Language Documentation: Values, Theories, Practices
Summary Report I

Jeff Good
Max Planck Institute for Evolutionary Anthropology
and The Open Language Archives Community
goodeva.mpg.de
My perspective

- Primary interest: The technological side of language documentation
- What this means here
  - I have no values
- More precisely: Good “tech. support” for language documentation requires the community, not the technician, to have a clear vision of what they need the technology to do
Values

• Two (broad) visions
  • Language documentation for linguistics
  • Language documentation for “the larger human community”

• What is this community?
  • All interested people (especially academics)
  • All speakers of a language
Theories

• The two visions have parallels in two competing theoretical stances in linguistics

• Linguistics as a historical discipline (the “new” typology)

• Linguistics as a science

• Aside: I noted a lot of discussion of values and practices, but little of theories
NSF/NEH’s take

- Each endangered language embodies unique local knowledge of the cultures and natural systems in the region in which it is spoken.
- These languages are among the few sources of evidence for filling in the record of the human past.
- The great variety of these languages represents a vast, largely unmapped terrain on which linguists, cognitive scientists, and philosophers can chart the full capabilities—and limits—of the human mind.
Practices

• Two sets of values, two sets of practices

• Document and develop

• Describe

• Prototypes

• Transcribed recordings and “developmental” materials

• Grammar, dictionary, texts
Practices

• Documentation for the “human community”
• Lots of genres
• “Rich” recordings (e.g., lots of naturalistic video)
• Collaboration with communities
• Mobilized resources (linguistics as language development? Is this documentation?)
Practices

• Documentation for linguistics

• Focus on creation of the “Holy Trinity”: Grammar, dictionary, text

• Targeted *description* of linguistically “interesting” features of the language

• Lots of specialized data formats (e.g., interlinear text, paradigms)
The good news

• The two sets of values/practices can be complementary

• For example, different genres reveal different grammatical features

• And linguistically-motivated data gathering finds patterns that might otherwise be lost (like Yak-naming conventions)
The mixed news

• We’ve also heard that community-linguist interaction can help achieve both goals

• However, I don’t think we really understand this very well yet—from a social or a technological perspective

• Should linguists try to be historians, scientists, and language developers? Can we be?
The mixed news

• Take, as an example, mobilization

• Mobilization of language resources is undeniably a good thing

• Is it the linguist’s duty to mobilize resources (which often means creating unarchivable materials)?

• Or just to produce mobilizable resources (or portable resources in Bird and Simons’ sense)?
The bad news

- Resources are limited
- In 2005, we cannot be
  - Ideal historians
  - Ideal scientists
  - Ideal educators
- So what should we try to be?
What should we be?

• I heard compelling arguments that we need to be historians and scientists

• Being good historians causes us to find scientifically interesting things

• Being good scientists helps us decide what to record for history
What should we be?

• I think we need to think harder about linguist as language developers

• Is it part of the *practice of linguistics* to serve our communities *or*

• Do we serve communities because it’s “ethical” *or*

• Do we serve communities because we as individuals value doing this?
Ethics/Values

• Paraphrasing Michael Krauss...

• Linguistics can be ethical only
  • If it documents languages before they disappear
  • It documents them in a way which helps keep them from disappearing

• Do we all agree?
Cautionary tale

• It became hard to do documentary work within Australia
• So, the linguists went elsewhere
• With 6,000+ languages to be documented, this seems to be a perfectly rational response...
• ...if we have a responsibility to all humanity not only to speaker communities
Advice needed

- So, what tools and standards do we need?
- Do we stop at basic recording and transcription tools?
- Do we build tools for scientific research?
- Do we build tools to create community resources?
- What kind of training do we need?
Advice needed

• Of course, our attitudes will evolve over time

• Assuming history, science, and community support are laudable goals but that we can’t do it all...

• What do the linguists do now?