Teaching XLIFF to translators and localisers

Lucía Morado Vázquez
Silvia Rodríguez Vázquez
University of Geneva
Contents

- Rationale
- Previous considerations
- Teaching methodology
- Module contents
- Case study -> Standards seminars at UAB
- Lessons learnt and future work
Our objective

Empower the translator

He/she needs to be in control of (and understand) the files he/she is manipulating.
**Respondents: 372**

<table>
<thead>
<tr>
<th>Tool</th>
<th>2010</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>OmegaT</td>
<td></td>
<td>32 (6%)</td>
</tr>
<tr>
<td>Déjà Vu (any version)</td>
<td></td>
<td>117 (10%)</td>
</tr>
<tr>
<td>SDL Trados 2007</td>
<td></td>
<td>354 (31%)</td>
</tr>
<tr>
<td>SDL Trados Studio</td>
<td></td>
<td>356 (32%)</td>
</tr>
<tr>
<td>WordFast Classic</td>
<td></td>
<td>168 (13%)</td>
</tr>
<tr>
<td>WordFast Pro</td>
<td></td>
<td>33 (7%)</td>
</tr>
<tr>
<td>memoQ</td>
<td></td>
<td>352 (31%)</td>
</tr>
<tr>
<td>SDLX</td>
<td></td>
<td>50 (4%)</td>
</tr>
<tr>
<td>Star Transit</td>
<td></td>
<td>36 (3%)</td>
</tr>
<tr>
<td>Across</td>
<td></td>
<td>47 (4%)</td>
</tr>
<tr>
<td>Cafetrans</td>
<td></td>
<td>59 (5%)</td>
</tr>
<tr>
<td>memSource</td>
<td></td>
<td>21 (1%)</td>
</tr>
<tr>
<td>Fluency</td>
<td></td>
<td>7 (0%)</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>123 (10%)</td>
</tr>
</tbody>
</table>

**Respondents: 458**

<table>
<thead>
<tr>
<th>Tool</th>
<th>2010</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>OmegaT</td>
<td></td>
<td>93 (8%)</td>
</tr>
<tr>
<td>Déjà Vu (any version)</td>
<td></td>
<td>117 (10%)</td>
</tr>
<tr>
<td>SDL Trados 2007</td>
<td></td>
<td>354 (31%)</td>
</tr>
<tr>
<td>SDL Trados Studio</td>
<td></td>
<td>356 (32%)</td>
</tr>
<tr>
<td>WordFast Classic</td>
<td></td>
<td>168 (13%)</td>
</tr>
<tr>
<td>WordFast Pro</td>
<td></td>
<td>33 (7%)</td>
</tr>
<tr>
<td>memoQ</td>
<td></td>
<td>352 (31%)</td>
</tr>
<tr>
<td>SDLX</td>
<td></td>
<td>50 (4%)</td>
</tr>
<tr>
<td>Star Transit</td>
<td></td>
<td>36 (3%)</td>
</tr>
<tr>
<td>Across</td>
<td></td>
<td>47 (4%)</td>
</tr>
<tr>
<td>Cafetrans</td>
<td></td>
<td>59 (5%)</td>
</tr>
<tr>
<td>memSource</td>
<td></td>
<td>21 (1%)</td>
</tr>
<tr>
<td>Fluency</td>
<td></td>
<td>7 (0%)</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>123 (10%)</td>
</tr>
</tbody>
</table>

**Respondents: 1135**

<table>
<thead>
<tr>
<th>Tool</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>OmegaT</td>
<td>93 (52%)</td>
</tr>
<tr>
<td>Déjà Vu (any version)</td>
<td>185 (26%)</td>
</tr>
<tr>
<td>SDL Trados 2007</td>
<td>93 (13%)</td>
</tr>
<tr>
<td>SDL Trados Studio</td>
<td>89 (4%)</td>
</tr>
<tr>
<td>WordFast Classic</td>
<td>25 (3%)</td>
</tr>
</tbody>
</table>

**None at all, dammit!**
<table>
<thead>
<tr>
<th>Tool Name</th>
<th>XLIFF 1.2 Support</th>
<th>Own custom extension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alchemy Catalyst 10</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Araya XLIFF Editor</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Microsoft LBA 5.3/MLP 5.3</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>MemoQ 5.0.64</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>MultiTrans Prism</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Okapi Framework M16</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>OmegaT 2.3.0</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>SDL Trados Studio 2011 SP1</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Solas v1</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Swordfish Translation Editor</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Translation Factory 4.5.14</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Translation Workspace 1.10</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>XTM 6.2</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

XLIFF support in CAT tools. Filip & Morado Vázquez (2013)
Rationale

Advantages of XLIFF
(..)

Tool independent:
XLIFF allows the creation of better quality tools as the tool providers are concentrating their work on improving the tool rather than on providing filters for different formats. It also has the advantage of allowing localizers to be expert in a small number of tools they know very well rather than being less expert with a wide range of tools.

XLIFF 1.2 White Paper (2007), p.6
Initial considerations

- Where could it be taught?
- Students’ technical background
Initial considerations

Where could it be taught?
In a translation related degree:

- **Localisation (and project management)**
  - Undergraduate level: BA in Translation
  - Postgraduate level: MA in Translation

- **XML and multilingual documents**
  - Postgraduate level: MA in Translation

- **Localisation standards seminar**
  - Postgraduate level: MA in Translation Technology
Initial considerations

Students’ technical background

- Mix of technical and non technical backgrounds (specially at MA level).
  
  -> We need to start from zero, but taking into account also the advanced students.

- Markup languages’ knowledge is required (at least basic html).

- Basic principles of XML are also needed.
Teaching Methodology

Theory is combined with hands-on sessions and labs.

Students have to develop their own problem solving techniques
-> Constructivism approach
Module Contents

- XML basic concepts (reminder)
- Introduction: standards of localisation
- History and development of XLIFF
- Extraction-merge paradigm
- Advantages of using XLIFF
- Support in CAT tools
- Main XLIFF elements and attributes
Module Labs and Hands-on sessions

Creation of a (minimal) XLIFF file
Validation using XLIFF Checker

```xml
<xml version="1.0" encoding="UTF-8"?>
<xlfiff version="1.2" xmlns="urn:oasis:names:tc:xliff:document:1.2"
<file original="hello.txt" source-language="en" target-language="es">
  <header>
    <note>Traduce con mucho cuidado</note>
  </header>
  <body>
    <trans-unit id="1" translate="yes" approved="yes">
      <source>Hello World!</source>
      <target>Hola mundo!</target>
    </trans-unit>
    <alt-trans>
      <source>Hello World!</source>
      <target xml:lang="fr">Bonjour le monde!</target>
    </alt-trans>
  </body>
</file>
</xlfiff>
```
Module Labs and Hand-on sessions

XLIFF Quiz
Inspect an XLIFF file
Answer to the questions

Localisation et gestion de project 2014

Moodle ▶ L10n14 ▶ Tests ▶ XLIFF Exercise ▶ Tentative 1

XLIFF Exercise - Tentative 1

1 What is the XLIFF version?
Points: --/1
Réponse:
Module Labs and Hand-on sessions

Corrupted XLIFF files
Inspect the code
Create a bug report
Fix the file
They have to develop their own problem solving techniques

<table>
<thead>
<tr>
<th>Bug #</th>
<th>Location</th>
<th>DESCRIPTION/COMMENT</th>
<th>ORIGINAL</th>
<th>CORRECTED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Corrupted1: Ligne 7</td>
<td>header n'est pas fermé. Il manque la balise &quot;&quot;&gt;&quot;</td>
<td>&lt;/header</td>
<td>&lt;/header&gt;</td>
</tr>
<tr>
<td>3</td>
<td>corrupted3: Ligne 1</td>
<td>&quot;xliFF version&quot; manque 1.2</td>
<td>&lt;xliFF version=&quot;&quot;</td>
<td>&lt;xliFF version=&quot;1.2&quot;</td>
</tr>
<tr>
<td>3</td>
<td>corrupted3: Lignes 9 - 17</td>
<td>&quot;corps&quot; doit être en anglais.</td>
<td>&lt;corps&gt; &lt;/corps&gt;</td>
<td>&lt;body&gt; &lt;/body&gt;</td>
</tr>
<tr>
<td>4</td>
<td>corrupted4: Ligne 23</td>
<td>Il manque la balise &lt;/xliFF&gt; a la fin.</td>
<td>&lt;/xliFF&gt;</td>
<td>&lt;/xliFF&gt;</td>
</tr>
<tr>
<td>5</td>
<td>corrupted5: Lignes 2 - 18</td>
<td>&quot;archivo&quot; doit être en anglais.</td>
<td>&lt;archivo &lt;/archivo&gt;</td>
<td>&lt;/file&gt;</td>
</tr>
<tr>
<td>5</td>
<td>corrupted5: Ligne 3</td>
<td>&quot;body&quot; doit être déplacé de la ligne 3 à la ligne 4</td>
<td>&lt;header&gt; &lt;body&gt;</td>
<td>&lt;header&gt; &lt;header&gt;</td>
</tr>
<tr>
<td>5</td>
<td>corrupted5: Ligne 17</td>
<td>&quot;header&quot; doit être déplacé de la ligne 17 à la ligne 3</td>
<td>&lt;header&gt; &lt;body&gt;</td>
<td>&lt;header&gt; &lt;header&gt;</td>
</tr>
</tbody>
</table>
Standards seminar at UAB
XLIFF y estándares de localización

MA in Translation Technologies (Màster en Tradumàtica)

- 8 hours seminar.
- It belongs to the module (2nd term): Web, multimedia and videogame translation.
- February 2014.
Module objectives

Acquire advanced knowledge on

- Theoretical background on main localisation standards.
- Technical aspects of XML based formats
# Module structure and contents

## Contents

<table>
<thead>
<tr>
<th>XML</th>
<th>Labs</th>
<th>Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) XML and localisation</td>
<td>Creation of a XML file</td>
<td>Exchanger XML Editor</td>
</tr>
<tr>
<td>b) Filters for XML</td>
<td>Filter for XML in a CAT tool</td>
<td>SDL Trados 2011</td>
</tr>
</tbody>
</table>

---

**Image:**
- **Exchanger XML Editor**
- **SDL Trados 2011**
Module structure and contents

Contents

- Localisation standards
  a) LISA standards
  b) OASIS XLIFF

Labs

- Division of a TMX file.
- Creation of a minimal XLIFF file
- XLIFF Quiz
- Corrupted XLIFF files

Tools

- Notepad++
- XLIFF Checker
- Any tool they want...
Module structure and contents

Contents

- Standards in OS
- Gettext system
- Tinygettext

Labs

Linguistic QA testing of a game, fixing the errors in the po file.

Tools

- SuperTuxKart
- Virtaal
- Notepad++
Feedback from students

General appreciation  9.3/10

The learning process was adequate

https://docs.google.com/spreadsheet/ccc?key=0AqulmYUd46OOdG9talF4MXp6cG9OQlg5R25NYWh1X3c&usp=sharing#gid=0
Feedback from students

The content of the seminar has fulfilled my expectations

- Yes: 18
- More or less: 2
- No: 0
Feedback from students

The objectives of the seminar were achieved

- Yes: 18
- Partially: 1
- No: 0
Feedback from students

The materials of the seminar were adequate
Feedback from students

The difficulty level was adequate

- Yes, it was.
- No, it was not. Something was missing.
- No, it was not. Many things were missing.
Feedback from students

Do you think that you could apply the acquired knowledge to your professional life?

- Yes: 16
- No: 2
- I do not know: 0
Feedback from students

Other comments

Creo que esta asignatura debería darse más pronto y tener más horas, ya que hemos visto mucha materia en poco tiempo. Además, mucha de lo que hemos dado es básico para entender mejor cómo hacer filtros o corregir archivos ocultos, temas que sería mejor conocer más al inicio del curso.

I think this seminar should be taught earlier and it should have more hours, as we have seen a lot of topics in a short period of time. Moreover, what we have learnt is basic to understand how to create filters or fix hidden files, it would be better to know how to do that earlier.
Feedback from students

Other comments

*Es muy útil de cara a nuestro futuro profesional.*

*It is very useful for our future professional career.*
Feedback from students

Other comments

Quizás fue mucha información concentrada en poco tiempo. Sin embargo, las actividades prácticas fueron muy interesantes y nos sirvieron para entender mejor en qué consiste un estándar como XLIFF.

Maybe it was too much information in a short period of time. However, the practical labs were very interesting and they helped us to better understand how a standard like XLIFF works.
Feedback from students

Other comments

Ha sido una asignatura demasiado comprimida. Deberíamos haber tenido alguna sesión más.

The seminar was too “compressed”. We should have had some more lessons on the topic.
Lessons learnt and future work

- XML background should be a must.
- XLIFF should be taught at the end of the semester.
- The smaller the group, the better iteration (hands-on session)

- Gather more real data and problems related to XLIFF
- Introduce XLIFF 2.0
References


