University of Zurich[UZH]

Department of Computational Linguistics

LREC 2026 Palma

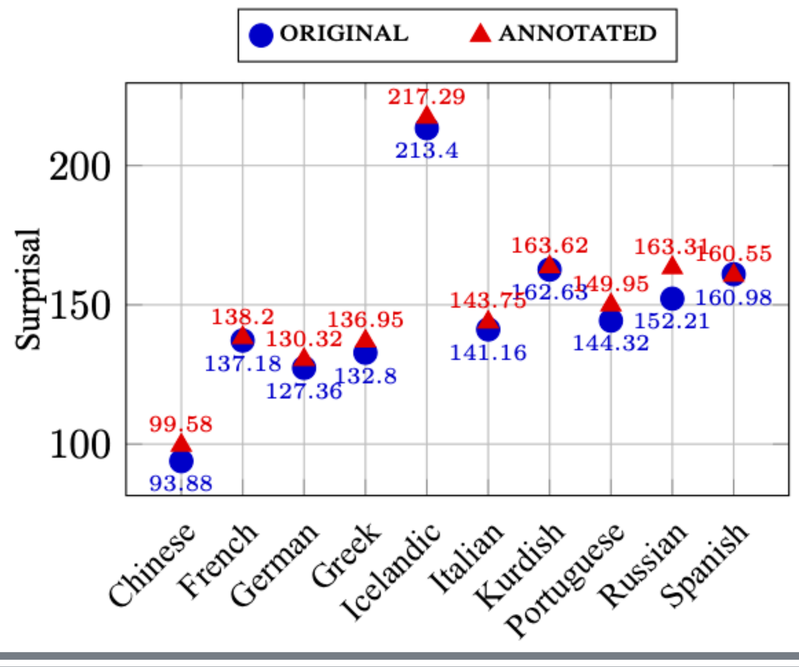
~~A Multilingual Evaluation of Loanword Identification~~
A Many-Tongued Weighing of Word-Loan Spotting

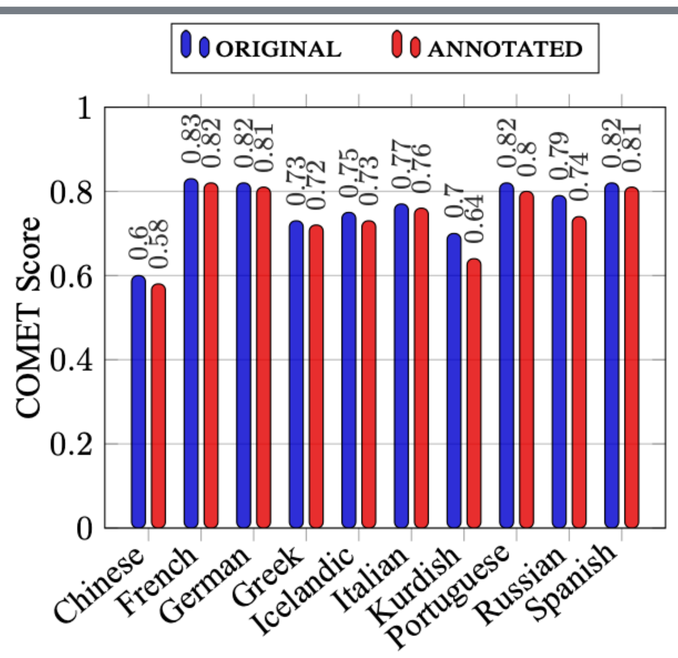
Mérilin Sousa Silva **Sina Ahmadi**
University of Zurich

Motivation

- Loanwords have been studied for decades in linguistics, but underexplored in NLP/CL
- **Prior work (Ahmadi et al., ACL 2025)**: ConLoan showed that LLMs prefer loanwords (**lower surprisal**) and **NMT performs worse** on native alternatives
- **This paper: If models prefer loanwords, can they identify them when explicitly asked?**
- Why it matters: low-resource NLP, language education, code-switching detection

We already showed that LLMs show LOWER surprisal for loanword sentences



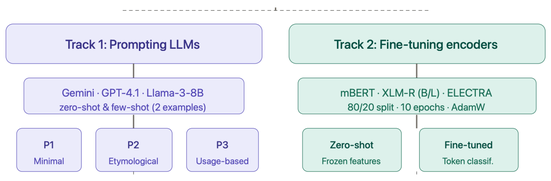


We also showed that NMT performs WORSE when translating native alternatives

Methodology

- **Task:** BIO sequence labeling for loanword identification across 10 languages using ConLoan
- **Two evaluation tracks:** (1) prompting LLMs and (2) fine-tuning multilingual encoders for token classification

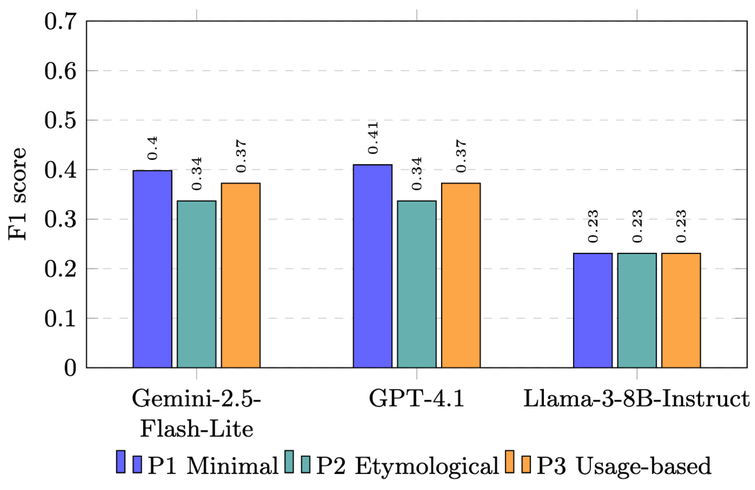
Podemos vender-te um ***franchise*** *disto por 3000$.*
(We can sell you a franchise of this for 3,000 dollars.)
Podemos vender-te um ***franquia*** *disto por 3000$.*



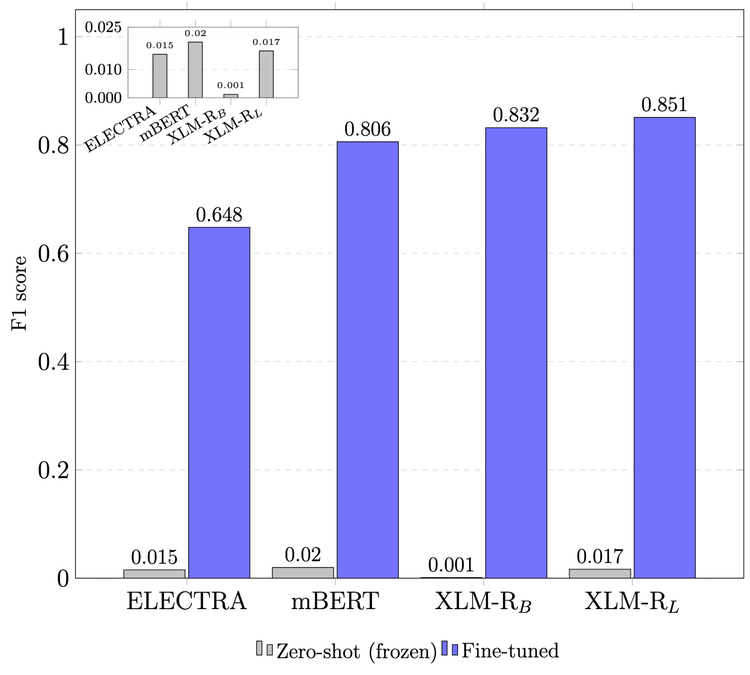
- **Prompt 1 (Minimal)**: No definition; the model is simply instructed to identify loanwords in the given sentence.
- **Prompt 2 (Etymological)**: a word that entered a language's lexicon through borrowing at some point in its history.
- **Prompt 3 (Usage-based)**: Extends P2 by distinguishing loanwords from code-switches

Key Findings

Finding 1: LLMs perform poorly on loanword identification



Finding 2: Fine-tuned encoders substantially outperform LLMs, but pretrained encoders perform near-zero



Finding 3: Misclassification patterns persist across models

- **Code-switches mistaken for loanwords**
  Il nous appartient, dans la mesure du possible – et je m'y emploie – de faire en sorte que ce qui est globalement un bon ***deal*** *entre les Américains et les Chinois, soit un aussi bon* ***deal*** *pour les Européens.*
- **Named entities flagged as loanwords**
  An der Spitze der internationalen Rangliste laut der letzten ***PISA-Studie*** *steht der Shanghai-Distrikt von China.*
- **Greco-Latin terms missed as loanwords**
  [. . .] þar sem nítröt geta breyst í nítrít og nítrósamín, og hvatti til þess að teknar yrðu upp góðar starfsvenjur í landbúnaði til þess að tryggja eins lágt nítratmagn og kostur er.

Conclusion

- Loanword identification is far from solved, even fine-tuned models err systematically
- LLMs are borrowing-blind: F1 < 0.5 even with explicit definitions
- Fine-tuning helps dramatically (XLM-R-L: 0.85 F1) but doesn't fix the conceptual gap
- Models confuse loanwords with code-switches, named entities, and Greco-Latin terms
- Implication: current NLP tools cannot reliably support minority language preservation against lexical pressure