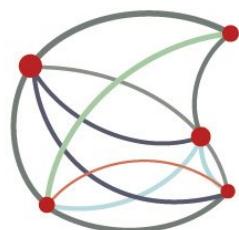


Integrating automatic subtitling and translation into MOOCs and OER repositories with UPV tools

Machine Learning and Language Processing research group

www.mllp.upv.es



MLLP

2nd World OER Congress



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

19 September 2017

Contents

The MLLP research group	3
EMMA: Reducing the effort for multilinguality	4
Transcription and translation of videos	5
Translation of course texts	6
MLLP Platform demo	7
Quality and time saving measurements	9
Conclusions	15



- Research group at *Univ. Politècnica de València* (Spain)

- **Recent projects**

- [2011–14] **transLectures**: Transcription and translation of video lectures (*EU FP7*)
- [2014–16] **EMMA**: European Multiple MOOC Aggregator (*EU CIP*)
- [2016–18] **MORE**: Multilingual Open Resources for Education (*Spanish gov.*)
- [2017–20] **X5gon**: Cross-{Modal,Cultural,Lingual,Domain,Site} Global OER Network (*H2020*)

- **Language technologies at the MLLP**

- Automatic Speech Recognition (subtitling):
 - * EN, ES, CA, FR, DE, IT, NL, PT, SL, ET
- Machine Translation (for subtitles, documents and websites):
 - * {ES, CA, FR, IT, NL, PT, ET} → EN
 - * EN → {ES, CA, FR, DE, IT, PT, SL}
- Text-to-Speech Synthesis:
 - * EN, ES, CA
 - * E.g., EN → {ES, CA} voice-to-voice translation

EMMA (2014 – 2016)

European Multiple MOOC Aggregator

<https://europeanmoocs.eu>

- **Main goal:** To provide multilingual access to European MOOCs
- **Motivation:**
 - Language barrier is keeping many learners from taking MOOCs
 - **EMMA** uses **transLectures** tools to translate videos and texts
- **Cost of manually translating MOOCs:**
 - Example: 6-week course, 2 hours of video, 75,000 words
 - * Videos:
 - 1) Manual transcription of the videos (10 RTF)
 - 2) Translation of the videos (30 RTF)
 - Effort for 2 hours of video: 0.5 PM
 - * Texts:
 - Manual translation rate: approx. 2500 words per day
 - Effort for 75,000 words: 1.5 PM
 - Solution to lower costs:
 - * ASR and MT: effort can be reduced by over 50% (2 PM → 0.6 PM)

Transcription and translation of videos

1. Generation of **automatic transcriptions** (subtitles) from video

2. Manual review of automatic transcriptions

The screenshot shows a video player interface for a lecture titled "S2V1 Arômes du vin". The video frame displays a man speaking with the subtitle "Le vocabulaire des arômes" overlaid. Below the video, a transcript is shown with several lines of text and their corresponding timestamps. The transcript discusses the complexity of wine aromas and how they reflect chemical complexity.

Text (French)	Text (English)	Timestamp (s)
impact sur l'arôme du vin. Ces composés d'arômes	impact on the aroma of the wine. These aromatic compounds	12.1 cps
sont acheminés par l'air inspiré vers les récepteurs olfactifs,	are carried, by the air you breathe in, to the olfactory receptors,	12.4 cps
soit par la voie orthonasale, soit par la voie rétronasale.	either by the orthonasal pathway, or by the retrorhinal pathway.	11.8 cps
La complexité du vocabulaire des arômes du vin	The complexity of the vocabulary of wine aromas	16.0 cps
est à l'image de leur complexité chimique.	is the reflection of their chemical complexity.	15.8 cps
Afin de simplifier un tel lexique, il est courant d'organiser le	In order to simplify such a lexicon, it is common to organize the	16.7 cps
descripteur en familles plus ou moins précises.	descriptors into more or less accurate families.	17.5 cps
Ces lexiques se présentent en général sous la forme d'une roue	These glossaries are generally presented in the form of a wheel	16.2 cps
des arômes où les familles sont organisées en trois niveaux.	of aromas, where the families are organized into three levels.	15.5 cps
Un premier niveau assez général, un deuxième niveau un peu plus spécifique	A first, rather general level, a second, slightly more specific level.	15.7 cps

3. Generation of **automatic translations** from transcriptions

4. Manual review of automatic translations

This screenshot shows the same video and transcript as the previous one, but with the English translation applied to the French text. The English text is displayed in blue boxes at the bottom of the transcript, indicating where the translation has been made. The timestamps remain the same as in the original transcript.

Text (French)	Text (English)	Timestamp (s)
impact sur l'arôme du vin. Ces composés d'arômes	impact on the aroma of the wine. These aromatic compounds	12.1 cps
sont acheminés par l'air inspiré vers les récepteurs olfactifs,	are carried, by the air you breathe in, to the olfactory receptors,	12.4 cps
soit par la voie orthonasale, soit par la voie rétronasale.	either by the orthonasal pathway, or by the retrorhinal pathway.	11.8 cps
La complexité du vocabulaire des arômes du vin	The complexity of the vocabulary of wine aromas	16.0 cps
est à l'image de leur complexité chimique.	is the reflection of their chemical complexity.	15.8 cps
Afin de simplifier un tel lexique, il est courant d'organiser le	In order to simplify such a lexicon, it is common to organize the	16.7 cps
descripteur en familles plus ou moins précises.	descriptors into more or less accurate families.	17.5 cps
Ces lexiques se présentent en général sous la forme d'une roue	These glossaries are generally presented in the form of a wheel	16.2 cps
des arômes où les familles sont organisées en trois niveaux.	of aromas, where the families are organized into three levels.	15.5 cps
Un premier niveau assez général, un deuxième niveau un peu plus spécifique	A first, rather general level, a second, slightly more specific level.	15.7 cps

Translation of course texts

- Course text is ingested into machine translation system
- Advanced web interface:
 - Review source and target texts side-by-side
 - Preview of source and target documents

The screenshot shows a web-based machine translation interface. At the top, there are three buttons: 'View HTML (original)', 'Edit translation' (with a checked checkbox), and 'View HTML (translation)'. Below these buttons is a 'Hide HTML tags' link. The main area is divided into two columns: 'Original (es)' on the left and 'Translation (en)' on the right. A vertical 'Status' column is on the far right.

Original (es)	Translation (en)	Status
1 En cualquiera de los casos en que introduzcamos valores de tipo texto en las distintas celdas vemos que éstos se nos ajustan por defecto:	In any of the cases in which we introduce text type values in the different cells we see that they are adjusted by default:	
2 A la izquierda.	To the left.	
3 A la derecha.	To the right.	

MLLP Platform demo

<https://ttp.mllp.upv.es/channel/demo>

The screenshot shows the MLLP Platform demo interface. On the left, there is a sidebar with navigation links for VIDEO TRANSCRIPTION, ONLINE TRANSCRIPTION, DOCUMENT TRANSLATION, and MY ACCOUNT. The main area displays a grid of video thumbnails, each with its title, duration, and subtitle language options (en, fr, it, ca, sl, es). The titles include "MLLP Multilingual Web Translation Demo", "EIT Digital: Driving Europe's Digital Transformation", "Introducing the EIT Digital Master School - Application 2...", "co2logic", "Introducing the EIT Digital Master School - Application 2...", "EIT Digital: Driving Europe's Digital Transformation", "Vilnius", "Prague", "Introduction to gender statistics", "Administrative data collection on gender-based violence", and "Promo_Pan_European_Seal_Foca". At the bottom, there is a navigation bar with page numbers (1, 2, 3, 4, 5) and a footer containing the UPV logo, the text "UNIVERSITAT POLITÈCNICA DE VALÈNCIA", the DSIC logo, and the copyright notice "© MLLP 2016".

<https://media.upv.es/#/portal/video/641465a3-eadb-0545-9dc4-9222052190a0>



Subtitle options at the bottom right
("speech balloon" icon).

Quality and time saving measurements

Video transcription

Language	Word Error Rate		ASR time saving w.r.t. 10 RTF
	EMMA	YouTube Δ	
Spanish	15	+ 52%	– 67%
Estonian	27	N/A	N/A
English	39	+ 41%	– 47%
French	21	+ 55%	– 33%
Italian	17	+ 85%	– 61%
Dutch	25	+ 68%	– 42%
Portuguese	43	+ 31%	N/A
Average	27	+ 55%	–38%

Video translation

Language pair	Quality (BLEU)			MT time saving w.r.t. 30 RTF
	EMMA	Google	Δ	
English → Spanish	43	– 10%		– 74%
English → Italian	53	– 16%		– 68%
Spanish → English	41	– 15%		– 70%
Italian → English	48	– 22%		– 62%
Estonian → English	28	– 5%		N/A
French → English	33	– 7%		– 23%
Dutch → English	42	– 18%		– 68%
Portug. → English	56	– 18%		N/A
Average	43	– 14%		– 61%

Improving automatic video transcription

- Applying state-of-the-art techniques:
 - BLSTM Recurrent Neural Networks
 - Cross-word triphone modelling
 - Multilingual language models
- More speech data for system training

Language	Data (hours)		Word Error Rate			RTF (Processing)	
	2016	Sept 2017	2016	Sept 2017	Δ	2016	Sept 2017
Spanish	200	800	15.0	11.7	- 22%	1.5	0.8
Catalan	50	2430	17.8	16.6	- 7%	1.2	0.7
English	430	2550	26.7	22.2	- 17%	2.0	1.5
German	300	880	27.0	22.5	- 17%	1.2	1.0
French	80	670	21.3	16.5	- 23%	2.1	1.0
Italian	130	740	27.8	25.4	- 9%	0.7	0.7
Dutch	300	620	27.5	26.7	- 3%	1.5	0.9
Portuguese	120	530	34.4	23.3	- 32%	1.0	1.0
Slovene	90	90	38.6	33.9	- 12%	1.9	1.1
Average	190	1030	26.2	22.1	- 16%	1.4	0.9

Integration of the tools: Web Service + API

<http://ttp.mllp.upv.es/doc/>



2. VIDEO TRANSCRIPTION REQUEST

After clicking on “Save”, EMMA recognizes the number of videos uploaded on that page and Teacher (or co-author) can request a video transcription for each video by clicking on the specific button

Course title
Transcription workflow

Choose the title for your courses and remember that this is what the search engines will pick up on and the first step to engaging students.

Course structure
Overview Learning objectives Outcomes

EMMA MOOC: Cultural Heritage, Social Innovation and New Citizenship

TRANSCRIPTION AND TRANSLATION ON EMMA

Edit translation: English
select language

To be published
22-10-2015

Closing date
22-11-2015

StartEndDate
Self-Paced

Coming soon course

Archived

Transcriptions
Ask for transcription of video 1

Course Image Cover
Set a cover image for this course through your Media Library. Please use only .jpg or .png square images.

Open Media Library

Coauthors

A red arrow points from the 'Ask for transcription of video 1' button to the 'Transcriptions' section.

MLLP Platform

<http://ttp.mllp.upv.es>

The screenshot shows the MLLP Platform interface. On the left, there is a sidebar with navigation links: VIDEO TRANSCRIPTION (My videos, Upload media, Statistics), ONLINE TRANSCRIPTION (Dictation (Beta)), DOCUMENT TRANSLATION (My documents, Upload document, Web translation), and MY ACCOUNT (My account, My groups, FAQ, API, Log out). The main area displays a grid of video thumbnails, each with subtitle language options (e.g., en, fr, it, ca, sl, es) and a timestamp. The videos include:

- MLLP Multilingual Web Translation Demo (03:59, Jul 21, 2016)
- EIT Digital: Driving Europe's Digital Transformation (01:41, Jun 28, 2016)
- Introducing the EIT Digital Master School - Application 2... (01:25, Jun 28, 2016)
- co2logic (03:58, Jun 27, 2016)
- Introducing the EIT Digital Master School - Application 2... (01:25, Jun 8, 2016)
- EIT Digital: Driving Europe's Digital Transformation (01:41, Jun 7, 2016)
- Vilnius (04:00:02, Feb 27, 2016)
- Prag (20:16, Feb 18, 2016)
- Introduction to gender statistics (03:20, Feb 18, 2016)
- Administrative data collection on gender-based violence (02:40, Feb 18, 2016)
- Promo_Pan_European_Seal_Foca (03:04, Feb 18, 2016)

At the bottom, there is a navigation bar with page numbers (1, 2, 3, 4, 5) and a footer containing the University of Valencia logo, the text "UNIVERSITAT POLITÈCNICA DE VALÈNCIA", the DSIC logo, and the copyright notice "© MLLP 2016".

MLLP experiences

- UPV Mèdia / Polimèdia
 - Over 18,000 videos subtitled: Spanish, English, Catalan, French...
 - Over 3600 hours total
- edX & UPV[X]: 45 UPV MOOCs on edX
- EMMA: Full MOOCs (video+text) from 8 universities (7 countries, 8 languages)

- UPV Mèdia / Polimèdia
 - Over 18,000 videos subtitled: Spanish, English, Catalan, French...
 - Over 3600 hours total
- edX & UPV[X]: 45 UPV MOOCs on edX
- EMMA: Full MOOCs (video+text) from 8 universities (7 countries, 8 languages)
- **Service contracts:**
 - Univ. Carlos III de Madrid (2014–)
 - VideoLectures.NET (2011–2014; 2016–)
 - EU CdT: Translation Centre for the Bodies of the European Union (2015–2016)
 - Open Universiteit Nederland (2017)
 - Univ. de Lisboa (2017–)

Conclusions

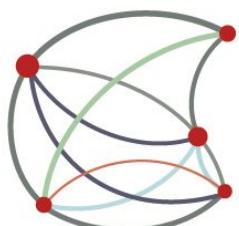
<https://tpp.mllp.upv.es>

- Multilingual access **boosts visibility** (in EMMA: +66% enrolled students)
- The cost of manually translating courses and OER is high
- Automatic transcription + translation can **reduce effort** by 50%
- Accuracy of automatic transcription and translation depends on several factors:
 - Languages, recording conditions, specificity of the course, use of course-related annotated data resources
- Designing multilingual courses and OER should also take into account:
 - Slides, images, application interfaces (demos), bibliography...
- **Current and future work** (MORE, X5gon...):
 - Improving auto transcription accuracy for noisy recordings (i.e., classroom recordings)
 - Improving machine translation with new Neural Machine Translation techniques
 - Text-to-speech technology for “auto dubbing” of video lectures
 - Multilingual website translation
 - Live translation in forums
 - Virtual multilingual teaching assistant

Integrating automatic subtitling and translation into MOOCs and OER repositories with UPV tools

Machine Learning and Language Processing research group

www.mllp.upv.es



MLLP



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

2nd World OER Congress

19 September 2017