

AI · 동료 · 교수자의 단계적 피드백이 대학생의 글쓰기 수행에 미치는 효과

이윤빈 덕성여자대학교 차미리사교양대학 조교수

- I. 서론
- II. 선행 연구 검토 및 단계적 피드백 방안의 구성
- III. 연구 방법
- IV. 연구 결과
- V. 결론

I. 서론

대학 글쓰기 교육에서 학습자에게 어떤 주체의 피드백을 어떠한 방식으로 제공해야 하는가는 지속적으로 고민되어 온 문제다. 전문성을 지닌 교수자의 피드백은 학습자의 수준에 기반한 맞춤형 지원을 통해 쓰기 능력과 동기를 증진할 수 있는 가장 효과적인 교육 방안의 하나로 여겨진다(Renati, Renata, Hidayah, & Indriani, 2024). 그러나 한 명의 교수자가 다수 학습자의 글쓰기에 일대일로 개입하기에는 현실적으로 큰 부담이 따른다. 그래서 교수자들은 다양한 대안적 피드백 방안을 사용해 왔다. 학습자의 글에 공통적으로 나타나는 문제를 가진 대표 사례 글을 선정하여 피드백을 하거나, 교수자의 안내에 따라 학습자가 자기 글을 점검하게 하거나, 동료 피드백을 시행하는 방안 등이 그것이다.

그러나 대표 사례 글 중심의 피드백은 개별 학습자의 문제에 정확하게 대응할 수 없어 실제 수정으로의 연결성이 떨어지고, 자기 점검은 필자의 자기 중심성(egocentrism)을 극복하고 타인의 관점에서 글을 객관화해 보는데 한계가 있다. 동료 피드백은 필자의 눈높이에 맞는 언어와 방식으로 제공

되어 실제 수정으로 이어질 가능성이 높다는 점에서 가장 효과적인 대안적 피드백 방안으로 활용되고 있지만(정희모·이재성, 2008; Oley, 1992), 동료의 쓰기 수준에 따라 잘못된 피드백이 제공될 수 있는 문제가 지속적으로 지적되고 있다(이윤빈·정희모, 2014; Hovardas, Tsivitanidou, & Zacharia, 2014; Topping, 1998).

최근에는 생성형 AI의 대중화에 따라, AI 피드백의 활용 가능성에 대한 관심이 높아지고 있다. AI 피드백은 학습자가 시간과 장소에 구애받지 않고 신속하고 반복적으로 제공받을 수 있어 학습자의 자율적인 글쓰기 점검을 지원하는 도구로 주목받고 있다(Wang & Han, 2022). 그러나 문단 구성이나 문법 오류 등 글의 구조 및 표현 측면에 대한 피드백에서는 안정적인 타당도를 유지하는 반면, 내용의 타당성이나 맥락 적합성과 같은 내용적 측면에 대한 심층적 판단에는 한계를 보인다는 지적이 잇따르고 있다(박혜림·이의연, 2025; 이윤빈, 2025; Bogolepova, 2025; Yoon, Miszoglad, & Pierce, 2023).

이 연구는 각각의 강점과 한계를 지닌 교수자, 동료, AI의 피드백을 상호 보완적으로 활용한다면 학습자의 글쓰기를 보다 효과적으로 지원할 수 있으리라는 문제의식에서 기획되었다. 이 연구의 목적은 대학 글쓰기 교육에서 AI·동료·교수자가 역할을 분담한 단계적 피드백 방안을 제안하고, 이를 실제 수업 현장에 적용하여 그 효과를 탐색하는 데 있다. 구체적으로, 이 연구는 AI는 글의 구조 및 표현을 포함하는 형태 중심 피드백을, 동료는 독자 반응의 일환으로서 내용 중심 피드백을, 교수자는 형태와 내용을 아우르는 심층 통합적 피드백을 순차적으로 제공하여, 학습자가 이를 차례로 점검 및 반영하며 자신의 논증문을 수정하게 할 것이다. 그리고 각 단계에서 학습자가 각 주제별 피드백에 어떻게 반응하며, 글의 질은 어떻게 변화하는지를 살펴 단계적 피드백 방안의 효과를 탐색적으로 확인하고자 한다.

이 연구가 다루는 구체적인 문제는 다음과 같다.

1. AI·동료·교수자의 단계적 피드백에 대한 대학생 학습자의 반응 양상은 어떠한가?
 - 1-1. 단계적 피드백에 대한 전반적인 반응 양상은 어떠한가?
 - 1-2. 피드백 제공 단계별 학습자의 반응 양상은 어떠한가?
2. AI·동료·교수자의 단계적 피드백은 대학생 학습자가 작성한 글의 질에 어떠한 영향을 미치는가?
 - 2-1. 글의 전반적인 질은 어떻게 변화하는가?
 - 2-2. 피드백이 제공된 영역별 글의 질은 어떻게 변화하는가?

Ⅱ 장에서는 교수자, 동료, AI 피드백 관련 연구 동향에 대한 검토를 바탕으로, 단계적 피드백 방안의 구성 원리 및 적용 절차를 제안한다. Ⅲ 장에서는 단계적 피드백 방안을 실제 대학 글쓰기 교육 현장에 적용하여 그 효과를 검토하고자 한다.

II. 선행 연구 검토 및 단계적 피드백 방안의 구성

1. 교수자, 동료, AI 피드백에 대한 연구 동향

글쓰기 피드백의 본질이 텍스트를 매개로 이루어지는 대화적 교섭과 지식의 순환에 있다고 할 때(공성수, 2022) 피드백의 제공 주체가 누구인가는 학습자가 피드백에 반응하고 성장하는 방식에 큰 영향을 미친다. 그래서 국내외의 연구자들은 피드백 제공 주체에 따른 피드백의 강점과 한계를 다각도로 탐색해 왔다.

피드백 제공 주체에 따른 피드백 효과에 대한 연구는 주로 교수자 피드백과 동료 피드백을 비교하는 데 집중되어 왔으나(김연희·김남국, 2013;

정희모·이재성, 2008; Gielen, Peeters, Dochy, Onghena, & Struyven, 2010; Paulus, 1999; Yang, Badger, & Yu, 2006), 2022년 말 생성형 AI가 대중화된 이후로는 AI 피드백의 효과를 탐색하는 연구들(박혜림·이의연, 2025; 안효원·나현수·하채희·송용창, 2025; 정유남, 2024; AlGhamdi, 2024; Escalante, Pack, & Barrett, 2023)과 더불어, 교수자 피드백과 AI 피드백의 비교(김은선, 2025; 심지연, 2024; Steiss et al., 2024), 동료 피드백과 AI 피드백의 비교(이유나·손은지·이상수, 2024; Banihashem, Kerman, Noroozi, Moon, & Drachsler, 2024; Chan, Lo, & Wong, 2025; Ji, Zhao, & Wu, 2025), 또는 교수자·동료·AI 피드백 모두를 비교(김민정, 2024; 이윤빈, 2025; Zeevy-Solovey, 2024)하는 연구들이 활발하게 진행되고 있다. 이러한 연구들에서 반복적으로 확인되는 주체별 피드백의 강점과 한계를 정리하면 <표 1>과 같다.

<표 1> 제공 주체별 피드백의 강점과 한계

구분	강점	한계
교수자 피드백	<ul style="list-style-type: none"> 전문성과 신뢰성을 바탕으로 학습자의 수준과 맥락을 고려한 맞춤형 조언 글의 전체 영역(내용·구조·표현)을 아우르는 심층적·통합적 조언 	<ul style="list-style-type: none"> 시간과 노력 소요가 커 대규모 수업에서 접근성과 즉시성에 제약 수정 요구 수준이 높아 실제 반영으로 이어지기 어려울 수 있음
동료 피드백	<ul style="list-style-type: none"> 학습자 관점의 실용적 조언으로 실제 수정으로 이어질 가능성 높음 공감과 정서적 지지를 통해 쓰기 동기 및 효능감 증진 	<ul style="list-style-type: none"> 동료의 쓰기 수준에 따라 피드백의 타당도에 편차 발생 심층적·통합적 조언은 어려워 글의 질 향상에 일정한 한계 존재
AI 피드백	<ul style="list-style-type: none"> 시간·장소에 구애받지 않고 즉각적·반복적으로 활용 가능 구조나 표현 등 형태적 측면에서 일정 수준 이상의 타당도 확보 	<ul style="list-style-type: none"> 글쓰기 맥락이나 필자의 의도, 정서적 요소에 대한 이해 부족 과도한 수정 시 학습자의 비판적 사고와 저자성 약화 우려

<표 1>에서 볼 수 있듯이, 교수자, 동료, AI의 피드백은 각기 다른 강점과 한계를 지니고 있어, 어느 한 방식에만 의존하기보다는 이들을 상호 보완적으로 활용하는 방안이 필요하다. 그러나 주체별 피드백의 효과를 밝히는

연구에 비하여, 이들을 효과적으로 결합하여 활용하는 방안을 제시한 연구는 상대적으로 드물다.

국내에서는 박혜림·이의연(2025)이 대학생의 디지털 글쓰기에 AI 피드백을 제공한 효과를 분석하면서, AI 피드백은 글의 맥락적·정서적 요소를 충분히 반영하지 못하므로 교수자·동료 피드백과 병행해 이를 보완할 필요가 있다고 주장했다. 이윤빈(2025)에서도 교수자·동료·AI 피드백의 특성과 타당도를 비교한 결과에 대해 논의하며, 앞으로의 대학 글쓰기 교육에서는 세 주체의 피드백을 결합하여 활용할 필요가 있다고 제안한 바 있다. 그러나 이러한 사례를 제외하면, 대학 글쓰기 교육 현장에서 세 주체의 피드백을 실제로 결합하여 적용하고 효과를 분석한 연구는 아직 찾아보기 어렵다.

국외에서는 단일 주체의 피드백에 비해 여러 주체의 혼합 피드백이 학습자의 쓰기 능력을 효과적으로 신장한다는 연구 결과들이 지속적으로 보고되고 있다. Jiang(2025)은 AI 피드백만을 제공한 대학생 집단과 AI와 교수자 피드백을 함께 제공한 대학생 집단의 글쓰기 수행을 비교하여, 혼합 피드백 제공 집단에서 글쓰기 수행 능력이 유의미하게 향상되었다고 밝혔다. AI 피드백의 즉시성과 교수자 피드백의 풍부한 상호작용 및 맞춤형 조언이 상승(synergy) 효과를 발생시킨 결과였다. Mirsanjari & Moradi(2024)도 교수자·동료·AI 피드백을 모두 제공한 대학생 집단에서 쓰기 능력뿐 아니라 비판적 사고력과 자기 성찰, 협력적 태도 측면 모두의 유의미한 상승이 관찰되었다고 보고했다.

이러한 결과를 바탕으로, Sperber, MacArthur, Minnillo, Stillman, & Whithaus(2025)은 동료와 AI 피드백을 결합한 'PAIRR(Peer and AI Review + Reflection)' 모델을 제안하고 그 효과를 검증했다. 대학생 집단이 초고 완성 후 동료 및 AI 피드백을 제공받고, 두 피드백을 비교 및 비판적으로 평가하여 주체적으로 수정 방안을 도출하게 하는 모델이었다. 대학생들은 동료와 AI 피드백을 평가하는 과정에서 AI 리터러시와 자기 주도적 글쓰기 역량의 신장을 경험했다. 그러나 일부 학습자들은 동료와 AI가 상이한 피드백을

제공할 때 어떤 피드백을 수용해야 할지 판단하지 못해 어려움을 겪었다. 즉, 상이한 피드백을 비판적으로 비교·선별할 수 있는 메타인지적 역량이 충분하지 않은 학습자에게는 다양한 피드백이 오히려 혼란을 초래할 수 있다는 점이 이 모델의 한계인 것으로 평가된다. 또한 이러한 한계는 교수자의 중재가 부재한 채 학습자에게 비교와 판단을 전적으로 맡긴 PAIRR 모델의 구조적 특성과도 관련되는 것으로 판단된다.

한편, AI·동료·교수자의 피드백을 순차적으로 제공한 사례로는 Zhang & Hyland(2022)가 있다. 연구자들은 중국 대학에서 영어 글쓰기 강좌를 수강하는 학습자들에게 자동화된 AWE 시스템·동료·교수자 피드백을 차례로 제공하며 단계적인 수정을 수행하게 했다. 그 결과, AI 피드백은 주로 문법이나 문장 구조와 같은 형태적 오류 수정에 기여하고, 동료 피드백은 공감적이나 비교적 제한된 내용 조언에 머무르며, 교수자 피드백은 글의 논리에 대한 심층적인 조언으로 작용했다. 그러나 각 주체의 피드백이 갖는 강점을 살리도록 역할을 분담하여 제공한 것은 아니며, 실제 적용 과정에서 결과적으로 기능 차이가 드러났다. 또한 학습자가 각 주체별 피드백을 비판적으로 성찰하고 수용하는 과정도 충분히 요구되지 않았다. 아울러 이 연구는 L2 맥락에서 이루어져, 모국어 기반의 글쓰기 환경과는 본질적인 차이를 갖는다.

이처럼 다양한 주체의 피드백을 제공하는 교육 방안을 제안한 소수의 연구 사례에서도, 교수자의 중재 부재로 인한 혼란이나 주체별 피드백의 고유한 강점을 살리지 못한 한계, 학습자의 성찰 계기 부족과 같은 문제점이 나타났다. 이러한 한계를 극복하기 위해서는 각 주체의 강점을 전면화하도록 역할을 분담하고, 피드백에 대한 학습자의 비판적 성찰을 유도하며, 학습자가 피드백 수용과 수정 과정에서 경험하는 문제를 교수자가 조력할 수 있는 새로운 방안이 필요하다. 그러한 방안의 구성 원리와 구체적인 절차는 2절과 3절에서 제시하고자 한다.

2. 단계적 피드백 방안의 구성 원리

이 연구에서 제안하는 단계적 피드백 방안은 다음 세 가지 원리를 기반으로 구성되었다.

첫째, AI·동료·교수자가 역할을 분담하여 단계적으로 피드백을 제공함으로써 학습자가 각 주체의 강점을 상호 보완적으로 활용하여 초점화된 수정을 수행할 수 있도록 했다.

Sperber et al.(2025)의 PAIRR 모델에서 확인할 수 있듯이, 각 주체가 글의 전(全) 영역에 대해 피드백을 제공할 경우 검토해야 하는 피드백의 분량이 상당할 뿐 아니라, 동일한 문제에 대해 서로 충돌하는 피드백이 빈번히 제공되어 학습자가 혼란을 경험할 수 있다. 이를 방지하기 위해 이 연구는 피드백 제공 주체별 강점과 한계에 대한 선행 연구들의 결과(〈표 1〉)를 바탕으로, AI는 형태 중심 피드백을, 동료는 내용 중심 피드백을, 교수는 심층 통합적 피드백을 담당하여 이를 순차적으로 제공하게 했다.

이때 피드백 대상 영역은 글의 구조와 표현을 ‘형태’로, 의미적 요소를 ‘내용’으로 규정한 Ellis(2001)와 Fatmah, Warsono, & LB(2022)의 논의에 기반하여 〈표 2〉와 같이 구분했다.¹⁾ 이를 통해, (1) 먼저 AI가 글이 서론·본

1) 글의 ‘형태(形態, form)’와 ‘내용(內容, content)’에 해당하는 요소를 구분하고자 할 때 ‘구조’를 어느 영역에 포함시킬 것인가에 대한 시각은 연구자마다 상이하다. 영어 교육(EFL/L2) 분야에서는 형태 중심 피드백의 범위에 문법, 철자, 어휘, 문장 구조 등 표현의 정확성만을 포함하는 시각(Lightbown & Spada, 1990)이 우세해 왔으나, 서론·본문·결론의 구성이나 문단 구분 등 글의 외형적 구조를 형태에 포함하는 시각(Ellis, 2001; Fatmah et al., 2022)도 함께 존재한다. 이 연구에서는 다음 두 가지 이유에서 후자의 입장을 취한다. 첫째, 이 연구에서 ‘구조’는 글의 외형적 구획뿐 아니라, 논증문으로서 갖추어야 할 최소한의 구성 요소들이 갖추어져 있는지를 점검하는 것까지 포함한다. 이는 개별 내용의 전개 방식이나 설득력을 판단하는 것이 아닌, 논증문이라는 장르가 요구하는 형식적 조건의 충족 여부를 확인하는 것으로 ‘형태’ 범주에 포함하는 것이 보다 적합하다. 둘째, 구조는 AI가 문법, 철자 등 표현 영역과 더불어 타당도 높은 피드백을 제공할 수 있는 영역으로 보고되어 왔으며(Escalante et al., 2023; Wang & Han, 2022), 이를 AI 피드백이 담당하는

론·결론 구조로 이루어지고 논증 구조의 요소들을 갖추고 있는지, 문장 및 출처 표기는 정확한지와 같은 글의 외형을 먼저 정비하고, (2) 다음으로 동료 가 논증이 명확하고 설득력 있게 전달되었는지에 대한 독자 반응을 제공하 여 소통 오류가 발생하는 부분을 점검하게 한 뒤, (3) 최종적으로 교수자가 글의 형태와 내용을 아우르는 심층적이고 통합적인 피드백을 통해 글의 전 반적인 완성도를 제고하도록 유도했다.

〈표 2〉 논증문에 대한 피드백 대상 영역의 구분

대상 영역	설명	구분
1. 주장의 명확성	필자의 주장이 독자에게 명확하고 일관되게 전달되는가	내용
2. 이유의 타당성	주장을 뒷받침하는 이유가 타당하여 설득력이 있는가	내용
3. 근거의 신뢰성	이유를 뒷받침하는 근거가 신뢰할 만한 것인가	내용
4. 반론·재반론의 적절성	예상되는 반론을 인식하고 이에 효과적으로 재반론하였는가	내용
5. 글 구조의 체계성	서론·본론·결론 구조 및 논증 구조의 요소들을 갖추고 있는가	형태
6. 문단 구조의 적절성	문단이 하나의 소주제와 이를 뒷받침하는 문장들로 이루어졌는가	형태
7. 문장 표현의 정확성과 적절성	문장 표현이 문법적으로 정확하고 공식 글쓰기에 적합한가	형태
8. 출처 표기의 적절성	인용한 내용의 출처 표기를 규범에 맞게 하였는가	형태

둘째, 제공된 피드백에 반응하고 수정을 수행하는 전 과정에 대해 학습 자가 자신의 판단을 메타 인지적으로 성찰하게 함으로써 학습자의 저자성 (authorship)이 강화될 수 있도록 조력했다.

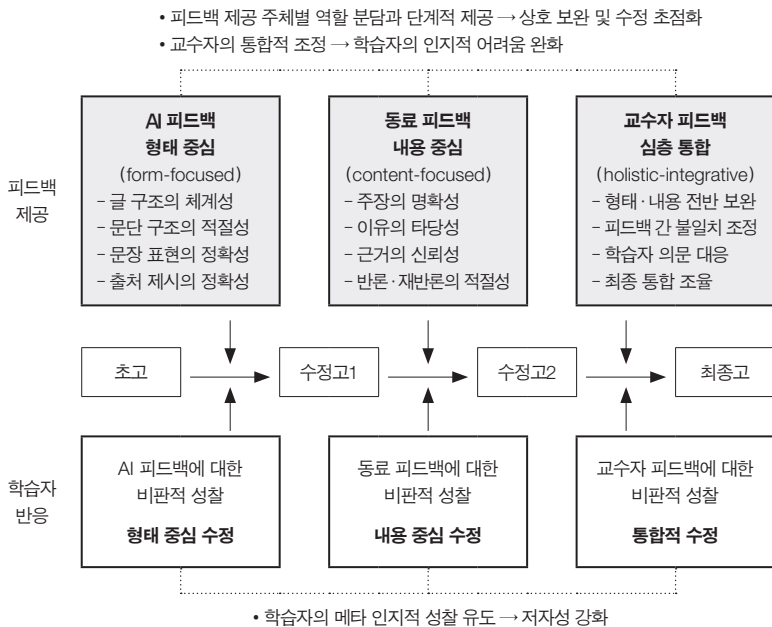
‘형태’에 포함시킬 경우 학습자가 글의 외형적 정비를 마친 후 의미적 소통의 문제에 보다 집중하게 할 수 있는 교육적 이점이 있다.

Carless & Boud(2018)는 학습자에게 적절한 피드백을 제공하는 일 이상으로 중요한 것은 학습자가 피드백을 이해하고 주체적으로 판단하며 행동으로 옮기는 능력인 ‘피드백 리터러시(feedback literacy)’를 함양하게 하는 일이라고 강조한 바 있다. 즉, 피드백이 교육적으로 의미를 갖는 순간은 피드백이 제공되는 때가 아닌, 학습자가 이를 비판적으로 판단하고 주체적으로 반영하며 글에 대한 책임 있는 저자로서의 정체성을 형성해 나가는 때다. 이에 이 연구는 학습자가 모든 수정 단계에서 각 피드백에 대한 자신의 비판적 수용 여부 및 수정 내용을 언어화하여 기록하게 했다. 이 과정에서 학습자가 저자로서 자신의 판단과 목소리(voice)(Ivanič, 1998)를 메타 인지적으로 성찰하며, 나아가 이를 강화할 수 있도록 지원했다.

셋째, 최종 수정 전 교수자가 피드백 수용 및 수정 과정에서 학습자가 갖게 된 의문에 대응함으로써 학습자가 주체적으로 해결하기 어려운 인지적 곤란을 해소하고, 최종 수정의 질을 제고하고자 했다.

피드백을 수용하고 수정하는 과정에서 학습자는 다양한 문제를 경험할 수 있다. 타당도가 높은 피드백이 제공되었지만 그 의미를 충분히 이해하지 못할 수도 있고, 이해는 했으나 실제 수정에서 난관에 부딪힐 수도 있다. 또한 피드백 제공 주체별 강점을 전면화한 구조에서도, 타당도가 낮거나 서로 충돌하는 피드백이 제공될 가능성은 여전히 존재한다. 특히 쓰기 수준이 다양한 동료 3인에 의해 제공되는 동료 피드백의 경우, 동료의 쓰기 수준에 따라 부정확하거나 상충하는 조언이 제시되어 학습자가 판단에 어려움을 겪을 수 있다. 이 연구에서는 학습자가 독자적으로 해결하기 어려운 문제에 대해 수정고에 ‘의문’ 메모를 남기게 하고, 교수자가 피드백 과정에서 이를 확인하여 응답하며 통합적 조율을 수행했다. 이를 통해 학습자의 인지적 어려움을 지원하고, 최종고의 완성도를 보다 높이고자 했다.

이상의 원리를 적용한 단계적 피드백 방안의 구성은 <그림 1>과 같다.



〈그림 1〉 AI·동료·교수자의 단계적 피드백 방안 구성

3. 단계적 피드백 방안의 적용 절차 및 방법

AI·동료·교수자의 단계적 피드백 방안은 총 여섯 단계로 이루어진다. 세 피드백 주체가 담당 영역에 대한 조언을 제공하는 단계(1, 3, 5)와 학습자가 그 수용 여부를 주체적으로 판단하여 수정을 수행하고 성찰 내용을 기록하는 단계(2, 4, 6)가 교차하며 진행된다.

1단계에서는 학습자가 자신이 작성한 초고에 대해 AI의 형태 중심 피드백을 제공받는다. 이는 글의 외형적 완결성을 높이기 위한 점검으로, 〈표 2〉에서 ‘형태’ 측면으로 구분한 네 가지 항목(글 구조의 체계성, 문단 구조의 적절성, 문장 표현의 정확성, 출처 표기의 적절성)을 중심으로 이루어진다.

이때 학습자가 AI에게 입력하는 피드백 요청 프롬프트는 교수자가 자신의 교육 목적에 부합하게 설계하여 제공한다. AI 피드백은 프롬프트의 설

계에 따라 그 내용과 형식이 크게 달라지기에, 교수자가 수업에서 강조한 바람직한 글의 구조와 표현에 대한 기준을 반영하기 위해서다. 교수자는 [프롬프트1]에서 AI의 역할(‘대학생의 논증문을 점검하는 전문적인 글쓰기 튜터’)을 규정하고, 네 가지 피드백 대상 영역에 대해 설명한 후 [프롬프트2]~[프롬프트5]를 통해 각 항목에 대한 구체적인 점검 지침을 제시한다. 또한 [프롬프트6]에서는 글의 형태적 측면에 총평 및 개선 방안을 항목화하여 정리하게 함으로써 학습자의 성찰 및 수정 작업이 효율적으로 이루어지도록 돕는다.

2단계에서는 학습자가 AI 피드백을 비판적으로 검토하여 수용 여부를 판단하고, 이를 반영하여 초고의 형태적 측면에 대한 수정을 수행한다. 이때 워드 프로세서의 ‘메모’ 기능을 사용하여, 피드백이 제공된 원고의 부분에 관련 피드백의 내용과 그에 대한 자신의 판단 및 이유를 간략히 기록한다. 수정을 유보했거나 거부한 부분에도 반드시 메모를 붙인다. 또한 판단하기 어려운 부문에 대해서는 [의문], 피드백과 무관한 자기 주도적 수정이 이루어진 부문에 대해서는 [자기 수정]이라는 말머리를 붙여 관련 내용을 작성한다. 메모 유형 및 작성 예시는 <표 3>과 같다.

<표 3> 학습자의 반응 메모 유형 및 작성 예시

유형	설명	예시
수용	피드백의 타당성을 인정하여 수정에 반영함	(수용/AI) 참고문헌 배열이 잘못되었음 → 필자 이름 순서로 배열하지 않아 수정함
보류	피드백의 타당성은 인정하지만, 수정이 어렵거나 부담되어 유보함	(보류/AI) 이유가 1개밖에 없어 보완 필요함 → 동의하지만 적절한 이유 및 이를 뒷받침할 근거가 떠오르지 않아 보류함
거부	피드백이 타당하지 않다고 판단하여 수정에 반영하지 않음	(거부/AI) 문장이 길어 단문으로 나눌 것 → 현재 문장이 자연스럽고 리듬감 있다고 판단하여 반영하지 않음
의문	피드백에 대해 판단하기 어려운 경우 교수자에게 질문함	(의문/AI) 결론 문단이 부재하므로 추가 필요함 → 예상 반론과 재반론으로 결론 문단을 구성했는데 별도의 문단이 또 있어야 하나요?
자기 수정	피드백이 언급하지 않은 수정을 수행한 경우	(자기 수정) 문장 구조가 어색해서 능동태 문장으로 바꿈

3단계에서는 학습자가 2단계에서 작성한 수정고1(초고에 대한 AI 피드백을 반영하여 수정한 글)의 내용적 측면에 대한 동료 피드백 활동을 수행한다. 동료 피드백은 4인 1조로 구성된 모둠 단위로 시행하며, 각 모둠은 학기 초 실시한 사전 평가 결과를 바탕으로 쓰기 능력 상·중상·중하·하 수준의 학생이 균형 있게 분포하도록 구성한다. 학습자가 쓰기 수준이 낮은 동료로부터 피드백을 받게 될 가능성을 방지하기 위해서다. 학습자는 동료 3인으로부터 피드백을 받고, 자신도 동료 3인의 글에 대해 피드백을 제공한다.

동료 피드백은 <표 2>에서 ‘내용’ 측면으로 구분한 네 가지 항목(주장의 명확성, 이유의 타당성, 근거의 신뢰성, 예상 반론·재반론의 적절성)을 중심으로 이루어진다. 교수자는 사전 교육을 통해 수업 시간에 다룬 논증 구조의 요소에 대해 다시 한번 설명하고, 예시 글에 대한 피드백 시연을 통해 피드백 제공 방법을 숙지하게 한다. 이때 교수자는 피드백의 목적이 오류 지적이 아니라, 학습자가 자신의 글을 객관화하고 성찰할 수 있도록 돕는 ‘독자 반응’임을 명확히 안내한다. 동료 3인은 각기 다른 색상의 필기구를 사용하여 학습자의 원고에 피드백을 작성하고, 원고 하단에는 독자로서 의미 전달이 잘 되지 않았던 부분 및 수정 제안을 항목화하여 정리함으로써 학습자의 성찰과 수정 작업이 보다 효율적으로 이루어지도록 돕는다.

4단계에서는 학습자가 3단계에서 제공된 동료 피드백을 검토한 뒤 수용 여부를 판단하고, 이를 반영하여 수정고1의 내용적 측면을 수정하여 수정고2를 작성한다. 방법 및 절차는 2단계와 동일하다. 먼저 동료 피드백의 주요 제안 사항들에 대한 자신의 입장을 정하고, 원고의 관련 부분에 메모를 달아 각 피드백에 대한 판단과 이유를 간략히 기록한다. 말머리는 피드백 유형([동료1], [동료2], [동료3])과 반응 유형([수용], [보류], [거부], [의문], [자기 수정])을 조합하여 구성한다. 메모 작성 방식은 <표 3>과 동일하나, 두 명 이상의 동료가 동일한 사항을 언급했을 때는 이를 하나의 피드백으로 간주하여 반응한다. 예를 들어, “[수용/동료2, 3] 이 근거가 이유와 무슨 관련이 있는지 이해되지 않음 → 다시 보니 관련성이 부족한 것 같아 근거를 교체

함.”과 같이 기록한다.

5단계에서는 학습자가 4단계에서 작성한 수정고2(수정고1에 대한 동료 피드백을 반영하여 수정한 글)의 심층 통합적 측면에 대한 교수자 피드백을 시행한다.

교수자는 학습자가 제출한 수정고2 파일의 글과 메모를 함께 확인하고, 다음 세 가지 측면을 중심으로 피드백을 제공한다. 첫째, 말머리에 [의문]이 붙은 메모에 답글을 달아 학습자의 혼란을 해소하고 판단을 돕는다. 둘째, AI나 동료가 제안한 피드백의 내용이 부적절했음에도 학습자가 이를 수용하여 글의 질이 저하된 경우, 해당 부분에 대한 보완 의견을 제시한다. 셋째, 글의 하단에 총평을 작성하여 글의 구조와 표현, 논리 전반을 아우르는 통합적인 수정 방향을 제시함으로써 앞선 단계에서 형태적·내용적 측면을 분리하여 피드백을 제공하는 과정에서 누락된 문제를 보완한다. 예를 들어, 글의 구조나 표현, 개별 논증 요소의 적절성에서는 큰 문제가 없지만 전체 글의 논지가 초반과 후반이 미묘하게 어긋나 있는 경우, 교수자는 이에 대한 보완 의견을 제시할 수 있다.

6단계에서는 학습자가 교수자의 피드백을 검토한 뒤 이를 토대로 최종고를 작성한다. 최종고에는 앞선 단계와 마찬가지로 학습자가 주체적으로 수용 여부를 판단하여 반응한 메모²⁾를 첨가한다. 또한 최종고의 본문 아래에는 그동안의 수정 과정에 대한 간단한 성찰을 작성하게 하는 것이 권장된다. 성찰문은 정해진 형식 없이 자유 서술 형태로 작성하며, 학습자가 각 피드백에 대한 자신의 반응과 글의 변화에 대해 총체적으로 성찰해 보게 하는 데 목적이 있다.

-
- 2) 이 단계에서도 학습자는 교수자 피드백과 관련한 의문 사항이 있을 경우 앞 단계에서와 마찬가지로 [의문] 메모를 사용하며, 교수자는 이에 대해 LMS의 과제 답글로 간단한 설명이나 추가 조언을 제공할 수 있다. 다만, 해당 문답의 내용에 대해서는 추가적인 수정을 수행하지 않으며, 학습자의 메타 인지적 성찰을 유도하는 보조적 수단으로 활용한다.

III. 연구 방법

1. 연구 참여자 및 자료 수집 절차

이 연구에는 서울 소재 A대학교에서 2025년 1학기에 교양 글쓰기 강좌를 수강한 대학생 34명이 참여했다. A대학교의 글쓰기 강좌는 2학년 대상 필수 교양 수업으로 운영되고 있어 34명은 전원 2학년 학생이었고, 전공은 다양했다(인문계열 8명, 사회계열 12명, 이공계열 11명, 예술계열 3명). 강좌의 수강 총원은 37명이었으나, 과제 불성실 수행자 및 미제출자 3명을 제외했다. 참여자들에게는 연구 목적과 자료 활용 방식에 대해 사전 안내하고 자발적 참여 동의를 받았다.

학습자들은 1~5주차 동안 글쓰기 과정(계획하기, 자료 찾기 및 출처 제시하기, 구성하기, 초고 쓰기) 및 논증하기에 대한 교육을 받은 상태였다. 단계적 피드백 및 수정 활동은 강좌의 6~10주차에 시행되었으며, II-3에서 제시한 절차와 방법에 따라 [초고 작성 → (1) AI 피드백(수업 중 실습) → (2) 수정1(과제) → (3) 동료 피드백(수업 중 실습) → (4) 수정2(과제) → (5) 교수자 피드백(LMS 제공) → (6) 최종 수정(과제)] 순서로 진행되었다.

먼저, 6주차에는 ‘시사적 이슈에 대한 논증문 쓰기’ 수업의 과제로 논증문 초고를 작성했다. 학습자들은 “우리 사회의 주요한 시사적 이슈 중 찬반 입장이 명확히 대립하는 이슈를 하나 골라, 그에 대한 자신의 주장을 타당하게 논증하는 글을 쓰시오.”라는 과제를 부과받고 일주일 동안 1500자 분량의 논증문을 작성하여 제출했다.

7주차에는 수업 시간에 교수자의 안내에 따라 AI 피드백³⁾을 제공받은

3) 이 연구에서는 모든 학습자가 OpenAI의 GPT-4o를 사용하여 AI 피드백을 제공받도록 했다. 생성형 AI 모델 및 버전에 따라 피드백의 양과 질이 달라지는 편차를 방지하고, 분석

뒤 개별 과제로 피드백 점검 및 수정고1 작성을 수행했다. 과제 부과에 앞서 교수자는 <표 3>을 중심으로 메모 작성 방법 및 사례를 안내했고, 학습자들은 7주차 종료 시까지 글의 구조 및 표현을 중심으로 한 첫 번째 수정을 수행하며 피드백에 대한 반응 및 수정 내역을 메모로 기록했다.

이후, 8주차 수업 시간에는 모듈별로 수정고1에 대한 동료 피드백 활동을 시행하고, 역시 개별 과제로서 피드백을 점검하며 수정고2를 작성했다. 동료 피드백 활동에 앞서 교수자는 논증 요소에 대한 복습 설명 및 예시 글에 대한 피드백 시연을 통해 피드백 활동이 원활히 이루어지도록 노력했다. 학습자들은 글의 의미 전달 문제에 집중하여 두 번째 수정을 수행하며, 관련 사항을 메모했다. 수정고2는 8주차 종료 시까지 제출되었다.

9주차에는 교수자가 학습자의 수정고2에 피드백을 수행했으며, 10주차 시작 시에 강좌의 LMS를 통해 피드백이 첨가된 파일을 제공했다. 학습자는 이를 검토한 뒤 10주차 종료 시까지 최종 수정을 수행했다. 이때 학습자는 최종고의 본문 아래에 전체 피드백 및 수정 과정을 진술하게 되돌아본 간략한 성찰문을 자유롭게 작성했다. 성찰문은 주요 분석 대상에는 포함하지 않고, 질적 검토하여 연구 결과 해석 과정에서 참고 자료로 활용했다.

이상의 과정을 통해, 논증문 초고 34편, 수정고1 34편, 수정고2 34편, 최종고 34편(이상 원고 총 136편) 및 수정고1에 포함된 AI 피드백 384건과 학습자 반응 420건, 수정고2에 포함된 동료 피드백 391건과 학습자 반응 443건, 최종고에 포함된 교수자 피드백 158건과 학습자 반응 176건(이상 피드백 총 933건, 학습자 반응 총 1,039건)⁴⁾이 수집되었다.

의 타당성을 제고하기 위해서였다. 또한 프롬프트는 교수자가 강좌의 교수·학습 내용을 반영하여 설계하였으며, AI의 역할과 임무를 설명하는 [프롬프트1], 네 가지 대상 영역에 대한 점검을 요청하는 [프롬프트2~5], 총평을 요구하는 [프롬프트6]으로 구성되었다. 모든 학습자가 동일한 프롬프트 세트를 사용하도록 하여 요청 방식에 따른 편차를 통제했고, 초고 텍스트는 [프롬프트1] 입력 후 AI가 역할을 이해했음을 확인한 뒤 [프롬프트2]와 함께 pdf 파일로 첨부하도록 안내하여 텍스트 제시 방식에 따른 반응 변동을 최소화했다.

- 4) 학습자 반응 수(1,039건)가 피드백 수(933건)보다 106건 많은 것은 학습자가 피드백에

이때 피드백 및 반응 건수는 학습자가 수정고에 작성한 메모의 말머리(예 [수용/AI], [보류/교수자])를 기준으로 산정했다. 또한 동료 피드백 부문에서 여러 동료가 동일한 조언을 하여 학습자가 이를 하나의 메모로 통합하여 작성한 경우(예 [거부/동료1, 2])에도 피드백 및 반응 건수를 각각 1건으로 산정했다. 이 연구의 목적이 동일한 피드백이 반복적으로 제공된 양상을 살피는 데 있는 것이 아닌, 피드백 내용에 대한 학습자의 반응 양상을 분석하는 데 있기 때문이었다.

2. 자료 분석 절차와 방법

수집된 단계별 원고 136편, AI·동료·교수자의 피드백 933건, 학습자 반응 1,039건은 개인정보 보호를 위해 익명화한 뒤 다음 절차와 방법에 따라 분석⁵⁾했다.

첫째, 단계적 피드백에 대한 학습자의 반응 양상을 분석했다. 먼저, AI·동료·교수자가 제공한 전체 피드백에 대해 학습자가 보인 반응을 <표 3>의 다섯 가지 유형(수용, 보류, 거부, 의문, 자기 수정)을 중심으로 빈도 분석했다. 단계적 피드백을 받고 수정하는 과정에서 나타난 학습자의 전반적인 반응 양상을 점검하기 위해서였다. 다음으로, 각 제공 주체별 피드백에 대한 학습자 반응 양상을 교차 분석하여, 학습자가 AI의 형태 중심 피드백, 동료의 내용 중심 피드백, 교수자의 심층 통합적 피드백 각각에 대해 어떠한 반응을 보였는지 세부적으로 확인했다.

둘째, 각 단계별 학습자의 글의 질이 변화한 양상을 분석했다. 이를 위해, 초고, 수정고1, 수정고2, 최종고 총 136편을 대상으로 평가자 3인⁶⁾이 9점

서 언급하지 않은 수정을 자율적으로 수행한 '자기 수정'이 포함되었기 때문이다.

5) 이 연구의 모든 통계 분석은 SPSS Ver. 29.0을 사용하여 수행했다.

6) 평가자 3인은 모두 대학에서 글쓰기 교육을 담당해 온 교수자로, 평가자1은 16년, 평가자2는 8년, 평가자3은 14년의 교육 경력을 가지고 있다.

척도를 사용⁷⁾하여 분석적 평가를 실시했다. 분석적 평가의 준거는 <표 2>에서 논증문의 피드백 대상 영역으로 구분한 항목 여덟 가지((1) 주장의 명확성, (2) 이유의 타당성, (3) 근거의 신뢰성, (4) 예상 반론·재반론의 적절성, (5) 글 구조의 체계성, (6) 문단 구조의 적절성, (7) 문장 표현의 정확성과 적절성, (8) 출처 표기의 적절성)를 동일하게 사용했다. 피드백이 제공된 영역에서 실제로 글의 질이 상승했는지 여부를 확인하기 위해서였다.

각 단계별 원고에 대한 평가자의 선입견을 배제하기 위해, 136편의 원고는 메모와 피드백을 모두 제거한 채 무작위로 섞여 전달되었다. 평가자들은 본 평가에 앞서 연구자가 주도한 온라인 사전 회의에서 샘플 글 3편에 대한 예비 평가를 수행한 결과를 비교하여 기준을 조율한 뒤 개별적으로 본 평가를 수행했다. 세 평가자 간 신뢰도는 총점 기준 ICC(2, k)=.87로 높은 수준이었다. 평가 영역별 ICC 또한 .74(예상 반론·재반론의 적절성)에서 .89(출처 표기의 적절성) 사이에 분포하여, 모든 영역에서 양호 이상의 신뢰도를 보였다.

단계별 글의 질 변화는 동일한 학습자가 4단계(초고, 수정고1, 수정고2, 최종고)를 거쳐 작성한 원고 점수를 반복 측정했으므로, 총점과 8개 영역별 점수를 각각 종속변수로 설정하여 반복측정 분산분석(Repeated Measures ANOVA)을 실시했다. 단계적 피드백이 글의 전반적인 질과 영역별 질에 미친 효과를 구분해 검토하기 위해서였다. 결과값에 유의한 차이가 나타난 경우에는 Bonferroni 방법을 사용하여 사후 검정을 수행했다. 또한 반복측정 분산분석의 가정 검토 과정에서 Mauchly의 구형성 가정이 충족된 것으로 확인되었으며($p>.05$), 이에 따라 구형성을 가정한 분석 결과를 보고했다. 반복측정 요인은 4수준($k=4$)으로 요인 자유도는 $k-1=3$ 이며, 참여자 수 34명

7) 9점 척도는 (1) 대학 글쓰기 평가에서 교수자들이 상·중·하의 세 수준을 다시 '+, 0, -'로 세분화하여 평가하는 방식(A+: 9, A0: 8, A-: 7, B+: 6, B0: 5, B-: 4, C+: 3, C0: 2, C-: 1)과 부합하며, (2) 5점 또는 7점 척도에 비해 글의 질 변화를 보다 미세하게 포착할 수 있다는 점에서 이 연구의 분석 목적에 적합하다고 판단하여 채택했다.

을 고려할 때 잔차 자유도는 $(k-1)(N-1)=3 \times 33=99$ 로 산출되었다.

IV. 연구 결과

1. 단계적 피드백에 대한 학습자의 반응 양상

1) 단계적 피드백에 대한 학습자의 전반적인 반응 양상

대학생 학습자 34명이 AI·동료·교수자의 피드백을 받고 수정을 수행하는 전체 단계에서 보인 반응 수는 총 1,039건(1인당 30.6건)으로, 유형별 분포는 <표 4>와 같이 나타났다.⁸⁾

<표 4> 수정 과정에서 나타난 학습자의 반응 유형 분포

반응 유형	전체 빈도(1인당 빈도)	백분율(%)
수용	627(18.4)	60.3
보류	97(2.9)	9.3
거부	119(3.5)	11.5
의문	90(2.6)	8.7
자기 수정	106(3.1)	10.2
합계	1,039(30.6)	100.0

먼저 주목되는 사항은 학습자가 피드백을 타당하다고 생각하고 실제 수

8) 이 절에서는 단계적 피드백에 대한 학습자의 반응 양상을 탐색적으로 파악하는 것을 목적으로 하였기에, 빈도와 백분율에 기반한 기술통계적 서술만을 수행하고 유형 및 단계별 분포 차이에 대한 통계적 검증은 별도로 실시하지 않았다.

정에 반영한 ‘수용’의 비율이 전체의 60.3%로 상당히 높게 나타났다. 학습자가 피드백의 타당성을 인정하더라도 그것이 언제나 수정으로 연결되는 것은 아니다. 피드백의 타당성 여부와는 별개로 수정 지침의 구체성이 부족하거나 수정의 난이도가 높아 부담이 될 경우, 학습자가 수정을 포기하는 일이 자주 발생한다(Silver & Lee, 2007). 그러나 이 연구에서는 실제로 수정을 수행한 ‘수용’의 비율은 높고(60.3%), 타당성을 인정하지만 수정으로 이행하지는 않은 ‘보류’의 비율은 상대적으로 낮게(9.3%) 나타났다.

이는 단계적 피드백 방안이 학습자의 인지적 부담을 경감하여 실제 수정으로의 이행률을 높이는 데 효과적으로 작용했을 가능성을 시사한다. 즉, AI·동료·교수자가 영역을 분담하여 순차적으로 피드백을 제공한 구조로 인해, 학습자가 인지적 과부하 없이 한 번에 하나의 영역에 집중하여 피드백을 해석하고 수정을 수행할 수 있었을 것으로 해석된다.

한편, ‘거부(11.5%)’와 ‘의문(8.7%)’의 비율은 그 자체로는 높지 않지만, 학습자가 피드백을 신중하게 평가하고 저자로서 책임 있는 판단을 내리고자 하는 태도를 드러낸다는 점에서 의미를 갖는다. ‘거부’는 학습자가 자신의 수사적 목적이나 글의 맥락을 고려해 피드백을 비판적으로 평가하는 반응이며, ‘의문’은 피드백의 타당성에 대해 즉시 평가하기 어려운 경우 교수자의 의견을 추가적으로 검토하여 판단하려는 반응이다. 이들은 모두 학습자가 피드백을 단순히 수용하는 것에서 나아가, 해석과 판단의 주체로서 능동적으로 반응하고 있음을 시사한다.

마지막으로, 피드백에서 언급하지 않은 문제에 대해 학습자 스스로 ‘자기 수정’을 수행한 비율이 10.2%에 이르렀다는 사실도 주목할 만하다. 제공된 피드백에 대한 반응과 달리, ‘자기 수정’은 의무적으로 부과된 것이 아니었는데도 학습자들은 1인당 평균 3.1건의 수정을 추가적으로 수행했다. 특히, 이 ‘자기 수정’은 피드백 단계별로 그 동기와 양상이 다르게 나타나, 단계적 피드백 방안이 학습자의 글 전반에 대한 재고를 유도하는 촉매(觸媒)로 작용했을 가능성을 보여준다. 이에 대해서는 2)항에서 보다 구체적으로 고

찰한다.

2) 피드백 제공 주체별 학습자의 반응 양상

대학생 학습자 34명이 (1) AI 피드백 제공 이후 수정고1 작성 단계, (2) 동료 피드백 제공 이후 수정고2 작성 단계, (3) 교수자 피드백 제공 이후 최종고 작성 단계에서 보인 반응 유형별 수(괄호 안: %)는 <표 5>와 같았다.

<표 5> 피드백 제공 단계별 학습자의 반응 유형 분포(괄호 안: %)

반응 유형 \ 단계	수정고1 (AI 이후)	수정고2 (동료 이후)	최종고 (교수자 이후)	합계(%)
수용	285(67.9)	253(57.1)	89(50.6)	627(60.3)
보류	28(6.7)	23(5.2)	46(26.1)	97(9.3)
거부	30(7.1)	76(17.2)	13(7.4)	119(11.5)
의문	41(9.8)	39(8.8)	10(5.7)	90(8.7)
자기 수정	36(8.6)	52(11.7)	18(10.2)	106(10.2)
합계(%)	420(100.0)	443(100.0)	176(100.0)	1,039(100.0)

먼저, AI 피드백 제공 이후 수정고1 작성 단계에서는 [수용(67.9%) > 의문(9.8%) > 자기 수정(8.6%) > 거부(7.1%) > 보류(6.7%)] 순서로 반응이 나타났다. 이중 ‘수용’과 ‘의문’의 비율은 다른 단계와 비교할 때 가장 높았고, ‘거부’와 ‘자기 수정’의 비율은 가장 낮았다.

AI 피드백이 수용률은 가장 높고 거부율이 가장 낮게 나타난 이유는 해당 피드백이 그 타당성을 상대적으로 명확히 판단할 수 있는 영역에 제공되었기 때문으로 해석될 수 있다. 즉, ‘본론의 문단 수가 분량에 비해 부족하다’거나 ‘문단에 소주제문이 존재하지 않는다’와 같은 형식 구조에 대한 피드백이나, ‘주어와 서술어의 호응이 맞지 않는다’거나 ‘인용 표현의 출처 표기에 오류가 있다’는 표현 관련 피드백은 해석의 다양성이 존재할 여

지가 적다. 그래서 학습자들은 수용 이유를 진술한 메모에서도 긴 설명 없이 ‘맞음’, ‘인정함’과 같은 간략한 표현과 함께 즉시 수정을 수행하는 양상을 보였다.

한편, 피드백에 대해 교수자에게 질문하는 ‘의문’의 비율이 AI 피드백에서 가장 높게 나타난 점은 흥미롭다. ‘의문’은 상호 충돌하는 조언이 발생하는 동료 피드백에 대한 반응으로 가장 빈번히 나타날 것으로 예상되었기 때문이다. 그러나 학습자들은 예컨대 “AI가 제 글에 예상 반론이 존재하지 않는다고 했는데, 2문단 끝부분의 내용은 예상 반론이라고 볼 수 없는 건가요 (학습자15)?”와 같이, 진위 여부가 분명하지 않은 사항에 대해서는 적극적으로 의문을 제기했다.

이러한 현상은 AI 피드백이 기계적·추상적 조언에 머무른 경우에 주로 발생한 것으로 추정된다. 선행 연구(이윤빈, 2025; Bogolepova, 2025; Yoon et al., 2023)에서 언급한 해당 한계를 최소화하기 위해, 피드백 요청 프롬프트에는 ‘글의 내용을 구체적으로 언급하며 피드백하라.’는 지시가 포함되어 있었다. 그러나 AI 피드백은 ‘글에 존재하는 내용에 문제가 있을 경우’에는 이를 구체적으로 언급했지만, 학습자가 의도 전달에 실패하여 ‘글에 있어야 할 내용이 부재하는 것으로 보이는 경우’에는 일반론적인 조언만을 제공했다. 교수자 피드백이 학습자의 의도를 맥락적으로 추론한 조언을 제공한 것과는 대조적⁹⁾이다. 이에 학습자들은 AI가 제시한 일반론적인 조언에 대해,

9) 예를 들어, 학습자15는 ‘대학의 전공 자율제도는 폐지되어야 한다.’는 주장을 논증하면서, 2문단 후반에 ‘진로를 스스로 결정하지 못한 학생도 있지만, 이는 임시 단계에서 고민해야 할 문제라고 본다.’라는 문장을 삽입했다. 학습자15는 이를 예상 반론으로 언급한 것이었으나 이는 적합한 반론이 아니었기에 AI는 해당 글을 예상 반론이 부재하는 글로 취급하여 일반론적인 피드백(“예상 반론에 대한 언급이 없어 글의 균형성이 부족해 보입니다. 독자의 반대 입장을 미리 제시하고 이에 대한 반박을 덧붙이면 주장의 설득력이 강화될 수 있습니다.”)만을 제공했다. 반면, 교수자는 학습자의 의도를 맥락적으로 파악하여, “이 문장은 개인의 책임을 지직한 것으로, 전공 자율제도 폐지 자체에 대한 반론으로 보기 어렵습니다. 예컨대 ‘청소년기에 충분한 진로 탐색을 하지 못한 학생들에게 대학이 진

자신의 구체적인 의도를 부각하여 교수자의 판단을 요청하는 방식으로 반응했다.

‘자기 수정’은 비율상으로는 다른 단계들에 비해 낮았으나, AI가 언급한 형태적 측면을 수정하는 과정에서 학습자가 판단하여 내용적 수정을 수행하는 연쇄 현상(ripple effect)이 나타났다는 점에서 주목된다. 예컨대 학습자 33은 AI 피드백을 반영하여 3문단의 소주제문 위치를 두괄식으로 조정한 이후, 앞뒤 문단과의 논리적 연결이 자연스럽게 돌아갔다는 판단하에 추가적으로 2, 4문단의 내용을 수정·보완했다. 이는 AI 피드백이 글의 형태적 측면만을 개선하는 데 그치지 않고, 학습자의 자발적인 점검을 유도하여 내용적 측면에 대한 개선을 촉발했을 가능성¹⁰⁾을 보여준다.

다음으로, 동료 피드백 제공 이후 수정고2 작성 단계에서는 [수용(57.1%) > 거부(17.2%) > 자기 수정(11.7%) > 의문(8.8%) > 보류(5.2%)] 순서의 반응이 나타났다. 이 중 ‘거부’와 ‘자기 수정’의 비율은 다른 단계와 비교할 때 가장 높았고, ‘보류’의 비율은 가장 낮았다.

이 단계에서는 학습자가 자신의 글에 대해 저자로서 책임 의식을 가지고 주체적으로 반응하는 이른 바 저자성(authorship)이 가장 뚜렷하게 발현되는 경향성이 나타났다. 이는 동료 피드백이 ‘이유의 설득력이 떨어진다’든지 ‘예상 반론에 대한 재반론이 적절하지 않다’와 같이 다양한 해석이 존재할 수 있는 내용 영역에 제공되었다는 점, 그리고 동료 3인의 피드백이 불일치하는 경우가 발생하여 학습자의 주체적인 판단이 요구되었다는 점에 기인하는 것으로 보인다.

학습자의 저자성은 다음 두 측면에서 특히 두드러지게 확인되었다. 첫째, 제공된 피드백에 대한 거부율이 가장 높은 데 그치지 않고, 거부의 이유

로 탐색 기회를 보장해야 한다.’와 같이 제도 자체의 필요성을 부각해주어야 합니다.”라는 피드백을 통해 왜 해당 글에 예상 반론이 실제로 부재하는가에 대한 구체적인 설명을 제공했다.

10) 실제 내용적 측면에서의 개선이 일어났는지에 대해서는 다음 절에서 논의한다.

를 적극적·비판적으로 제시하는 경향이 나타났다. 학습자들은 동료 피드백의 타당성이 부족하다고 판단했을 경우, 그것이 왜 문제이고 자신이 원래 작성한 내용이 왜 적절한가에 대해 상당한 분량을 할애하여 논쟁적으로 기술하는 경향을 보였다. 이는 동료 피드백 상황에서는 학습자가 자신의 원래 주장을 방어하거나 피드백의 타당성을 적극적으로 반박하는 방어적 저자성(defensive authorship)이 강하게 나타난다고 주장한 선행 연구의 결과(Nelson & Carson, 1998; Zacharias, 2007)와도 일치한다.¹¹⁾

둘째, 가장 높은 비율로 나타난 ‘자기 수정’ 중 다수가 학습자가 거부한 내용 관련 피드백과 관련된 표현 중심의 수정이었다는 점도 주목된다. 특히 학습자들은 2명 이상의 동료가 제공한 피드백을 거부했을 경우, 왜 자신의 의도를 복수(複數)의 독자들이 제대로 인지하지 못했는가를 점검하여 해당 부문의 표현을 명료하게 다듬는 양상을 보였다. 예를 들어, 학습자09는 2명의 동료가 ‘이유의 설득력이 부족해 교체가 필요하다’고 한 피드백을 적극적으로 반박하며 거부했다. 그러나 복수의 독자가 동일하게 오독(誤讀)했다면 자신의 표현 방식에 문제가 있을 수 있다고 판단하여, 이유를 교체하는 대신 해당 문장의 표현을 보다 이해하기 쉽게 수정했다. 이는 동료 피드백을 단순히 거부하는 데서 나아가, 오독의 원인을 성찰하고 조정하려는 책임 있는 저자성(responsible authorship)이 발현된 것으로 해석할 수 있다(Ivanič, 1998).

마지막으로, 교수자 피드백 제공 후 최종고 작성 단계의 반응은 [수용(50.6%) > 보류(26.1%) > 자기 수정(10.2%) > 거부(7.4%) > 의문(5.7%)] 순

11) 이러한 현상은 선행 연구들에서 반복적으로 언급한 것처럼 학습자와 동료가 전문성 면에서 동등한 위치에 있기에 발생하는 것으로 해석된다. 예컨대 학습자29는 성찰문에서 자신이 동료 피드백에 대해 적극적으로 반박한 이유에 대해 “교수님은 교수님이어서 맞으시겠지 싶고, AI는 인공지능이 나보다야 낫겠지 싶어 (피드백 내용에 대해 비판하기가-인용자) 망설여졌지만, 동료들은 나와 같은 수준에 있으니 좀더 자신 있게 내가 원래 했던 생각이 더 옳다고 반박할 수 있었다.”고 기술했다.

서였다. 이중 ‘보류’의 비율이 다른 단계 비율의 3배 이상으로 높았고, ‘수용’과 ‘의문’의 비율은 가장 낮았다.

‘보류’는 피드백의 타당성은 인정하지만 수정이 어렵거나 부담되어 수용하지 않는 것이다. 이러한 ‘보류’가 교수자 피드백에서 가장 높게 나타난 첫 번째 이유는 글의 형태적·내용적 측면을 아우르며 제시된 심층 통합적 피드백의 수정 난이도가 앞 단계 피드백에 비해 높은 경우가 많았기 때문으로 보인다. 특히 보류는 글의 논리에는 큰 문제가 없지만 과제가 요구한 것과 초점이 다른 글을 썼거나, 서론 및 결론에서 제시한 주장과 본문에서 논증한 내용이 미묘하게 어긋나 글의 전면 재구성이 요구된 경우¹²⁾에 발생 빈도가 높았다. 이러한 문제는 글의 형태적·내용적 측면 중 하나에 중점을 둔 앞 단계의 피드백에서는 쉽게 언급되지 않기에, 학습자 입장에서는 “타당성은 인정하지만, AI·동료 피드백을 받으며 이미 여러 번 수정을 했는데 최종 고 단계에서 교수님께 대공사를 해야 하는 피드백을 받아 난감하다(학습자 11).”며 수정을 포기하는 경우가 발생했다. 이는 글의 전반적인 방향성을 점검하는 교수자 피드백이 최종 단계에서 제공되는 단계적 피드백 방안의 구조적 한계로, 이를 보완하는 방안¹³⁾이 필요할 것으로 판단되었다.

또한 ‘보류’의 비율이 가장 높게 나타난 두 번째 이유는 ‘의문’의 비율이 가장 낮게 나타난 이유와 연동되어 있을 것으로 이해된다. 일반적으로 교수자 피드백은 학습자와 교수자의 눈높이가 달라 학습자가 의문을 가질 여지가 많은 것으로 여겨진다(Zacharias, 2007). 그러나 동시에 교수자가 갖는 압

12) 예를 들어, 학습자05는 ‘찬반 입장이 명확히 대립하는 이슈에 대한 논증문을 쓰라’는 과제 요구 사항과 달리, “장기기증에 대한 긍정적 인식을 확산시킬 필요가 있다.”는 주장을 설정하여, 예상 반론 및 재반론 없이 논증을 진행했다. 또한 학습자27이 작성한 글의 경우, 서론의 문제 제기과 결론에서 강조한 주장은 ‘고교학점제는 폐지되어야 한다’인데, 본문에서는 현행 고교학점제의 개선 방안에 중점을 둔 논의를 전개했다. 이들의 글은 과제 요구에 부합하기 위해 주장을 재설정하거나 주장과 부합하도록 논의를 재구성하는 전면 수정이 필요한 경우였으나, AI와 동료 피드백 단계에서는 해당 문제점이 언급되지 않았다.

13) 한계를 보완할 수 있는 구체적인 방안에 대해서는 결론에서 논의하고자 한다.

목적 권위로 인해 학습자가 AI 피드백의 경우에서처럼 즉각적으로 의문을 제기하기에 심리적 부담을 느꼈을 가능성도 존재¹⁴⁾한다. 즉, 학습자가 피드백을 제공한 주체인 교수자에게 직접적으로 의문을 제기하는 대신, 타당성은 인정하지만 수정은 유보하는 ‘보류’로 반응을 우회했을 가능성을 배제할 수 없다.

한편, 이 단계에서의 ‘자기 수정’이 학습자가 ‘보류’한 피드백에 대한 이차적 반응으로 수행된 경우가 빈번하다는 점도 흥미롭다. 교수자가 제안한 수준의 수정까지는 수행하지 못했으나, 글의 문제점을 인지하고 학습자의 수준에서 수행할 수 있는 수정을 시도한 사례가 자주 확인되었다. 예컨대 학습자11은 주장을 뒷받침하는 두 가지 이유를 재설정해야 한다는 교수자의 피드백은 적합한 이유가 떠오르지 않는다며 보류했으나, 현재 제시한 이유의 타당성을 높이기 위해 설명을 상세화하고 근거를 보완하는 자기 수정을 수행했다. 이는 교수자의 피드백이 본래의 목적을 온전히 달성하지 못한 경우에도, 학습자로 하여금 글의 문제를 인지하여 가능한 범위에서 대안을 탐색하고 주체적 수정을 시도하게 하는 촉발 요인으로 작용했음을 보여준다.

이처럼 학습자들은 단계적으로 제공된 AI·동료·교수자의 피드백을 전반적으로는 높은 비율로 ‘수용’하는 가운데, 각 피드백의 특성에 따라 다양한 반응을 보였다. 상대적으로 명확한 판단이 가능했던 AI 피드백에 대해서는 높은 수용률을 보임과 동시에 AI의 기계적·추상적 진술에 대해 적극적으로 의문을 제기했고, 해석의 다양성과 충돌이 나타난 동료 피드백에 대해서

14) 이러한 심리적 부담의 가능성은 ‘거부(7.4%: 13건)’의 경우에서도 확인되었다. 학생들은 교수자 피드백을 거부할 때 다른 주체의 피드백에서처럼 적극적·비판적으로 이유를 서술하기보다 “나는 지금이 최선인 것 같아서...(학습자16)”와 같이 소극적이고 모호하게 표현하는 경향을 보였다. 또한 반론을 제시하는 경우에도 “저는 3단락에 넣었다고 생각하는데 교수님이 못 보신 것 같아요.(학습자02)”와 같이 최대한 예의를 갖추어 간접적으로 의견을 밝히는 모습을 보였다.

는 방어적이면서도 책임 있는 저자성을 강하게 드러냈다. 교수자 피드백에 대해서는 수정 난이도와 부담으로 인한 보류가 많았으나, AI·동료 피드백과 마찬가지로 해당 피드백을 계기로 주체적인 자기 수정을 시도하는 양상이 나타났다. 다음 절에서는 이러한 반응을 통해 학습자의 글에 어떠한 질적 변화가 나타났는지 확인한다.

2. 피드백 제공 단계별 글의 질 변화 양상

1) 단계별 글의 전반적 질 변화 양상

대학생 학습자 34명이 작성한 (1) 초고, (2) AI 피드백 후 작성한 수정고1, (3) 동료 피드백 후 작성한 수정고2, (4) 교수자 피드백 후 작성한 최종고 각 34편의 전반적인 질의 변화 양상을 확인하기 위해, 분석적 평가를 시행한 8개 영역의 평균 점수를 사용하여 반복측정 분산분석(RM ANOVA)을 시행한 결과는 <표 6>과 같았다.

<표 6> 단계별 글의 전반적 질 변화에 대한 분산분석 결과

단계	평균(M)	표준편차(SD)	Bonferroni 사후검정(p)
초고	5.54	0.92	-
수정고1	6.52	1.10	초고<수정고1 (.012)**
수정고2	6.87	0.97	수정고1<수정고2 (.048)*
최종고	7.49	0.79	수정고2<최종고 (<.001)***

$F(3, 99) = 4.89, p = .004, \eta^2 = .129$

* $p < .05$, ** $p < .01$, *** $p < .001$

먼저, 각 단계에서 글의 질 평균은 [초고(5.54) < 수정고1(6.52) < 수정고2(6.87) < 최종고(7.49)]의 순서로 나타났으며, 모든 단계에서 점진적인 향상 경향을 보였다. 반복측정 분산분석 결과 단계 간 평균 차이는 통계적으

로 유의했고 중간 수준의 효과 크기($\eta^2 = .129$)가 확인되었다($F(3, 99)=4.89$, $p=.004$). 이는 단계가 진행됨에 따라 학습자의 글의 질이 전반적으로 상승하는 양상이 관찰되었음을 의미한다. 특히 초고 대비 최종고의 질은 9점 척도 기준 5.54(B-~B0 수준)에서 7.49(A-~A0 수준)으로 1.95점 향상되어, 단계적 피드백 방안을 통해 학습자가 작성한 글이 전반적인 질적 향상을 보인 것으로 해석되었다.

다음으로, 글의 질이 가장 많이 향상된 단계는 수정고1(6.52)로, 초고(5.54) 대비 0.98점 상승했다($p=.012$). 이는 앞 절에서 확인한 바, AI 피드백에 대한 수용률이 가장 높아(67.9%) 수정이 가장 많이 수행된 단계로, 이 단계에서 글의 질 또한 크게 개선되는 양상이 확인되었다. 또한 교수자 피드백의 경우, 수용률은 다른 단계 대비 가장 낮았지만(50.6%) 글의 질은 수정고2(6.87) 대비 최종고(7.49)에서 0.62점 상승하여, 수용된 피드백이 글의 질을 실질적으로 향상시키는 역할을 했음을 시사했다($p<.001$). 동료 피드백 이후 작성된 수정고2(6.87)는 수정고1(6.52) 대비 0.35점 상승하여 다른 단계에 비하여 다소 낮은 편이었으나($p=.048$) 세부 영역에서는 일부 항목에서 의미 있는 개선이 이루어졌다. 단계별 원고가 피드백이 제공된 세부 영역에서 어떠한 구체적인 변화를 보였는지에 대해서는 이어지는 2)항에서 확인한다.

2) 단계별 글의 피드백 제공 영역별 질 변화 양상

AI·동료·교수자가 역할을 분담하여 피드백을 제공한 8가지 세부 영역에서 초고, 수정고1, 수정고2, 최종고의 질이 변화한 양상을 확인하기 위해 영역별 반복측정 분산분석(RM ANOVA)을 시행했다. <표 7>에서 확인할 수 있는 주요 결과를 정리하면 다음과 같다.

〈표 7〉 단계별 글의 피드백 제공 영역별 질 변화에 대한 분산분석 결과

영역	초M (SD)	수1M (SD)	수2M (SD)	최M (SD)	F	p	η^2	Bonferroni (p)
1. 주장	5.40 (1.23)	6.14 (1.34)	6.69 (1.19)	7.35 (0.81)	3.12	.029	0.082	초<수1(.031)*, 수1<수2(ns), 수2<최(.006)**
2. 이유	5.28 (1.15)	5.60 (2.04)	6.36 (1.78)	7.67 (0.90)	7.83	<.001	0.201	초<수1(.002)**, 수1<수2(.01)*, 수2<최(<.001)***
3. 근거	5.14 (0.74)	5.71 (1.01)	6.84 (0.83)	7.18 (0.67)	6.51	.003	0.153	초<수1(.044)*, 수1<수2(.021)*, 수2<최(ns)
4. 반론	4.75 (1.08)	5.05 (0.76)	5.47 (0.54)	6.38 (1.02)	2.48	.064	0.047	-
5. 글	5.78 (1.44)	7.15 (0.93)	7.72 (1.21)	7.94 (0.54)	11.22	<.001	0.238	초<수1(<.001)***, 수1<수2(ns), 수2<최(ns)
6. 문단	5.44 (0.98)	7.56 (1.38)	7.28 (0.95)	7.78 (0.68)	4.90	.012	0.109	초<수1(<.001)***, 수1<수2(.043)*, 수2<최(.031)*
7. 문장	5.66 (1.37)	6.59 (1.12)	6.23 (0.88)	7.05 (0.71)	5.66	.006	0.127	초<수1(.006)**, 수1<수2(ns), 수2<최(.031)*
8. 출처	6.88 (0.58)	8.32 (0.73)	8.35 (0.69)	8.56 (0.34)	1.92	.101	0.039	-

df=(3, 99), * $p < .05$, ** $p < .01$, *** $p < .001$

첫째, AI 피드백 제공 이후 작성된 수정고1 단계에서는 글의 형태적 측면을 중심으로 큰 향상이 나타났고, 이후 단계에서는 형태적 변화가 상대적으로 적었다. 가장 큰 변화가 나타난 것은 ‘문단 구조의 적절성’ 영역(7.56)으로, 초고(5.44) 대비 2.12점이 상승했고($p<.001$), 이후 ‘출처 표기의 적절성(+1.44)’, ‘글 구조의 체계성(+1.37, $p<.001$)’, ‘문장 표현의 정확성과 적절성(+0.93, $p<.006$)’ 순으로 모두 높은 상승률을 보였다. ‘출처 표기의 적절성’

은 이미 이 단계의 점수가 8.32로 높게 상승하여, 이후 단계에서는 변화가 미미했으며(수정고2: +0.03, 최종고: +0.21), 나머지 세 영역에서도 수정고1 대비 최종고의 점수 변화가 크지 않았다(글 구조: +0.79, 문단 구조: +0.22, 문장 표현: +0.46). 이는 AI 피드백을 초기에 배치해 학습자들이 글의 외형을 먼저 정비하게 한 전략이 목적인 효과를 발생시켰음을 시사한다. 즉, 초기 수정을 통해 정비한 글의 형태가 이후에도 비교적 안정적으로 유지되며, 학습자가 내용 중심의 수정에 집중할 수 있는 기반이 마련되었다고 해석할 수 있다.

또한 수정고1 단계에서는 피드백이 제공된 형태적 측면뿐 아니라 내용적 측면에서도 일정한 향상이 일어났음이 확인되었다. 특히, ‘주장의 명확성’ 영역에서는 전(全) 단계에서 가장 높은 상승(+0.74, $p<.031$)이 나타났고, ‘근거의 신뢰성’ 영역에서도 수정고2 다음으로 높은 상승(+0.57, $p<.044$)을 보였다. 이때 ‘주장의 명확성’은 AI 피드백을 수용한 학습자가 전체 글과 문단의 구조를 정비하기 위해 주제문(주장)과 문단의 소주제문을 재배치하는 과정에서 개선된 것으로 관찰된다. 또한 ‘근거의 신뢰성’은 근거가 부재한 글에 대해 AI가 해당 사실을 지적하여 이를 보완함으로써 상승되었다. 요컨대 글의 형태적 측면에 대한 수정이 연쇄적으로 내용적 측면에서의 개선을 함께 발생시킨 것으로 해석된다.

둘째, 동료 피드백 제공 이후 작성된 수정고2 단계에서는 글의 내용적 측면을 중심으로 고른 향상이 나타났고, 특히 ‘근거의 신뢰성’(+1.13, $p<.021$)과 ‘이유의 타당성’(+0.76, $p<.01$) 영역에서 통계적으로 유의미한 상승이 확인되었다. 글의 내용에 대한 독자 반응의 일환으로 제공된 동료 피드백이 이유 및 근거에 대해 많이 언급했고, 학습자가 이에 반응하여 글을 수정하는 과정에서 두 영역의 완성도가 높아진 것으로 보인다.

흥미로운 사실은 학습자가 동료 피드백을 ‘수용’하여 수정한 경우에서만 아니라, 피드백을 ‘거부’하고 ‘자기 수정’을 수행한 경우에도 해당 영역에서의 점수 상승이 명확히 고찰되었다는 것이다. 예를 들어, 학습자05는 ‘제시된 이유 두 가지가 모두 설득력이 부족하다’는 동료 피드백에 대해 ‘두 이

유 모두 충분히 타당하다’며 거부했으나, 그 대신 각 이유들을 뒷받침하는 근거들을 전체적으로 보완하는 ‘자기 수정’을 수행했다. 그리고 그 결과 ‘근거의 신뢰성’ 점수가 수정고1 대비 3점 상승(4.67 → 7.67)했다. 이는 동료 피드백에 대한 학습자의 비판적 점검과 주체적인 저자로서의 개입(authorial agency)이 글의 질을 개선한 사례라고 할 수 있다.

한편, 수정고2 단계에서는 ‘문단 구조의 체계성(-0.28, $p<.043$)’과 ‘문장 표현의 정확성과 적절성(-0.36)’ 영역에서 수정고1 대비 점수가 소폭 하락하는 양상이 나타났다. 학습자가 앞 단계에서 정비한 글에는 없었던 새로운 이유나 근거를 추가하거나 내용을 수정하는 과정에서 새로운 문제가 발생한 것이다. 그러나 이 두 영역은 모두 교수자 피드백 이후 작성된 최종고에서 개선되어(문단: +0.50, 문장: +0.82), 수정고1보다도 점수가 소폭 상승했다. 예컨대 학습자30의 수정고2에서는 기존 문단에 있던 이유를 교체하는 과정에서 문단이 지나치게 길어지고 비문(非文)도 포함되어 두 영역 모두에서 점수가 하락하는 문제가 발생했다. 그러나 교수자 피드백을 통해 학습자가 해당 문제를 인지하여 이를 수정하면서, 결과적으로 ‘이유의 타당성’뿐 아니라, ‘문단 구조의 체계성’과 ‘문장 표현의 정확성과 적절성’이 모두 개선된 최종고를 작성했다. 이는 최종 단계에 배치된 교수자 피드백이 글의 질을 통합적으로 조율하는 역할을 효과적으로 수행했음을 보여준다.

셋째, 교수자 피드백 제공 이후 작성된 최종고 단계에서는 ‘이유의 타당성’과 ‘반론·재반론의 적절성’ 영역에서의 상승이 주목되었다. 이중 ‘이유의 타당성’은 8개 영역 중 초고 대비 최종고가 2.39점 높아져 가장 큰 개선을 보인 영역으로, 주체별 피드백의 누적 효과가 뚜렷하게 확인되었다. 즉, AI 피드백 후 글의 형태를 정비하는 과정에서 해당 영역의 점수가 소폭 상승(+0.32, $p<.002$)했고, 동료 피드백을 통한 초점화된 점검으로 본격적인 개선(+0.76, $p<.01$)을 보였다. 또한 최종적으로 교수자가 앞 단계에서 학습자가 제기한 의문을 해소하고 기존 피드백이 하지 못한 심층적인 논리에 대한 조언을 제공함으로써 가장 큰 향상(+1.31, $p<.001$)이 나타났다.

‘반론·재반론의 적절성’ 또한 ‘이유의 타당성’과 마찬가지로 학습자들이 가장 개선하기 어려워하는 영역이다. AI나 동료가 타당한 피드백을 제공하는 데 한계가 있는 것으로 알려져 있고(Bogolepova, 2025) 학습자들이 동료 피드백과 관련하여 가장 빈번히 ‘의문’을 제기한 영역이기도 하다. 이 연구에서도 수정고1, 수정고2 단계에서는 유의미한 상승이 발생하지 않았으나, 전문성을 지닌 교수자의 심층적 피드백이 제공된 최종고에서는 앞선 두 단계 대비 큰 폭의 상승(+0.91)이 나타났다.

이러한 현상은 앞서 살펴본 바, 수정고2 작성 과정에서 새롭게 발생한 문제(문단 구조 및 문장 표현 영역의 일시적 저하)가 교수자 피드백 이후 해결된 것과 더불어, 최종 단계에 배치된 교수자 피드백이 효과적으로 기능하고 있음을 시사한다. 즉, 교수자 피드백은 AI·동료 피드백의 한계를 보완하고, 학습자의 의문을 해소하며, 단계적 수정 과정에서 새롭게 발생하는 문제를 조정하는 통합적인 중재자로서의 역할을 수행함으로써 최종고의 질적 완성도를 제고했다.

V. 결론

이 연구는 AI·동료·교수자가 역할을 분담한 단계적 피드백 방안을 제안하고, 이를 대학 글쓰기 교육 현장에 적용하여 그 효과를 탐색했다. 각기 다른 강점과 한계를 지닌 세 주체의 피드백을 상호 보완적으로 활용하기 위해 (1) AI는 글의 구조와 표현을 포함하는 형태적 측면에, (2) 동료는 독자 반응의 일환으로 내용적 측면에, (3) 교수자는 심층 통합적 측면에 집중하여 순차적으로 피드백을 제공하고, 학습자가 이를 비판적으로 점검하며 세 차례의 논증문 수정을 수행하게 했다. 그리고 주체별 피드백에 대한 학습자의 반응 및 단계별 글의 질 변화 양상을 고찰했다.

이 연구를 통해 확인된 주요 결과는 다음과 같았다. 첫째, AI·동료·교수자 피드백에 대한 학습자의 수용률은 전반적으로 높게(60.3%) 나타났고, 초고 대비 수정고의 질 또한 통계적으로 유의미한 수준에서 상승(+1.95:35.2%)했다. 이는 각 주체가 영역을 분담하여 순차적으로 피드백을 제공한 구조로 인해, 학습자가 인지적 과부하 없이 단계별로 초점화된 수정을 수행할 수 있었기 때문으로 해석되었다.

둘째, AI 피드백은 주제별 피드백 중 학습자의 수용률이 가장 높았고(67.9%), ‘문단 구조의 적절성’, ‘출처 표기의 정확성’ 등 글의 형태적 측면에 대한 개선 효과가 큰 것으로 나타났다. 또한 수정고1 작성 단계에서 상승한 형태적 측면의 점수는 이후 단계에서 변화가 크지 않아, AI 피드백을 초기에 배치해 학습자가 글의 외형을 먼저 정비하고 내용 수정에 집중하게 한 전략이 목적인 효과를 발생시킨 것으로 판단되었다. 한편, 형태적 측면에 대한 피드백을 받고 수정하는 과정에서 학습자가 내용적 측면에 대한 ‘자기 수정’을 수행하는 연쇄 효과가 발생하여, 내용적 측면의 개선도 일부 이루어진 것으로 확인되었다.

셋째, 동료 피드백은 해석의 다양성이 존재하는 내용적 측면을 다루고 있어, 학습자의 거부율(17.2%)과 자기 수정율(11.7%)이 타 피드백 대비 높게 나타났다. 수정고2 작성 단계에서는 학습자의 방어적이면서도 책임 있는 저자성(authorship)이 강하게 발휘되어, 동료 피드백을 거부한 뒤 동료의 오독(誤讀)이 발생한 지점을 탐색하여 스스로 수정하는 양상이 관찰되었다. 그 결과, ‘근거의 신뢰성’ 및 ‘이유의 타당성’ 영역에서 유의미한 질적 개선이 이루어졌다.

넷째, 교수자 피드백은 학습자가 피드백의 타당성은 인정하지만 수정은 유보하는 보류율이 타 피드백 대비 높게(26.1%) 나타났는데, 이는 교수자 피드백이 논지의 재구성 등 난이도 높은 수정을 요구한 경우에 주로 발생했다. 다만, 학습자들은 교수자 피드백을 보류한 경우에도 가능한 수준에서 대안적 수정을 시도하는 양상을 보였다. 또한 최종고에서는 ‘반론·재반론의

적절성'과 같이, AI나 동료의 타당도 높은 피드백을 제공하기에 한계가 있는 영역에 대한 개선이 뚜렷하게 이루어졌으며, 앞 단계의 수정에서 발생한 문제가 조정되는 등 전반적인 질적 제고가 나타났다.

이러한 결과는 AI·동료·교수자의 피드백을 상호 보완적으로 결합하고, 그 역할과 제공 시점을 분담한 단계적 구조가 학습자의 글쓰기 수행을 실질적으로 향상시킬 수 있음을 시사한다. 특히 학습자들은 각 주체의 피드백을 단순히 수용하거나 거부하는 것에서 나아가, 자신의 판단을 명시적으로 언어화하며 능동적으로 대응했다. 이러한 반응 과정에서 드러난 저자성의 다양하고 복합적인 양상은 글쓰기 교육에서 피드백이 단순한 오류 수정을 위한 기제가 아니라, 필자의 정체성을 구성하는 촉매(觸媒)로 작용할 수 있음을 보여주었다. 아울러 형태 중심에서 내용 중심으로, 다시 통합적 성찰로 이어지는 피드백의 순차적 구조는 학습자에게 점진적인 성찰의 기회를 제공하며, 각 단계의 수정이 다음 단계를 준비하는 발판이 될 수 있다는 점에서 교육적 유용성을 갖는다고 할 수 있다.

한편, 이 연구는 단계적 피드백 방안을 구성하여 제안하고 그 효과를 탐색적으로 확인하는 데 목적을 두었기에, 통제집단을 설정하거나 다른 피드백 방안과의 비교를 통해 상대적 효과성을 검증하는 작업은 수행하지 않았다. 또한 연구 참여자가 단일 대학의 2학년 학습자 34명에 한정되어 있어, 연구 결과를 일반화하는 데는 일정한 한계가 존재한다. 후속 연구에서는 보다 광범위한 학습자 집단을 대상으로, 다른 피드백 방안과의 비교를 포함한 외적 타당도 검증을 통해 단계적 피드백 방안의 효과를 보다 정밀하게 검증하고자 한다.

또한 AI, 동료, 교수자 피드백의 순서로 배치된 단계적 피드백 방안의 구조는 전체적으로 높은 피드백 수용률과 글의 질 개선을 이끌어냈으나, 교수자 피드백이 글의 논지를 재구성하는 전면 수정을 요구할 경우 학습자의 보류율이 높아지는 경향이 나타났다. 이러한 한계를 보완하기 위해서는 단계적 피드백 시행에 앞서 학습자가 자신의 글이 과제 요구와 설정한 논지에

부합하는지를 자가 진단할 수 있도록 간단한 체크리스트 기반의 ‘사전 논지 점검 활동(preliminary thesis check)’을 도입하는 것이 효과적일 수 있다. 이 점검 결과에 따라 중대한 방향성 오류가 의심되는 경우, 교수자가 간단한 확인 및 권고를 통해 초기 논지 방향을 조정할 수 있도록 선택적으로 개입한다면 교수자의 부담을 줄이면서도 후속 피드백 단계에서의 전면 수정 가능성을 줄일 수 있을 것이다.

글쓰기 교육에서 피드백은 ‘무엇을 조언할지’뿐 아니라 ‘누가, 언제, 어떻게 개입할지’를 함께 설계하려는 교육적 고민을 토대로 제공될 때 학습자의 성장을 이끌어낼 수 있다. 이 연구는 AI·동료·교수자의 피드백을 역할과 순서를 분담하여 구조화함으로써 학습자의 저자성과 글의 질을 효과적으로 제고할 수 있음을 탐색적으로 확인했다. 앞으로 피드백의 주체와 시점, 방식을 유기적으로 조율하는 방안에 대한 논의가 글쓰기 교육 현장에서 더욱 활발히 이루어지기를 기대한다.

* 본 논문은 2025.10.14. 투고되었으며, 2025.11.10. 심사가 시작되어 2025.12.03. 심사가 종료되었음.

참고문헌

- 공성수(2022), 「대학 글쓰기 피드백 연구의 경향과 시사점」, 『리터러시연구』 13(3), 153-214.
- 김민정(2024), 「고차원적 사고 유형 과제에 대한 교수자 피드백, 동료 피드백, 생성형 AI 피드백의 비교」, 『교육방법연구』 36(4), 337-359.
- 김연희·김남국(2013), 「과정 중심의 작문 수업에서 교사 피드백과 동료 피드백의 유형 및 특징」, 『언어학연구』 29, 1-26.
- 김은선(2025), 「교사와 인공지능 글쓰기 피드백에 대한 초등학교생의 반응」, 『한국초등국어교육』 80, 5-35.
- 박혜림·이의연(2025), 「AI 피드백을 활용한 디지털 글쓰기 프로그램의 효과 분석」, 『사물인터넷융복합논문지』 11(2), 15-22.
- 심지연(2024), 「생성형 AI를 활용한 한국어 쓰기 피드백 방안」, 『Journal of Korean Culture』 67, 45-79.
- 안효원·나현수·하채희·송용창(2025), 「AI 코스웨어의 글쓰기 피드백에 대한 교사 인식 연구」, 『새국어교육』 142, 183-214.
- 이유나·손은지·이상수(2024), 「ChatGPT 피드백과 동료 피드백이 중학생 글쓰기 능력에 미치는 영향」, 『교육방법연구』 36(2), 111-137.
- 이윤빈·정희모(2014), 「대학생 글쓰기에서 동료 피드백의 양상 및 타당도 연구」, 『작문연구』 20, 299-334.
- 이윤빈(2025), 「대학 글쓰기 교육에서 교수자·동료·AI 피드백의 특성 및 타당도 비교 - 대학생 글쓰기의 전형적 문제를 포함한 논증적 글을 중심으로」, 『교양교육연구』 19(3), 19-33.
- 정유남(2024), 「대학 글쓰기 수업에서 학생 인식 기반 AI 피드백의 효용성 연구」, 『교양교육연구』 18(5), 159-173.
- 정희모·이재성(2008), 「대학생 글쓰기의 수정 방법에 관한 실험 연구-자기첨삭, 동료첨삭, 교수첨삭의 효과를 중심으로」, 『국어교육학연구』 33, 657-685.
- AlGhamdi, R. (2024), "Exploring the impact of ChatGPT-generated feedback on technical writing skills of computing students: A blinded study", *Education and Information Technologies* 29(14), 18901-18926.
- Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024), "Feedback sources in essay writing: peer-generated or AI-generated feedback?", *International Journal of Educational Technology in Higher Education* 21(1), 23.
- Bogolepova, S. V. (2025), "Potential of artificial intelligence tools for text evaluation and feedback provision", *Professional Discourse & Communication* 7(1), 70-88.
- Carless, D. & Boud, D. (2018), "The development of student feedback literacy: enabling uptake of feedback", *Assessment & Evaluation in Higher Education* 43(8), 1315-1325.

- Chan, S., Lo, N., & Wong, A. (2025), "Leveraging generative AI for enhancing university-level English writing", *Cogent Education* 12(1), 2440182.
- Ellis, R. (2001), *Task-based language learning and teaching*, OUP.
- Escalante, J., Pack, A., & Barrett, A. (2023), "AI-generated feedback on writing: Insights into efficacy and ENL student preference", *International Journal of Educational Technology in Higher Education* 20(57), 1-20.
- Fatmah, E. N., Warsono, W., & LB, D. A. (2022), "Teachers' Practice of Form-Focused Feedback in Writing Classroom Teaching Writing At SMK N 1 Blado, Batang", *English Education Journal* 12(2), 284-292.
- Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010), "Improving the effectiveness of peer feedback for learning", *Learning and instruction* 20(4), 304-315.
- Hovardas, T., Tsivitanidou, O. E., & Zacharia, Z. C. (2014), "Peer versus expert feedback: An investigation of the quality of peer feedback among secondary school students", *Computers & Education* 71, 133-152.
- Ivanič, R. (1998), *Writing and identity* (Vol. 10), Amsterdam: John Benjamins Publishing Company.
- Jiang, Y. (2025), "Interaction and dialogue: Integration and application of artificial intelligence in blended mode writing feedback", *The Internet and Higher Education* 64, 100975.
- Ji, J., Zhao, H., & Wu, Y. (2025), "Feedback Sources in Argumentative Writing: AI Feedback or Peer Feedback?", In *Proceedings of the 19th International Conference of the Learning Sciences-ICLS 2025*, pp. 1242-1246. International Society of the Learning Sciences.
- Lightbown, P. M. & Spada, N. (1990), "Focus-on-form and corrective feedback in communicative language teaching: Effects on second language learning", *Studies in Second Language Acquisition* 12(4), 429-448.
- Mirsanjari, Z. & Moradi Abbasabady, M. (2024), "The Effectiveness of Blended Feedback in EFL Learners' Essay Writing and Their Perceptions", *International Journal of Research in English Education* 9(4), 17-32.
- Nelson, G. L. & Carson, J. G. (1998), "ESL students' perceptions of effectiveness in peer response groups", *Journal of second language writing* 7(2), 113-131.
- Oley, N. (1992), "Extra credit and peer tutoring: Impact on the quality of writing in introductory psychology in an open admissions college", *Teaching of Psychology* 19(2), 78-81.
- Paulus, T. M. (1999), "The effect of peer and teacher feedback on student writing", *Journal of second language writing* 8(3), 265-289.

- Renati, D., Renata, D., Hidayah, N., & Indriani, L. (2024), "The impact of teacher's feedback on students' writing ability", *Journal of English Teaching and Research* 9(1), 105-113.
- Silver, R. & Lee, S. (2007), "What does it take to make a change? Teacher feedback and student revisions", *English Teaching: Practice and Critique* 6(1), 25-49.
- Sperber, L., MacArthur, M., Minnillo, S., Stillman, N., & Whithaus, C. (2025), "Peer and AI Review+ Reflection (PAIRR): A human-centered approach to formative assessment", *Computers and Composition* 76, 102921.
- Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024), "Comparing the quality of human and ChatGPT feedback of students' writing", *Learning and Instruction* 91, 101894.
- Topping, K. (1998), "Peer assessment between students in colleges and universities", *Review of educational Research* 68(3), 249-276.
- Wang, Z. & Han, F. (2022), "The effects of teacher feedback and automated feedback on cognitive and psychological aspects of foreign language writing: A mixed-methods research", *Frontiers in Psychology* 13, 1-16.
- Yang, M., Badger, R., & Yu, Z. (2006), "A comparative study of peer and teacher feedback in a Chinese EFL writing class", *Journal of second language writing* 15(3), 179-200.
- Yoon, S. -Y., Miszoglaid, E., & Pierce, L. R. (2023), "Evaluation of ChatGPT feedback on ELL writers' coherence and cohesion" (arXiv Preprint), arXiv.
- Zacharias, N. T. (2007), "Teacher and student attitudes toward teacher feedback", *REL C Journal* 38(1), 38-52.
- Zeevy-Solovey, O. (2024), "Comparing peer, ChatGPT, and teacher corrective feedback in EFL writing: Students' perceptions and preferences", *Technology in Language Teaching & Learning* 6(3), 1482-1482.
- Zhang, Z. V. & Hyland, K. (2022), "Fostering student engagement with feedback: An integrated approach", *Assessing Writing* 51, 100586.

AI·동료·교수자의 단계적 피드백이 대학생의 글쓰기 수행에 미치는 효과

이윤빈

이 연구는 AI·동료·교수자가 역할을 분담한 단계적 피드백 방안을 제안하고, 이를 대학 글쓰기 교육 현장에 적용하여 그 효과를 탐색했다. 세 주체의 피드백을 상호 보완적으로 활용하기 위해 (1) AI는 글의 구조와 표현을 포함하는 형태적 측면에, (2) 동료는 독자 반응의 일환으로 내용적 측면에, (3) 교수는 형태와 내용을 아우르는 심층 통합적 측면에 집중하여 순차적으로 피드백을 제공하고, 학습자가 이를 비판적으로 점검하며 세 차례의 논증문 수정을 수행하게 했다. 그리고 주체별 피드백에 대한 학습자의 반응과 단계별 글의 질 변화 양상을 고찰했다.

연구 결과, 학습자들은 전반적으로 높은 수준으로 피드백을 수용했고, 단계별 글의 질 또한 유의미하게 개선되었다. AI 피드백 단계에서는 글의 외형이 정비되어 형태적 완성도가 높아진 후, 이 효과가 이후 단계에서도 안정적으로 유지되었다. 동료 피드백 단계에서는 다양한 독자 반응을 비판적으로 해석하는 학습자의 저자성이 부각되었고, 교수자 피드백 단계에서는 앞 단계의 한계를 보완하는 통합적 질 제고가 이루어졌다. 또한 세 피드백은 각기 다른 양상으로 학습자의 자기 수정을 유도하는 촉매로 작용했다. 이러한 결과는 단계적 피드백 방안이 학습자의 인지적 부담을 완화하여 초점화된 수정을 유도하고, 나아가 비판적 수용과 저자성 발달을 함께 촉진할 수 있음을 보여준다.

핵심어 대학 글쓰기, 단계적 피드백, AI 피드백, 동료 피드백, 교수자 피드백, 피드백 수용

ABSTRACT

Effects of Sequential Feedback from AI, Peers, and Instructors on College Students' Writing Performance

Lee Yunbin

This study proposes a stepwise feedback model involving AI, peers, and instructors, each with a distinct role, and examines its effects in a college writing context. AI addressed structural and linguistic forms; peers contributed content-level reader responses; and instructors provided integrated feedback that combined both aspects. Students revised their argumentative essays three times while actively engaging with each feedback source.

The findings indicate high feedback uptake and significant writing improvement at each stage. AI feedback improved formal accuracy with sustained effects, peer feedback enhanced authorial agency through critical reader interpretation, and instructor feedback was consolidated and refined before revisions. Each feedback type uniquely fostered self-revision, suggesting that the model reduces cognitive load, supports focused revisions, and promotes both critical reception and authorial voice.

KEYWORDS College writing, stepwise feedback, AI feedback, peer feedback, instructor feedback, feedback uptake