

# WAKING A SLEEPING BEAUTY

INCREASING THE USABILITY OF THE DOBeS-ARCHIVE FOR VARIOUS AUDIENCES

*Julia C. Miller & Gabriele Müller*

BEAVER-TEAM

The rapidly growing DoBeS-archive is steadily accumulating linguistic and cultural knowledge on endangered languages and their speakers. However, access to the corpus, navigation through the archive and data retrieval are as yet skills limited to the collectors of the data and the creators of the archive. Our talk addresses the necessity of creating entry points and browsing facilities for the archive that will suite the needs and interests of different user audiences and might even attract curious visitors – thus raising awareness for endangerment of languages.

Presently, there are two main ways for finding data in the corpus: Through metadata search, respectively clicking through the tree structures or through annotation search which is limited to the transcribed parts of the data. Both these options are difficult to employ for people not yet acquainted with the data and the corpus and even for linguist familiar with the archive finding data on e.g. lexical tone might prove to be time-consuming. Furthermore, language documentation is a very large concept and the resulting archives are as different as the projects that have been working on them, making it hard to search the data with different degrees and focusses of annotation. By enhancing the DoBeS-website in certain ways, the archive could become more attractive for a broader community of interested users and easier to handle for researchers with specific interests.

For non-specialist users, it would be very helpful to create entertaining ways to get acquainted with the corpus, like a tour of community materials and projects. An entry path both for specialists and laymen could be a topical browser for the whole corpus (e.g. counting, motherese, words for "yes" and "no") and a story browser could similarly be useful for ethnologists while still entertaining non-specialists with browse-themes: e.g. creation, death, spirits, animal stories, biographies.

A third browse option – browsing by elicitation set-up – would make the retrieval of data for scientific purposes more structured. Thus starting from certain stimulus material (e.g. MPI-stimulus-kits, reciprocals videos, maptasks) browsing the corpus would offer comparable data on certain topics. This will be illustrated by maptask-data, that can be easily recorded and could be useful to researchers on intonation.

Furthermore, specialist users could be guided through the archive to their points of interest by profiles for each project, indicating research priorities, data distribution for topics or content types, annotation practices, etc. A sample profile will be presented. For members of new DoBeS-teams, access to older archives should be provided for ideas and orientation if needed.

These and other ideas will be put to discussion, hoping to increase the visits in the DoBeS-archive from a variety of users.