OPENING THE BLACK BOX of Software Localization
PLAN OF TALK

- Why Localize?
- What could go wrong?
- First Principle of Localization
- Opening The Black Box
- Making Localization Work
- Tools for Agile and Continuous Delivery
- Discussion
WHY LOCALIZE?
Localization is the only way to reach global markets

- Seeking increased foreign markets share
- Multilingual is a requirement, not a nice to have
  - Compliance, regulations, mandates
  - User-generated content
  - Customer satisfaction and experience
QUICK FACTS

• 2012 Ditmo study: downloads for apps in foreign markets increased by 128 when translated into target language

• 2014 Common Sense Advisory Study finds 75% of global consumers would choose native targeted app over similar non-targeted; more than half would prefer poorly translated app to non-native

• 2016, Asia-Pacific region surpasses North and South America as leading rev generator in app sales

• John Yunker @ Global Design: Do businesses in the US ask "what's the return on investment for creating a website for my business?"
FIRST PRINCIPLE OF LOCALIZATION

• Separate Content from Code
• Benefit 1: protects code from corruption
• Benefit 2: opens the pool of human translation talent
WHAT COULD GO WRONG?

- Poor quality translation (missed context)
- Unexpected Costs (employee time suck)
- Missed deadlines
- Mangled code
- You build it, but they don’t come (SEO)
Translated App

Black Box

English App

Translated App
THE BLACK BOX
DEFINITIONS

**Translation**: replacing text in one language with text of equivalent meaning in another language.

**Localization (l10n)**: includes all of the changes needed to adapt content for a particular market, to take into account things like local time, date, and currency formats; cultural preferences about content; UI and design choices.

**Internationalization (i18n)**: Preparing software and websites to ensure a smooth and error-free localization process.
WHAT ARE THE CORE TECHNOLOGIES THAT PROFESSIONAL TRANSLATORS USE?

• The CAT Tool
• Translation Memory
• Term Base
• QA
• Pseudo Translation
COMPUTER ASSISTED TRANSLATION (UI)
LEADING CAT TOOLS

- Desktop vs. Cloud
- SDL Studio, MemoQ, SmartCAT, Wordfast
TRANSLATION MEMORY

• Stores prior translations by the “segment”
• Available to translator when processing translation
• Increases efficiency, reduces costs
• 100%, context, fuzzy matches
• Need to be editable, sharable,
• Meta data for memory management
• TMX Standard
• Impact of TM on translator’s process
ECONOMICS OF TRANSLATION MEMORY
TERM BASE

- Ensures consistent terminology
- Stores by single words and short phrases
- Available to translators in process
- Very useful in final stage QA
- Helpful for accurate translation of UI
Computer algorithms check for patterns that suggest error:

• Inconsistency in source
• Numeric mismatch
• Empty segment
• Missing tags
• Source and Target identical
### Error Description

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Source</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untranslated Segment</td>
<td>strings to translate xls.txt</td>
<td>Hoja1</td>
<td><strong>Untranslated</strong></td>
</tr>
<tr>
<td></td>
<td>strings to translate xls.txt</td>
<td>Hoja2</td>
<td><strong>Untranslated</strong></td>
</tr>
<tr>
<td></td>
<td>strings to translate xls.txt</td>
<td>Hoja3</td>
<td><strong>Untranslated</strong></td>
</tr>
<tr>
<td>Inconsistency in Target</td>
<td>strings to translate xls.txt</td>
<td>Support for bidirectional languages</td>
<td>Support para lenguas bidireccionales</td>
</tr>
<tr>
<td></td>
<td>strings to translate xls.txt</td>
<td>Support for bidirectional languages</td>
<td>Compatibility para lenguas bidireccionales</td>
</tr>
<tr>
<td>Inconsistency in Source</td>
<td>strings to translate xls.txt</td>
<td>Archive Log Backup</td>
<td>Copia de seguridad de la base de datos</td>
</tr>
<tr>
<td></td>
<td>strings to translate xls.txt</td>
<td>Database Backup</td>
<td>Copia de seguridad de la base de datos</td>
</tr>
<tr>
<td>Tag Mismatch</td>
<td>strings to translate xls.txt</td>
<td><code>&lt;p&gt;...&lt;br&gt;</code></td>
<td><code>&lt;p&gt;...&lt;br&gt;</code></td>
</tr>
<tr>
<td>Numeric Mismatch</td>
<td>strings to translate xls.txt</td>
<td>Tepe Handling - 7 Day</td>
<td>Manejo de cinta - 8 días</td>
</tr>
</tbody>
</table>

- 3: Inconsistency in Source
- Copia de seguridad de la base de datos
- Database Backup
- Database/Archive Log Backup
PSEUDO TRANSLATION

• Produces generated, fake translation for testing purposes
• Confirm character encoding OK
• See how UI does with text expansion, right to left
• Confirm all translatable content has been captured
• Prototyping workflow
ENSURE TRANSLATABLE TEXT IS CAPTURED
LOCALIZATION STEPS

- Pseudo-Translation
- Translation
- QA
- Pre-live testing
- Delivery
MAKING LOCALIZATION WORK
OVERVIEW

• Technical Considerations
  • Consider automation needs
  • CMS Choice/Localization Platform
  • Confirm Unicode Encoding
  • Best practices for handling textual content

• Content and Design Decisions
  • Simplifying Content
  • Text Expansion
  • Right-to-left
  • Color Choices
  • SEO/ASO

• Testing
  • Pseudo Localization
  • Pre-live Testing
AUTOMATION?

• How often do you release content updates?

• Are you on an agile or continuous delivery cycle?

• If updates occur more than a few times a year, strongly consider adopting a localization platform so that content changes are captured and pushed to translation teams automatically.

• (we’ll discuss localization platforms in a little more detail, later)
CMS

• How well i18n-ized?

• Wordpress, Drupal, Joomla, Any CMS, Squarespace, Wix

• **General rule of thumb:** Platforms that require less tech savvy generally present higher problems for localization

• Has to be responsive. Most foreign markets more mobile than US.

• Needs to export translatable content in some form of XML

• **Note:** We’ll talk about dev environments and localization platforms later.
Most modern CMS support Unicode (UTF-8), but make sure, and make sure encoding us maintained throughout the cycle of localization.
• Leave room for text expansion
• Keep translatable text together in resource files
• Don’t hard code international date, time, measurement
• Be careful with use of concatenated and dynamic text, as grammar and word order rules changes across locales
CONTENT AND DESIGN

• Simple and clear language in source
• One term, one meaning
• Avoid slang, references to local culture (unless necessary)
• Transcreation for marketing content and keywords
RIGHT TO LEFT (MIRRORMING)
DO THESE DRUGS MAKE HIM SICK OR VERY HEALTHY?
COLOR CHOICES
**SEO/ ASO**

- Be careful that your localization choices don’t end up crippling your international SEO
- Pay attention to how information architecture choices can impact performance (ccTLD, sub-domain, sub-directory)
- Follow latest practices regarding meta information (hreflang, region-specific language locales…)
- Consider transcreation for apps descriptions and keywords
TESTING

• As good and useful as some in-context editing can be, there is no substitute from having someone actually use and experience the localized software

• pre-live testing should always be considered in parallel with a pseudo translation/localization process before the translation stage has completed

• pre-live testing should not be catching issues that pseudo translation could have caught
TOOLS FOR AGILE AND CONTINUOUS DELIVERY
DISTINGUISHING FEATURES/LIMITATIONS

**Features**
- File Management
- Workflow Automation
- Translating Environment
- In-context Translation
- SEO

**Limitations**
- File Incompatibility
- Project Management
- Translation Environment
- Freelancer Access
- SEO
MARKET LEADERS

• Tool or platform?
• Proxy or CDN?
• Crowdin, Phraseapp, Lingohub, Onesky, Easyling, Localizejs, Smartling, Onesky, Transifex, Serge
SMARTLING

Key Features

• Good at capturing translatable content with little developer intervention
• Global CDN (fast content delivery)
• Relatively Strong TM Management
• Excellent SEO tools
• Tailored connectors to different CMS

Drawbacks

• It’ll cost you
LOCALIZEJS

Key Features

- Little tech capability required
- “Push-button” human translation
- Inexpensive

Drawbacks

- Little or no SEO
- Poor quality control
- Limited TM management
PHRASEAPP

Key Features

• Largely agnostic to file type
• Github support
• Relatively inexpensive
• Translation Progress Dashboard

Drawbacks

• Limited translation memory management
• Buggy file conversions
• Incomplete Locale Selectors
ALCHEMY CATALYST

Key Features
• Supports all aspects of localization cycle, translation, engineering, testing and project mgmt.
• Advanced translation memory
• Long-time market leader

Drawbacks
• Requires employee training
• Needs in-house resources; local deployment
• Windows-centric; no iOS support
• A little crusty
WHAT TO CONSIDER

- Development Environment
- Budget
- Project management resources
- Level of automation needed
- SEO Expectations
- Future: enabling spoken interaction in multiple languages
SO....

• Opening the localization black box helps to bridge gaps between technologists, linguists and PMs.

• The sooner developers develop good i18n habits, the smoother the l10n process can go.
SHARE YOUR EXPERIENCES

• How often do you release content updates?
• What is the scope and cadence of releases?
• Are you aiming for “sim-ship”? (Simultaneous shipment across all markets incl. foreign)
• Think about your localization goals: agile or continuous delivery?
• Is it acceptable to distribute localized versions on a schedule other than the US releases?
• How do you ensure that source/English content is prepped for translation? How much is context needed?
• How do you budget for localization expenses? Is it tied to a customer’s order or market research? First-to-market?
• What does your localization team look like? Who is responsible for what?
THANK YOU!

Merci — Gracias — Danke
Grazie — Obrigado
شكرا
谢谢
당신을 감사하십시오
ありがとうございます
www.mtmlinguasoft.com