Engaging Content
Engaging People

Machine Translation for Subtitling – a way to improve access to MOOCs

Ke Hu, Sharon O'Brien, Dorothy Kenny

ADAPT Centre, SALIS, Dublin City University

Ireland's European Structural and Investment Funds Programmes 2014-2020
Co-funded by the Irish Government and the European Union
European Union
European Regional Development Fund
SFI Science Foundation Ireland

The ADAPT Centre is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

Overview

- Why MT?
- Research question & hypothesis
- Methodologies
  - Reception model
- Experiment
  - Participants
  - Procedure
  - Results
- Conclusion

Why MT?

Removing the language obstacle → Creating translated subtitles → Fansubbing

Large quantity

Considerable demand

Machine Translation!!

Research question and hypothesis

Main research question:

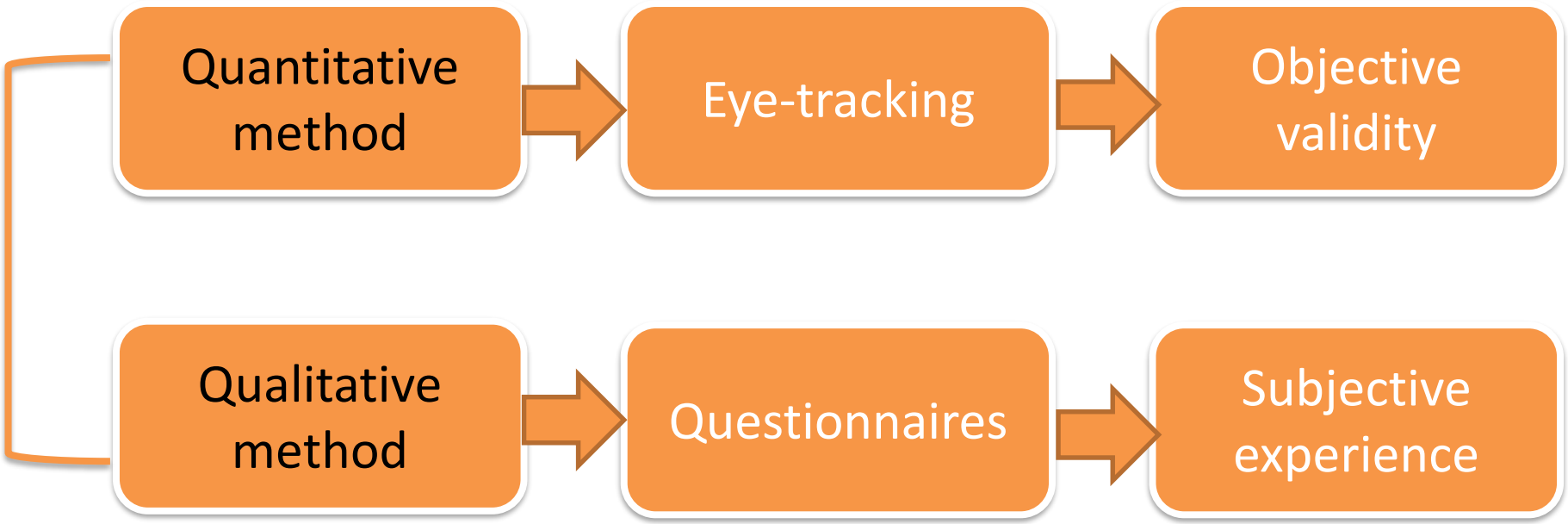
Is there a difference in reception between participants who are offered raw MT subtitles and those who are offered full PEMT and HT subtitles?

Main hypothesis:

Participants who are offered full PEMT subtitles and HT subtitles will score higher on our reception metrics compared with those who are offered raw MT subtitles.

Methodologies

A mixed-methods approach



Reception model

Element	Related to	Reflected in	Measured by
Response	Perceptual decoding	Attentional processes	Eye-tracking
Reaction	Psycho-cognitive issue	Processing effort and comprehension	Eye-tracking and comprehension testing
Repercussion	Attitudinal issues and sociocultural dimensions	Attitudes and beliefs	Attitude questions

Based on Gambier's model (Gambier, 2009)

Experiment

- Two Chinese universities , October 2017
- Video: “What is physical activity?” (6”59’) under the MOOC “Sit Less, Get Active” on Coursera.
- MT system: Google Translate (EN-ZH)
- Three versions of subtitles: RAW vs. PEMT vs. HT

Raw MT subtitles
(BLEU: 42.05%)

Full PEMT subtitles
(TER: 19.69%)

Participants

- MOOCs: university students, 18-25 years old
- China: 55.96 out of 100, English Proficiency Index 2017 by EF

- Ideal participants: Chinese undergraduates with low English level

- 66 participants (three groups)

Gender	34 male, 32 female
Age	18-23
Education	All undergraduates
English level	Mean=14.72 (full score 25) -> Level B1 (threshold)

Procedure

Step 1: Pre-recruitment questionnaire & Online English test (Cambridge English Language Assessment)

Step 2: Watching MOOC video with eye-tracker (SMI REDn Scientific)

Step 3: Post-task questionnaire: comprehension testing (multiple choice) and attitude survey (five-point Likert scale)

Attitude survey results per group

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
Group PE (24)	33.04% (114)	49.28% (170)	13.04% (45)	4.35% (15)	0.29% (1)
Group RAW (22)	17.21% (53)	57.79% (178)	18.51% (57)	6.49% (20)	0
Group HT (15)	27.05% (56)	39.61% (82)	24.16% (50)	9.18% (19)	0

The attitude of Group PE was overall better than Group RAW and Group HT.

Comprehension testing score per group

	Max	Min	Mode	Mean	SD
Group PE (24)	13	6	10, 11	9.58	1.74
Group RAW (22)	11	6	8	8.55	1.37
Group HT (15)	12	6	9, 10, 11	9.47	1.85

Group PE performed the best.

- **ANOVA**: there is a statistically significant difference between the comprehension testing score of the three groups.

- **LSD test**

Comparison 1: Group PE vs. Group Raw: significantly different

Comparison 2: Group Raw vs. Group HT: not significantly different

Comparison 3: Group HT vs. Group PE: not significantly different

Eye-tracking analysis

		AOI	Visit Count	Fixation Count	Visit Duration	Mean Fixation Duration
ANOVA	Raw vs. PE	SUB	N	N	N	Y
		IMA	N	N	N	N
	Raw vs. HT	SUB	N	Y	Y	Y
		IMA	N	N	N	Y
	PE vs. HT	SUB	N	Y	N	N
		IMA	N	N	N	N
Means	RAW vs. PE & HT	SUB	> √	> x	> √	> √
		IMA	< x	< √	< √	> x
	PE vs. HT	SUB	> √	> √	> √	> x
		IMA	< x	< x	< √	< x

"Y" means there is a statistically significant difference between the two groups.
"N" means the opposite.
"√" means the result supports the corresponding hypothesis.
"x" means the result does not support the corresponding hypothesis.

Conclusion

Participants who were offered full PEMT subtitles scored almost the same on our reception metrics as those who were offered HT subtitles.

RAW vs. PE ?
RAW vs. HT ?

Premise: the quality of full PEMT subtitles was lower than HT subtitles

??

-> Quality assessment
-> Frequency analysis

Thank you for your attention!

Ke Hu: ke.hu2@mail.dcu.ie
Sharon O'Brien: sharon.obrien@dcu.ie
Dorothy Kenny: dorothy.kenny@dcu.ie