

LLM-based Translation Across 500 Years. The Case for Early New High German

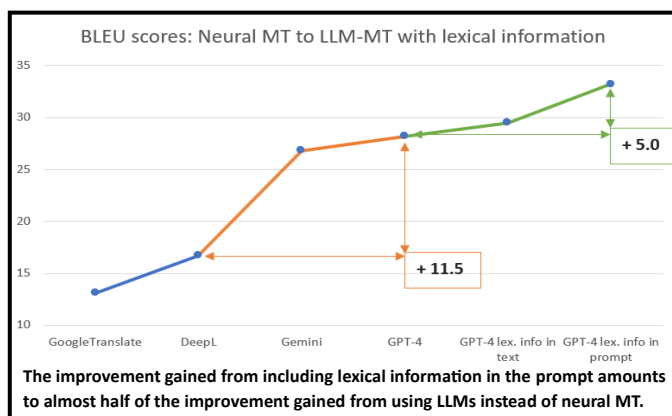
Martin Volk, Dominic P. Fischer, Patricia Scheurer, Raphael Schwitter, Phillip B. Ströbel

In A Nutshell

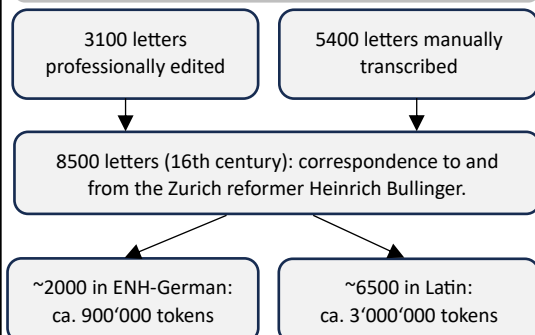
Original	<i>Ich weiß nitt, kans och nitt erfahren, wo si sind, dann sy an keinem ort sich summend²⁸. [28: verweilen]</i>			
Ref. DE	Ich weiss nicht und kann auch nicht herausfinden, wo sie sich aufhalten, da sie nirgendwo lange bleiben.			
Ref. EN	I don't know and can't find out where they are, as they don't stay anywhere for long.			
	Neural MT	LLM-based MT	LLM-based MT enhanced with lexical footnotes	BLEU score on test set
Google Trans.	Ich weiß es nicht und kann es auch nicht herausfinden, wo sie sind, dann brummen sie nirgends.			13.1
DeepL	Ich weiß es nicht, ich kann nicht herausfinden, wo sie sind, dann sind sie nirgendwo brummen.			16.7
Gemini	Ich weiß nicht, und kann es auch nicht erfahren, wo sie sich befinden, da sie sich an keinem Ort aufhalten.			26.8
GPT-4	Ich weiß nicht, kann auch nicht herausfinden, wo sie sind, denn sie zeigen sich an keinem Ort.			28.2
GPT-4 Lex. footnotes	in text*	Ich weiß nicht, kann auch nicht erfahren, wo sie sind, denn sie verweilen an keinem Ort.		29.5
	in prompt**	Ich weiß nicht, kann auch nicht erfahren, wo sie sind, denn sie verweilen an keinem Ort.		33.2

* Footnotes replaced directly in text: *Transfer this letter from old German into modern German: „Ich weiss nitt, kans och nitt erfahren, wo si sind, dann sy an keinem ort sich verweilen.“*

** Translation suggestion for footnotes in prompt: *Transfer this letter from old German into modern German. Translate „summend“ as „verweilen“: „Ich weiss nitt, kans och nitt erfahren, wo si sind, dann sy an keinem ort sich summend.“*



The Corpus:



The Use of Lexical Footnotes: Motivation and Challenges

- Exploit lexical footnotes in historic texts or use historic-modern dictionaries; thus improving LLM-based MT by incorporating human expertise.
- Lexical (esp. content word) footnotes are a good indicator for translation quality: Lowest BLEU meant lowest (10%), highest BLEU meant highest percentage of footnotes (72%) as well as highest amount of content words within.

The big challenge: What does the footnote refer to?

Item den Fellix Müller, den düch scherer ³⁸ .	38:
<i>Like Felix Müller; the cloth shearer:</i>	Tuchscherer
... und den doff ⁷⁰ [u]nnd daß nacht mall den	70: die Taufe;
heren, wie maß ⁷¹ hie zû Zürich brucht	71: man es
<i>... and the baptism and the last supper of the lord, as one uses it here in Zurich</i>	

Conclusion

We argue that LLMs like GPT-4 and Gemini are the first useful systems to translate ENH-German into modern German.

We show how to exploit human information from editions to achieve higher quality translations.

Limitations

- Test set based on LLM pre-translations instead of human translations. We counterbalance the approach by having three people check and correct the pre-translations.
- Disregarding regional variants within ENH-German. In future work we will exploit the sender locations to cluster the ENH-German letters.
- Lexical footnotes resemble the use of a bilingual dictionary, but the latter might contain multiple senses for a given word which requires disambiguation.