Professional Technology for Interpreter Education
About

Productivity software for translation revision that

- speeds up the revision process
- guarantees consistency and objectivity, and
- works independent of language or domain
What is translationQ not?

Computer-Assisted Translation (CAT)

Machine Translation (MT)
We are a CAR, not a CAT
(but we can live together as you can see...)

CAR = computer assisted revision
Revision today...

In het kader van het bovengenoemde dossier, heb ik de eer u mede te delen dat ik een verzoek aanhangig gemaakt heb bij de Franse Centrale Autoriteit ten gunste van mevr. D C B, gedomesticeerd... in in België, pm een onmiddellijke terugkeer te verkrijgen, naar België, van haar dochter, D C O.

Uit de meegedeelde informatie door de eiserses blijkt dat mevr. D C B, geboren op ... te ..., en de heer C M, geboren op ... te B, een korte relatie onderhouden hebben waaruit een kind voortgekomen is, D C O, geboren op ... te P.

De partijen zouden in mei uit elkaar gegaan zijn, namelijk voor de geboorte van het kind.

Na deze scheidin, zou D C gehuisvest geweest zijn op de woonplaats van haar moeder, mevr. D C M, in Parijs in Frankrijk.

Nadat duidelijk werd dat ze een kind verwachtte, heeft mevr. D C de heer C over de situatie ingelicht. Deze laatste zou geprobeerd hebben haar te overtuigen om vrijwillig de zwangerschap te ondernemen, wat de eiserses zouden hebben geweigerd.

De eiserses heeft bijgevolg de zwangerschap bewust alleen voor haar rekening genomen.

Na de geboorte zou de heer CH nooit betrokken geweest zijn in het leven van O, ondanks de pogingen van de moeder om een band tussen de vader en het kind te sturen. Ook zou hij gewijzigd gedraagd hebben jegens de moeder.

Gezien deze situatie en het feit dat zij buiten haar moeder geen banden had met Frankrijk, zou mevr. D C besloten hebben om in april weer in België te komen wonen.

Door deze verblijfsverandering heeft de heer C een verzoek aanhangig gemaakt bij het gerecht te P. om een besluit te nemen over het gezag over het kind. Toch zijn, volgens de eiserses, alle aktes met betrekking tot die procedure betrokken, aan de woonplaats van haar vader zelfs al was de heer C op de hoogte van haar verblijfplaats in België.

Bijgevolg is een beslissing gewezen op ... zonder dat mevr. D C aanwezig kon zijn of vertegenwoordigd kon zijn.

Op grond van deze beslissing, heeft het Tribunal de Grande Instance te P. de verblijfplaats van het kind gevestigd op de woonplaats van de vader.

Opmerking: wijziging van verblijfplaats.
Revision today...

Impossible to get overviews and statistics and evolutions...
Meet the translationQ pilot audience

- 96 admin accounts
- 45 organisations
- 22 universities
- 11 countries
- 803 users
How did we gather the “pilot” input?

3 surveys for pilot customers
1 student survey
1 Special Interest Group
Numerous e-mails, conversations and visits
Event and follow-up surveys

- Which are the top 3 features we should add to translationQ as soon as possible for you?
  - Need to be able to go back and edit assignments already published (e.g., due date, title of file)
  - Possibility to customize the error categories
  - Automatic segmentation
  - Ability to import users lists (Excel)
  - Possibility to change the maximum score after revision process
  - Possibility to overwrite the “system” score
  - Link error memory to user group
  - Import different file types, including .docx, .sdlxliff, .sliff, .xlf,
  - Positive feedback to translators
  - Adding generic feedback
  - Statistics
  - Uploading multiple (Word) documents for revision
Follow-up surveys

- Which are the top 3 features we should add to translationQ as soon as possible for you?

- Final translated text next to source text and revised translation
- Uploading source text with original formatting (incl. paragraph structure) and possible image
- Sharing of error / revision memories across a course team
- Adding multiple revisers to a revision project
Follow up survey 1

• Which error categories do you use today?
  – TAUS DQF
  – MeLLANGE
  – Custom set of categories
  – Grid used by the Italian Translators' Association (AITI) + the grid we used at the University of Trieste
  – ...

• Conclusion: allow customisation, but take into account data sharing
Follow up survey 1

• **How do you score translations today?**
  – A score from A to F to 4 questions: style, terminology, fluency, general opinion
  – PIE-scores
  – I have my own system, and need to be able to change the maximum score afterwards to continue using my system
  – I establish how serious errors are, I grade them from -0.25 to -2, I give points from +0.25 to +2 for good solutions, I decide the pass/fail threshold for each source text translation and then try to apply it consistently
  – Max score per section (Target language, Meaning, Specialized Meaning and Functional Adequacy)
  – ...

• **Conclusion**: allow customisation
Conclusion

- translationQ is the result of **co-creation** between KU Leuven, Televic and all the participants in the academic pilot project
- Your input and feedback is crucial to help us achieve our goal: **building the best revision tool**.

Thank you!
But there is more...

- Big Wrong Data
With translationQ ...

Get **insights** in
Every translator’s performances and evolutions
Common errors
Common error categories

...
Reports available in translationQ

- Insights in a translator’s performance
Reports available in translationQ

- Translation insights
What we learn from the error memories

• Data from UNI Padova (*grazie mille!*)
  – Number of different errors
  – Average frequency on an error
  – # of unique errors vs. # of repeated errors

• Impressive numbers of errors
  – E.g. Hogeschool Zuyd (NL)
    • # 2 revisions
      – 1 with 10 submitted translations
      – 1 with 9 submitted translations
    • 405 different errors defined
Revision memory insights

The magic of dynamic tables...

- Frequency of errors (→ efficiency of translation)
- Score distribution
  - with real error view
  - consistency of scoring
- Scoring per error category
What you can learn from Big Wrong Data

- Most frequent translation errors
- Strengths and weaknesses of each translator
- Averages and evolutions per translation agency
- Level of difficulty of a source text
- ...
What we learned from our database

- Error categories per translator
  - View per translator
    - E.g. Speranza
  - View per category
    - Ranking of categories
    - Who’s the category queen/king?
  - Relation between suggested/accepted/rejected errors
    - Is this the path to “generic errors”? More research needed!
What we learned from the revisors

• Different revision “habits”
  – E.g. Some use “Feedback” instead of “Correction”
  – Very different scoring systems
    • Some score on 100
      – With scores ranging from -20 to +2
    • Some score on 10
      – With scores ranging from -0.25 to 0

• Too many revisors use “None” or “Uncategorised” as error category
  – *Should we prevent it?*
What we want to research in the near future

• How reusable are revision memories?
  – Differences per language pair
  – Within same/across domains
  – Within one/across institution(s)
  – With multiple revisors (how to do it efficiently?)

• Relationship between revision memories and translation memories
  – What happens if revision memories are applied to TMs?
  – How can revision memories help to improve the quality of TMs?
  – What if CAT tools link their revision module to translation memories?
  – What if we compare the errors made by CAT and made by humans?
What we want to research in the near future

- Relationship between translation and interpreting
  - What we apply translation to interpreter errors? Interpreting Error Memories?

- Allow evaluators to bookmark interpreting errors
- Allow evaluators to categorize interpreting errors and adding
- Having overviews and samples of
Ongoing research

• Time and effort spent in
  – human revision vs.
  – car

• Intra-rater reliability

• Inter-rater reliability
  – scores
  – error memories (do they find the same errors?)
  – how consistent are they (same errors, same penalties?)

• Measure the influence of (teaching) activities (and feedback)
What is translationQ used for?

- Translation student evaluation and revision.
- Evaluating and ranking large groups of translators for selection and recruitment.
- To find and reuse common errors to create adaptive exercises.
- To screen and rank the quality of external translators and freelancers and translation agencies.
- Organize and objectively score high stakes translation exams.
- Define the difficulty level of source texts.

Use translationQ to train students to become professional revisors “a tool to teach revision in Revision technique class.”

1st language text evaluation and revision (not translations, eg. arbitration law).
“Reverse” revision: finding good translations

• Preselected Item Evaluation
  – Perfectly possible with this tool
  – Finding good translations of key sentences, words or full segments
  – Positive scoring
  – 100% consistent scoring
  – Psychometric analysis

  – Examples in literature, medical, legal texts...
• PIE steps

A human revisor preselects items/keywords in the source text or question

100 % objective/consistent
How do I prepare for this procedure [PI 3]?

Prior to beginning the study, you will be asked to drink enough water [PI 4] or fluids to fill your bladder. You may start filling your bladder from home providing your travelling distance is within one hour [PI 5]. It is important that the bladder should feel full (to a point that you would normally pass urine) but not be uncomfortably full. Do not overfill the bladder [PI 6].

When the bladder feels full, you will be asked to urinate [PI 7] into the special uroflow study container. It is important that you pass urine in as normal a pattern (for you) as possible. If you think that this urination was not normal (for you) tell the nurse [PI 8] and you may be asked to repeat the study.

If you have any mobility problems, or any other special needs, please let us know in advance so that we can make any necessary arrangements to accommodate you.

What happens after the procedure?

Immediately after the uroflow, an ultrasound probe will be passed over your abdomen. The nurse performing the procedure [PI 9] will lightly press the probe into the lower part of your abdomen and measure the amount of urine left in your bladder.

The results of these studies [PI 10] will be discussed with you on your next visit.
You want to buy a new or used car. Before you make the decision to choose a particular car, you want to take a test drive. But what happens if you have an accident during this test drive? Who is responsible for the damage?

If you are in error yourself, things are more difficult. The car seller must have a standard vehicle insurance (with minimum coverage requirements) for the car. It is legally mandatory. This insurance will compensate for the damage you caused to others involved in the accident. On the other hand, the damage to the car you drive is not reimbursed by this insurance. You are therefore risking ... turning yourself up. This would be covered by a Comprehensive insurance. A Third Party Property Damage insurance does not.
• PIE Steps
Calculation of the scores of the candidates on the preselected items

1. Raw score
   (number of hits/total number of PI’s)

2. If enough candidates (50+), calculation of
   - the item **difficulty level (p value)** and
   - the **discrimination index (d index)** of the items
• PIE steps
  ▪ the difficulty level (p value): adequate p value: between 0.27 and 0.79
  ▪ the discriminatory power (d index) of the items: adequate d index: higher than 0.29

⇒ Eliminating “invalid” items

▪ Bringing the raw score back to the final PIE score
What does the PIE method tell you?

NOT an evaluation of the full answer quality

But: a reliable ranking of all your candidates

100% consistent
Objective
Much faster than full text correction

And a reliable ranking of your items