GALMA
Galician Observatory
for Media Accessibility

Multimedia
Technology
Group (GTM)

Universida de Vigo

Automatic Galician Subtitles:
Towards the Creation of a Live Subtitling Tool

María Rico-Vázquez, Laura Docío-Fernández, Carmen García-Mateo

INTRODUCTION

Audiovisual Material

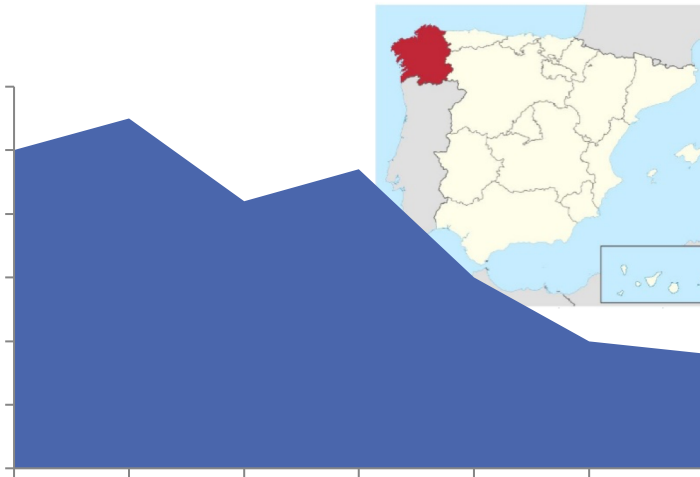**Live Subtitling**

Speech Recognition

Automation

GALICIA. GALICIAN LANGUAGE

- Few speakers (youth)

- Limited presence in mass media

- Vehicular language

Audiovisual Media Accessibility in Galicia

GALMA
Galician Observatory for Media Accessibility

Multimedia Technology Group (GTM)

»

Speech Recognition software for live subtitling in Galician

SR SOFTWARE. Functioning and Potential

EARLY STAGE

↓

NO punctuation

NO capitalization

NO subtitle format

Language Models

Software training

SR SOFTWARE. Materials and Assessment

21 samples

TELEVISIÓN DE GALICIA

7 samples: news

7 samples: weather

7 samples: sports

NER model

$$Accuracy = \frac{N - \cancel{X} - R}{N} \times 100$$

N = Number of words, E = Editing errors, R = Recognition errors

EXCELLENT	>99.5%
VERY GOOD	99%-99.5%
GOOD	98.5%-99%
ACCEPTABLE	98%-98.5%
POOR	<98%

SR SOFTWARE. (Initial) Results and Discussion

Errors

	🎤	🌧	🎾	Total
Minor	1733	2163	1970	5866
Standard	175	122	328	625
Serious	2	5	6	13
Total	1910	2290	2304	6504
Accuracy Rate (AR)	95.24%	94.23%	94.22%	**94.59%**

- Most errors would not affect viewers' comprehension
- Unclear speech creates errors that affect comprehension
- Few errors that omit information/misinform
- Average AR below quality threshold (NER)

Punctuation Errors Excluded

	🎤	🌧	🎾	Total
Minor	753	981	1049	2783
Standard	171	115	321	607
Serious	2	2	6	10
Total	926	1098	1376	3400
Accuracy Rate (AR)	97.49%	97.10%	96.25%	**96.95%**

- Punctuation errors represent a large percentage of the total
- Minor errors are significantly reduced
- More severe errors are slightly reduced
- Average AR improves considerably

Punctuation + Capitalization Errors Excluded

	🎤	🌧	🎾	Total
Minor	486	599	487	1572
Standard	171	114	320	605
Serious	2	2	6	10
Total	659	715	813	2187
Accuracy Rate (AR)	98.10%	98.00%	97.47%	**97.86%**

- Capital letter in proper names, place names, institutions, players or teams...
- Minimal changes in more severe errors
- Average AR almost reaches the quality threshold (NER)

SR SOFTWARE. Strengths and Weaknesses

- Most errors are **Minor**: they should not affect viewers' comprehension

  However, a large number of **Minor** errors in a sentence may hinder reading and understanding

- The number of **Serious** errors is low: meaning of original ideas hardly changes in written text

  However, quite a few errors (**Standard**) disrupt the flow/meaning of the original text and might cause surprise

- Quite **sensitive software**: most words are recognized

  However, not only words but **other sounds are (mis)recognized**

Recognition should be further improved

- Software **not** ready to be used for subtitling provision through **respeaking**

  - Needs to be activated for live use
  - Need to recognize punctuation marks and capital letters

  How far are we?

  Software may be used for subtitling provision with **human post-editing**

  - Punctuation and Capitalization must be implemented
  - Time to correct errors: ≈10 sec.
    - Total Errors: 52 min. post-editing (per 10-min. sample)
    - Without P. Errors: 27 min. post editing (per 10-min. sample)
    - Without P. + C. Errors: 17 min. post-editing (per 10-min. sample)

Final Thoughts

SR Software: **first attempt** to promote access to audiovisual media in Galicia through live subtitling.

The software has proved to be a good recognizer of pre-recorded files, but **some aspects must be improved** in order to generate high quality subtitles: punctuation, capitalization, amount of text displayed, recognition capabilities...

Subtitling provision through **Respeaking** is not feasible yet with this SR software.
Post-edition would be a possibility to start using it in real contexts if abovementioned aspects improve.

A **commitment to minority languages by subtitling providers** is essential.

More material is needed to feed and train the recognizer: a low presence of audiovisual and textual material in Galician limits the advance in terms of speech technology.

Further work is needed to contribute to the **promising future** of this SR software.

Thank you!

María Rico-Vázquez, Laura Docío-Fernández, Carmen García-Mateo

7th International Symposium on Live Subtitling and Accessibility 2020