

The PASSAGE project: Standard German Subtitling of Swiss German TV content

Pierrette Bouillon, Johanna Gerlach, Jonathan Mutal, Marianne Starlander

✉ {pierrette.bouillon, johanna.gerlach, jonathan.mutal, marianne.starlander}@unige.ch

🏛️ Faculty of translation and interpreting, University of Geneva

SCAN ME



Motivation

Swiss German

- A collective term for all the regional dialects
- Spoken everyday language
- No standard written form

60% of the population speak Swiss German



Swiss TV

- Swiss German is mainly used in Swiss TV
- Contents in Swiss German are subtitled into Standard German



Systems

Neural machine translation (NMT)
transformer architecture with copy attention

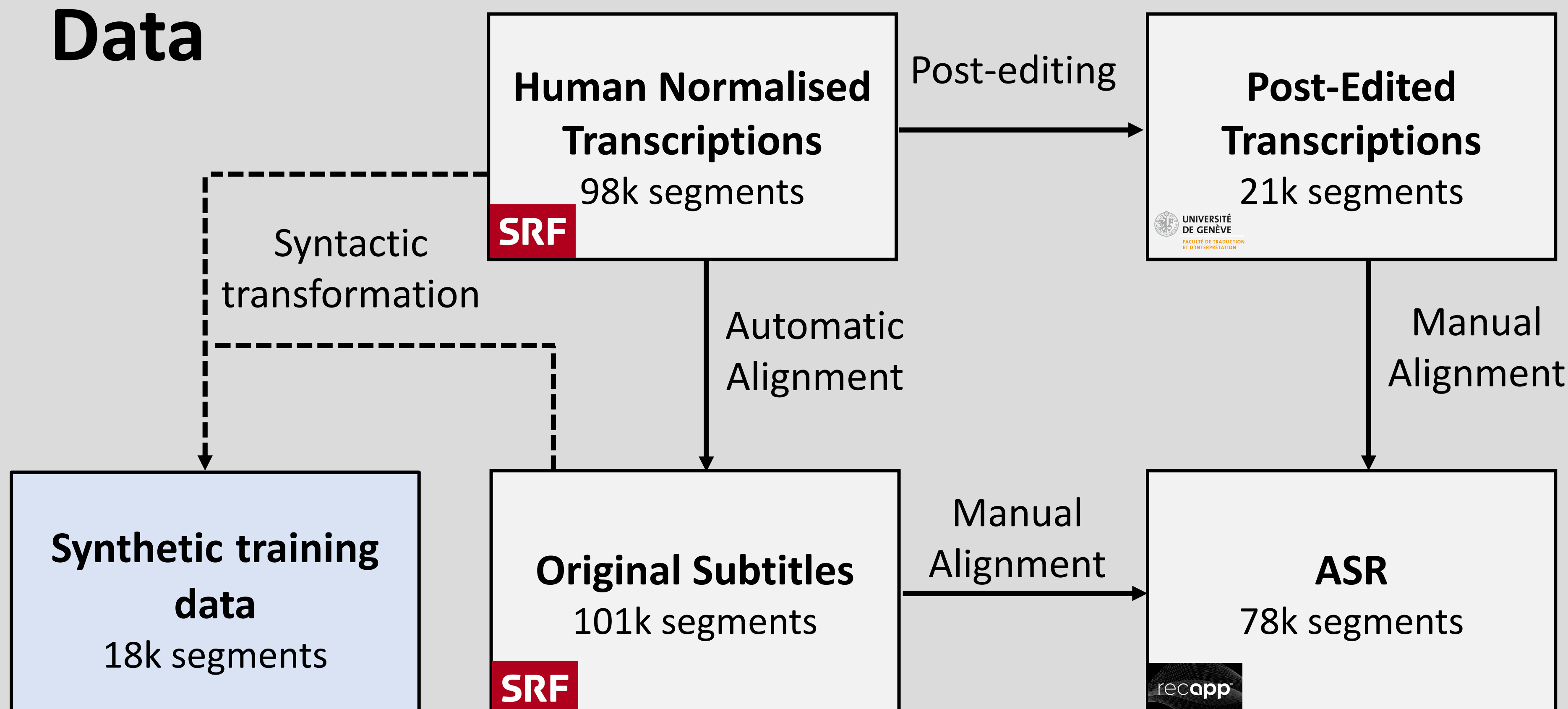
Automatic Post-Editing (APE)
system predicts the type of edits instead of the word

Test Data

Human Normalised Transcriptions
-
Post-Edited Transcriptions

Test Suite
50 ASR segments with syntactic phenomena

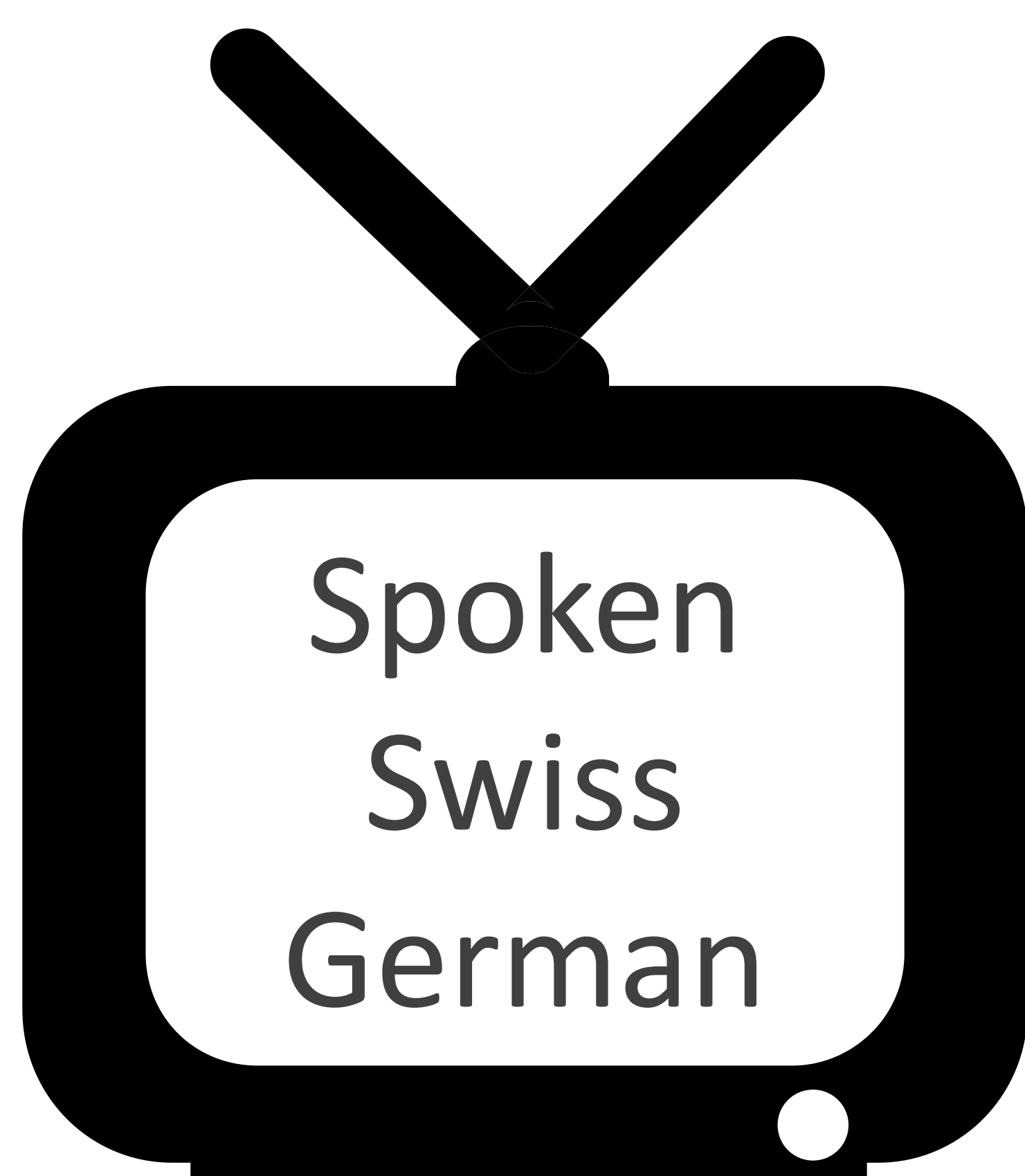
Data



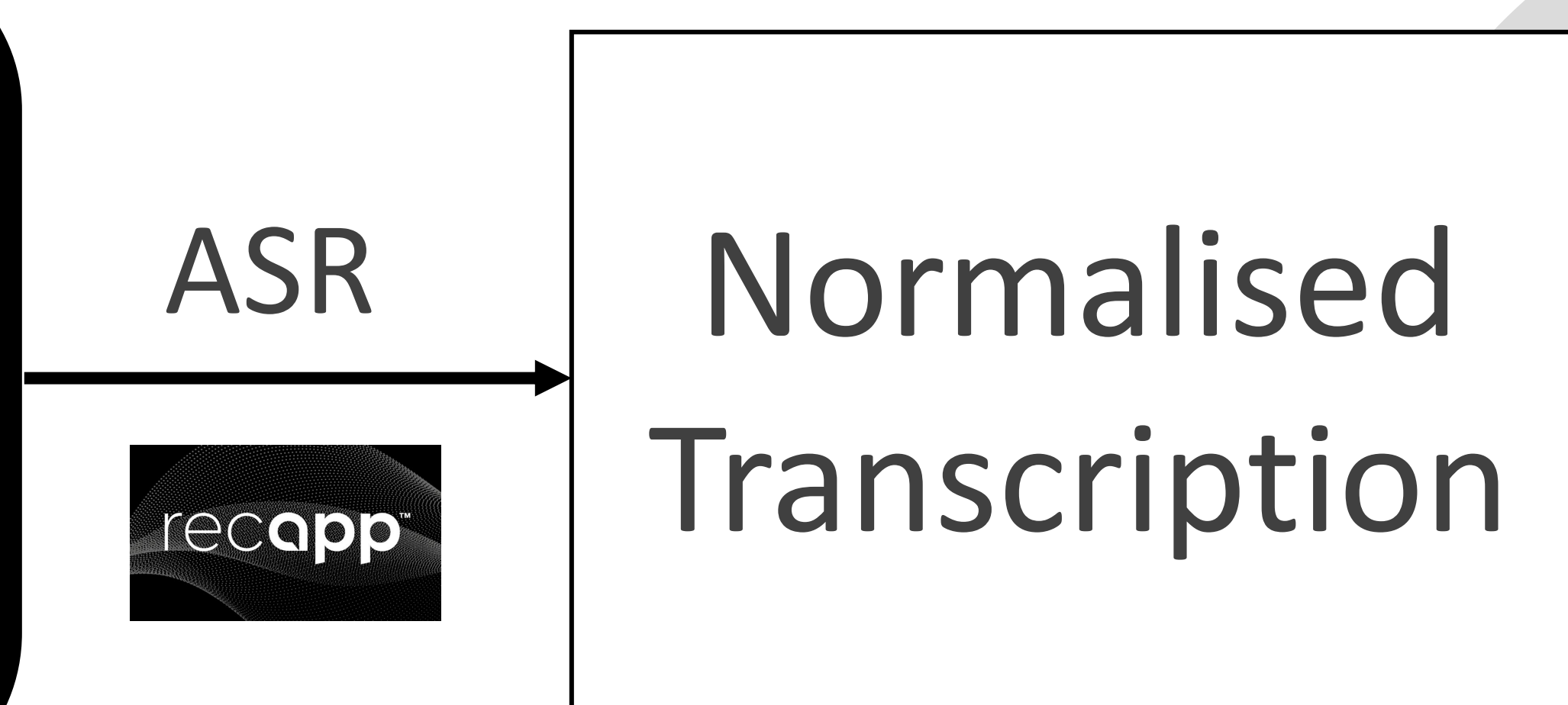
Results

For NMT, the addition of targeted synthetic training data improves the result

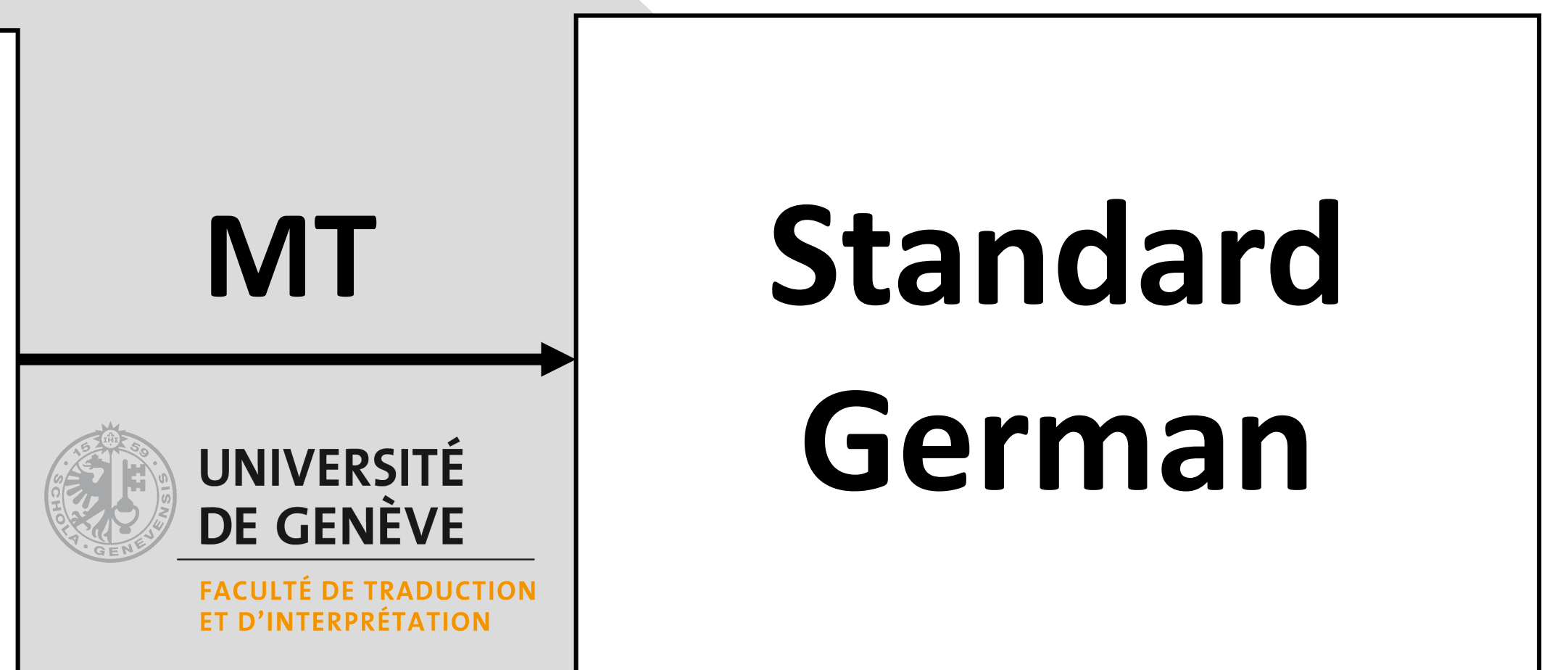
NMT system was slightly better than the APE system (BLEU 64.91 vs 61.49)



und d'Regierig hät no wiiteri Idee zum d'Stüüre abetue.



Und die Regierung hat noch weitere Ideen zum die Steuern senken.



Und die Regierung hat noch weitere Ideen, um die Steuern zu senken.