

국립발레단과 유니버설발레단 유튜브 댓글의 온라인 반응 구조 비교 분석

김민지* 숙명여자대학교

초록 본 연구는 국립발레단과 유니버설발레단 관련 유튜브 콘텐츠 댓글을 대상으로 온라인 이용자 반응의 차이를 비교·분석하는 데 목적이 있다. 이를 위해 2015년 1월 1일부터 2025년 12월 31일까지 게시된 관련 콘텐츠에서 총 902개 동영상과 10,621개의 원댓글을 수집하였으며, 단어 빈도분석과 LDA 토픽모델링, 토픽 간 거리 지도 및 토픽-키워드 네트워크 분석을 실시하였다. 분석 결과, 두 발레단 댓글에서는 공연, 무용수, 무대, 작품, 감동, 응원 등 발레 감상과 관련된 공통 반응이 확인되었다. 그러나 국립발레단 댓글에서는 예술성, 무용수의 기량, 해외 발레와의 비교, 특정 무용수에 대한 팬심이 상대적으로 두드러졌으며, 유니버설발레단 댓글에서는 「호두까기인형」을 중심으로 한 연말 공연 기대, 가족 관람, 정서적 기억, 무용수의 훈련과 직업적 현실에 대한 공감에 두드러졌다. 즉 국립발레단은 예술적 기준과 기량을 중심으로 '평가의 대상'으로 수용되는 경향을 보인 반면, 유니버설발레단은 시즌 레퍼토리와 관람 경험을 중심으로 '경험의 대상'으로 수용되는 경향을 보였다. 다만 이러한 구분은 두 발레단의 고정적 속성이나 모든 이용자의 인식을 의미하는 것이 아니라, 본 연구에서 수집한 댓글에 나타난 상대적 반응 경향을 설명하기 위한 분석적 개념이다. 본 연구는 동일한 발레 장르 안에서도 발레단별 레퍼토리 구성, 콘텐츠 특성 및 축적된 단체 이미지와 온라인 이용자 반응이 서로 다른 방식으로 관련될 가능성을 보여준다는 점에서 의의가 있다. 또한 유튜브 댓글을 공연예술의 디지털 수용 자료로 활용하고, 두 발레단의 온라인 반응 구조를 비교함으로써 무용학 분야의 디지털 관객 연구에 기초자료를 제공하였다.

주요어 : 국립발레단, 유니버설발레단, 유튜브 댓글, 텍스트마이닝, 온라인 반응 구조

I. 서론

1. 연구의 필요성 및 목적

디지털 전환은 공연예술의 향유 방식과 유통 구조를 근본적으로 변화시키고 있다. 공연은 더 이상 물리적 공연장에 한정되지 않으며, 유튜브와 같은 온라인 플랫폼을 통해 반복적으로 소비되고 재해석되는 문화콘텐츠로 확장되고 있다(백민제, 2021; 장지영, 2021). 특히 공연예술 분야에서는 공연 실황, 하이라이트 영상, 리허설, 인터뷰, 홍보 콘텐츠 등 다양한 형태의 영상 콘텐츠가 온라인 플랫폼을 통해 제공되고 있으며, 이용자들은 이를 시청하고 댓글을 통해 감상과 평가, 감정, 정보 등을 공유하고 있다(김태연, 2024; 나보리, 이슬기, 2023; 정종은, 2021; 현보람, 2022). 이러한 변화는 공연예술 소비가 단순한 관람 경험을 넘어 온라인 플랫폼 기반의 상호작용과 재해석 과정으로 확장되고 있음을 보여준다.

발레 역시 유튜브를 중심으로 다양한 영상 콘텐츠가 활발히 유통되고 있으며, 이에 대한 이용자 반응 또한 댓글의 형태로 지속적으로 축적되고 있다(염지현, 이해준, 2026; 조연위, 안남일, 2023). 유튜브 댓글은 공연예술에 대한 이용자의 인식과 감상이 집합적으로 드러나는 디지털 텍스트 자료로서, 변화하는 문화예술 향유 방식을 보여주

* 숙명여자대학교 무용학과 강사, 0426mjk@gmail.com

는 중요한 데이터라 할 수 있다(김선정, 2025; 김가은, 홍미성, 2020). 과거 공연 감상이 개인적이고 일회적인 경험에 가까웠다면, 디지털 플랫폼 환경에서는 감상과 평가, 감정 표현이 텍스트 형태로 외부화되어 타인과 공유된다. 따라서 댓글은 온라인 이용자의 반응과 해석이 집합적으로 드러나는 비정형 텍스트 데이터로서, 공연예술 이용자 반응 구조의 특성을 분석하는 데 유의미한 자료로 활용될 수 있다(김선영, 이의신, 2017; 김태연, 2024; 안성우, 2025). 나아가 댓글은 단순한 감상문의 집합을 넘어 특정 해석과 반응이 반복적으로 공유·확산되는 사회적 의미 형성 과정으로 이해될 수 있다(김경식, 2025; 정원준, 2021). 그러나 이러한 온라인 반응 구조가 어떠한 양상으로 형성되는지에 대한 비교 연구는 아직 충분히 이루어지지 않았다.

최근 무용학 분야에서도 온라인 플랫폼 데이터를 활용한 연구가 증가하고 있으며, 이에 따라 비정형 텍스트 기반 분석 방법의 활용 가능성도 확대되고 있다. 특히 텍스트마이닝과 토픽모델링은 댓글과 같은 비정형 텍스트에서 반복적으로 나타나는 단어 패턴을 분석하고, 그 이면에 존재하는 잠재적 주제 구조를 파악하는 데 유용한 방법으로 활용되고 있다(Blei, 2012). 장유정(2023)은 코로나19 이후 무용의 비대면화와 관련된 온라인 미디어 데이터를 분석하여 사회적 담론 변화 양상을 논의하였고, 김태연(2024)은 공공무용단 유튜브 콘텐츠의 제목, 본문, 댓글을 분석하여 콘텐츠 유형별 이용자 반응 특성을 도출하였다. 또한 이윤경(2023)은 국립현대무용단 유튜브 댓글을 통해 대중의 관심 분야와 선호 콘텐츠 유형을 분석하였으며, 염지현, 이해준(2026)은 성인발레 유튜브 콘텐츠를 대상으로 온라인 반응 구조의 특징과 변화를 살펴보았다. 이처럼 기존 연구들은 온라인 플랫폼 데이터를 활용하여 공연예술 이용자의 반응과 콘텐츠 소비 양상을 분석하였다는 점에서 의의를 지닌다. 특히 동일한 장르에 속한 공연예술단체라 하더라도 운영 기반과 콘텐츠 특성에 따라 이용자 반응이 서로 다르게 나타날 가능성이 있다. 그러나 기존 연구는 주로 특정 기관이나 콘텐츠 자체에 초점을 맞추어 이루어져, 공연예술단체의 운영 특성과 온라인 이용자 반응의 관계를 비교·분석한 연구는 상대적으로 부족하다.

공연예술단체의 운영 방식과 자원 구조는 콘텐츠 기획과 유통 전략, 관객 소통 방식과 밀접하게 연관되어 있으며(유진주, 2025), 이는 온라인 플랫폼에서 나타나는 콘텐츠 유형과 이용자 반응의 차이와도 관련될 가능성이 있다. 또한 공연예술단체의 대표 레퍼토리와 브랜드 전략은 이용자의 감상 방식과 관심 주제 형성에 영향을 미칠 수 있다. 국립발레단과 유니버설발레단은 동일한 발레 장르에 속하지만, 각각 공공적 운영 기반과 민간 운영 기반을 지닌다는 점에서 대비되는 사례이다. 국립발레단은 정부 재정지원을 기반으로 운영되는 공공 예술단체인 반면, 유니버설발레단은 민간 재원을 기반으로 비교적 독립적인 운영 구조를 갖는다. 특히 국립발레단은 공공성과 대표성을 중심으로 운영되는 반면, 유니버설발레단은 민간 재원을 바탕으로 브랜드 구축과 관객 확대 전략을 수행해 왔다는 점에서 차이를 보인다(국립발레단, 2026; 양영은, 2016; 유니버설발레단, 2026; 정옥희, 2025). 이러한 운영 기반과 콘텐츠 특성의 차이는 온라인 플랫폼에서 형성되는 이용자 반응 양상과도 관련될 가능성이 있다. 그럼에도 불구하고 기존 연구는 발레단의 영상 콘텐츠 활성화나 특정 콘텐츠에 대한 반응 분석에 주로 초점을 맞추어 왔으며, 서로 다른 특성을 지닌 발레단의 온라인 반응 구조를 비교한 연구는 충분히 이루어지지 않았다. 특히 단일 기관 중심의 분석은 온라인 반응 구조의 공통점과 차이점을 비교하고, 그 형성 맥락을 입체적으로 설명하는 데 한계를 지닌다. 국립발레단과 유니버설발레단을 비교하는 것은 동일한 발레 장르 안에서 서로 다른 운영 기반과 레퍼토리 및 콘텐츠 특성을 지닌 두 전문예술단체의 온라인 수용 양상을 대조함으로써, 단일 기관 분석만으로는 파악하기 어려운 공통 반응과 차별적 반응을 함께 밝힐 수 있다는 점에서 필요하다. 또한 이러한 비교는 공연예술의 디지털 수용 양상을 보다 입체적으로 이해하고, 각 발레단의 온라인 콘텐츠 운영과 관객 소통 방향을 구체화하는 데 활용할 수 있는 기초자료를 제공한다는 점에서 학술적·실천적 의의를 지닌다.

이에 본 연구는 동일한 발레 장르에 속하지만 운영 기반과 콘텐츠 특성이 상이한 국립발레단과 유니버설발레단을 대상으로, 유튜브 댓글에 나타난 온라인 반응 구조의 차이를 비교·분석하는 데 목적이 있다. 이를 위해 국립발

레단과 유니버설발레단 관련 유튜브 콘텐츠 댓글을 수집하고, 텍스트마이닝과 LDA 토픽모델링, 토픽 간 거리 지도 및 토픽-키워드 네트워크 분석을 활용하여 핵심 키워드와 잠재적 주제 구조를 도출하고자 한다. 또한 두 발레단 댓글에서 공통적으로 나타나는 관심 주제와 차별적으로 나타나는 반응 양상을 비교함으로써, 두 발레단의 운영 기반과 콘텐츠 특성이 온라인 이용자 반응과 어떠한 관련성을 보이는지를 탐색하고자 한다.

2. 연구문제

본 연구의 목적을 달성하기 위하여 설정된 연구문제는 다음과 같다.

첫째, 국립발레단과 유니버설발레단 유튜브 콘텐츠 댓글에 나타난 주요 단어의 특성은 어떠한가?

둘째, 국립발레단과 유니버설발레단 유튜브 콘텐츠 댓글의 주요 토픽 구조는 어떠한가?

셋째, 국립발레단과 유니버설발레단 이용자의 온라인 반응 양상에는 어떠한 차이가 나타나는가?

II. 연구방법

1. 분석 대상

본 연구는 국립발레단과 유니버설발레단 관련 유튜브 콘텐츠의 댓글을 분석 대상으로 선정하였다. 두 발레단 관련 동영상 댓글을 수집하기 위하여 YouTube Data API v3를 활용하였다. 데이터 수집 기간은 두 집단 간 비교 가능성을 확보하기 위하여 2015년 1월 1일부터 2025년 12월 31일까지로 설정하였다. 데이터 수집을 위한 검색어는 국립발레단의 경우 단체의 공식 명칭인 '국립발레단'으로 입력하였으며, 유니버설발레단의 경우 대중적 표기에서 '유니버설발레단'과 '유니버설발레단'이 혼용되는 점을 고려하여 두 검색어를 함께 사용하였다. 데이터 수집량은 각 검색어와 연관성이 높은 순으로 최대 500개의 동영상을 수집하고, 각 동영상에 해당되는 댓글 중 연관성이 높은 상위 500개를 수집하는 방식으로 설정하였다. 수집된 동영상은 video ID를 기준으로 중복 제거하였으며, 댓글 비활성화 영상과 공개 댓글이 존재하지 않는 영상은 최종 댓글 분석 자료에서 제외하였다. 그 결과 국립발레단 관련 동영상 402개와 댓글 6,379개, 유니버설발레단 관련 동영상 500개와 댓글 4,242개가 수집되어 총 902개 동영상과 10,621개의 댓글이 원자료로 구축되었다.

2. 데이터 정제

수집된 데이터에는 한글과 영문, 숫자, 특수문자 등 다양한 텍스트가 포함되어 있다. 본 연구에서는 국립발레단과 유니버설발레단 관련 유튜브 댓글 중 한글로 작성된 댓글만을 분석 대상으로 선정하였다. 외국어 댓글은 언어별 형태소 분석체계와 의미 표현 방식이 상이하여 동일한 전처리 기준을 적용하기 어렵고, 번역 과정에서 의미가 변형될 가능성이 있다. 따라서 두 발레단의 댓글을 동일한 기준으로 비교하고 분석의 일관성과 비교 가능성을 확보하기 위하여 한글 댓글만을 분석 대상으로 설정하였다. 이후 한글 텍스트를 기준으로 명사와 형용사를 추출하고, 분석결과의 정확도를 높이기 위하여 데이터 정제 과정을 거쳤다. 데이터 정제는 Python 기반 한국어 형태소 분석기와 연구자가 구축한 사전(Dictionary)을 활용하여 수행하였다.

데이터 정제 과정으로는 먼저 유의어를 설정하였다. 유의어는 발레 관련 유튜브 댓글에서 자주 등장하는 표현 중 같은 대상을 지칭하지만 서로 다른 표기로 사용되는 단어들을 하나의 대표 단어로 변환하여 추출하는 것이다. 예를 들어 '유니버설발레단'과 '유니버설발레단'은 동일한 단체를 의미하므로 '유니버설발레단'으로 통일하였으며,

‘국립 발레단’과 같이 띄어쓰기가 포함된 표현은 ‘국립발레단’으로 변환하였다. 또한 ‘호두까기 인형’, ‘백조의 호수’, ‘로미오와 줄리엣’ 등 두 개 이상의 단어가 결합되어 하나의 작품명이나 고유명사로 사용되는 표현들이 각각 분리되어 추출되지 않도록 지정어를 설정하였다.

마지막으로 유튜브에서 흔하게 사용하는 홍보성·플랫폼성 단어인 ‘구독’, ‘좋아요’, ‘댓글’, ‘영상’, ‘유튜브’와 같이 분석 결과에 중요한 영향을 미치지 않는 단어들은 제외어로 처리하였다. 또한 빈도는 높지만 특정 주제의 의미 구성에 기여하기보다 일반적 감탄 표현으로 반복되는 단어는 제외하였다. 다만 ‘예쁘다’, ‘힘들다’, ‘아프다’, ‘재밌다’, ‘대단하다’, ‘멋있다’ 등은 댓글에서 무용수의 미적 표현, 훈련 과정, 직업적 고충 및 공연 감상 반응을 설명하는 데 사용된 표현으로 판단하여 분석어로 유지하였다. 즉 본 연구에서는 품사 자체를 일률적인 제외 기준으로 적용하기보다, 해당 단어가 토픽의 의미를 구성하고 해석하는 데 기여하는지를 기준으로 분석어 포함 여부를 결정하였다. 분석 시 결과의 정확도를 높이기 위하여 한 음절 단어는 제외한 후 두 음절 이상의 단어만을 추출하도록 설정하였으며, 정제 후 토큰 수가 2개 미만인 댓글, 한글 비율이 낮은 댓글, 두 집단에 중복 수집된 동영상의 댓글은 최종 분석자료에서 제외하였다.

3. 분석 방법

본 연구는 연구문제 1과 관련하여 국립발레단과 유니버설발레단 관련 유튜브 콘텐츠 댓글에서 주로 사용되는 단어를 살펴보기 위하여 단어 빈도분석(Word Frequency)을 실시하였다. 또한 두 집단의 댓글에서 나타나는 핵심 단어를 한눈에 비교할 수 있는 시각화 자료를 제시하기 위하여 워드 클라우드(Word Cloud)를 실행하였다.

이후 연구문제 2와 관련하여 국립발레단과 유니버설발레단에 대한 이용자들의 관심 주제를 살펴보기 위하여 각 집단별로 토픽모델링(Topic Modeling)을 실시하였다. 토픽모델링은 텍스트 마이닝(Text Mining) 분석방법 중 하나로, 대량의 비정형 텍스트 자료 속에서 잠재적으로 존재하는 주제 구조를 탐색하는 내용 분석 방법이다(Maier et al., 2018). 그중 본 연구에서 사용한 LDA(Latent Dirichlet Allocation) 알고리즘은 문서 집합 내에서 단어들이 함께 출현하는 확률적 구조를 바탕으로 문서에 내재된 주제를 추출하는 방법이다(왕서기, 김유미, 2022). 본 연구에서는 LDA 토픽모델링을 이용하여 국립발레단과 유니버설발레단 관련 유튜브 콘텐츠를 시청한 대중들의 댓글을 주제 분석함으로써, 두 발레단에 대해 대중들이 주목하는 주요 이슈를 도출하고자 하였다.

LDA 토픽모델링 기법은 토픽의 수를 몇 개로 설정하느냐에 따라 분석 결과에 영향을 미치게 되므로, 적절한 토픽 수를 결정하는 것이 중요하다. 본 연구에서는 적절한 토픽 수를 탐색하기 위하여 토픽 수를 2개부터 10개까지 설정하여 반복 분석하였으며, 토픽 수 선정의 객관성을 높이기 위해 c_v coherence, u_mass coherence, c_npmi coherence 및 $\log$ perplexity를 함께 산출하였다. c_v 와 c_npmi 는 값이 높을수록 토픽 내 주요 단어의 의미적 응집성이 높은 것으로 해석하였으며, u_mass 는 0에 가까울수록 양호한 것으로 보았다. $\log$ perplexity는 모델의 예측 적합도를 확인하기 위한 보조 지표로 활용하였다. 이후 복수 평가 지표와 토픽의 해석 가능성 및 두 집단 간 비교 가능성을 종합적으로 검토하여 최종 토픽 수를 결정하였다. 특히 토픽 수가 지나치게 적을 경우 세부 주제가 통합되고, 지나치게 많을 경우 유사한 주제가 분산될 가능성이 있으므로, 본 연구에서는 4개에서 8개 사이의 토픽 수를 중심으로 해석 가능성을 함께 검토하였다.

도출된 토픽은 각 토픽에 할당된 주요 단어 상위 10개를 참고하고, 해당 단어가 포함된 원댓글의 맥락을 함께 파악한 후 귀납적 추론에 근거하여 명명하였다(최원민, 이지향, 2020). 마지막으로 연구문제 3과 관련하여 국립발레단과 유니버설발레단에서 도출된 토픽을 비교하였다. 비교는 각 집단에서 공통적으로 나타나는 관심 주제와 특정 집단에서 상대적으로 두드러지는 차별적 주제를 구분하는 방식으로 수행하였다. 이를 통해 유튜브 댓글에 나타난 두 발레단의 대중 인식, 콘텐츠 소비 양상, 브랜드 이미지의 공통점과 차이점을 해석하고자 하였다.

토픽모델링 결과에서 도출된 토픽 간 구조적 관계와 최종 토픽 수의 해석 가능성을 보완적으로 확인하기 위하여 토픽 간 거리 지도를 제시하였다. 토픽 간 거리 지도는 각 토픽의 단어 가중치 분포를 기준으로 토픽 간 유사도와 차이를 산출하고, 이를 2차원 공간에 배치하는 방식으로 구성하였다. 본 연구에서는 토픽별 상위 키워드의 가중치 분포를 이용하여 Jensen-Shannon distance를 계산한 뒤, 다차원척도법(Multidimensional Scaling, MDS)을 적용하여 토픽 간 상대적 거리를 시각화하였다. 지도에서 원의 크기는 해당 토픽의 비중을 의미하며, 원 사이의 거리가 가까울수록 두 토픽의 핵심 키워드 구성이 유사하고, 멀수록 서로 다른 주제 구조를 갖는 것으로 해석하였다. 또한 토픽과 핵심 키워드 간의 연결 구조를 직관적으로 파악하기 위하여 토픽-키워드 네트워크 분석을 실시하였다. 토픽-키워드 네트워크 분석은 LDA 토픽모델링에서 도출된 각 토픽과 상위 10개 키워드의 연결 관계를 바탕으로 수행하였다. 네트워크는 토픽과 키워드로 구성된 이원 네트워크(two-mode network)로 구축하였으며, 토픽 노드는 사각형, 키워드 노드는 원으로 표시하였다. 토픽과 키워드 사이의 링크는 해당 키워드가 특정 토픽에서 차지하는 조건부 확률값 $p(\text{word}|\text{topic})$ 을 기준으로 설정하였고, 이 값을 키워드 가중치로 사용하였다. 가중치가 높을수록 해당 키워드가 특정 토픽을 설명하는 상대적 중요도가 높은 것으로 해석하였다.

III. 연구 결과

1. 단어 빈도분석

국립발레단과 유니버설발레단에 관련된 유튜브 콘텐츠 댓글에서 대중들이 자주 언급하는 단어들을 파악하기 위하여 단어의 출현 빈도를 분석하였다. 그 결과 각 집단의 댓글에서 가장 많이 등장한 상위 50개의 빈출 단어를 <표 1>과 같이 나타냈다. 또한 단어 빈도분석 결과를 바탕으로 워드 클라우드(Word Cloud)를 실행하여 전체 단어 중 가장 자주 언급한 단어들을 한눈에 파악할 수 있도록 시각화하여 <그림 1>과 같이 제시하였다.

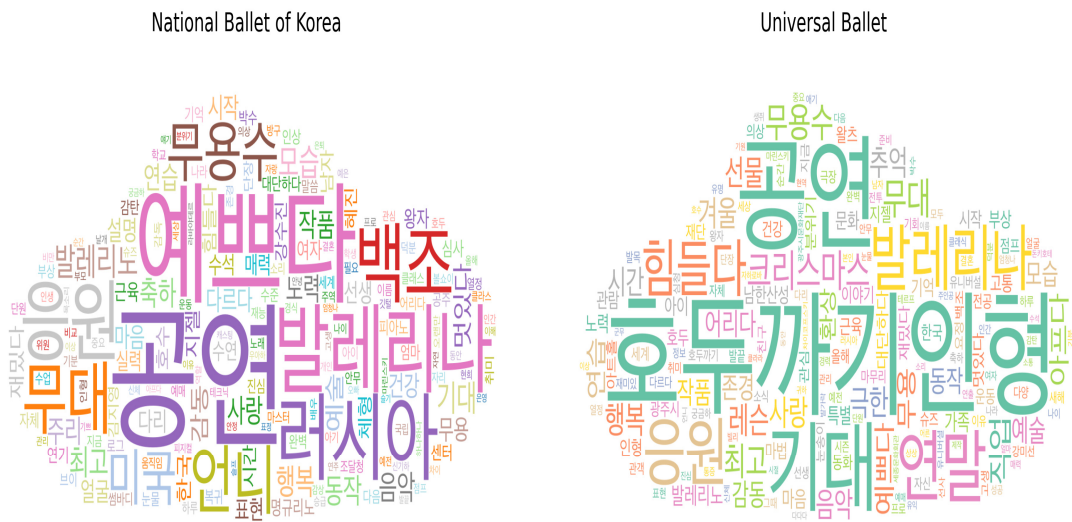


그림 1. 국립발레단과 유니버설발레단 관련 유튜브 댓글의 상위 50개 빈출단어

단어 빈도분석을 실시한 결과, 국립발레단 관련 댓글에서는 ‘공연’, ‘예쁘다’, ‘발레리나’, ‘러시아’, ‘응원’, ‘백조’,

‘무대’, ‘언니’, ‘무용수’, ‘발레리노’, ‘미국’, ‘모습’, ‘기대’, ‘최고’, ‘행복’, ‘동작’, ‘예술’, ‘음악’, ‘축하’, ‘감동’, ‘작품’, ‘멋있다’, ‘사랑’, ‘주리’, ‘연습’, ‘노력’, ‘다르다’, ‘건강’, ‘시작’, ‘무용’, ‘마음’, ‘설명’, ‘한국’, ‘재밌다’, ‘지젤’, ‘얼굴’, ‘다리’, ‘힘들다’, ‘표현’, ‘선생’, ‘호수’, ‘체형’, ‘남자’, ‘시간’, ‘매력’, ‘수석’, ‘혜진’, ‘실력’, ‘여자’, ‘강수진’이 상위 50개 빈출 단어로 나타났다.

유니버설발레단 관련 댓글에서는 ‘호두까기인형’, ‘공연’, ‘기대’, ‘연말’, ‘응원’, ‘발레리나’, ‘힘들다’, ‘크리스마스’, ‘무대’, ‘무용수’, ‘직업’, ‘최고’, ‘행복’, ‘레슨’, ‘겨울’, ‘극한’, ‘선물’, ‘사랑’, ‘아프다’, ‘예쁘다’, ‘추억’, ‘무용’, ‘연습’, ‘음악’, ‘시간’, ‘어리다’, ‘작품’, ‘예술’, ‘감동’, ‘동작’, ‘환상’, ‘모습’, ‘존경’, ‘발레리노’, ‘가족’, ‘대단하다’, ‘마음’, ‘노력’, ‘특별’, ‘문화’, ‘아이’, ‘기억’, ‘남한산성’, ‘분위기’, ‘고통’, ‘관람’, ‘부상’, ‘멋있다’, ‘시작’, ‘재밌다’가 상위 50개 빈출 단어로 나타났다.

이러한 결과를 통해 국립발레단 댓글에서는 공연 감상, 무용수, 해외 발레, 응원과 관련된 단어가 두드러지는 반면, 유니버설발레단 댓글에서는 ‘호두까기인형’, ‘연말’, ‘크리스마스’, ‘선물’ 등 특정 레퍼토리와 시즌 공연을 중심으로 한 단어가 상대적으로 높게 나타났음을 확인할 수 있다.

표 1. 국립발레단과 유니버설발레단 관련 유튜브 댓글 빈도분석 결과

순번	국립발레단 단어	국립발레단 빈도	유니버설발레단 단어	유니버설발레단 빈도
1	공연	285	호두까기인형	737
2	예쁘다	254	공연	628
3	발레리나	250	기대	321
4	러시아	244	연말	185
5	응원	202	응원	162
6	백조	195	발레리나	147
7	무대	160	힘들다	141
8	언니	144	크리스마스	131
9	무용수	127	무대	126
10	발레리노	108	무용수	109
11	미국	108	직업	103
12	모습	106	최고	100
13	기대	99	행복	83
14	최고	97	레슨	83
15	행복	91	겨울	76
16	동작	86	극한	75
17	예술	84	선물	74
18	음악	83	사랑	73
19	축하	82	아프다	73
20	감동	79	예쁘다	72
21	작품	78	추억	69
22	멋있다	75	무용	68
23	사랑	74	연습	67
24	주리	72	음악	65
25	연습	69	시간	63
26	노력	69	어리다	63

순번	국립발레단 단어	국립발레단 빈도	유니버설발레단 단어	유니버설발레단 빈도
27	다르다	63	작품	62
28	건강	62	예술	62
29	시작	61	감동	60
30	무용	59	동작	59
31	마음	58	환상	57
32	설명	56	모습	56
33	한국	55	존경	55
34	재밌다	54	발레리노	54
35	지젤	53	가족	51
36	얼굴	49	대단하다	48
37	다리	48	마음	48
38	힘들다	48	노력	46
39	표현	47	특별	46
40	선생	46	문화	45
41	호수	45	아이	44
42	체형	45	기억	42
43	남자	43	남한산성	39
44	시간	43	분위기	39
45	매력	42	고통	39
46	수석	42	관람	38
47	혜진	41	부상	38
48	실력	40	멋있다	38
49	여자	39	시작	38
50	강수진	39	재밌다	37

2. 토픽모델링

토픽모델링 분석에 앞서 적절한 토픽 개수를 선정하기 위하여 토픽모델링 평가(Evaluation of Topic Models)를 실행하였다. 본 연구에서는 토픽 수를 2개부터 10개까지 설정하여 각 집단별로 반복 분석하였으며, 토픽 수 선정의 객관성을 높이기 위하여 c_v coherence, u_mass coherence, c_npmi coherence, log perplexity를 함께 검토하였다. 토픽의 일관성 지표인 c_v 와 c_npmi 는 값이 높을수록 토픽 내 주요 단어의 의미적 응집성이 높은 것으로 해석하였으며, u_mass 는 0에 가까울수록 양호한 것으로 보았다. log perplexity는 모델의 예측 적합도를 확인하기 위한 보조 지표로 활용하였다.

토픽모델링 평가 결과는 <그림 2>와 같이 나타났다. c_v coherence를 기준으로 할 경우 국립발레단은 토픽 수 3개에서 0.417로 가장 높은 값을 보였으며, 유니버설발레단은 토픽 수 2개에서 0.655로 가장 높은 값을 보였다. u_mass 와 c_npmi 역시 전반적으로 낮은 토픽 수에서 상대적으로 양호한 값을 보였으나, log perplexity는 토픽 수가 증가할수록 개선되는 경향을 보였다. 이처럼 지표별 최적 지점이 일치하지 않았기 때문에 본 연구에서는 단일 지표만으로 토픽 수를 결정하지 않고, 복수 평가 지표와 토픽의 해석 가능성, 그리고 두 집단 간 비교 가능성을 종합적으로 검토하였다.

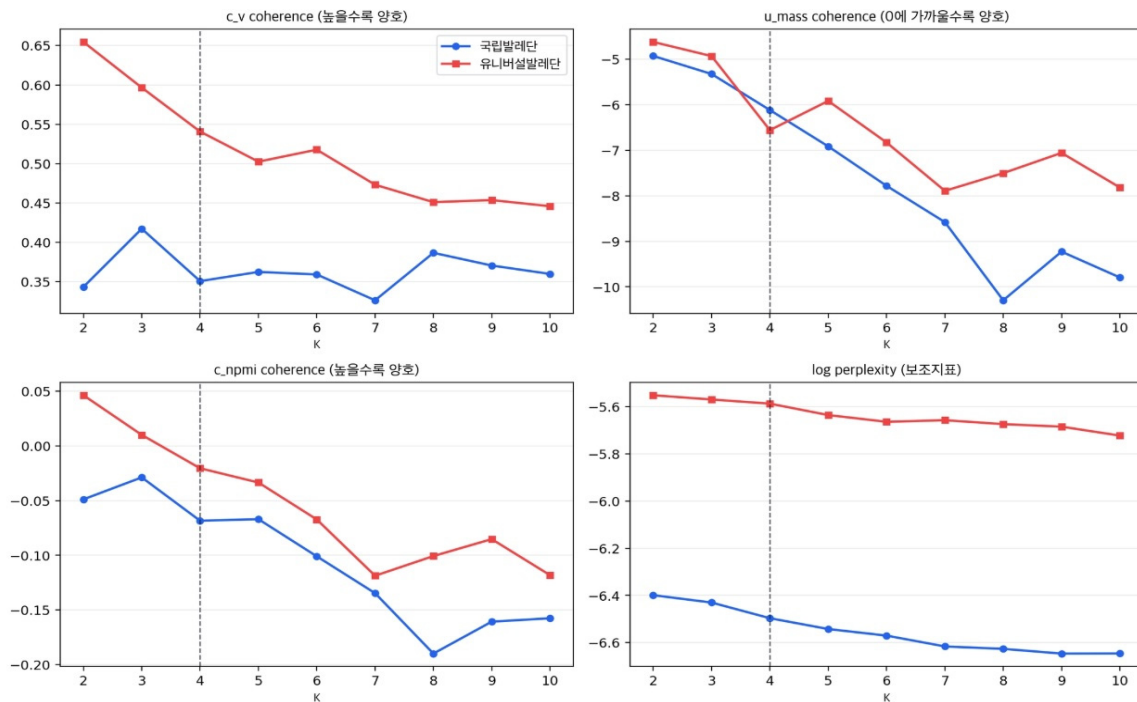


그림 2. 토픽모델링 복수지표 평가 결과

특히 토픽 수가 2개 또는 3개일 경우 세부 주제가 과도하게 통합되는 경향이 나타났다. 유니버설발레단의 경우 토픽 수 2개에서는 ‘호두까기인형’과 ‘공연’ 관련 주제가 하나의 큰 축으로 묶이고, 나머지 주제가 포괄적으로 결합되어 세부 관심 주제를 파악하는 데 한계가 있었다. 반대로 토픽 수가 5개 이상으로 증가할 경우 유사한 주제가 여러 토픽으로 분산되는 경향이 확인되었다. 이에 따라 본 연구에서는 복수 지표의 결과와 원댓글 맥락 검토를 종합하여 국립발레단과 유니버설발레단 모두 최종 토픽 수를 4개로 결정하였다. 최종 모델의 c_v 값은 국립발레단 0.351, 유니버설발레단 0.541로 나타났다.

최종 토픽 수의 해석 가능성을 보완적으로 확인하기 위하여 토픽 간 거리 지도를 <그림 3>과 같이 제시하였다. 토픽 간 거리 지도는 각 토픽의 단어 확률분포를 기준으로 토픽 간 거리를 산출한 뒤 이를 2차원 공간에 배치한 것이다. 국립발레단의 4개 토픽은 서로 과도하게 중첩되지 않고 비교적 구분되어 나타났다. 유니버설발레단에서도 4개 토픽이 전반적으로 구분되었으나, 일부 토픽은 상대적으로 인접하게 나타났다. 이러한 결과는 최종적으로 선정한 4개 토픽이 두 집단의 주요 반응 구조를 비교하고 해석하는 데 적절한 수준의 구분 가능성과 해석 가능성을 갖는 것으로 판단할 수 있음을 보여준다.

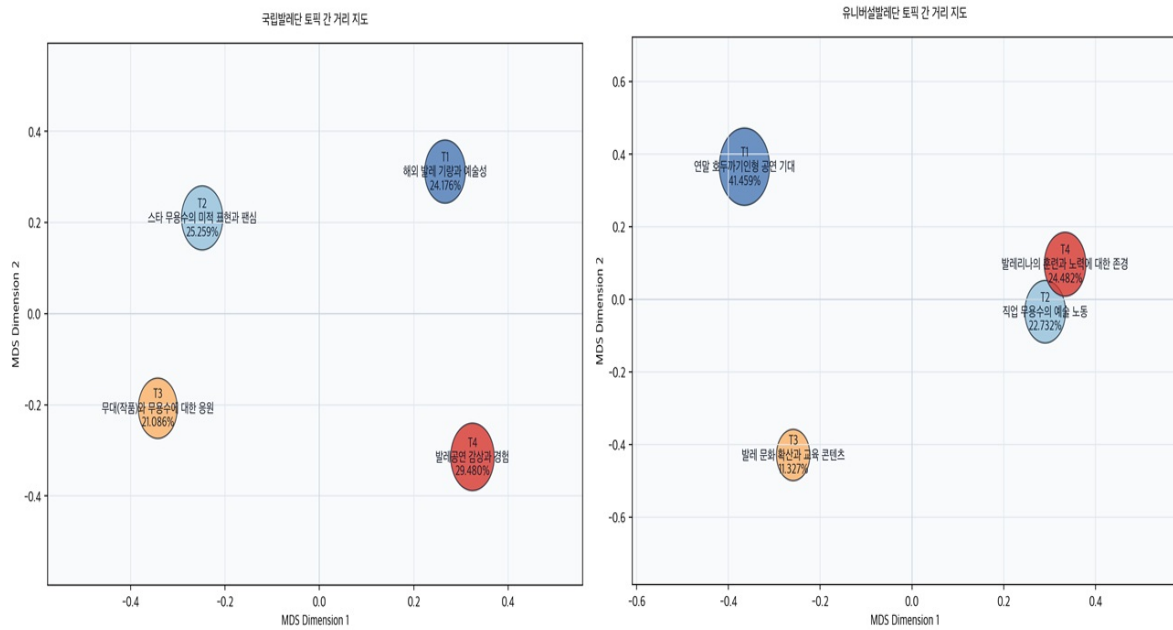


그림 3. 국립발레단과 유니버설발레단 토픽 간 거리 지도

토픽모델링을 실행한 후 각 토픽에 할당될 확률이 높은 상위 10개의 단어를 중심으로 토픽을 해석하였다. 도출된 토픽모델링 결과에서 토픽명의 타당성을 높이기 위하여 각 토픽에 포함된 주요 단어와 해당 단어가 포함된 원 댓글의 맥락을 함께 확인하였다. 이후 무용학 박사과 데이터분석 전문가의 협의를 거쳐 각 토픽의 의미를 검토하였으며, 귀납적 추론에 근거하여 최종 토픽명을 명명하였다.

그 결과 국립발레단 댓글에서는 총 4개의 토픽이 도출되었다. 가장 높은 비중을 차지한 토픽은 ‘발레공연 감상과 경험’으로 29.480%를 차지하였으며, 다음으로 ‘스타 무용수의 미적 표현과 팬심’ 25.259%, ‘해외 발레 기량과 예술성’ 24.176%, ‘무대(작품)과 무용수에 대한 응원’ 21.086%의 순으로 나타났다.

유니버설발레단 댓글에서도 총 4개의 토픽이 도출되었다. 가장 높은 비중을 차지한 토픽은 ‘연말 호두까기인형 공연 기대’로 41.459%를 차지하였으며, 다음으로 ‘발레리나의 훈련과 노력에 대한 존경’ 24.482%, ‘직업 무용수의 예술 노동’ 22.732%, ‘발레 문화 확산과 교육 콘텐츠’ 11.327%의 순으로 나타났다.

이러한 결과를 통해 국립발레단 댓글은 공연 감상, 무용수 팬덤, 해외 발레와의 비교, 무대와 무용수에 대한 응원 등 비교적 다양한 주제로 분산되어 나타난 반면, 유니버설발레단 댓글은 ‘호두까기인형’을 중심으로 한 연말 공연 기대와 감상 반응이 가장 두드러지는 특징을 보였다.

3. 토픽-키워드 네트워크 그래프

토픽모델링 결과를 보다 직관적으로 확인하기 위하여 각 집단별 토픽과 상위 키워드 간의 연결 관계를 네트워크 그래프로 제시하였다. 연결선의 두께는 토픽과 키워드 간의 연결 가중치를 나타내며, 선이 두꺼울수록 해당 키워드가 특정 토픽을 설명하는 상대적 중요도가 높은 것으로 해석하였다. 또한 네트워크에서 나타난 연결 관계를 정량적으로 확인할 수 있도록 각 토픽별 상위 키워드와 가중치를 표로 함께 제시하였다.

1) 국립발레단 토픽-키워드 네트워크

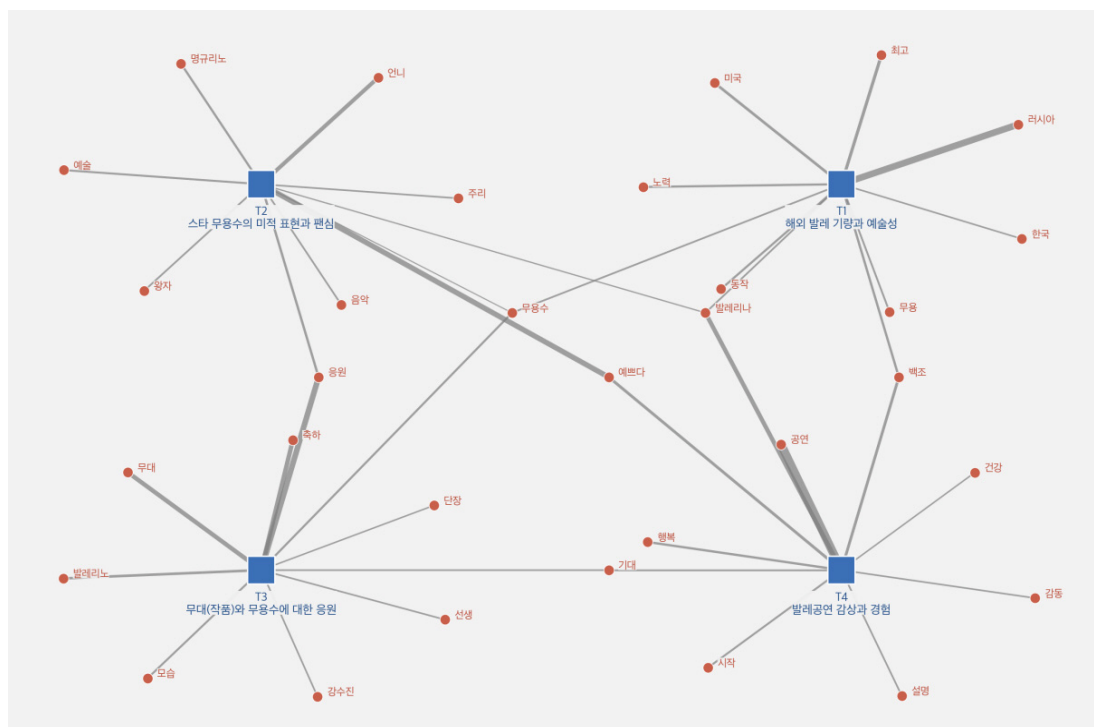


그림 4. 국립발레단 토픽-키워드 네트워크

국립발레단의 토픽-키워드 네트워크 분석 결과는 <그림 4>와 같이 나타났다. 네트워크에서는 ‘공연’, ‘발레리나’, ‘응원’, ‘러시아’, ‘백조’ 등의 키워드가 주요 토픽을 구성하였다. 특히 공연 감상과 발레 경험 관련 토픽의 비중이 가장 높게 나타났으며, 스타 무용수에 대한 미적 감상과 팬심, 해외 발레 기량 및 예술성에 대한 평가, 무대와 무용수에 대한 응원이 비교적 균형 있게 분포하였다.

국립발레단 관련 유튜브 댓글의 토픽모델링 결과는 <표 2>와 같이 총 4개의 토픽으로 도출되었다. 가장 높은 비중을 차지한 토픽은 ‘발레공연 감상과 경험’으로 29.480%를 차지하였으며, 다음으로 ‘스타 무용수의 미적 표현과 팬심’ 25.259%, ‘해외 발레 기량과 예술성’ 24.176%, ‘무대(작품)와 무용수에 대한 응원’ 21.086%의 순으로 나타났다.

표 2. 국립발레단 관련 유튜브 댓글 토픽모델링 결과

토픽	국립발레단 토픽명	비중	주요 키워드(가중치)
1	해외 발레 기량과 예술성	24.176%	러시아(0.0544), 최고(0.0276), 미국(0.0240), 백조(0.0213), 노력(0.0199), 동작(0.0195), 무용(0.0175), 발레리나(0.0163), 무용수(0.0163), 한국(0.0150)
2	스타 무용수의 미적 표현과 팬심	25.259%	예쁘다(0.0443), 언니(0.0341), 응원(0.0216), 명규리노(0.0202), 예술(0.0181), 왕자(0.0167), 음악(0.0165), 주리(0.0156), 발레리나(0.0125), 무용수(0.0111)
3	무대(작품)와 무용수에 대한 응원	21.086%	응원(0.0481), 축하(0.0387), 무대(0.0371), 발레리노(0.0229), 무용수(0.0193), 모습(0.0185), 기대(0.0157), 강수진(0.0155), 선생(0.0152), 단장(0.0142)
4	발레공연 감상과 경험	29.480%	공연(0.0715), 발레리나(0.0370), 예쁘다(0.0252), 백조(0.0250), 행복(0.0206), 시작(0.0181), 기대(0.0168), 설명(0.0160), 감동(0.0150), 건강(0.0142)

2) 유니버설발레단 토픽-키워드 네트워크

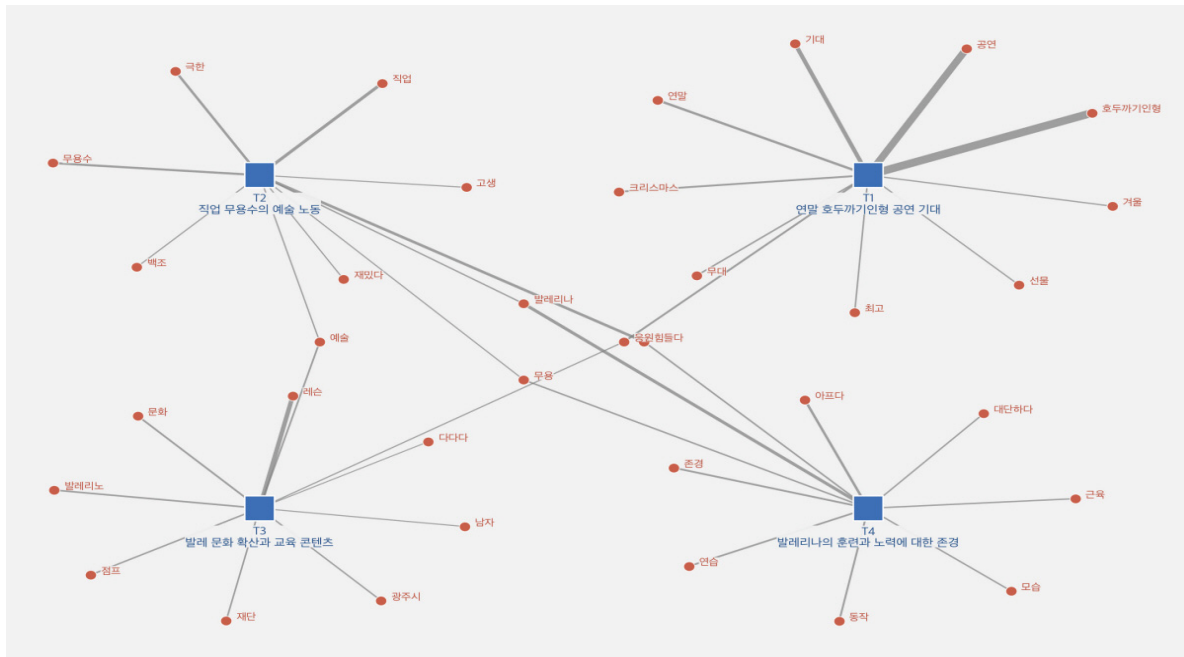


그림 5. 유니버설발레단 토픽-키워드 네트워크

유니버설발레단의 토픽-키워드 네트워크 분석 결과는 <그림 5>와 같이 나타났다. 네트워크에서는 ‘호두까기인형’, ‘공연’, ‘기대’, ‘연말’, ‘크리스마스’ 등의 키워드가 가장 강하게 연결되었다. 이는 유니버설발레단 댓글에서 연말 시즌 레퍼토리인 「호두까기인형」에 대한 기대와 감상 반응이 중심으로 나타났음을 보여준다. 또한 직업 무용수의 노동과 고충, 발레리나의 훈련과 노력에 대한 존경, 레슨 및 교육 콘텐츠를 통한 발레 문화 확산이라는 주제가 함께 도출되었다.

유니버설발레단 관련 유튜브 댓글의 토픽모델링 결과는 <표 3>과 같이 총 4개의 토픽으로 도출되었다. 가장 높은 비중을 차지한 토픽은 ‘연말 호두까기인형 공연 기대’로 41.459%를 차지하였으며, 다음으로 ‘발레리나의 훈련과 노력에 대한 존경’ 24.482%, ‘직업 무용수의 예술 노동’ 22.732%, ‘발레 문화 확산과 교육 콘텐츠’ 11.327%의 순으로 나타났다.

표 3. 유니버설발레단 관련 유튜브 댓글 토픽모델링 결과

토픽	유니버설발레단 토픽명	비중	주요 키워드(가중치)
1	연말 호두까기인형 공연 기대	41.459%	호두까기인형(0.1399), 공연(0.1137), 기대(0.0601), 연말(0.0352), 응원(0.0255), 크리스마스(0.0249), 무대(0.0189), 최고(0.0159), 선물(0.0141), 겨울(0.0140)
2	직업 무용수의 예술 노동	22.732%	직업(0.0457), 힘들다(0.0404), 극한(0.0334), 무용수(0.0312), 발레리나(0.0155), 예술(0.0130), 백조(0.0121), 무용(0.0117), 재밌다(0.0115), 고생(0.0107)
3	발레 문화 확산과 교육 콘텐츠	11.327%	레슨(0.0637), 문화(0.0226), 발레리노(0.0208), 예술(0.0197), 점프(0.0183), 재단(0.0171), 광주시(0.0161), 응원(0.0122), 남자(0.0104), 다다다(0.0094)
4	발레리나의 훈련과 노력에 대한 존경	24.482%	발레리나(0.0468), 아프다(0.0323), 존경(0.0240), 연습(0.0223), 동작(0.0217), 힘들다(0.0193), 무용(0.0186), 모습(0.0168), 근육(0.0164), 대단하다(0.0162)

IV. 논의

본 연구는 국립발레단과 유니버설발레단의 유튜브 콘텐츠 댓글을 대상으로 온라인 이용자 반응의 차이를 비교·분석하였다. 분석 결과, 두 발레단 모두 공연, 무용수, 무대, 작품, 감동, 응원 등 발레 감상과 관련된 공통 반응이 확인되었다. 이는 유튜브 댓글이 단순한 감상 표현을 넘어 공연예술에 대한 감정, 평가, 지지, 비교, 기억이 축적되는 디지털 수용 자료로 기능할 수 있음을 시사한다. 나아가 이용자들은 발레 콘텐츠를 수동적으로 소비하는 데 그치지 않고, 작품과 무용수의 수행을 해석하고 평가하는 반응을 보였다.

먼저 국립발레단 관련 댓글에서는 예술성, 무용수의 기량, 해외 발레와의 비교, 특정 무용수에 대한 팬심이 두드러지게 나타났다. 단어 빈도분석에서는 ‘공연’, ‘발레리나’, ‘러시아’, ‘백조’, ‘무용수’, ‘예술’, ‘실력’ 등이 상위에 위치했으며, 토픽모델링에서도 ‘해외 발레 기량과 예술성’, ‘스타 무용수의 미적 표현과 팬심’, ‘무대와 무용수에 대한 응원’, ‘발레공연 감상과 경험’이 도출되었다. 이러한 결과는 국립발레단에 대한 온라인 반응이 고전발레의 형식미, 무용수의 기술적 완성도, 국제적 비교 의식, 특정 무용수에 대한 팬심이 결합된 방식으로 형성되고 있음을 시사한다. 특히 ‘동작’, ‘실력’, ‘예술’, ‘무용수’ 등 수행과 완성을 평가하는 언어가 함께 나타난 점은, 댓글이 단순한 호감 표현을 넘어 예술적 기준에 기반한 평가 담론을 포함하고 있음을 보여준다. 이는 발레공연 관객만족도에서 안무와 무용수 기량이 핵심 요인으로 논의된 나보리, 금현섭(2022)의 연구와도 맥락을 같이한다.

이러한 양상은 국립발레단의 제도적 위치와 상징성을 고려할 때 하나의 해석적 맥락에서 이해될 수 있다. 양영은(2016)은 국립발레단을 국가 재정지원을 받는 대표 발레단으로 규정하며, 해외 레퍼토리 수용과 창작발레 개발을 병행해 온 단체로 설명하였다. 이 관점에서 댓글에 나타난 ‘러시아’, ‘미국’, ‘예술’, ‘실력’ 등의 키워드는 이용자들이 국립발레단을 국제적 기준과 비교하여 인식하는 경향이 나타났음을 시사한다. 다만 본 연구는 댓글 텍스트 분석에 기반하므로, 제도적 성격이 이용자 반응을 직접적으로 형성하였다고 단정하기보다는 이러한 요소들이 해석의 맥락으로 작용했을 가능성으로 이해하는 것이 적절하다.

반면 유니버설발레단 관련 댓글에서는 「호두까기인형」을 중심으로 한 시즌 레퍼토리 반응이 뚜렷하게 나타났다. 단어 빈도분석에서 ‘호두까기인형’, ‘연말’, ‘크리스마스’, ‘겨울’, ‘가족’, ‘추억’ 등이 높은 빈도를 보였으며, 토픽모델링에서도 ‘연말 호두까기인형 공연 기대’가 핵심 주제로 도출되었다. 이는 「호두까기인형」이 단일 공연을 넘어 시즌성, 가족 관람, 정서적 기억이 결합된 문화적 경험으로 수용되고 있음을 의미한다. 노영재(2010)가 해당 작품을 크리스마스 시즌과 결합된 사회문화적 레퍼토리로 설명한 점 역시 이러한 해석을 뒷받침한다.

유니버설발레단의 댓글 반응은 레퍼토리 구성과 반복된 공연 경험과 관련하여 형성된 수용 양상으로 해석할 수 있다. 정옥희(2025)는 유니버설발레단의 형성을 무용수, 레퍼토리, 관객이 상호작용하며 구축된 과정으로 설명하였다. 이러한 관점에서 보면 「호두까기인형」과 연말 공연에 대한 높은 기대는 반복 공연과 관련하여 특정 작품이 이용자의 정서적 기억과 관람 경험 속에 자리 잡은 양상으로 해석할 수 있다. 다만 이러한 해석 역시 특정 브랜드 전략이나 관객 개발 정책의 직접적인 효과로 단정하기보다는, 콘텐츠 특성과 공연 경험이 복합적으로 반영된 양상으로 해석하는 것이 적절하다. 또한 분석 대상 콘텐츠의 레퍼토리 구성 자체가 특정 주제의 빈도를 높였을 가능성도 함께 고려할 필요가 있다.

신문기사와 공연평은 직접적인 분석 대상이 아니라 두 발레단의 차이와 댓글 결과를 이해하기 위한 보조적 맥락 자료로 활용하였다. 관련 보도와 비평은 국립발레단을 공공성, 예술적 완성도, 국제 경쟁력의 측면에서 다루는 반면, 유니버설발레단은 「호두까기인형」의 시즌성, 시각적 화려함, 가족 관람 경험을 중심으로 설명하는 경향을 보인다(이은성, 2026; 장광열, 2022; 장병호, 2022; 박성준, 2025; 최희정, 2025). 이러한 담론은 본 연구에서 확인된 댓글 반응과 일정 부분 유사한 양상을 보이나, 생성 주체와 목적이 상이하므로 직접적인 영향관계로 해석하기

보다는 보조적 해석 맥락으로 이해할 필요가 있다.

종합하면 두 발레단은 모두 고전발레와 무용수를 중심으로 반응이 형성되었지만, 이용자가 주목하는 감상 지점에서는 차이가 나타났다. 국립발레단 댓글은 작품 구조, 동작, 기량, 예술성, 국제 비교에 집중된 반면, 유니버설발레단 댓글은 시즌성, 가족 관람, 정서적 기억과 긴밀하게 결합되어 있었다. 이러한 결과는 동일한 발레 장르 내에서도 레퍼토리 구성과 콘텐츠 특성에 따라 이용자가 주목하는 감상 요소가 달라질 수 있음을 시사한다. 본 연구에서 ‘평가 중심 수용’은 기량, 예술성, 수행 수준 및 비교와 관련된 반응이 상대적으로 강조되는 양상을 의미하며, ‘경험 중심 수용’은 시즌성, 관람 기억, 가족 관람 및 정서적 공감이 상대적으로 강조되는 양상을 의미한다. 이에 본 연구에서는 국립발레단의 수용 양상을 ‘평가 중심 수용’, 유니버설발레단의 수용 양상을 ‘경험 중심 수용’으로 개념화하였다. 다만 이러한 구분은 절대적인 특성이 아니라, 본 연구 자료에서 나타난 상대적 경향을 설명하기 위한 분석적 구분이다.

무용수에 대한 반응에서도 공통성과 차이가 동시에 확인되었다. 두 발레단 모두 무용수의 기량과 노력에 대한 관심이 나타났다는 점에서, 이용자들은 발레를 작품 단위에 한정하지 않고 수행 과정까지 함께 인식하고 있었다. 그러나 국립발레단 댓글에서는 특정 무용수의 실력, 외형적 아름다움, 미적 표현, 특정 무용수에 대한 응원과 팬심이 강조된 반면, 유니버설발레단 댓글에서는 훈련, 부상, 고통 등 직업적 현실에 대한 공감이 상대적으로 두드러졌다. 이는 무용수의 신체가 예술적 표현의 매체이자 동시에 노동의 결과물로 인식된다는 한혜원, 김영미(2025)의 논의와도 연결된다.

또한 본 연구는 유튜브가 무용공연의 감상과 해석, 비교와 평가가 이루어지는 디지털 환경으로 확장되고 있음을 보여준다. 김태연(2024)은 공공무용단 유튜브 콘텐츠의 구성 요소와 댓글을 분석하여 이용자 반응 특성을 도출하였으며, 이는 본 연구의 접근과 맥을 같이한다. 더불어 홍미성, 김가은(2020)의 연구에서 나타난 기능학습 및 작품 분석 중심 이용 양상은 본 연구 댓글에서 확인된 ‘동작’, ‘실력’, ‘작품’ 등의 반응과도 연결된다.

한편 본 연구는 댓글 텍스트 분석에 기반한 탐색적 연구로서, 제도적 특성이나 레퍼토리 전략과 이용자 반응 간의 관계를 직접적인 인과로 설명하기보다는 맥락적 관련성의 수준에서 이해할 필요가 있다.

실천적 측면에서 두 발레단의 온라인 콘텐츠 전략은 이러한 반응 차이를 반영해 차별화될 필요가 있다. 국립발레단은 예술성, 기량, 고전 레퍼토리, 국제 비교에 대한 관심이 높은 만큼 작품 해설, 레퍼토리 계보, 무용수 인터뷰, 리허설 과정 등 전문적 이해를 확장하는 콘텐츠를 강화할 필요가 있다. 이는 공공 발레단으로서 예술적 가치와 교육적 기능을 동시에 수행하는 전략이 될 수 있다. 반면 유니버설발레단은 「호두까기인형」을 중심으로 형성된 시즌성과 정서적 경험을 기반으로 가족 관객 대상 콘텐츠, 관람 가이드, 작품 이해 자료, 무용수의 직업 세계를 보여주는 콘텐츠를 강화할 필요가 있다. 이는 레퍼토리 기반 관람 경험을 온라인에서 지속적으로 확장하는 데 기여할 수 있다. 다만 이러한 제언은 댓글 분석에서 나타난 관심 주제를 바탕으로 한 가능성 제시이며, 실제 효과는 이용자 조사와 콘텐츠 성과 분석 등을 통해 검증할 필요가 있다.

V. 결론 및 제언

본 연구는 국립발레단과 유니버설발레단 관련 유튜브 콘텐츠 댓글을 대상으로 온라인 이용자 반응의 차이를 비교·분석하는 것을 목적으로 하였다. 이를 위해 2015년부터 2025년까지 게시된 관련 유튜브 콘텐츠에서 총 902개 동영상과 10,621개의 원댓글을 수집하였다. 분석방법으로 단어 빈도분석과 LDA 토픽모델링을 실시하고, 토픽 간 거리 지도와 토픽-키워드 네트워크 분석을 통해 주요 키워드, 잠재적 주제 구조 및 토픽과 핵심 키워드 간의 연결

관계를 분석하였다.

본 연구의 결과를 연구문제에 따라 정리하면 다음과 같다.

첫째, 주요 단어의 특성에서는 두 발레단의 공통점과 차이가 함께 확인되었다. 국립발레단과 유니버설발레단 댓글에서는 공연, 무용수, 무대, 작품, 감동, 응원 등 발레 감상과 관련된 단어가 공통적으로 나타났다. 이는 유튜브 댓글이 단순한 감상 표현의 공간을 넘어 공연예술에 대한 감정, 평가, 지지 및 기억이 축적되는 디지털 수용 자료로 활용될 수 있음을 시사한다. 한편 국립발레단 댓글에서는 ‘러시아’, ‘백조’, ‘예술’, ‘실력’ 등 고전발레의 형식미, 무용수의 기량 및 해외 발레와의 비교를 반영하는 단어가 두드러졌다. 반면 유니버설발레단 댓글에서는 ‘호두까기인형’, ‘연말’, ‘크리스마스’, ‘가족’, ‘추억’ 등 특정 레퍼토리와 계절적 관람 경험을 나타내는 단어가 상대적으로 자주 나타났다.

둘째, 토픽 구조에서도 두 발레단의 상대적인 차이가 확인되었다. 국립발레단 댓글에서는 ‘해외 발레 기량과 예술성’, ‘스타 무용수의 미적 표현과 팬심’, ‘무대와 무용수에 대한 응원’, ‘발레공연 감상과 경험’이 주요 토픽으로 도출되었다. 이는 국립발레단 관련 댓글에서 예술적 완성도, 무용수의 기술적 수행, 국제 비교 및 특정 무용수에 대한 팬심이 상대적으로 강조되었음을 보여준다. 반면 유니버설발레단 댓글에서는 ‘연말 호두까기인형 공연 기대’, ‘발레리나의 훈련과 노력에 대한 존경’, ‘직업 무용수의 예술 노동’, ‘발레 문화 확산과 교육 콘텐츠’가 주요 토픽으로 나타났다. 유니버설발레단 관련 댓글에서는 시즌 레퍼토리와 정서적 경험, 무용수의 훈련과 신체적 부담 및 발레 문화의 접근성에 대한 관심이 상대적으로 두드러졌다. 또한 토픽 간 거리 지도를 통해 국립발레단의 네 개 토픽은 비교적 구분되어 나타났고, 유니버설발레단은 일부 토픽이 2차원 공간상에서 상대적으로 인접하게 나타났다.

셋째, 온라인 반응 양상에서는 두 발레단 간 상대적인 차이가 나타났다. 국립발레단 관련 댓글에서는 예술적 기준, 무용수의 기량 및 국제 비교를 중심으로 한 평가적 반응이 상대적으로 두드러졌다. 반면 유니버설발레단 관련 댓글에서는 시즌 레퍼토리, 가족 관람, 정서적 기억 및 무용수의 직업적 현실에 대한 경험적·공감적 반응이 상대적으로 두드러졌다. 본 연구에서 제시한 ‘평가 중심 수용’과 ‘경험 중심 수용’은 두 발레단을 절대적으로 구분하거나 모든 이용자의 인식을 대표하는 개념이 아니라, 본 연구에서 수집한 댓글에 나타난 반응의 상대적 강조점을 설명하기 위한 분석적 개념이다. 또한 이러한 차이는 발레단의 제도적 기반이나 브랜드 전략이 댓글 반응을 직접 결정했음을 의미하기보다, 각 발레단의 레퍼토리 구성, 콘텐츠 특성 및 축적된 단체 이미지가 온라인 이용자 반응과 복합적으로 관련되었을 가능성을 시사한다.

본 연구의 의의는 다음과 같다.

첫째, 동일한 발레 장르에 속한 국립발레단과 유니버설발레단을 비교함으로써 단일 기관 분석만으로는 파악하기 어려운 온라인 반응의 공통점과 차이점을 함께 제시하였다. 다만 본 연구는 두 발레단의 제도적 위치나 브랜드 전략이 댓글 반응의 차이를 직접 발생시켰음을 검증한 것이 아니라, 서로 다른 제도적·콘텐츠 맥락을 지닌 두 발레단의 온라인 수용 양상을 비교한 탐색적 연구라는 데 의미가 있다.

둘째, 유튜브 댓글을 공연예술 이용자의 디지털 반응 자료로 활용하여 발레 관객 연구의 범위를 공연장 중심의 관람 경험에서 온라인 플랫폼 기반의 감상과 평가로 확장하였다.

셋째, 텍스트마이닝과 LDA 토픽모델링을 활용하여 발레단별 온라인 반응 구조를 체계적으로 비교함으로써 무용학 분야의 디지털 관객 연구에 기초자료를 제공하였다.

넷째, 토픽 간 거리 지도와 토픽-키워드 네트워크를 함께 활용하여 토픽 간 상대적 관계와 핵심 키워드의 연결 가치를 보완적으로 제시하고, 분석 결과를 다각적으로 해석하였다는 점에서 방법론적 의의가 있다.

실천적 측면에서 본 연구는 두 발레단의 온라인 콘텐츠 운영 방향을 구체화하는 데 참고자료로 활용될 수 있다. 국립발레단은 예술성, 무용수의 기량, 고전 레퍼토리 및 국제 비교에 대한 이용자의 관심이 상대적으로 높게 나타

난 점을 고려하여 작품 해설, 레퍼토리의 역사와 계보, 무용수 인터뷰 및 리허설 과정 등 전문적 이해를 지원하는 콘텐츠를 강화할 수 있다. 반면 유니버설발레단은 「호두까기인형」을 중심으로 나타난 시즌성과 가족 관람 및 정서적 경험을 고려하여 가족 관객 대상 관람 가이드, 작품의 줄거리와 등장인물 소개, 연말 시즌 연계 콘텐츠 및 무용수의 직업 세계를 보여주는 콘텐츠를 확대할 수 있다. 이러한 제언은 두 발레단의 고정적인 콘텐츠 전략을 규정하는 것이 아니라, 본 연구에서 확인된 댓글 반응의 상대적 특징을 바탕으로 제시한 실천적 방향이다. 따라서 해당 콘텐츠 전략의 실제 효과는 향후 이용자 만족도, 콘텐츠 이용행태 및 관람 의도 등을 통해 검증할 필요가 있다.

그러나 본 연구는 다음과 같은 한계를 지닌다.

첫째, 유튜브 검색 결과와 댓글 수집 방식은 플랫폼의 검색·추천 알고리즘과 콘텐츠 노출 방식의 영향을 받을 수 있으므로 수집된 자료가 전체 온라인 반응을 대표한다고 보기는 어렵다. 아울러 본 연구는 유튜브 플랫폼만을 대상으로 하였으므로 다른 플랫폼에서 나타나는 이용자 반응의 특성을 반영하지 못하였다. 댓글 역시 적극적으로 참여한 일부 이용자의 의견이 상대적으로 많이 반영되는 자기선택적 자료라는 한계가 있다.

둘째, 본 연구는 한글 댓글만을 분석 대상으로 하였으므로 해외 이용자의 반응과 다국어 댓글의 특성을 충분히 반영하지 못하였다. 외국어 댓글은 언어별 형태소 분석체계 차이와 번역 과정에서의 의미 변형 가능성으로 인해 분석에서 제외되었으며, 이는 분석의 일관성을 위한 선택이지만 전체 반응 구조를 설명하는 데에는 한계가 있다.

셋째, 공식·비공식 채널, 공연 실황, 홍보, 인터뷰, 교육 콘텐츠 등 채널 및 콘텐츠 유형과 작품별 구성 비중을 세분화하여 분석하지 못하였다. 이에 따라 특정 키워드와 토픽의 집중도가 콘텐츠 구성과 관련되었을 가능성을 배제하기 어렵다.

넷째, 댓글 작성자의 인구사회학적 특성과 실제 관람 여부를 확인할 수 없어 결과를 전체 관객 인식으로 일반화하는 데에는 한계가 있다.

다섯째, LDA 토픽모델링의 특성상 토픽 명명 과정에는 연구자의 해석이 개입될 수 있으며, 이에 대한 신중한 해석이 요구된다.

후속연구에서는 채널 및 콘텐츠 유형에 따른 반응 차이를 세분화하고, 감성분석 및 시계열 분석을 결합하여 온라인 반응의 정서와 변화 양상을 입체적으로 분석할 필요가 있다. 또한 다국어 분석과 다양한 플랫폼 자료를 포함함으로써 공연예술의 디지털 수용 양상을 보다 포괄적으로 설명할 수 있을 것이다. 나아가 설문조사나 인터뷰를 병행하여 온라인 댓글과 실제 관람 경험 간의 관계를 규명하는 연구도 요구된다.

종합하면 국립발레단 관련 댓글에서는 예술성과 기량을 중심으로 한 평가적 반응이, 유니버설발레단 관련 댓글에서는 시즌 레퍼토리와 관람 경험을 중심으로 한 경험적 반응이 상대적으로 두드러졌다. 이는 두 발레단의 고정적 속성이나 제도적 기반의 직접적인 효과를 의미하는 것이 아니라, 본 연구 자료에서 확인된 상대적 경향이다. 본 연구는 공연예술 수용 연구의 범위를 온라인 플랫폼에서 축적되는 감상과 평가로 확장했다는 점에서 의의가 있다.

참고문헌

- 김가은, 홍미성(2020). 포스트 코로나 시대 '라이브 스트리밍(live streaming)' 무용공연의 시청현황과 개선방안. **한국체육학회지** 59(5) : 309-321.
- 김경식(2025). 유튜브 댓글의 도핑 담론에 대한 의미연결망 및 토픽모델링 분석. **융합과 통섭** 8(2) : 1-16.
- 김선영, 이의신(2017). 관객개발을 위한 공연예술 소비자 빅데이터 분석 사례 고찰. **한국산학기술학회논문지** 18(12) : 286-299.
- 김선정(2025). 유튜브 무용공연 시청동기, 시청만족, 지속시청의도 간 구조관계. **한국무용과학회지** 42(2) : 33-50.
- 김태연(2024). 공공무용단의 유튜브(YouTube) 콘텐츠 표현유형에 대한 이용자 반응 분석. **한국무용학회지** 24(1) : 123-140.
- 나보리, 금현섭(2022). 발레공연의 만족과 충성도 연구. **예술경영연구** 62 : 41-65.
- 나보리, 이슬기(2023). 온라인 공연예술 콘텐츠 지원정책에 대한 정책수요와 제도적 개선과제에 관한 연구. **문화정책논총** 37(2) : 161-191.
- 노영재(2010). 발레「호두까기 인형」의 변천에 따른 사회문화적 의미. **무용예술학연구** 29(29) : 1-26.
- 백민제(2021). 온택트(On tact) 시대, 공연예술 산업의 온라인 플랫폼 활용에 관한 연구. **연기예술연구** 22(2) : 59-71.
- 안성우(2025). Mnet 유튜브 댓글 분석을 통한 대중 인식 연구. **한국무용과학회지** 42(2) : 11-24.
- 양영은(2016). 문화예술정책과 국립발레단의 상호작용으로 형성·변화하는 국가 정체성. **무용예술학연구** 60(3) : 87-107.
- 왕서기, 김유미(2022). 텍스트 마이닝(Text-mining)을 이용한 댓글 분석 연구 - BTS의 'Butter'와 'Permission to Dance' 뮤직비디오 유튜브 댓글을 중심으로 -. **언어사실과 관점** 56 : 157-181.
- 염지현, 이해준(2026). LDA토픽모델링을 적용한 성인발레 콘텐츠 동향분석: 2018~2025 유튜브 데이터를 중심으로. **한국무용학회지** 25(4) : 73-85.
- 유진주(2025). 디지털전환시대의 국내 무용공연 경영·마케팅 연구에 대한 메타분석. **스포츠사이언스** 43(2) : 151-159.
- 이윤경(2023). 국립현대무용단에 관한 유튜브 콘텐츠 댓글 분석. **한국체육학회지** 62(4) : 43-52.
- 장유정(2023). 빅데이터를 통해 살펴본 코로나 팬데믹으로 인한 무용의 비대면화에 대한 사회적 담론 변화. **한국무용학회지** 22(4) : 69-77.
- 장지영(2021). 코로나시대 공연예술의 패러다임 전환 - 비대면 환경의 쟁점과 새로운 담론들. **학술대회** 2021(6) : 35-48.
- 정원준(2021). 코로나19 위험 커뮤니케이션 탐색 연구: 소셜 미디어 의존 이론을 응용한 정부와 일반인 유튜브 콘텐츠와 댓글 내 담론 분석. **PR연구** 25(2) : 59-101.
- 정옥희(2025). 발레단은 어떻게 만들어지는가 : 유니버설발레단의 '발레단 만들기'에 대한 역사쓰기. **한국무용학회지** 24(4) : 325-340.
- 정중은(2021). 코로나 시대, 공연 제작자와 관객의 '공연영상 플랫폼' 활용 경험: 설문조사 분석을 중심으로. **문화정책논총** 35(3) : 37-64.
- 조언위, 안남일(2023). 한국 국립발레단의 영상 콘텐츠 제작 활성화 방안. **한국예술연구** 41 : 195-214.
- 최원민, 이지향(2020). An Analysis of Research Trends in Korean Dance Psychology. **Research in Dance and Physical Activity**, 4(1) : 45-55.
- 한혜원, 김영미(2025). 빅데이터 기반 텍스트 마이닝을 활용한 무용학에서의 몸 담론에 대한 지식구조 분석 ; 2000년 이후 연구물 중심으로. **한국무용학회지** 24(4) : 107-123.
- 현보람(2022). 유튜브를 활용한 무용인식 분석. **한국융합과학회지** 11(3) : 244-258.
- 홍미성, 김가은(2020). Z세대 무용전공생의 유튜브(YouTube) 이용실태 분석. **한국무용과학회지** 37(2) : 1-20.
- 박성준(2025.10.20). 다시 돌아온 크리스마스 시즌 대표작 유니버설발레단의 '호두까기인형'. 『세계일보』, <https://www.segye.com/newsView/20251020509066?OutUrl=naver>
- 이은성(2026.04.27). 국립발레단 '백조의 호수'...당진에서 완성도 높은 공연 펼친다. 『쿠키뉴스』, https://m.kukinews.com/article/view/kuk202604270084#_enliple
- 장광열(2022.11). 창단 60주년에 바라 본 대한민국 국립발레단의 향방: '발레'를 통한 국가 이미지 고양과 질 높은 공공성. 『춤웹진』, http://www.koreadance.kr/board/board_view.php?view_id=677&board_name=rating
- 장병호(2022.10.17). '창단 60주년' 국립발레단, 100년 비전 찾을 포럼 연다. 『이데일리』, <https://www.edaily.co.kr/News/Read?newsId=01233286632493904&mediaCodeNo=257>

- 최희정(2025.12.23). 어른의 동심 vs 아이의 동화...두 발레단이 그린 '호두까기인형' [객석에서]. 『뉴스』. https://www.newsis.com/view/NISX20251222_0003451145
- 국립발레단(2026.04.04). 발레단 소개-경영공시, from <https://www.korean-national-ballet.kr/ko/introduction/ir>.
- 유니버설발레단(2026.04.04). 발레단 소개, from https://www.universalballet.com/kr/bbs/content.php?co_id=aboutubc.
- Blei, D. M. (2012). Probabilistic Topic Models. *Communications of the ACM*, 55 (4), 77-84.
- Maier, D., Waldherr, A., Miltner, P., Weidemann, G., Niekler, A., Keiner, A., Pfetsch, B., Heyer, G., 140 Reber, U., Haussler, T., schmind-Petri, H., & Adam, S(2018), "Applying LDA Topic Modeling in Communication Research: Toward a Valid and Reliable Methodology", *Communication Methods and Measures* 12(2-3) : p.93-118

ABSTRACT

A Comparative Analysis of Online Response Structures in YouTube Comments Related to the Korean National Ballet and Universal Ballet

Minji Kim* Sookmyung Women's University

This study aims to examine differences in online audience responses to YouTube content related to the Korean National Ballet and Universal Ballet. A total of 902 videos and 10,621 top-level comments posted between January 1, 2015, and December 31, 2025, were collected for analysis. The study employed word frequency analysis, Latent Dirichlet Allocation (LDA) topic modeling, inter-topic distance mapping, and topic-keyword network analysis. The findings indicate that comments on both ballet companies exhibit shared response patterns related to performances, dancers, stage production, repertoire, emotional impressions, and expressions of support. However, comments on the Korean National Ballet show a stronger emphasis on artistic excellence, dancers' technical proficiency, comparisons with international ballet companies, and interest in specific performers. In contrast, comments on the Universal Ballet are more closely associated with expectations surrounding year-end performances of *The Nutcracker*, family attendance, affective memories, and empathy toward dancers' training processes and professional challenges. These results suggest that the Korean National Ballet tends to be received primarily in evaluative terms based on artistic standards and technical excellence, whereas the Universal Ballet is more often received as an experiential cultural event centered on seasonal repertoire and audience engagement. This distinction should not be interpreted as an inherent or fixed characteristic of either company, nor as representative of all audience perceptions; rather, it serves as an analytical construct for explaining the relative patterns observed in the dataset. This study contributes to the field by demonstrating that variations in repertoire composition, content characteristics, and accumulated organizational image may be associated with differing patterns of online audience responses, even within the same genre of ballet. Furthermore, by utilizing YouTube comments as digital reception data and comparatively analyzing the response structures of two ballet companies, this research provides foundational data for digital audience research in dance studies.

Key words : Korean National Ballet, Universal Ballet, YouTube comments, text mining, online audience responses

논문투고일: 2026.05.31

논문심사일: 2026.07.11

심사완료일: 2026.07.20

* Lecturer, Department of Dance, Sookmyung Women's University.