

# SQ: Basic computing skills for linguists

April 2010

<b>Lecturer</b>	Luis Vicente
<b>Email</b>	vicente@uni-potsdam.de
<b>Office hours</b>	Tuesdays 10:00–12:00, but spontaneous walk-ins are also welcome
<b>Course website</b>	<a href="https://moodle.uni-potsdam.de/course/view.php?id=3448">https://moodle.uni-potsdam.de/course/view.php?id=3448</a>
<b>Place</b>	II.24.1.79 (Computerpool)
<b>Time</b>	Tuesdays, 10:00–12:00

## 1 General information

This course will introduce students to the  $\text{\LaTeX}$  typesetting system, UNIX shell scripting, and basic natural language applications of Python. These tools are becoming more and more common in linguistic work (as well as in the general academic and professional sphere), and therefore familiarity with their usage is an asset to students.

## 2 Topics

1.  **$\text{\LaTeX}$** : document layout; bibliography management; installation of new packages; utilities of interest to linguists.
2. **Command line**: commands for system management and navigation; command piping and output redirection; regular expressions.
3. **Python**: basic programming methods; corpus methods with NLTK.

## 3 Required software

In order to practice the techniques we cover in this course, students must have access to a computer running the following software:

- A UNIX command line.
- Python version 2.6 (or higher) with the NLTK library installed.
- A complete  $\text{\LaTeX}$  system.

The machines in the Computerpool are already equipped with this software, but you are advised to install and use it in your home computer too. Ask me if you need help.

## 4 Reading

Since this course is heavily focused on getting hands-on experience with these tools, there will be little reading. Nonetheless, students will be required to consult standard manuals in order to complete some of the homework. The following are good sources of information.

**General** In order to get into the right mindset for this course, students should read Neal Stephenson's essay *In the beginning... was the command line*, which is freely available from <http://www.cryptomicon.com/beginning.html>. The German translation is entitled *Die Diktatur des schönen Scheins*, but it only exists in book format (no free download, but copies sell for cheap on [amazon.de](http://amazon.de)). Unfortunately, the University library doesn't own a copy, and neither do the Potsdam or Berlin Stadtsbibliotheken. Students who find this essay engaging are encouraged to also read Stephenson's novel *Cryptomicon*.

### LaTeX

- Oetiker, Partl, Hyna, and Schlegl. 2006. *The not-so-short introduction to LaTeX 2 $\epsilon$* . Available online.
- Mittelbach and Goosens. 2004. *The LaTeX Companion, 2<sup>nd</sup> edition*. Incredibly comprehensive. Available at the library.

### Shell scripting

- Garrells. 2009. *Bash guide for beginners*. Available online.
- Cooper. 2009. *Advanced bash-scripting guide: an in-depth exploration of the art of shell scripting*. Available online.

### Python and NLTK

- The online Python Tutorial: <http://docs.python.org/2/tutorial/index.html>
- Bird, Klein, and Loper. 2009. *Natural Language Processing with Python*. Available at <http://www.nltk.org/book>.