LSA Archiving Tutorial
January 2005

Archives, linguists, and language speakers

THE HANS RAUSING
Endangered Languages Project
Because every last word means
another lost world...

Credits

Peter K. Austin, Endangered Languages Academic Programme, School of Oriental and African Studies (SOAS)

Robert Munro, Endangered Languages Archive, SOAS

David Nathan, Endangered Languages Archive, SOAS

www.hrelp.org

THE HANS RAUSING
Endangered Languages Project
*Because every last word means
another lost world...*

Overview

- ❖ What is archiving of language materials?
- ❖ What kinds of language archives are there?
- ❖ Who uses language archives?
- ❖ Why should you work together with a language archive?
- ❖ What can a language archive offer you?
- ❖ What can you archive and how do you archive your materials?
- ❖ Archives Networks and resources

THE HANS RAUSING
Endangered Languages Project
Because every last word means
another lost world...

What is a archiving of language materials?

❖ Preparing materials in a structured form for long-term preservation

❖ Archiving is *not* backup

❖ Archiving is *not* publication

❖ Beware: "I'll archive my data on the web"

❖ Archiving pays attention to intellectual property rights, and cultural, social, religious and individual restrictions - access and distribution are *separate from* preservation

THE HANS RAUSING
Endangered Languages Project
*Because every last word means
another lost world...*

What is a language archive?

❖ A trusted repository created and maintained by an institution with a commitment to permanence and the long-term preservation of archived resources

❖ Will have clear policies for materials acquisition, cataloguing, dissemination, quality assurance, forward migration to new digital formats, disaster recovery

THE HANS RAUSING
Endangered Languages Project
*Because every last word means
another lost world...*

Kinds of language archives

❖ Many cross-cutting classifications:

❖ Indigenous vs outsider, eg. Squamish Nation

❖ Regional vs international, eg. AILLA, Paradisec; DoBeS, ELAR

❖ Associated with research institute, eg. AIATSIS, UCB California Survey, ANLC

❖ Granter-funded, eg. DoBeS, ELAR, OTA

❖ Physical vs digital vs mixed, eg. DoBeS vs Vienna Sound Archive, British Library

THE HANS RAUSING
Endangered Languages Project
*Because every last word means
another lost world...*

Who uses language archives?

- ❖ Speakers and their descendants - 95% of users of AIATSIS and UCB are community members

- ❖ Individual linguists or teams (depositors) - to create or renew materials

- ❖ Other researchers - comparative/historical linguists, typologists, theoreticians, anthropologists, historians, musicologists etc etc

- ❖ Journalists - wanting copy and images, sounds

- ❖ General public - interested in exotic languages

THE HANS RAUSING
Endangered Languages Project
*Because every last word means
another lost world...*

Why work with a language archive?

- ❖ To ensure long-term preservation of your work
- ❖ To ensure long-term accessibility to the data
- ❖ For the potential benefit of the language community now and in the future
- ❖ For safekeeping your own work and establishing your intellectual property rights
- ❖ To make your data available to other interested people

THE HANS RAUSING
Endangered Languages Project
*Because every last word means
another lost world...*

What can a language archive offer you?

- ❖ *Security* - keep your electronic materials safe
- ❖ *Preservation* - store your materials for the long term
- ❖ *Discovery* - help others to find out about your materials
- ❖ *Protocols* - respect and implement any sensitivities or restrictions
- ❖ *Sharing* - share the results of your work with others, if appropriate
- ❖ *Acknowledgement* - create citable acknowledgement of your research
- ❖ *Mobilisation* - help to create usable language materials for communities
- ❖ *Quality and standards* - advice for assuring your materials are of the highest quality and robust standards

What can you archive?

- ❖ Any linguistic materials, including:
- ❖ Media - sound and video
- ❖ Graphics - pictures and images
- ❖ Documents, including fieldnotes, grammars, description and analysis
- ❖ Structured data, eg. time-aligned annotated transcriptions, databases
- ❖ Metadata, including structured data about your materials, typically in written text form

THE HANS RAUSING
Endangered Languages Project
*Because every last word means
another lost world...*

How can you archive?

❖ Contact a relevant archivist *before* you begin your project and ask for advice - if they can't help you they will suggest someone who can

❖ The archive will want you to prepare your materials in an appropriate informative and explicit format (encoded and stored in a particular way), record relevant metadata, and specify intellectual property and protocol issues including conditions on access and use

THE HANS RAUSING
Endangered Languages Project
*Because every last word means
another lost world...*

Archives Networks

- ❖ Several networks of archives exist:
- ❖ Digital Endangered Languages and Archives Network (DELAMAN) - see www.delaman.org
- ❖ Open Language Archives Community (OLAC) - see Gary Simons presentation to follow
- ❖ Archivists are nice friendly people and working together with them from the beginning of your project can have all sorts of benefits
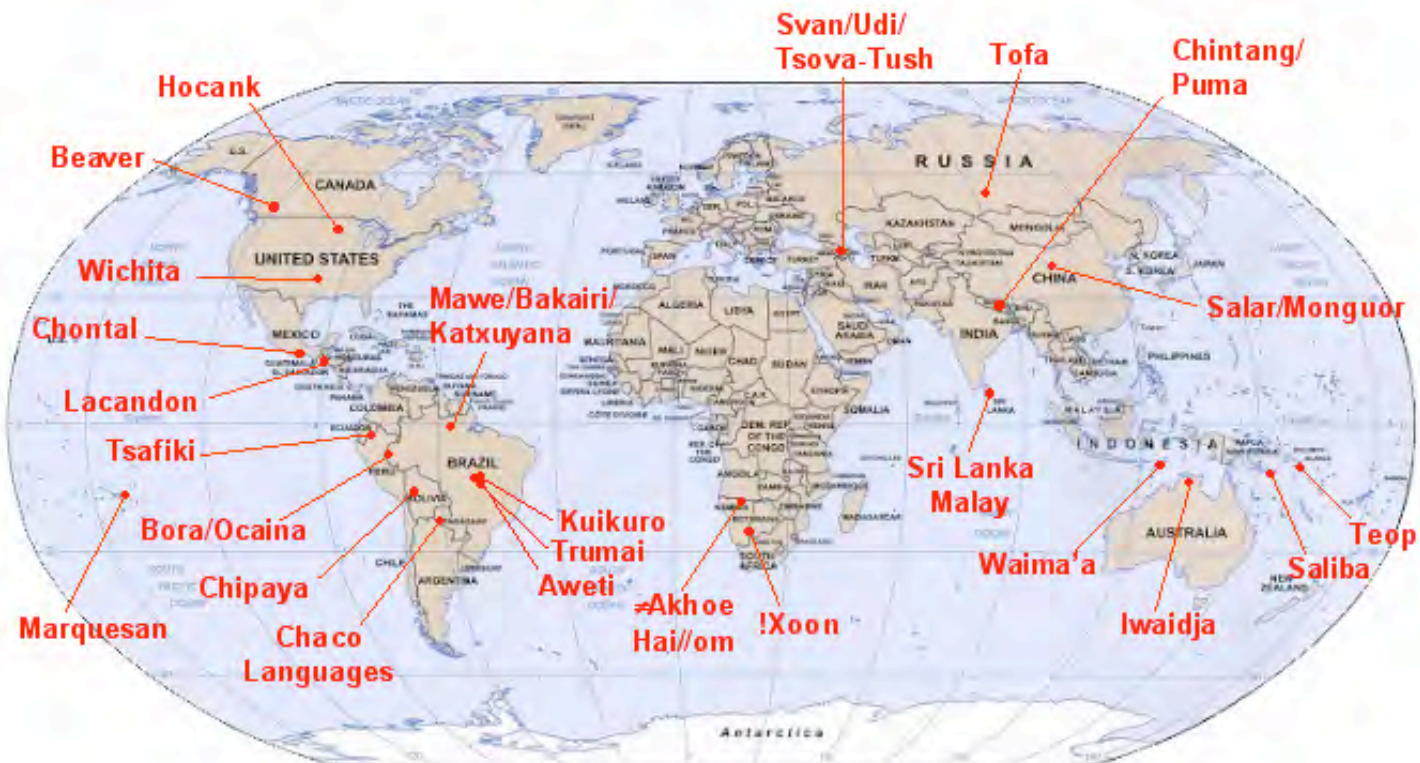
Any questions or comments?

THE HANS RAUSING
Endangered Languages Project
*Because every last word means
another lost world...*

DoBeS - Volkswagen Foundation

THE HANS RAUSING
Endangered Languages Project
Because every last word means
another lost world...

SOAS - ELAR