# National Science Foundation
## Workshop for
# Documenting Endangered Languages



October 14 - 17, 2007

The New England Center,

University of New Hampshire, Durham, NH

# BABEL –

## Better Analyses Based on Endangered Languages

Peter Austin
HRELP

- framework offered by the European Science Foundation (ESF) to promote and stimulate European collaborative research of the highest quality, innovative and investigator driven
- BABEL is one of 6 themes chosen in 2007 for development (only Humanities theme). If funded it will support:
  - salaries (PhDs, post-docs, senior scientists, technicians, etc.)
  - equipment and consumables
  - publications and general dissemination
  - access to large infrastructures
  - participation in conferences and other travel
  - networking (workshops, conferences, summer schools etc.)

Overview

- recent concern for language endangerment means linguistics has started to take language diversity properly into account

- dramatic change in amount and nature of language data that is being analysed has had and will continue to have a profound influence on our insights into the human language faculty

- need for storage of data for generations to come has triggered a reflection on data collection, representation and storage

- shift towards balanced corpora containing natural discourse and text data along with judicious use of stimuli, questionnaires and elicitation
- strengthened bottom-up approach to generalisations about linguistic behaviour encouraging linguists to let the data speak
- endangered languages research provides us with new sounds, new conceptual structures, new categories, counter-examples to hypothesized universals, and also shifts the balance of what we see as common or rare in the languages of the world.
- crucial difference of BABEL from existing documentation initiatives: bringing newly gathered data into development of linguistic theory

1. what is relationship between linguistic theory (including typology and formal theories) and the data and analyses arising from documentation of endangered languages?

   – how can theory inform documentation?

   – how can documentation inform theory?

- *example*: 1. theory of noun incorporation crucially built upon analysis of data from endangered languages (Mark Baker 2007)

- *example*: 2. theoretical and typological understanding of causative and applicative constructions (adding arguments to verbs) informs work on describing individual languages

2. documentation wears **interdisciplinarity** on its sleeve (eg. involvement of anthropologists, biologists, geographers, IT specialists in projects) but can this be achieved and if so how is it best conceptualised?

3. can we develop **multimedia interfaces** to rich language documentations to serve the needs of different user communities and if so what might these look like?

4. what metrics can be developed to determine the **quality** of language documentation outcomes, going beyond current concerns for archives and technical data standards?

5.  how can we change research practices in linguistics through collaboration and training?

6.  what impact does work with endangered languages communities have on linguistics from a moral and ethical philosophy perspective? Can we look beyond politically crafted guidelines and statements to explore the impact on traditions of research practices and institutional frameworks?

Proposal developers

- Peter Austin, Hans Rausing Endangered Languages Project, SOAS, UK
- Michael Fortescue, University of Copenhagen, Denmark
- Nikolaus Himmelmann, Ruhr-Universität Bochum, Germany
- François Jacquesson, LACITO, CNRS, France
- Maarten Mous, Leiden University, Netherlands
- Mauro Tosco, L'Orientale, Naples, Italy