Tailor-made Quality-controlled Translation

Sergio Penkale    Andy Way
Lingo24
sergio.penkale@lingo24.com
Outline of Presentation

1. Background
2. The Case for Customisable Levels of Quality
3. Quality Levels in Human and Machine Translation
4. Customisable Quality Levels in Coach
5. Conclusions
Business requirements are changing

- Explosion in online content
- User-Generated Content
- High-quality professional translation not feasible in every case
- MT improving
- New workflows emerging
Quality and fitness for purpose

- Lifespan: how soon will the content become redundant? *Perishability*
- Target Audience: Expectations from users of the content
- Volume of content that needs translating
- Urgency with which is needed
- ...

- Style and fluency vital for press release
- Not so much on a technical manual. Accuracy always essential
- Real-time conversation: speed is main priority
Different needs, different quality levels

Level 1
Bulk of the website (60%)
- Support documentation
- Less popular products
- Rarely visited pages

Level 2
General content (30%)
- Main product pages
- General information pages
- Terms and conditions

Level 3
Critical content (10%)
- Landing pages
- Promotional pages
- Press releases
Different needs, different quality levels

Automated Quality Checks - Level 1 Preset

- Completeness Check
  - Tag mismatch
- Formatting Check
- Punctuation
- Number Formatting
- Number Values
- Spellcheck
- Empty Translation
- Terminology Check
- Inconsistent TM Matches
- Inconsistent Project Match

lingo24
love your language
Different needs, different quality levels

**Level 1**

**Bulk of the website (60%)**
- Support documentation
- Less popular products
- Rarely visited pages

**Level 2**

**General content (30%)**
- Main product pages
- General information pages
- Terms and conditions

**Level 3**

**Critical content (10%)**
- Landing pages
- Promotional pages
- Press releases
'One-size-fits-all' outdated, but . . .

One recent trend is the offering of various ‘quality levels’, something professional translators cannot and will not do. **For us there is only one quality level: professional, publication-ready quality.** (Rose Newell)
Not all (translators) would agree!

- At least two levels for some time: Light vs. full post-editing...
- Bellos (2011: 335): “not all [translators] are great at their job”, so whole idea of “one quality level” inherently flawed.
- Indeed, rather than criticising “just about every bulk translation agency”, Newell is dismissive of the PEMT work that many thousands of her fellow translators perform.
- Clearly these “professional translators” can and do offer different levels of quality which are fit for purpose...
- Olga Beregovaya (LocWorld 2013, London): “the old quality models may no longer be the answer when applied to post-edited output used for new content delivery methods”
What do users find useful?

- Zetzche: “translation quality will remain a contentious topic of discussion, maybe more so than as a matter of implementation”
- Gives specific examples of MT-ed and human-translated Help files on Microsoft’s knowledge-base
- “A translator who compares the translation quality of the two articles will immediately have a visceral response: one is ’good quality’ and the other seems to scream out its ’poor, machine-translated quality.’ But the users? They find them both (virtually) equally helpful”!
- “The perception of quality needs to be a lot more dynamic. There is certainly room for quality metrics and standards, but we need to accept that these don’t apply to everything. And some of the translation buyers have long figured that out.”
What is 'good quality'?

- Fluency: cf. 'light' vs. 'full' post-editing
- Checking consistent glossary usage will impact in both quality and cost
- Same for style, TM adherence, etc.
We need a *dynamic* model of translation quality evaluation!

- **Text-related issues**: content type, context
- **Purpose**: end-user requirements, communicative function, perishability
- **Mode of translation creation**: qualified human translator, unqualified volunteer, MT or TM system or some combination thereof
- **Traditional problems with evaluation**: subjectivity, time, inappropriate use of linguistic resources, learning curve, technology
- **Emerging problems with evaluation**: budgetary constraints, new notions of 'text' – tweets, blog postings, multi-media, UGC, – new technology
Evaluating MT quality

- Traditional MT evaluation done comparing with ’perfect’ human translation (e.g. BLEU)

- Manual fluency and adequacy marks

- Ranking, task-based, etc.

- Need to use most appropriate metric to ensure fitness for purpose

- Quality Estimation will play a vital role
Customisable quality levels in Coach

**Managed tasks**
- Terminology use
- Tone of voice
- Fluency

**Automated quality checks**
- Inconsistent TM matches
- Measurement Units
- Empty segment check
- Inconsistent spaces
- Tag matches
- Non-matching brackets
- Entire capitalisation
- Initial capitalisation
- Spell checking
- Non-matching question marks
- Consecutive spaces
- Punctuation
- Numerical values
- Symbols

lingo24
love your language
Customisable quality checks

- Live highlighting of spelling mistakes and multiple spaces,
- Live terminology check,
- Inconsistent translations,
- Language-specific QA checks: Number formatting & Punctuation,
- Mandatory vs. Optional.
Filtering translatable content

- Translator access to QA report.
- Information may be shared, depending on workflow.
Propagating changes & repetitions

Do you wish to update subsequent repetitions?

- Yes, for this file
- Yes, across all my active files
- No
Enhanced internal matching

- In-Context Repetitions
- Fuzzy Repetitions
Measuring real vs. expected progress

<table>
<thead>
<tr>
<th>Progress</th>
<th>Task/Batch Deadline</th>
<th>Overall time</th>
<th>Editing distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>2013-04-22 11:00:00</td>
<td>00:30:00</td>
<td>11.4%</td>
</tr>
<tr>
<td>100%</td>
<td>2013-04-19 17:00:00</td>
<td>00:11:38</td>
<td>0.27%</td>
</tr>
</tbody>
</table>

230 words translated, 230 expected by now
Sampling workflow
Sampling workflow (2)

Was the sampling successful?

- YES
  - Final document
  - The final translated document

- NO
  - Experienced proofreader
  - Sampling 1
  - Sampling 2
  - Up to cost saved from full proofreading 60%

Was the sampling successful?

- YES
  - Final document
  - The final translated document

- NO
  - Experienced proofreader
  - Sampling 1
  - Sampling 2
  - Sampling 3
  - Up to cost saved from full proofreading 40%
API integration

- API support is essential
- Reduction in overhead costs
- Reduction in delivery times
- Ability to select required quality level
- Automatic allocation of most appropriate resources (e.g. Automatic Subject Detection)
Summary

- ‘one-size-fits-all’ shifting towards fitness for purpose
- For industry to take advantage, tools enforcing quality levels are required
- COACH has the capability of being a game-changer in the localisation and translation industries.
- Makes translators *demonstrably* more efficient.
- Delivers translations at predictable levels of Quality and Price
Thanks for your attention!

sergio.penkale@lingo24.com