OLPC Localization Strategy Proposal

Edward Cherlin
Earth Treasury

Author

- Peace Corps, Korea

- Buddhist seminary training

- High-tech market analyst

- Unicode contributor

- Free Software developer and manager

- API tech writer

- Simputer Evangelist

Language Support Issues

- Localization of strings in software
- Translations of Wiki, textbooks, content
- Dictionaries
- Keyboards, IMEs, fonts, rendering
- Text-to-speech
- Speech recognition

Proceeding Normally

- Spelling dictionaries
- Keyboards, IMEs, fonts, rendering
- Text-to-speech
- Speech recognition

Not Started

- Adding support for needed languages in text-to-speech

Pootle lists more than 60 language projects

Afrikaans, Amharic, Arabic, Aymara, Basque, Bengali, Bengali (India), Bulgarian, Catalan, Chinese (China), Chinese (Hong Kong), Chinese (Taiwan), Czech, Danish, Dari, Dutch, Dzongkha, English, English (South African), English (US), Finnish, French, Friulian, Fula, Galician, Georgian, German, Greek, Gujarati, Hausa, Hindi, Icelandic, Igbo, Italian, Japanese, Khmer, Kinyarwanda, Korean, Kreyol, Macedonian, Malayalam, Maltese, Marathi, Mongolian, Nepali, Pashto, Persian, Polish, Portuguese, Portuguese (Brazil), Punjabi, Quechua, Romanian, Russian, Serbian, Sinhala, Slovenian, Sotho, Spanish, Swedish, Tamil, Telugu, Templates, Thai, Turkish, Ukrainian, Urdu, Vietnamese, Wolof, Yoruba

European languages

Basque, Bulgarian, Catalan, Czech, Danish, Dutch, English, Finnish, French, Friulian, Galician, German, Greek, Icelandic, Italian, Maltese, Macedonian, Polish, Portuguese, Russian, Serbian, Slovenian, Spanish, Swedish, Ukrainian

Other languages

- Chinese (Taiwan)

- Georgian

- Japanese

- Korean

- Turkish

Languages of target regions

- North Asia: Mongolian

- India: Bengali (India), Gujarati, Hindi, Malayalam, Marathi, Punjabi, Tamil, Telugu, English (Oriya and others are missing.)

- South Asia: Dari, Dzongkha, Nepali, Pashto, Sinhala

- Southeast Asia: Bengali, Khmer, Thai, Vietnamese, French

- East Asia: Chinese (China), Chinese (Hong Kong), English

- Africa: Afrikaans, Amharic, Arabic, Fula (Fulfulde), Hausa, Igbo, Kinyarwanda, Sotho, Wolof, Yoruba

- Caribbean: Kreyol

- Latin America: Spanish, Portuguese (Brazil), Quechua

Target Countries: Purchases

- Uruguay

- Peru

- US (Birmingham, Alabama)

Target countries & languages: G1G1

- Haiti (Kreyól Ayisyen, French)

- Rwanda (Kinyarwanda)

- Ethiopia (Amharic, Oromo, Tigrinya, Afar, Gamo-Gofa-Dawro, Hadiyya, Sidamo, Somali, Wolaytta)

- Cambodia (Khmer)

- Afghanistan (Dari, Pashto, Hazaragi, Aimaq)

- Mongolia (Mongolian: Cyrillic, Traditional)

- Nepal (Nepali, Bhojpuri, Maithili)

Nigerian national languages

- Edo

- Efik

- Adamawa Fulfulde

- Hausa

- Idoma

- Igbo

- Central Kanuri

- Yoruba

- English

Other major Nigerian languages

- Anaang
- Nigerian Fulfulde
- Izon
- Tiv

Largest language communities

- 347 languages with more than a million speakers (Ethnologue)

- 93.8% of the world's population

- That leaves us 280 more to do.

- I have no way to determine how many should be current priorities.

Language Infrastructure

- Pootle for localization

- Git for keyboards, etc.

- Nothing for content library

- Nothing for tracking licenses and permissions

- Wiki translations just go on the Wiki, but there is no visible mechanism for synchronizing.

How do we decide what's next?

- **Little usable public information on which countries to target.**

Green pilot countries

- Ethiopia G1G1
- Rwanda G1G1
- Libya MOU, no purchases
- Nigeria Pilot, no purchases
- USA one city
- Argentina Pilot, no purchases
- Brazil trying to buy
- Paraguay No
- Peru OK

- Uruguay OK
- Mongolia G1G1
- Thailand No
- India Turned it down
- Nepal OK
- Pakistan ?
- Romania Turned it down
- Russia No

How do we decide what's next?

- Little usable public information on which countries to target.

- **Within target countries, we have no public information on which language communities to target.**

How do we decide what's next?

- Little usable public information on which countries to target.

- Within target countries, we have no public information on which language communities to target.

- *There is no visible planning or outreach to language communities.*

Starting a localization project I

- Research
  - country
  - XO distribution plan
  - languages
  - government

  - NGOs
  - Linux distro?
  - LUGs
  - literature
  - textbooks

Starting a localization project II

- Recruiting

- Training

- Glossary

- Coordination

- Documentation

What's missing

- Information
- Communication
- Management
- Infrastructure
- More volunteers
- More paid staff

Job reqs

- Localization coordinator and recruiter

- Doc lead, several tech writers, and writers who understand practical Constructionist education.

- Yes, we need the CEO and COO. They need to understand the tech, the business, the mission, and above all how to communicate, including how to listen.

Job reqs

- Localization coordinator and recruiter

- Doc lead, several tech writers, and writers who understand practical Constructionist education.

- Yes, we need the CEO and COO. They need to understand the tech, the business, the mission, and above all how to communicate, including how to listen.

I'm listening.

Questions and suggestions, please.