

생성형 AI 기반 음악 창작의 구조 모델 연구: Planning-Process-Output Layer의 3단계 프레임워크와 교육 현장 적용

이원빈, 조태선*

명지전문대학 실용음악과, *청운대학교 실용음악학과

e-mail:mamawindows@naver.com

*entheos@chungwoon.ac.kr

A Structural Model of Generative AI-Based Music Creation: A Three-Layer Framework of Planning-Process-Output and Its Application in Educational Settings

Won-Vin Lee, Tae-Seon Cho*

Dept. of Applied Music, Myongji College

*Dept. of Applied Music, Chungwoon University

요약

본 연구는 생성형 AI 환경에서 음악 창작 구조를 프롬프트 설계(Planning Layer) → 창작 수행(Process Layer) → 결과물(Output Layer)의 3단계 모델로 이론화하고, 화성의과대학교 교양수업 'AI 뮤직 크리에이터'의 실습 사례를 통해 현장 적용 가능성을 탐색하는 데 목적이 있다. 해당 수업은 4단계 AI 이중 도구 결합 파이프라인을 기반으로 운영되었으며, 30명의 학습자가 7개 조로 편성되어 기획, 스토리보드, 작사·작곡, 영상 편집의 역할을 분담하고 각 조는 주제에 맞는 뮤직비디오 콘텐츠를 제작하였다. 사례 분석 결과, 음악적 감성과 주제 의도를 시각 언어로 전환하는 프롬프트 설계 수준이 이후 창작 수행의 체계성과 결과물 완성도에 중요한 영향을 미치는 것으로 나타났으며, 이는 본 연구의 창작 구조 모델이 교육 현장에서 유효한 분석틀로 기능할 수 있음을 시사한다. 본 연구는 생성형 AI 기반 창작교육의 핵심 역량이 설계·수정·판단 능력의 통합적 구조로 재편되고 있음을 논의하며, AI 음악교육 및 뮤직 크리에이터 교육과정 개발을 위한 개념적 기반을 제공한다.

1. 서론

1.1 연구 배경

디지털 플랫폼의 확산과 생성형 AI 기술의 발전은 음악 창작 환경을 근본적으로 재편하고 있다. AI 기반 작곡 도구는 작곡·편곡·믹싱·마스터링 전 과정에 개입하여 기존의 협업 중심 제작 구조를 '1인 제작 시스템'으로 전환시키고 있으며, '프롬프트 디자인(prompt design)' 능력이 창작의 핵심 역량으로 부상하고 있다 [1]. 또한 생성형 AI는 이미지·영상·모션·편집 영역과 결합하면서 음악 중심 콘텐츠를 멀티미디어 창작물로 확장시키고 있다.

이러한 흐름 속에서 화성의과대학교는 2026년 교양수업 'AI 뮤직 크리에이터'를 개설하고, Gemini Nano·Grok·Google Video AI·CapCut의 4가지 도구를 순차적으로 연결하는 파이프라인 기반 시네마틱 뮤직비디오 제작 실습을 운영하고 있다.

본 수업은 본 연구에서 제안하는 창작 구조 모델이 실제 교육 현장에서 어떻게 구현되는지를 검토할 수 있는 적합한 사례를 제

공한다. 그러나 기존 연구들은 개별 교육 효과 중심의 논의에 그치며, 프롬프트 설계-창작 수행-결과물 품질의 구조적 관계를 통합적으로 설명하는 연구는 충분하지 않다.

본 연구는 두 가지 목적을 지향한다. 첫째, 생성형 AI 기반 음악 창작의 구조를 Planning Layer-Process Layer-Output Layer의 3단계 모델로 이론화한다. 둘째, 'AI 뮤직 크리에이터' 수업 사례를 통해 해당 모델의 현장 적용 가능성을 탐색한다.

1.2 이론적 배경

음악 제작 환경은 MIDI 도입(1980년대), DAW 대중화(2000년대)를 거쳐 생성형 AI 도구의 등장으로 새로운 전환기를 맞이하고 있다. AI 작곡 도구는 텍스트 기반 프롬프트를 통해 멜로디·화성·편곡·보컬을 포함한 완성도 높은 음악을 자동 생성함으로써, 기존 도구들과 질적으로 구별되는 창작 환경을 제공한다 [1]. Suno, Udio 등 상용 서비스는 음악 제작을 전문 영역에서 대중적 창작 영역으로 확장시키고 있으며 [2][3], 창작자의 역할을 기술 수행 중심에서 기획·설계·선택 중심으로 이동시키는 구조적 전환을 수반한다.

이러한 맥락에서 뮤직 크리에이터는 콘텐츠 기획, 퍼스널 브랜딩, 플랫폼 이해, 수익화 전략 등을 통합적으로 수행하는 창작 주체로 재정의되며[4], AI 창작 환경에서 이러한 역량은 구체적으로 '프롬프트 설계 능력'으로 수렴된다. 실제로 감각적 맥락과 기술적 조건을 복합 구조화한 프롬프트는 비구조적 입력보다 높은 일관성과 완성도를 보이는 결과물로 이어지는 경향이 있으며[1], 이는 프롬프트 설계의 구체성과 맥락성이 결과물 품질을 좌우하는 핵심 요인임을 보여준다.

본 연구의 3단계 창작 구조 모델은 이러한 선행 논의를 토대로, AI 기반 음악 제작의 절차적 흐름—① 프롬프트 입력, ② AI 생성 처리, ③ 결과물 수정 및 반복, ④ 편집 및 배포—을 창작자의 인지적·기획적 역량이 투입되는 구조로 재해석한 것이다[1]. Planning Layer는 창작 의도를 언어적으로 구조화하는 단계이며, Process Layer는 AI 생성·결과물 평가·반복 수정·최종 선택이 순환적으로 이루어지는 수행 단계로 메타인지적 역량이 핵심으로 작용한다. Output Layer는 최종 결과물을 음악 완성도·콘텐츠 전달력·타겟 적합성의 차원에서 판단하는 단계이다.

2. 본 론

본 연구는 화성의과학대학교 2026학년도 1학기 교양수업 'AI 뮤직 크리에이터'의 프로젝트 실습을 분석하는 탐색적 사례연구이다. 연구 대상은 의료경영·간호·바이오헬스·데이터사이언스·라이프스타일·스포츠과학·의생명 등 다양한 전공의 수강생 30명으로, 총 7개 조로 편성되었다. 각 조는 기획·스토리보드·작사·작곡·영상 제작의 역할을 분담하였다. 분석 자료는 수업계획안, 조별 기획안 및 프롬프트 산출물, 중간 결과물, 최종 제출물, 수업 관찰 기록으로 구성하였으며, 사례분석과 과정추적(process tracing)을 결합한 질적 분석 접근을 택하였다. 본 연구는 후속 양적 연구를 위한 모델 검토 및 현장 적용 가능성 탐색에 중점을 둔다.

2.1 수업 개요 및 실습 구조

'AI 뮤직 크리에이터' 수업은 Gemini Nano → Grok → Google Video AI → CapCut의 4단계 AI 이종 도구 결합 파이프라인으로 구성되었다. 수업의 목표는 '음악적 감성의 시각적 프롬프트 변환', '4가지 AI 도구의 특성 이해 및 연결', '고화질 소스 확보 및 박사 기반 영상 편집의 세 가지 역량을 통합적으로 습득하는 데 있었다.

2.2 창작 구조 모델의 실습 적용 분석

실습 사례에서 창작 구조 모델의 각 층위는 다음과 같은 양상으로 구현되었다. Planning Layer에 해당하는 Step 1(Gemini Nano) 단계에서 학습자는 음악 주체의 감성과 콘셉트를 AI가 처리 가능한 시각적 언어로 변환하는 프롬프트 설계를 수행하였다. 관찰 결과, 감정선과 장면적 맥락을 구체적으로 설계한 조일수록 이후 단계에서 더 일관된 시각 결과를 확보하는 경향이 나타난 반면, 프롬프트가 추상적인 경우에는 주제와 어울리지 않는 결과가 반복 생성되었으며, 이는 Planning Layer의 질이 후속 단계의 효율성과 직결됨을 시사한다. Process Layer에 해당하는 Step 2(Grok)와 Step 3(Google Video AI) 단계에서 학습자는 복수의 스틸컷 후보를 생성·비교하여 최종 이미지를 선별하는 순환적 판단 과정을 거쳤으며, 이후 'cinematic panning', 'slow zoom-in' 등의 카메라 컨트롤 프롬프트를 적용하여 모션 클립을 생성하였다.

초기 프롬프트에서 색감·공간감·분위기를 명확히 기술한 조는 이 단계에서의 수정 횟수가 상대적으로 적었던 반면, 구조성이 낮은 경우에는 재생성이 반복되는 양상이 관찰되었다. Output Layer에 해당하는 Step 4(CapCut) 단계에서 학습자는 오디오 파형의 킥 드럼 타이밍에 맞추어 컷을 전환하고, 후렴구 구간에 Audio Spectrum 및 Light Glitch 효과를 적용하였다. 이 단계에서 결과물의 완성도는 이미지 품질 단독이 아니라 음악-영상의 리듬적 정합성·주제 전달력·플랫폼 적합성의 복합적 관계에 의해 형성되었으며, 이는 Output Layer의 질이 기술적 완성도와 함께 창작 기획의 적합성과 연동됨을 보여준다.

2.3 실습 사례의 창작 구조 모델 검토

사례를 종합하면, Planning Layer에서의 프롬프트 설계 수준은 이후 과정 전체의 방향성과 효율성을 좌우하는 핵심 변수로 작용하였다. 프롬프트를 구조적으로 설계한 조는 이미지 생성 및 최종 편집 단계에서 수정의 일관성이 유지된 반면, 비구조적 설계의 경우 Process Layer에서 반복적 재생성이 장기화되고 음악-영상 일체감도 약화되었다.

의료경영·간호·데이터사이언스 등 비음악 전공 학습자들이 프롬프트 설계 능력을 중심으로 창작물을 완성한 점은, AI 환경에서 설계 능력이 보편적 창작 역량으로 기능할 수 있음을 시사한다.

3. 결론 및 시사점

본 사례 분석은 AI 환경에서 프롬프트 디자인 능력이 창작의 핵심 역량으로 부상한다는 논의를 교육 현장을 통해 지지한다. 수업 사례에서 창작자의 역할은 결과물을 직접 생성하는 주체에서

AI 도구를 연결·조율하고 결과물을 선별·수정하는 크리에이티브 디렉터로 전환되는 양상을 보였으며, 이는 생성형 AI 시대의 창작이 기술 수행 중심에서 설계와 판단 중심으로 이동하고 있음을 보여준다. 아울러 비음악 전공 학습자들이 프롬프트 설계 능력을 중심으로 시네마틱 뮤직비디오를 완성하였다는 사실은, Planning Layer가 특정 전공 지식보다 보편적 언어 기획 역량에 크게 의존할 수 있음을 시사하며, 실용음악 교육의 핵심이 프롬프트 디자인 능력·콘텐츠 기획력·멀티미디어 융합 역량으로 확장되어야 한다는 명제를 교육 현장 차원에서 뒷받침한다[1].

이는 AI 기반 창작 교육이 음악 전공 교육을 넘어 교양 교육 차원에서도 유효한 모형으로 기능할 가능성을 보여준다. 다만 본 연구는 단일 수업 사례의 관찰 분석에 기반하므로 결과를 일반화하는 데 한계가 있으며, 가설의 통계적 검증까지 도달하지 못하였다. 향후 연구에서는 복수의 교육기관 사례를 비교하고 프롬프트 유형별 결과물 품질 차이를 정량적으로 검증하는 연구가 필요하며, Suno·Udio 등 음악 생성 도구를 포함한 이종 도구 파이프라인 환경으로 연구 범위를 확장할 필요가 있다.

본 연구는 생성형 AI 기반 음악 창작을 Planning Layer-Process Layer-Output Layer의 3단계 구조로 개념화하고, 'AI 뮤직 크리에이터' 수업의 실습 사례를 통해 그 현장 적용 가능성을 검토하였다. 사례 분석 결과, Gemini Nano(프롬프트 설계) → Grok(이미지 생성·선별) → Google Video AI(모션 생성·수정) → CapCut(결과물 조립)의 파이프라인은 창작 구조 모델과 정합적으로 대응하였으며, Planning Layer에서의 프롬프트 설계 수준이 이후 창작 수행과 최종 결과물 품질에 중요한 영향을 미치는 양상이 관찰되었다.

본 연구의 학술적 기여는 세 가지로 정리된다. 첫째, 생성형 AI 기반 음악 창작을 프롬프트 설계-창작 수행-결과물의 구조 관계로 모델화하였다. 둘째, AI 기반 음악 제작에 관한 선행 실무 중심 논의를 창작 구조의 이론적 모델과 연결하여 실무와 이론 사이의 설명 체계를 보완하였다. 셋째, 실제 수업 사례를 통해 AI 기반 창작교육 모형의 구체적 구현 가능성을 제시하고 실천적 시사점을 제공하였다.

생성형 AI 시대에 창작 역량의 핵심은 전통적 작곡 기술을 넘어, 프롬프트 설계 능력·반복 수정 능력·결과물 판단 능력이 결합된 설계 중심 역량으로 전환되고 있다. 따라서 향후 AI 음악교육은 이러한 구조적 전환을 반영한 교육모형의 체계적 개발과 검증으로 나아갈 필요가 있다.

참고문헌

[1] 이원빈, 조태선. (2025). 기술 발전에 따른 음악 제작 환경의 변화 연구 — AI 작곡을 중심으로 —. 한국산학기술학회. 춘계 학술발표논문집.

- [2] 서미경, 안성희. (2026). AI 시대 음악 창작의 도구 변화와 학습 효과: Suno를 활용한 교육 데이터 중심으로. 문화산업 연구, 26(1).
- [3] 조영일, 박재완. (2026). 상용 생성형 AI 음악 서비스의 음악적 아티팩트 탐지: 설명 가능한 머신러닝 접근을 중심으로. 아시아태평양융합연구교류논문지, 12(3).
- [4] 이원빈, 조태선. (2025). 뉴미디어 발전에 따른 뮤직 크리에이터 K-pop 콘텐츠 분석: 유튜브 채널을 중심으로. 한국산학기술학회, Vol26 No.9.
- [5] 오희숙. (2023). 모방과 창의성 사이의 테크놀로지: 코프(D. Cope)의 초기 AI 작곡 모델 에미(Emmy) 연구. 음악이론연구, 40권.
- [6] 유공호. (2025). AI 작곡 도구를 활용한 실용음악 1인 창작자의 음악 콘텐츠가 팬층의 공감, 충성도에 미치는 영향 연구: 창의적 인성의 매개효과를 중심으로. 콘텐츠와 산업, 7(3).
- [7] 한수현, 한경훈. (2025). 초등 음악창작 교육에서 매개변수형 AI 작곡 수업이 음악 자기효능감과 흥미에 미치는 영향. 음악교육연구, 54(3).
- [8] 설황룡. (2026). AI 작곡 시스템을 활용한 협업 창작이 음악 전공 대학생의 협력적 자기효능감을 통한 문제해결능력에 미치는 영향 연구. 콘텐츠와 산업, 8(1).
- [9] 신재준, 김영삼. (2025). AI 작곡 기술의 영화음악 적용 가능성 연구. 영화교육연구, 19(3).
- [10] 한수현, 한경훈. (2024). 구성주의 학습 환경 설계모형 (CLEs) 기반 인공지능(AI) 작곡 프로그램을 활용한 음악창작 수업 개발. 예술교육연구, 22(4).
- [11] 신혜련, 이용배. (2026). AI 도구 활용이 음악 수업의 흥미도 및 자기효능감에 미치는 영향. 정보교육연구, 4(1).