Facebook i18n Workshop

Kevin Scannell
Saint Louis University
21 May 2013
Trendsetting l10n efforts

- What are some of the most widely used pieces of software today?
- Some good news: many of these are among the most widely localized
  - Google search: 152 locales, 143 languages
  - Facebook: 105 locales, 94 languages
  - Firefox: 94 locales, 87 languages
  - Microsoft: 106 (LIPs)
Facebook l10n

- Crowdsourced translation (mostly)
- Some in-house translators
- Smaller languages 100% volunteer
- Spanish was first non-English locale, Jan. 2008
- Minority languages: Irish, Welsh, Northern Sámi, Basque, Faroese, Frisian, Rumantsch
- Indigenous languages: Cherokee, Aymara
I18n Workshop

- See http://borel.slu.edu/pub/fb-l10n.html
- Imagine you work for Facebook's Int'l Team
- How would you present these strings to localizers for translation?
Why do this?

- Important to understand i18n and l10n
- See just how broken widely used systems are!
- At the same time see how hard the issues are
- Crowd-sourced translations in particular
- Understand the obstacles other languages face
- No settling!
- Motivate potential solutions, e.g. Mozilla's l20n
Facebook's Translation App

- About 100 languages supported
- Innovative “inline” translation platform
- Quality control via “voting”
- Relatively large number of strings
- 67518 translations, 82623 votes by top 500
- 74% complete :/
Software as a Service

- Will discuss open source, esp. for Irish
- Has resulted in a rich set of (free) offerings
- Looming danger: Gmail, Google Docs, etc.
- Also Facebook, Twitter, etc.
- Losing control of the online linguistic landscape
The next 6900 languages

- Facebook: no new languages in 2+ years
- Google's GIYL similarly shuttered
- Among minority languages, piecemeal offerings
- Irish is “lucky”: FB, Twitter (though not Gmail)
- Scottish Gaelic, none of the above
Options?

- Pressure from each individual language group?
- Legislation, rights-based approaches?
- Develop your own social network (Irish e.g. 2x!)
- Don't use, or let your kids use, Facebook, etc.?
Workaround

- An idea of Neskie Manuel, for Secwepemctsín
- Use Greasemonkey (JavaScript plugins)
- Creates an “overlay” in the user's browser
- Doesn't require Facebook's permission
- Carrying on project in Neskie's memory
- Generalized to work with any language
- Vow to manage all linguistic complexities
Unofficial localizations

- Tremendous response, 60+ locales underway
- http://goo.gl/WEj1c
- Majority have never localized software before
- Majority are endangered according to UNESCO
- Capped number of messages at 200
- Each locale can choose a “default” language
Terminology Wiki

- Most languages have had to coin new terms
- Technical terms ("mobile upload", "cookie")
- Facebook-specific ("like", "poke", "news feed")
- Western concepts ("privacy", "advertising", etc.)
- Please add your own!!
Sharing metaphors

- “Computer” (sorter, lightning brain)
- “Link” (bridge)
- “Domain” (secret initiation camp)
- “Cookie” (cuckoo's egg)
- “Cache” (leftovers)
- “Keyboard” (fingerboard vs. musical instrument)
- “De-interlacing” (de-cross-weaving)
Examples from the Wiki

- “Activity Log”
- “Chat”
- “Change Cover”
- “See Friendship”
- “Unlike”