

A Semi-Automated Live Interlingual Communication Workflow Featuring Intralingual Respeaking: Evaluation and Benchmarking

Language Resources and Evaluation Conference (LREC), June 2022

Tomasz Korybski, Elena Davitti,
Constantin Orasan, Sabine Braun



Research
England



MATRIC (*Machine Translation and Respeaking in Interlingual Communication*)

- Comparing output of different workflows for **real-time, interlingual communication**

Product: live text (subtitles or transcript) in a different language

Process: different methods and workflows to achieve same output

Human-centric



Interlingual Respeaking

SMART

Simultaneous Interpreting +
Intralingual Respeaking

Semi-Automated



Intralingual Respeaking +
Machine translation

MATRIC

Simultaneous Interpreting +
Automatic Speech Recognition

Fully Automated

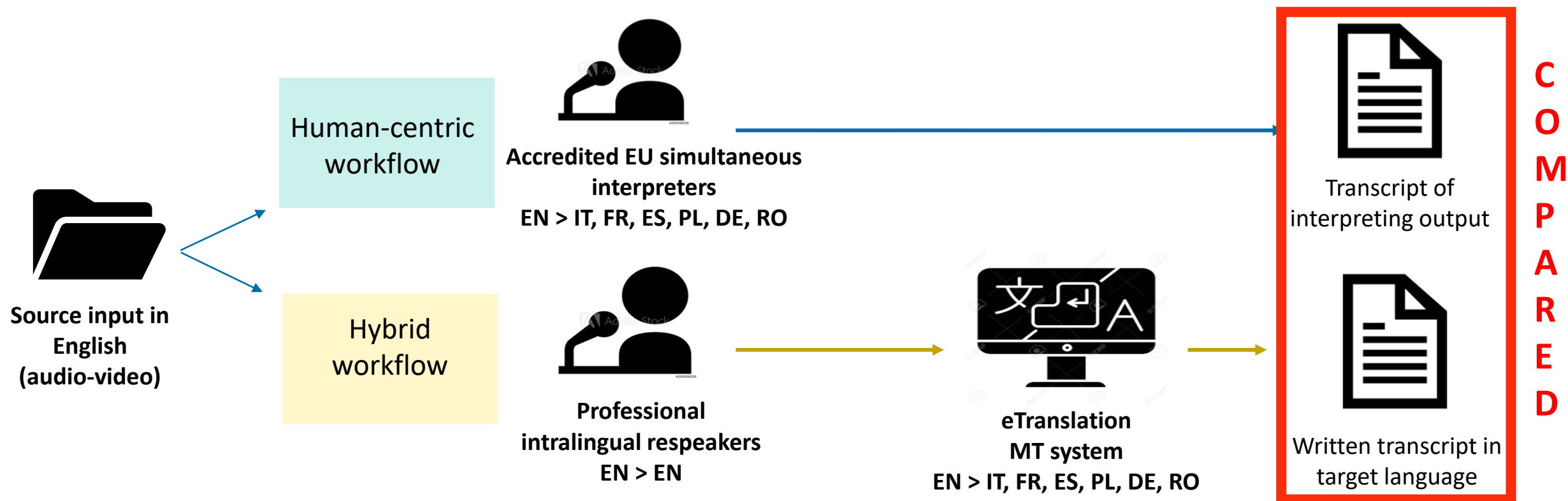


Automatic Speech Recognition +
Machine Translation

Why implement this project?

- Both ASR and MT technologies have made **huge advances**, but there are still **challenges** for them
- Both ‘struggle to **generalize across use cases**’ (Aksënova et al., 2021)
- ASR: the challenges of **ambient noise** and **speaker / environment variability**
- NMT: poor **out-of-domain performance**, problems with **rare word** translation / **long sentence translation, terminological consistency**
- These and other challenges mean that **full automation in live interlingual communication** is still burdened with many risks
- **Multimodal** communication is **understudied**

MATRIC workflows explained



MATRIC evaluation scoring grid

NTR SCORING												Speech03: [title]		VERBATIM SCORING						
NTR SCORING												Source transcript (EN)	Target (INTE output)	TRANSLATION - CONTENT (omission, addition, substitution)			TRANSLATION - FORM (style, correctness)			
NTR SCORING														Effective editions	MinT(Cont-errors)	MajT(Cont-errors)	CritT(Cont-errors)	MinT(Form-errors)	MajT(Form-errors)	
1	0.00												Right so dear Chair dear Pascal Honourable members ladies and gentlemen	Buenas tardes, siento retraso.						
2	0.00	1											First of all good afternoon and I apologise about the connection issues	Los interpretes también tenemos problemas, la conexión de la Comisaría no se escucha perfectamente, haremos lo que podamos.	The interpreter mentions that the audio quality is not very good, and that she/he will do her/his best.					
3	-0.25					1							One month ago when I was here with you again on the impact discussing the impact of covid 19 we said at the time the better we will see better days ahead	Bien, hace unos días hablábamos del impacto de esta pandemia y decíamos que habría días mejores y sigo convencida de ello	S - "a month ago" is conveyed as "a few days ago". Minor.					
4	-0.50							1			1		And I still believe that very firmly Today we are one week on from the Commission's adoption of the ambitious Farm to Fork strategy	Bueno, esta esa estrategia que aprobamos, son ambiciosa, la de la granja al plato y luego ese programa sanitario EU for health, que es aún más ambicioso.	□ - The idea of the strategy being adopted a week ago is not in the target speech. This might no be an issue. □ - "And a day later" is also missing, but it doesn't affect the general message.				ST - The reformulation here could be a bit clearer. C - "son ambiciosa" should be "es ambiciosa"	
5													And a day later on the adoption of the new stand-alone and massively more ambitious health programme EU for Health as we have called it							
6													It's part of the historic economy package which I must say we are beginning to look forward to	Forma parte de ese paquete económico histórico.						

MATRIC evaluation scoring grid

NTR SCORING														Speech02: [title]		VERBATIM SCORING					
NTR reductions														Source transcript (EN)	Target (INTE output)	TRANSLATION - CONTENT (omission, addition, substitution)			TRANSLATION - FORM (style, correctness)		
EE	MajT(Corr-OMISS)	MinT(Corr-OMISS)	CyrT(Corr-OMISS)	MajT(Corr-ADD)	MinT(Corr-ADD)	CyrT(Corr-ADD)	MajT(Corr-SUBS)	MinT(Corr-SUBS)	CyrT(Corr-SUBS)	MajT(ForM-STYLE)	MinT(ForM-STYLE)	MajT(ForM-CORR)	MinT(ForM-CORR)	Effective editions	MinT(Cont-errors)	MajT(Cont-errors)	CritT(Cont-errors)	MinT(For-errors)	MajT(For-errors)		
-0.75							1			1			1								
11															S- unique animals replaced with primera vez (...)				sentences are not well liked with the previous part nor between them "pérdidas de viviendas, pérdidas de vidas". ST - "y es posible 1000"		
12																					
13																					
14																					
-1.50	3						1			1					O - "Beyond" is missing, which has an impact in the structure of the sentence.						

MATRIC initial findings

Table 2: NTR scores for each speech

	SPEECH 1		SPEECH 2		SPEECH 3	
	SIM INTE	HYBRID	SIM INTE	HYBRID	SIM INTE	HYBRID
Spanish	98.6%	98.8%	98.2%	98.7%	98.2%	99.2%
Italian	98%	98.8%	99.3%	98.6%	99.4%	98.2%
French	99.5%	99%	99.3%	99.3%	99.7%	99.3%
Polish	99%	98.7%	98.7%	98.5%	98.4%	98.7%

What's behind the scores?

- More **omissions** in SIM INTE outputs, frequently of minor severity
- More **effective editions** in SIM INTE outputs, driving INTE output success

Source (EN)	Target (FR)	EE
Right so dear Chair dear Pascal Honourable members ladies and gentlemen.	Cher Pascal, chers députés, mesdames et Messieurs <i>(Dear Pascal, dear members, ladies and gentlemen.)</i>	Omission of oral spoken marker 'right so'
Source (EN)	Target (ES)	EE
And I believe that both Farm to Fork in their own ways and of course EU for Health gives us this opportunity	Yo creo que estas dos estrategias nos dan la oportunidad para ello. <i>(I believe that these two strategies give us the opportunity for it.)</i>	Proper names replaced and summarised by 'these two strategies' which is clear on the basis of context.

Limitations, challenges and next steps

- Larger-scale research with different genres necessary
- Cross-comparison with a different evaluation approach
- Comparison with a fully-automated workflow
- Integration of other MT solutions in the workflow
- A positive outcome: establishing a replicable methodological approach

Acknowledgements

MATRIC Evaluators:

- Francesco Saina
- Michal Gornik
- Megan Stockwell
- Eimee Brown
- Eloy Rodriguez Gonzalez

References

Romero-Fresco, P. and F. Pöchhacker (2017) “Quality assessment in interlingual live subtitling: the NTR model.” *Linguistica Antverpiensia, New Series: Themes in Translation Studies* 16: 149-167.

Eugeni, C. (2020). Human-Computer Interaction in Diamesic Translation. Multilingual Live Subtitling. In Dejica, D, Eugeni, C and Dejica-Carțiș, A, (eds.) *Translation Studies and Information Technology - New Pathways for Researchers, Teachers and Professionals*. Translation Studies Series. Editura Politehnica: Politehnica University Timișoara, (pp. 19-31).

Fantinuoli, C., Prandi, B. (2021) Towards the evaluation of automatic simultaneous speech translation from a communicative perspective - arXiv preprint arXiv:2103.08364, doi: 10.48550/arXiv.2103.08364, 16 September 2021

Thank you for your attention!



E.Davitti@surrey.ac.uk
T.Korybski@surrey.ac.uk