

What is the Language Bank of Finland?

The Language Bank of Finland is a collection of services built especially for researchers. The Language Bank receives, archives and redistributes electronic research materials for further use. Our services are free of charge for researchers and students.

The Language Bank of Finland is developed and maintained by the FIN-CLARIN consortium that consists of a number of Finnish universities, research organizations and CSC – IT Center for Science. FIN-CLARIN is part of the international research infrastructure CLARIN ERIC.

Why are research materials and data archived?

The collection and analysis of research materials takes time. Archived materials enable researchers to repeat a study or to search the material efficiently. This makes the research results more reliable. The same research material can be reused in new studies, for education and in theses and dissertations.

Once a language materials are archived as corpora in the Language Bank of Finland, they will remain usable and available for research. They also help preserve language and culture for future generations. In time, all research materials will become sources of historical research.

In case you are participating in a study whose material is to be archived in the Language Bank, any information you have provided will be permanently stored on the servers used by FIN-CLARIN. If you do not wish your data to be permanently archived, please ask the researcher for the option to participate without archiving.

For what purposes does the Language Bank allow access to corpora?

The Language Bank provides both public and restricted corpora. The public corpora are available for any purpose. The restricted corpora are distributed either for research and education purposes, or for individual use only.

Please check the research brochure or the consent form for the purposes for which your data will be distributed from the Language Bank of Finland. Ask the researcher for more information if necessary.

Who will have access to the corpora in the Language Bank? How do you ensure that the corpora are used for proper purposes?

The public corpora are openly available for everyone and no registration is required. The restricted corpora are made available to registered customers only.



The customers must commit to the general terms of use as well as to the specific conditions regarding the use of each corpus. According to these, the customer must not compromise the privacy of the subjects or other individuals when using the material or in research publications.

In case the material is only available for individual use, the registered customer is required to specify the intended purpose. He or she will be required to provide, e.g., a description of the research topic, or the name of the course he or she is attending.

In case of misconduct or improper use, the Language Bank of Finland has the right to close the customer's user account and to report the incident to the creators of the corpus in question, to the customer's employer or university and funding organ, and also to law enforcement, if necessary.

What kind of identifying information is included in the research material that is archived?

The archived corpora will only include those identifying details that are necessary in order to understand the contents of the material. Unless a more thorough anonymization process has been agreed on, the Language Bank will always remove the contact information and personal identity numbers of the research subjects. In written parts of the material, the researcher will most often remove person names. When necessary, other identifying information may be removed as well.

Please note that in the case of audio and video material, it is usually not possible to remove the identifying information completely, since this would make the recordings unusable for research.

Please check the research brochure or the consent form for the identifiers that will be removed from the archived corpora. Ask the researcher for more information if necessary.

More information about the Language Bank

In case you wish to know more about the processes and services of the Language Bank of Finland, please contact FIN-CLARIN. We will be happy to answer your questions. You can find the contact information at the bottom of the page. If you are interested, you can even explore the public corpora that are available in the Language Bank by visiting the address <https://www.kielipankki.fi/language-bank/>.

Privacy practices of the Language Bank of Finland:

<https://www.kielipankki.fi/language-bank/privacy-practices/>

