Evaluating Machine Translation of Latin Interjections in the Digital Library of Polish and Poland-related News Pamphlets

Maciej Ogrodniczuk, Katarzyna Krynska

Institute of Computer Science / Institute of Polish Language
Polish Academy of Sciences

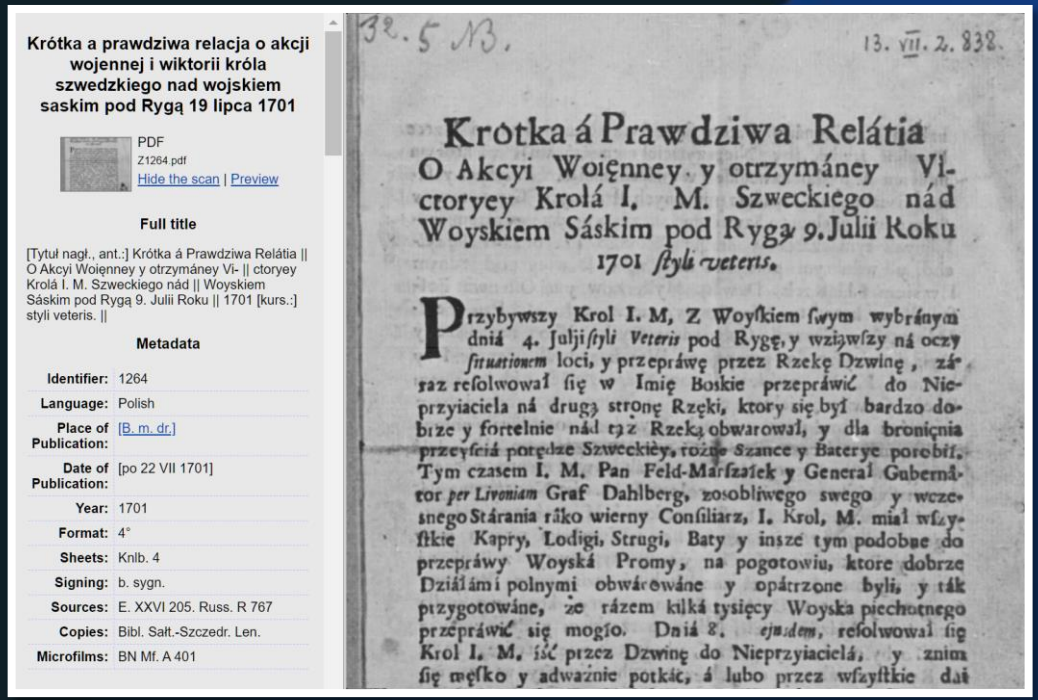
The 24th International Conference on Asia-Pacific Digital Libraries (ICADL 2022)
December 1, 2022

Paper summary

- we have a digital library of Polish of a middle Polish pre-press prints
- some print texts contain Latin interjections
- some items contain manually created Latin–Polish dictionaries
- why not translate them all automatically?
- let’s evaluate different solutions (out-of-the-box one vs. language model-based)
- by using automatic and human metrics: BLEU vs. adequacy and fluency

The Digital Library of Polish Prints from the 16th–18th Centuries

- a thematic DL with over 2 000 prepress documents
- managed by Eprints
- soon 200+ items to be extended with texts
- available at https://cbdu.ijp.pan.pl



Krótka a prawdziwa relacja o akcji wojennej i wiktorii króla szwedzkiego nad wojskiem saskim pod Rygą 19 lipca 1701

PDF
Z1264.pdf
Hide the scan | Preview

Full title

[Tytuł nagł., ant.:] Krótka á Prawdziwa Relátia || O Akcyi Woiẹnney y otrzymáney Vi- || ctoryey Krolá I. M. Szweckiego nád || Woyskiem Sáskim pod Rygą 9. Julii Roku || 1701 [kurs.:] styli veteris. ||

Metadata

Identifier:	1264
Language:	Polish
Place of Publication:	[B. m. dr.]
Date of Publication:	[po 22 VII 1701]
Year:	1701
Format:	4°
Sheets:	Knlb. 4
Signing:	b. sygn.
Sources:	E. XXVI 205. Russ. R 767
Copies:	Bibl. Sałt.-Szczedr. Len.
Microfilms:	BN Mf. A 401

32.5 M3. 13. VII. 2. 838.

Krótka á Prawdziwa Relátia
O Akcyi Woiẹnney y otrzymáney Victoryey Krolá I. M. Szweckiego nád Woyskiem Sáskim pod Rygą 9. Julii Roku 1701 *styli veteris.*

Przybywszy Krol I. M, Z Woyſkiem ſwym wybránym dniá 4. Julji *ſtyli Veteris* pod Rygę, y wziąwſzy ná oczy *ſituationem* loci, y przeprawę przez Rzekę Dzwinę, zá raz reſolwował ſię w Imię Boskie przeprawić do Nieprzyiaciela ná drugą ſtronę Rzęki, ktory ſię był bardzo dobrze y fortelnie nád tą z Rzeką obwarował, y dla broniẹnia przeyſciá porobił Szweckiey, rożne Szance y Baterye porobił. Tym czasem I. M. Pan Feld-Marſzałek y General Gubernátor *per Livoniam* Graf Dahlberg, zoſobliwego swego y wczesnego Stárania ráko wierny Conſiliarz, I. Krol, M. miał wſzyſtkie Kapry, Lodigi, Strugi, Baty y inſze tym podobne do przeprawy Woyská Promy, na pogotowiu, ktore dobrze Działámi polnymi obwárowáne y opátrzone były, y ták przygotowáne, że rázem kilká tyſięcy Woyska piechotnego przeprawić ſię mogło. Dniá 8. *ejusdem*, reſolwował ſię Krol I. M. iść przez Dzwinę do Nieprzyiacielá, y zniim ſię męſko y adwáżnie potkáć, á lubo przez wſzyſtkie dni

The Latin content

- Latin interjections constitute 7–8% of the print texts
- while still understandable for people familiar with Latin and experts, may be difficult to decipher by a non-specialist
- manually created dictionaries of Latin terms and phrases were added to some prints
- due to financial constraints of the original project, only 20+ prints were annotated in this way
- the ‘Latin dictionaries’ contain approx. 600 entries

Data preparation

- dictionary entries were pre-processed to remove ‘editorial content’
- but: will larger contexts help the translation? let’s check!
- transcriptions of 18 prints containing Latin available in TEI-encoded *The Electronic Corpus of the 17th and 18th Century Polish Texts* (https://korba.edu.pl)
- Latin parts marked with `<foreign>` tags
- some additional cleaning needed (inserting parts of content removed from translations, expanding abbreviations)
- final translation dataset: full sentences containing Latin interjections

Evaluation settings

1. MT approach:
    - Google Translate
    - GPT-3
2. Scope:
    - context-less (isolated)
    - context-aware
3. Evaluation method:
    - BLEU
    - manual (White's 5-point scaleof adequacy and fluency)

Results of automatic evaluation

Setting	Engine	Cumulative BLEU score			
		1-gram	2-gram	3-gram	4-gram
Isolated	Google	26.22	18.77	**14.74**	**12.58**
	GPT-3	26.27	16.45	11.35	8.33
Context-wise	Google	23.95	15.46	11.63	8.85
	GPT-3	**31.83**	**20.18**	13.45	9.54

Manual evaluation

Value	Adequacy	Fluency
5	all meaning	flawless Polish
4	most meaning	good Polish
3	much meaning	non-native Polish
2	little meaning	disfluent Polish
1	none	incomprehensible

Adequacy is related to the adherence of the translation to the source text.
Fluency grades the quality of the target text only.

Results of human evaluation

Setting	Engine	Adequacy	Fluency
Isolated	Google	3.11	4.63
	GPT-3	2.87	4.52
Context-wise	Google	2.61	4.02
	GPT-3	3.00	4.32

Some conclusions

- Even though untranslated passages (whether left in Latin or omitted) were rated lowest on the evaluation scale (1 – none or incomprehensible) in both adequacy and fluency categories, the average is still quite high (3 – conveys much meaning).
- In most cases, a high degree of fluency can be observed.
- Context helps most in the case of uni- and bigrams which is probably related to Polish inflection.
- Language models are good at smoothing the translation result.
- Still to early to use MT automatically but it proves to be a useful translator's aid.

Thank you!