TOOLS FOR QUALITY ASSURANCE AND TRANSLATION MEMORY MAINTENANCE

Tuomas Kostiainen (tuomas@jps.net)
www.finntranslations.com

---

1. Translation Quality Assurance (QA) in CAT tools

- SDL Trados Studio 2011 (see Figure 1):
  - Document Verifier
  - QA Checker
  - Terminology Verifier
  - Tag Verification
  - XML Validation
  - see also this blog article about terminology verification in Trados Studio
- memoQ 5.0 (see Figure 2)
  - The QA configuration pane consists of five tabs: Segments and terms; Numbers; Punctuation; Spaces, capitals, character; and Inline tags.
  - QA check can also be done for translation memories
  - Quality Assurance in memoQ webinar
- Wordfast 3 Pro (see Figure 3)
  - Transcheck: Edit > Preferences > Transcheck (Forbidden characters, Partial, Punctuations, Terminology, Untranslatable). See also Using Transcheck.
- SDL Trados 2007
  - Available in TagEditor (Tools > Plug-ins > Verifiers > SDL Trados QA Checker 2.0)

2. Translation Quality Assurance (QA) and Translation Memory Maintenance with QA tools

- ErrorSpy (see Figure 4): "Trados unclean files or text files which contain both languages alongside each other, sentence by sentence (source sentence and translation separated by any separator you wish). Similarly you can check Translation Memories (Trados TXT format or TMX files), Trados TTX files, Transit XV files or XLIFF files." Includes an internal editor, a spell-checking function (for a very limited number of languages) and an XML validation check.
- QA Distiller (see Figure 5): TTX, XLIFF (incl. SDL and memoQ XLIFF), STF, TMX. Editing in an internal editor or TagEditor.
- Verifika (see Figure 6): TTX, SDLXLIFF, XLF/MQXLIFF/MQXLZ (memoQ), TXML (WordFast), TMX (additional bilingual formats will be added). Includes an internal editor and a spell-checking function (MS Word or Hunspell).
- ApSIC Xbench (see Figure 7): XLIFF, TMX, TTX, SDLXLIFF, TXML, ITD, etc. New version also includes a spell-checker (Hunspell; limited number of languages). *Project Checklist* and *Personal Checklist.* Limited number of checks compared to others. However, has a very powerful search function for bilingual files. No internal editor but can be linked to external editors.

- Okapi CheckMate (see Figure 8): any bilingual or multilingual file format supported by the Okapi filters, for example: XLIFF, TMX, PO, TS, TTX, Trados-Tagged RTF.

3. Translation Memory Maintenance in CAT tools

Built-in functions in CAT tools (such as Trados Studio, memoQ and Wordfast Pro) for editing, searching, filtering, importing/exporting TMs

Trados Studio 2011
- Format: SDLTM
- Import: TMX, SDLXLIFF, TTX, ITD
- Export: TMX
- Transfer: SDLTM
- Translation Memories view > File >>
- Translation memory editor
- Blog articles about TM maintenance and TM fields

memoQ 5.0
- Format: MTX (several files) and MTM (one file)
- Import: TMX, CSV
- Export to TM: from LiveDocs corpus (bilingual DOC/RTF, TTX, XLIFF)
- Export: TMX
- Transfer: TMX
- Resource console > Translation memories
- Translation memory editor

Wordfast Pro 3.0
- Format: TXT (Wordfast Pro format)
- Import: no import function but TM content can be copied and pasted in a text editor
- Export: TMX, TXT
- Transfer: TXT
- Translation memory > New/Select TM > Translations > Translation Memory > Local
- Translation memory editor: TM Administration perspective
  - somewhat limited: no filters, no regular expressions (regex)

4. Translation Memory Maintenance in Text Editors

Translation memories in TMX and TXT format can be edited in text editors that can handle large files, such as UltraEdit. UltraEdit has also good find and replace functions that make it easy to do repetitive and complex searches and replacements quickly.

5. Use of Regular Expressions in QA Functions/Tools

Sources of regex information:
- http://www.qa-distiller.com/files/CheatSheet.pdf
- http://www.codeproject.com/Articles/9099/The-30-Minute-Regex-Tutorial
- http://www.regular-expressions.info/
- http://regexlib.com and http://regexlib.com/CheatSheet.aspx
- http://atril.com/en/content/13-regular-expressions-regex (Dejavu)
- http://catology.boisset.eu/Regular+expressions
- http://kilgray.com/webinars/regular-expression-based-filters-1700-cet-0 (memoQ)
- http://kilgray.com/webinars/regex-masses-english-011749 (memoQ)
- http://msdn.microsoft.com/en-us/library/az24scfc
- ApSIC Xbench User Guide

Use of regex in CAT and QA tools:
- **Trados Studio:** QA, file filters, find & replace, display filtering, segmentation rules
- **memoQ:** used in Auto-translation rules (incl. page ranges, unit conversions), Segmentation rules, Cascading filters (for file types), File filters, Internal RegEx Tagger (Format > Run regex tagger), QA settings (Check auto-translatables)
- **Wordfast:** not in QA features; only in file filters
- **ApSIC Xbench, CheckMate, ErrorSpy, QA Distiller** and **Verifika:** extensive use of regex

Examples for the presentation:
- `(\d{2})/(\d{2})/(\d{4})`   05/19/2012 vs. 5/19/2012
- `(\d{1,2})/(\d{2})/(\d{4})`   05/19/2012 and 5/19/2012
- `[0-9]{1,}%`   75% vs. 75 %
- `[0-9]{1,}\smm`   25 mm vs. 25mm
- `\d[\s]*-[\s]*\d`   5 - 6 vs. 5-6
- `ISO\s\d+`   ISO 3000 vs. ISO3000
- `^nut`   Nutritious food...
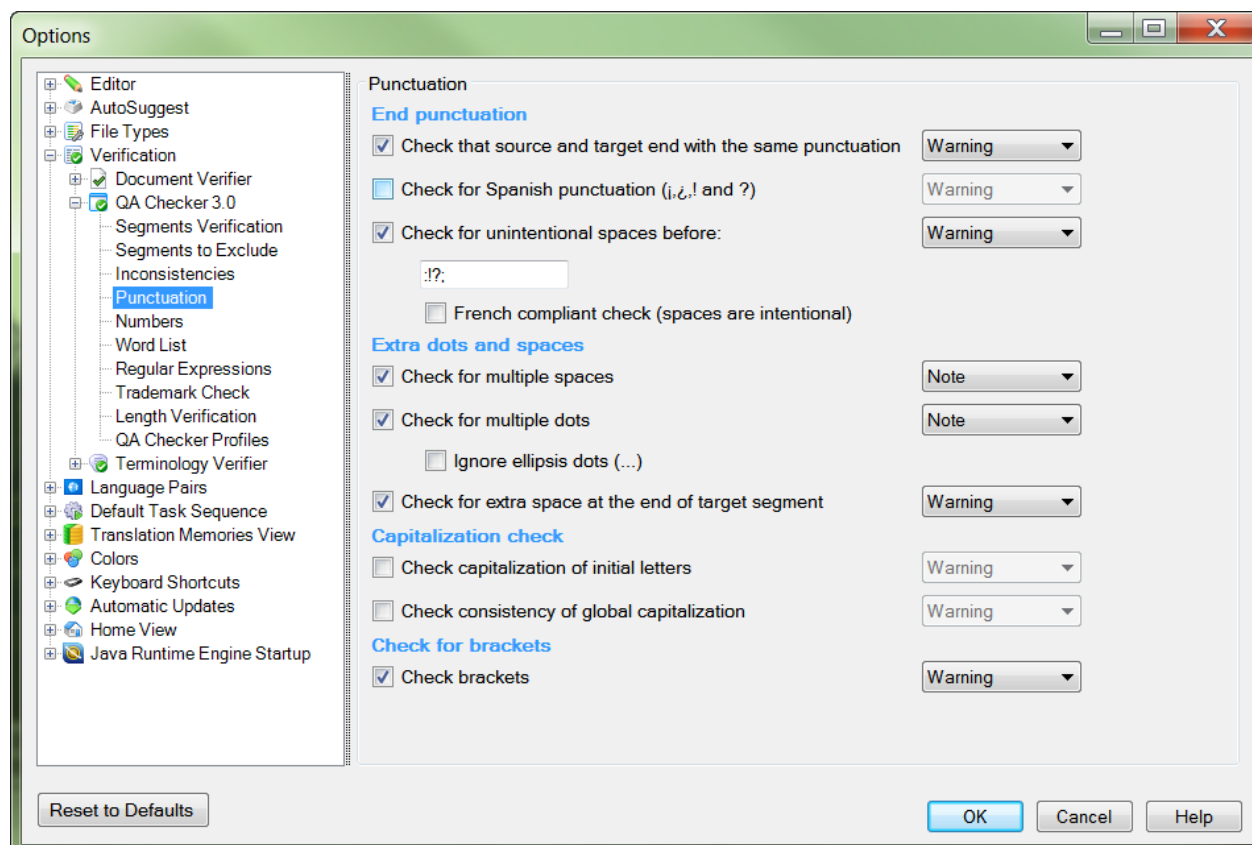- `nut$`   ...and a hazelnut
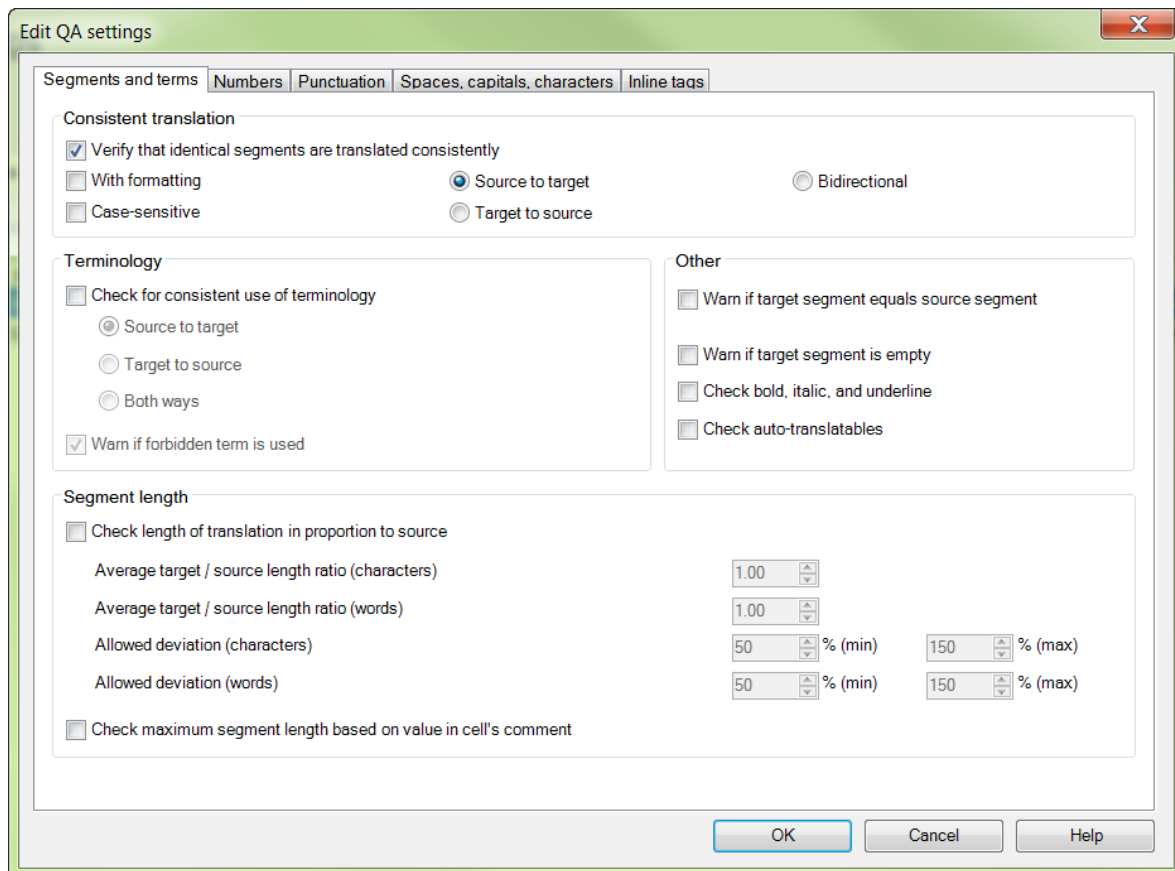- `nut\s|nut$`

Figure 1. Trados Studio 2011 QA settings

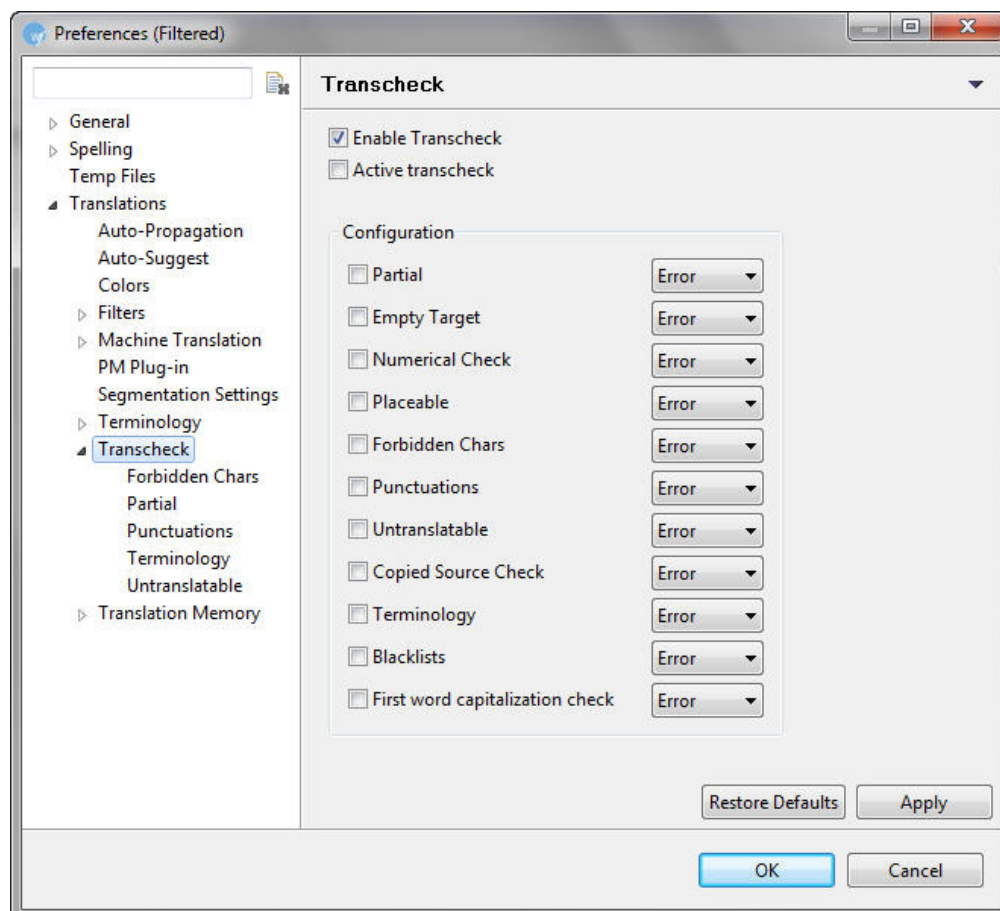Figure 2. memoQ 5.0 QA settings
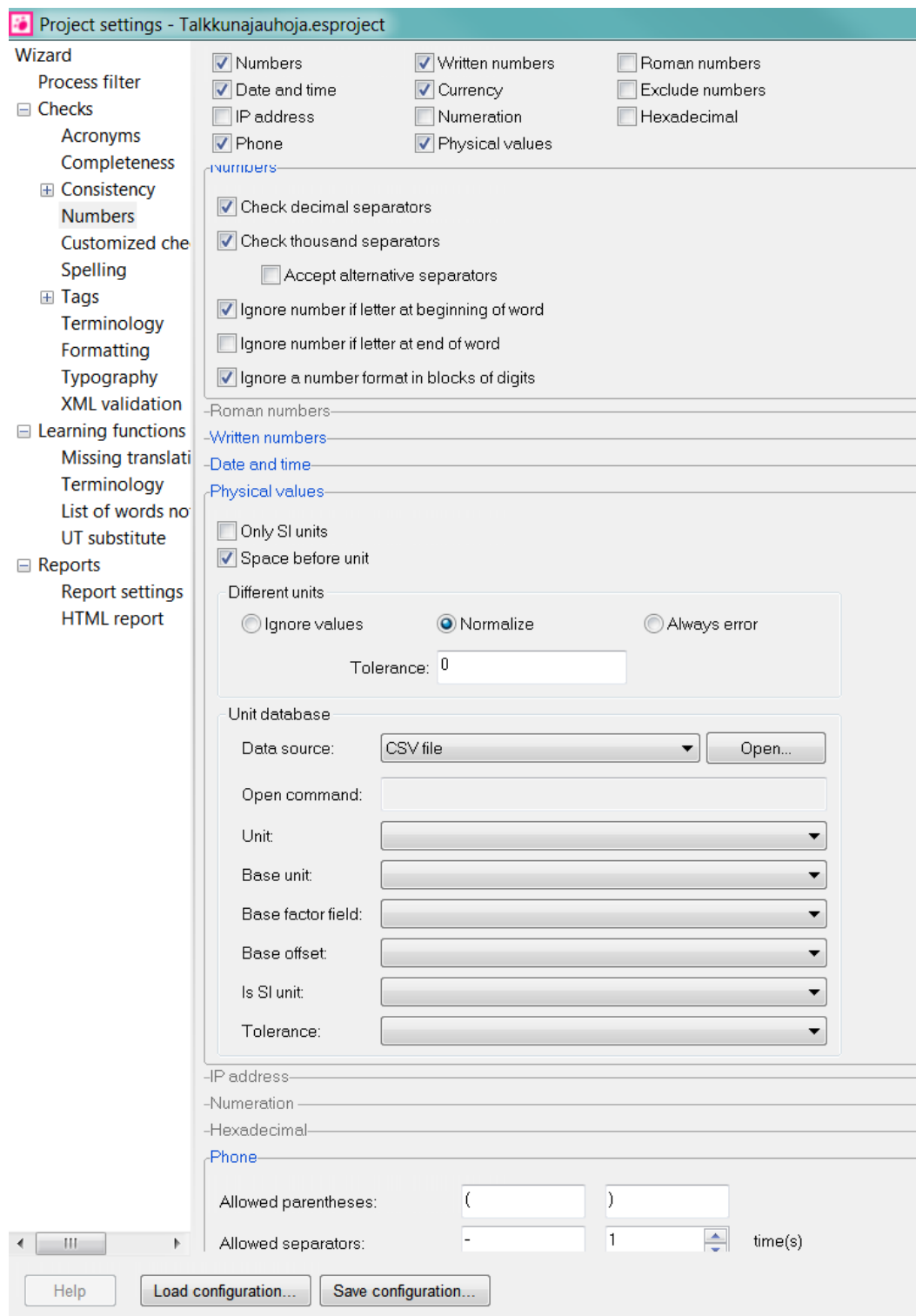
Figure 3. Wordfast Pro 3 QA settings (Transcheck)
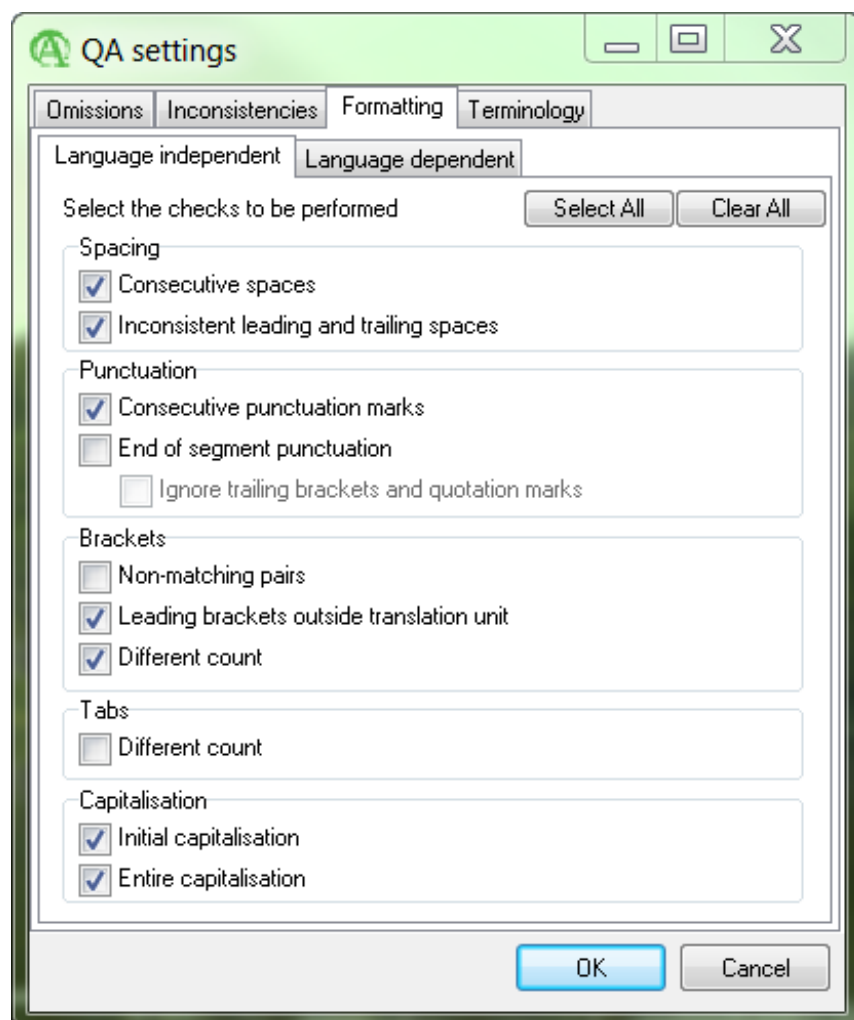
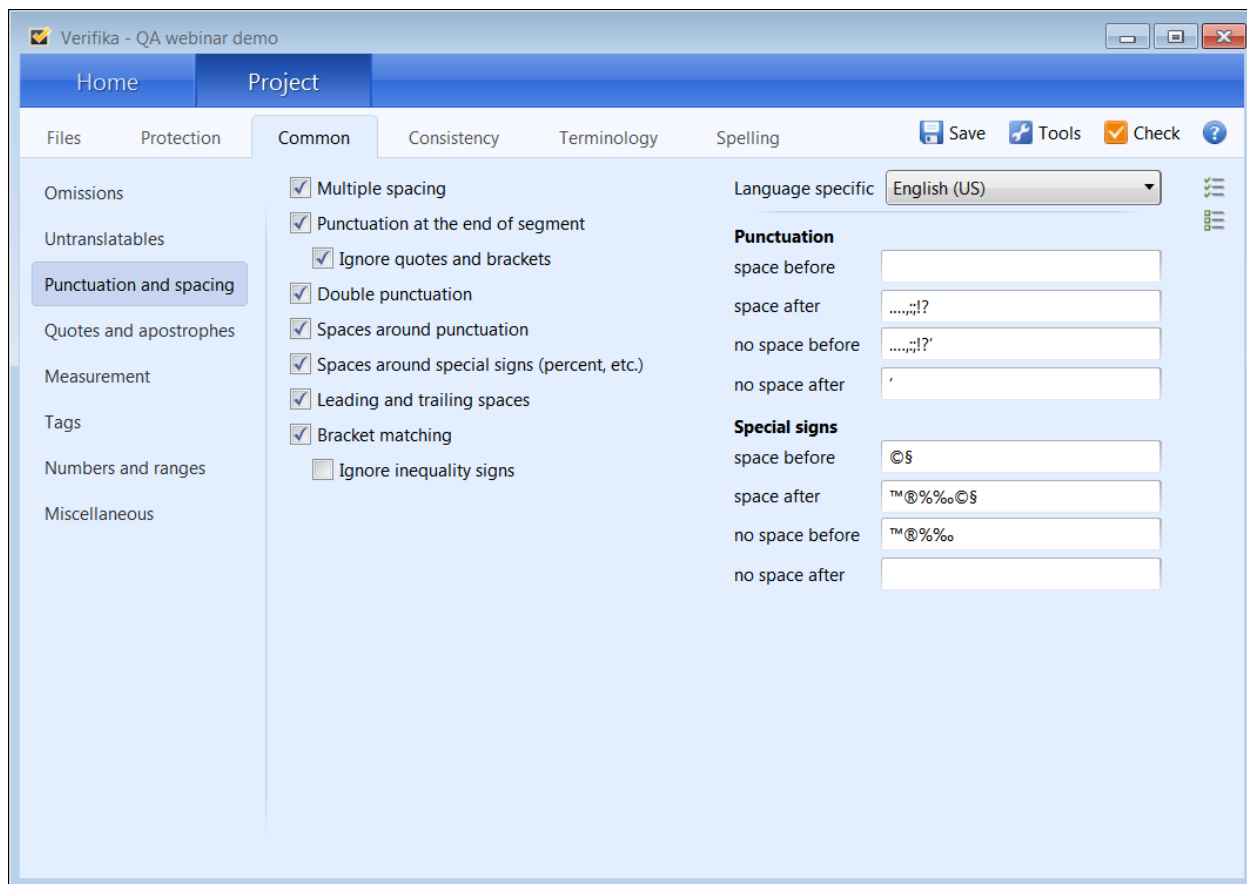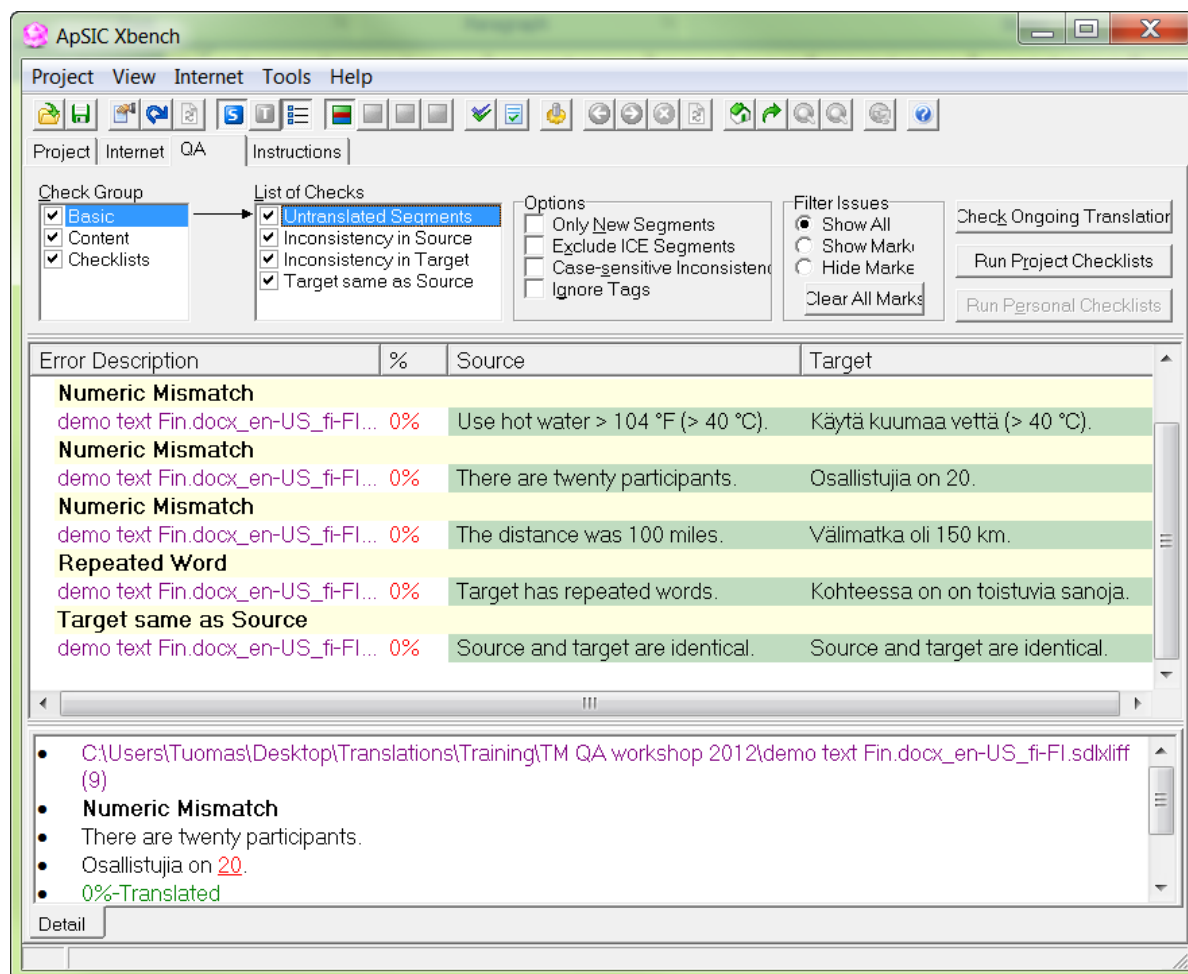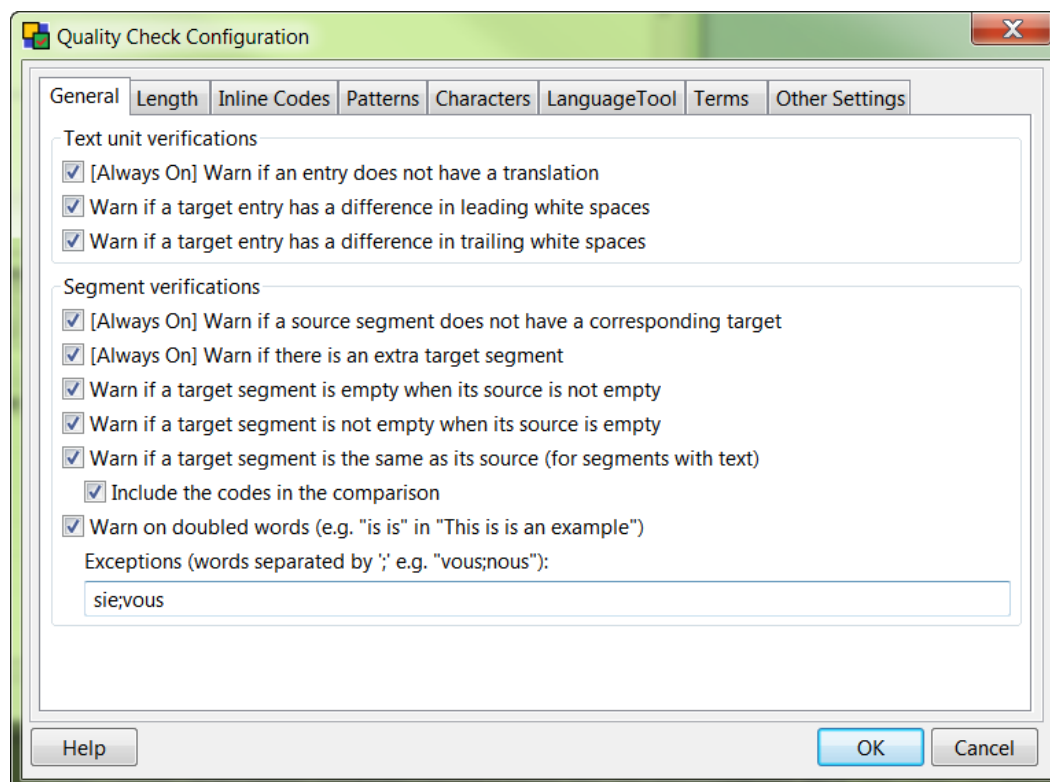Figure 4. ErrorSpy 6.2

Figure 5. QA Distiller 8.5

Figure 6. Verifika 1.2

Figure 7. ApSIC Xbench 2.9

Figure 8. Okapi CheckMate