

A Rising Tide Lifts All Boats?

Quality Correlation between Human Translation (HT) and Machine Assisted Translation (MAT)

EVELYN YANG GARLAND, CT RONY GAO, CT

Introduction

> Does the human who produces the best translation without MT also produce the best translation with the assistance of MT?

>Are translation and post-editing completely different skills?

>Is "human + machine" always better than machine alone in terms of quality?



This Photo by Unknown Author is licensed under CC BY-SA-NC

Methodology

- Main hypothesis: positive correlation between HT quality and MAT quality
- Subjects: 8 volunteers (English-to-Chinese translation practice group)
- Source Texts: two 250-word passages in English similar in style and difficulty level
 - Passage A. Human Translation (HT): MT tools NOT allowed
 - Passage B. Machine-Assisted Translation (MAT): one MT version provided as reference; all other MT tools allowed
- Quality Evaluation: ATA grading framework

Results 1: Correlation between HT and MAT?



| Descriptive Statistics | Passage A (HT) | Passage B (MAT) | p-value |
|------------------------|-------------------|--------------------|---------|
| Mean | 21.8 | 27.9 | 0.025 |
| Median | 20.5 | 26.3 | |
| Range | 37.5 | 32.0 | |
| Standard Deviation | 10.9 | 9.2 | |
| N | 8 | 8 | |
| Correlation | | | |
| | | t-stat | p-value |
| Pearson | 0.85 | 3.99 | 0.007 |

Proceedings of the 18th Biennial Machine Translation Summit, Virtual USA, August 16 - 20, 2021, Volume 2: MT Users and Providers Track

Results 2: How do MAT and raw MT compare?



Proceedings of the 18th Biennial Machine Translation Summit, Virtual USA, August 16 - 20, 2021, Volume 2: MT Users and Providers Track

Discussion 1

Does the human who produces the best translation without MT also produce the best translation with the assistance of MT?

> Are translation and post-editing completely different skills?

>Is "human + machine" always better than machine alone in terms of quality?





Discussion 2



ΗT

Limitations

➤Small sample size

- Uncontrolled before-and-after study
- >No personal or professional information collected

>Lack of empirical data to confirm the difficulty levels of the two passages

>Only one evaluation criterion: quality score under the ATA grading framework

Time, productivity, or cost not measured

Acknowledgement

- American Translators Association (ATA)
- Chinese Language Division of ATA
- ➢ Jessie Lu, Tianlu Redmon, Larry Bogoslaw & Geoffrey Koby
- Most importantly, translators and evaluators who most generously donated their time and expertise