

## 한국어 읽기 관련 LLM attention과 인간 시선추적 데이터 비교

최재원

(한국외국어대학교, 한양대학교)

남윤주

(한양대학교)

국내에는 다양한 형식의 언어 자원이 존재하지만, 영어나 일본어와 달리 안구 운동 및 뇌파 정보가 포함된 인지신경학적 코퍼스는 전무하다. 안구운동 추적 데이터는 인간의 언어 처리 과정에서 작동하는 인지적 요인을 규명할 수 있다는 점에서 큰 의의를 가진다. 남윤주(2023-2028)는 한국연구재단 보호연구지원사업의 지원으로 한국어 처리 관련 인지신경데이터베이스 구축(뇌파 및 시선추적 데이터베이스)을 진행 중이며, 본 연구에서는 이 중 일부 데이터를 활용하여 연구를 진행하였다.

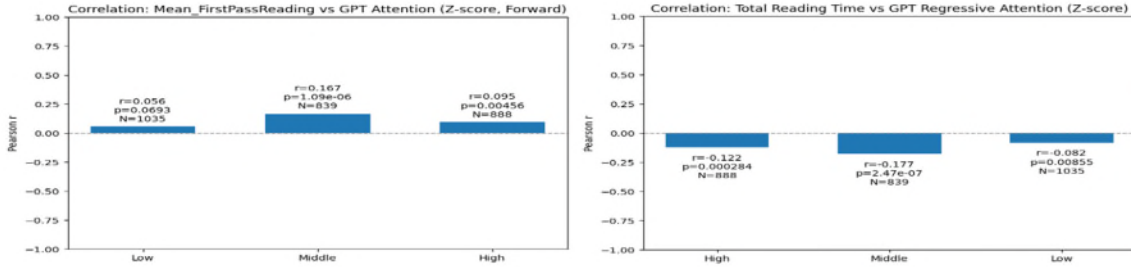
문장을 읽는 동안 어디에 주의를 할당할 것인가는 인간의 읽기뿐만 아니라 LLM의 언어 처리에 있어서도 매우 중요한 문제이다. 지금까지 진행된 알파벳 언어에 대한 연구들에 따르면 독자는 문장이나 절의 마지막 단어를 처리하는 데 더 많은 시간을 할애하는 경향이 있으며, 이를 흔히 '마무리 효과(wrap-up effect)'라고 한다(Rayner et al., 2000). 그러나 최근 중국어 안구운동추적연구에 의하면 자연스럽게 띄어쓰기가 없는 표의어인 중국어에서는 wrap-up effect가 보고되지 않았으며, 이는 문자 및 언어적 특성에 따라 초점이 맞춰지는 어휘 혹은 언어 성분이 달라질 수 있으며, 마지막 어휘에서 읽기 시간이 길어지는 wrap-up effect의 발생 여부 역시 달라질 수 있음을 의미한다. 그리고 만약 그렇다면, transformer model의 attention 역시 언어에 따라 달라져야만 효과적인 자연어 처리가 가능할 것이라고 유추할 수 있다.

이러한 문제의식을 바탕으로 본 연구에서는 한국어로 구성되고, 사전 난이도 검증용 결친 텍스트를 처리하는 70명의 한국인 모국어 화자의 안구운동을 수집·분석하여 인간의 시선 처리 지표와 언어 모델의 Attention 값을 비교·분석함으로써, 1) 한국어에서의 wrap-up effect 발생 여부 및 2) 인간 및 LLM의 Attention 일치 여부를 검증하고자 하였다. 단방향 모델의 경우 회귀 과정이 없으므로 초기 읽기 시간과의 상관을 비교하였고, 회귀 유사 모델의 경우 회귀 과정이 있으므로 전체 읽기 시간과의 상관을 비교하였다.

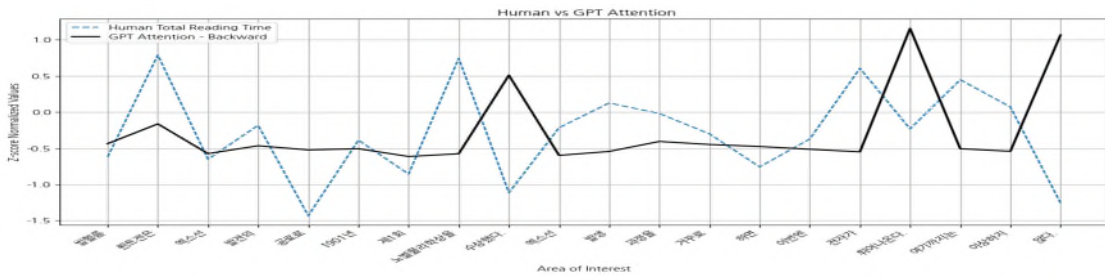
실험 결과, 단방향 방식의 “KoGPT2”는 인간의 읽기 관련 안구운동 중 초기 읽기 시간과 약한 양의 상관 관계를 보였으며((1a) 그림 1 참고), 회귀적 시선을 모사한 “KLUE-BERT(masking 기반 회귀 유사 처리로 일부 변형)”는 전체 읽기 시간과 약한 음의 상관 관계를 나타냈다((1b), 그림 2 참조). 세부 문장 단위 분석에서는, 인간의 시선은 “아인슈타인”, “노벨물리학상” 등 고유명사\*전문 용어에서 읽기 시간(Fixation time)이 길게 나타나 낮은 단어 처리 과정에서 인지 부하가 발생하였음을 확인할 수 있다. 반면에 GPT attention은 해당 구간에서 뚜렷한 변화가 없고, 오히려 “튀어나온다/않다”와 같은 문장의 끝 부분에서 급격히 값이 증가하는 술어 중심 attention 혹은 wrap-up effect와 유사한 패턴을 보였다((2) 그림 3). 이는 곧 한국어 독자가 보여주는 읽기 중 주의 할당 혹은 처리의 중요성 부여와 GPT 모델에서의 attention이 상반되고 있음을 의미한다.

추가적으로, 회귀 지표를 활용하여 비교한 결과, 모델의 회귀 빈도와 인간의 시선 고정 횟수 간에는 중간 수준의 양의 상관 관계가 나타났으며, 모델의 회귀 강도와 인간의 총 읽기 시간 간에는 유의한 양의 상관 관계를 확인하였다. 이는 앞서 비교했던 attention 분석과는 달리 모델의 회귀 유사 행동은 인간의 회귀와 유사하게 작용을 하고 있음을 시사한다.

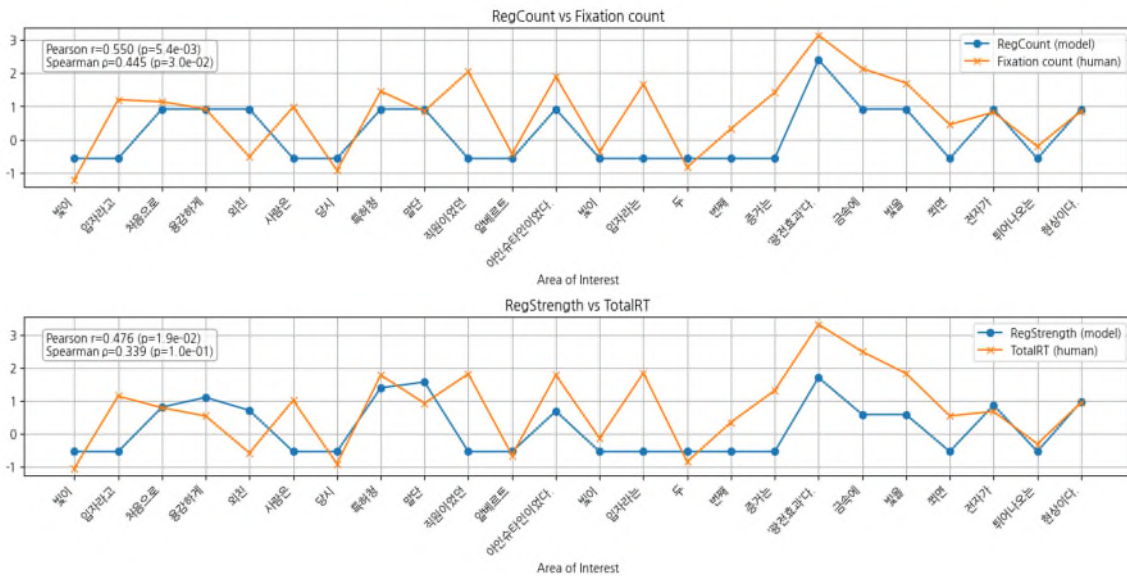
본 연구는 현재까지 텍스트 전체 단위의 상관 비교에 초점을 두었기에 향후 구문 정보나 의미적 빈도 등을 고려하여 더욱 세부적인 분석을 진행할 예정이다. 나아가 인간의 시선 처리와 GPT attention 간 차이를 규명함으로써, 한국어 기반 인지 특성을 반영한 새로운 언어모델을 개발하는데 있어 기반을 마련할 수 있을 것이다.



- (1) a. 그림 1. KoGPT2 attention과 인간 초기 읽기 시간의 상관관계
- b. 그림 2. KLUE-BERT attention과 인간 전체 읽기 시간의 상관관계



(2) 그림 3. 문장별 인간 시선과 GPT attention 패턴 비교



(3) 그림 4. 모델 회귀 지표와 인간 시선 지표의 상관관계

**참고문헌**

Keith Rayner, Gretchen Kambe, and Susan A. Duffy. 2000. The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology Section A*, 53(4):1061–1080.

Xue J, Borogjoon Bao Y, Xie D, Chen G, Kuperman V. 2025. No wrap-up effect in Chinese reading: Evidence from eye movements. *Q J Exp Psychol (Hove)*. Epub ahead of print. PMID: 40410941.