BREAK-OUT DISCUSSION RELATED TO L10N TOOLS

MARCH 16, 2012

Participants

Elena Rudeshko (moderator), ELEKS
- background: L10N of software products with software development and linguistic knowledge
- interested in: L10N process automation and tools development for L10N needs

Charles McCathieNevile, Opera
- background: tools development
- interested in: how to trick ordinary developers

Jan Anders Nelson, Microsoft
- background: L10N tools development for the company
- interested in: relationship between terminology, concept, fuzzy matching, content will be reused

Iulianna Van Der Lek, post graduate student
- background: learning about the tools
- interested in: CAT tools, introduction of the new tools

Matjaž Horvat, Mozilla
- background: developer of presented on Mar 15, 2012 Pontoon tool for websites
- interested in: general discussion, integration between tools

Margie Foster, Intel
- background: PgM for platform
- interested in: how to translate Web Apps

Fokke Sluiter, DGT unit of EC
- background: translation for EU institution; a lot of editing activities
- interested in: look from the linguistic perspective

Richard Ishida, W3C
- background: I18N activity lead
- interested in: processes standardization

Kis Hajnalka, LTC
- background: new for the company
- interested in: post-edition platform to post content to the final web

Poul Anderssen, Web Translation unit of EC
- background: systematical translation
- interested in: reducing the difficulties in every day life related to translation by tools

Discussed items

1. Software development process + developers' peculiarities
2. The necessity of standards implementation
3. Translation, translators and their needs
4. Software development/L10N interaction
5. Translation trends: TM, MT, professional translators or community?
6. Education
7. Software I18N aspects
Participants input

Charles:
[1] developers think about themselves during SW development process
  ▪ translation tool of custom building: started to develop in Spanish
  ▪ works with Spanish speaking developers -> difficulty to find developer with good English
[2] to have the same kit from one company is a death
[3] ways to shift staff from one tool from another smoothly
[4] start from scratch
  ▪ how to build the tool which helps to understand other issues?
  ▪ L10N is a problem for SW developers
  ▪ communities should have their own tools to talk to each other -- important for L10N process
[5] should community replace prof. translators?
  ▪ perfect translation or good enough -> depends on the situation when it is used
  ▪ it's very expensive to have translators who copy-paste during their work

Jan:
[1] XLIFF 1.2 usage is strict
  ▪ vendors should pay their attention whether the tools rely on the needs of the process:
    ○ be free to invest in their internal tooling to innovate on their process efficiency
    ○ still be capable of a clean hand/off - hand/back cycle with developers
  ▪ each business unit in the company has their unique business needs -> they look for the layers where
    I18N engineering can be common and try to differentiate only where the businesses dictate
[4] start talking with the stakeholders is important:
  ▪ early interaction will surface and remove I18N defects before they can leak into the L10N work cycle
    and “explode” across the number of shipping languages requiring a lot more time to correct
  ▪ each business unit in the company has their unique business needs:
    ○ they look for the layers where I18N engineering can be common
    ○ and try to differentiate only where the businesses dictate
  ▪ internal tool development teams meet to determine the common engineering requirements to collaborate
    on
  ▪ communication between different groups: “We need to converse”
  ▪ DEV tools: e.g. MS VS10 ships in 9 languages (http://msdn.microsoft.com/en-us/library/th647yhh.aspx):
    ○ this is a critical mess for Chinese due to string writing skills of developers
    ○ strings need to be rewritten from skilled writers prior to translation to assure the best results
    ○ developers and their writers need the ability to focus on the app in their native languages
[5] translation looks in the content -> the tool should optimize the difference -> we need to look at this
  ▪ Thai example:
    ○ 80% of visible UI have been translated -> excellent results for the market
    ○ visibility of UI strings is accomplished by instrumentation that users opt in to provide
  ▪ TM/MT/crowdsourcing feedback:
    ○ it is important to have trusted translators + MT usage
    ○ JP MT example: an increased expectation that the quality of translation gets better with time -> this
      evolves to an increased state where post MT human editing is needed
  ▪ healthy self-enforcing community -> it must be able to exist on its own or it will not be sustainable
  ▪ next steps: to move forward with HTML files, legacy, all old content
[7] current implementation of WR Tenet == 2 years -> good results in the engineering system -> look forward
    to further improvements over future release cycles
  ▪ quality improvement: the focus was created by looking at historical top bug areas and targeting them
  ▪ items to implement I18N: QA management + recognition from top that I18N is a part of the production
    process + skilled people == permanent process control
Richard:
[2] • standardization is required  
  • to build the tool to bring together known formats and the standards should be brought  
[3] • nowadays translators should permanently learn  
[4] • current tools need is to have interoperability between different formats

Fokke:
[3] • monolingual English vs. languages with wide strings (e.g. Bulgarian, Greek)  
[4] • constraints in the tools  
  • tools are not a problem, language is a problem  
  • L10N DEV tools: terminology is in English

Iulianna:
[3] • cross language problem is present  
[6] • there is a lack in the university education to supply L10N needs  
  • the number of studied CAT tools is limited  
  • all knowledge comes from work experience or desire of the person to investigate L10N and translation related processes  
  • there is a necessity to improve education process from L10N industry perspective

Matiaž:
[1] • policy of using English in SW DEV companies in Slovenia  
  • a lot of English DEV terms are slavisized in the professional developers jargon  
[4] • Pontoon is an attempt to improve translation of web site contents:  
  • live website L10N tool  
  • is still under development  
  • only one developer is involved

Margie:
[7] • I18N is still the secondary process from DEV perspective  
  • there is a number of defects which are postponed from release to release

Elena:
[1] • it is a real problem to have a good English speaking developer on board in Ukraine, Russia:  
  • technical staff usually has a problem with languages  
  • they can read tech documentation in English, but no more  
  • local companies improve English knowledge of their staff by organizing internal English language courses  
  • the language inside professional developer’s community is a continuous mix of slavisized English tech terms  
[3] • a lot of SW DEV related technical terminology is not translated and used in everyday life in English -> this caused a difficulty for tech translators  
  • problem with wide strings in APPs for Slavic languages in comparison with logical English  
[5] • a common trend of transnational corporations is to have a good enough language rather than a perfect one:  
  • e.g. Facebook and their local language communities  
  • Thai Facebook has been fully translated without involvement of professional translators -> they prefer to have a good enough Thai  
[6] • there is no any L10N/I18N related subjects in Ukrainian universities  
  • local companies trains their staff by themselves
I18N issues are usually a complexity during the work with the customer:
  - the chain is vendor -> customer’s L10N department -> customer’s DEV team and sometimes DEV teams are based in different locations
  - standardization of I18N processes is required internally in the organization

**Kis:**

[3] • the typical internal translation process is TM -> MT -> web-site -> post-editing

**Poul:**

[3] • consistently using Trados for re-use the translation
  - avoiding translation of XLS-files

**Common input:**

- non-English speakers is still an issue
- we who involved in L10N industry can speak on more than one language
- if you build the culture you should start speak English
- culture is built on our workplaces everyday

**CONCLUSIONS**

1. L10N is a problem
2. users are a problem
3. lack of education and training is a problem
4. exchange formats should be all aligned
5. standard format to use
6. new language -> to break the interchange -> to fix the interchange not the word
7. software developers don't think about L10N and I18N needs starting from college