Big and not so big data in linguistics
We, the corpus linguists of the Zurich Center for Linguistics *(ZüKL) and the* URPP Language and Space, *are experts in collecting, processing, storing and querying data for linguistic research. No matter if you are a linguist at the UZH and want to work with data in your BA paper or apply for a 5 Mio. grant or anything in between, we can support you.*

ZüKL *and* the URPP CorpusLab *work in collaboration, providing complementary expertise. The main focus of the* CorpusLab *support is on data-driven methodology and general data management.* ZüKL *is in charge of providing hardware and software for web-related data management. Both provide help in running specialized language processing tools. They are reachable through a joint contact point.*

How to get support
Simply send a request to the head of the CorpusLab, Tanja Samardzic*. She will help you further specify your needs and actions to be taken. She will also coordinate the work with* Alexandra Bünzli *(ZüKL) and any other potential collaborators who might get involved in working with your data.*

Here are some of the tasks we might be able to help you with:

Collecting data
Depending on the nature of your data and of your research question, tools that allow for much more efficient ways of working are available. For instance, transcribing a recording of spoken language is a frequently encountered task for which Microsoft Word is frequently used. While Microsoft Word seems straightforward and easy to use, a proper transcription tool will do much better in aligning your transcription with the sound source, encoding pauses and any special mark-up in a standard way. With our help, you will get to use the more specialized tools with little or no effort. Setting up your data collection in a way that is in line with your research question will save you a lot of time. In this respect, the URPP is your first contact and can advise you. If server-based tools are needed, the *ZüKL* provides the infrastructure in the form of server capacity and maintenance of databases on the server.

Processing data
Do you need to annotate your data with grammatical or semantic features? There are tools to perform this task efficiently, and–depending on your needs–there are tools to do this (semi-)automatically or facilitate the manual (post-)processing greatly. Some of them are installed on the *ZüKL* server and we can thus provide easy access to them. The *ZüKL* server administrator is also open for suggestions or can offer support for the installation of tools on your local machine.

Data formatting, querying and evaluation
You have downloaded linguistic data from some corpus or created your own and now want to process them further. How do you get the data into the proper form? The *URPP* is knowledgeable in this field and can help you convert your data into a format that can be evaluated as a stand-alone database or integrated in one of the querying tools available on the *ZüKL* server. If you need support querying your data, the *ZüKL* can assist you with tools installed on their server or point you to other service providers at the UZH experienced in field, such as statistics. We are all also highly experienced in querying languages, RegEx, statistics, normalization etc.

Data storage
Your thesis has been accepted, but your publisher wants the readers to have access to the data? You created a corpus that you want to make available to the public in a browser or as raw data? The *ZüKL* has the infrastructure for you. If you followed our advice when creating your corpus, we can even make your data searchable via existing browsing tools such as *CQP*, *ANNIS* or *NoSketch*.

Application for grants
We can help you create a budget for a grant application and to formulate a methodology that has a chance to be accepted by reviewers. Once accepted, we can support you in formulating the requirements you need to

find experts that can perform the tasks described in the application. Please do not forget to budget the manpower and resources you need to achieve the various technical tasks in your project.

Project organization
If you need a homepage for your project or a document storage facility, you might be better off with an *UZH CMS* page and with *OLAT* if all your staff is registered at the UZH. However, if your project covers more than one university, you might need a tailor-made solution. The *ZüKL* can support you in formulating your needs to find a suitable programmer and we can provide storage. Please keep in mind that the URL of such an Internet page must end in .uzh.ch.

What we cannot provide
Our *ZüKL* server is intended to host data that have to be processed by an installed application or that have to be made available online. If you need to work on a Word document together with your department's team, or temporarily store a document that is too big to be attached to an e-mail, please contact IT support at your department. They have better solutions for these kinds of tasks. Furthermore, the central IT-services of the UZH offer some very useful services:

- Send or receive big files which can not be attached to an e-mail
- You want to store your files in the cloud and share it with others? Do not use the publically available Dropbox but SwitchDrive! You already have access with your UZH-AAI-Login.