

# DOCUMENTARY LINGUISTICS I

prof. Nicole Nau, UAM winter 2018/2019

**Second lecture**  
09 October 2018

# TOPICS OF THE DAY

- ❖ What is (a) language documentation (LD) and what is it good for?
- ❖ Short summary: how and why languages were documented in earlier times
- ❖ Principles of documenting languages in the 21st century
- ❖ Contents of a LD: primary data and metadata
- ❖ Literature on language documentation and documentary linguistics

# WHAT IS LANGUAGE DOCUMENTATION?

«linguists generally view language documentation [...] as **work** undertaken by an amateur or professional linguist that **results** in some **record** of a language for the **purpose** of furthering the study and understanding of that language.» (Berge 2010)

«documentation» in Polish: (a) dokumentowanie, (b) dokumentacja

Why and by whom is this work done?

Who may need the result?

# EARLIER DOCUMENTATION OF LANGUAGES (SEE 1ST LECTURE)

Example: Tulu, a Dravidian language of India

Grammar:

Brigel, J. 1872. *A Grammar of the Tulu Language*. Mangalore: C. Stolz. Basel Mission Book & Tract Depository

Dictionary:

Männer, A. 1886-1888. *Tulu-English and English-Tulu Dictionary*. Mangalore: Basel Mission Press.

# LANGUAGE DOCUMENTATION IN THE 19TH-20TH C: CHARACTERISTICS

- ❖ done by linguists (often only 1 person)
  - ❖ done by outsiders (not part of the community)
  - ❖ results published on paper (usually in small editions)
  - ❖ results not made accessible to the speech community
  - ❖ «Boasian standard»: a dictionary, a grammar, and a collection of texts
  - ❖ text collection: most often narrative texts (traditional stories)
  - ❖ grammar: contains isolated sentences, 20th c often elicited/constructed
- => what was wrong with that?

# DOCUMENTARY LINGUISTICS TODAY

«**Documentary Linguistics** [...] is concerned with the making and keeping of **records** of the world's **languages** and their **patterns of use**.» (Woodbury 2003)

# WHY DOCUMENT A LANGUAGE? FOR WHOM?

To have a record for the future (and the present) -

- ❖ for **speakers** of the language: to preserve the language as part of the culture, the community's heritage
- ❖ for **researchers**: to have material for their investigations
  - general linguistics, language typology, historical linguistics, ...
  - psychology, ethnography, ...
- ❖ for **educators**: to have material for teaching the language and about the language
- ❖ for people and purposes we don't know yet!

# WHAT IS **A** LANGUAGE DOCUMENTATION?

«a comprehensive corpus of primary data which leaves nothing to be desired by later generations wanting to explore whatever aspect of the language they are interested in» (Himmelman 2006: 3)

Of course, this is an **ideal**!



# PRINCIPLES OF LANGUAGE DOCUMENTATION (BEGINNING OF 21ST CENTURY)

- ❖ done in interdisciplinary teams: linguists + other specialists,
- ❖ outsiders and insiders: speech community is involved not only as „informants”
- ❖ done for various purposes
- ❖ published electronically (and on paper)
- ❖ stored in digital archives
- ❖ contains written, audio and video material
- ❖ documents the use of the language in various situations

# DOCUMENTARY LINGUISTICS AND LANGUAGE DOCUMENTATION TODAY


«**Documentary Linguistics** [...] is concerned with the making and keeping of records of the world's languages and their patterns of use.» (Woodbury 2003)

«**language documentation** is the **creation, annotation, preservation** and **dissemination** of transparent records of a language» (Woodbury 2011)

# QUESTIONS

If your language were endangered, what would YOU want to document for the future? What records would you create?

Given the limits of time, money, and persons: what should be documented first?



«a language documentation should strive to include as many and as varied records as practically feasible [...] Ideally, then, a language documentation would cover all registers and varieties» (Himmelman 2006)

# CONTENTS OF A LD: BASIC DISTINCTIONS

- ❖ Primary data

- ❖ Metadata

- ❖ Annotation

---

- ❖ Description (results of analysis: grammar, dictionary)

# DATA, ANNOTATION AND METADATA IN AN EARLY RECORD (THE LORD'S PRAYER IN LAVIAN IN MÜNSTER'S COSMOGRAPHIA 1588)

## Von dem Teutschen Landt.

## Notiz

Ein andere Statt in Lifflandt mit nammen Derpst/hat ire Rauffmanschafft mit den Die Statt Derpst. Reuel. Reussen. Reuel aber ist weiter gegen Mitnacht zu dem Ostsee gelegen/gebraucht sich nicht weniger der Rauffmanschafft als Riga/ist auch schier derselb brauch da. Sie haben Lübeckische Recht/welcher sich die andern Stettlein auch gebrauchen. Es sind viel Sprach in Lifflandt/deren sich keine mit der andern vergleicht. Zu Reuel vnd zu Derpst reden sie ein Sprach/heißt Estensch/vñ vmb Riga braucht man Lnuisch/da ist die recht Liffländische Sprach/vnd betten daselbst die Bawren/so Gottes Wort bericht haben/auff folgen de weiß ihr Vatter vnser:

Vatter vnser der du bist in Himmeln/ geheiligt werd dein Nam/ zutomm vns dein  
Täbes mus kas tu es eckschan debbessis/ Schwetis tows waark/ enack mums tows  
Reich dein will geschehe wie in Himmeln also auff Erden. Vnser täglich  
walstibe/ tows praats bus ka eckschan Debbes/ta wursan semmes. Musse denische  
Brot gib vns heut vergib vns vnser schuld als wir vergeben vnsern  
Mäyße düth mums schodeen/ pammat mums müsse gräke kames pammat müsse  
Schuldigern/ nicht eynfähr vns bösen versuchung sonder behüt vns vor allem bösen.  
parradueken/ Ne wedde mums louna badeckle/ pett passarga mums nu wüsse louna.

Dis alles oder zum größern theil/hab ich von dem erfarnen Johann Hasentödter/ so manich jar in Lifflandt an den Herren Höfen vñ Cansleyen gewesen/vñ viel erfahre hat.

# THE LORD'S PRAYER IN A SORBIAN DIALECT, RECORDING WITH ANNOTATION IN ELAN – WHAT IS THE PRIMARY DATA?



00:00:00.000 Selection: 00:00:00.000 - 00:00:00.000 0

KSF-006-2012072...

00:00:00.000 00:00:01.000 00:00:02.000 00:00:03.000 00:00:04.000 00:00:05.000 00:00:06.000 00:00:07.000 00:00:08.000 00:00:09.000 00:00:10.000 00:00:11.000 00:00:12.000 00:00:13.000 00:00:14.000 00:00:15.000 00:00:16.000

00:00:00.000 00:00:01.000 00:00:02.000 00:00:03.000 00:00:04.000 00:00:05.000 00:00:06.000 00:00:07.000 00:00:08.000 00:00:09.000 00:00:10.000 00:00:11.000 00:00:12.000 00:00:13.000 00:00:14.000 00:00:15.000 00:00:16.000

topics. [1]  
spch.d [13]  
ft.deu [13]  
**ft.eng** [13]  
nt.deu [3]

Lord's Prayer in the dialect of Schleife; German translation: current ecumenical version, Arbeitsg

|                     |                 |                  |                                  |
|---------------------|-----------------|------------------|----------------------------------|
| Wóšce naš, kiž sy   | swěćone bydž    | pćidž k nam Twój | Twója wóla se stanjo kaž na nj   |
| Vater unser im Him  | geheiligt werde | Dein Reich komm  | Dein Wille geschehe, wie im Hi   |
| Our Father in heav  | hallowed be yo  | your kingdom co  | your will be done, on earth as i |
| 00:00:03# der du bi |                 |                  |                                  |

# PRIMARY DATA, 1. = WHAT CAN BE OBSERVED

«a language documentation, conceived of as a lasting, multipurpose record of a language, should contain a large set of primary data which provide evidence for the language(s) used at a given time in a given community (in all of the different senses of “language”). Of major importance in this regard are specimens of **observable linguistic behavior**, i.e. examples of how the people actually communicate with each other.» (Himmelman 2006)



# OBSERVABLE LINGUISTIC BEHAVIOR

What is a communicative activity? Give examples!

What constitutes a communicative activity / a speech event?

What part of an activity / event is **linguistic** behavior?

What part(s) should be contained in a documentation?

Does one speech event / activity contain one language?

<http://www.sorosoro.org/en/category/video-of-the-week/>

# PRIMARY DATA, 2. METALINGUISTIC KNOWLEDGE

What is metalinguistic knowledge?

Give examples or show your metalinguistic knowledge!

How can metalinguistic knowledge be accessed and documented?

# EXAMPLES FOR METALINGUISTIC KNOWLEDGE

- ❖ I know that a Christian prayer ends with the word «Amen».
- ❖ I know that «pyra» is used in the Poznan dialect of Polish.
- ❖ I know where the stress is on Polish words.
- ❖ I know that «I know that ...» is a possible and normal way to start a sentence in English, but «I know if ...» is not.
- ❖ I know that Polish words ending in «(i)arnia» designate places such as shops and cafes.



# NALESNIKARNIA GRAMOFON

パンケーキ CREPES БЛИНЫ PFANNKUCHEN CREPELLE 煎饼



# (META)LINGUISTIC KNOWLEDGE VS. ENCYCLOPEDIC AND CULTURAL KNOWLEDGE

**Piwiarnia** – lokal gastronomiczny przeznaczony głównie do spożywania piwa.

Oprócz funkcji gastronomicznej piwiarnie pełnią funkcję społeczną, integrując bywalców. Lokale te służą swobodnym pogawędkom i zabawie. Często w piwiarniach istnieje możliwość gry w karty, kości, darty, czy piłkarzyki. W piwiarniach często można usłyszeć lekką muzykę (piosenka biesiadna).

Do wielu gatunków piw, nalewanych z beczek, jak również podawanych w butelkach, serwuje się zwykle ciepłe i zimne przekąski.

(<https://pl.wikipedia.org/wiki/Piwiarnia>)

# CONTENTS OF A LD: BASIC DISTINCTIONS

## ❖ Primary data

- ❖ Observable linguistic behavior, speech events, communicative activities = how people use their language
- ❖ Metalinguistic knowledge = what people know about their language and its use

## ❖ Metadata

## ❖ Annotation

- 
- ❖ Description (results of analysis: grammar, dictionary)