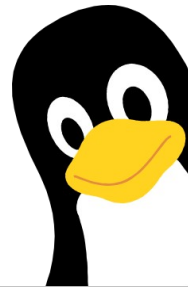


Globalization in Free Software Communities



@JobavaL10n



What?



“It's about the cultural adaptation of software”

Jeff Beatty – Localization head at Mozilla

We aren't talking about mere translation, but the cultural adaptation of software. This involves not just language, but also local customs and country. For example, Spanish is spoken in Latin America, but each country has a distinct dialect and even different styles of address and expression. For that reason, Firefox is localized not into Spanish, but Spain-Spanish, Argentinian, Chilean, Mexican etc.

What?

i18n, L10n, g11n

Here I ask the audience if they know what the three terms displayed are.

I also ask the audience if they've been involved directly or indirectly in any stage of a globalization (internationalization, localization).

What?

i18n -> internationalization

What?

i18n -> internationalization

L10n -> localization

What?

i18n -> internationalization

L10n -> localization

i18n + L10n -> **globalization**

Why?

- globally-relevant applications need to be in the user's language
- small % of audience actually comfortable in English
- largest growth markets not English-speaking
- a chance to build a community around your project
- **it's easy!** (but can be tricky)

Tools

Pootle (2004) -> django + redis + SQL

- Mozilla Firefox, websites, Libreoffice
- Includes translation memory
- Older and more mature tool



Tools

Pontoon (2011) -> django + postgres + heroku

- Websites, Mozilla Firefox (Slovenia)
- Uses external TM Amagama
- Notable feature: in-website localization



Before we dive in

Languages have different plural forms (cardinals)

- English: 2 (singular, plural)
- Romanian: 3 (singular, 2 plurals)
- Japanese, Chinese: 1 form
- Russian and most Slavic languages: 4 forms
- Gaelic languages: 4 or 5 forms
- Arabic: 6 forms

Wrong ASSUMPTIONS

اور بازار سے لے آئے اگر ٹوٹ گیا

ساغرِ جِحم سے میرا جامِ سِفال اچھا ہے

“I can get another if I break it / so a clay cup trumps a grail”

(Mirza Ghalib)

Every language is written left-to-right

Text in the image is from a poem by Urdu poet Mirza Ghalib. It means “I can get another if I break it / so a clay cup trumps a grail”

Wrong ASSUMPTIONS

;

?

“ ” „ » « 「 」

Every language has the same punctuation

The sign on the top left is a Greek question mark
The sign on the right is not a question mark

Wrong ASSUMPTIONS

| | | | | | | | | | | |
|----------------------|---|---|---|---|---|---|---|---|---|---|
| Western Arabic | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Eastern Arabic | ٠ | ١ | ٢ | ٣ | ٤ | ٥ | ٦ | ٧ | ٨ | ٩ |
| Perso-Arabic variant | ۰ | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ | ۷ | ۸ | ۹ |
| Urdu variant | ۰ | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ | ۷ | ۸ | ۹ |

Every language has the same numeric symbols

Some cultures have traditional style numerals. Those depicted here are also practically applicable since they also follow the RTL text orientation of their language, so text containing traditional numerals will be consistent as far as orientation.

Start with

A flexible UX

- make sure strings **2-3 times as long** fit into the interface (many languages longer than English)
 - Romanian: 20% on average longer
 - Italian, Russian: 25%
 - French: 32%
 - Chinese, Japanese, Korean: -50%

The average figures matter less than individual string variations which can be longer.

Continue with

A glossary of terms

- app-specific terms
- generic technical terms that need to stay consistent
- trademarks and other proper names

The app-specific terms and trademarks should involve the developer or product side people.

Know also

Localization is not translation

- Clarity > strict accuracy
- Space considerations
- Idioms and pop culture don't translate well (**bring me a shrubbery!**)
- Localization is **adaptation** and **translation**

Clarity trumps strict translation accuracy. The purpose is not to have a good translation, but to have an interface that conveys the meaning and intent the developers had in mind. In that respect, sometimes we use English variants of technical terms when the audience understand what we mean. Sometimes the translation of a word from English produces a result that's homonymic and thus ambiguous (for example: the English words „core” and „kernel” both translate to the same Romanian word: „nucleu”).

How?

Typical workflow:

- Mark strings in source code i18n
- Extract strings into a localization format i18n
- Manage extracted strings (pref w/ a tool) L10n
- Translate the strings (pref w/ GUI tool) L10n
- Insert strings back into the application L10n

This describes the workflow from i18n to g11n.

The gl1n process

1. Mark strings in source code (gettext example)

```
printf("Hello World\n");  
window.set-title("Teh Awesome");
```



```
printf(gettext("Hello World\n"));  
window.set-title(gettext("Teh Awesome"));
```

The g11n process

2. Extract strings into a localization format

#comment to string 1

msgid "Hello World\n"

msgstr ""





#comment to string 2

msgid "Teh Awesome"

msgstr ""

The glIn process

3. Manage extracted strings

| Name ▾ | Progress ▾ | Total ▾ |
|--|------------------------|---------|
|  Firefox 47 (Aurora) | <div><div></div></div> | 9662450 |
|  Firefox for Android (Aurora) | <div><div></div></div> | 642218 |
|  Firefox for iOS 4.0 | <div><div></div></div> | 47792 |
|  Firefox Health Report | <div><div></div></div> | 60288 |
|  Firefox OS 2.5 | <div><div></div></div> | 1598508 |
|  Firefox OS (master) | <div><div></div></div> | 1718898 |
|  Lightning | <div><div></div></div> | 120750 |
|  Mozilla.org | <div><div></div></div> | 2773066 |

The g11n process

4. Localize (translate, adapt)

watch for: string length, consistency

each language should have style guide

```
msgid "Hello World\n"
```

```
msgstr "Witaj świecie\n"
```

The style guide is the localizer's job if one doesn't exist already. You can usually just build on top of the work done by similar projects out there, or software for that particular language.

What localizers need

- Terminology control (app-specific or like L10n.ro/Glosar)
- Good search functionality
- Translation memories (app-specific or like **Transvision**)
- Machine translation (eg. Google or Bing translate)
- General purpose dictionaries (eg. hallo.ro)
- Good UX, low technical barriers

I skip this slide if time is short so I can allocate time for live translation.

Make it easier on them

- **Context matters**
 - use whole sentences
 - group strings together semantically
- **Idioms and pop-culture translate badly**
 - don't be too punny
- **Give credit**
- **Swag** – stickers, t-shirts, pins, posters etc.

I skip this slide if time is short so I can allocate time for live translation.

Potential gotchas

- **RTL**
 - right-to-left layouts and bi-di text
- **Top-to-bottom**
 - Chinese, Japanese

【二〇〇六年七月四日訊】北京時間15:27 (UTC時間03:56:27)，在中國河北文安發生了一次地震，震中為北緯39度，東經116度（中國地震局數據），中國地震局測定震級為二級（美國地質測量局測定為二級）。

中國地震局隨後發表的消息說，北京、天津普遍有震感，河北、山西、山東等地的部分地區也有震感。目前暫無人員傷亡的報告。

中國總理溫家寶，副總理回良玉迅速做出了指示，要求地震部門加強監測、分析工作，瞭解震情，保證群眾生命財產安全，維護社會穩定。並且特別要求注意首都圈的情況。

I skip this slide if time is short so I can allocate time for live translation.

Pitch

To project leaders, decision makers and programmers

- Opportunity to create a community
- Application needs to be globally relevant
 - **highest-growth markets are non-English**
- More testers
- More users

I skip this slide if time is short so I can allocate time for live translation.

Pitch

To students and applicants

- **Gain experience** volunteering and working for open-source projects
- Few or no technical skills needed
- Be near **technically excellent people**

I skip this slide if time is short so I can allocate time for live translation.

Pitch

To free software and people in general

- Grow the ecosystem
- Contribute and get in touch with great communities
- Promote and impact your local language and culture
- Scratch an itch

I skip this slide if time is short so I can allocate time for live translation.

Conclusions

Not daunting if you know where to begin

Can be tricky depends on language and f/w

Can be messy working with people

Fun

Live Translation

FIREFOX

Libre Office

mozilla.org

GNOME

I like to dedicate at least 10 minutes on this stage, so will compress the presentation to fit this in.

It's important to have audience participation and coax a few of them to actually contribute a few translations.

It's even better if there are non-English, non-Romanian speakers in the audience, who can then contribute a suggestion to a different locale.

Questions?

Twitter: @JobavaL10n

Blog: medium.com/@Jobava/latest

Email: jobava@mozilla.ro

