

## 생성형 AI의 한국어 시제 이해 및 생성 능력 연구

황은하 (Eunha Hwang), 배재대학교

이 연구는 생성형 AI의 한국어 시제에 대한 이해와 생성 능력에 대한 평가와 더불어, 생성형 AI의 시제 문법 능력 향상 방안을 모색하는 것을 목적으로 한다. 이를 위해, 본 연구는 기본 문법서의 예문을 중심으로 평가 데이터 세트로 구축하고, 국내외 생성형 AI(ChatGPT 3.5, ChatGPT 4.0, Gemini, CLOVA X 등)의 한국어 시제 이해 및 생성 능력을 평가하며, 오류 분석 및 생성형 AI 간 비교 분석을 진행한다.

최근 생성형 AI의 언어 능력에 대한 다양한 연구가 시도되고 있는데, 남길임·황은하·송현주·안의정(2024)은 한국어의 어미와 부정, 두 문법 범주를 중심으로 생성형 AI의 문법적 능력을 평가한 결과 영어의 문법 능력에 훨씬 못 미치는 성능을 보였으며, 이를 한국어의 형태론적 특성에서 기인하는 것으로 보았다. 이 연구의 연장선상에서 본 연구는 생성형 AI가 한국어 시제에 대해서는 어느 정도의 문법 능력을 보유하고 있는지를 평가하고 오류 분석을 통해 생성형 AI가 시제를 이해하고 생성하는 기제를 유추하며, 나아가 AI의 한국어 시제 능력을 높일 수 있는 프롬프트 엔지니어링 방안에 대해서도 논의를 시도한다.

시제(時制, tense)는 시간과 관련된 문법 범주의 하나로 언어 보편적으로 존재하는 동시에 언어마다 개별성을 지닌다. 한국어에서는 '-었-', '-는-', '-겠-' 등 시간의미를 나타내는 선어말어미, 즉 문법 형태소가 '과거-현재-미래'의 대립을 이루며 시제 문법 범주를 이룬다. 본 연구는 남기심·고영근 외(2019)의 『새로 쓴 표준 국어문법론』의 '시제' 관련 장절의 예문 100여 개로 평가 세트를 구축하되, 평가 문장에서 지나치게 빈도가 낮거나 난이도가 높은 어휘 등은 교체하며, 마땅한 정문이 없는 비문은 제외하거나 교체하는 등의 방식으로 평가 세트를 정교화한다. 평가는 이해 능력 측면의 문법성 여부를 판단하는 테스트와, 생성 능력 측면의 비문을 정문으로 교정하는 테스트를 동시에 수행한다. 테스트 결과에 대해서는 시제 실현의 유형별로 나누어 면밀한 오류 분석을 수행하며, 여러 AI 간 비교 분석을 수행한다.

특히 앞선 AI 평가 연구들이 단일 프롬프트로 단 1회씩 실험을 수행하여 AI의 사용에서 프롬프트 엔지니어링의 중요성이 간과된 측면이 있다면, 본 연구에서는

생성형 AI의 작동 메커니즘을 고려하여 프롬프트를 다양하게 설계하고 입력하며 멀티 턴의 상호작용을 통한 방식으로 평가를 수행하고자 한다. 이로써, 생성형 AI의 문법 능력에 대한 연구를 수행하는 동시에 생성형 AI의 문법 능력 향상을 위한 방안 모색에도 기여할 수 있기를 기대한다.