Year: 2024

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

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

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
University of Zurich
Department of Computational Linguistics
`volk@cl.uzh.ch`

Abstract

The recently developed large language models (LLMs) show surprising translation capabilities for modern languages. In contrast, this paper investigates the ability of GPT-4 and Gemini to translate 500-year-old letters from Early New High German into modern German. We experiment with a corpus from the 16th century that is partly in Latin and partly in ENH-German. This corpus consists of more than 3000 letters that have been edited and annotated by experts from the Institute of Swiss Reformation Studies. We exploit their annotations for the evaluation of machine translation from ENH-German to German. Our experiments show that using the lexical footnotes by the editors in the prompts or directly injected into the text leads to high quality translations.

1 Introduction

Early New High (ENH) German marks the period in the history of the German language between the mid-14th and the mid-17th century. During this time the language experienced significant linguistic, cultural, and social changes that lay the foundation for modern German. Several characteristics distinguish ENH-German from both its predecessor, Middle High German, and its successor, New High German.

Native German speakers can grasp texts in ENH-German after some time of training or customization. Often, the overall gist is clear, although some words remain puzzling because they were spelled differently, have shifted meaning considerably or have fallen out of use. For example:

- ENHG: *unverschempter fravel* → unverschämter Frevel (outrageous offence)

- ENHG: *fast yfrig* (= literally: fest eifrig) → sehr entschieden (very determined)

- ENHG: *in die harr liden* → auf die Dauer ertragen (bear in the long run)

Translating Early New High German into modern German looks easy at first sight. But how good are machine translation systems on this task? It is a positive property of recent MT systems that they are robust against spelling variations (Bergmanis et al., 2020) which abound in ENH-German (Dipper and Schaffer, 2021). Subword segmentation and subsequent embeddings have resulted in MT systems that can handle learner language and dialectal spellings. Therefore, neural MT systems like Google Translate or DeepL can translate e.g. Swiss German dialect tweets or texts with word order variation and spelling errors into well-formulated English. If we then reverse the translation direction, we will obtain correctly spelled and well-worded German texts. And LLMs are even better at this text rewriting task than MT systems.

So this all indicates that MT for ENH-German into modern English or German should be possible. We compare DeepL and Google Translate with GPT-4 and Gemini in different configurations: adding lexical information to the prompts and inserting lexical information into the ENH-German sentences. We conclude with some translation experiments on sentences with ENH-German - Latin code-switching that abound in our corpus.

2 Related Work

2.1 Previous Work on NLP for ENH-German

For many years, the processing of historical languages depended on a normalization step. All spelling variants of a word were mapped to a normalized word form (e.g. ENH-German *wyn*, *win* were mapped to modern German *Wein* (wine)). Bollmann et al. (2017) applied an encoder-decoder architecture (a form of character-based rather than word-based neural MT) for text normalization of ENH-German. Their best models had an average word accuracy of 82.7%.

Schulz and Kuhn (2016) presented a Part-of-Speech tagger for historical German texts and evaluated it on a small testset. Ortmann (2021) developed a chunker for various stages of historical German including ENH-German. She evaluated the recognition of four phrase types (noun, prepositional, adjective, and adverb phrases). Sapp et al. (2023) built a parser for ENH-German by exploiting cross-dialect training from a middle low German treebank. Detecting code-switching between Latin and ENH-German was introduced by Volk et al. (2022).

2.2 Previous Work on MT for Ancient Languages

The first attempts to exploit MT for ancient languages aimed to normalize spelling variations. For an example, see Hämäläinen et al. (2019) for the normalization of Early English letters. But more recently, directly using neural MT for historical languages has been studied in various directions. Wang et al. (2023) organized a shared task for MT of Ancient Chinese. Park et al. (2022) worked on neural MT of ancient Korean.

Volk et al. (2024) investigate LLM-based MT for Latin and ENH-German. They focus on Latin but also touch upon MT and summarization for ENH-German. They compare GPT-4 based translations from ENH-German to German and English against human-written summaries. Despite low overlap scores with the summaries, they argue that GPT-4 is clearly better in translating ENH-German into English and German than Google Translate.

We found no other literature on MT from ENH-German to modern languages. We believe there exist no dedicated MT systems for ENH-German as source language. Our paper is the first systematic study of MT for ENH-German.

3 Our Corpus of Letters in ENH-German

We work on a large corpus of 16th-century letters (Volk et al., 2022; Ströbel et al., 2024). 3100 letters have been professionally edited by the Institute for Swiss Reformation Studies[1], and another 5400 have been manually transcribed. Three quarters of the letters are in Latin, the rest is in ENH-German, many letters contain code-switching between the two languages. This means we have a corpus of roughly 900,000 tokens in ENH-German (and 3 million tokens in Latin). In addition, our corpus

comprises 2500 letters that have automatic transcriptions produced by Handwritten Text Recognition (HTR).

The letters include historical characters like ẹ, ů, å, ỏ, ů. Abbreviations have been spelled out by the editors and transcribers, for instance *U a w b* is spelled out as *U[wer] a[ller] w[illige] b[rüeder]* (your all devoted brothers). In our translation experiments we use the spelled-out words without the brackets.

Paragraph boundaries are set by the transcribers, sentence boundaries have been automatically added. We automatically assigned a language tag to each sentence based on a self-trained language identifier that is able to distinguish between ENH-German and Latin with high accuracy (Volk et al., 2022).

The letters are part of the correspondence to and from the Zurich reformer Heinrich Bullinger. They deal with politics, theological debates, regional and European news as well as education and family matters. Bullinger's correspondence network extended from Zurich throughout Europe.

The 3100 edited letters have been published in 20 volumes of a "critical edition" (Gäbler et al., 1974–2020). They come with 81,573 footnotes in German that contain various types of comments by the editors. For instance, we were able to classify 3740 of these as biographical footnotes. They contain biographical information on some person mentioned in the letter (e.g. date and place of birth or profession). And – this is of relevance for the current work – we marked 8000 footnotes as lexical, most of them with the translation of a word or a short phrase from ENH-German to modern German. See Table 1 for examples.

For high precision, we marked only footnotes with one or two words as lexical. One-word footnotes account for 83% of the marked footnotes. 12% are two-word footnotes with a phrase (*in Richtung, gestern Abend, zu gehen* (in direction, last night, to go)), and 5% are two words separated by a comma which denote translation alternatives (*Gewehr, Waffen; abermals, erneut; verbergen, vorenthalten* (gun, weapon; again, anew; hide, withhold)).

Even though the footnote information is concise, the automatic application of this lexical information is challenging. Example footnote 19 in Table 1 shows the simple case of one modern German word that corresponds to one word in the ENH-German sentence. But footnote 38 has a German compound

---

Sentence in ENH-German	Footnote in modern German
... min hehen[19] heigind inn můßen dar zů zwingen *... my lords had to force him to do this*	19: Herren
Item den Fellix Müller, den důch scherer[38]. *Like Felix Müller, the cloth shearer*	38: Tuchscherer
... und den doff[70] [u]nnd daß nacht mall den heren, wie maß[71] hie zů Zürich brucht ... *... and the baptism and the last supper of the lord, as one uses it here in Zurich*	70: die Taufe; 71: man es
Alles dis ich üch verschiben[81] han, dem ist aso. *Everything that I recorded for you, is so.*	81: verschrieben, aufgezeichnet

Table 1: Sentences with examples of lexical footnotes, taken from letter 794 in (Gäbler et al., 1974–2020), Konrad Wirz to Heinrich Bullinger, April 1536. The letter was published in volume 6 of the edition in 1995.

that corresponds to two words in ENH-German. Footnote 70 is a two-to-two words correspondence, whereas 71 is a one-to-two mapping. Footnote 81 lists two alternative translations for the ENH-German word in the text.

4 LLM-based MT for ENH-German

4.1 Evaluating against Lexical Footnotes

As a first experiment, we translated 50 random ENH-German sentences into modern German with GPT-4 and checked whether the given German words from the lexical footnotes were in the translations. For instance, when we translate the second example sentence from Table 1 which has the lexical footnote *Tuchscherer*, we check whether this target word is in the modern German translation. The hypothesis is that the presence of these target words are evidence for good translations. This hypothesis is supported by the fact that the lexical footnotes comment on "non-intuitive" or difficult ENH-German words, as deemed by the editors.

We translated our ENH-German letters by prompting GPT-4 with: Transfer this letter from old German into modern German, sentence by sentence: [letter here]. Output only the transferred sentences, one by one. Do not use any numbering.

We found that 198 out of 743 target words (27%) are in the GPT-4 output of the 50 letters. This is considerably higher than the 10% of target words that we find when we translated via the pivot language English with Google Translate (ENH-German → English → German; cf. section 5).

4.2 Creating a Testset for MT Evaluation

In order to evaluate the MT quality for the ENH-German letters we need human reference translations. To create such a a testset efficiently we randomly selected 10 ENH-German letters from our corpus and pre-translated them with GPT-4 into modern German with the prompt: Translate the following letter from Early New High German to modern German.

We realized that GPT-4 preserves the historic style of the letters, and we therefore translated the German output again with the prompt: Reformuliere folgende Sätze in flüssigem, modernem Deutsch. (Reformulate the following sentences in fluent, modern German.) See Table 2 for an example of how the sentence changes in the two translation steps.

We then asked an expert in medieval linguistics (Latin and ENH-German) to correct the second output, which we then regarded as the gold standard human reference translation.

We realize that this method biases the human reference translation towards GPT-4. This approach, however, enabled us to produce reference translations for 10 letters (201 sentences) with a reasonable effort.

To counteract the bias, we asked another expert to correct the same sentences. The evaluation showed that comparing GPT-4's translations against the two different references yielded only minimal discrepancies.

4.3 Evaluating against our Testset

We translated the test set, letter by letter, from ENH-German to modern German by using GPT-4 (through the API) with the same prompt as in Sec-

Original ENH-German	**Ich weiß nitt, kans och nitt erfaren, wo si sind, dann sy an keinem ort sich summend[28]. [28: verweilen]**
Human Reference German	Ich weiss nicht und kann auch nicht herausfinden, wo sie sich aufhalten, da sie nirgendwo lange bleiben.
English	I don't know and can't find out where they are, as they don't stay anywhere for long.

MT System	**Automatic Translation**
GPT-4	Ich weiß nicht, kann auch nicht herausfinden, wo sie sind, denn sie zeigen sich an keinem Ort.
GPT-4 with lexical info in prompt	Ich weiß nicht, kann auch nicht erfahren, wo sie sind, denn sie verweilen an keinem Ort.
GPT-4 with lexical info inserted in text	Ich weiß nicht, kann auch nicht erfahren, wo sie sind, denn sie verweilen an keinem Ort.
GPT-4 two-step translation	Ich weiß nicht und kann auch nicht herausfinden, wo sie sich aufhalten, denn sie bleiben nirgendwo lange.
Google Gemini	Ich weiß nicht, und kann es auch nicht erfahren, wo sie sich befinden, da sie sich an keinem Ort aufhalten.
DeepL	Ich weiß es nicht, ich kann nicht herausfinden, wo sie sind, dann sind sie nirgendwo brummen.
GoogleTranslate	Ich weiß es nicht und kann es auch nicht herausfinden, wo sie sind, dann brummen sie nirgends.

Table 2: An ENH-German sentence taken from a letter of Berchtold Haller to Heinrich Bullinger, 28.10.1535, translated to modern German by different systems.

tion 4.1. For every letter, we computed the lowercase BLEU score and then averaged the scores. This results in a BLEU score of 28.2.

For comparison, we also translated the test set letters with Google Gemini (through the website). Just asking it to translate the letter resulted in boilerplate additions. Therefore, we sharpened the prompt to: `The following letter is in old German (Early New High German). Please translate it into modern German line by line. Please provide only the translation in German. No explanations.`

This worked for eight out of the ten files from our test set and resulted in an average BLEU score of 26.8. Gemini refused to translate the other two files, without any reasonable explanation. In repeated attempts Gemini did not produce any output for these letters.

4.4 Adding Lexical Information to the Prompt

Similar to the integration of terminology to a prompt (as in Bogoychev and Chen (2023)), we add the translation suggestions from the lexical footnotes to the prompt. Our general prompt is: `Translate the following letter from Early New High German into modern German. For` instance, when we translate the first sentence from Table 1, we add to the prompt `Translate 'hehen' as 'Herren'`. Since we do not know to how many ENH-German words a lexical footnote item refers, we use the heuristic that we specify the same number of words as in the lexical footnote. This means we add `Translate 'scherer' as 'Tuchscherer'.`, `Translate 'den doff' as 'die Taufe'.`, and `Translate 'wie maß' as 'man es'`. Unfortunately, this introduces some noise into the translation suggestions.

In a first evaluation we checked how often the desired target words (which were specified in the lexical footnotes) are in the automatic translation. We found that 533 out of 743 target words (72%) are contained in the translations. This count is based on exact matching the words from the lexical footnotes in the translations. Inflected forms would not match. The two human reference translations contained 52 resp. 53% of the lexical footnotes, while prompting GPT-4 without lexical information contained 27% (cf. Section 4.1).

In a second evaluation, we compared the GPT-4 output of our 10 letter test set with the human reference translation. This resulted in an average BLEU score of 33.2 (the scores range from 29.9 to

System Configuration	BLEU
GPT-4	28.2
GPT-4 with lexical info inserted in text	29.5
GPT-4 with lexical info in prompt	33.2
GPT-4 two-step translation*	51.9
GoogleTranslate (ENHG → EN → DE)	13.1
DeepL (ENHG → EN → DE)	16.7
Google Gemini	26.8

Table 3: Averaged BLEU scores (computed with the SacreBLEU tool) on the test set (10 ENH-German letters) when translating ENH-German to modern German. *The two-step translation served as the basis for the human reference translation.

38.9).

One may view adding lexical information to a prompt as an unrealistic setting since ENH-German texts do not usually come with specific translation suggestions. We argue that this setting resembles the use of a bilingual dictionary[2] (ENH-German to modern German) as an information source for steering the LLM translation.

4.5 Inserting Lexical Information into the Sentence

Rather than adding the lexical information as translation suggestions to the prompt, we now insert them directly into the source sentence by replacing the original word with the modern target word. This means we replace "hehen" with "Herren" in our ENH-German example sentence from Table 1 which then looks like "... min Herren heigind inn můßen dar zů zwingen" before we feed it to GPT-4 for translation.

We evaluated in the same way as above, both against the lexical footnotes and the test set. Interestingly, the evaluation with the target words from the lexical footnotes showed that fewer of them occurred in the translations: 467 out of 743 (63%). This means that adding the lexical translation information to the prompt preserves this information better than inserting it into the source sentence, which, in turn, suggests a better translation.

This result is confirmed by our evaluation against the test set (cf. Table 3).

The lexical footnotes in our corpus suggest target words for content words and function words. We would argue that the correct translation of content words is more important. Therefore, we automati-

cally classified all our lexical footnotes into content vs. function words. Two-word footnotes were split and their parts classified.

When translating with the above simple prompt, the percentage of content words in the contained footnotes is 11% lower than in the missing footnotes (62.2% vs. 73.2%). With the footnotes directly inserted into the text when prompting, the difference is only marginal with the percentage of content words in the contained footnotes being 0.7% lower (69.4% vs. 70.1%). Finally, with the footnotes included in the prompt: the percentage of content words is 11.1% higher in the contained footnotes (71.4% vs. 60.3%). This shows that the quality of the contained footnotes increases when the lexical information is included in the prompt.

5 Comparison to Neural MT Systems

We cannot directly compare our LLM-based MT results to neural MT systems like DeepL or Google Translate since they do not offer ENH-German as source language. But we can pretend that the input is German and ask for a translation into some other language. We chose English as the pivot language. If we subsequently reverse the translation direction, we will get a modern German version.

When we applied this two-step translation with DeepL (and UK-English as pivot) for the 10 ENH-German letters in our test set, we obtained an average BLEU score of 16.7 (ranging from 11 to 21.9 for the 10 files). We see that DeepL interprets the words on the surface, e.g. translating ENH-German "rowen" into English as "rowing" instead of "robbing" or "stealing". DeepL allows the integration of a glossary which we did not use since we would need ENH-German to English correspondences, while our lexical footnotes provide ENH-German to modern German mappings.

---

[2]For examples see Frühneuhochdeutsches Wörterbuch at https://fwb-online.de/ or the Reference Corpus Early New High German at https://www.linguistics.ruhr-uni-bochum.de/ref/

Original ENH-German and Latin	**Philippum nostrum amicissime salutabis;** dices illi, das man ein latinische, schöne, wol yngebundne bibel gäbe umb try guldin.
Human Reference English	Give my warmest greetings to our Philipp; tell him that a beautifully bound Latin Bible can be purchased for three florins.
Translated by GPT-4	You will greet our dear Philip most kindly; tell him that one can get a Latin, beautiful, well-bound Bible for three guilders.
Translated by Gemini	Greet our Philip most kindly; tell him that a beautiful, well-bound Latin Bible is offered for three guilders.
Original ENH-German and Latin	Ich han ein pflägeri im huß; deren gib ich alle wuchen 1 fl. (**sic et alii**), on spyß und tranck;
Human Reference English	I have a servant at home; I give her 1 fl. every week without food and drink (so do others as well).
Translated by GPT-4	I have a care facility in the house; to which I give 1 florin every week (and others do the same), without food and drink;
Translated by Gemini	I have a nurse in the house; I give her 1 florin every week (and so do others), not including food and drink.
Original ENH-German and Latin	Die **seniores illius ecclesiae** habend inn bschickt.
Human Reference English	The leaders of his church have sent him.
Translated by GPT-4	The elders of that church have been sent in.
Translated by Gemini	The elders of that church have put them in charge.

Table 4: Sentences with code-switching (i.e. mixing ENH-German and Latin) taken from our letter collection, translated by Open AI's GPT-4 and Google Gemini.

We also observe that occasional Latin sentences in our ENH-German letters are left untranslated by DeepL.

We are aware that DeepL offers a rewriting system ("DeepL Write") in addition to their MT system. In principle, this rewriting system can turn ENH-German texts into modern German. It allows one to select among four styles (simple, business, academic, easy) and four tones (enthusiastic, friendly, sovereign, diplomatic). It is unclear which style and tone combination would be most suitable for our letters. Rewriting also restructures the text, leading to additional challenges for evaluation, which is why we did not evaluate this system.

6 Evaluating Sentences with Code-Switching

So far, we have concentrated on sentences that are exclusively in ENH-German. But our corpus contains many sentences with code-switching. Therefore, we selected 61 sentences with a mix of Latin and ENH-German from our corpus and had them translated into English by an expert. We used English as the target language here because we know from previous experiments (Volk et al., 2024) that it results in high quality translations from Latin

source texts.

We then asked the LLMs to translate these sentences (without context) into English with the prompt: `The following sentences are a mixture of Latin and old German (Early New High German). Translate them into modern English line by line.`

For GPT-4, this resulted in a BLEU score of 23.2 and a ChrF score of 46.3. Gemini scores slightly higher with a BLEU score of 25.4 and a ChrF score of 48.2 The online MT systems are unable to handle a mixture of the two languages in question here.

Table 4 shows two example sentences with impressive translations, slightly more fluent and idiomatic in the Gemini output than in in GPT-4. But we should keep in mind that sometimes the translation for presumably simple sentences has serious errors, as in our third example where GPT-4 translates an active sentence with a passive one, and thus gets the agent wrong, and Gemini produces a plural pronoun where the input pronoun is in singular.

7 Conclusion

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

showed how to exploit footnotes that specify lexical information in an edition of letters from the 16th century. These lexical footnotes map "non-intuitive" ENH-German words from the letters to modern German words (and thus provide translation suggestions). We used these lexical footnotes to evaluate the translations and then to steer the translations. We show that a two-step translation process with GPT-4 leads to high-quality translations in modern German.

We limited our work by automatically identifying only the most apparent lexical footnotes, i.e. footnotes with only one or two words. In future work we will identify and use lexical footnotes that are longer. A glance at our corpus reveals that there will be more than 1000 such footnotes which are more informative but also more complicated to exploit. It is often unclear to how many tokens from the ENH-German sentence they correspond.

Our study focused on commercial MT systems and multilingual LLMs. In future work we will also investigate open LLMs like LLaMA which we can then finetune to our specific needs.

Limitations

The most obvious limitation is our choice of building a test set based on LLM pre-translations. Independent human translations would be better (but are more time-consuming to produce). We counterbalance this approach by having three persons check and correct the pre-translations.

Secondly, we are aware that we regard ENH-German as a static block, although there are likely personal or regional variants that differ in distance to modern German and are thus harder to translate. In future work we will exploit the sender locations to cluster the ENH-German letters.

Thirdly, we argue that using lexical footnotes resembles the use of a bilingual dictionary. This is a simplification since these footnotes contain translation suggestions that were selected by the editors. A bilingual dictionary might contain multiple senses for a given word which must be disambiguated for use in translation.

Ethics Statement

Given the age of the ENH-German texts, its use does not pose an ethical challenge.

References

Toms Bergmanis, Artūrs Stafanovičs, and Mārcis Pinnis. 2020. Robust neural machine translation: Modeling orthographic and interpunctual variation. In *Human Language Technologies – The Baltic Perspective*.

Nikolay Bogoychev and Pinzhen Chen. 2023. Terminology-aware translation with constrained decoding and large language model prompting. In *Proceedings of the Eighth Conference on Machine Translation (WMT)*, pages 890–896, Singapore. Association for Computational Linguistics.

Marcel Bollmann, Joachim Bingel, and Anders Søgaard. 2017. Learning attention for historical text normalization by learning to pronounce. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 332–344, Vancouver, Canada.

Stefanie Dipper and Jan Christian Schaffer. 2021. Identifikation von Vorkommensformen der Lemmata in Quellenzitaten frühneuhochdeutscher Lexikoneinträge. In *Proceedings of the 17th Conference on Natural Language Processing (KONVENS 2021)*, pages 62–72, Online.

Ulrich Gäbler, Endre Zsindley, Kurt Maeder, Matthias Senn, Kurt Jakob Rüetschi, Hans Ulrich Bächtold, Rainer Heinrich, Alexandra Kess, Christian Moser, Reinhard Bodenmann, Judith Steiniger, and Yvonne Häfner, editors. 1974–2020. *Heinrich Bullinger Briefwechsel*. Heinrich Bullinger Werke. Theologischer Verlag Zürich.

Mika Hämäläinen, Tanja Säily, Jack Rueter, Jörg Tiedemann, and Eetu Mäkelä. 2019. Revisiting NMT for normalization of early English letters. In *Proceedings of the 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, pages 71–75, Minneapolis, USA. Association for Computational Linguistics.

Katrin Ortmann. 2021. Chunking historical German. In *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa)*, pages 190–199.

Chanjun Park, Seolhwa Lee, Jaehyung Seo, Hyeonseok Moon, Sugyeong Eo, and Heuiseok Lim. 2022. Priming Ancient Korean neural machine translation. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference (LREC)*, pages 22–28, Marseille, France.

Christopher Sapp, Daniel Dakota, and Elliott Evans. 2023. Parsing Early New High German: Benefits and limitations of cross-dialectal training. In *Proceedings of the 21st International Workshop on Treebanks and Linguistic Theories (TLT, GURT/SyntaxFest 2023)*, pages 54–66, Washington, D.C. Association for Computational Linguistics.

Sarah Schulz and Jonas Kuhn. 2016. Learning from within? Comparing PoS tagging approaches for historical text. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 4316–4322, Portorož, Slovenia.

Phillip Benjamin Ströbel, Lukas Fischer, Raphael Müller, Patricia Scheurer, Bernard Schroffenegger, and Martin Volk. 2024. Multilingual workflows in Bullinger Digital: Data curation for Latin and Early New High German. *Journal of Open Humanities Data*, 10.

Martin Volk, Dominic P. Fischer, Lukas Fischer, Patricia Scheurer, and Phillip B. Ströbel. 2024. LLM-based machine translation and summarization for Latin. In *Proceedings of the Third Workshop on Language Technologies for Historical and Ancient Languages – LT4HALA (at LREC/COLING)*, Torino.

Martin Volk, Lukas Fischer, Patricia Scheurer, Raphael Schwitter, Phillip Ströbel, and Benjamin Suter. 2022. Nunc profana tractemus. Detecting code-switching in a large corpus of 16th century letters. In *Proceedings of LREC-2022*, Marseille.

Dongbo Wang, Litao Lin, Zhixiao Zhao, Wenhao Ye, Kai Meng, Wenlong Sun, Lianzhen Zhao, Xue Zhao, Si Shen, Wei Zhang, and Bin Li. 2023. EvaHan2023: Overview of the first international Ancient Chinese translation bakeoff. In *Proceedings of ALT2023: Ancient Language Translation Workshop*, pages 1–14, Macau SAR, China. Asia-Pacific Association for Machine Translation.