REAL-TIME SUBTITLING BY RESPEAKING IN HBBTV

CONTENTS

0. PRE-RECORDED SUBTITLING

Pre-recorded subtitling 0

Pre-recorded subtitling, also known as off-line or pre-prepared subtitling, is the standard mode of subtitling. In this case, subtitles are prepared beforehand “with precise in and out timecodes, based on the speech in each time period” (EBU 2004:10). Increasingly, the respeaking technique is also used to produce this kind of subtitles, as we speak faster than we write, which means that producing pre-recorded subtitles through re-speaking can help save time and increase productivity.

1. SEMI-LIVE OR AS-LIVE SUBTITLING

Semi-live or as-live subtitling 1

It is often used for live programmes heavily scripted and include pre-edited segments such as interviews, archive footage or items from previous bulletins. In this case, “the subtitler creates a list of subtitles, without time-codes, and during transmission cues these manually in sync with the programme” (EBU 2004:10). Semi-live subtitling is often used in news programmes (when the scripts are available in advance), but also in other contexts such as the theatre and the opera.

2. REAL-TIME SUBTITLING

Real-time subtitling 2

It is, according to the European Broadcasting Union (EBU) (2004:10), the most ambitious type of subtitling, where the respeaker “seeks to under-stand the context of the programme in advance” and then, on transmission, “creates the subtitles in real-time”.

3. SPEECH RECOGNITION

Speaker-dependent speech recognition software is used for real-time subtitling by most TV stations worldwide. It could be described as a software licence that can be used only by one single user at a time and that 'learns' continually about the specific way in which a given user dictates.

Given the rapid evolution of speaker-independent speech recognition (also called automatic speech recognition: ASR), it makes sense to anticipate that subtitling companies will try to resort to this technology once it has reached optimum levels of accuracy. First of all, this software may be used with the intervention of a human operator, who can correct misrecognitions and errors of punctuation and speaker identification before sending the subtitles on air. And, perhaps in a more distant future, human intervention may be excluded from the process altogether.

EXAMPLE SPEECH RECOGNITION

4. RESPEAKING

A technique in which a respeaker listens to the original sound of a live programme or event and respeaks it, including punctuation marks and some specific features for the deaf and hard of hearing audience, to a speech recognition software, which turns the recognized utterances into subtitles displayed on the screen with the shortest possible delay (Romero-Fresco 2011). Although most of the subtitles broadcast on TV enable deaf and hard-of-hearing viewers to have access to the audiovisual content being broadcast, live respoken subtitles are not perfect, and they typically include recognition errors, edition errors and delay.

EXAMPLE RESPEAKING

5. RECOGNITION ERRORS

These are usually misrecognitions caused by mispronunciations/mishearing or by the specific technology used to produce the subtitles.

Example of recognition error (I)

Example of recognition error (II)

Example of recognition error (III)

Example of recognition error (IIII)

6. EDITION ERRORS

Edition errors 6 _____

Another category of errors are edition errors. They are the result of the respeaker's judgement or decision to, for example, omit a piece of information, often because the original speech rate is too fast. The following video features a fast speech rate. It is an example of why respeakers often find it impossible to deliver a verbatim rendition of what is being said and must therefore edit down the speaker's words while trying not to lose information. Compare it with the original speech and check how much has been edited and how much key information has been lost.

EXAMPLE EDITION ERRORS

7. DELAY

Delay 7 _____

Real-time subtitles are produced just moments before they can be transmitted. Consequently, they will always lag behind the original audio signal. This delay is perceived by deaf and hard-of-hearing viewers as their main disadvantage, as it makes it impossible to combine listening and viewing (e.g. lip reading). In the UK, live subtitles are displayed word-by-word in order to reduce the delay as much as possible:

In countries, such as Germany, Spain and Switzerland, live subtitles are displayed in blocks. Whereas this display mode increases delay, since the text is not transmitted automatically until the maximum number of characters in a subtitle row (or the maximum number of defined subtitle lines) has been filled, it is easier to read than subtitles displayed in scrolling mode. In other countries, such as Belgium, it is considered legal to delay the video and audio signal in order to resynchronize both of them with the live subtitles which are presented in block format matching the visual information.

EXAMPLE DELAY