[=]LTA
LiveTextAccess

UNIT 2

Element 1 – Learning outcome 3

TRANSCRIPT: ASSESSMENT MODELS

Table of contents

1   Slide 1

LiveTextAccess: Training for real-time intralingual subtitlers.

2   Slide 2

This is Unit 2. Linguistic competence. Element 1. Functionality which is made of accuracy, readability and legibility.

3   Slide 3

Accuracy assessment models created by UAB, SSLM and SDI.

4   Slide 4

The learning outcomes for this video lecture are to produce accurate transcriptions in terms of spelling, grammar and meaning and to produce legible and readable transcriptions both while creating the transcriptions, and after the live situation by applying readability and eligibility indicators.

5   Slide 5

Agenda. In this unit we will be dealing with the following points:

From quantity to quality, Accuracy key areas. Accuracy assessment models and we will finish with a Summary.

6   Slide 6

From quantity to quality.

7   Slide 7

According to the EFHOH "it is well and good having subtitles but if they are not accurate then they are of little or no use to the consumer". This is a particular problem with regards to live subtitling".

8   Slide 8

Accuracy key areas.

9   Slide 9

There are three key areas which will determine the accuracy of a real-time subtitle lexical, which means correct pronunciation and/or spelling. Grammar, which means correct syntactic structure of the unit and semantic, which means that the meaning of the source text is conveyed. Errors in these elements can be due to a human failure or due to a malfunctioning of the software and they may or may not affect the understanding of the meaning the speaker intends to deliver.

10   Slide 10

Accuracy assessment models.

11   Slide 11

To measure accuracy there are two main models, the NER model based on the Word Error Rate System and the IRA model which is a conceptual system. The NER model is especially effective when measuring the accuracy of verbatim subtitles because it's based on a World Error Rate system in which all utterances of the speaker are respoken or typed. Therefore the source text from the speaker can be compared with the output text of the subtitles.

12   Slide 12

The NER model, developed by Romero-Fresco in 2011 is a model to assess the accuracy by analysing the extent to which errors affect the coherence of the subtitle text or modify the content.

13   Slide 13

In the NER model accuracy is measured following a formula in which "N" is the total number of words in the subtitles, "E" are the addition errors and "R" the recognition errors. So "N" the total number of words of the subtitles minus "E" edition errors, minus "R" recognition errors will be divided by "N" and multiplied by 100. This will give us an accuracy rate saying how accurate the subtitles are but this is only an indication of the quality. The final rate should not be taken as an end result. There is also a section where correct additions can be included and an assessment can be made. In any case the target should not be less than 98%. This is the established threshold for subtitles that present an acceptable quality this model is being used and can be used in any type of subtitling techniques and working context as it is based on a word edition.

14   Slide 14

The IRA model is the idea units rendition assessment model, ideas are any sentence type conveying a meaning regardless of its degree of synthetic dependency. Idea units are divided in rendered and non-rendered idea units rendered idea units are those represented in the output text while non-rendered units are those in which a message is not conveyed due to omission of the unit itself. Non-rendered idea units are omissions which occur when the software fails to reproduce a conceptual unit. Misrepresentations are due to semantic errors which compromise the decoding of the source text. Rendered units are repetitions which means units which are highly comprehensible due to its faithful verbatim reproduction of the words added in the search text, and alterations, which means that the unit contains minor errors which however do not deteriorate the comprehensibility of the message for the end user.

15   Slide 15

There are two types of alterations expansions and reductions. Expansion occurs when a conceptual unit is explained or disambiguated using a higher number of characters than those used in the source text. Reduction occurs when a unit is characterized by partial omission and/or partial or total reformulation. There are two kinds of reduction, omission and compression.

16   Slide 16

Reduction of the original message is necessary for two main reasons. First to adapt the message to the reading speed of the viewers and secondly to avoid delay.

17   Slide 17

If we center the attention to alterations produced by reductions. Real time subtitlers should develop editing skills to reduce the message without losing key information.

18   Slide 18

Reduction relies on two main strategies omission which involves deleting redundant information and compression, which involves rewording in a shorter form. Omission is possible when the audience or when a speaker uses redundant words.

19   Slide 19

Main types of omission.

20   Slide 20

According to Romero-Fresco the main types of words omitted by live subtitlers in the case of redundant words are discourse markers: so, well, I mean, you know. Connectors: and, but, though. This means that respoken sentences are notably shorter than the source text sentences, given that the absence of these conjunctions often entail the beginning of a new sentence. Also intensifiers: really, much more, well. Or repetitions and unimportant asides: "you were saying" or "worth bearing in mind".

21   Slide 21

Compression is possible when a word or a group of words can be substituted by a shorter equivalent also long grammatical structures can be substituted by one word prepositions, conjunctions or verbs.

22   Slide 22

The main aim of a real-time interlingual subtitler should be to make sure that subtitles are clear, accurate and accessible for the user.

23   Slide 23

Summary.

24   Slide 24

As a summary we would like to stress the following points: The NER model assesses verbatim accuracy and the IRA model assesses sensatim accuracy. Output text will present alterations based on production techniques such as omission and compression and real time subtitlers have to make sure that subtitles are clear, accurate and accessible for the user.

25   Disclaimer, acknowledgement and copyright information

Authors: Pilar Orero, Estella Oncins

26   Additional metadata[1]

Title or file name of resource	U2_E1_LO3_Accuracy-assessment-models_TRANSCRIPT
Version	FINAL
Description of resource	Audio transcripts for the video lecture about assessment models
Keywords	Real-time intralingual subtitlers, respeakers, velotypists, verbatim, sensatim
Author/Contributor	UAB, SSML, SDI München
File type	Text
Length (pages, actual length of audio or video)	8 pages
Copyright holder	UAB, SSML, SDI München
Other related pedagogical assets	PowerPoint files and videos in the Youtube playlist of the unit
Related subject or skill area	Unit 2 Linguistic competence
Publisher	LiveTextAccess
Format	PDF file
Issued	22/01/2020
Language	English
Licence	CC BY-SA 4.0
Accessibility	PDF checked with TingTun
Certified by	ECQA: http://ecqa.org/

---

[1] The scheme used is an adaptation of the format provided by the EU project 2014-1-DE01-KA203-000679