical Number	Any Plural Singular	*		

Further, the user can select the number of result that will be displayed on a page and there is an option available to clear the complete form.

For example, the user wants to search for all the words in the Dance domain. Then he has to select 'Dance' in the pragmatic section (Domain). (78 entries will be found). If he wants to select only the verbs of the Dance domain he adds a constraint on Part-of-speech in the syntax section :

Pragmation	cs		🔇 Syntax	
Chronology	– Any – Neologism Oldfashioned		Valency	Any Di Mono
Connotation	- Any Euphemistic Jocular Offensive Pejorative		Transitivity	Tri - Any Intransitive Transitive
Domain	Dance Diplomacy Ecology Economy Electronics_electricity Engineering Fashion Fishing	^	Part of Speech	Any Adjective Adverb Noun Other Verb
	Folklore Gastronomy Geography Geology Golf Herladry	~		

These selection criteria will give 14 results:

S Cornetto Demo

Simple Search	Advanced Search	Synset Search Visua	lisation	Help		
Advanced Search	h Matching Lexic	al Entries				
Matching Lexical Entries						
Download Selection	(Max. 500)			1 to 14 from 14	Previous	Next
ID	Written Form	Semantics	PoS	Examples	Pragmatics	Syntax
accentueren-v-4	accentueren	kracht, nadruk geven aan (bepaalde noten van een muziekstuk)	verb	accentueren van melodieën	music, dance	
afstemmen-v-4	afstemmen	een muziekinstrument zuiver stemmen	verb	snaren op elkaar afstemmen	music, dance	
buikdansen-v-1	buikdansen	een buikdans doen	verb	elke avond is er buikdansen in het Turkse restaurant op de hoek	dance	
hakken-v-7	hakken	gabberstijl dansen	verb	op Nike Air Max schoenen hakken	dance	
inspelen-v-4	inspelen	een muziekpartij spelen voor opname	verb	een gitaarsolo inspelen (1 more)	music, dance	
schuifelen-v-2	schuifelen	slijpen	verb	schuifelen op langzame muziek	dance	
stijldansen-v-1	stijldansen	stijldansen uitvoeren	verb		dance	
swingen-v-1	swingen	dansen op popmuziek	verb	een avondje lekker swingen	dance	
tapdansen-v-1	tapdansen	tappen	verb		dance	
trillen-v-3	trillen	met vibrato ten gehore brengen	verb	een snaar laten/doen trillen (1 more)	music, dance	
twisten-v-2	twisten	de twist dansen	verb		dance	
variëren-v-3	variëren	(een thema) bewerken	verb	de regels in een liedje variëren	music, dance	
volksdansen-v-1	volksdansen	volksdans beoefenen	verb		dance	
walsen-v-3	walsen	een wals dansen	verb	een hele avond romantisch walsen	dance	

Clicking on a row of the table shows the details of the lexical entry. Click on the left file tab to return to the form. There is also an option available to download the selected items as LMF-XML.

Synset Visualization

- A wordnet groups lexical entries into sets of synonyms called synsets. It records relations between these synsets of which the most important are hypernym and hyponym relations. Hypernym relations point at broader concepts (e.g. *animal* is a hypernym of *horse*) and hyponym relations point at a narrower concept (e.g. *rose* is a hyponym of plant.)
 - A synset is a set of lexical entries that are synonyms of each other. Synsets are the basic units of a wordnet and they are linked to each other by all kinds of relations.
 - A lexical entry is a combination of a word and its meaning. So, the Dutch word *toren* in its meaning of building (*tower*) one lexical entry and *toren* in its meaning of chess piece (*rook*) in is another one.

With the synset visualization tool you can look up words in the Dutch wordnet. It is very easy to use the visualization tool: type in the word and press enter to see the word web. The grey squares indicate the Synset-IDs and the colored •••• dots indicate the Lexical Entries (LEs). The colors correspond with the part-of-speech. For example, if you search for 'arm', you'll see blue words that indicate the nouns and yellow words that indicate the adjectives (see screenshot below). The **striped lines** between the nodes indicate the hyponym relations and the **orange lines** indicate the hyponym relations. **Yellow lines** indicate other relations, for example near synonyms. It is possible to **drag the lines** to a different location on your screen by grabbing the nodes. It is also possible **to click on a grey node** to see the visualization for that particular synset. **Move your cursor** on the nodes to show more information about the LEs or the synsets. If you want to **zoom in** for a close-up on parts of the results then use the scroll wheel ; and if you want to **pan or move** the figure across the screen then drag the background.



When you look up a word that has different meanings, you will only see a representation of of its first meaning. For the visualization of the other meanings you have to select the part-of-speech and/or the desired Sense number in the form. You can find the sense numbers via the Simple Search or just try one.

Word Form:	knop
Part of Speech:	Noun ¢
Sense Number:	3 ‡
Search	

If you have selected a synset with a lot of relations (you could try *plant*), you may get confused by the amount of information you get. The other synset representation available through the Simple Search modus may be more useful in that case.

