“Technologically challenged languages”

Justus Roux

Joint COCOSDA & ICCWLRE Workshop
LREC 2004
Lisbon
• In principle: All languages are `technologically challenged`.

• Some languages are more `technologically challenged` than others.

• Matter of scale and environment.
• Number of factors determining whether a language can be regarded as `technologically challenged`:

- Current phase of development
  • Some languages have a longer `technological history` than others
  • Availability of resources & expertise
– Degrees of development

• Some languages may still be regarded as technologically challenged if they do not have full-fledged Speech-to-Speech translation capabilities – others if they do not have electronic dictionaries etc
– The need for being `developed``
  - Ideal (academic)
  - Reality
    – Can we really expect all languages to be `technologically developed`?
    – Cost factor
    – Level of available expertise (linguistic and technological) in a particular context
      » Roadmap implies availability of expertise – not necessarily the case
      » Question: Human capacity building as a goal in the roadmap strategy? (Integrated training programmes)
• Cut off point for development of “technologically challenged languages”? 

  Example: Case of SA English
Language Situation
Mother tongue division (n=40.5 mil speakers)

- Zulu 22%
- Xhosa 18%
- Afrikaans 16%
- N Sotho 10%
- English 9%
- Tswana 7%
- Swati 3%
- Tsonga 4%
- Venda 2%
- S Sotho 7%
- Ndebele 2%
• What priorities should be set in developing languages at technological level?
  – Economic factors?
    • Commercial viability?
  – Socio-political factors?
    • Access to information
      – Eg. Language specific speech based systems empowering pre-literate communities? (Given the growth of mobile communications in Africa.)