

Involving users and collaborating between disciplines in making cultural heritage accessible for research

Johanna Berg, Rickard Domeij, Jens Edlund, Gunnar Eriksson, David House, Zofia Malisz, Susanne Nylund Skog & Jenny Öqvist

Tilltal@isof.se

CLARIN Annual Conference 2017

Budapest, 18–20 sept



SWE-CLARIN

Contents

1. The Tilltal Project (Eng: SpeakingUp)
2. Involving and collaborating with users
3. Theory and methods for studying usage
4. Goals, challenges and implications for design

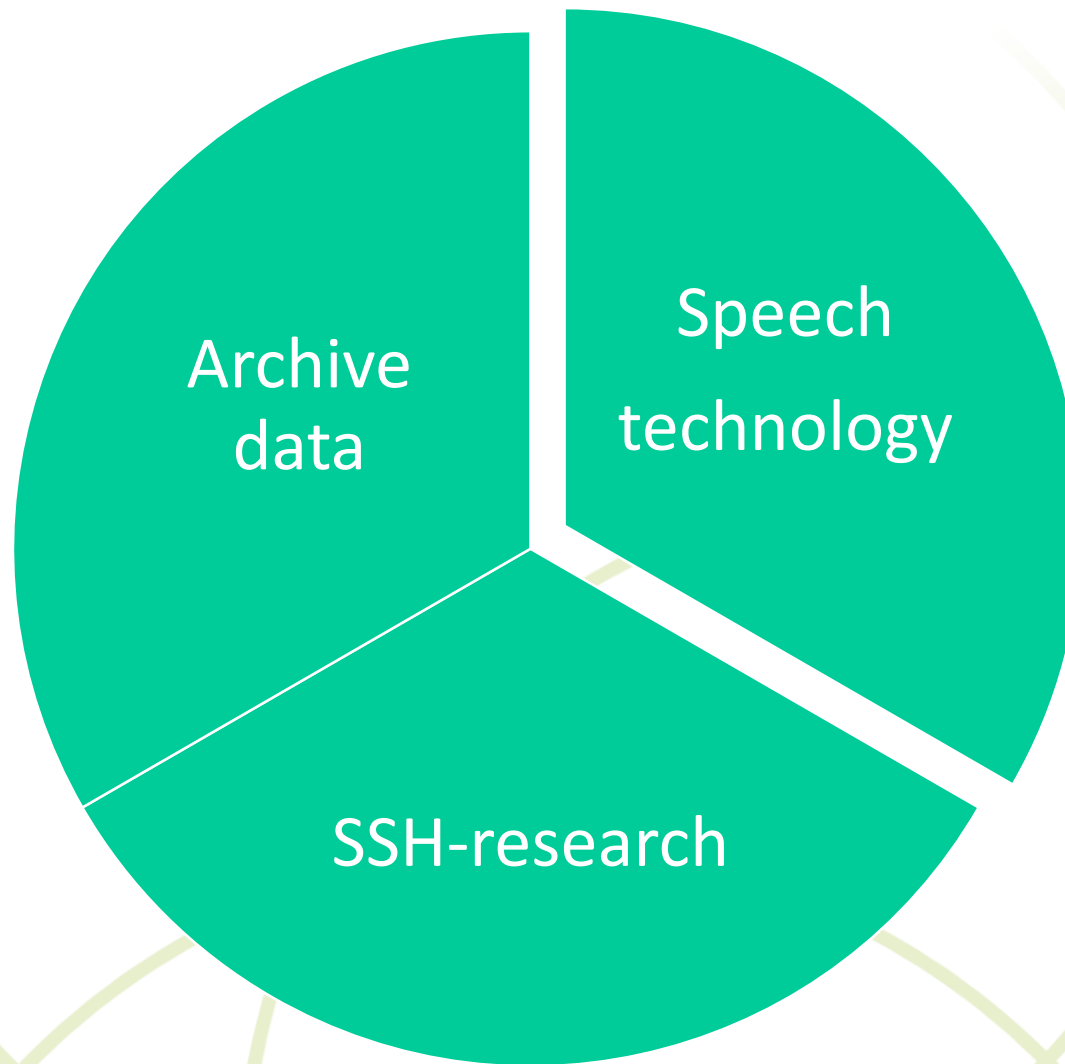


The Tiltal project

- examines how speech technology methods can make speech recordings at public memory institutions more accessible to researchers
- Funded by the Swedish Foundation for Humanities and Social Sciences 2017–2020. Nearly 1 milj EUR.
- Collaboration between three SWE-CLARIN partners: The Swedish National Archives, KTH Royal Institute of Technology, Institute for Language and Folklore (ISOF)
- ISOF = Language council + Dialect and folklore archives in Uppsala and Gothenburg with large amounts of data in text and speech (recorded interviews).



The idea of Tiltal



Involving users from different disciplines

- SWE-CLARIN workshop in 2015 to involve researchers and data holders in project cooperations centered around researchers' needs.
- Suggesting ways of using speech technology on large speech data sets for facilitating the research process.
- Three studies conceived as use cases in the Tilltal project:
 1. Use text and speech technology as tools to explore a large folklore collection containing both text and speech data.
 2. Use speech analysis for finding linguistic variation in recordings.
 3. Apply methods from interaction analysis to large quantities of speech recordings (found data) to study interaction patterns.



Theory and methods for studying usage

- What we call **the usage study** of the Tilltal project is an overarching investigation and design process that involves all three use cases.
- Based on **activity theory** and **use case modelling**

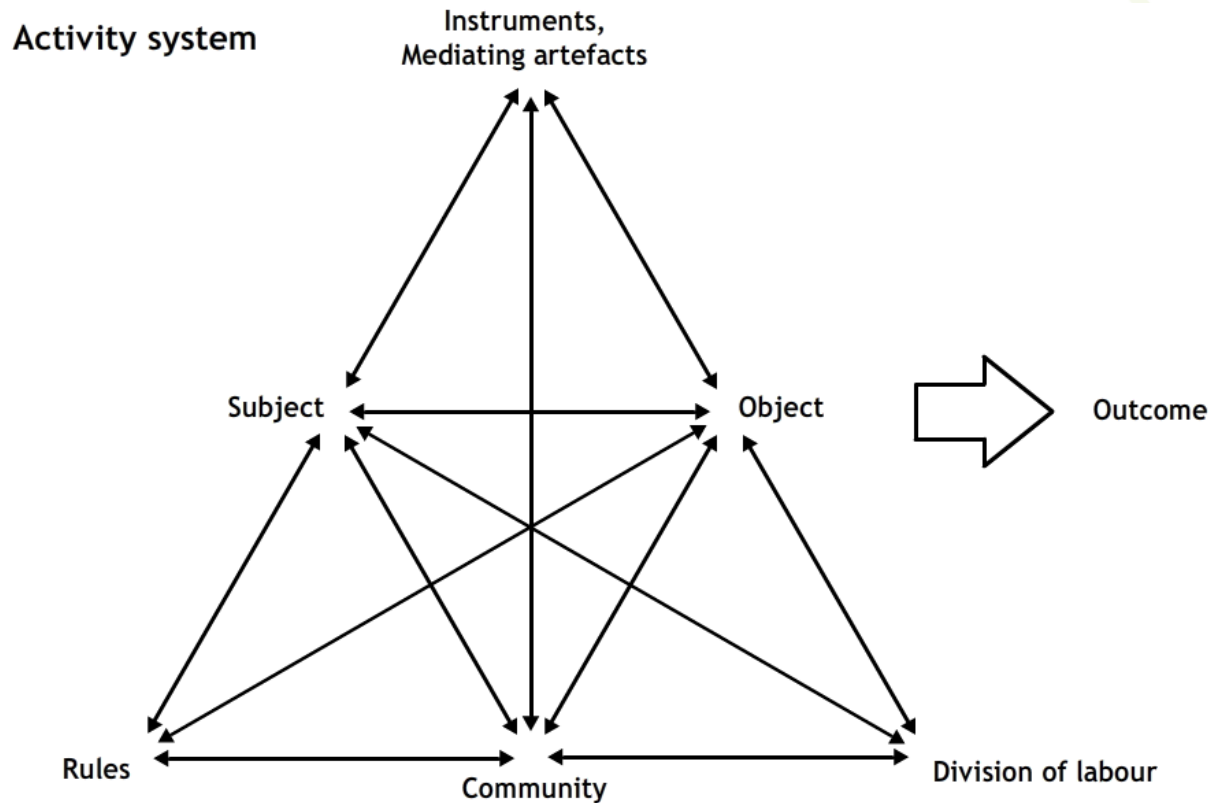


Activity theory (e.g. Kaptelinin, 2013)

- a framework for studying the activities of collecting, using and making speech data available for research at ISOF
- Investigating key factors and their relations:
 - tools and other artefacts used for the purpose
 - the community in which the activities take place
 - the division of labor between actors
 - the rules and conventions that govern the activities (as intellectual property rights, privacy laws and research ethics)
 - contradictions within and transformation of the system
- Field diaries, observations, interviews, group discussions and workshops are used for data collection.



Engeströms activity system model (1987)



 © Matt Bury 2012



Use case analysis

- is applied on collected data to further specify and model researchers' needs of information and usable tools for seeking information in speech material
- will help us formulate detailed requirements for evaluating existing tools and suggesting future solutions.
- The most prominent information needs for each use case will be identified and modelled with the help of use case theory and practices taken from Jacobson et al. (1992, 2011).



Goals, challenges and implications for design

- Provide guidance and constraints to the future development of usable tools for research based on speech data.
- By combining use case modelling with activity theory, we hope to do this without losing sight of the larger context within which digital tools and artefacts are used.
- By actively involving researchers and archivists in participatory design (e.g. Kensing & Blomberg 1998), we ensure that results will meet the needs and practices of SHR researchers and will be as usable as possible for them.
- Folklorists request systems that can handle different types of related data resources as one connected collection.

