



GLOBAL ALLIANCE
OF SPEECH-TO-TEXT CAPTIONING

Introducing the NER Evaluation Method Developed by Pablo Romero-Fresco

Who is Dr. Pablo Romero-Fresco?

Pablo Romero-Fresco is Ramon y Cajal researcher at Universidade de Vigo (Spain) and Honorary Professor of Translation and Filmmaking at the University of Roehampton (London, UK). He is the author of the book *Subtitling through Speech Recognition: Respeaking and Accessible Filmmaking*. He is the leader of the international research centre Galician Observatory for Media Access (GALMA), for which he is currently coordinating several international projects on media accessibility, including Interlingual Live Subtitling for Access (ILSA). Pablo is also a filmmaker. His first documentary, *Joining the Dots* (2012), was screened during the 69th Venice Film Festival and was used by Netflix as well as film schools around Europe to raise awareness about audio description.

Basic Requirements for a Model of Assessment

- Functional and easy to apply
- Includes the basic principles of Word Error Rate (WER) calculations in SR (Speech Recognition)
- Different programs, different editing
- Possibility of edited and yet accurate respeaking/voice/steno
- Compare subtitles (captions) with original spoken text
- Include other relevant info (delay, position, speed)

WHAT IS THE WER MODEL?

WER measures the number of words in a sentence that are:

- a. Inserted
- b. Deleted
- c. Replaced in the sentence

The next slide illustrates the explanation of the calculation and how it is applied to a sentence.

TRADITIONAL WER METHODS

Per US National Institute of Standards and Technology

$$\text{Accuracy Rate} = \frac{N - \text{Errors}}{N} \times 100 = \%$$

Where	is	the	whole	wheat	flour
		D	S	S	I
Where	is		hole	we	eat
					S
					flower

D represents deleted word

S represents mistranslated words

I represents inserted word

Broken down, the sentence “Where is **hole we eat flower**”

= Number of errors 5 divided by the number of words x 100 = %

What Does NER Stand for?

Number of words in the live captions

Edition errors

Recognition errors

Application of Natural Language Processing (NLP) that processes and understands large amounts of unstructured human language.

(Continued on next slide)

What is Natural Language Processing?

Natural language processing helps computers communicate with humans in their own language and scales other language-related tasks. For example, NLP makes it possible for computers to read text, hear speech, interpret it, measure sentiment and determine which parts are important.

Also known as entity identification, entity chunking and entity extraction.

NER Model Assessing Guidelines

$$\text{Accuracy} = \frac{N - E - R}{N} \times 100 = \%$$

Correct editions:
Assessment:

Target = 98%

Feedback from Consumers on Types of Errors

“There are errors, yes, but you can easily figure out what the correct word was meant to be. I’m deaf, not stupid.”

“Thanks to live subtitling, I’m now bilingual. I can speak English and teletext.”

“As far as I’m concerned, they are not errors but lies.”

TYPE OF ERRORS

Independent idea unit: Often a sentence which may be composed of several dependent idea units, and it makes sense as a full message.

Dependent idea unit: Often a complement, it provides information about the “who,” “what,” “when,” “where,” “how” of an independent idea unit.

EDITION ERRORS

Serious error (1.00 deduction) – An edition that causes change of meaning – (misinformation or considered a lie).

Standard error (0.50 deduction) – Omission of a full independent idea unit or omission of a dependent idea unit that renders the whole unit meaningless (nonsensical).

Minor error (0.25 deduction) – Omission of a dependent idea unit that doesn't render the whole unit meaningless or nonsensical.

MINOR RECOGNITION ERRORS

What great goal **by a** Ryan
Giggs!

For people are still missing
following Sunday's tornado.

What a great goal **by Ryan**
Giggs!

Four people are still missing
following Sunday's tornado.

STANDARD RECOGNITION ERRORS

He's a **buy you a bull**
asset.

He's a **valuable** asset.

Is it really **attend Tatian**?

Is it really **a temptation**?

SERIOUS RECOGNITION ERRORS

Public funding for universities
has been cut by **15%** this year.

Public funding for universities
has been cut by **50%** this year.

He never talks **dirty.**

He never talks **to Rudy.**

Who Uses the NER Model?

1. Adopted by the Canadian Governmental Regulator (CRTC) in the new guidelines.
2. Included in the official UNE subtitling guidelines in Spain.
3. Adopted by the Brazilian Regulator ABNT in its updated official guidelines, NBR15290.
4. Used by the British Governmental Regulator to assess subtitles in the UK.

Who Uses the NER Model?

1. Used by the BBC, ITV, Channel 4, Channel 5, Sky in the UK.
2. Used by subtitling companies for training and assessment purposes (Ericsson, Ai-Media, Deluxe, Verbavoice, SwissTXT).
3. Used by other regulators, broadcasters, subtitling companies and in other countries, such as France, Italy, Belgium, Holland, Poland, Austria, Germany and South Africa.

QUESTIONS?

Please feel free to contact us:

Chris Ales: chriales@capplus.co

Steve Clark: sclarkrt@gmail.com

Global Alliance: info@spechtotextcaptioning.org