

문법 학습 비계 설정자로서 ChatGPT 활용을 위한 과제와 방향

박종미 | 국립국어원 연구원

- * 이 논문은 제83회 국어교육학회 학술발표대회(2024.12.7.)에서 발표한 것을 수정·보완한 것이다.

- I. 머리말
- II. ChatGPT의 교육 활용 동향
- III. 문법 비계 설정자로서 ChatGPT 분석
- IV. ChatGPT 활용을 위한 과제와 방향
- V. 맺음말

I. 머리말

인공지능(AI)에 대한 관심이 전 세계적으로 확산됨에 따라, 정치, 사회, 경제 등 거의 모든 분야에서 AI 기술이 활발히 도입되고 있다. 교육 분야 또한 예외가 아니다. AI는 오랜 기간 교육계에서 제안된 이상적 목표들을 실현 할 수 있는 핵심 기술로 주목받고 있다. 예를 들어, 서술형 평가의 중요성은 교육계에서 지속적으로 강조되어 왔으나, 채점 업무의 과중함 등 현실적 제약으로 인해 널리 도입되지 못했다. 그러나 최근 개발되고 있는 AI 기반 쓰기 자동 채점 시스템은 이러한 문제를 해결하여 서술형 평가 확대의 가능성 을 제시하는 기술로 부각되고 있다. 또한, 학습자의 수준과 특성에 맞춘 1:1 맞춤형 비계 설정은 가장 효과적인 학습 방식으로 평가받아 왔으나, 교사 한 명이 다수의 학생을 지도해야 하는 현실에서는 실현이 어려운 이상으로 여겨져 왔다. 그러나 기술의 발전으로 AI가 교사를 보조하여 개인화된 학습을 지원하는 것이 더 이상 불가능하지 않다. 다만, 이를 현실화하기 위해서는 이미 개발된 AI 기술을 교육 상황에 적용해 보고, 그 과정에서 발생하는 문제 점과 한계를 지속적으로 분석하여, AI를 교육적 목적에 부합하도록 개선해

나가는 노력이 필요하다. 본 연구는 대규모 언어 모델(LLM) 기반 AI 챗봇인 ChatGPT가 문법 학습에서 학생 개인별 맞춤형 비계 설정자로서의 역할 수행 가능성을 검토하고, 향후 이 기술을 문법 교육에 효과적으로 활용하기 위해 필요한 과제와 방향을 제시하고자 한다.

- 연구 문제 1: 현재의 ChatGPT는 문법 교육에서 개인 맞춤형 비계 설정자 역할을 수행할 수 있는가?
- 연구 문제 2: 문법 교육에서 ChatGPT 활용을 위해 해결해야 할 과제와 나아갈 방향은 무엇인가?

II. ChatGPT의 교육 활용 동향

1. ChatGPT의 개요

ChatGPT는 OpenAI에서 개발한 대규모 언어 모델(Large Language Model, LLM) 기반의 AI 챗봇이다. 언어 관련 과제를 수행하며 사용자의 요구에 맞춘 맞춤형 응답을 생성할 수 있다. 2022년 OpenAI에서 ChatGPT를 공개한 이후, 사람들은 AI가 자연어를 사용해 인간처럼 대화하는 모습에 큰 충격을 받았으며, 공개 일주일 만에 사용자 수가 100만 명을 넘어섰다(이학무, 2023).

ChatGPT는 범용 기술(General-Purpose Technology, GPT)에 해당한다. 범용 기술은 세계 경제 성장을 이끄는 근본적 기술로, 인쇄 기술과 증기 엔진 기술 등이 그 예에 해당한다(노대원, 2023; 이승윤, 2023). ChatGPT가 이러한 혁신적 기술들과 견줄 만한 기술로 평가받는 이유는, 전문 영역과 생활 영역을 막론하고 거의 모든 분야에서 인간과 대화하며 필요한 정보를 제

공하고 작업을 수행해 업무의 효율성을 높이기 때문이다.

교육의 영역에서도 마찬가지이다. 현재는 ChatGPT가 교육의 혁신을 가시적으로 보여주고 있는 단계는 아니지만, 학습자의 수준과 요구에 맞춘 개인화된 학습을 지원하거나 교사의 업무를 경감하는 등 그 활용 가능성이 무한정하다. 우리나라에서도 2023년을 기점으로 이와 관련한 연구들이 활발하게 진행되고 있다.

2. 국어 교육에서의 ChatGPT 활용 선행 연구

국어 교육에서 ChatGPT의 활용에 대한 연구는 화법, 작문, 독서, 문학, 문법의 전 영역에서 논의되고 있다. 특히 문학과 작문 영역에서의 논의가 활발하다. 강우규(2023)는 문학 교육에서 ChatGPT를 활용한 텍스트 재해석과 창작 활동 지원 방안을 제안하며, 〈춘향전〉 다시 쓰기 사례를 통해 그 효과를 검토하였다. 노대원(2023) 역시 문학 교수학습 과정(강의 준비, 강의 진행, 창작 활동 등)에서 AI 활용 방안을 제시하며 AI 적용 가능성을 논의하였다. 또한 양수연·고유라·신효정·주세형(2023)는 ChatGPT가 문학 작품 〈춘향전〉과 〈클라라와 태양〉을 기반으로 생성한 교육적 질문을 교사가 생성한 질문과 비교·분석하여, ChatGPT의 교육적 질문 생성 역할에 대한 가능성을 검토하였다.

쓰기 영역에서도 ChatGPT의 활용 가능성은 활발히 연구되고 있다. 사실 쓰기 영역에서는 ChatGPT의 등장 이전부터 AI 기술에 주목해 왔는데, 일찍이 이용상·시기자·박도영·윤경아·구슬기·임황규(2013) 등의 연구에서 쓰기 자동 채점 시스템을 연구한 바 있으며, 이 기술에 대한 연구는 이진·정진경·김한샘(2024) 등을 통해 현재까지 진행되고 있다. 최근 ChatGPT의 등장으로 쓰기 피드백 방안이나 쓰기 문항 개발과 관련한 연구도 활발히 이루어지고 있다. 최숙기·박종임(2024)은 ChatGPT가 교사의 쓰기 문항 개발 과정을 지원함으로써 교사의 업무 부담을 줄이는 데 기여할 수 있음을 제안

하였다. 이 연구는 ChatGPT 활용을 위해 교사에게 요구되는 새로운 역량에 대한 고민을 제기하였으며, 그 역량으로 프롬프트 개발 역량과 ChatGPT 응답에 대한 비판적 재구성 역량을 제시한 점에서 의의가 있다.

또한 문학과 쓰기 영역 외에도 이소희·김갑수(2023)와 같이 초등학교 국어 교과의 어휘 학습을 지원하는 AI 챗봇을 직접 개발하는 등 문법과 다른 영역들에서도 ChatGPT 관련 연구들이 진행되고 있다.

3. 비계 설정자로서의 ChatGPT 활용 선행 연구

앞서 언급한 바처럼 ChatGPT는 범용 기술로서, 적용 영역과 적용 목적에 제한이 없다. 본 연구에서는 ChatGPT가 수행할 수 있는 무한한 역할 중에서도 특히 ‘비계 설정자(Sccaffolder)’로서의 역할에 주목하고자 한다.

2024년, ChatGPT는 멀티모달(Multimodal) 기술의 적용으로 성능이 한층 향상되었다. 기존에 텍스트만 입출력이 가능했던 것에 비해, 현재는 텍스트뿐 아니라 이미지, 오디오, 비디오에 대해서도 입출력이 가능해지며 학습 지원 도구로서의 잠재력이 더욱 확대되었다(권미선, 2024). OpenAI에서는 2024년 멀티모달 기술이 적용된 ChatGPT를 공개하면서 ChatGPT가 학습자의 수학 문제 해결을 단계적으로 안내하는 영상(OpenAI, 2024)을 활용 사례로 제시하였다. 이 영상에서는 학습자의 수준에 맞춘 음성 대화를 통해 학습자가 스스로 문제를 해결할 수 있도록 질문하고 수행을 요청하며 유도하는 ChatGPT의 기능이 강조되었다. 이는 ChatGPT가 수학 학습 상황에서 맞춤형 비계 설정자로서의 역할을 수행할 수 있음을 보여준다.

이러한 ChatGPT의 기능과 관련하여 국내에서도 수학 교육 연구들이 진행되고 있어 주목할 만하다. 손태권(2023)은 초등학교 3학년 학생의 분수 문제 해결 과정을 피드백하는 예비 교사 2인의 녹음 담화를 분석하여, 교사의 피드백 적절성을 전문가와 ChatGPT에게 각각 평가하도록 하였다. 비교 결과, 전문가와 ChatGPT의 평가에는 유사점과 차이점이 동시에 발견되었

는데, 이 연구에서는 ChatGPT가 학생의 지식과 사고를 파악하고 효과적인 교수·학습을 지원하는 교사의 보조 도구로 활용될 가능성이 있음을 제안하였다(손태권, 2023: 108).

안도연·손태권·이광호(2023)는 초등학교 6학년 24명을 대상으로 ChatGPT의 효과성에 대한 실험을 진행하였다. 학생들이 수리 논리 과제를 해결하는 과정에서 ChatGPT가 제공한 힌트가 문제 해결에 미치는 영향을 평가하였고, ChatGPT가 제공한 힌트의 적절성을 비계 설정 전략의 관점에서 분석하였다. 실험 결과, ChatGPT의 힌트를 받은 학생들은 문제 풀이 수준이 유의미하게 향상되었으며, ChatGPT가 제공한 힌트는 비계 설정 전략 중 ‘재구조화’ 전략을 가장 많이 사용하였고, 학습자의 응답 수준에 맞게 적절히 조정되었다. 이러한 결과는 ChatGPT가 초등학생의 수리 학습에서 비계 설정 도구로 활용될 수 있음을 시사한다.

이처럼 수학 교육에서 ChatGPT가 비계 설정자 역할을 수행할 가능성 이 논의되고 있는 것과 같이, ChatGPT는 문법 교육에서도 학습자의 문법 문제 해결을 돋는 비계 설정자 역할에 대한 수행 가능성이 있다. 다만, 한국어를 대상으로 한 학교 문법 지식을 다루는 교과 영역의 특수성을 고려할 때, 영어에 최적화된 ChatGPT의 한국어 성능(학교 문법 기준)을 확인하고 개선 방안을 탐색하는 연구가 필요하다. 이에 본 연구에서는 ChatGPT의 문법 문제 해결 과정에서의 비계 설정자 역할 수행 가능성을 확인하고, 향후 ChatGPT를 문법 교육에 활용하기 위해 해결해야 할 과제를 제시하고자 한다.

III. 문법 비계 설정자로서 ChatGPT 분석

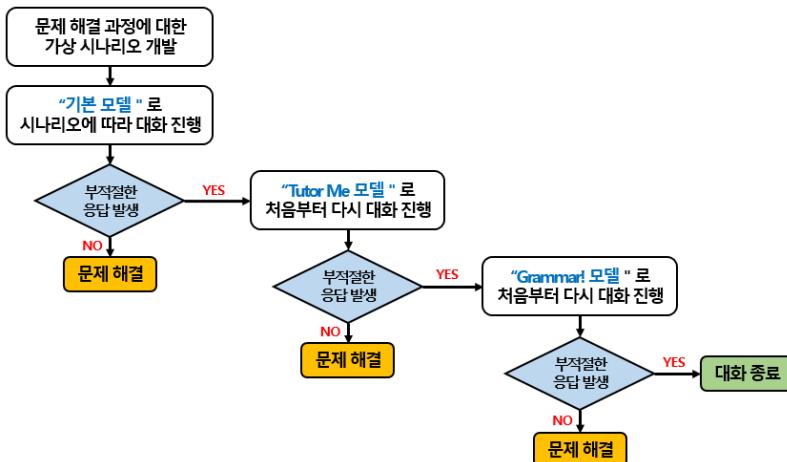
1. 시연 과정

본 장에서는 멀티모달 기능이 적용된 ChatGPT-4o 모델을 대상으로 문법 문제 해결을 목적으로 한 대화를 시연함으로써, 현재 ChatGPT 모델의 문법 비계 설정자 역할 가능성과 한계점을 분석하였다. 통상적으로 AI 모델의 성능 검증에는 별도로 마련된 평가용 문제가 활용되지만, ChatGPT는 문법 비계 설정자라는 특정 역할을 수행하도록 개발된 모델이 아니며, 이에 대한 성능 검증 문항이나 방법이 아직 체계적으로 논의된 바 없다. 따라서 본 연구는 엄밀한 성능 검증보다는 ChatGPT가 이러한 역할을 수행할 가능성을 탐색적으로 확인하는 데 중점을 두었다. 향후 문법 비계 설정자 역할을 수행할 수 있는 AI 시스템의 개발과 함께, 이를 체계적으로 검증할 별도의 평가 문항과 척도 개발이 필수적으로 요구된다.

시연은 학습자가 고등학교 『언어와 매체』 교과서에 등장하는 문법 학습 활동을 해결하는 과정을 가상 시나리오로 작성한 후, 해당 시나리오를 바탕으로 ChatGPT-4o와 대화를 진행하는 방식으로 이루어졌다.

시연에는 총 세 가지 ChatGPT 모델이 사용되었다. 가장 먼저 사용된 모델은 “기본 모델(“<https://chatgpt.com/>”에 접속하여 바로 사용할 수 있는 모델)”이며, 대화 도중 기본 모델이 비계 설정자로서 부적절한 응답을 할 경우 대화를 종료하고 두 번째 모델인 “Tutor Me 모델”로 대화를 재시작하였다. Tutor Me 모델은 미국의 비영리 교육 단체인 Khan Academy에서 기본 모델에 ‘튜터’의 역할을 학습시킨 GPTs 모델이다. 이 모델은 기본 모델에서 발생한 문제를 GPTs 모델을 통해 극복 가능한지 확인하기 위해 활용되었다. Tutor Me 모델에서도 비계 설정자로서 부적절한 응답이 나타날 경우, 대화를 종료하고 세 번째, “Grammar! 모델”로 대화를 재시작하였다. Grammar!

모델은 연구자가 직접 기본 모델이 한국어 문법 학습의 튜터로서 역할을 수행할 수 있도록 학습시킨 GPTs 모델로, Tutor Me 모델에서 발생한 문제를 극복하기 위한 대안으로 개발되었다. 시연 과정을 그림으로 나타내면 다음과 같다.



〈그림 1〉 ChatGPT 시연 과정

이 과정에서 한 가지 모델로 대화를 끝까지 진행한 후 다른 모델을 적용하지 않고, 문제가 발생하면 그 자리에서 대화를 종료한 이유는, 대화 과정에서 ChatGPT가 엉뚱하거나 잘못된 응답을 할 경우 그 응답에 기반하여 문제 해결을 위한 대화를 지속하기 어렵기 때문이다. 따라서 특정 모델에서 부적절한 응답이 발생하면 대화를 중단하고 다른 모델로 재시작하는 방식을 채택하였다. 학습 방식이 다른 세 가지 모델을 순차적으로 적용한 시연 과정을 통해, 향후 ChatGPT를 문법 학습에서 비계 설정자로 활용하기 위해 기본 모델을 기반으로 어떤 방식으로 AI를 학습시키고 모델을 개발해야 하는지에 대한 시사점을 얻을 수 있다.

시연에 사용된 학습 활동은 『언어와 매체』(이관규·박경희·신호철·신희

성 · 이동석 · 정지현 외, 2019) 교과서, 문장의 짜임 단원에 등장하는 다음의 문제이다.

3 안은문장과 안긴문장에 대해 알아보자.

- Ⓐ 우리는 그가 정당했음을 깨달았다.
- Ⓑ 그는 이마에 흐르는 땀을 닦았다.
- Ⓒ 토끼는 앞발이 짧다.

(1) Ⓛ~Ⓒ에서 절을 찾아 그 종류를 쓰고, 각 절의 원래 홀문장을 적어 보자.

	절	원래의 홀문장
Ⓐ	그가 정당했음을 명사절	그가 정당했다.
Ⓑ		
Ⓒ		

〈그림 2〉 문법 학습 활동(이관규 외, 2019)

위 학습 활동은 안은문장과 안긴문장에서 ‘절’을 찾아 그 종류를 쓰고, 해당 절의 원래 홀문장 형태를 적어 보는 활동이다. 이 활동을 선정한 이유는 다음과 같다.

첫째, 문제 해결을 위해 ‘절’이라는 선행 지식이 필수적이라는 점에서, 선행 지식이 부족한 학습자가 문제 해결에 어려움을 겪는 상황을 시나리오에 포함함으로써 ChatGPT가 학교 문법에 대한 적절한 지식 정보를 학습자에게 제공할 수 있는지를 확인할 수 있다.

둘째, 학습자가 ‘절’이라는 선행 지식을 단순히 기억하는 수준을 넘어 새로운 문장에 적용해야 하는 과정을 요구한다. 이 과정에서 어려움을 겪는 학습자의 상황을 시나리오에 반영함으로써, ChatGPT가 단순히 지식을 제공하는 것을 넘어 학습자의 지식 적용 과정을 평가하고 필요한 비계를 제공할 수 있는지를 검토할 수 있다.

셋째, 이 활동은 우리 국어의 특수성에 해당하는 서술절을 포함하고 있

다는 점에서 의의가 있다. 이는 영어에 최적화된 ChatGPT가 한국어의 고유한 문법적 특성을 다룰 수 있는지를 평가할 기회를 제공한다. 특히, 서술절을 포함한 안은문장에 대한 ChatGPT의 정보 제공 능력을 시험함으로써 한국어 문법 학습에서 ChatGPT의 활용 가능성을 평가할 수 있다.

이 활동에 대한 학습자의 문제 해결 과정을 다룬 시나리오는 다음 〈표 1〉과 같다. 시나리오는 예상되는 발화 전체를 상세히 작성하지 않고, 전체 대화의 흐름에서 반드시 포함해야 할 상황만 반영하는 방식으로 설계되었다. 이는 시나리오의 유연성과 실제 활용 가능성을 높이기 위함이다.

〈표 1〉 학습자의 문제 해결 과정 대화 시나리오

단계	시나리오
1단계	학습자: 흔자 해결하기 어려운 문제가 있어. 문제를 업로드하면 도와줄래? ChatGPT: ... 학습자: (문제 사진 업로드) 이 문제를 어떻게 풀어야 할지 모르겠어. ChatGPT: ...
2단계	학습자: '절'이 뭔지 아직 잘 모르겠어. '절'이 뭐야? ChatGPT: ...
3단계	학습자: 두 번째 문장에서 절을 찾아볼게. 먼저 주어부터 찾으면 '그는'. 주어가 '그는' 한 개뿐이니까 이 문장은 홀문장인 것 같은데? 절이 없는 홀문장이야. ChatGPT: ...
4단계	학습자: '토끼는 앞발이 짧다.'가 서술절을 안은 문장이라는 걸 이해가 되는데 왜 홀문장으로 볼 수는 없는 걸까? '토끼의 앞발이 짧다.'에서 '토끼'를 강조하기 위해 '의' 대신에 '는'이라는 보조사가 사용된 것 같은데. ChatGPT: ...

시나리오의 1단계는 학습자가 ChatGPT에게 도움을 요청하는 단계이다. 이 단계에서 학습자는 해결해야 할 문법 문제를 업로드하고 어떻게 풀어야 하느냐고 질문한다.

2단계는 학습자가 문제 해결에 필요한 문법 개념인 절에 대한 선행 지식이 없어 정보를 요청하는 상황이다. 이 상황은 ChatGPT가 학교 문법 지식에 대한 정확한 정보를 제공할 수 있는지를 검토하기 위해 설정되었다.

3단계는 학습자가 문법 개념을 과도하게 단순화하여 이해함으로써 적용에 실패하는 상황이다. 이관희·조진수(2015)는 교사를 대상으로 “문법 개념의 과도한 단순화”에 해당하는 오개념 사례를 제시한 바 있는데 이를 참고하여 학습자가 가질 법한 오개념 시나리오를 작성한 것이다. 절은 주어와 서술어를 갖춘 단위로 정의되지만 실제로는 문맥에 따라 성분이 생략될 수 있다. 그러나 이 점을 학습자가 간과하고 절을 단순히 표면적으로 주어와 서술어를 갖춘 단위로만 인식하여 개념 적용에 실패한 상황을 시나리오에 반영하였다.

4단계는 학교 문법이 절대불변인 정답이 아니라 합의에 의한 점임을 충분히 이해하지 못한 학습자가 혼란을 겪는 상황을 가정한 것이다. 이관희·조진수(2015)는 이러한 유형의 오개념을 “학교 문법의 합의점에 대한 이해 부족”으로 명명한 바 있다. 또한 이 단계에서는 ‘서술절’이라는 한국어의 특수성에 대한 ChatGPT의 설명을 검토해 볼 수 있도록 의도되었다.

2. 분석 방법

ChatGPT 응답은 “비계 설정 전략”과 “학교 문법 지식”이라는 두 가지 범주에서 평가되었다. 각 범주에 대한 구체적인 분석 기준은 다음과 같다.

〈표 2〉 ChatGPT 응답 분석 방법

분석 범주	분석 기준	분석 결과
비계 설정 전략	ChatGPT는 Hogan & Pressley(1997)에서 제시한 14가지 비계 전략 중 하나 이상을 적절히 사용하여 응답하였는가? ¹⁾	적절/부적절
학교 문법 지식	ChatGPT는 정확한 학교 문법에 기반하여 응답하였는가?	환각 있음/환각 없음

1) Hogan & Pressley(1997)에서 제시한 14가지 비계 전략은 다음과 같다. 1) 학습자의 요구를 복돋우고 세분화된 목표 설정하기, 2) 목표 추구를 유지하기(수준에 따라 과제 수준

우선, 비계 설정 전략의 측면에서는 ChatGPT의 응답이 Hogan & Pressley(1997)에 제시된 14가지 비계 전략 중 어느 하나에 해당하며, 해당 상황에서 적절하다고 판단될 경우 ‘적절’, 비계 전략에 해당하지 않거나 상황에서 적용하기 부적절한 전략으로 판단될 경우 ‘부적절’로 분석하였다.

학교 문법 측면에서는 ChatGPT의 응답이 학교 문법 지식과 일치하는지를 평가하였다. AI가 잘못된 정보를 사실인 것처럼 제공하는 현상을 ‘환각(hallucination)’이라고 하는데, ChatGPT의 응답에 잘못된 학교 문법 지식이 하나라도 포함되어 있을 경우 ‘환각 있음’, 그렇지 않을 경우 ‘환각 없음’으로 분석하였다.

3. 분석 결과

대화는 가장 먼저 ‘기본 모델’을 대상으로 시연되었다. 대화의 1단계는 학습자가 문제를 제시하고 도움을 요청하는 단계로, 기본 모델은 ‘어려운 문제가 있어서 도와달라.’고 하는 학습자의 요청에 “네. 물론이죠! 문제를 올려 주시면 풀이 과정을 차근차근 도와드리겠습니다.”라고 응답하였다. 그러나 막상 학습자가 문제 이미지를 업로드하자 다음과 같이 문제에 대한 정답을 쏟아내기 시작했다.

을 조절하여), 3)상호작용에 있어 대화적 유형 사용하기, 4)학생의 인지적 정의적 상태와 능력의 수준을 알고자 노력하기, 5)학습자의 사고를 직접적으로 평가하지 않는 피드백 제공하기(간접적 지시), 6)사고 과정을 시범보이기, 7)학습자의 실습을 지도하고 조력하기, 8)설정된 조력을 제공하기, 9)필요에 따라 설명 제공하기, 10)학습자의 발언과 질문을 고무하고 이용하기, 11)학습자가 답을 유출하기 위해서 서로 협의하는 동안 인내하기, 12)정서적인 지지와 양성의 높은 수준을 보여주기, 13)학습자의 실수에 유동적으로 반응하기, 14)교사의 지도와 강도 줄이기. Hogan & Pressley(1997)의 비계 전략을 활용한 것은 비교적 다른 연구에 비해 전략이 다양하고 상세하여 예측하기 어려운 ChatGPT의 응답을 분석하는 초기의 기준으로 적절하다고 판단하였기 때문이다.

☞ 올려주신 이미지를 보니, 안은문장과 안긴문장에 대한 문제입니다. 함께 풀이 과정을 살펴보겠습니다.

문제 풀이 과정:

1. 문제 1번 풀이

- 각 문장에서 절을 찾아 그 종류를 쓰고, 각 절의 원래 풀문장을 작성해야 합니다.

(1) 예시 문장

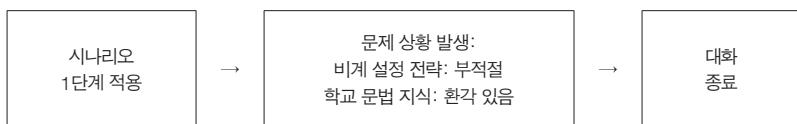
- ① "우리는 그가 정당했음을 깨달았다."
 - 절: 그가 정당했음 (명사절)
 - 원래 풀문장: 그가 정당했다.
- ② "그는 이미에 흐르는 땀을 닦았다."
 - 절: 이미에 흐르는 (관형절)
 - 원래 풀문장: 땀이 이미에 흐른다.
- ③ "토끼는 앞발이 짧다."
 - 절: 앞발이 짧다 (명사절)
 - 원래 풀문장: 앞발이 짧다.

이렇게 각 절의 종류와 원래 풀문장을 작성해 보세요.

〈그림 3〉 ChatGPT4.0와의 대화(2024.11.8.)

위와 같은 응답 방식은 Hogan & Pressley(1997)의 비계 설정 전략 중 어느 하나에도 해당하지 않는 부적절한 것이다. 따라서 비계 설정 전략 범주에서 '부적절'로 평가되었다. 또한, 응답에 부정확한 학교 문법 지식도 포함되어 있었다. 기본 모델은 "토끼는 앞발이 짧다"에서 "앞발이 짧다."를 명사절로 분류하고 있으나, 학교 문법에서는 이를 서술절로 분류한다. 따라서 학교 문법 지식 측면에서도 '환각 있음'으로 판정되었다.

이처럼 기본 모델은 대화의 1단계에서 비계 설정자로서의 역할 수행이 어렵다는 점이 확인되었고 대화는 종료되었다.



〈그림 4〉 '기본 모델'의 시연 결과

다음 시연에 사용된 모델은 'Tutor Me'이다. Tutor Me는 기본 GPT 모델에 기반하여 칸아카데미에서 '튜터'라는 특수한 역할을 부여한 GPTs 모델이다. Tutor Me 모델은 아래 그림과 같이 문제 풀이를 도와달라는 학습자의 요청에 대해 학습자가 첫 번째로 해야 할 일을 제안하며 학습자가 스스로 해결해 볼 수 있도록 유도하였다.

이 문제를 어떻게 풀어야 할지 모르겠어.

 좋아! 문제를 같이 풀어보자. 먼저, 문제의 핵심은 주어진 문장에서 "절"을 찾아서 그 종류를 분류하고, 원래의 흐문장을 원성하는 것이야.

첫 번째로 할 일:

1. "절"이 무엇인지 생각해 보자.
 - "절"은 주어와 서술어가 있는 단위로, 문장 안에서 더 큰 역할(명사, 형용사, 부사 등)을 해.
2. 각 문장에서 절을 찾아봐. 예를 들어, ①에서는 "그가 정당했음"이 절로 보이지?
3. 이 절이 어떤 역할을 하고 있는지(명사, 형용사, 부사)를 생각해 봐. 그리고 절이 원래는 어떤 독립된 문장이었는지도 살펴봐야 해.

이제 ②와 ③을 보며 질문을 하나씩 해 볼게:

- ②에서, "그는 이미 흐르는 땅을 밟았다." 이 문장에서 **흐르는 땅**은 어떤 절일까? 그리고 그 원래 문장은 어떻게 구성되었을까?

어떤 생각이 떠오르는지 알려줘! 😊

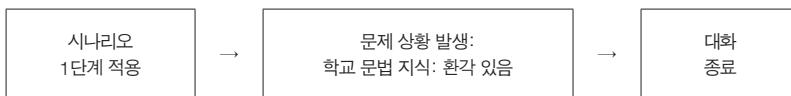
〈그림 5〉 Tutor Me와의 대화(2024.12.19.)

Tutor Me는 학습자에게 '절'이 무엇인지 생각해 보도록 요청하며, 문제에서 예시로 제시된 문장 "우리는 그가 정당했음을 깨달았다."에서 "그가 정당했음"이 절에 해당한다고 예를 들어 설명하였다. 이후, 학습자에게 다른 문장에서 절을 찾아 그 역할을 파악해 보라고 안내하였다. 실제 교사가 학생에게 제공하는 지도에 비해서는 다소 미흡한 점이 있지만, Hogan & Pressley(1997)에서 제시한 스캐폴딩 전략 중 '세분화된 목표 설정하기' 전략을 활용하여 학습자가 단계적으로 문제에 접근할 수 있도록 지원한 점은 비계 설정자로서의 역할을 수행하려는 시도를 보여준다. 또한, 기본 모델과는 달

리 정답을 즉시 제공하지 않아 개선된 모습을 보여주었다.

그러나 Tutor Me 모델의 응답을 살펴보면 학교 문법 지식에 대한 환각 문제가 여전히 존재한다. Tutor Me는 ‘절’을 “주어와 서술어가 있는 단위로, 문장 안에서 더 큰 역할(명사, 형용사, 부사 등)을 한다.”라고 설명하였다. 이 설명은 ‘더 큰 역할’이라는 표현이 불분명하고, 학교 문법에서는 설정하고 있지 않은 ‘형용사절’을 상정하고 있다는 점에서 부적절하다. 또한, 응답의 마지막 부분에서 “이 문장에서 ‘흐르는 땀’은 어떤 절일까?”라고 학습자에게 질문하고 있는데, 실제로 문장에서 절에 해당하는 정확한 부분은 “이마에 흐르는”이다.

이처럼 Tutor Me 모델은 튜터 역할에 적합한 지침을 학습한 덕분에 비계 설정 문제는 발생하지 않았으나 여전히 학교 문법 지식의 환각 문제가 발견되어 대화가 종료되었다.



〈그림 6〉 Tutor Me 모델 시연 결과

마지막으로 시연에 사용된 모델은 Grammar!이다. Grammar!는 기본 모델에 기반하여 연구자가 직접 문법 학습에서 비계 설정자로서의 역할을 수행할 수 있도록 튜닝한 모델이다. 기본 모델에서 나타난 부적절한 비계 설정 전략 문제를 해결하기 위해 TutorMe 모델과 같이 튜터 역할을 수행하도록 별도의 지침 문서를 생성하여 학습시키는 동시에, 학교 문법 지식에 대한 환각 문제를 극복하기 위해, 문제 풀이에 필요한 지식을 정리하여 검색 증강 생성(Retrieval-Augmented Generation)에 활용함으로써 응답의 정확성을 향상시키고자 한 모델이다. 다음은 시나리오 1단계를 적용한 Grammar!와의 대화 장면이다.

이 문제를 어떻게 풀어야 할지 모르겠어.



좋아, 올려준 문제를 함께 풀어보자! 문제는 문장에서 "절"을 찾고 그 종류를 쓰며, 원래의 흔문장을 적는 것이야. 이미 첫 번째 문장(ㄱ)에 대한 예시가 제공되었으니 이를 참고하며 나머지를 풀어보자. 차근차근 단계별로 접근해 보자!

첫 번째 단계: 절을 찾기

문장에서 "절"은 주어와 서술어가 함께 있는 단위를 의미해. 예를 들어:

- ㄱ. "그가 정당했음" → "그"가 주어, "정당했음"이 서술어.

그럼 이제 두 번째 문장(ㄴ)부터 시작해 볼까?

- ㄴ. 그는 이마에 흐르는 땀을 닦았다에서, 문장의 서술어는 무엇인지 먼저 찾아볼래?

☞ ☞ ☞ ☞

〈그림 7〉 Grammar!와의 대화(1)(2024.12.19.)

Grammar! 모델은 문제 풀이를 요청한 학습자에게 '세분화된 목표 설정하기' 전략을 활용하여 학습자를 효과적으로 안내하였다. 동시에 학교 문법 지식 측면에서도 환각이 나타나지 않았다. 1단계에서 문제가 발생하지 않았기 때문에 대화는 2단계 시나리오로 이어졌다. 2단계는 학습자가 '절'의 개념을 잘 이해하지 못하여 선행 지식 부족을 드러내고 이에 대한 설명을 요청하는 상황이다.

"절"이 뭔지 아직 잘 모르겠어. 절이 뭐야?



좋은 질문이야! 🌟 그럼 ***절***이 무엇인지부터 차근차근 설명해 줄게.

절이란?

절은 주어와 서술어를 갖춘 단위야. 즉, 하나의 문장처럼 주어와 서술어가 갖추어진 덩어리를 말해.

예를 들어:

- "그가 정당했음"을 깨달았다."

여기서 **"그가 정당했음"**이 하나의 절이야. 왜냐하면 여기에는 "그가"(주어)와 "정당했음"(서술어)이 있기 때문이야.

절과 문장의 차이

- 문장: 독립적으로 완전한 의미를 가지는 말
- 절: 문장 안에서 주어와 서술어가 있는 단위

예시로 다시 살펴볼까?

〈그림 8〉 Grammar!와의 대화(2) (2024.12.19.)

이 상황에서 Grammar! 모델은 학습자에게 ‘절’의 정의를 비교적 명확히 제공하였으며 그 내용은 학교 문법 기준에서 적절하였다. 또한 “예시로 다시 살펴볼까?”와 같은 질문을 통해 학습자의 이해도를 확인하고자 하는 모습도 보였다. 이는 ‘학생의 인지적 정의적 상태와 능력의 수준을 알고자 노력하기’ 전략에 해당한다고 할 수 있다.

2단계에서도 문제가 발생하지 않았으므로 대화는 3단계 시나리오로 이어졌다. 3단계는 학습자가 절의 개념을 잘못 적용하여 오류를 범하는 상황이다.

“그는 이마에 흐르는 땀을 닦았다.”에서 절을 찾아볼게. 절은 주어와 서술어를 갖춘 데어리니까... 서술어는 ‘흐르는’인데 주어가 없네. 이거 훌문장이야?



오! 정말蛍い觀察だね! 🌟 주어를 찾아보려는 너의 시도가 정말 좋아. 그런데 주어가 눈에 보이지 않는다고 해서 절이 아닌 것은 아니야. 왜냐하면 절의 주어가 생략되는 경우도 있기 때문이야.

다시 생각해보자!

문장: “그는 이마에 흐르는 땀을 닦았다.”

- 1 **흐르는**은 동사(서술어)야.
2 주어는 보이지 않지만, 문맥상 **땀**이 '흐르는'의 주어 역할을 하고 있어.
즉, **흐르는**은 “땀”을 꾸며주고 있고, **땀**이 흐르다**로 바꿀 수 있어.

정리해보면:

- 절: 이마에 흐르는
- 주어: (생략된) 땀
- 서술어: 흐르는

이 절은 “땀”을 꾸며주는 관형절이야.

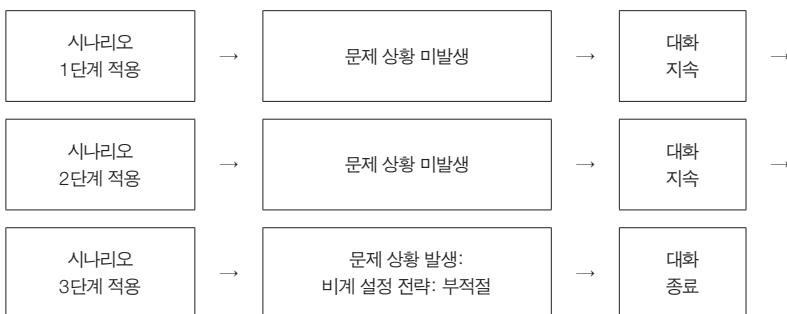
혹시 지금까지 이해되지 않는 부분이 있니? 아니면 원래 훌문장을 적어볼 수 있을까? 😊🌟

〈그림 9〉 Grammar!와의 대화(3) (2024.12.19.)

위의 대화에서 학습자는 단순히 절의 주어와 서술어가 표면적으로 드러나지 않았다는 이유로 “그는 이마에 흐르는 땀을 닦았다”를 훌문장으로 분류하는 오류를 범하였다. 이에 Grammar! 모델은 학습자의 시도를 긍정적으로 평가하며, 절의 주어가 생략될 수 있다는 점을 정확히 지적하고 다

시 생각해 보기 를 요청하였다. ‘정서적인 지지와 양성의 높은 수준을 보여 주기’와 ‘학습자의 실수에 유동적으로 반응하기’의 비계 전략을 사용한 것으로 보인다.

그러나 대화의 마지막 부분을 보면 Grammar! 모델은 ‘이마에 흐르는’ 을 관형절이라고 먼저 언급해 버려 학습자 대신 정답을 이야기하는 실수를 범하였다. 퓨터로서 정답을 이야기하면 안 된다는 지침을 학습시켰음에도 불구하고 의도대로 작동하지 않은 것이다. 이에 Grammar! 모델과의 대화는 3단계에서 종료되었다.



〈그림 10〉 Grammar! 모델 시연 결과

세 모델과의 대화 결과를 정리하면 다음과 같다.

〈표 3〉 ChatGPT 응답 분석 결과

ChatGPT 모델	시연 결과		
	시나리오 적용 단계	비계설정전략	학교문법지식
기본 모델	1단계	부적절	환각 있음
Tutor Me	1단계	적절	환각 있음
Grammar!	1단계	적절	환각 없음
	2단계	적절	환각 없음
	3단계	부적절	환각 없음

결과적으로, 세 모델 모두 대화를 완전히 종료하지 못했으며, 세 모델 모두 문법 비계 설정자 역할을 수행하기에는 적절하지 않은 것으로 평가된다. 그러나 기본 모델과 비교했을 때, 튜터 역할에 적합한 지침을 학습한 모델들은 비계 설정 전략 활용에서 기능이 개선되었으며, 학교 문법 지식에 대해 검색 증강 방식을 적용한 모델은 환각 현상이 감소하는 것을 확인할 수 있었다.

IV. ChatGPT 활용을 위한 과제와 방향

본 장에서는 Ⅲ장의 분석 결과에 근거하여 향후 ChatGPT를 문법 교육에서 맞춤형 비계 설정자로 활용하기 위한 과제를 제시하였다.

1. 과제(1): 문법 문제 풀이 데이터 구축

기본 모델과 Tutor Me 모델에서 나타난 학교 문법 지식의 환각 현상을 해결하기 위해, 본 연구에서는 문법 문제 풀이에 필요한 문법 지식을 정리한 데이터를 검색 증강 생성(Retrieval-Augmented-Generation) 방식으로 활용한 Grammar! 모델을 개발하였고, 실제로 환각 현상이 감소함을 확인하였다. 이러한 결과는 향후 ChatGPT를 포함한 AI가 문법 교육에서 학습자 맞춤형 비계 설정자로서의 역할을 수행하기 위해 검색 증강 생성 방식에 활용할 데이터베이스 구축이 필요함을 시사한다.

검색 증강 생성 방식은 AI 모델이 데이터를 직접 학습하는 데 그치지 않고, 실시간으로 외부 데이터베이스 문서를 검색하여 응답 생성에 활용하는 기술이다. 본 연구에서는 비전문가 수준에서 텍스트 기반 풀이 과정 데이터를 작성하여 검색에 활용하였으나, 향후 학교 문법 문제 전반에 대한 학습자

예상 수준별 문제 풀이 과정 데이터가, AI가 효율적이고 효과적으로 검색할 수 있는 전문적인 지식 표현 수준으로 생성되고 축적될 필요가 있다.

문법 문제 풀이와 관련된 다양한 예제, 규칙, 가능한 예외 사항 등이 데 이터베이스에 포함되어야 하며, 학습자가 자주 직면하는 문제 상황에 대한 정보도 포함되면 AI의 비계 설정자로서의 성능을 제고할 수 있을 것이며 문법 교육 전문가와 데이터 전문가의 협업이 필요한 영역이다.

이러한 데이터 구축 과정을 통해 검색 증강 방식의 효과를 극대화할 수 있으며, ChatGPT와 같은 AI 모델이 문법 교육에서 더욱 정확하고 유용한 비계를 제공하는 데 효과적인 학습데이터로 활용될 수 있을 것이다.

2. 과제(2): 교사의 비계 설정 발화 말뭉치 구축

AI의 기능 향상을 위해 대규모 말뭉치의 구축은 필수적이다. 기본 모델은 부적절한 비계 설정 전략을 사용하는 대화 방식이 문제로 지적되었으며, 이를 개선하기 위해 Tutor Me 모델과 Grammar! 모델에서는 튜터 역할에 대한 지침을 학습시키는 방식을 채택하였다. 이러한 접근은 일정 수준의 성과를 보였으나, 실제 교사의 역할을 완전히 대체할 만큼의 기능 향상에는 한계가 있었다.

AI가 학생의 문법 풀이 과정을 이해하고, 상황에 적합한 비계를 제공하기 위해서는 교사가 학생과의 상호작용에서 활용하는 비계 설정 발화를 체계적으로 수집하여 말뭉치로 구축하는 작업이 필요하다. 이러한 말뭉치는 교사가 학생의 문법 문제 해결 과정을 지원하는 실제 대화 데이터를 포함하며, 문법적 설명, 질문 유도, 힌트 제공 등 다양한 전략을 포함해야 한다. 이렇게 구축된 말뭉치를 AI의 학습데이터로 활용함으로써 AI의 비계 설정자 역할에 대한 성능 향상을 기대할 수 있다. 말뭉치 구축 과정에서는 다음과 같은 요소를 고려해야 한다.

첫째, 학습자의 수준별 비계 설정 발화 자료 구축이 필요하다. 비계 설

정은 학습자의 수준과 특성에 따라 개인별 맞춤형으로 진행되어야 한다. 따라서 말뭉치 자료도 학습자의 수준에 따라 체계적으로 구축되어야 한다. 동일한 문제에 대해서도 다양한 학습자의 문법 성취 수준에 따라 교사가 다르게 진행한 비계 설정 발화들을 수집함으로써, 이를 학습한 AI가 학습자의 응답 수준에 따라 적절한 비계를 설정할 수 있도록 학습시키는 것이 중요하다.

둘째, 교사의 비계 설정 발화 유형 구분과 효과성 태그(tag) 부여가 필요하다. 비계와 관련한 선행 연구에서 제시된 비계 전략은 매우 다양하며, 학습자의 응답 상황에 따라 적절한 전략이 선택되어야 한다. 따라서 교사의 비계 설정 발화를 비계 전략 유형별로 구분하고, 각 전략의 효과성에 대한 태그를 포함해야 한다. 이러한 태그 정보는 AI가 보다 효과적인 비계 전략을 선택하고 응답의 질을 높이는 데 기여할 수 있다.

이러한 말뭉치 구축 과정에서도 문법 교육 전문가와 데이터 전문가 간의 적극적인 협력이 필수적이다. 이러한 협력은 데이터의 질을 보장하고, AI 모델이 실질적으로 유용한 문법 비계 설정 도구로 발전하는 데 기여할 것이다.

현재 국립국어원에서는 모두의 말뭉치 사업을 통해 “일상 대화 음성 말뭉치”, “글쓰기 원시 자료 말뭉치” 등 다양한 우리말 말뭉치를 축적하고 있다. 특히 “추론 확신성 분석 말뭉치”나 “속성 기반 감성 분석 말뭉치” 등 문법 단위 간의 의미 관계를 기반으로 태그 작업한 말뭉치들은 AI 학습뿐만 아니라 평가용으로도 정비되어 활용되고 있다(김한샘·함영균·박승희·송상현·임경태·유현조 외, 2022). 이러한 맥락에서 “교사의 문법 비계 설정 발화 말뭉치” 등을 체계적으로 구축할 경우, 문법 교육용 AI 개발에 크게 기여할 수 있을 것이다.

V. 맷음말

본 연구는 ChatGPT를 문법 학습에서 비계 설정자로 활용할 가능성 을 탐구하기 위해 세 가지 모델(기본 모델, TutorMe, Grammar!)을 대상으로 시연을 진행하고, 각 모델의 응답을 분석하였다. 연구 결과, ChatGPT는 기본 모델에서 시작하여, 추가적인 지침과 데이터 투닝을 통해 비계 설정 전략과 학교 문법 지식 측면에서 점진적으로 개선된 응답을 제공하였다. 특히, Grammar! 모델은 검색 증강 생성 방식을 활용하여 학교 문법 지식의 환각 현상을 감소시키는 성과를 보였다.

그러나 세 모델 모두 학습자가 문제를 스스로 해결할 수 있도록 돋는 비계 설정자로서 완전한 역할을 수행하기에는 부족함이 있었다. 이는 특히, 특정 상황에서 잘못된 문법 지식을 제공하거나 정답을 직접 제시하는 문제에서 드러났다. 이러한 한계는 ChatGPT가 문법 교육에서 효과적으로 활용되기 위해 해결해야 할 과제를 제시한다. 첫째, 학교 문법의 특수성과 복잡성을 더 정교하게 반영한 검색 증강 데이터 구축이 요구된다. 둘째, 학습자의 수준과 요구에 맞춘 비계 설정을 위하여 AI 학습을 위한 교사의 비계 설정 발화 말뭉치 구축이 필요하다.

본 연구는 ChatGPT를 문법 학습에서 활용하기 위한 기초 연구로서, 문법 교육에서 AI 활용 가능성과 한계를 제시함으로써 향후 연구와 실무적 적용에 기여하고자 한다. 다만, 본 연구에서는 정확한 학교 문법 지식을 적용해야 하는 문법 탐구 문제에 한하여 연구가 진행되었는데, 향후 발산적 사고, 창의적 사고를 요구하는 문법 탐구에서도 ChatGPT가 활용될 수 있을지에 대한 후속 연구가 필요하다.

* 본 논문은 2025.01.24. 투고되었으며, 2025.02.11. 심사가 시작되어 2025.03.10. 심사가 종료되었음.

참고문헌

- 강우규(2023), 「생성형 AI를 활용한 고전소설 교육방안 모색」, 『돈암어문학』 44, 35-69.
- 권미선(2024), 「GPTs 기반 문제해결 맞춤형 챗봇 제작 및 수학적 성능 분석」, 『초등수학교육』 27(3), 303-320.
- 김한샘·함영균·박승희·송상현·임경태·유현조·정유남·송영숙·나승훈·여진영(2022), 『2022년 인공지능의 언어 능력 평가 체계 운영 및 말뭉치 정비』, 서울: 국립국어원.
- 노대원(2023), 「소설 가르치는 로봇-생성 AI를 활용한 문학교육 교수·학습 전략과 사례」, 『비평문학』 90, 35-65.
- 손태권(2023), 「ChatGPT의 수학교육 활용 가능성 탐색 분수 문제에 관한 학생의 산출물과 예비교사의 담화 사례를 중심으로」, 『初等數學教育』 26(2), 99-113.
- 안도연·손태권·이광호(2023), 「비계설정 도구로서의 ChatGPT 초등학생의 수학 논리 문제 해결 능력에 미치는 영향 평가」, 『Brain, Digital, & Learning』 3(2), 183-196.
- 양수연·고유라·신호정·주세형(2023), 「생성형 인공지능은 교사의 교육적 질문 생성 역할을 대신할 수 있는가—ChatGPT가 생성한 읽기 질문의 활용 가능성에 대한 교사 인식 연구」, 『새국어교육』 136, 117-165.
- 이관규·박경희·신호철·신희성·이동석·정지현·하성욱(2019), 『언어와 매체』, 서울: 비상교육.
- 이관희·조진수(2015), 「문법 교사의 오개념 유형화 연구」, 『새국어교육』 102, 107-152.
- 이소희·김갑수(2023), 「ChatGPT 기반 초등학교 6학년 국어 교과 챗봇 개발—어휘 학습 지원을 중심으로」, 『정보교육학회논문지』 28(3), 341-350.
- 이승윤(2023. 4. 1.), 「생성 AI가 노동시장에 미치는 영향: 오픈AI - 펜실베니아 대학 공동 연구 소개」, 『SPRI AI Brief 특집호』, 1-10.
- 이용상·시기자·박도영·윤경아·구슬기·임황규(2013), 「쓰기 자동채점 알고리듬의 성능 비교: 최대 엔트로피 기법과 서포트 벡터 회귀 기법」, 『교육과정평가연구』 16(3), 147-165.
- 이진·정진경·김한샘(2024), 「한국어 학습자 쓰기 자동채점을 위한 어휘 다양도 지표의 자질 활용 가능성 탐색」, 『문법 교육』 50, 193-229.
- 이학무(2023. 1. 1.), 「AI 성공 키워드는 산업별 최적화한 경험과 지식, 조선일보, 검색일자 2024. 12. 3., 사이트 주소 <https://it.chosun.com/news/articleView.html?idx-no=2022123001361>.
- 최숙기·박종임(2024), 「생성형 AI를 활용한 현직 국어교사의 서·논술형 평가 문항 개발 양상 분석」, 『청람어문교육』 97, 243-270.
- Hogan, K. & Pressley, M. (1997). Scaffolding Scientific Competencies within Classroom Communities of Inquiry. In K. Hogan, & M. Pressley (Eds.), *Scaffolding Student Learning: Instructional Approaches and Issues*, Cambridge, Massachusetts: Brookline Books.
- OpenAI(2024. 5. 14.), 「Math problems with GPT-4o」, YouTube, 검색일자: 2024. 12. 3.,

사이트 주소: https://www.youtube.com/watch?v=_nSmkyDNulk&list=PLOX-w6f10VTv8VOvPNVQ8c4D4NyMRMotXh&index=19&t=22s.

문법 학습 비계 설정자로서 ChatGPT 활용을 위한 과제와 방향

박종미

본 연구는 ChatGPT를 문법 학습의 비계 설정자로 활용할 가능성을 탐구하기 위해 세 가지 모델(기본 모델, TutorMe, Grammar!)을 시연하고 응답을 분석하였다.

분석 결과, 기본 모델은 부적절한 비계 설정 전략을 사용하였으며, 학교 문법 지식에서 환각 현상이 발생하였다. TutorMe 모델은 비계 설정 전략에서 개선된 응답을 제공했으나, 학교 문법 지식의 부정확성이 여전히 존재하였다. Grammar! 모델은 검색 증강 방식을 통해 문법 지식의 정확도를 높였으나, 정답을 제공함으로써 비계 설정자 역할에는 한계가 있었다.

ChatGPT를 문법 학습에 효과적으로 활용하기 위해서는 비계 설정을 지원하는 AI 학습용 말뭉치 구축과 학교 문법 문제 풀이 데이터를 기반으로 한 검색 증강 생성 기술의 개선이 필요하다.

핵심어 ChatGPT, 문법 교육, 비계 설정, 검색 증강 생성, 학교 문법, AI 활용

ABSTRACT

Utilizing ChatGPT as a Scaffolder in Grammar Learning: Challenges and Future Directions

Bak Jongmi

This study investigated the feasibility of using ChatGPT as a scaffold for grammar learning by conducting demonstrations and analyzing responses from three models: the basic model, TutorMe, and Grammar!

The analysis revealed that the basic model employed inappropriate scaffolding strategies and displayed hallucinations regarding school grammar knowledge. The TutorMe model provided improved scaffolding responses but still exhibited inaccuracies in grammatical knowledge. The Grammar! model enhanced the accuracy of grammar knowledge through Retrieval-Augmented Generation; however, it showed limitations in its role as a scaffold by providing direct answers rather than guiding the learning process.

To effectively utilize ChatGPT in grammar learning, developing AI training corpora that support scaffolding strategies and improve search-augmented generation capabilities based on school grammar problem-solving data.

KEYWORDS ChatGPT, grammar education, scaffolding, RAG, school grammar, AI application