

## The IMS workbench - a introductory case study by Roger Levy

Local tutorial: <http://www.stanford.edu/dept/linguistics/corpora/cas-tut-cqp.html>

Other sites: <http://bulba.sdsu.edu/cqphelp.html>

Question: Is there any semantic difference between "into+VPing" and "to+VP" as verbal complements, or is the difference between "get X to/\*into VP" and, say, "sucker X \*to/into VPing" strictly syntactic subcategorization?

Using IMS to solve this problem:

1. Go to your home directory under UNIX (e.g. on cardinal.stanford.edu)
2. Add the following to your *.cshrc*:

```
setenv PATH /afs/ir.stanford.edu/data/linguistic-data/mnt/mnt23/IMS/bin:"$PATH"  
setenv CORPUS_REGISTRY /afs/ir.stanford.edu/data/linguistic-data/mnt/mnt23/IMS/registry  
setenv CQP_LOCAL_CORPUS_DIR /tmp
```

3. Log in anew, or enter

```
source ~/.cshrc
```

so that the changes made in the *.cshrc* take effect.

4. Run *cqp*:

```
cqp
```

5. When you see the prompt

```
[no corpus]>
```

enter a corpus, e.g.

```
[no corpus]> BROWN;
```

or if you don't know which corpora are available just enter

```
[no corpus]> show;
```

which will give you a list of the form

*System corpora:*

```
BNC  
BROWN~  
BROWN  
DICKENS
```

6. Now you can enter searches will be done on the selected corpus, e.g.

```
BROWN> "into" [tag="V.G"];
```

```
14187: 's wife for talking him <into letting> me play ". Bob_Burnes ,
```

```
24818: the song just whispering <into being> . It is greatly to Berman
```

120819: *been misled by this fact <into minimizing> the degree of change whic*  
 205236: *kes who put their hearts <into making> the tragic scene . I 've*  
 BROWN> BNC;  
 BNC> "into" [pos="V.G"];  
 ...  
 99581310: *tive . Do n't be deluded <into thinking> that they are locked beca*  
 99590192: *or a breeze can fool you <into thinking> the sun is weaker than it*  
 99599287: *, a lot of work has gone <into improving> the quality of care that*  
 ...

7. You may also save the subcorpus which is defined through your search pattern:

BNC> Into=Last;  
 BNC> save Into;  
 BNC> Into;  
 BNC:Into[2228]> show;  
 System corpora:  
   BNC  
   BROWN~  
   BROWN  
   DICKENS  
 Local Corpora:  
 BNC:Into (2228 intervals): *loaded, saved, no update needed*

8. Finally, CQP allows you do look for collocations, for example you could sort the collocates of "into" which are within 4 words left of the "into" and are some type of verb.

BNC:Into[2228]> set Into collocate rightmost [pos="V.\*"] within left 4 words from match exclusive;  
 BNC> set Last collocate rightmost [pos="V.\*"] within left 4 words from match exclusive;  
 BNC> group Last match word;

The output will be a list of words (in this case verbs) with the number of collocations following it, e.g.:

...  
                   pushed 6  
 ...