Facebook in all six Celtic Languages (mar dhea)

Kevin Scannell
Saint Louis University
May 31, 2012

Social media and endangered languages

- Social media: Facebook, Twitter, blogs, podcasts (audio and video), "content communities"

- 1000's of speakers of 100's of endangered languages are using their languages in social media

- Indigenous Tweets and Indigenous Blogs

- Connects scattered speaker populations, diaspora

- Informal register; use the bits of language you know

- Make up for lack of "traditional" media

- Appealing to kids, the most important demographic!

Facebook for language revitalization

- I spent the past year working with Twitter helping improve support for small languages

- Both Twitter and FB have full Unicode support; create content in virtually any written language

- Facebook has some advantages for revitalization:

- Better support for multimedia

- No 140 character limit

- Clearer "threading" of conversations

- Everyone is already on it!

The Celtic Languages on Facebook

- Thousands of learners, native speakers, and orgs use Facebook daily to share news in all 6 Celtic langs

- "Groups" aimed specifically at languages: "Facebook e brezhoneg", "Siarad Cymraeg", "Keskowethyans an Taves Kernewek", "Úsáidim Facebøok as Gaeilge" (sic), "Gàidhlig na h-Alba", "COWAG", …. many more!

- Reasonably complete translations of the site (links, navigation, help) into Irish and Welsh; mixed quality

Facebook's Translation Platform

- Translated entirely by users of the site; volunteers

- Voting system to vote down bad translations, vote up good ones (amazingly, it seems to work)

- Ability to translate the site "live", while using Facebook as you normally would

- Sophisticated terminology support, quality checks

- About 100 languages available, including Irish, Welsh, Guaraní, Aymara, Cherokee, N. Sámi, Basque, Frisian, Quechua, Xhosa, Rumantsch, ...

The Problem

- Facebook have not added any new languages for quite a while, despite petitions and much agitation

- Same situation with Google's search interface

- Have heard secondhand about issues with spam

- At the end of the day, neither Facebook nor Google really cares about minority languages, they want to make money

- Upshot: they have great services, but it's a bad idea to rely on the goodwill of any for-profit company in devising a technology strategy for your language

The Solution, I

- My friend Neskie Manuel came up with a way to take Facebook out of the equation and translate the site into his language of Secwepemctsín

- Neskie passed away tragically last May; I'm carrying this work forward in his honor

- Uses a technology called "Greasemonkey"

- Users can install a bit of JavaScript that "overlays" translations of any site, entirely in their browser

- Game changing idea! No need to rely on Facebook or Google to "allow" us to translate their sites

The Solution, II

- I rewrote Neskie's code to work with Facebook's new design, and generalized to all languages

- To get Facebook in your language, you just have to send me translations of about 150 short phrases

- Overwhelming response: nearly complete translations into 21 languages; 30+ more coming

- Installs with one click in Google Chrome

- Can set a reasonable "default language" for untranslated strings (e.g. French for Breton, Spanish for Nawat, Russian for Chechen)

Terminology development

- Part of my aim was to reach out to language groups that have never translated software or websites

- Terminology can be a major problem; one has to make up everything! (GAELIC-L in the early 90's)

- I asked each translator to also provide "back translations" into English, for the benefit of others looking for the right term or metaphor

- Great opportunity for endangered language groups to share "different ways of thinking and knowing"

Back Translations Wiki

- These are all available on a wiki that new translators can consult; e.g. under the term "Activity Log":

- French: "Personal history"

- Gundjeihmi: "What have you been doing"

- Irish: "Activity history"

- Nawat: "See what has happened"

- Chichewa: "Frequent activities"

Gàidhlig (Michael Bauer)

Gaelg (Adrian Cain, Chris Sheard)

Breton (Fulup Jakez)

Cornish (coming soon...)