



## ChatGPT의 피드백이 장르별 글쓰기 역량 인식에 미치는 영향

최윤희 (충신대학교)

Received: 10 April 2025  
Revised: 18 April 2025  
Accepted: 23 April 2025

**Choe, Yoonhee. (2025). Effects of ChatGPT-Generated Feedback on Genre-Specific Writing Competence Perception. *Modern English Education*, 26, 197-212.**

### Keywords

ChatGPT, genre, cause and effect writing, comparison and contrast writing, argumentative writing, writing competence  
챗지피티, 장르, 인과문, 비교문, 논증문, 쓰기 역량

### Yoonhee Choe

Professor  
Department of English Language Education  
Chongshin University  
yhchoi@chongshin.ac.kr  
ISNI: 0000-0004-6463-2998

\* This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea(NRF-2020S1A5A8047751).

### Abstract

This study examined how pre-service English teachers would utilize ChatGPT in three different English writing genres and perceive their writing competences. Participants were pre-service English teachers in the Department of English Education in Seoul, Korea. They initially wrote essays in cause and effect, comparison, and argumentative genres. They then received ChatGPT-generated feedback. To assess their perception in writing competences, macroscopic aspects (Thesis Statement, Idea Generation, Idea Development, and Idea Organization) and microscopic aspects (Topic Sentence, Coherence, Lexical Diversity, and Syntactic Complexity) were measured using a 5-point Likert scale. Responses collected before and after ChatGPT feedback were analyzed using Multivariate Analysis of Variance (MANOVA). Findings indicated that ChatGPT use led to perceived improvements in both macroscopic and microscopic aspects regardless of genre, while no significant differences were found between genres. This study highlights the potential of using ChatGPT to enhance English writing skills. It provides insights for AI-integrated writing education.

## 서론

최근 인공지능(artificial intelligence) 기술의 발전은 영어 교육, 특히 영어 글쓰기 교육(writing instruction)과 관련된 연구에 많은 영향을 미치고 있다(Choe, 2023a; 2023b; Hwang et al., 2024; Kim & Jeong, 2025; Kohnke & Zou, 2024; Mahapatra, 2024). 예를 들어, 자동화된 글쓰기 평가(Automated Writing Evaluation) 및 AI 기반 피드백 시스템은 기존의 인간 교사의 피드백을 보완하는 역할을 하며, 즉각적이고 반복적인 피드백을 제공할 수 있는 장점을 가지고 있다(Cao & Zhong, 2023). 특히 2022년 말 등장한 생성형 AI(generative AI) 모델인 ChatGPT는 기존의 AI 피드백 시스템과 달리 단순한 문법 교정을 넘어, 학습자가 원하는 대로 텍스트와 이미지를 생성해 내고, 피드백을 제공하고 있다. 또한, 어휘적 다양성(lexical diversity), 통사적 복잡성(syntactic complexity)과 문장간 응집성(cohesion) 등과

같은 고차원적인 글쓰기 요소까지 평가하고 조언할 수 있는 능력을 갖추고 있다고 한다(Athanassopoulos et al., 2023). 이에 따라 생성형 AI를 활용한 영어 글쓰기 교육이 점차 확산되고 있으며, 연구자들은 AI 피드백이 학습자의 글쓰기 수행에 미치는 영향을 분석하기 위한 다양한 연구를 수행하고 있다(e.g., Choe, 2023a; 2023b; Choe, 2024). 예를 들어, Mahapatra(2024)는 ChatGPT 기반 피드백이 영어 학습자의 문법 정확성(grammatical accuracy), 응집성(cohesion), 통사 복잡성(syntactic complexity)에 미치는 긍정적인 영향을 보고하였으며, Athanassopoulos et al.(2023)은 AI 피드백이 학습자의 어휘적 다양성(lexical diversity)과 아이디어 전개(idea development)를 향상시킬 수 있음을 밝혔다.

이러한 피드백은 단순히 글의 완성도를 높이는 데 그치지 않고, 학습자가 자신의 글쓰기 과정을 비판적으로 성찰하도록 유도함으로써 메타인지적 인식과 자기성찰을 촉진하는 역할도 수행한다. 최근에는 AI 도구를 단순히 수용하는 것을 넘어, 학습자가 AI의 피드백을 비판적으로 해석하고 활용할 수 있는 역량, 즉 비판적 AI 리터러시(critical AI literacy, CAIL)의 중요성이 강조되고 있다(Wang & Wang, 2025). CAIL은 학습자가 AI의 조언을 무비판적으로 수용하기보다는, 그 적절성과 타당성을 스스로 평가하고 선택적으로 수용하는 과정을 포함한다. 이러한 역량은 글쓰기에서 메타인지적 인식(meta-cognitive awareness) 및 자기성찰(self-reflection)과 밀접하게 연관되어 있으며, AI 피드백의 효과를 극대화하기 위한 핵심 요인으로 작용한다. 따라서 AI 기반 피드백의 교육적 효과를 온전히 이해하기 위해서는, 기술적 정확성뿐만 아니라 학습자의 인지적 수용 과정과 성찰적 태도를 함께 고려해야 한다(Mahapatra, 2024; Michel et al., 2025).

그러나 기존 연구들은 AI 피드백이 제공하는 정보가 학습자의 메타인지적 사고를 얼마나 유도하는지에 대한 분석이 부족하다(Choe, 2024). 즉, AI 피드백이 학습자의 글쓰기 과정에 대한 인식을 증가시키는 효과를 보고하였지만(Choe, 2023a; 2023b), AI 피드백이 학생들의 비판 및 비평적 사고(critical thinking)와 자기조절(self-regulation) 학습에 어떻게 기여하는지는 명확하게 검토하지 않았다. 또한, Wang and Wang(2025)의 연구에서는 AI 피드백이 문장 수준의 개선에는 효과적이지만, 학습자가 AI의 조언을 비판적으로 평가하고 자신의 의견을 반영하는 과정이 부족할 수 있음을 지적하였다. 이러한 연구 결과는 AI 피드백을 효과적으로 활용하기 위해서는 학습자의 자기성찰과 메타인지적 인식을 함께 고려해야 한다는 점을 시사한다. 이와 더불어, 현재까지의 연구는 AI 기반 피드백이 학습자들의 글쓰기 수행에 미치는 영향을 검토해왔지만, 대부분 단일 장르(single-genre) 중심의 연구에 국한되어 있으며(Heng et al., 2023), AI 피드백이 장르별로 동일한 효과를 가지는지 여부에 대한 연구는 미비한 상황이다(Choe et al., 2020; Choe et al., 2022). 기존 연구들은 주로 교사 또는 채점 전문가가 학습자의 글쓰기를 평가하는 방식으로 진행되었으며, AI 피드백을 받은 학습자가 스스로 자신의 글쓰기 역량이 향상되었는지 평가하는 연구는 부족하다. 특히, 예비 영어 교사들이 AI 피드백을 활용하여 자신의 글쓰기 역량 향상을 스스로 인식하는 과정은 기존 연구에서 이루어지진 하였으나, 다양한 장르에 대한 연구는 매우 부족한 상황이다(Choe, 2023a; 2024).

따라서, 본 연구에서는 ChatGPT 기반 AI 피드백이 예비 영어 교사들의 글쓰기 역량의 인식에 어느 정도 영향을 미치는지를 장르별로 분석하고자 한다. 구체적으로, 인과문, 비교문, 논증문의 세 가지 주요 장르에서 ChatGPT 피드백 전후의 글쓰기 역량 향상 정도를 거시적(macroscopic) 및 미시적(microscopic) 측면에서 비교하고자 한다. 이에, 본 연구의 연구 문제는 다음과 같다.

첫째, 영어글쓰기 장르에 따라 ChatGPT 피드백 전후, 예비 영어 교사의 영어 글쓰기 역량에 대한 인식은 거시적 측면(논제문 작성, 아이디어 생성, 아이디어 발전 및 아이디어 구성)에서 유의미한 변화가 있는가?

둘째, 영어글쓰기 장르에 따라 ChatGPT 피드백 전후, 예비 영어 교사의 영어 글쓰기 역량에 대한 인식은 미시적 측면(주제문 작성, 일관성, 어휘적 다양성 및 통사 복잡성)에서 유의미한 변화가 있는가?

## 이론적 배경 및 선행연구

생성형 AI(genAI)의 발전은 영어 글쓰기 교육의 방식에 근본적인 변화를 가져오고 있다. ChatGPT와 같은 AI 도구는 자동화된 피드백, 문법 수정, 어휘와 표현의 다양성 및 논리적 구조 향상을 지원하며, 학습자의 글쓰기 역량 함양 및 학습자의 자율성(learner autonomy)을 촉진하는 것으로 보고되고 있다. 그러나 AI 기반 피드백이 글쓰기의

거시적(macroscopic) 및 미시적(microscopic) 측면에 미치는 영향, 그리고 장르별로 ChatGPT의 효과가 달라질 가능성에 대한 연구는 아직 발전 단계에 있다. ChatGPT 등 생성형 AI 도구를 활용한 글쓰기 피드백의 효과에 관한 최근 연구들을 살펴보면 다음과 같다.

## ChatGPT의 피드백이 글쓰기에 미치는 영향

최근 ChatGPT를 비롯한 생성형 AI를 활용한 글쓰기 피드백 연구들이 증가하고 있으며, 초기 연구들은 ChatGPT 피드백의 부정적 측면 보다는 긍정적 효과를 보고하는 연구들이 많이 수행되어 왔다(e.g., Mahapatra, 2024). 구체적으로 살펴보면, 다음과 같다. Michel et al. (2025) 연구는 독일어 L2 학습자의 협업 글쓰기(collaborative writing) 과정에서 AI 피드백이 어떻게 활용되는지를 분석하였다. 연구자들은 ChatGPT 피드백이 기능적 적절성(functional adequacy)을 향상시키며, 문장의 조직력을 강화하는 데 기여한다고 밝혔다. 그러나, 학생들이 AI 피드백을 그대로 받아들이는 경향이 있어 비판적 사고(critical thinking)가 저하될 가능성이 있으며, AI 피드백을 받은 텍스트가 논리적 깊이가 부족하다고 지적하였다. 또한, Wang and Wang (2025)는 APSE 모델(awareness, positionality, strategies, evaluation)을 개발하여 AI 피드백이 학습자의 비판적 AI 리터러시(critical AI literacy) 향상에 미치는 영향을 분석하였다. 연구 결과, AI 피드백이 학습자들에게 실시간 교정 및 수정 기회를 제공하지만, AI 피드백을 비판적으로 평가하는 역량이 중요하다고 결론지었다. AI를 학습 도구로 활용할 경우, 단순한 피드백 수용이 아니라, AI의 응답을 평가하고 조정하는 메타인지적 접근(metacognitive approach)이 필요하다고 주장하였다. Mahapatra (2024) 연구는 ESL 학습자를 대상으로 ChatGPT를 형성적 피드백(formative feedback) 도구로 활용하여 ESL 대학생들의 학술적 글쓰기 능력 향상에 미치는 영향을 중재 연구(intervention study)를 수행하여 살펴본 결과, ChatGPT 피드백을 받은 집단에서 글쓰기 기술이 유의하게 향상되었음을 확인하였다. 특히, ChatGPT 피드백이 학습자의 아이디어 생성(idea generation)과 글쓰기 조직력(organization skills)을 강화하는 데 효과적임을 실험적으로 입증하였다. 즉, AI 피드백이 제공된 후 학생들의 글에서 논리적 전개, 문장 응집성, 어휘 선택이 유의미하게 향상되었다. 또한, Yang and Lin (2025) 연구는 EFL 학습자들이 AI 번역 및 페러프레이징 기능을 활용하여 언어적 격차를 보완하는 방식을 분석하였다. 연구 결과, 언어 능력이 낮은 학습자는 AI에 과도하게 의존하는 경향이 있으며, 이는 글쓰기 능력의 독립성을 약화시킬 가능성이 있다고 결론지었다. 이상의 연구들은 ChatGPT와 같은 생성형 인공지능 도구를 활용한 피드백은 단순한 텍스트 생성 기능을 넘어, 고품질의 글을 생성할 수 있는 능력(Gao et al., 2022)뿐만 아니라, 텍스트의 조직, 언어 사용, 오류 수정에 대한 피드백 제공(Ohio University, 2023), 논리적 구성과 적절한 세부 정보 및 결론 제시(Fitria, 2023) 등의 측면에서 외국어 학습자의 글쓰기에 대체로 긍정적인 영향을 끼치지만 과잉 의존 또는 피상적인 수준의 피드백 등의 부정적인 측면도 있으니, 이를 교수방법에 고려할 필요가 있음을 시사한다(Yao et al., 2025).

## ChatGPT의 피드백이 장르에 따른 글쓰기 역량에 미치는 영향

위에서는 ChatGPT의 피드백이 영어 글쓰기에 전반적으로 어떠한 영향을 미치는지를 살펴본 연구들을 검토하였다. 좀 더 구체적으로 생성형 AI가 구체적으로 글의 장르에 따라 어떠한 영향을 미치는지를 분석한 연구들을 살펴보면 다음과 같다. 먼저, 국내 연구로, Choe(2024)는 인과문(cause and effect writing), 비교문(comparison and contrast writing), 논설문(argumentative writing)의 세 가지 장르에서 ChatGPT 피드백의 효과를 검증한 연구를 수행하였다. 참가자들은 각 장르별로 에세이를 1차 작성한 후 ChatGPT 피드백을 받아 수정본을 작성하였으며, 연구자는 글쓰기의 거시적 측면(논제문 작성, 아이디어 생성, 전개, 구성) 및 미시적 측면(문단 주제문 작성, 문장 일관성, 어휘 다양성, 통사 복잡성)을 평가하였다. 다변량 분산분석(MANOVA) 결과, ChatGPT 피드백 후에 전반적인 글의 질이 향상된 것으로 나타났다. 구체적으로, 모든 장르에 공통적으로 글의 일관성(coherence), 어휘적 다양성(lexical diversity), 통사 복잡성(syntactic complexity) 지표에서 유의미한 향상이 확인되었으며, 일부 거시적 요소 중 아이디어 전개력도 피드백 전보다 향상되었다고 보고하였다. 흥미로운 점은, 장르 자체의 효과나 장르와 피드백의 상호작용 효과는 대부분 유의하지 않게 나타났다는 것이다. 이는 ChatGPT의 피드백 효과가 특정 장르에 국한되지 않고 전 장르에 걸쳐 고르게 나타났음을 의미한다. 다시 말해, 비교문, 인과문 및 논증문이나 ChatGPT를 활용하면 학습자 글쓰기의 여러 측면을 개선시키는 데 전반적으로 도움이 되었던 것이다. 이러한 결과는 생성형 AI가 학습자의 장르 인식과 글쓰기 능력을 향상시키는 데 잠재력이 있음을 보여준다. 동시에 Choe(2024)는 ChatGPT

피드백이 주로 글의 피상적인 측면(문장 수준의 표현 등)에서 개선을 보여줄 수 있으므로, 장르별 특성과 학습자 수준을 고려한 보완적 지도가 필요하다고 지적하였다. 이와 같은 연구들은 ChatGPT가 학습자의 글쓰기 능력 향상에 기여할 가능성이 높으며, 장르별 특성을 고려한 보완적 지도 전략과 함께 활용될 경우 더욱 효과적일 수 있음을 시사한다. Choe(2024)의 연구에서처럼, ChatGPT는 학습자의 글쓰기 능력에 영향을 미치는데, 이는 글의 거시적(macroscopic) 측면(논제문, 아이디어 생성, 아이디어 전개 및 구성)과 미시적(microscopic) 측면(문단의 주제문, 일관성, 어휘적 다양성, 통사 복잡성)으로 구분하여 분석하였다는 점에서 의의를 볼 수 있으나, 연구 참가자들의 인터뷰나 설문 분석을 통해 양적인 향상의 과정과 인식을 동시에 살펴보는 못했다는 한계가 있다.

글의 거시적 및 미시적 측면에서의 ChatGPT 피드백 효과를 살펴본 해외 연구로, Markey et al. (2024)는 ChatGPT가 생성하는 대규모 텍스트를 분석하며, AI가 생성하는 글이 정보 밀도가 높지만, 창의적 사고나 논증의 깊이가 부족하다는 점을 강조하였다. Markey et al. (2024)는 AI 피드백이 지나치게 형식적이거나 일반적인 피드백에 그칠 경우, 학생들이 비판적으로 사고하는 능력을 약화시킬 위험이 있다고 경고하였다. 또한, Redmann (2024)는 독일어 L2 수업에서 장르 기반 글쓰기 교육(genre-based writing pedagogy)을 연구하며, AI 피드백이 장르 학습을 지원할 수 있지만, 일반적인 내용 생성에 그칠 가능성이 높다고 지적하였다. 즉, AI 피드백이 장르 학습을 지원하여 논증적 글쓰기에서 유용할 수 있으나, 일반적인 패턴을 반복할 가능성이 있다라고 하였다. 그의 연구 결과에서 학생들은 AI 피드백을 받으면 문장 수준의 개선이 이루어졌지만, 장르 특유의 논리적 흐름을 형성하는 데는 한계가 있었다고 보고하였다. 또 다른, ChatGPT의 장르별 피드백 차이에 대한 연구로, De Oliveira & Santos (2025)는 AI 생성 멘토 텍스트(mentor texts)가 장르 학습을 지원하는 데 중요한 역할을 할 수 있음을 입증하였다. 연구자는 AI 피드백을 활용하면 학습자가 장르의 주요 요소를 파악할 수 있지만, AI 피드백에 과도하게 의존하지 않도록 교수 전략이 필요함을 강조하였다.

이상의 연구들을 종합해 보면, ChatGPT 등의 생성형 AI 도구를 사용한 피드백의 영어 글쓰기 교육 시 교수학습 방법에 여러 가지 시사점을 제안한다. Mahapatra (2024)가 언급하였듯이, AI 도구의 적절한 활용이 학습자의 글쓰기 자신감 향상에 기여할 수 있으나, AI 글쓰기 도구 남용 시 비판적 사고 저하 가능성(Markey et al., 2024)도 고려해야 한다. 따라서, 비평적 AI 리터러시 교육을 병행할 필요가 있다(Michel et al., 2025, Wang & Wang, 2025). 또한, 장르에 따른 ChatGPT를 활용한 피드백은 일반적으로 문단의 일관성, 어휘 다양성, 통사 복잡성을 향상시키는 데 기여하는 것으로 나타났다. 또한, 외국어 학습자들이 인지적으로 어려움을 느끼는 논증문(argumentation)에서 효과적이며, 비교 및 인과적 글쓰기에서는 제한적 효과를 보였다고 하나, 통계적으로 유의미한 차이를 보이지는 않는 것으로 나타났다(Choe, 2024). 이러한 맥락에서 기능적 적절성(functional adequacy) 개념은 AI 피드백의 교육적 효과를 평가하는 핵심 요소로 부각되고 있다. Michel et al. (2025)은 기능적 적절성에 기반하여, 각 장르의 담화적 요구에 적합한 장르별 차별화된 피드백 전략이 필요하다고 제안하였다. 그러나 기존의 선행연구들은 대부분의 연구가 소규모 실험으로 진행되어 결과를 일반화하기에 어려움이 있었으며, 글쓰기에서 중요한 메타인지와 자기성찰적인 측면인 학습자의 글쓰기 과정과 인식에 대한 자기 평가에 대한 분석은 제한적이었다. 또한, 학술적인 장르에 따른 ChatGPT의 피드백이 글의 질적인 요소에 어떠한 영향을 끼친다고 인식하는지를 조사한 연구는 매우 부족하였다.

따라서, 본 연구에서는 이러한 한계를 보완하고자 예비 영어 교사를 대상으로 장르별 ChatGPT 기반 AI 피드백의 영향을 분석하고자 한다. 장르는 구체적으로 예비교사들이 필수적으로 수강하는 ‘영어교과논리 및 논술’ 교과목에서 다루는 주요 세 가지 장르인 인과문, 비교문, 논증문에서 AI 피드백 전후의 글쓰기 수행을 거시적 측면(논제문, 아이디어 생성, 아이디어 개발 및 아이디어 구성)과 미시적 측면(주제문, 일관성, 어휘적 다양성 및 통사 복잡성)에서 비교 분석하고자 한다. 또한, AI 피드백이 단순한 문법 교정 기능을 넘어, 학습자의 자기성찰과 메타인지적 사고를 촉진하는지에 대한 실증적 연구가 필요하므로, 생성형 AI의 피드백의 예비영어교사들의 쓰기 역량 인식에 어떠한 영향을 미치는지, 교육적 활용 가능성을 탐색하고자 한다.

## 연구 방법

### 연구 대상

본 연구는 서울에 위치한 한 대학의 영어 교육과의 교과목 중 ‘영어 교과 논리 및 논술’ 과목을 2023년 1학기

2024년 1학기에 수강한 예비 영어교사 47명을 대상으로 수행되었다. 연구 참여자는 대부분 3학년(43명)으로 구성되었으며, 2학년과 4학년이 각각 2명씩 포함되었다. 연구 참가자들의 기본 배경은 표1과 같다.

**TABLE 1**  
*Participants' Demographic Information (N = 47)*

Gender		Majors		Proficiency levels <sup>1</sup>			Total
Male	23	English Education	42	B1	B2	C1	47
Female	24	Christian Education	5	7	36	4	

## 연구 절차

연구 참여자들은 학기 초반 주제문(topic sentence) 작성과 문단 내 일관성(coherence)에 대한 학습을 진행하였으며, 중간고사 이후에는 인과문(causative writing), 비교문(comparative writing), 논증문(argumentative writing)이라는 세 가지 주요 장르의 특징을 이해하고 장르별 글쓰기를 수행하였고, ChatGPT의 피드백을 받고 수정하는 활동을 하였다. 즉, 연구 참여자들은 글쓰기 과제 수행 전, 표2에 제시한대로, 1~8주차 동안 ‘Topic Sentence, Coherence, Lexical Diversity, Syntactic Complexity’에 대한 명확한 개념 학습 및 사례 분석 활동을 수행하였다. 이를 통해 주요 평가 기준에 대한 이해도를 높이는 사전교육이 이루어졌다.

연구 참가자들이 한 학기간 학습하며 활동한 것에 대한 내용은 표 2에 제시하였다. 본 연구는 마지막 주인 15주차에 실시되었고, 연구 참가자들은 각 장르별로 제시된 주제에 따라 독립적으로 작문한 후, ChatGPT의 피드백을 제공받고 이를 반영하여 수정하는 과정을 거쳤다. 참여자들은 중간고사 이전에는 단락 쓰기 중심의 학습만 수행하였으며, ChatGPT 피드백은 중간고사 이후 장르별 글쓰기 수행부터 적용하였다. 또한, 참가자들은 다음과 같은 공통 프롬프트를 사용하여 ChatGPT에 피드백을 요청하였다: “Please give feedback on my English writing in terms of macroscopic aspects(thesis statement, idea generation, idea development, idea organization) and microscopic aspects(topic sentence, coherence, lexical diversity, and syntactic complexity). Suggest detailed feedback.” 프롬프트는 강의자가 사전에 안내한 고정된 형태로 통일되었으며, 개별적 변경 없이 동일한 문장을 복사하여 사용하였다.

**TABLE 2**  
*Data Collection Process: Weekly Writing Activities*

Week	Topics	Lectures & Activities
Week 1~8	Paragraph writing	Topic sentence, coherence, paraphrasing, summarizing
Week 9~14	[Genre-based writing]	
Week 9~10	Cause & Effect writing	Genre Analysis, 1 <sup>st</sup> draft & Revision
Week 11~12	Comparison & Contrast writing	Genre Analysis, 1 <sup>st</sup> draft & Revision
Week 13~14	Argumentative writing	Genre Analysis, 1 <sup>st</sup> draft & Revision
Week 15	Data Collection	[Given Topics] - The causes and effect of smartphone addiction among teenagers - Compare and contrast the advantages or disadvantages of Face-to-Face(F2F)and Online(i.e., zoom) Classes - Should students be allowed to use smartphones in schools?

본 연구에서 글의 주제에 대한 친숙도(topic familiarity)나 특정 장르의 인지적 난이도가 글쓰기 수행에 영향을 미칠 가능성을 최소화하기 위해, 각 장르별 글쓰기 과제는 무작위로 배치되었으며, 장르 제시 순서 또한 참여자 간에 상호 보완적으로(counterbalanced) 조정되었다. 연구 참여자들은 ChatGPT 피드백을 반영하여 개별적으로 수정된

<sup>1</sup> B1 = Intermediate Low, B2 = Intermediate High, C1 = Advanced (CEFR 기준)

초안을 작성하였으며, 모든 초안과 수정본을 LMS(learning management system)에 업로드하였다. 연구자는 이를 검토하여 연구 절차가 적절히 진행되었음을 확인 및 검토 하였다.

## 자료 수집 및 평가 방법

본 연구가 기존 연구들과의 차별점은 글쓰기 평가를 외부 평가자가 수행하는 것이 아니라, 연구 참여자가 자신의 글쓰기 역량을 직접 평가(self-assessment)하는 방식을 적용했다는 점이다. 연구 참여자들은 ChatGPT 사용 전후 자신의 글쓰기 역량이 어떻게 변화했는지를 주관적으로 평가하고, 그 차이를 분석하는 자기평가지(can-do self-assessment) 방식을 활용하였다. 참여자들은 자신의 글쓰기 역량을 평가하기 위해 다음과 같은 거시적(macroscopic) 및 미시적(microscopic) 평가 항목을 기준으로 응답하였다. 글쓰기 능력에 대한 자기 인식 분석을 위해 종속변인을 거시적 및 미시적 수준으로 구분한 것은 글 전체의 구조적 완성도와 문장 수준의 언어적 정확성이라는 두 차원의 글쓰기 평가 기준에 기반하며, Weigle(2002)와 Barkaoui(2010)의 연구에 이론적 근거를 둔다. 거시적 요소는 글의 전개, 논제의 명료성, 아이디어의 발전 및 조직력 등 전체 구성과 관련된 능력을 포함하며, 미시적 요소는 주제문 구성, 문장 간 연결성, 어휘 사용의 다양성, 문법적 복잡성 등 세부 표현 수준의 요소를 포함한다. 이러한 구조적 구분은 글쓰기 교육과 평가에서 일반적으로 받아들여지고 있다 (Barkaoui, 2010; Weigle, 2002). 예를 들어, TOEFL iBT 및 TOEIC Writing과 같은 표준화된 영어 쓰기 평가에서는 아이디어 전개, 조직력, 논리적 구성과 같은 담화 수준의 요소들을 거시적 요소로 간주하며, 이에 더해 문법적 정확성, 어휘 사용, 문장 간 일관성과 같은 요소들은 미시적 요소로 평가한다(Educational Testing Service, 2025). 이러한 이중 평가 체계는 글의 구조적 논리성과 언어적 완성도를 종합적으로 고려하여 글의 전반적인 질을 판단한다는 점에서, 본 연구의 분석틀과 이론적 정합성을 가진다. 실제로 Weigle(2002)는 평가자들이 글의 ‘전반적(global)’ 요소와 ‘국지적(local)’ 요소를 모두 고려하여 평가하는 경향이 있음을 언급하였으며, Barkaoui(2010) 또한 이러한 요소 구분이 평가자 간의 평가 일관성을 높이는 데 기여한다고 보고하였다. 또한, 최근 연구에 따르면, ChatGPT는 문법, 어휘, 연결어 사용 등 표면적 요소에서 뛰어난 피드백을 제공하나, 아이디어 구성, 논리적 구조화, 글의 응집성과 같은 심층적 요소에서는 인간의 피드백에 미치지 못하는 경향이 있다 (Algaraady & Mahyoob, 2023; Zhou et al., 2023). 따라서 본 연구의 변수 구조 구분은 표3과 같이, ChatGPT 피드백의 효과를 보다 정밀하게 분석하기 위한 기능적 분석 틀로서도 유용하며, 특히 표면적(surface-level) 요소와 심층적(in-depth) 요소 간의 차이를 설명하는 데 이론적 및 실용적 정당성을 제공한다.

**TABLE 3**  
*Evaluation Criteria for Macroscopic and Microscopic Aspects of Writing*

Macroscopic aspects	Thesis statement
	Idea generation
	Idea development
	Organization
Microscopic aspects	Topic sentence
	Coherence
	Lexical diversity
	Syntactic complexity

이에 따라, 본 연구에서는 참여자들이 자신의 글쓰기 역량을 자가 평가할 수 있도록, 거시적 및 미시적 평가 기준에 기반한 1~5점 척도의 리커트(Likert) 척도를 사용하여 역량 향상 정도를 측정하였다. 평가 시점은 GPT 사용 전(pre-GPT) 평가가 중간고사 직후, 즉 ChatGPT 피드백을 처음 사용하기 전의 시점에서 이루어졌으며, post-GPT 평가는 모든 장르별 글쓰기가 종료된 후, 학기 마지막 주(15주차)에 이루어졌다.

다음 그림1은 연구 참가 학생 중 한 명이 자신이 작성한 논증문과 그에 대한 ChatGPT의 피드백을 제시한다.

3. Genre: Argumentative Essay: at least 4 paragraphs

Topic: Should students be allowed to use smartphones in schools?

Organization (Block or **Point-by-point**)

Although there are undoubtedly numerous advantages to using smartphones, I firmly believe that teachers should require students to turn off their phones while in school.

Some proponents argue that smartphones are beneficial in the classroom as they provide access to a wealth of academic materials and aid in comprehension. However, smartphone can rather distract the class distracting attention of students. Because teenagers are still weak in controlling their desires to do another thing, they can use smartphone to get in another site and this can distract their attention.

Furthermore, advocates claim that smartphones facilitate communication and collaboration among students. Despite increased opportunity to connect and share each other's idea, I think using smartphone can rather lead them to misuse it or an escalation in **cyberbullying** within the classroom. Allowing students unrestricted access to smartphones may result in some students using them for non-academic conversations, potentially causing disruptions and excluding those who are not part of these private exchanges. And using controlled computer installed in school instead of smartphone can help them communicate and fairly share ideas about the class.

In conclusion, considering these two primary drawbacks of using smartphones in the classroom, I think students have to participate in class without smartphone. By eliminating smartphone usage, teachers can create an environment conducive to focused learning and minimize the potential negative consequences associated with their misuse.

Language Use:

- Grammatical accuracy: Review your essay for grammatical errors such as subject-verb agreement, verb tense consistency, and proper use of articles. For instance, "Although so many advantages that can be gathered with using smartphone" should be "Although there are many advantages to using smartphones."
- Appropriate lexical choices: While your choice of vocabulary is generally suitable, consider incorporating more precise or varied terms. Instead of repeating certain phrases like "using smartphone," use synonyms or related terms like "utilizing mobile devices" or "employing cell phones" to enhance the quality of your writing.
- Syntactic complexity: To improve the syntactic structure of your essay, vary your sentence structures by including complex and compound sentences. Use subordinating conjunctions (e.g., "although," "because," "if") and relative pronouns (e.g., "that," "which") to create more sophisticated sentence structures and convey your ideas more effectively.
- Punctuation: Pay attention to punctuation errors, such as missing or misplaced commas, periods, and capitalization. Ensure that your punctuation is accurate and consistent throughout the essay.

By implementing these changes, you can address the criteria of the rubric and improve the overall quality of your essay. Strengthening the content by developing creative ideas, providing evidence, and ensuring logical coherence will make your argument more persuasive. Organizing your essay with clear introductions, topic sentences, and transitional devices will improve the overall structure and flow. Additionally, focusing on grammatical accuracy, appropriate lexical choices, syntactic complexity, and punctuation will enhance the clarity and readability of your essay, making it more compelling for the reader.

Regenerate response

Send a message. ➤

FIGURE 1

Example of ChatGPT-Generated Feedback Provided on a Student's Argumentative Writing

## 분석 방법

수집된 데이터는 반복측정 다변량 분산분석(repeated measures multivariate analysis of variance, MANOVA)을 통해 분석되었다. MANOVA의 타당성을 확보하기 위해 다음과 같은 사전 점검을 수행하였다. MANOVA는 종속변인 간 과도한 다중공선성을 허용하지 않으므로, Pearson 상관계수를 분석하였다. 거시적 측면(thesis statement, ideag, idea development, organization)에서 종속변인 간 상관계수는 .42~.58 수준으로, 과도한 상관(>.90)을 보이지 않으며, 서로 상이한 구성 개념임을 보여주었다. 또한, 미시적 측면(topic sentence, coherence, lexical diversity, syntactic complexity)에서 부분 상관계수는 .40~.81 사이이며, 특히 어휘 다양성(lexical diversity)과 통사적 복잡성(syntactic complexity) 간 상관( $r = .82$ )은 다소 높은 편이나, 이는 글의 문장의 어휘 다양성과 구조 복잡성 간 개념적 연관성을 반영함에 기인하므로, 공선성 문제가 발생할 수준은 아닌 것으로 보고 분석하였다.

MANOVA 분석을 통해 참여자들이 ChatGPT 사용 전후 자신의 글쓰기 능력에 대한 인식이 어떻게 변화했는지를 측정하고, 장르별 차이를 검토하였다. 이를 통해 연구자는 AI 도구가 한국 대학생 예비 영어교사들의 자기평가를 기반으로 한 글쓰기 역량 변화에 미치는 영향을 보다 체계적으로 분석하였다. 기존 연구들(e.g., Choe, 2024)이 AI 피드백이 외부 평가자에 의해 객관적으로 평가된 학습자의 글쓰기 성과에 미치는 영향을 분석하는 데 초점을 맞춘 반면 본 연구는 참여자들이 자신의 글쓰기 능력을 직접 평가하는 방식을 적용하여, AI 피드백이 학습자의 자기 인식과 메타인지적 인식(metacognitive awareness)에 미치는 영향을 검토하였다. 이를 통해 본 연구는 생성형 AI인 ChatGPT가 학습자의 글쓰기 수행 능력뿐만 아니라 자기평가(self-assessment) 과정에서의 인식 변화에도 영향을 미칠 수 있음을 실증적으로 규명하고, 향후 영어 글쓰기 교육에서 AI 도구의 활용 가능성을 보다 폭넓게 논의하고자 한다.

## 결과 및 논의

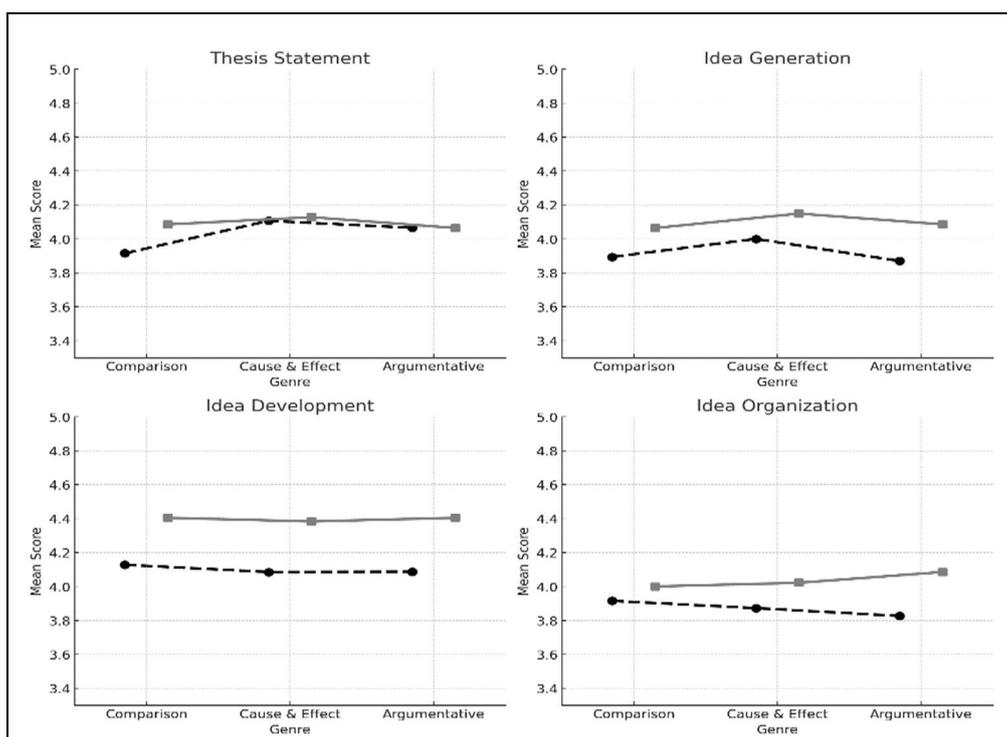
### 예비영어교사의 ChatGPT 피드백과 장르에 따른 글의 거시적 측면의 역량 인식 변화

예비영어교사의 세 가지 서로 다른 장르의 글쓰기에 대한 ChatGPT의 피드백 후 그들의 글쓰기 역량의 인식에 어떠한 변화가 있는지를 살펴보기 위해 반복측정 다변량분산분석을 실시하였다. 그 결과 표4는 ChatGPT 피드백 전후의 거시적 글쓰기 요소에 대한 평균(M) 및 표준 편차(SD)를 세 가지 장르별로 제시하고 있다.

**TABLE 4**  
*Means and SD for Macro Aspects by Genre and ChatGPT*

		Pre-GPT		Post-GPT	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Thesis statement	Comparison & Contrast	3.91	0.95	4.09	1.08
	Cause & Effect	4.11	0.89	4.13	1.01
	Argumentative writing	4.07	0.98	4.17	0.97
Idea generation	Comparison & Contrast	3.89	0.81	4.06	0.96
	Cause & Effect	4.00	0.86	4.15	0.93
	Argumentative	3.87	0.98	4.16	0.93
Idea development	Comparison & Contrast	4.13	0.74	4.40	0.85
	Cause & Effect	4.09	0.75	4.38	0.90
	Argumentative	4.09	0.86	4.42	0.91
Idea organization	Comparison & Contrast	3.91	1.02	4.00	0.93
	Cause & Effect	3.87	1.06	4.02	0.97
	Argumentative	3.83	1.06	4.10	0.86

위의 결과를 종합하면, ChatGPT 피드백 후 모든 장르에서 거시적 글쓰기 요소의 평균 점수가 향상되었으며, AI 기반 피드백이 학습자의 글쓰기 능력 인식의 향상에 긍정적인 영향을 미쳤음을 시사한다. 특히, 아이디어 발전(idea development)의 점수 증가폭이 가장 크며, 세 장르 중에서도 논증 장르에서  $\Delta M = 0.33$ 으로 가장 높은 향상을 보였다. 논증문에서 ‘아이디어 발전’ 점수의 향상이 가장 컸던 이유는 해당 장르가 요구하는 인지적 부담(cognitive load)이 높기 때문으로 보인다. 논증문은 주장, 근거, 반론 제시 등 고차원적인 사고를 요하므로, 학습자들은 ChatGPT의 구조화된 피드백, 예를 들면, 구체적 근거 제안, 논리적 전개 방식 설명 등을 더 유용하게 받아들였을 가능성이 있다. 이러한 피드백은 학습자에게 보다 명확한 아이디어 구성 전략을 제시함으로써, 결과적으로 아이디어 발전에 대한 자기 인식을 높였을 것으로 볼 수 있다. 한편, 아이디어 조직 측면에서도 논증문에서 가장 뚜렷한 향상을 보였는데, 이는 ChatGPT가 명확한 논리적 전개와 글의 구조화를 촉진하는 구체적인 피드백을 제공한 결과로 볼 수 있다. 반면 논제문 작성과 아이디어 생성 측면에서는 다른 요소에 비해 상대적으로 향상폭이 작았으나, 전 장르에 걸쳐 일관된 개선이 나타난 것으로 나타났다. 위 표4를 그래프로 도식화하면 다음 그림2와 같다.



**FIGURE 2**  
Macroscopic Writing Aspects by ChatGPT Feedback and Genre

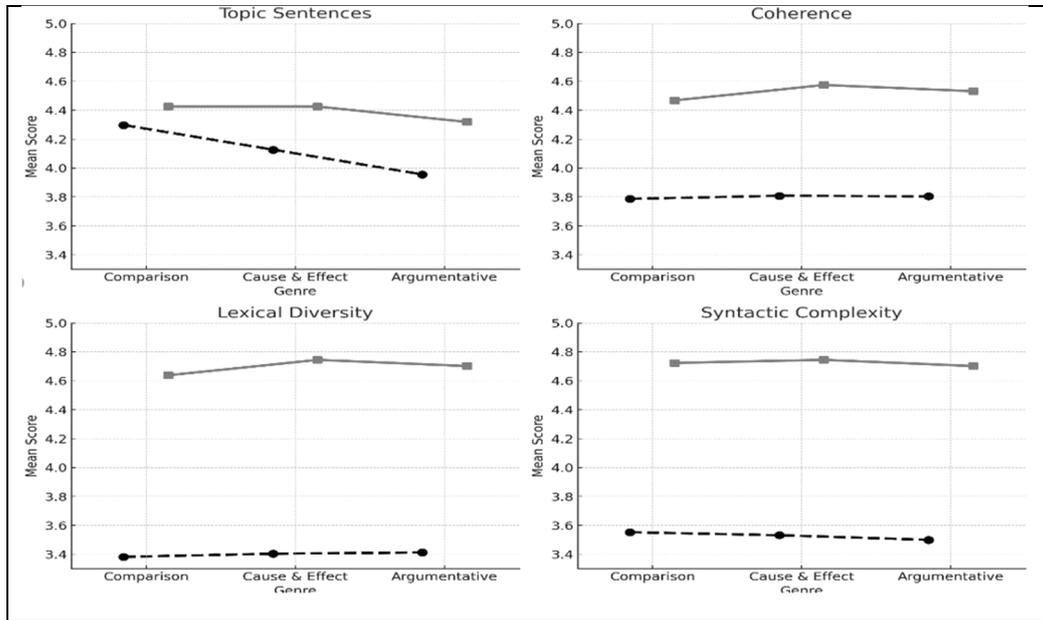
### 예비영어교사의 ChatGPT 피드백과 장르에 따른 글의 미시적 측면의 역량 인식 변화

다음 표5는 예비영어교사의 ChatGPT 피드백 전후의 미시적 글쓰기 요소에 대한 인식의 평균(M) 및 표준편차(SD)를 세 가지 장르(비교, 인과, 논증)별로 제시하고 있다.

**TABLE 5**  
Means and SD for Micro Aspects by Genre and GPT

		Pre-GPT		Post-GPT	
		M	SD	M	SD
Topic sentence	Comparison & Contrast	4.30	0.81	4.43	0.71
	Cause & Effect	4.13	0.82	4.43	0.77
	Argumentative	3.96	0.99	4.30	0.81
Coherence	Comparison & Contrast	3.79	0.93	4.47	0.72
	Cause & Effect	3.81	0.90	4.57	0.62
	Argumentative	3.8	0.96	4.52	0.70
Lexical diversity	Comparison & Contrast	3.38	0.82	4.64	0.82
	Cause & Effect	3.40	0.71	4.74	0.67
	Argumentative	3.41	0.83	4.69	0.74
Syntactic complexity	Comparison & Contrast	3.55	0.95	4.72	0.68
	Cause & Effect	3.53	0.95	4.74	0.67
	Argumentative	3.50	0.96	4.70	0.68

문단의 주제문(topic sentence) 작성 측면, 문장 간 일관성, 어휘의 다양성 및 통사적 복잡성 측면에서, ChatGPT 사용 전보다 ChatGPT의 피드백 후 모든 장르에서 글쓰기 역량의 인식이 향상되었다.



**FIGURE 3**  
*Microscopic Writing Aspects by ChatGPT Feedback and Genre*

그림3의 결과를 살펴보면, 모든 장르에서 주제문 작성, 문장 간 일관성(coherence), 어휘 다양성(lexical diversity), 통사 복잡성(syntactic complexity)의 글쓰기 역량 인식이 현저히 향상되었다. 특히 어휘 다양성과 통사 복잡성 측면에서의 향상폭이 가장 컸으며, 모든 장르에서  $\Delta M$  값이 1 이상으로 나타났다. 이는 ChatGPT가 학생들에게 다양한 어휘와 복잡한 문장 구조를 직접적으로 제안하고, 이를 통해 학생들이 보다 정교하고 풍부한 언어 표현을 사용할 수 있게 된 결과로 볼 수 있다. 또한 문장 간 일관성에서도 모든 장르에서 뚜렷한 향상이 나타났으며, 이는 ChatGPT의 피드백이 문장 간 연결성과 논리적 흐름을 강화하는 데 효과적으로 기여했음을 시사한다. 특히 논증문에서의 미시적 측면 향상이 가장 두드러졌으며, 이는 논증문이 보다 높은 언어적 정교성과 논리적 일관성을 요구하기 때문으로 볼 수 있다.

### 예비영어교사의 ChatGPT와 장르에 따른 글쓰기 역량 인식에 대한 다변량분산분석 결과

본 연구는 GPT 피드백 사용 여부(Pre-GPT vs. Post-GPT)와 장르가 작문의 거시적(macro) 및 미시적(micro) 측면에 미치는 영향을 알아보기 위해 반복측정 다변량 분산 분석(MANOVA)을 실시하였다. 분석에서 독립변인은 GPT 사용 여부와 장르였으며, 종속변인은 거시적 측면(thesis statement, idea generation, idea development, idea organization)과 미시적 측면(topic sentences, coherence, lexical diversity, syntactic complexity)의 총 8개 하위 요소로 설정되었다. 먼저, 거시적 측면에 대한 반복측정 다변량 분산분석 결과는 표6과 같이, GPT 피드백 사용 여부에 따른 주효과는 통계적으로 유의미하였다(Wilks'  $\Lambda = .74$ ,  $F(4, 43) = 3.76$ ,  $p = .010$ ,  $\eta^2 = .26$ ). 이는 GPT 피드백이 전반적인 거시적 작문 능력 향상에 유의미한 영향을 미쳤음을 의미한다. 단변량 분석에서 아이디어 전개( $F(1, 46) = 10.85$ ,  $p = .002$ ,  $\eta^2 = .19$ )와 아이디어 구성/조직( $F(1, 46) = 5.01$ ,  $p = .030$ ,  $\eta^2 = .10$ )이 통계적으로 유의미하게 향상된 반면, 논제문 작성과 아이디어 생성 측면은 통계적으로 유의미하지 않았다( $p > .05$ ). 장르에 따른 주효과는 유의하지 않았다(Wilks'  $\Lambda = .95$ ,  $F(8, 82) = 0.49$ ,  $p = .854$ ), 또한 GPT 사용 여부와 장르 간 상호작용 효과도 통계적으로 유의하지 않았다(Wilks'  $\Lambda = .91$ ,  $F(8, 82) = 0.98$ ,  $p = .457$ ).

**TABLE 6***Repeated-measures MANOVA Results for Macro Aspects*

Effect	Wilks' Lambda ( $\Lambda$ )	<i>F</i>	df1	df2	<i>p</i> -value	$\eta^2$ (Effect Size)
GPT	.74	3.76	4	43	.010	.26
Genre	.95	0.49	8	82	.854	—
GPT × Genre	.91	0.98	8	82	.457	—

미시적 측면에 대한 반복측정 다변량 분산 분석 결과는 표7에서와 같이 GPT 피드백 사용 여부의 주효과가 매우 유의미하였다(Wilks'  $\Lambda = .55$ ,  $F(4, 43) = 8.77$ ,  $p < .001$ ,  $\eta^2 = .45$ ). 이 결과는 GPT 피드백이 미시적 작문 능력을 전반적으로 현저히 향상시켰음을 나타낸다. 단변량 분석 결과, 주제문 작성( $F(1, 46) = 6.23$ ,  $p = .016$ ,  $\eta^2 = .12$ ), 문장 간 일관성( $F(1, 46) = 7.10$ ,  $p = .011$ ,  $\eta^2 = .13$ ), 어휘적 다양성( $F(1, 46) = 15.84$ ,  $p < .001$ ,  $\eta^2 = .26$ ), 그리고 통사적 복잡성( $F(1, 46) = 17.92$ ,  $p < .001$ ,  $\eta^2 = .28$ ) 모두 GPT 피드백 후 유의미하게 향상된 것으로 나타났다. 그러나 장르에 따른 주효과는 유의하지 않았으며(Wilks'  $\Lambda = .93$ ,  $F(8, 82) = 0.65$ ,  $p = .729$ ), GPT 사용 여부와 장르 간 상호작용 효과도 유의하지 않았다(Wilks'  $\Lambda = .89$ ,  $F(8, 82) = 1.32$ ,  $p = .246$ ). 이상의 결과를 종합하면, GPT 기반 피드백은 특히 미시적 작문 능력(어휘 다양성, 문법적 복잡성 등)에 강력한 영향을 미쳤으며, 거시적 측면에서도 특정 요소(예: 아이디어 전개와 아이디어 구성)의 향상에 효과적이었다. 그러나 GPT 피드백 효과가 장르에 따라 다르게 나타나지는 않았음을 확인할 수 있다.

**TABLE 7***Repeated-measures MANOVA Results for Micro Aspects*

Effect	Wilks' Lambda ( $\Lambda$ )	<i>F</i>	df1	df2	<i>p</i> -value	$\eta^2$ (Effect Size)
GPT	.55	8.77	4	43	< .001	.45
Genre	.93	0.65	8	82	.729	—
GPT × Genre	.89	1.32	8	82	.246	—

이는 ChatGPT 피드백이 특정한 장르에 국한되지 않고, 모든 장르에서 비슷한 수준으로 글쓰기 역량 향상에 기여했음을 의미한다. 이러한 결과는 Redmann(2024)의 결과와 유사한 관점에서 논의 될 수 있다. Redmann(2024)은 독일어를 제2언어로 학습하는 학습자를 대상으로 한 장르 기반 글쓰기 수업에서, 생성형 AI가 논증문과 같은 구조화된 장르에서 학습자들의 글쓰기 수행을 지원할 수 있음을 시사하였다. 특히 그는 AI가 학습자의 문장 구성 능력이나 문법적 정확성 향상에는 도움을 줄 수 있지만, 그 피드백이 장르 고유의 담화 구조나 복잡한 의미 전개 방식까지 세밀하게 지도하는 데에는 한계가 있다고 언급하였다. 이는 AI 피드백이 전체적인 구조보다는 문장 수준의 표면적 개선에 집중되는 경향을 반영한다. 본 연구에서는 장르 간 유의미한 차이가 없었으며, 이는 ChatGPT 피드백이 특정 장르에 특화된 것이 아니라, 장르에 관계없이 일정 수준의 향상을 이끌어낼 수 있음을 의미하면서도, ChatGPT의 피드백에 장르적 특성을 반영하여 활용하는 능력이 병행되지 않을 경우, 피상적인 수준의 학습 효과에 그칠 위험도 존재함을 시사한다고 볼 수 있다.

이번 연구 결과는 기존 연구와 비교하여 논의하면 다음과 같다. Choe(2024)의 연구에서는 ChatGPT 피드백이 실제 글쓰기 역량 중 문장 간 일관성(coherence), 어휘적 다양성(lexical diversity), 통사 복잡성(syntactic complexity) 향상에 기여한다고 보고하였는데, 본 연구에서도 ChatGPT 피드백 후 전체적으로 글쓰기 능력이 향상되었음이 확인되며, 이는 영어 학습자의 영어 쓰기 역량을 영어 평가자들이 채점하여 글의 향상 정도를 평가했던 Choe(2024) 연구 결과와 같은 선상에 있다고 보인다. Michel et al. (2025) 연구는 ChatGPT가 기능적 적절성(functional adequacy)을 높이는 데 기여할 수 있다고 주장했는데, 본 연구에서도 ChatGPT가 장르에 관계없이 글의 거시적 측면의 구성력과 미시적 측면의 문장 수준의 표면적인 향상을 개선하는 데 기여한 것으로 보인다. 그러나, 논리적 전개나 구성 등에서는 상대적으로 덜 효과적인 것으로 나타나서, 기존의 연구(Hwang et al., 2024, Redmann, 2024)에서처럼 AI의 피드백은 단순한 문장 수준이며, 피상적이고, 학습자의 깊이있는 논리적 사고와 비판적 사고 체계에까지는 인간 피드백에 비해, 상대적으로 영향력이 미비한 것일 수도 있음을 제시한다.

또한, 장르에 따른 ChatGPT의 피드백이 예비영어교사의 글쓰기 역량 인식 향상에 미치는 영향 측면에서, 기존 연구와는 차이가 보인다. Redmann (2024) 연구는 독일어 쓰기 수업에서 ChatGPT 피드백이 논증적 글쓰기에서 어느

정도 효과적이라고 보고한 반면, 본 연구에서는 논문에서 예비영어교사들이 글쓰기 역량이 상대적으로 다른 장르보다 더 많이 향상된 것으로 인식을 하는 것으로 나타났으나, 장르에 따른 ChatGPT 피드백이 예비영어교사의 글쓰기 역량 인식 향상에 통계적으로 유의미한 차이를 보이지는 않았다. 이러한 원인 중 하나로, 모든 장르별 글쓰기는 본 연구 참가자들이 수강했던 교과목의 후반부에 집중되었고, ChatGPT 피드백도 장르별 글쓰기가 끝난 뒤 한 번에 진행되었다. 이에 따라, 장르에 따른 인식 차이가 통계적으로 유의미하지 않게 나타났을 가능성이 있으며, 이로 인해 피드백 효과가 장르 간 차별화되지 않았다는 결과가 도출되었을 수 있다. 이와 더불어, 장르별 차이가 나타나지 않은 또 다른 이유로는 영어 쓰기에서는 AI 피드백이 특정 장르에 제한되지 않고 전반적으로 효과적일 수 있음을 나타내거나 또는, AI의 피드백이 비판적 논증 발전에 미치는 영향력이 미비하였기 때문일 수도 있다. De Oliveira and Santos (2025) 연구에서 ChatGPT가 장르별 멘토 텍스트로 활용될 수 있다고 주장했지만, 본 연구에서는 장르별 차이가 나타나지 않았으므로 앞으로 장르 기반 글쓰기 교육에서 장르 특화된 AI의 최적화된 피드백 전략이 필요함을 제시한다.

이러한 결과를 종합적으로 요약하면, ChatGPT 피드백은 장르를 가리지 않고 전반적인 글쓰기 개선에 효과적이었다고 연구 참가자들은 인식하였다. 이러한 결과는 ChatGPT를 활용하는 영어 학습자 또는 교사들이 영어 글쓰기를 가르칠 때, 어떤 유형의 글을 작성하든 AI 기반 피드백이 도움이 될 가능성이 높다는 것을 시사한다. 그러나 장르별 피드백의 효과를 인식하기 위해서는 장르별 특성을 반영한 추가적이고 최적화된 학습 전략을 제시하고 가르칠 필요가 있어 보인다. 이러한 결과를 통해, AI 피드백을 효과적으로 활용하기 위해서는 학습자가 비판적 사고 능력을 갖추고, AI 피드백을 선택적으로 활용하는 전략이 필요함을 의미한다. 향후 연구에서는 AI 피드백을 적용할 때 학습자들이 실제로 어떤 방식으로 수정 결정을 내리는지(revision decision-making process)도 심층적으로 분석할 필요가 있다.

## 결론

본 연구는 예비 영어 교사들이 ChatGPT 피드백을 활용하여 영어 글쓰기를 수행한 후, 거시적(macroscopic) 및 미시적(microscopic) 측면에서의 글쓰기 역량에 대한 스스로의 인식의 변화를 분석하였다. 기술통계 분석 결과, 모든 장르에서 ChatGPT 피드백 후 스스로 역량이 향상되었다고 인식하는 것으로 나타났으며, 특히 아이디어 발전(idea development)과 어휘 다양성(lexical diversity)의 향상이 두드러졌다. 다변량 분산분석(MANOVA) 결과, ChatGPT 피드백의 주효과는 유의미한 차이를 보였으며, 글쓰기 장르 간 차이는 통계적으로 유의하지 않았다. 또한, ChatGPT 피드백과 글쓰기 장르 간의 상호작용 효과도 유의하지 않은 것으로 나타났다. 이는 ChatGPT 피드백이 장르에 관계 없이 전반적인 글쓰기 역량을 강화하는 데 기여할 수 있음을 시사한다. 기존 연구에서도 AI 기반 피드백이 문장 간 일관성, 어휘의 다양성, 구문 복잡성 향상에 긍정적인 영향을 미친다고 보고되었으며 (Choe, 2024; Redmann, 2024), 본 연구의 결과 역시 이러한 경향과 일치한다. 특히, 본 연구에서 확인된 어휘 다양성의 향상은 기존 연구(Michel et al., 2025)보다 더욱 두드러진 결과로, ChatGPT가 다양한 표현을 제공함으로써 학습자의 언어적 확장성을 촉진할 가능성을 시사한다.

본 연구는 기존의 AI 피드백 연구와 차별화되는 몇 가지 의의를 가진다. 기존 연구들은 주로 AI 피드백이 특정 글쓰기 장르에서 어떻게 활용될 수 있는지, 또는 AI가 생성한 텍스트와 인간이 작성한 텍스트를 비교하는 데 초점을 맞춘 반면 (De Oliveira & Santos, 2025; Markey et al., 2024), 본 연구는 예비 영어 교사들이 GenAI를 교실에서 활용하여 피드백을 제공받고 글쓰기를 학습하고 가르칠 수 있도록, 스스로 자기 평가 방식으로 AI 피드백의 효과성을 자기평가하여 분석했다는 점에서 차별성을 가진다. 또한, AI가 특정 장르에서 더 효과적인지에 대한 논의가 계속되고 있는 상황에서(Choe, 2024; Redmann, 2024), 본 연구는 ChatGPT 피드백이 특정 장르에 국한되지 않고 전반적인 글쓰기 능력 향상에 기여할 가능성이 있음을 제시하였다. 또한, 기존 연구(Choe, 2024)가 외부 평가자에 의한 글쓰기 평가에 초점을 맞춘 반면, 본 연구는 학습자의 자기평가(self-assessment)를 중심으로 ChatGPT 피드백의 효과를 분석하였다. 이는 학습자의 메타인지적 인식과 자기성찰을 중시하는 최근 영어교육 연구의 흐름과 맞닿아 있다. 이러한 점에서 본 연구는 생성형 AI 기반 글쓰기 피드백에 대한 학습자 인식 연구의 영역을 확장시킨다. 더 나아가, AI 피드백이 학습자들의 자기 조정 전략(self-regulated learning)에 미치는 영향에 대한 연구가 필요한 상황에서, 본 연구는 AI가 학습자의 자기 평가 능력을 어떻게 강화할 수 있는지를 탐색하는 데 기여하였다.

본 연구는 자기 평가 기반의 설문 데이터를 활용하여 학습자의 인식을 측정하였다는 점에서 기존 연구(Choe, 2024)에서 글쓰기 역량 향상을 객관적으로만 측정했던 연구를 보완하였지만 객관적 데이터와 함께 살펴보기 못했다는 한계가 존재한다. 따라서 후속 연구에서는 자기 평가 데이터를 보완할 수 있는 객관적 평가 방식을 병행할 필요가 있다. 예를 들어, 전문 평가자(expert raters)의 채점, AI 기반 자동 평가(automated writing evaluation, AWE), Coh-matrix 등의 텍스트 분석 기법 등을 활용하여 학습자들이 인식한 변화가 실제 글쓰기 성과 향상으로 이어지는지를 검증하는 것이 필요하다. 자기 평가 데이터와 객관적 평가 데이터를 함께 분석할 경우, ChatGPT의 글쓰기 피드백 효과를 보다 정밀하게 규명할 수 있을 것이다. 또한, 특정 대학의 예비영어교사들만을 대상으로 하여 일반화하기에는 다소 무리가 있다. 향후 연구에서는 AI 피드백을 활용한 장기적인 실험 연구 및 교사 피드백과의 비교 연구가 필요하며, 더 나아가, 학습자의 인지적, 정서적 반응 분석도 질적으로 함께 살펴볼 필요가 있다. 이를 통해 AI 기반 글쓰기 교육의 최적화 방안을 도출할 필요가 있다.

## References

- Algaraady, J., & Mahyoob, M. (2023). ChatGPT's Capabilities in Spotting and Analyzing Writing Errors Experienced by EFL Learners. *Arab World English Journal (AWEJ) Special Issue on CALL*, (9), 3–17. <https://dx.doi.org/10.24093/awej/call9.1>
- Athanassopoulos, S., Manoli, P., Gouvi, M., Lavidas, K., & Komis, V. (2023). The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom. *Advances in Mobile Learning Educational Research*, 3(2), 818–824.
- Barkaoui, K. (2010). Variability in ESL essay rating processes: The role of the rating scale and rater experience. *Language Assessment Quarterly*, 7(1), 54–74. <https://doi.org/10.1080/15434300903464418>
- Cao, S., & Zhong, L. (2023). Exploring the effectiveness of ChatGPT-based feedback compared with teacher feedback and self-feedback: Evidence from Chinese to English translation. arXiv. <https://arxiv.org/abs/2309.01645>
- Choe, Yoonhee. (2023a). Exploring perceptions of Korean pre-service English teachers on using ChatGPT in the English writing process. *Journal of the Korea English Education Society*, 22(2), 243–262.
- Choe, Yoonhee. (2023b). Exploring ChatGPT's impact on the English summary writing of pre-service English teachers. *Multimedia-Assisted Language Learning*, 26(2), 104–132. <https://doi.org/10.15702/mall.2023.26.2.104>
- Choe, Yoonhee. (2024). The influence of English genres and ChatGPT on Korean EFL learners' writing performance. *English Language Assessment*, 19(2), 131–149. <https://doi.org/10.37244/ela.2024.19.2.131>
- Choe, Yoonhee., Ryu, Jisu., & Jeon, Moongee. (2020). Effects of source texts' types on Korean EFL college learners' summary writing performance. *The Journal of Linguistic Science*, 95, 331–355. <https://doi.org/10.21296/jls.2020.12.95.331>
- Choe, Yoonhee., Ryu, Jisu., & Jeon, Moongee. (2022). Effects of pre-task planning and source texts on Korean EFL college learners' summary writing. *The Journal of AsiaTEFL*, 19(1), 109–124. <https://doi.org/10.21296/jls.2023.06.105.1>
- De Oliveira, L. C., & Santos, A. E. (2025). Using AI-generated mentor texts for genre-based pedagogy in second language writing. *Journal of Second Language Writing*, 67, 101184. <https://doi.org/10.1016/j.jslw.2025.101184>
- Educational Testing Service. (n.d.-b). *Scoring Guides (Rubrics) for Writing Responses*. Retrieved April 7, 2025, from <https://www.ets.org/pdfs/toefl/toefl-ibt-writing-rubrics.pdf>
- Fitria, T. N. (2023). The use of ChatGPT in assisting EFL writing: A practical overview. *Journal of English Teaching and Linguistics*, 4(1), 77–86.
- Gao, C., Howard, C. D., & Li, M. (2022). How ChatGPT is shaping the future of writing. *Journal of Artificial Intelligence and Education*, 5(3), 112–123.
- Heng, R., Pu, L., & Liu, X. (2023). The effects of genre on the lexical richness of argumentative and expository writing by Chinese EFL learners. *Frontiers in Psychology*, 13, Article 1082228. <https://doi.org/10.3389/fpsyg.2022.1082228>
- Hwang, Myungwhan., Jeons, R., & Lee, Hee-Kyung. (2024). Exploring learner prompting behavior and its effect on ChatGPT-assisted English writing revision. *The Asia-Pacific Education Researcher*. Advance online publication. <https://doi.org/10.1007/s40299-024-00930-6>
- Kim, Sung-yeon & Jeong, Nam-sook. (2025). English for Specific Purpose (ESP) with ChatGPT: ESP learners' reading proficiency and perceptions. *Modern English Education*, 26, 84-99.
- Kohnke, L., & Zou, D. (2024). The impact of ChatGPT on English language learners' writing skills: An assessment of its effectiveness. *International Journal of Interactive Mobile Technologies*, 18(3), 4–15.
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11(1), 9. <https://doi.org/10.1186/s40561-024-00295-9>
- Markey, B., Brown, D. W., Laudenbach, M., & Kohler, A. (2024). Dense and disconnected: Analyzing the sedimented style of ChatGPT-generated text at scale. *Written Communication*, 41(4), 571–600. <https://doi.org/10.1177/07410883241263528>
- Michel, M., Bazhutkina, I., Abel, N., & Strobl, C. (2025). Collaborative writing based on generative AI models: Revision and deliberation processes in German as a foreign language. *Journal of Second Language Writing*, 67, 101185.

- Ohio University. (2023). *Center for Teaching, Learning, and Assessment hosting activities focused on ChatGPT*. Retrieved April 7, 2025, from <https://www.ohio.edu/writing-center> Center for Teaching, Learning, and Assessment hosting activities focused on ChatGPT.
- Redmann, D. (2024). Genre-based writing in the German classroom in the age of generative AI. *Die Unterrichtspraxis*, 57(1), 45–67.
- Wang, C., & Wang, Z. (2025). Investigating L2 writers' critical AI literacy in AI-assisted writing: An APSE model. *Journal of Second Language Writing*, 67, 101187. <https://doi.org/10.1016/j.jslw.2025.101187>
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511732997>
- Yan, B. (2023). ChatGPT and academic writing: Opportunities and challenges. *Asian Journal of Education and e-Learning*, 11(2), 50–57.
- Yang, W., & Lin, C. (2025). Translanguaging with generative AI in EFL writing: Students' practices and perceptions. *Journal of Second Language Writing*, 67, 101181. <https://doi.org/10.1016/j.jslw.2025.101181>
- Yao, Y., Zhu, X., Xiao, L., & Lu, Q. (2025). Secondary school English teachers' application of artificial intelligence-guided chatbot in the provision of feedback on student writing: An activity theory perspective. *Journal of Second Language Writing*, 67, 101175. <https://doi.org/10.1016/j.jslw.2025.101175>
- Zhou, T., Cao, S., Zhou, S., Zhang, Y., & He, A. (2023). Chinese Intermediate English Learners outdid ChatGPT in deep cohesion: Evidence from English narrative writing. *System*, 118, 103141. <https://doi.org/10.1016/j.system.2023.103141>

**부록**

**Can-Do Self-Assessment (자기 평가 설문지)**

※ 다음은 여러분이 영어로 다양한 장르의 글을 작성할 때 어느 정도 할 수 있는지 살펴보기 위한 설문입니다. 솔직하게 응답해주세요.

1: 전혀 아니다. 2: 약간 아니다. 3: 보통. 4: 약간 그렇다. 5: 매우 그렇다.

**Cause & Effect(원인/결과)**

문항	1	2	3	4	5
1. 나는 원인과 결과 장르의 Thesis Statement를 잘 작성할 수 있다.	1	2	3	4	5
2. 나는 원인과 결과 장르의 idea를 잘 생각해 낼 수 있다.	1	2	3	4	5
3. 나는 원인과 결과 장르의 idea를 서론, 본론, 결론으로 잘 발전시킬 수 있다.	1	2	3	4	5
4. 나는 원인과 결과 장르의 idea를 Block 또는 Chain으로 잘 구성할 수 있다.	1	2	3	4	5
5. 나는 원인과 결과 장르의 글을 쓸 때, 각 문단의 topic sentence를 잘 쓸 수 있다.	1	2	3	4	5
6. 나는 원인과 결과 장르의 글을 쓸 때, 문장 간의 유기적(Coherence)으로 잘 연결할 수 있다.	1	2	3	4	5
7. 나는 원인과 결과 장르의 글을 쓸 때, 어휘를 잘 생각해내고 적절하게 쓸 수 있다.	1	2	3	4	5
8. 나는 원인과 결과 장르의 글을 쓸 때, 통사적(문법적)으로 문장을 잘 구성할 수 있다.	1	2	3	4	5

**Comparison & Contrast(비교/대조)**

문항	1	2	3	4	5
1. 나는 비교/대조 장르의 글을 쓸 때, Thesis Statement를 잘 작성할 수 있다.	1	2	3	4	5
2. 나는 비교/대조 장르의 idea를 잘 생각해 낼 수 있다.	1	2	3	4	5
3. 나는 비교/대조 장르의 idea를 서론, 본론, 결론으로 잘 발전시킬 수 있다.	1	2	3	4	5
4. 나는 비교/대조 장르의 idea를 Block 또는 Point-by-Point로 잘 구성할 수 있다.	1	2	3	4	5
5. 나는 비교/대조 장르의 글을 쓸 때, 각 문단의 topic sentence를 잘 쓸 수 있다.	1	2	3	4	5
6. 나는 비교/대조 장르의 글을 쓸 때, 문장 간의 유기적(Coherence)으로 잘 연결할 수 있다.	1	2	3	4	5
7. 나는 비교/대조 장르의 글을 쓸 때, 어휘를 잘 생각해내고 적절하게 쓸 수 있다.	1	2	3	4	5
8. 나는 비교/대조 장르의 글을 쓸 때, 통사적(문법적)으로 문장을 잘 구성할 수 있다.	1	2	3	4	5

**Argumentative writing(논설/논증)**

문항	1	2	3	4	5
1. 나는 논증문 장르의 Thesis Statement를 잘 작성할 수 있다.	1	2	3	4	5
2. 나는 논증문 장르의 idea를 잘 생각해 낼 수 있다.	1	2	3	4	5
3. 나는 논증문 장르의 idea를 서론, 본론, 결론으로 잘 발전시킬 수 있다.	1	2	3	4	5
4. 나는 논증문 장르의 idea를 Block 또는 Point-by-Point로 잘 구성할 수 있다.	1	2	3	4	5
5. 나는 논증문 장르의 글을 쓸 때, 각 문단의 topic sentence를 잘 쓸 수 있다.	1	2	3	4	5
6. 나는 논증문 장르의 글을 쓸 때, 문장 간의 유기적(Coherence)으로 잘 연결할 수 있다.	1	2	3	4	5
7. 나는 논증문 장르의 글을 쓸 때, 어휘를 잘 생각해내고 적절하게 쓸 수 있다.	1	2	3	4	5
8. 나는 논증문 장르의 글을 쓸 때, 통사적(문법적)으로 문장을 잘 구성할 수 있다.	1	2	3	4	5

## 내가 ‘Cause &amp; Effect’ 글을 쓸 때, ChatGPT의 피드백은

문항	1	2	3	4	5
1. Thesis Statement를 잘 작성할 수 있게 도움 주었다.	1	2	3	4	5
2. idea를 잘 생각해 낼 수 있게 도왔다.	1	2	3	4	5
3. idea를 서론, 본론, 결론으로 잘 발전시킬 수 있게 도움 주었다.	1	2	3	4	5
4. idea를 Block 또는 Chain으로 잘 구성할 수 있게 도움 주었다.	1	2	3	4	5
5. 각 문단의 topic sentence를 잘 쓸 수 있게 도움 주었다.	1	2	3	4	5
6. 문장 간의 유기적(Coherence)으로 잘 연결할 수 있게 도움 주었다.	1	2	3	4	5
7. 어휘를 잘 생각해내고 적절하게 쓸 수 있게 도움 주었다.	1	2	3	4	5
8. 통사적(문법적)으로 문장을 잘 구성할 수 있게 도움 주었다.	1	2	3	4	5

## 내가 ‘Comparison &amp; Contrast’ 글을 쓸 때, ChatGPT의 피드백은

문항	1	2	3	4	5
1. Thesis Statement를 잘 작성할 수 있게 도움 주었다.	1	2	3	4	5
2. idea를 잘 생각해 낼 수 있게 도왔다.	1	2	3	4	5
3. idea를 서론, 본론, 결론으로 잘 발전시킬 수 있게 도움 주었다.	1	2	3	4	5
4. idea를 Block 또는 Point-by-Point으로 잘 구성할 수 있게 도움 주었다.	1	2	3	4	5
5. 각 문단의 topic sentence를 잘 쓸 수 있게 도움 주었다.	1	2	3	4	5
6. 문장 간의 유기적(Coherence)으로 잘 연결할 수 있게 도움 주었다.	1	2	3	4	5
7. 어휘를 잘 생각해내고 적절하게 쓸 수 있게 도움 주었다.	1	2	3	4	5
8. 통사적(문법적)으로 문장을 잘 구성할 수 있게 도움 주었다.	1	2	3	4	5

## 내가 ‘Argumentative writing’을 쓸 때, ChatGPT의 피드백은

문항	1	2	3	4	5
1. Thesis Statement를 잘 작성할 수 있게 도움 주었다.	1	2	3	4	5
2. idea를 잘 생각해 낼 수 있게 도왔다.	1	2	3	4	5
3. idea를 서론, 본론, 결론으로 잘 발전시킬 수 있게 도움 주었다.	1	2	3	4	5
4. idea를 Block 또는 Point-by-Point으로 잘 구성할 수 있게 도움 주었다.	1	2	3	4	5
5. 각 문단의 topic sentence를 잘 쓸 수 있게 도움 주었다.	1	2	3	4	5
6. 문장 간의 유기적(Coherence)으로 잘 연결할 수 있게 도움 주었다.	1	2	3	4	5
7. 어휘를 잘 생각해내고 적절하게 쓸 수 있게 도움 주었다.	1	2	3	4	5
8. 통사적(문법적)으로 문장을 잘 구성할 수 있게 도움 주었다.	1	2	3	4	5