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

Johanna Berg
The Swedish National Archives, Sweden
Johanna.Berg@riksarkivet.se

Rickard Domeij
The Language Council Institute for language and folklore, Sweden
rickard.domeij@isof.se

Jens Edlund
Speech, Music & Hearing KTH Royal Institute of Technology, Sweden
edlund@speech.kth.se

Gunnar Eriksson
The Language Council Institute for language and folklore, Sweden
gunnar.eriksson@isof.se

David House
Speech, Music & Hearing KTH Royal Institute of Technology, Sweden
davidh@speech.kth.se

Zofia Malisz
Speech, Music & Hearing KTH Royal Institute of Technology, Sweden
malisz@kth.se

Susanne Nylund Skog
The Folklore Archives The Institute for Language and Folklore, Sweden
susanne.nylund.skog@isof.se

Jenny Öqvist
The Dialect Archives Institute for language and folklore, Sweden
jenny.oqvist@isof.se

Abstract

This paper presents the project design and initial experiences of involving users and collaborating between disciplines within the newly started project Tilltal (Berg et al. 2016). The long-term goal of the Tilltal project is to make speech data at the Swedish memory archives more accessible to SSH researchers. We achieve this not only by describing methods by which speech technology can be used to reach SSH research goals, but also by providing fruitful examples of involving users, studying usage and collaborating between disciplines using the approach of participatory design (e.g. Kensing & Blomberg 1998). We use activity theory to survey the research activities surrounding the archival materials (e.g. Nardi 1996). We model characteristic situations of use following ideas in Hansen et al. 2014, propose language technology solutions and assess their usefulness in practice by means of use cases (Jacobson et al. 1992, 2011).

1 Introduction

Currently, the large amounts of recorded speech available at Swedish memory institutions are rarely used due to the lack of effective methods for handling archival sound material. The aim of the project Tilltal is to examine how speech technology methods can make speech recordings at public memory institutions more accessible to researchers (cf. Edlund & Gustafson 2016). The project is a collaboration between Riksarkivet (RA), The Institute of Language and Folklore (ISOF), the Royal Institute of Technology (KTH) and SWE-CLARIN. It is funded by the Swedish Foundation for Humanities and Social Sciences from 2017 to 2020.

The ISOF archives in Gothenburg, Umeå, and Uppsala houses one of Sweden's largest folklore collections, as well as extensive dialect transcriptions and recordings. The result of more than a hundred years of data collection is archived; written, illustrated, and audio recorded narratives of everyday

lives and festivities. The material encompasses multiple genres including life stories, different types of folk poetry and music, traditions and beliefs. The Institute is dedicated to preserving and continuously adding to the dialect and folklore collections, and to the research and edification of Swedish folk culture and dialects.

The Tilltal project explores how speech technology methods and tools can be adapted and developed to process large amounts of historical voice recordings from the archives of ISOF. To make this possible, language technologists, SSH researchers and data holders work in close cooperation within the project.

2   Background – The Tilltal project

In a SWE-CLARIN workshop we had in 2015, the idea was to put together groups of people with three different roles: an SSH researcher, a speech technologist, and a data holder. We wanted to take the SSH researchers' current work procedure as the starting point and give them the opportunity to describe their work in small groups, allowing data holders and speech technologists to suggest ways in which the research process could be facilitated by large speech data sets and speech technology. The hands-on task was to come up with suggestions for research projects. As a result, the three sub-studies within the Tilltal project were conceived as use cases.

Three full research agendas from different fields and aimed at different types of speech analysis make up the following use cases. *Use Case 1: From personal experience narratives to cultural heritage* builds on ethnology and aims at providing more easy access and more entries to a well-known, multimodal collection. *Use Case 2: Linguistic variation in time and space* is predominantly sociolinguistic and aims for considerably more efficient and objective access to pronunciation varieties in recordings. *Use case 3: Interaction patterns over time and type of conversation* extends previous work within interaction analysis by adding large quantities of speech data in a field that has so far been almost exclusively tethered to purpose-recorded data sets.

Tilltal is intended as a framework in which existing speech technology methods are applied to speech materials that are found in archives. We are confident that these methods can be used fruitfully to some extent, but exactly what will work and what will not is impossible to tell without experimentation. In order to keep the investigations and their results valid and meaningful, we avoid artificial experimental tasks and focus instead on real, current research questions. The involvement of researchers and real research questions also allows us to add another key element, a research question on a different level: to study and document how the archives are used by the SSH researchers involved and how new methods can be introduced and tested in close cooperation with them.

3   Methods for studying usage

What we call the usage study of the Tilltal project should be seen as an overarching investigation and design process that involves all three use cases. It has two parts using activity theory and use case modelling respectively as methods for investigation.

3.1   Activity theory

The first part is to understand and describe the bigger picture of (using tools for) collecting, processing and making spoken narratives available for research at the Institute of Language and Folklore. Activity theory (e.g. Kaptelinin, 2013) is used in the project as a theoretical framework for understanding the use of tools and other artifacts in making spoken narratives available for research at the Institute. This involves considering contextual factors surrounding the core activities (see fig. 1): the community in which the activities take place (at the different archives), the division of labour between actors (e.g. IT staff, researchers and archivists) and the rules and conventions that govern the activities (as intellectual property rights, privacy laws and research ethics). Field diaries, observations, interviews, group discussions and workshops are used for data collection.
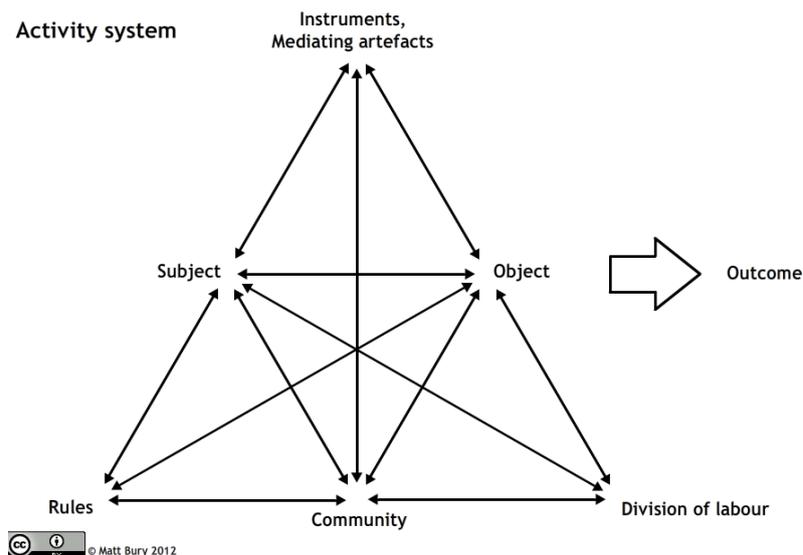
Figure 1. Engeström's activity system model.

Activity theory is a conceptual framework originating from the socio-cultural tradition in Russian psychology from the 1920s. Since then, the theory has been developed most notably by Leontyev (1978) and Engeström (1987) to become a theoretical tool for studying and describing human activities in social contexts focusing on the six related factors illustrated in Engeström's activity system model in fig 1. From the 1990s, activity theory has been used in Human Computer Interaction (e.g. Nardi, 1996) for -- in Kaptelinins words -- "understanding and designing technology in the context of purposeful, meaningful activities" instead of only focusing on individual tasks.

3.2   Use case modelling

The second part of the usage study applies use case analysis on collected data. The aim is to further specify and model researchers' needs of information and usable tools for seeking information in speech material: What kind of information is needed to answer a certain research question? To what extent can the language technology of today provide that information? Is the quality level of the delivered information sufficient for the researcher? What would be the optimal way to report the results to the researcher? What degree of interactivity with the result reporting system is convenient for the researcher? The use case analysis will help us formulate detailed requirements for evaluating existing tools and suggesting future solutions.

Use case modelling has been widely used in software engineering for requirement analysis since Ivar Jacobson introduced it in 1992. A use case is a list of actions or event steps, typically defining the interactions between an actor and a system, to achieve a goal. A collection of such use cases will serve as a way to identify common traits between researchers with different information needs as well as factors that differentiate between them. Such a factor may be the kind of information sought: whether it is found in the recording situation, in the relation between the participants, in the linguistic expression, or found in the topics and themes of the recording. 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) and Hansen et al (2014).

4   Goals, challenges and implications for design

The goal of the usage study as a whole is to provide guidance and constraints to the future development of usable tools for research based on speech data. By combining use case modelling with activity system analysis in the vein of Constantine (2009), we hope that this can be done without losing sight of the larger context within which digital tools and artefacts are and will be used. Furthermore, by actively involving researchers and archivists in the project on equal terms as in participatory design projects (e.g. Kensing & Blomberg 1998), we ensure that the results will meet the needs and practices of SHR researchers and, in the end, will be as usable as possible for them.

The approach of participatory design is very dynamic and requires an open and flexible process, as was demonstrated at a workshop in December 2016 on legal and ethical matters in making recordings generally available for research. Not only did it uncover more views, experiences and interests that influence the design process than expected. It also pinpointed great challenges with implications for system design. In particular, the folklore researchers requested a system that can handle different types of related data resources – recorded interviews, letters, notes and resumes from data collectors and researchers, query lists, etc. – as parts of one connected collection, not as isolated resources. Otherwise, it will not be possible to understand a recorded interview in context, in relation to other parts and overall goals of the collection. We will have the opportunity to explore this path further in *Use Case 1: From personal experience narratives to cultural heritages*.

5 Conclusion

The project Tilltal is a direct result of efforts to boost SSH collaborations with speech technologists in SWE-CLARIN, and we hope that its inauguration marks the beginning of a long line of fruitful, CLARIN-inspired and CLARIN-supported collaborations. The direct long-term goal is to make the Swedish speech archives more usable and accessible to SSH researchers in particular. We have shown how we achieve this not only by describing methods by which speech technology can be used to reach SSH research goals, but also by providing examples of fruitful interdisciplinary collaborations to make the results as usable as possible.

6 Acknowlegement

Reference

Berg, J., Domeij R., Edlund, J., Eriksson, G., House, D., Malisz, Z., Nylund Skog, S. & Öqvist, J. (2016). *Tilltal – making cultural heritage accessible for speech research*. Paper for the Proceedings of the CLARIN Annual Conference 2016 in Aix-en-Provence, France, 26–28 oct.

Constantine, L. L. (2009). Human Activity Modeling: Toward a pragmatic integration of activity theory and usage centered design. In Seffah, Vanderdonckt, & Desmarais (Eds.), *Human-Centered Software Engineering*. (pp. 27–51). London: Springer.

Edlund, J., & Gustafson, J. (2016). *Hidden Resources ─ Strategies to Acquire and Exploit Potential Spoken Language Resources in National Archives*. In N. C. (Conference Chair), K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, ... S. Piperidis (Eds.), Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016). Paris, France: European Language Resources Association (ELRA).

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Orienta-Konsultit Oy.

Hansen, P., Järvelin, A., Eriksson, G., & Karlgren, J. (2014). A use case framework for information access evaluation. In G. Paltoglou, F. Loizides, & P. Hansen (Eds.), *Professional Search in the Modern World: COST Action IC1002 on Multilingual and Multifaceted Interactive Information Access* (pp. 6–22). Springer.

Jacobson, I., Christerson, M., Jonsson, P., & Overgaard, G. (1992). *Object-Oriented Software Engineering: A Use Case Driven Approach. Addison-Wesley*.

Jacobson, I., Spence, I., & Bittner, K. (2011). *Use Case 2.0: The Guide to Succeeding with Use Cases*. Ivar Jacobson International.

Kaptelinin, V. (2013). *Activity Theory*. In M. Soegaard & R. F. Dam (Eds.), Encyclopedia of Human-Computer Interaction. Interaction Design Foundation.

Kensing, F., & Blomberg, J. (1998). Participatory design: Issues and concerns. Computer Supported Cooperative Work (CSCW), 7(3-4), 167-185.

Leontyev, A. N. (1978). *Activity, consciousness, and personality*. Pergamon Press.

Nardi, B. (Ed.). (1996). *Context and Consciousness: Activity Theory and Human-Computer Interaction*. Cambridge, MA, US: MIT Press.