Katarzyna Frąckowiak

Linguist · LLM Evaluation · Polish / English

For a non-standard CV
kasia420.github.io

Poznań, Poland · katarzynaglowne@gmail.com · kasia420.github.io · github.com/Kasia420

SUMMARY

MA student in Language, Mind & Technology at Adam Mickiewicz University, working at the intersection of linguistics and LLM evaluation. I design Polish-language test sets that probe how language models handle irony, implicature and grammar, and my MA research examines how NLP models handle non-native English. Native Polish speaker with C1 English and hands-on experience writing precise test items, judging model outputs and analysing results in Python and R.

PROJECTS

Polish LLM Evaluation Test Sets — independent project · Oct 2026 – present

- Designed four 50-item test sets probing Polish language understanding in LLMs: irony comprehension, implicature, case marking vs word order (who did what to whom), and English–Polish false friends
- Built minimal pairs and short dialogues with closed-format answers; piloted items on several models to remove ambiguous or trivially solvable cases
- Sets shared with a SpeakLeash community member for evaluating the Bielik LLM

Lost in Translation — course research project, extended · 2026 · kasia420.github.io/lost_in_translation

- Compared Claude's answers to 25 translation-matched English/Polish prompts on Polish culture, measuring cross-lingual similarity (LaBSE, median cosine .81) and sentiment (XLM-RoBERTa) against an English/Polish Wikipedia baseline
- Found Polish answers significantly less negative (Wilcoxon p = .0015, r = .70); a reproducible re-run with a pinned model snapshot, length control and a 1,000-pair classifier calibration showed a similar gap in Wikipedia itself; results presented on an interactive project website

MA thesis (in progress) — How LLMs process non-native (Polish L1) English input

EXPERIENCE

Research Intern — Psycholinguistics & Neurolinguistics Lab, UAM · 2024 – 2025

- Ran EEG/ERP sessions under a strict experimental protocol, keeping conditions consistent so data were comparable across participants
- Worked with semantic priming and lexical decision paradigms; trained in EEG signal basics (frequency bands, artefact types)

Volunteer — CADAAD 9th International Conference, UAM · July 2024

- Coordinated guest registration and on-site support for international participants at a conference on discourse analysis

English Tutor — self-employed · 2020 – 2023

- One-to-one tuition for Polish-speaking learners; daily work with L1-transfer errors (articles, prepositions, word order), the phenomenon I now study in LLMs

EDUCATION

MA, Language, Mind & Technology — Adam Mickiewicz University, Poznań · 2025 – present

BA, English Studies: Literature & Culture — Adam Mickiewicz University, Poznań · 2022 – 2025

SKILLS

- **Programming:** Python (pandas, NumPy, SciPy, matplotlib), R (tidyverse, ggplot2), Git / GitHub, Docker, REST APIs
- **NLP / LLM:** HuggingFace transformers, sentence-transformers, multilingual embeddings (LaBSE), LLM evaluation, benchmark design, prompt engineering
- **Data & statistics:** nonparametric tests, effect sizes, mixed-effects regression, data cleaning
- **Web & visualization:** HTML, CSS, JavaScript, interactive data visualization, GitHub Pages
- **Automation:** n8n

LANGUAGES

Polish — native · English — C1 (academic) · Spanish — A1 · German — A1/A2

CERTIFICATIONS

Generative AI in Business — University of Cambridge (Cambridge Advance Online) · 2025

DataCamp — Introduction to R, Introduction to the Tidyverse, Data Visualization with ggplot2, Exploratory Data Analysis in R · 2026