



FRIAS

FREIBURG INSTITUTE  
FOR ADVANCED STUDIES

frequenz  
effekte



graduierertenkolleg 1624

# Workshop *Oral History Meets Linguistics*

03–04 December 2015

FRIAS, Albertstr. 19, großer Seminarraum

Abstracts

# Talks

Annette Gerstenberg, FU Berlin

## How to make sense of historical key words in biographical interviews

Oral History and (socio-)linguistics have profoundly different approaches and as a result, the textual data gained differ in many aspects. However, facing the challenge of combining sources from these two disciplines is very rewarding.

In order to illustrate the differences, but also to highlight the possibilities that such an integrated approach offers us, we will analyse two resources: (1) The linguistic corpus LANGAGE is comprised of biographical interviews with French men and women aged between 70 and 94 (in 2005); (2) the French section of the large Oral History database which is titled “ZWANGSARBEIT 1939–1945”. This is made up of interviews with French speakers (at the time of the interviews, 2006, aged between 81 and 86) who were forced to work in the program of SERVICE DU TRAVAIL OBLIGATOIRE (STO) as it was called in French.

We present these two resources considering the choice of participants, the interview methods and the resulting type of interaction. Then we look at the different methods of exploitation, based on the conventions of transcription, digitization, the possibilities of accessing the data online, and word-audio-video alignment.

Furthermore, we illustrate the core research questions applicable to these different types of corpus: While Oral History resources are designed to reconstruct the meaning and the sense attributed to important historical events by different individuals and social groups, linguistic corpora allow for the analysis of patterns of language use.

We formulate our research question exploring how the historical key word “Zwangsarbeit” / “Service du Travail Obligatoire” is used in the interviews, and what meaning can be attributed to it in the different linguistic contexts. Comparing extracts from both resources, we show how historical reference and individual connotation emerge in discourse.

**Gerstenberg, Annette.** 2009. The Multifaceted Category of ‘Generation’: Elderly French Men and Women Talking about May ‘68. *International Journal of the Sociology of Language* 200. 153–170.

**Niethammer, Lutz.** 1985. Frage – Antworten – Fragen. Methodische Erfahrungen und Erwägungen zur Oral History. In Lutz Niethammer & Alexander v. Plato (eds.). *Lebensgeschichte und Sozialstruktur im Ruhrgebiet 1930–1960, Band 3: “Wir kriegen jetzt andere Zeiten”. Auf der Suche nach der Erfahrung des Volkes in nachfaschistischen Ländern*, 392–445. Bonn: Dietz.

**Plato, Alexander v.** 2007. ‘Es war moderne Sklaverei’. Erste Ergebnisse des lebensgeschichtlichen Dokumentationsprojekts zur Sklaven- und Zwangsarbeit. *BIOS* 20(2). 251–290.

**Thompson, Paul.** 1980. The New Oral History in France. *Oral History* 8(1). 14–20.

**Documenting oral histories in the Russian Far East for multiple aims and uses**

This presentation will discuss aims and ways of documenting oral histories among indigenous peoples in the Russian Far East. It will focus on experiences from recording and reviewing Itelmen, Koryak and Even texts since the mid-1990s in Kamchatka for potential multiple uses. Among the questions that will be raised are: Should oral history be recorded in the indigenous language or in the vernacular? Although the first option should be preferred whenever possible, one should be aware of risks and shortcomings in both approaches. It will be discussed how certain recording techniques can influence or bias the information. Another critical question is if – or to which degree – oral traditions should be documented and transmitted in written form. It opens up the discussion if these might be preserved within the communities by preference in their original form, simulated on DVD, or by supplemented – linguistic or practical – scripts. This leads to the probably principal question, what the specific aims for the documentation of oral history are. In sum I will point to the obvious advantages and mutual feedbacks in approaches for the documentation of oral history that aim to produce well-integrated outcomes which bring together or combine potential multiple uses. More exact and complete data can be obtained if oral history is recorded in the given indigenous language. As electronic learning tools, oral histories can stimulate particular interest in the preservation of endangered languages and indigenous knowledge among younger generations. As younger people have often no longer sufficient competence in their indigenous language, the required Russian subtitles in the film and in supplemented print editions serve additional purposes even for linguistic research, as larger amounts of indigenous language data must be transcribed and translated.

**From dialect corpus to interactive open-access database: FRED online**

FRED – the Freiburg Corpus of English Dialects – is a monolingual spoken-language dialect corpus including oral history interviews with speakers from nine larger dialect areas in England, Scotland, Wales, the Hebrides and the Isle of Man. The corpus consists of sound recordings and orthographic transcripts, spanning approximately 2.5 million words and 300 hours of speech (cf. Hernández 2006, Szmrecsanyi & Hernández 2007). A central aim in creating FRED was to provide a solid geographically balanced database for investigations into morphosyntactic variation in British English dialects. Up to this point, such investigations using the full corpus have been restricted to on-site research in Freiburg, with off-site uses being limited to a small subset of the corpus. This talk outlines how FRED is now made available to the whole research community by publishing it online via FreiDok. We will demonstrate how FRED transcripts and audio files (in a first step, of the subset “FRED-S”, comprising approximately 1 million words) can be accessed online. Furthermore, we introduce an interactive research database which facilitates searching for words in the texts and sorting them by social parameters of the speakers, such as age, sex and dialect area. In sum, this data presentation aims to show that FRED online offers multiple opportunities for research and teaching – not only for linguistics, but also, for example, for studies of oral history.

**Hernández, Nuria.** 2006. *User's guide to FRED: Freiburg Corpus of English Dialects*. English Dialects Research Group. University of Freiburg. Url: <https://www.freidok.uni-freiburg.de/data/2489>.

**Szmrecsanyi, Benedikt & Hernández, Nuria.** 2007. *Manual to accompany the Freiburg Corpus of English Dialects Sampler (“FRED-S”)*. English Department. University of Freiburg. Url: <https://www.freidok.uni-freiburg.de/data/2859>.

Cord Pagenstecher, FU Berlin

**History and memory. Curating and analyzing large-scale interview collections with survivors of Nazi persecutions**

Traditionally oral historians have analyzed individual interviews, usually conducted by themselves. Now, digital technologies support the creation and curation of large-scale collections making thousands of interviews accessible and searchable. The Center for Digital Systems at Freie Universität Berlin is hosting several major collections with testimonies focusing on Nazi Forced Labor and the Holocaust. The Visual History Archive of the USC Shoah Foundation ([www.vha.fu-berlin.de](http://www.vha.fu-berlin.de)), the Online Archive “Forced Labor 1939–1945” ([www.zwangsarbeit-archiv.de](http://www.zwangsarbeit-archiv.de)), and the British “Refugee Voices” collection ([www.refugeevoices.fu-berlin.de](http://www.refugeevoices.fu-berlin.de)) contain thousands of audio-visual life-story interviews, for which online platforms with sophisticated mapping, searching and annotating tools were developed. Therefore, possible overlaps between historical and linguistic questions, approaches and methods are of great concern, but still need to be elaborated.

Studying interviews, historians are usually interested in two layers: the facts that are being told by the interviewee, and the way these facts are being remembered and narrated. On the level of factual history, historians want to extract the names of places or persons or the dates of events. In the case of Holocaust or forced labor survivors, we try to identify and verify the names of concentration camps or forced labor factories mentioned in the interviews in order to contextualize the individual stories with other documents on these cases.

On the level of memory, we are looking for narrative patterns, specific metaphors, unfinished sentences, silences etc. For example, I looked at when and how the interviewees in the “Forced Labor 1939–1945” collection used the symbolic words slavery or slave labor in order to describe their forced labor experience in Nazi Germany. Linguistic methods can help in analyzing the role of the interviewer in the creation of the interview, in comparing interviews conducted in different time periods, and in analyzing gendered memory patterns.

Speech recognition and alignment tools can help in the curation of interview collections through supporting human transcription procedures or through creating a ‘dirty’ transcript for search purposes. It can also point the researcher to some important non-verbal communication structures.

The paper will present various exemplary questions and research topics that need interdisciplinary efforts between historians and linguists.

Stefan Pfänder, University of Freiburg

**Recording voice – giving voice – analysing voice: Autobiographical narratives as corpora**

In a well-known invited talk in the former GDR, Pierre Bourdieu addressed the task of the scientist who investigates social interaction. On the one hand, this task consists in listening to people very carefully and collecting their testimonies; an effective means here being the sound recording of autobiographical narratives. The researcher of the social, however, also has the task of giving those people a ‘voice’ who are not heard in public dialogue. This admonition by Bourdieu inspired a book series of multimodal corpus publications, with the release of one new volume per year since 2002. These small corpora gather memories of the 20th century, i.e. testimonies, which can be assigned to the category of oral history. In my talk, I would like to 1. introduce some of the volumes of this series, 2. show how the series editors give a ‘voice’ to the narrators and 3. illustrate how these corpora have been used and could be used for linguistic research.

Christoph Purschke, University of Luxembourg

**“None of the speakers knew that they were speaking for the Führer”:  
Opportunities and challenges in exploring politically motivated audio corpora  
using the example of the “Lautdenkmal reichsdeutscher Mundarten”**

Between 1936 and 1938, the Reichsbund of German civil servants initiated a large-scale recording series dealing with German vernaculars with technical support from the company Telefunken Schallplatte: the so-called “Lautdenkmal reichsdeutscher Mundarten zur Zeit Adolf Hitlers” (Sound Memorial of Vernaculars in the German Reich at the Time of Adolf Hitler). The scientific part of the survey was essentially conceived by the Marburg linguists Walter Mitzka and Bernhard Martin. This collection, encompassing 400 recordings of dialect speakers throughout the entire German Reich (including Austria and the Sudetenland) and compiled as a birthday present for Adolf Hitler, has been left largely unexploited to the present day, mainly due to the conditions under which it was created and preserved. In collaboration with the research centre “Deutscher Sprachatlas” in Marburg, the collection is currently being reconstructed, explored and edited, including the available context materials, in order to finally make these exciting data accessible to the public.

The talk gives an impression of the ongoing work on tapping and reconstructing the Lautdenkmal. This comprises insights into the state of the material and problems in processing it as well as examples of how this type of audio data can be exploited with regard to both linguistic and historico-cultural research questions.

Michael Rießler & Joshua Wilbur, University of Freiburg

## **Oral histories in endangered language corpora**

A number of endangered language documentation projects affiliated with the Freiburg Research Group in Saami Studies have been building language corpora for endangered Uralic languages for several years. While the data gathered on these minority languages (Pite Saami, Skolt Saami, Kildin Saami and Izhva Komi) corresponds to a wide variety of genres, a common type of recording can clearly be considered oral history. Our main motivation as linguists doing such recordings is to collect non-elicited, unplanned examples of the target language in the spoken modus on topics that speakers can relate to in a comfortable, relaxed, natural way, which aligns well with expectations of oral history. It is precisely such recordings which can prove to be valuable sources for other disciplines as well, particularly anthropology.

In our talk, we will present our approach to language documentation, and show examples of oral history recordings in our corpora. We believe that our heavily annotated recordings guarantee intellectual access even for researchers who do not speak the languages we document because these include translations into the respective national languages as well as English and/or Russian. Our extensive metadata provide relevant background information on speakers, the recording session itself, as well as project details, which thus allows non-linguists not only to contextualize recordings, but also to filter for potentially relevant categories such as gender, age or other topical keywords.

The original recordings (audio and frequently video), annotations and metadata are available to the research community via international language archives. While much work still needs to be done, we envision creating extensively linked documentations. This means that the metadata and annotations for individual recordings are tagged with keywords on relevant subjects, thus allowing searches both within a single archive, but also in connection with external archives and/or search engines, for both linguistics and other topics. Last but not least, our documentations can be used to provide endangered language communities with valuable linguistic and cultural resources to be utilized for instance for teaching, for local or even family history.

Jonnie Robinson, British Library, London

**Spectacular Vernacular: Oral history recordings at the British Library as a corpus of British English**

The British Library has in its sound archives recordings which capture spoken English over more than 100 years. The collections include the complete audio archives of internationally acclaimed surveys such as the *Survey of English Dialects*, *Survey of Anglo-Welsh Dialects*, *British National Corpus* and *BBC Voices*.

In addition to its prestigious language and dialect holdings the Library's extensive oral history collections represent a rich and relatively unexploited resource for linguists. Of particular interest to dialectologists and sociolinguists are substantial collections with nationwide coverage, such as the *Millennium Memory Bank* and *Listening Project*, both of which feature speakers from diverse social and geographic backgrounds across the UK.

In recent years the Library has enjoyed considerable success in developing access to this unique audio content. The Sounds website at <http://sounds.bl.uk/> includes entire collections and selected excerpts from others, and the educational resource *Sounds Familiar* (<http://www.bl.uk/soundsfamiliar>) presents linguistic analysis of *Survey of English Dialects* and *Millennium Memory Bank* content for a school, undergraduate and teacher audience.

This paper provides an overview of the Library's oral history collections and evaluates their relevance for studying historical and present-day varieties of British English. Examples from recordings will illustrate how these collections can be used to investigate regional and social variation, change over time and support research, teaching and learning in a variety of contexts.

Katja Roller, University of Freiburg

**Exploring the quantitative side of dialect perceptions by means of oral history interviews from Wales: The Millennium Memory Bank**

In the late 1990s, the BBC and the British Library launched a project of recording memories of people from a broad range of locations all over England, Wales, Northern Ireland and Scotland, which was later to become Europe's largest oral history archive: the Millennium Memory Bank (MMB). The collection, aimed at providing "a 'snapshot' of Britain at the turn of the millennium" (Perks 2001: 95), includes more than 6,000 interviews with speakers from different age groups and with diverse occupational and educational backgrounds. As well as serving as a substantial source for historico-cultural research, the MMB carries a lot of potential as a database for linguistic analyses, offering huge amounts dialect speech from different corners of the UK (cf. Haigh 2015, Roller 2015).

This talk outlines how MMB recordings from Wales can be used for linguistic research. The study presented explores (1) how consciously people from Wales and from London perceive different grammatical features of Welsh English (determined through questionnaires) and (2) whether the features perceived more consciously are more frequent in spoken Welsh English. This second analysis is (partly) based on transcribed MMB interviews. In these data, the frequencies of occurrence of different non-standard grammatical features were determined and contrasted with the perception data from the questionnaires. In the context of these analyses, I discuss some advantages and challenges of using oral histories for dialectological research, and finally point to examples of how linguistic approaches to MMB data can provide new insights for oral historians.

**Haigh, Sarah.** 2015. *Investigating regional speech in Yorkshire: Evidence from the Millennium Memory Bank*. MPhil thesis, University of Sheffield.

**Perks, Rob.** 2001. The Century Speaks: A public history partnership. *Oral History* 29(2). 95–105.

**Roller, Katja.** 2015. Towards the 'oral' in oral history: Using historical narratives in linguistics. *Oral History* 43(1). 73–84.

Susanne Wagner, University of Mainz

**Of cider making, wool mills and ironworks or: From oral history to dialect corpus  
– the making of FRED**

When work began on FRED in 2000, hundreds of archives, museums and other likely (and unlikely) resources had already been contacted, mostly with limited success. Three of us, starting out on our PhDs at the time, divided England into three areas, making me responsible for the Southwest. In the following years, I travelled through the Southwest to scout museums, libraries and private collections, some in beautiful surroundings but abysmal data preservation states. One oral history collection each in Somerset, Wiltshire and Cornwall as well as an “outlier” in the Midlands proved to be the most fruitful for FRED, and I will briefly relate how and what was acquired.

In this talk, I will focus on the practical data collection stage first, before discussing subsequent data preparation and presentation. Aspects to be highlighted include practical difficulties, technical issues, access issues, aspects of privacy and copyright, as well as transcription decisions and subsequent work with the oral history data. I will also briefly look at similar data/projects that might have been prepared much more systematically but are hardly used at all (as well as the reasons behind that), possibly leading to a more general discussion of “big data”, its (dis)advantages and pitfalls.

## Posters

Udo Baumann, University of Freiburg

### Using the Listening Project for a new corpus of spoken English

The Listening Project is an audio archive of intimate conversations between friends and relatives. The conversations are recorded by the BBC and stored in the Oral History archive of the British Library (<http://sounds.bl.uk/Oral-history/The-Listening-Project>). Participants from all areas of the United Kingdom are invited to contribute to the Listening Project by sharing a conversation. Each conversation offers a unique glimpse of friendships or family backgrounds of our time. Up to now, hundreds of conversations have been recorded and archived and the project is still going on.

However, the Listening Project is not only a fascinating cultural-historical document, but can also be used as a corpus of contemporary spoken British English: Many of the conversations come close to naturally occurring face-to-face conversations, which means that they are a valuable resource for the linguist interested in the study of naturally occurring language data. For my dissertation about ongoing changes in the use of the English progressive construction, I am currently compiling a new corpus of contemporary spoken English. For this purpose, I am planning to transcribe several conversations of the Listening Project ( $\sim 100,000$  words) and to use the transcripts to study progressive use in face-to-face conversation. Comparing the results to mid and late 20th-century data will offer the chance to track language change in progress.

**Fading memories and linguistic fossils in audio records of Kerala Jews**

In the years 2008–9, a project of language documentation of Jewish Malayalam was conducted in Israel under the auspices of the Ben Zvi Institute, guided by theoretical and practical guidelines for recording languages on the verge of extinction (Gippert et al., 2006; Messineo, 2008). Jewish Malayalam is a castolect of Malayalam, one of the four major Dravidian languages of South India and a Jewish language. Jewish Malayalam speakers migrated en masse to Israel in 1954, followed by the decline in the use of Jewish Malayalam. It is only as late as the 2000s, that their language became the focus of study and documentation (Zacharia and Gamliel, 2005; Gamliel, 2009; 2014).

Language documentation projects of fading languages like Jewish Malayalam exemplify the interface between linguistics and oral history; they contain first hand accounts of historical events, ethno-histories, memoirs and family anecdotes besides linguistic data on a language that may otherwise leave no records for future generations. They constitute a corpus of both linguistic data and oral genres applicable for multi-purpose research concerns (Himmelman, 1998). In the past two decades, documentary linguistics emerged as a field generating unique multi- and inter-disciplinary analytic approaches and scholarly practices (Austin, 2014). Digital archives worldwide resulted from the efforts to document fading languages (Austin, 2015). These archives offer opportunities for both linguists and anthropologists to describe and analyze little known areas of human heritage of cognition and culture. The paper presents the Jewish Malayalam documentation project as a case study dealing with the linguistic fossils revealed while documenting the fading memories of Kerala Jews in Israel.

**Austin, Peter.** 2014. Language Documentation in the 21st Century. In *JournalLIPP* 3. 57–71.

—. forthcoming. Language Documentation 20 Years on. In Martin Putz and Luna Filipovic (eds.). *Endangered Languages Across the Planet: Issues of Ecology, Policy and Documentation*. Amsterdam: John Benjamins.

**Gamliel, Ophira.** 2009. Jewish Malayalam. In *International Journal of Dravidian Linguistics* 38(1). 147–175.

—. 2014. Voices Yet to Be Heard: On Listening to the Last Speakers of Jewish Malayalam. In *Journal of Jewish Languages* 1(1:1). 135–167.

**Gippert, Jost, Nikolaus P. Himmelman and Ulrike Mosel** (eds.). 2006. *Essentials of Language Documentation*. Berlin: Mouton de Gruyter.

**Himmelman, Nikolaus P.** 1998. Documentary and Descriptive Linguistics. In *Linguistics* 36. 161–95.

**Messineo, Cristina.** 2008. Fieldwork and Documentation of Speech Genres in Indigenous Communities of Gran Chaco: Theoretical and Methodological Issues. In *Language Documentation and Conservation* 2(2). 278–279.

**Zacharia, Scaria and Ophira Gamliel** (eds. and trs.). 2005. *Kārkulali – Yefefiah – Gorgeous! Jewish Women’s Songs in Malayalam with Hebrew Translations*. Ben Zvi Institute, Jerusalem (in Hebrew and Malayalam).

Sonya Kinsey, University of Freiburg

## Owls, Sasquatch and Tsun' Dye: Uncovering indigenous Englishes through oral storytelling

Traditional storytelling practices are at the heart of many indigenous communities, and are important modes of transmitting history and culture from one generation to the next (Atloe & Fitznor, 2010; Morin, 2011). In the case of the Witsuwit'en people, located in north central British Columbia, a dramatic loss in fluent Witsuwit'en speakers has led to English becoming the dominant language of the reserves, including Moricetown. Language decline, stimulated by the residential school system and a declining number of fluent speakers, has reduced the number of Witsuwit'en speakers by 80% over twenty years (Hargus, 2007). Although language revitalization efforts are underway, including elementary school instruction, and a dictionary, the language has not gained any younger, fluent speakers (Hargus, 2007). English has therefore become the primary language through which oral histories are transmitted to children. Even the community's own textbook uses English as the main language, and presents all the traditional stories, told by bilingual Elders, in English (Morin, 2011). My project focuses on defining and understanding the features and origins of Witsuwit'en English. In order to uncover the most authentic, community centred English dialect, I asked all study participants to tell me about their lives and the area's history. This included any traditional stories that the participants had heard from their own parents or grandparents. When instances of non-standard English were counted, a correlation was found between those with frequent non-standard features and knowledge of local stories. Perhaps unsurprisingly, the participants who had the lowest number of non-standard features in their speech, also were unable to tell any of these stories. In the case of my fieldwork, a break in dialectal language transmission often coincided with a break in cultural transmission.

- Atloe, M. R., & Fitznor, L.** 2010. Aboriginal Educators Discuss Recognizing, Reclaiming, and Revitalizing Their Multi-Competences in Heritage/English-Language Use. *Canadian Journal of Native Education*. 13–34.
- Hargus, S.** 2007. *Witsuwit'en Grammar: Phonetics, phonology, morphology*. Vancouver: UBC Press.
- Morin, M. H.** 2011. *Niwhts'ide'ni Hibi'it'en The Ways of Our Ancestors Witsuwit'en History & Culture Throughout the Millennia*. Smithers: Friesens.