Success of automatic speech recognition

in the field of subtiling

WHO WE ARE

A Swiss company

- A Swiss public limited company
- Owned by the public broadcaster SRGSSR
- Four sites in Switzerland, headquarter in Biel
- Multilanguage
- Main focus: Editing, conversion and distribution of audiovisual content

WHAT ARE ACCESS SERVICES?

For whom?

- Enable people with sensory disabilities to access audio-visual content
- Mainly:
    - Subtitling ➔ for hard of hearing
    - signed programs ➔ for deaf
    - audio description ➔ blind
    - audio subtitles ➔ people without foreign language skills and dyslexia
    - content in simplified language ➔ people with a mental disability

SUBTITLING

The most important access service

- We provide subtitles for more than 35 years
- About 30'000 hours a year on 7 TV channels in three languages
- Staff of 100 people

Subtitles are created
- by editing on a keyboard for pre-recorded content
- or by respeaking for live content. We use Dragon 15 Group

Main challenges
- Increase the subtitling quota on television and any other media distribution channel
- Keep costs under control: it takes 10 working hours to subtitle one hour of content

SUBTITLING

What makes a good subtitle?

Some main rules
- Be in sync
- Do not exceed a certain reading speed
- Do not drop below a minimum duration, even for a "No!"
- Decomplex nested sentences
- Avoid metaphors
- Indicate speaker changes

- Subtitles are not a timecoded transcription!

SUBTITLING

A cost-intensive service

- Speech recognition becomes better and better
- But subtitling from automatic speech recognition remains incomprehensible
- And is therefore not an acceptable solution for us

- But automatic speech recognition with post-editing?
- Racu = Reconnaissance automatique avec correction ultérieure

SUBTITLING

This is how we proceeded 1

- Base: our workflow system mediahub
- Ability to integrate any speech recognition engine
- Fast and simplify editing process: we accept that the quality might drop slightly

- how are subtitles from ASR made?
    - Video is uploaded
    - text is recognized
    - the text is divided into chunks (captions)

SUBTITLING

This is how we proceeded 2

Any correction is made on the text level, so before it is divided.
The subtitles are divided by an algorithm.

Advantages:
- The corrector can focus on the text
- The person won't be distracted by formatting issues
- Faster
- Almost any person will be able to correct

Disadvantages:
- We lose autonomy over formatting
- and duration

SUBTITLING

This is how we proceeded 3

SUBTITLING

This is how we proceeded 4

We are currently in an intense test period
- especially with external people
- How do they deal with the editor?

- What can we improve in the division of the subtitles in order to increase readability?
- How can the mediahub be integrated into our existing system?
- What kind of content is most appropriate for this kind of subtitle creation?

- How could translation engines be integrated into the system?

THANK YOU FOR YOUR ATTENTION

Any questions?

Gion Linder
Head of Access Services
gion.linder@swisstxt.ch