

Introduction

- In GEC, it is crucial to ensure the user's comprehension of a reason for correction
- Existing studies do not directly explain why the correction is needed
- Generating such explanation is not simple even for LLMs due to its complexity
 - Aligning input and output tokens -> identifying correction points -> presenting corresponding explanations consistently
- We introduce (1) **a method controlling generation with Prompt Insertion (PI)** and (2) **create an Explainable GEC (XGEC) dataset of correction reasons**

Prompt Insertion (PI) Method

The lack of creative **qualified** human **resource** hampers the process.

LLM

The lack of creative **qualified** human **resources** hampers the process.

qualified -> -

resource -> resources

LLM

PI splits complex task into simple sub-tasks to explicitly provide positions that should be generated explanations as follows:

1. Correction

2. Extraction of edit pairs

- We compute the token alignment between the input and output text and extract the edits based on the rules

We can omit "qualified" because ...

- "Qualified" usually means having the necessary skills, often certified by tests.

Use "Resources" is in the plural form because ...

- "Human resources" refers to a workforce which indicates a group of people capable of work.

3. Explanation for each pair

- The extracted edits are given as an instruction to the LLM one by one as additional input, causing the LLM to generate an explanation

XGEC Dataset

- The XGEC dataset includes incorrect texts, correct texts, and explanations for each edit



Experiments

The **BERTScore** between system generated explanation and human reference on the XGEC test datasets

		Precision	Recall	F1
ChatGPT	w/ IP	83.2	85.5	84.3
	w/o IP	62.1	79.6	70.0
GPT-3.5	w/ IP	81.2	83.8	82.4
	w/o IP	61.2	79.4	69.1

- PI improves the performance in all scores on both datasets

Human evaluations of GPT-3.5 and ChatGPT with and without PI on the XGEC test dataset

- The validity perspective refers to the accuracy and usefulness of grammatical information of explanations for learners
- The coverage perspective means that the LLM-generated explanation mentions all grammatical corrections

		Validity	Coverage
ChatGPT	w/ IP	1.5	2.0
	w/o IP	1.2	1.4
GPT-3.5	w/ IP	1.4	2.0
	w/o IP	1.1	1.5

- PI improves the performance in all scores on both datasets
- The PI makes it clear to LLM the corrections that need to be explained, and allows for specific explanations tied to each correction, improving the quality of LLM's explanations.