

**Studying words with the WordNet lexical reference**

# LEXICAL CONNECTIONS

The WordNet lexical refer-

ence maps connec-  
tions between words.

Check out this fascinat-  
ing tool based on lan-  
guage data from two  
decades of research. **BY DMITRI POPOV**

**W**ordNet is a reference tool that lets you study the connections between words. In the words of its developers, WordNet is "... an online lexical reference system whose design is inspired by current psycholinguistic theories of human lexical memory. English nouns, verbs, adjectives and adverbs are organized into synonym sets, each representing one underlying lexical concept. Different relations link the synonym sets." (<http://wordnet.princeton.edu/>). In other words, since all the en-

tries in WordNet are organized into synonym sets (synsets) and they contain definitions and examples, WordNet can be used both as a thesaurus and a conventional dictionary.

However, what makes WordNet a unique reference tool is that every synset is connected to other synsets via a number of relations. This means that for each word in WordNet, you can retrieve not only its synonyms, but also hypernyms, hyponyms, meronyms, and holonyms.

A *hypernym* describes the *x is a kind of y* relationship between words. For example, in the relationship *an oak is a kind of tree*, *tree* is a hypernym, or, in other words, *tree* is a superordinate of *oak*. *Hyponym* also describes the *x is a kind of y* relationship, but in reverse. In the previous example, *oak* is a hyponym, or a subordinate, of *tree*. *Meronym* denotes a constituent part of, or a member of something. For example, *engine* is a meronym of *airplane*. *Holonym* is a meronym in reverse. In the example above, *air-*

*plane* is a holonym of *engine*. There are a few other terms that are used in WordNet, but these four are enough to give you an indication that WordNet is more than an ordinary digital dictionary.

## Installing WordNet

Most distributions provide a packaged version of WordNet, and you can install

### WordNet Commands

`wn word -over` provides an overview similar to a dictionary word article. The overview includes a number of senses, definitions, synonyms, and example sentences.

`wn word -syns {n / v / a / r}` returns a list of synonyms for the specified word, where n=noun, v=verb, a=adjective, r=adverbs. For example, if you want to see synonyms for the noun *monkey*, use `wn monkey -synsv`, which returns:

Sense 1

tamper, fiddle, monkey  
=> manipulate

Sense 2

putter, mess around, potter, ↗  
tinker, monkey, monkey around, ↗  
muck about, muck around  
=> work

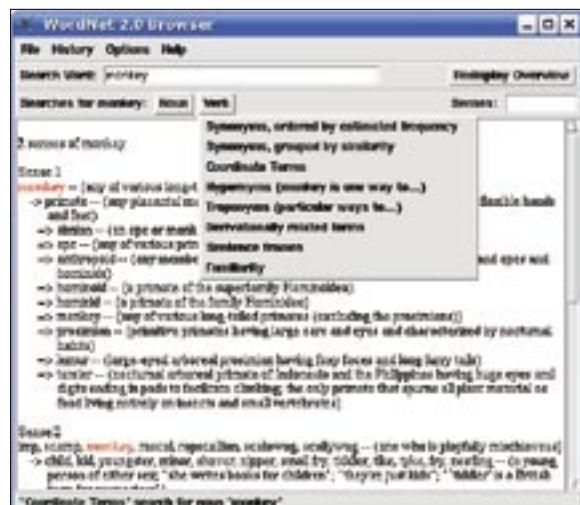


Figure 1: The WordNet browser provides a simple graphical interface to the WordNet reference system.

it using your package management tool. If your Linux distro doesn't include the WordNet package, you can download a tarball from WordNet's official website (<http://wordnet.princeton.edu/obtain>) and install it using the standard installation routine:

```
./configure
make
make install
```

The WordNet browser requires the Tcl/Tk packages, which must be installed before you build and install WordNet. Finally, if you use a Live CD Linux distro that supports klik, you can install WordNet from <http://wordnet.klik.atekon.de/>.

## WordNet's Basic Commands

Once you have completed the installation, you are ready to explore WordNet. The man pages provide an exhaustive overview of WordNet's commands. See the box labeled "WordNet Commands" for a summary of command examples.

**-hype {n | v}** and **-hypo {n | v}** display hypernyms and hyponyms respectively. For example, *wn monkey -hypo* returns the output shown in Listing 1.

**-tree {n | v}** performs a recursive search that finds the hyponyms of each hyponym. For example, *wn monkey -tree* returns the output show in Listing 2.

In this particular case, this command displays a

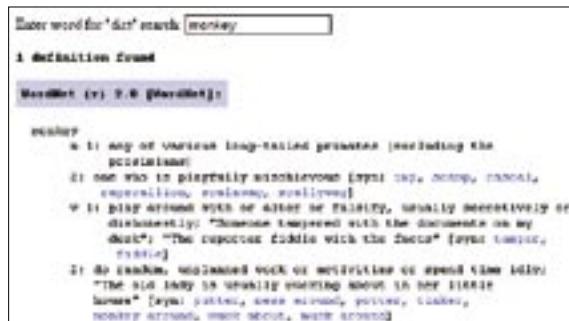


Figure 2: dictd and a PHP script allow you to run a simple WordNet server on a LAN.

## Listing 1: Monkey hypernyms

```
01 Sense 1
02 monkey
03      => primate
04          => placental, placental mammal, eutherian, eutherian
               mammal
05          => mammal
06              => vertebrate, craniate
07                  => chordate
08                      => animal, animate being, beast, brute,
                           creature, fauna
09                          => organism, being
10                              => living thing, animate thing
11                                  => object, physical object
12                                      => entity
```

list of monkey species. If you want of different types of airplanes, run the *wn airplane -tree* command.

## Graphical WordNet Tools

If using the command line version of WordNet is not your cup of tea, you can opt for a graphical tool. WordNet includes its own graphical browser. Although the browser has a rather simplistic interface, it does allow you to access basic WordNet features. Looking up a word in the WordNet browser is a two-step process. First, enter a word into the *Search Word* field and press Enter. The application then returns an overview of the search term similar to the *-over* parameter. The *Seaches for* bar displays buttons for each syntactic category the found word belongs to, and you can use them to view more detailed information such as synonyms, coordinate terms, domains (for adjectives), etc.

Another graphical application based on WordNet, and a rather interesting one, is wnconnect (<http://dingo.sbs.arizona.edu/~sandiway/wnconnect/>).

Enter two words, and wnconnect finds the shortest path or all possible connections between them. For example, enter the words *apple* and *monkey*, and wnconnect finds a connection between them and presents the final result as a

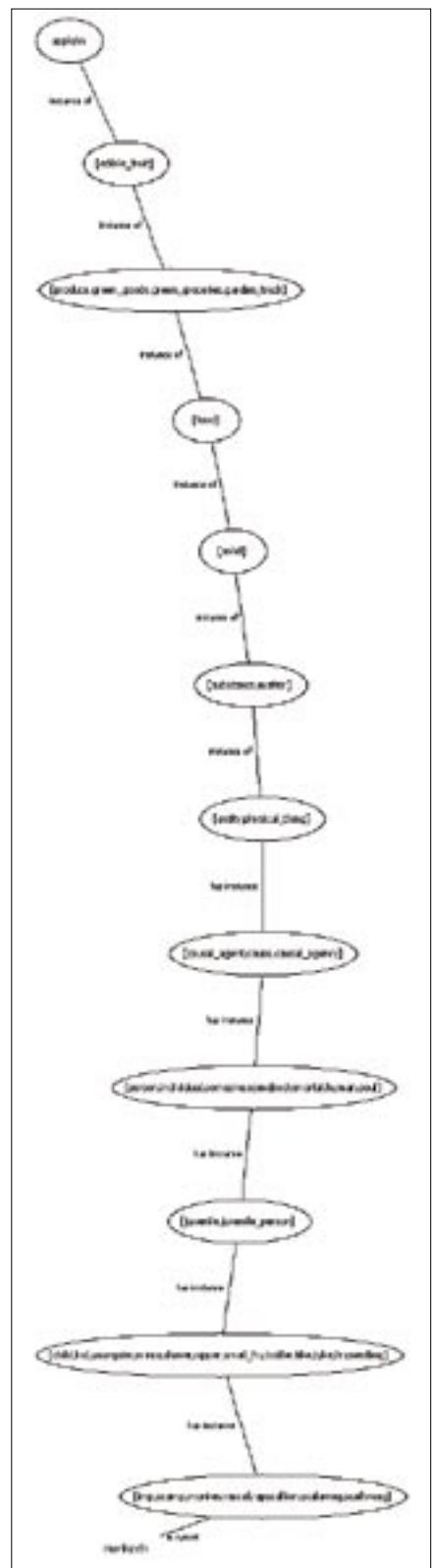
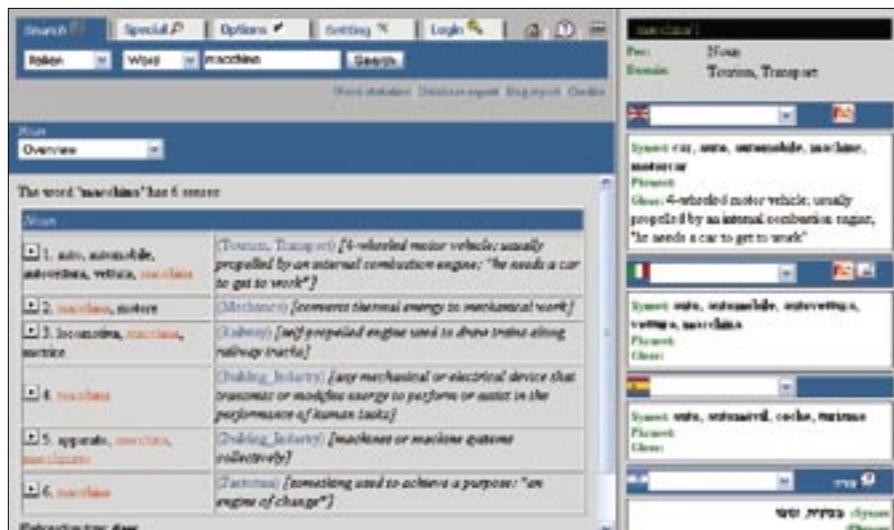


Figure 3: wnconnect charts the short path of all connections between words.



**Figure 4:** MultiWordNet Online provides a multi-lingual version of WordNet.

graphical chart in PNG or PDF formats. Actually, finding connections between words can be quite addictive, and you can even turn it into a game. Just pick two random words and try to map a connection between them, then use `wnconnect` to see whether you've got it right.

## WordNet on the Network

If you have WordNet installed on your Linux server, you can access the application via Telnet or SSH. However, you can also install a full-blown local network dictionary server accessible via a web interface.

The easiest way to provide local network users with access to WordNet is to install a dictd server and a pre-formatted WordNet database on your local server. Both components are available at <http://www.dict.org>. Installing dictd is a rather standard process. Make sure that the flex, bison, and byacc packages are installed, then do:

### **Listing 2: Recursive Hyponyms for Monkey**

```

01 monkey
02      => Old World monkey, catarrhine
03          => guenon, guenon monkey
04          => talapoin, Cercopithecus talapoin
05          => grivet, Cercopithecus aethiops
06          => vervet, vervet monkey, Cercopithecus aethiops
          pygerythrus
07          => green monkey, African green monkey, Cercopithecus
          aethiops sabaeus
08          => mangabey
09          => patas, hussar monkey, Erythrocebus patas

```

To make sure that everything works properly, switch to the `/usr/local/sbin` directory and execute the `dictd` command as root. Then use the `dict` client to look up a word:

```
dict monkey
```

If everything works as it is supposed to, you can add a web interface to the dictionary server. Start with installing the Apache web server and the `apache_mod_php` module. Create a new text file, copy the PHP script from <http://www.arachnid.com/linux/dict.php.html> and paste it into the file. In some cases (for example, on PCLinuxOS), you may need to enter the correct path to the `dictd` in the following line:

```
exec("/usr/bin/dict $query >
2>&1",$output,$error);
```

Save the file as `wn.php` in the `/var/www/html` directory. Now launch your browser and point it to the created page to check whether everything works properly.

## WordNet on the Web

Princeton University maintains a bare-bones online version of Wordnet. However, for the ultimate web-based version of WordNet, look no further than MultiWordNet On-line (<http://multiwordnet.itc.it/online/>). This is an implementation of WordNet for five different languages: English, Italian, Spanish, Hebrew, and Romanian. More impressive, however, is that you can view all these languages side by side, which makes MultiWordNet a quite unique language reference tool.

## Conclusion

Born as an academic project, WordNet has become one of the most exciting and useful language reference tools available for the average user. WordNet's major advantage is its versatility: you can use it as a thesaurus and dictionary, but it also provides a fascinating insight into the world of language. This article gives you just a glimpse of WordNet's possibilities, and if you want to know more, WordNet's website (<http://wordnet.princeton.edu/>) and the WordNet book (<http://www.amazon.com/gp/product/026206197X/>) is a good starting point. ■