Hauptseminar *Using NLP to Foster Language Awareness in SLA* ISCL, Universität Tübingen
Detmar Meurers Sommersemester 2008

## Exercise sheet 4

(Submit by email to dm@sfs.uni-tuebingen.de before class on Thursday, June 26)

Write the little programs below, naming them nr1.py, nr2.py, etc. Save them in a folder ex4-*userid* (e.g., ex4-dm) and send me an email with a tar.bz2 archive of that directory. Please make sure that it unpacks to a directory ex4-*userid*, not the individual files.

1. Loosely based on the idea that creating language awareness is one of the relevant issues in iCALL, write a web form on which a user can type a text and can then request to see the text displayed in one of the following ways:

   (a) with the nouns marked in blue

   (b) with the adjectives marked in green

   (c) with the finite verbs marked in orange

   (d) with the non-finite verbs marked in red

   (e) with all the above markings together

   Some notes:

   - On aticall the TreeTagger is installed (under /usr/local/lib/treetagger) including various language models) and you can conveniently call it from python using the treetaggerwrapper module, which is installed on aticall as well:

```
import treetaggerwrapper

#1) build a TreeTagger wrapper:
tagger = treetaggerwrapper.TreeTagger(TAGLANG='en',TAGDIR='~/TreeTagger')

#2) tag your text.
tags = tagger.TagText("This is a very short text to tag.")

#3) use the tags list... (list of string output from TreeTagger).
print tags

# this prints the following tagged text representation:

['This\tDT\tthis',
 'is\tVBZ\tbe',
 'a\tDT\ta',
 'very\tRB\tvery',
 'short\tJJ\tshort',
 'text\tNN\ttext',
 'to\tTO\tto',
 'tag\tVB\ttag',
 '.\tSENT\t.']
```

   - Use the color attribute of font tags to colorize the html output (style sheets would be a more general solution, but that's not the point of this exercise).