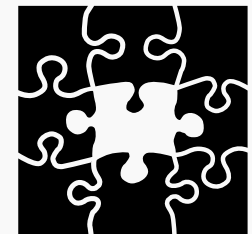# Can LLMs Really Learn to Translate a Low-Resource Language from One Grammar Book?

Seth Aycock

ILLC / LTL, University of Amsterdam

With David Stap, Di Wu, Christof Monz, Khalil Sima'an

Machine Translation from One Book

> ½ of languages have a grammar book available

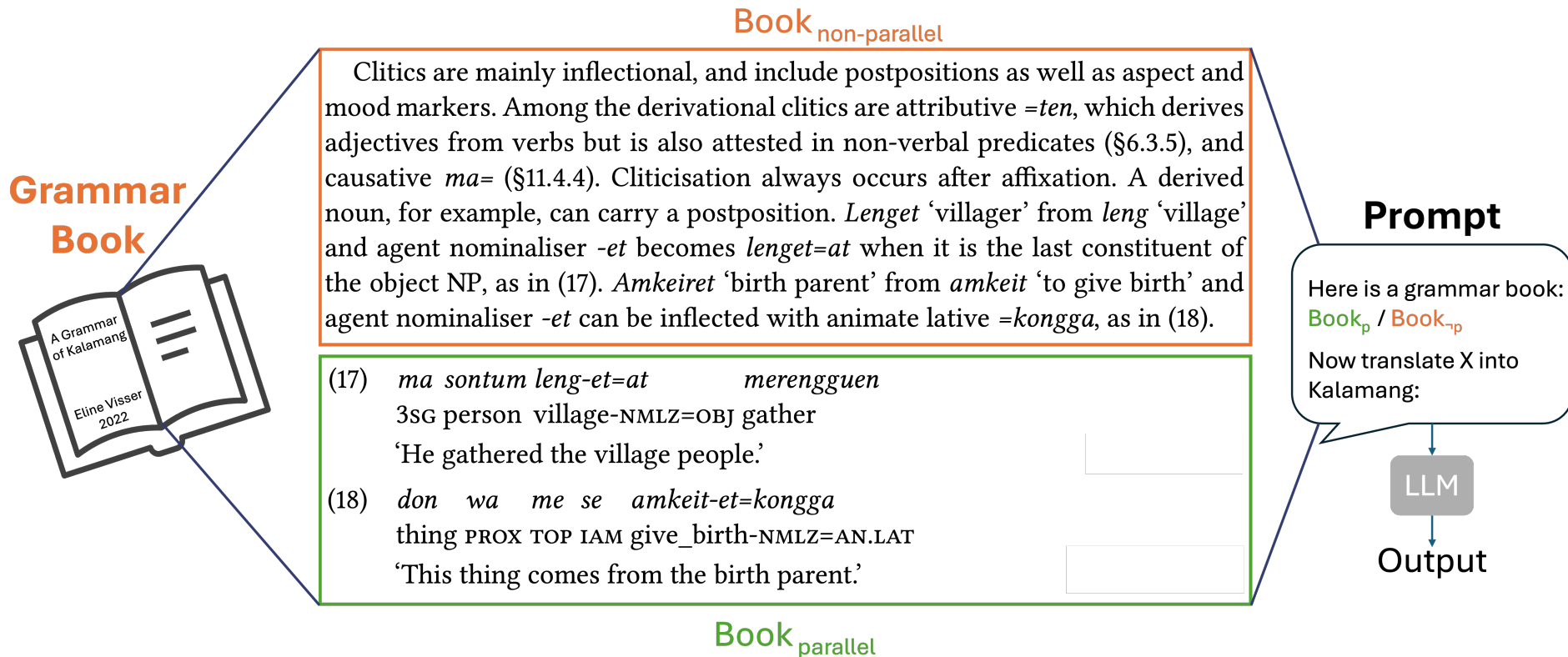Kalamang (kgv) – 200 speakers, first grammar published in 2022
Therefore **unseen** by LLMs during training

Tanzer et al. (2024):

Claim: Given the grammar book in-context, LLMs can learn to translate kgv

Really?

Linguistic knowledge rarely helps NLP task performance (Opitz et al. 2024) – is this an exception?

Book₍non-parallel₎

Clitics are mainly inflectional, and include postpositions as well as aspect and mood markers. Among the derivational clitics are attributive =ten, which derives adjectives from verbs but is also attested in non-verbal predicates (§6.3.5), and causative ma= (§11.4.4). Cliticisation always occurs after affixation. A derived noun, for example, can carry a postposition. *Lenget* 'villager' from *leng* 'village' and agent nominaliser -et becomes *lenget=at* when it is the last constituent of the object NP, as in (17). *Amkeiret* 'birth parent' from *amkeit* 'to give birth' and agent nominaliser -et can be inflected with animate lative =kongga, as in (18).

Grammar Book

A Grammar of Kalamang

Eline Visser 2022

(17) *ma sontum leng-et=at   merengguen*
      3SG person village-NMLZ=OBJ gather
      'He gathered the village people.'

(18) *don wa   me se   amkeit-et=kongga*
      thing PROX TOP IAM give_birth-NMLZ=AN.LAT
      'This thing comes from the birth parent.'

Book₍parallel₎

Prompt

Here is a grammar book:
Bookₚ / Book₍¬p₎

Now translate X into Kalamang:

LLM

Output

Research Questions

1. **Where** does LLMs' translation ability stem from: **explanations** or **examples**?

2. **Which tasks** does grammatical information help?

3. **What kind** of grammatical information is most useful?

Q1: Explanations or Examples for translation?

For Q1: We separate parallel sentences + glosses (Book$_{parallel\ (p)}$) from grammatical explanations (Book$_{non\text{-}parallel\ (\neg p)}$) for prompting and fine-tuning

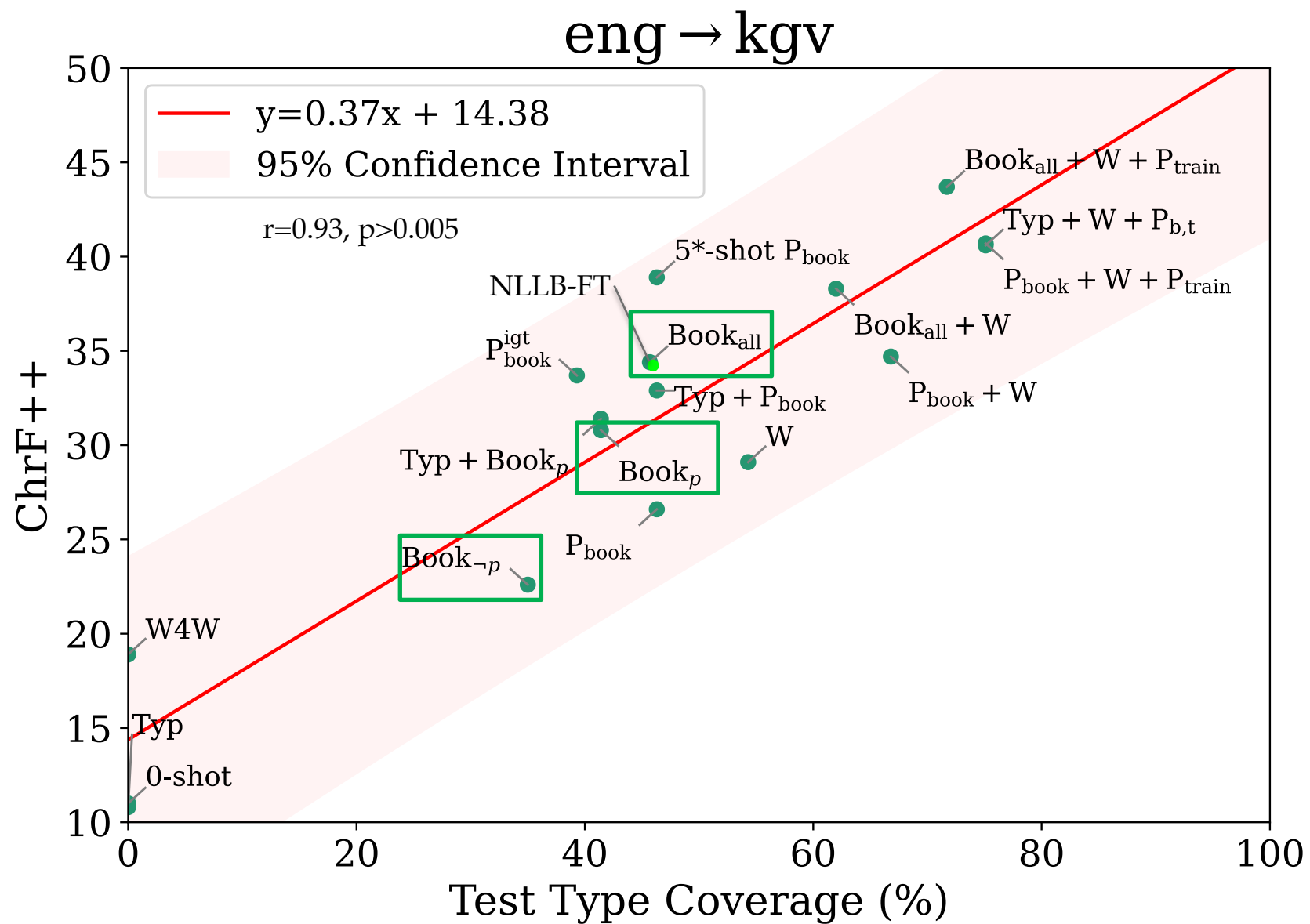Test set: Kalamang—English (100 examples); Nepali and Guarani FLORES (see paper)

Evaluation: ChrF++ for character and word-level information

Models: Gemini-1.5-Flash, Llama-3.1-8B (see paper), NLLB-1.3B (fine-tuned)

Translation with $Book_{parallel}$ beats $Book_{non\text{-}parallel}$

In kgv–eng, $Book_{parallel}$ also beats $Book_{all}$

LLMs fail to exploit grammar for translation

English to Kalamang Translation with Gemini-1.5-Flash

Which tasks?

Grammaticality Judgment

Prompt LLM to choose original sentence given:

      Grammar subset +

      Kalamang sentence

      Corrupted Kalamang sentence – swapped adjacent words, random words, or shuffled

Interlinear Gloss Prediction

Prompt LLM to generate Interlinear Gloss given:

      Grammar book subset +

      Segmented Kalamang sentence

      English translation

What kind of grammar?

Grammatical explanations of rules and exceptions, Book_{non-parallel}

Typology: describes high-level grammatical features of a language

ID	Grambank Typological Feature	Specification
GB020	Are there definite or specific articles?	No
GB130	What is the order of S and V in intransitive clauses?	SV
GB159	Are nouns reduplicated?	Yes

Summarise features in textual form as a prompt replacement for Book_{non-parallel}

For both gloss prediction and grammaticality judgment,
Book_{parallel} > Book_{all} >> Book_{non-parallel}

Best performing prompt is Typology + Book_{parallel}

So LLMs *can* exploit grammar for linguistic tasks

Summary

LLMs fail to exploit grammatical explanations for translation

Parallel examples are most helpful for translation, and fine-tuning on parallel data matches long-context LLMs

Typological prompting outperforms explanations (Book_{non-parallel}), for linguistic tasks

Task-appropriate data is key

Contact me:

Read the paper:

s.aycock@uva.nl