

A photograph of a village with several small huts with thatched roofs, surrounded by a dense forest of palm trees. The ground is sandy and there are a few people walking around. The image has a slightly faded, vintage feel.

Less-Networked Speaker Communities and Digital Language Archives

Lise M. Dobrin
University of Virginia
dobrin@virginia.edu

DH 2012 Hamburg

The Arapesh Grammar and Digital Language Archive Project

Archive

- Catalog
 - Recorded Events
 - Participants
- About the Archive
 - The Recordings
 - Standards & Infrastructure
 - How to Apply for Access
 - Funders
 - Staff

Arapesh Language

- Social & Historical Context
 - Varieties & Affiliation
 - Vitality
- Grammar
 - Phonology
 - Morphosyntax
- History of Study
 - Bibliography

General Resources

- Sample Texts
 - Bethlehem
 - Ehiwac
- Reo F. Fortune's *Arapesh*
- Map
- Image Gallery
 - Wautogik village, 1997-99
 - Mead-Fortune fieldwork, 1931-32
- Field Notebooks
- Other Resources

Welcome

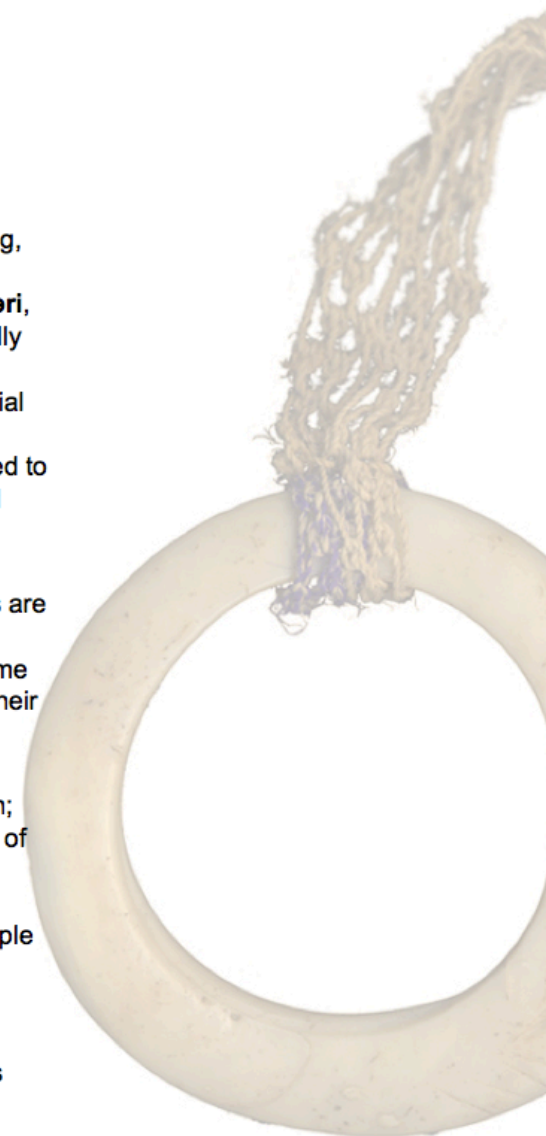
Welcome to the home page of the Arapesh Grammar and Digital Language Archive Project. The project was conceived as a way of preserving, integrating, and disseminating some of the rich documentary material that has been produced on the Arapesh languages - **Arapesh, Alpef, Arāpec, Bukiyip, Wəri, Ilahita, Abuʔ, Muhianj, "Apəkin Boran"** or "**our talk**" - languages traditionally spoken by some twenty five thousand villagers living on the coast and in the Torricelli hills of East Sepik Province, Papua New Guinea. Arapesh is of special linguistic interest because of its typologically unusual system of grammatical categorization that makes systematic reference to a noun's sound, as opposed to its meaning, in order to determine its class and hence what plural marker and agreement forms it will take.

Like numerous other minor languages the world over, the Arapesh languages are endangered. Many Arapesh villagers today use Papua New Guinea's creole lingua franca, Tok Pisin, as their medium of communication in daily life. In some cases, middle aged adults can understand but not speak the language, and their children have almost no knowledge of their ancestral tongue at all.

The core of the archive consists of transcribed audio recordings and a lexicon; these represent Arapesh varieties spoken on or near the coast in the villages of Wautogik, Woginara, Dogur, and Balam. It is hoped that these and other materials brought together in the archive will serve as a useful resource for anyone interested in Arapesh, including scholars, students, and Arapesh people themselves.

We are grateful to the National Endowment for the Humanities for financial support, through the Documenting Endangered Languages program, that has made this project possible.

Background image: The shell ring or yuwa (pl. yuwahas) is a highly valuable traditional Arapesh



Franz Boas (1911)

Handbook of American Indian Languages

Susan Gal (1979)

language shift on the Austria-Hungary border

Nancy Dorian (1981)

Scottish Gaelic language shift

Annette Schmidt (1985)

young people's Dyirbal

Joshua Fishman (1991)

reversing language shift

“Recent advances in information technology can magnify the effect of prompt and coordinated fieldwork. These advances make it possible not only to document endangered languages before they fall silent, but also to integrate and analyze that body of knowledge in unprecedented ways. Computerization of speech and universal Internet access is transforming the practice of linguistics in the area of endangered languages.”

**NSF Solicitation 11-554
Documenting Endangered Languages Program
“data, infrastructure and computational methods”**



E-MELD is a 5-year project with a dual objective:

- 1) To aid in the preservation of endangered languages data and documentation.*
- 2) To aid in the development of the infrastructure necessary for effective collaboration among electronic archives.*

LINGUIST List Projects



Members of the scientific community are faced with two urgent situations: the number of languages in the world is rapidly diminishing while the number of initiatives to digitize language data is rapidly multiplying. The latter might seem to be an unalloyed good in the face of the former, but there are two ways things may go wrong without adequate collaboration among archivists, field linguists, and language engineers. First, a common standard for the digitization of linguistic data may never be agreed upon; and the resulting variation in archiving practices and language representation would seriously inhibit data access, searching, and cross-linguistic comparison. Second, standards may be set without guidance from descriptive linguists, the people who best know the range of structural possibilities in human language. If linguistic archives are to offer the widest possible access to the data and provide it in a maximally useful form, consensus must be reached about certain aspects of archive infrastructure.

The primary goal of **E-MELD** is to promote this consensus.

School of Best Practices

Endangered Languages

Search for Information on an Endangered Language:

- ☐ Find Field Linguists
- ☐ Find Documentation

Choose a language

Browse or Search for Endangered and Extinct Languages

Search All Languages

Search OLAC Archives

Search for EL Documentation

Other searches

E-MELD Features

**EL resources created using office software
stored in a proprietary format
will be unsupported in 5-10 years**

**We need new open source tools
based on agreed-upon data models and
connected to portable data formats,
with all EL data housed in a network of
interoperative digital archives**

**Bird, Steven, and Gary Simons (2003). Seven Dimensions of Portability for
Language Documentation and Description. *Language* 79:557-582.**

In Canada and elsewhere, linguists have contributed to language teaching materials or to documentation such as dictionaries or texts for the communities within which they work. Others have assisted communities in developing language programs aimed at language revitalization, or have acted as consultants for studies on place names or ethnobotany.

Community-Based Language Research assumes that knowledge can and should be constructed for, with, and by community members, and that it is therefore not merely (or primarily) for or by linguists. In a model of this kind, linguists are not the sole researchers, nor are they necessarily the leading researchers; rather, they are partners working in a collaborative relationship with members of the language-using community.

Ewa Czaykowska-Higgins (2009)



Secwepemc Directions Phrase Book

Words

[List Alphabetically](#)
[List by Category](#)
[Search](#)

Phrases

[List by Phrase book](#)
[Search](#)

Songs
Stories

[Return to Portal Page](#)
[Return to Welcome Page](#)

Secwepemc

English

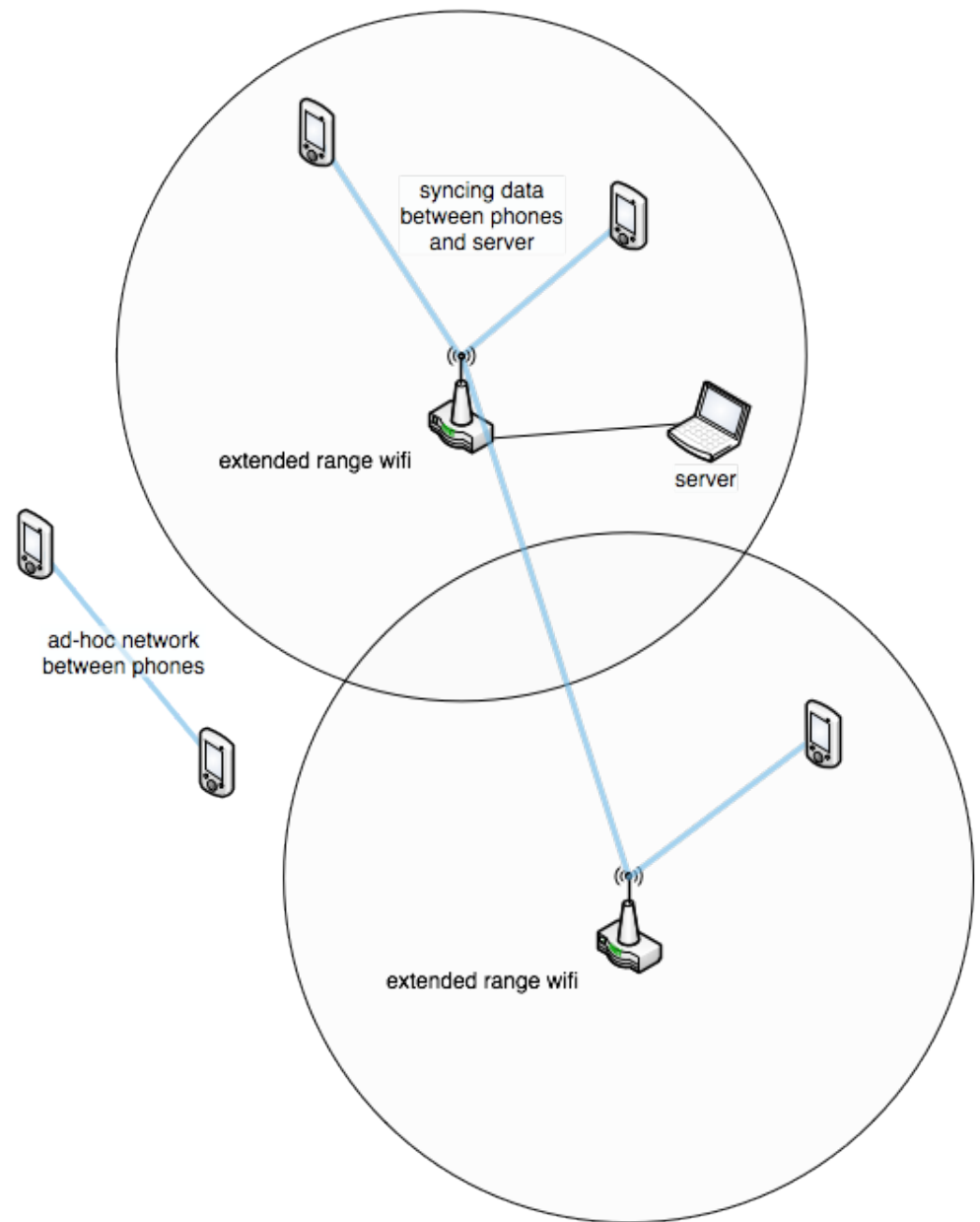
Ta7 k seckítscs ne s7í7illcw m-tspelqílcwes.	He couldn't catch up with the others, so he came back.
Tá7us ke7 síxelc tli7.	Don't move from that spot.
Tá7us ke7 splep.	Don't get lost.
Kectsétsme!	Give it to me!
Kekéwes k stseqmínc re tqweméke7.	Throw the ball far.
Kekéw m-witsínes te m-ste7ék.	They could be heard from afar as they approached.
Kekík7et ne newí7s re púltes re sqéxes.	Her dog lay close by her.
Melámen ri7 te séwllkwe.	It is a medicine lake.
Melcentéke re tqweméke7 tken7élye.	Kick the ball toward this way.
Yelpélc-ucw tken sts'kwe7ékst e pték-ucw te nuxwléws.	Turn left after the train tracks.
Nenús k tentécwes.	Put it over there.
Peqétkwe re stse7mét.stem tek setétkwe.	That lake is called White Lake.
Nerí7es k emút-ucw.	Sit right there.
Nerí7 le m-tntésés re spēke7s.	That was where he put his gloves.
Yerí7 le yegímkes ne stéckens re xpé7e.	It was grandfather's tether for his backpack.
Nerí7 wes re mútes kwemtús.	It always sits here.
Ne7élyus k tentécwes te ts'elcwílep.	Put the chair there.
Me7 t7ek-kt ne nexléwsten.	We will cross over the bridge.
Me7 tspelqíqelc-ken e tá7us e spenmímen.	I will return if I don't find it.
Me7 t7ek-kt tken7élye.	We will go this way.

Basic Oral Language Documentation

BOLD (PNG)

- ➔ **distribute cheap recorders**
- ➔ **collect data**
- ➔ **annotate orally**
- ➔ **archive and disseminate**







Endangered Archives Project: Digitizing the Buea Archives in western Cameroon

**Pierpaolo Di Carlo
Jeff Good**

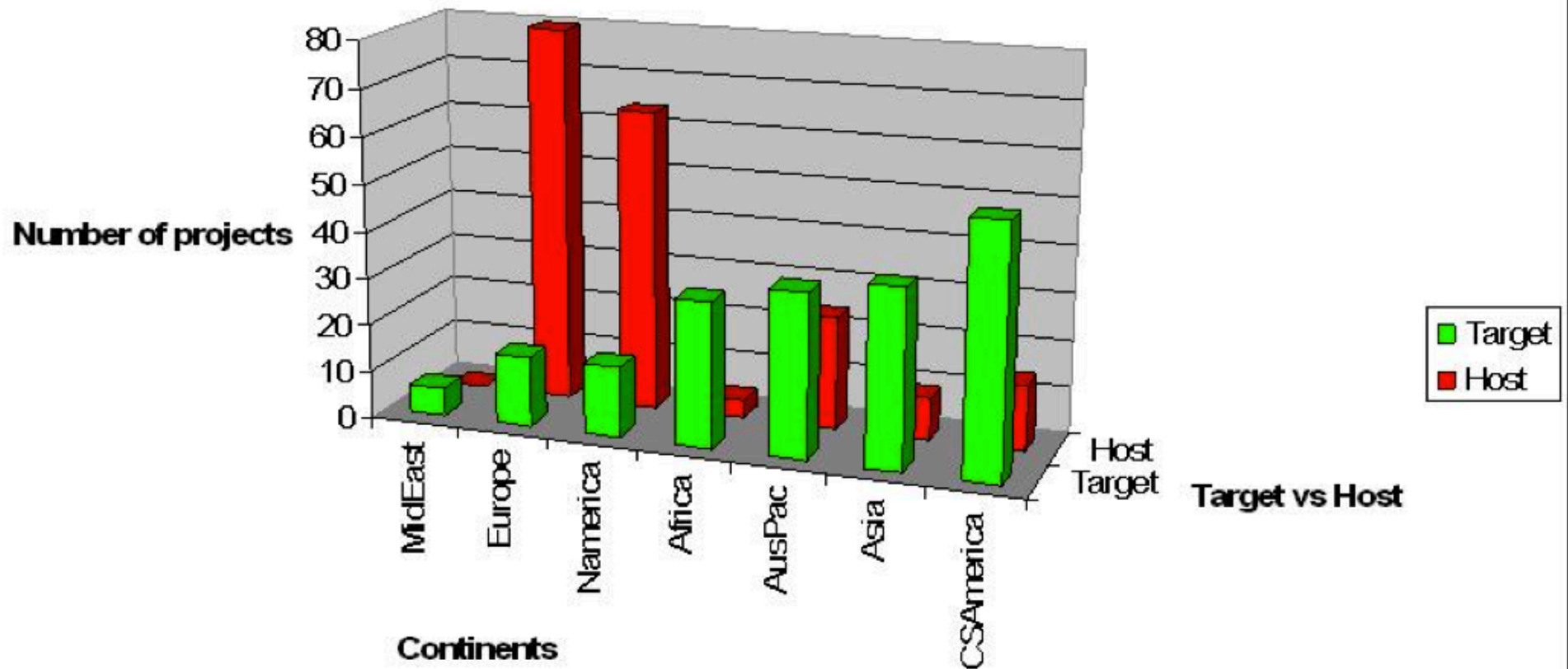


**Low-caste musician
Tashi Tsering plays
traditional music in
Lo Monthang,
Mustang, Nepal**

Katey Blumenthal



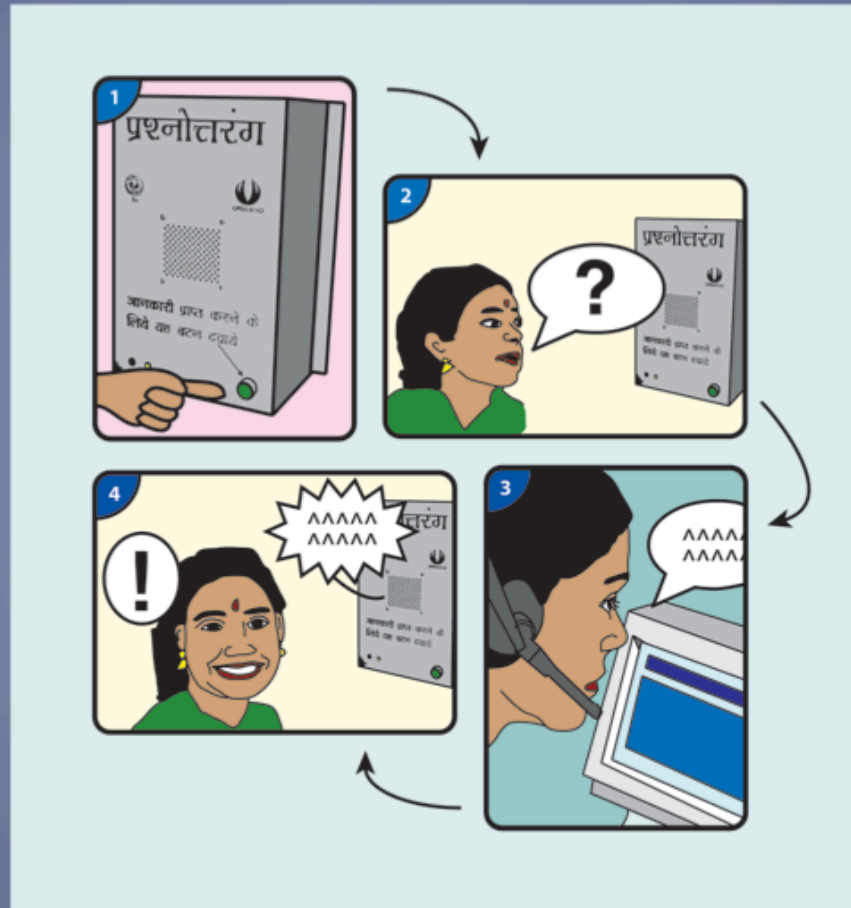
ELDP grants by continent



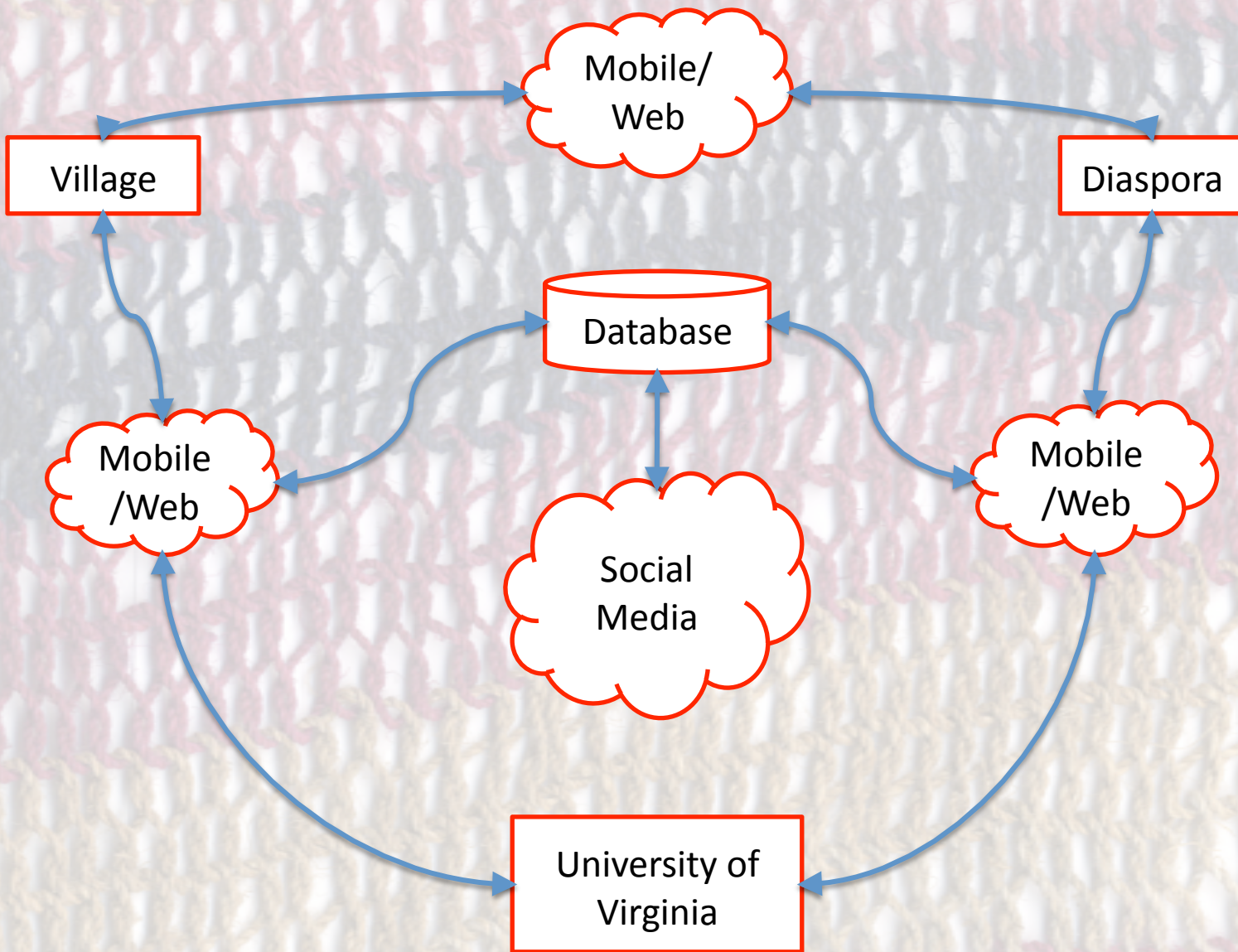
http://blogs.usyd.edu.au/elac/2010/09/where_does_the_dosh_go_peter_k.html



How Question Box Works:



Rose Shuman, Open Mind



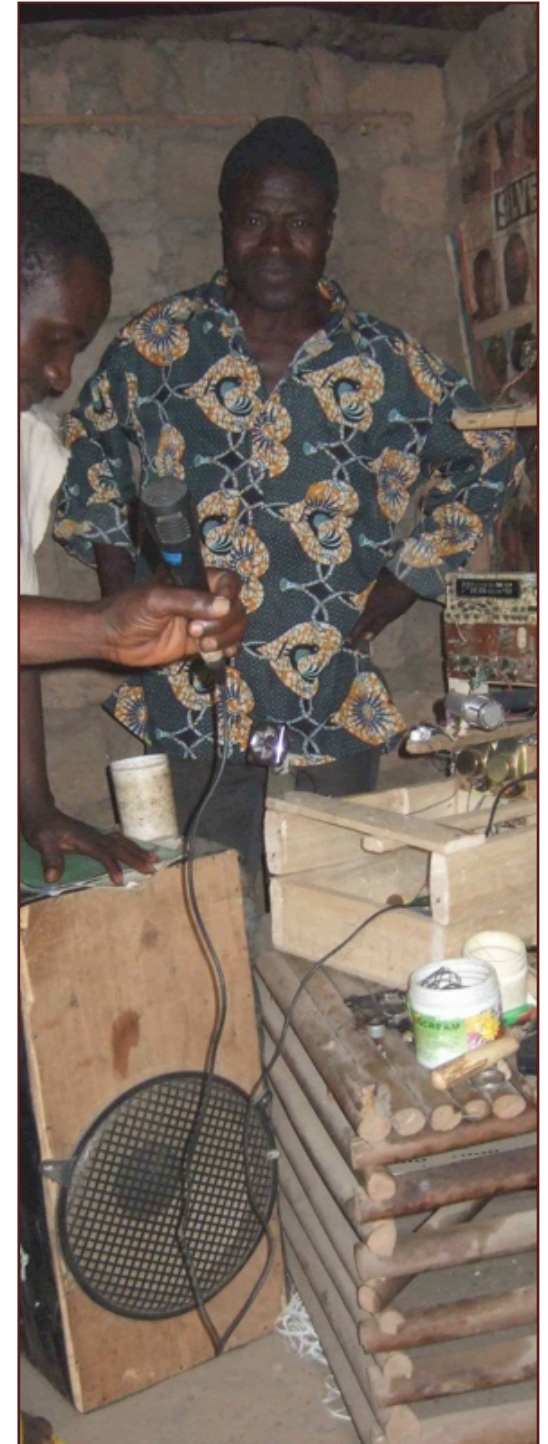
We have to be realistic about how technology is used locally, and not just project our hopes for how it could be used, or feed a community's hopes for how it could solve their problems.







- small, local radio station
- relatively cheap and can serve the whole area
- radio broadcasts would give the local languages an additional sphere of usage
- Standardization not necessary
- one linguistic moment is not locked in as 'definitive'
- how can an archive support this?



The interests of
those in the
village...





are not the same as those of people living in town.

A photograph of a tropical village scene. In the foreground, a dirt road curves from the bottom center towards the right. To the left of the road, there are several small huts with thatched roofs, surrounded by lush greenery and tall palm trees. To the right of the road, there are more trees, including some with bare branches. The sky is overcast with grey clouds. The text "Thank you" is written in a black, handwritten-style font across the upper middle part of the image.

Thank you

dobrin@virginia.edu