NERstar

___

Juan Martínez Pérez - SWISS TXT

NERstar – Evaluation and Post-Editing

- ■ New Features
  - Working Methods and Ways of Being Faster and More Efficient
  - Improve Quality of Subtitles in Terms of Accuracy, Display Mode, and Duration
  - Import of Video Files and Production of Sign Language Videos

New Features – Automated Processes

- Reality
  - Tight Deadlines
  - Automatic Translations on Text Level (Errors)
  - ASR-Files (Errors, High Text Density)

  → Correction Exclusively on Text Level
  → Focus on Text Veracity → ≠ Timecodes → ≠ Display Mode

NERstar - Telesguard

- **Workflow with Tight Deadline**
  - Broadcast: 17:40-17:50 → Romansh with Subs in DE
  - Retransmission: 20:00 → Romansh with Subs in IT
  - ≈ 90 Min. for correction of automatic translation
  - No video, correction is exclusively done on text level

NERstar – New Features

- Automated Processes → ● Import ● Editing ● Export
    1. Enforce Limits
    2. Special Features (e.g. Add Opener, Remove Tags and Duplicate Subs During Import)
    3. Recalculation of Timecodes

    4. Optimization of Duration
    5. Stabilization of Display Mode (Segmentation)
    6. Implementation of AI → Shorten Text, Automatic Translation, Simple Language → New Ground

NERstar + AI = New Ground

- Implementation of AI
  - Shorten Text (DE)
  - By-product Automatic Translation (DE-E, DE-IT)
  - Simple Language (In Process)

  → New Approach **→ Sentence- or Phrase-based Segmentation Instead of Subtitles Segmentation**

NERstar – New Approach

- **Why Sentence- or Phrase-based Segmentation? Assumptions:**
  - It is easier*
    - for AI to deliver better results when shortening or translating sentences instead of subtitles
    - to correct sentences than subtitles, in particular, sentences extending over one subtitle
    - to evaluate sentences than subtitles, in particular, sentences extending over one subtitle
    - to produce sign language videos based on sentences than subtitles
    *easier: is more representative of how we process information, and so does AI

NERstar – New Approach

- **Workflow (In Process)**
  - NERstar converts subtitle file into sentence-based file
  - Timecodes of subtitle file are not affected, but (temporarily) recalculated
  - This file then is used for correction and/or evaluation of subs, or production of sign language video
  - NERstar converts corrected sentence-based file into subtitle file again
  - This workflow will be part of an IICT research subproject

NERstar – New Features

- Automated Processes
  - Enforce Limits √
  - Recalculation of Timecodes √
  - Optimization of Duration → Formula → Redistribution of Time √
  - Display Mode (Segmentation) → Formula → Pyramid, Punctuation √
  - Implementation of AI → Shorten Text, Automatic Translation, Simple Language √
  - Sentence- or Phrase-based Segmentation √

NERstar – Sign Language Translation

- Workflow
  - Define text
  - Activate video recording
  - Record sign language
  - Save sign language video and corresponding *.srt

  → Material for improving sign language model = NERstar as sign language post-editing tool

Evaluation of Live/ASR/Translations

- **Workflows**
    1. Reference: Timecoded Transcription → 1. Evaluation
    2. Reference: ASR-Subtitles → At Once* → 1. Correction → 2. Evaluation
    3. Reference: Live-Subs → At Once* → 1. Correction → 2. Evaluation

    * "At once" meaning NER-Analysis on subtitle level is possible

NERstar – In The Nearest Future

- Planned Activities
  - Evaluate and compare automatic translation quality of NERstar-AI with e.g. Google, DeepL, etc.
  - Workshop for finding out possible programme formats NERstar could be used for production
  - Implementation in the different research projects, such as easier, 5GMediaHub, MediaVerse, IICT
  - In particular, evaluate sentence- or phrased based approach in the framework of the project IICT
  - Assess the quality of live subtitles in Switzerland
  - Look for synergies among the different access services and develop best practice

NERstar…

- …for Sign Language Post-Editing
- …as Transversal Post-Editor and for Simple Language
- …for Multilingual Post-Editing
- …for Citizen Journalism and Social Media Post-Editing

NERstar Academia

- **Free Access for Research Purposes**
  - NERstar editor provided as web-based system
  - Accessible via URL: https://nerstar.swisstxt.eu
  - Projects can be saved as files with extension *.ner, or exported as PDF
  - For general inquiries about NERstar and its functionalities: elson@speedchill.com
  - For using NERstar on SWISS TXT productive server: Gion.Linder@swisstxt.ch
  - For using NERstar in the framework of other research projects: Robin.Nachtrab-Ribback@swisstxt.ch

NERstar

- **Right Now**
  - Editing, Enforce Limits, Recalculation of Timecodes
  - NER-Analysis Using Transcription as Reference, Video Import
- **Next Week**
  - Optimization of Duration
  - Stabilization of Display Mode
  - Implementation of AI
  - NER-Analysis on Subtitle Level
  - Sign Language Video Recording

Thank you!

_____

Juan Martínez Pérez - SWISS TXT