

AI 기반 일·중 온라인 담론에 나타난 협한 정서 비교 분석

李 竣 瑞*

[요지]

본 연구는 인공지능 언어모델인 BERT를 활용하여 일본과 중국의 온라인 공간에서 나타나는 협한 담론 구조와 정서적 특징을 비교 분석하였다. 트위터, 웨이보, 바이두 등 양국을 대표하는 SNS 플랫폼에서 수집한 약 1만 4천 건의 자연언어 데이터를 기반으로, 각국 사용자들이 어떠한 방식으로 협한 감정을 구성하고 표현하는지를 계량적으로 탐색한 것이다.

분석 결과, 일본의 협한 담론은 ‘정치·외교적 한·일관계’, ‘감정적 혐오론’, ‘문화적 비교’라는 세 가지 주제로 분화되어 나타났다. 정치·외교적 한·일관계 담론은 보수 정당과 외교 갈등을 중심으로, 한국에 대한 체계적 비판을 담고 있고, 감정적 혐오론과 관련된 담론에서는 네티즌과 소비 보이콧이 결합된 혐오 표현이 강하게 나타났다. 반면, 문화와 일상이 혼합된 담론은 한국 콘텐츠나 미디어에 대한 평가 속에서 은근한 우월주의와 냉소가 드러났다.

중국의 협한 담론은 ‘정치·외교적 비판’, ‘젠더 갈등과 반페미니즘’, ‘역사·이념 대립’ 및 ‘문화 피로감’ 등 네 가지 축으로 구성되었다. 특히 중국의 경우 협한 감정이 젠더 이슈나 외교 갈등과 복합적으로 얽혀 나타났는데, 여성 혐오와 민족주의 정서가 함께 작동하는 사례도 다수 관찰되었다. 또한 한류 콘텐츠에 대한 소비 피로와 문화적 거리감이 협한 정서로 전이되는 양상도 발견되었다.

이와 같은 분석 결과는 일본과 중국의 협한 담론이 표면적으로는 유사한 양상을 보이지만, 그 내부 구조는 각국의 사회·문화적 맥락에 따라 서로 다른 방식으로 형성된다는 점을 시사한다. 특히 본 연구의 인공지능 기반의 토픽모델링은 단어 사이의 의미적 관계와 문맥적 흐름을 반영함으로써, 협한 감정이 어떻게 구조화되어 전개되고 있는지를 보다 입체적으로 파악할 수 있다. 이러한 분석은 일·중 언어 학습자가 단어 수준의 언어 의미의 이해를 뛰어넘어, 실제 온라인 담론에서 사용되는 표현의 감정적·사회문화적 함의를 파악하는 데 실질적인 도움이 될 수 있다. 나아가, 혐오 표현에 담긴 화용적 전략, 감정 구조, 담화 맥락 등을 분석하고 학습함으로써, 학습자는 비판적 언어 이해 능력과 더불어 다문화적 감수성을 함께 함양할 수 있다. 이는 궁극적으로 실질적인 의사소통 역량의 향상은 물론, 일본 및 중국 사회에 대한 심층적인 문화적 이해로도 이어질 수 있을 것이다.

키워드: 인공지능, 토픽모델링, BERT, 협한감정, 담론분석

1. 서론

최근 인공지능(Artificial Intelligence, AI)의 눈부신 발전은 인간의 지적 활동을 기계가 모방하고 구현하는 방식에 큰 변화를 가져오고 있다. 특히 AI의 하위 분야인 자연어처리(Natural Language Processing, NLP) 기술은 언어를 매개로 문화를 해석하는 분석 도구로 주목받고 있는데, 이는 언어교육을 포함한 다양한 인문·사회과학적 분야에 새로운 시야를 제공하고 있다.

딥러닝 기반의 사전학습(pre-training) 언어모델인 BERT(Bidirectional Encoder Representations from Transformers)는 문맥의 흐름을 양방향으로 학습함으로써, 단어의 의미와 문맥적 맥락까지도 보

* 성결대학교 글로벌물류학과 교수, 일본어학

다 정밀하게 분석할 수 있도록 설계된 것이다(박광현 외, 2019:1). BERT는 자가학습(Self-supervised Learning) 방식으로 방대한 코퍼스에서 의미 관계를 자동으로 습득할 수 있어, 언어를 통해 사회 구성원의 정체성이나 감정 구조, 문화 담론의 흐름을 탐색하는 데 효과적인 근거로 활용할 수 있다. 이러한 자연어처리 기술은 단순한 언어 정보 처리에 그치지 않고, 언어교육적 활용 가능성까지도 내포한다. 즉, BERT를 활용하면 특정 언어 표현에 내재된 감정적·화용론적 특성이나 사회문화적 맥락까지도 체계적이고 계량적인 방식으로 분석하는 데 기여할 수 있다.

언어는 단순한 의사소통의 수단을 뛰어넘어, 특정 집단의 가치관, 감정, 사고방식을 담아내는 문화적 매개체이다. 이러한 관점에서 BERT와 같은 AI 기반 언어모델은 문화적 표현 속에 담긴 언어적 맥락을 계량적인 방식으로 해석하는 데 탁월한 분석 도구라고 할 수 있다. 특히 본 연구에서 차용하는 다국어 사전학습 모델(mBERT)은 한국어, 일본어, 중국어 등 다양한 언어에 적용 가능하다는 점에서, 다국어 간 다문화 담론 분석에 매우 적합한 것이다(박재연 외, 2022:20).

본 연구는 인공지능 기반의 자연어처리 기술을 활용하여 일본과 중국의 온라인 공간에서 나타나는 혐한 담론 구조를 계량적으로 분석함으로써, 양국 사용자들이 혐한 관련 표현을 어떠한 감정 구조와 언어 전략을 통해 구현하고 있는지는 물론, 국가별 혐한 정서의 특성과 문화적 차이를 규명하는 것을 주요 목적으로 한다. 이를 통해 일·중 언어 학습자들이 해당 언어가 실제 사회·문화적 맥락에서 어떻게 사용되는지를 이해하는 데도 기여할 수 있어, 단순한 문법·어휘 습득을 넘어 화용론적 특성과 문화적 함의를 함께 파악하는데 실질적인 도움이 될 뿐만 아니라, 비판적 언어 이해 능력과 다문화 감수성 증진에도 긍정적인 영향을 줄 것으로 기대된다.

2. 본 연구의 목적과 이론적 배경

본 연구는 인공지능(AI) 기술을 활용하여 일본과 중국의 온라인 공간에서 나타나는 혐한 정서를 비교·분석하는 것을 목적으로 한다. 특히 본 연구는 자연어처리(Natural Language Processing, 이하 NLP) 기술의 한 축을 담당하고 있는 최신 언어모델 BERT(Bidirectional Encoder Representations from Transformers)를 활용하여, 일본어와 중국어로 작성된 방대한 SNS 데이터를 분석함으로써, 두 문화권에서 혐한 감정이 언어적으로 어떻게 표현되고, 감정적으로 어떤 구조를 갖추고 있는지를 심층적으로 탐색하고자 한다. 이를 위해 본 장에서는 인공지능과 자연어처리의 개념, 전통적인 언어 분석 방식의 한계, 그리고 BERT 모델이 등장하게 된 배경과 그 구조 및 장점을 소개함으로써 연구의 이론적 토대를 제공하고자 한다.

먼저, 인공지능은 인간의 사고, 학습, 판단 능력을 모방하여 기계가 자율적으로 문제를 해결하거나 언어를 이해하고 문장을 생성하는 기술을 말한다. 이러한 기술은 오늘날 스마트폰 음성비서, 온라인 번역기, 추천 알고리즘, 챗봇 등 우리 일상 곳곳에서 사용되고 있으며, 특히 인간 언어를 이해하고 처리할 수 있는 능력은 인공지능 기술 중에서도 가장 높은 수준의 인지적 처리 영역으로 간주된다. 이러한 인간 언어를 기계가 이해하고 처리하는 기술이 바로 자연어처리(NLP)이다. 자연어처리는 컴퓨터가 인간 언어의 구조, 의미, 감정 등을 분석할 수 있도록 하는 기술로, 음성 인식, 기계 번역, 문서 분류, 감정 분석 등 다양한 응용 분야에서 활용되고 있다. 과거에는 NLP 기술이 규칙 기반 방식(rule-based)이나 전

통적인 기계학습(machine learning)을 중심으로 발전해왔지만, 이들 방식은 문맥 정보를 포착하는 데 어려움이 있어 인간 언어의 복잡성과 애매성을 충분히 반영하지 못했다. 이러한 한계를 극복하고자 최근에는 딥러닝(deep learning) 기술을 바탕으로 한 사전학습 언어모델이 활발히 개발되었고, 그중에서도 BERT는 NLP 분야에서 혁신적인 전환점을 제시한 모델로 평가받는다.

BERT는 Google의 연구팀이 발표한 언어 이해 모델로서, ChatGPT¹⁾를 비롯한 최신 생성형 언어모델의 기반이 되는 트랜스포머(Transformer)라는 신경망 구조를 기반으로 하고 있다(Devlin et al., 2019). 트랜스포머는 문장 내 단어들 간의 관계를 처리하는 데 뛰어난 능력을 갖춘 구조로, 기존 모델들과 달리 문장의 양방향 문맥을 동시에 고려하여 정보를 처리할 수 있어, 단어 간 의미 관계, 문장 구조, 표현 방식 등을 훨씬 정밀하게 이해하는 데 기여한다.

BERT는 이러한 문맥 이해 능력을 갖추기 위해 두 단계의 학습 구조를 따른다. 첫 번째 단계는 사전 학습(pre-training)으로, 위키백과나 뉴스 기사 등과 같은 대규모 범용 텍스트 데이터를 통해 언어의 기본 구조와 표현 양상을 일반화하여 학습한다. 이때 BERT는 단어의 등장 위치, 주변 단어들과의 관계, 마스킹된 단어 예측 등을 통해 문맥과 구조에 대한 종합적인 이해 능력을 형성한다. 두 번째 단계는 파인튜닝(fine-tuning)²⁾으로, 이미 사전학습을 마친 언어모델에 특정 과제나 언어, 도메인에 특화된 소규모 데이터를 부가하여 추가 학습을 수행하는 과정이다. 이때 BERT는 특정 SNS 환경이나 주제에 최적화된 형태로 조정되어 보다 정확하고 유의미한 분석 결과를 도출할 수 있게 된다. 예를 들어, 일본어 SNS 상의 혐한 표현 분석에 활용될 경우, 일본어 특유의 어순과 표현법을 반영하여 혐한 정서의 뉘앙스를 보다 정확하게 포착할 수 있으며, 이는 단순한 단어 빈도 분석으로는 불가능한 수준의 정서적 감지와 문맥 파악을 가능하게 해준다. 이러한 이중 학습 구조는 데이터가 많지 않은 환경에서도 높은 분석 성능을 발휘할 수 있게 해주며, 다양한 언어와 주제에 유연하게 적용할 수 있다는 점에서 자연어처리 분야의 핵심 방법론으로 자리잡고 있다.

특히 BERT는 단일 언어 환경에만 국한되지 않고, 여러 언어를 동시에 학습할 수 있도록 설계된 확장 모델들도 존재한다. 대표적인 예로 본 연구의 mBERT(multilingual BERT)는 100개 이상의 언어를 동시에 학습한 다국어 사전학습 모델로, 영어, 한국어, 일본어, 중국어 등 다양한 언어권의 문장을 하나의 모델 안에서 비교 분석할 수 있게 해준다(Pires et al., 2019). 또한, 일본어와 중국어에 각각 특화된 Japanese BERT, Chinese BERT 등의 버전도 존재하며, 이들 모델은 각 언어의 고유한 문법, 어휘, 표현 방식까지 세밀하게 반영하여 높은 수준의 정밀한 분석을 가능하게 한다.

본 연구에서는 이러한 BERT 기반 언어모델을 활용하여 일본과 중국의 온라인 담론 속에서 나타나는 혐한 정서를 비교·분석하고자 한다. 기존의 텍스트 분석이 주로 특정 단어의 등장 빈도나 공기어 분석에 초점을 맞추었다면, BERT는 문장 전체의 의미 흐름과 문맥적 의미를 함께 고려하여 단어가 어떠한 감정을 담고 있는지를 정량적으로 파악할 수 있다. BERT는 언어 속에 내재된 감정과 의미를 단순한 키

1) OpenAI. (2024). ChatGPT (April 2024 version) [Large language model]. <https://chat.openai.com>

2) 사전학습(pre-training)은 대규모 범용 코퍼스를 기반으로 언어모델이 단어 간 의미 관계, 문장 구조, 문맥 정보 등을 일반화하여 학습하는 초기 단계를 의미한다. 언어에 대한 기본적인 이해 능력을 형성하는 과정이라고 할 수 있는데, 파인튜닝(fine-tuning)은 이러한 사전학습된 언어모델을 특정 언어에 맞춰 소규모 데이터를 이용해 추가 학습시킴으로써, 해당 과제에 최적화된 성능을 도출하는 단계이다. 이와 같은 이중 학습 구조는 적은 양의 데이터로도 높은 성능을 구현할 수 있다는 점에서 자연어처리 분야에서 효과적인 접근 방식으로 평가받는다(Lee et al., 2020:86)

워드 수준이 아니라, 문맥 속 의미의 흐름과 정서의 방향성, 표현의 세기 등을 종합적으로 판단할 수 있게 해주며, 이는 사회적 인식과 문화적 맥락이 반영된 담론 분석에 매우 유효한 기술이라고 할 수 있다.

본 연구는 이러한 BERT의 언어 감성 인식 기능을 바탕으로 일본과 중국 온라인 담론에서 나타나는 혐한 정서의 구조적 차이, 표현의 방식, 정서의 강도 등을 비교하고, 양국에서 혐한 감정이 어떻게 언어적으로 형성되고 어떤 사회문화적 맥락에서 전개되는지를 실증적으로 규명하고자 한다. 이를 통해 단순한 SNS 텍스트의 분석을 넘어, 언어를 통해 구성되는 사회적 인식, 문화 간 차이, 그리고 국가 간 감정의 흐름을 객관적이고 정량적으로 조명하고자 하는 것이다.

3. 일·중 혐한 담론 분석

3.1 데이터 개요

본 연구는 일본과 중국의 온라인 공간에서 형성되는 혐한(嫌韓) 담론 구조를 비교 분석하고자, 각국의 대표적인 소셜미디어 플랫폼에서 생성된 자연언어 텍스트 데이터를 기반으로 두 개의 코퍼스를 구축하였다. 혐한 감정은 정치적 사건이나 사회적 이슈에 따라 빠르게 확산되는 경향이 있는데, 이는 단순한 부정적 감정 표현은 물론, 특정 국가나 집단에 대한 편견 및 고정관념 등 다양한 방식으로 나타난다.

이러한 특징을 정량적으로 분석하기 위해, 본 연구는 언어적·문화적 차이를 반영하는 담론 구조의 비교 및 분석이 가능한 일본어와 중국어 혐한 관련 데이터를 수집하여 이를 분석의 기초 자료로 삼았다.

일본어 코퍼스는 2024년 5월 1일부터 2025년 4월 20일까지 약 1년간 트위터(X)³⁾에서 수집된 총 7,000건의 트윗으로 구성되어 있다. 트위터는 일본 내에서 정치·사회 이슈에 대한 실시간 반응과 감정 표현이 활발히 이루어지는 대표적 플랫폼으로, 본 연구의 목적에 적합한 데이터 수집 공간이라 판단하였다.

중국어 코퍼스는 같은 기간 동안 웨이보(微博)⁴⁾와 바이두(百度)⁵⁾를 통해 수집되었다. 웨이보는 중국 내 SNS 담론의 흐름을 잘 반영하는 공간이며, 바이두는 포털 기반의 게시판 중심 커뮤니티로, 서로 다른 담론 구조를 비교 분석하는 데 도움이 된다.

데이터 수집 시, 일본어 코퍼스에는 ‘혐한(嫌韓)’, ‘반한(反韓)’ 등의 키워드를, 중국어 코퍼스에는 ‘仇韩(혐한)’, ‘反韩(반한)’ 등 대응되는 표현을 중심으로 필터링 작업을 수행하였다. 각 문서에는 텍스트 본문 외에도 작성 날짜, 익명화된 사용자 ID, 사용 기기 정보 등 메타데이터가 포함되어 있으며, 이는 향후 시기별 경향성이나 사용자 유형 분석에 활용할 수 있는 보조 자료가 된다.

수집된 텍스트는 언어별 특성에 따라 맞춤형 전처리 과정을 거쳤다. 일본어는 형태소 분석을 통해 문장을 단어 단위로 분절한 뒤, 불용어 제거 및 특수기호 정제 작업을 진행하였다. 일본어와 중국어가 띄어쓰기가 명시적이지 않다는 점을 고려해 토큰나이징(tokenizing)을 먼저 수행하고, 이어서 문장 부호 및 불필요한 기호들을 제거하는 방식으로 처리하였다.

3) 본 연구에서 분석 대상으로 삼은 SNS 플랫폼은 2022년 말 일론 머스크(Elon Musk)의 인수 이후, 2023년에 ‘X’로 개칭되었다. 이전 명칭인 트위터(Twitter)는 본문 전반에서 병기하거나 혼용하여 사용한다.

4) <https://weibo.com>

5) <https://www.baidu.com>

이와 같은 전처리 작업을 통해 두 언어의 코퍼스가 분석 가능한 형태로 정제되었으며, 후속 토픽 모델링, 감성 분석, 시각화 작업의 기반이 마련된 것이다.

<표 1> 주요 일본어 토픽

	일본어 코퍼스	중국어 코퍼스
사용언어	일본어	중국어
수집 플랫폼	트위터(X)	웨이보, 바이두
수집 기간	2024.05.01.~2025.04.20	2024.05.01.~2025.04.20
총 문서 수	약 7,000건	약 7,000건
주요 키워드	‘혐한’, ‘반한’	‘仇韩’, ‘反韩’
포함된 항목	텍스트, 날짜, 익명 사용자 ID, 기기 정보 등	텍스트, 날짜, 익명 사용자 ID, 기기 정보 등
전처리 방식	형태소 분석, 불용어 제거 등	토큰나이징, 특수문자 제거 등
활용 목적	일본인 혐한 담론 구조 분석	중국인 혐한 감정 담론 구조 분석

3.2 BERT 기반 토픽모델링⁶⁾

SNS 데이터를 분석하는 데 있어서, 개별 문서가 내포하고 있는 주제와 감정의 흐름을 직관적으로 파악하는 것이 매우 중요하다. 이를 위해 본 연구는 인공지능 기반 사전학습 언어모델인 BERT를 활용해 각 문서를 임베딩하고, 그 결과를 바탕으로 비선형 차원 축소 기법인 UMAP(Uniform Manifold Approximation and Projection)과 밀도 기반 클러스터링 알고리즘인 HDBSCAN(Hierarchical Density-Based Spatial Clustering of Applications with Noise)을 적용하여(Grootendorst, M., 2022) 각각의 담론을 시각적으로 분류하였다.

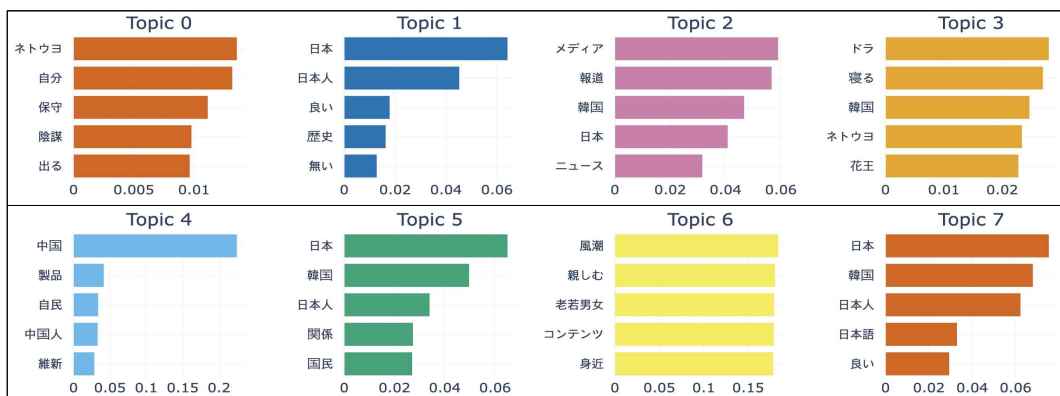
일본어와 중국어 데이터를 대상으로 한 본 연구에서는 다국어 분석에 특화된 사전학습 모델인 multilingual BERT(mBERT)를 이용해 문서의 의미를 고차원 벡터 공간에 임베딩하였다. 또한, 임베딩 UMAP을 사용하여 임베딩 벡터를 2차원 공간으로 축소하고, HDBSCAN 알고리즘을 적용하여 의미적으로 유사한 문장들을 하나의 군집(cluster)으로 분류하였다. 이러한 과정을 통해 각각의 문서가 어떤 의미적 맥락으로 특정 군집에 속하는지를 시각적으로 확인할 수 있는 것이다. 이와 같은 방식으로 도출된 결과는 단순한 키워드 분석을 뛰어넘어, 문맥적 의미를 중심으로 한 담론의 구조적 특성을 한 눈에 파악할 수 있다는 점에서 매우 유용하다. 특히 BERT 기반 임베딩은 단어 간 의미 관계뿐만 아니라 정서적 맥락까지도 포착할 수 있어, 본 연구의 혐한 담론처럼 복합적인 감정이 얹힌 텍스트 분석에 매우 유리한 것이다.

6) 기존의 LDA(Latent Dirichlet Allocation) 방식은 단어의 동시 출현 확률에 기반한 확률적 토픽모델링 기법으로, 문서 내 주제를 추론하는 데 유용하지만, 문맥 정보나 단어 간 의미적 관계를 반영하는 데 한계가 있다. 반면, BERT 기반 토픽모델링은 사전학습 언어모델의 문맥 인식 능력을 토픽 추출에 접목함으로써, 동일한 단어라도 문맥에 따라 의미가 달라질 수 있는 자연어의 복잡성을 보다 정밀하게 포착할 수 있다(김선욱, 2022:100).



<그림1> 일본어 혐한 담론에 대한 BERT 기반 클러스터링 결과⁷⁾

<그림1>은 일본 트위터 데이터에서 도출된 주요 혐한 토픽을 2차원 임베딩 공간상에 시각화한 결과이다. 각 점은 개별 트윗을 의미하며, 동일 색상의 점들은 유사한 의미 구조를 공유하는 텍스트들로 구성된 하나의 토픽 클러스터를 나타낸다. 이와 같은 시각적 분석 시각화 결과를 통해 일본 내 혐한 담론이 단일한 방향성을 지니기보다는, 다양한 사회·문화·이념적 맥락에 따라 다각도로 분화되어 있다는 점을 뚜렷이 파악할 수 있는 것이다. 본 연구는 이들 클러스터를 중심으로 핵심 토픽과 주요 키워드를 추가적으로 분석함으로써 일본 내 혐한 정서의 구조적 특징을 보다 심층적으로 탐색할 수 있다.



<그림2> 일본어 토픽별 주요 키워드

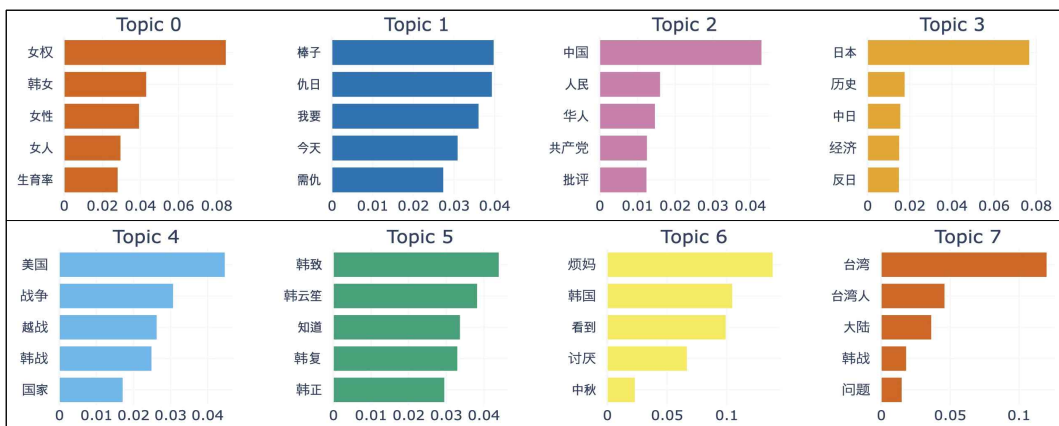
7) BERT 기반 토픽모델링 결과를 UMAP(Uniform Manifold Approximation and Projection) 차원 축소 기법을 활용하여 2차원 평면에 시각화한 것이다. 이를 통해 온라인 혐한 담론 내에서 주요 주제들이 어떻게 분포하고 군집화되는지를 시각적으로 확인할 수 있다.

<그림2>는 일본어 혐한 담론에서 도출된 각각의 주요 토픽을 구성하는 키워드 조합을 표 형식으로 제시한 것으로, 각 토픽의 중심 주제와 감정적 경향성을 파악할 수 있다.



<그림3> 중국어 혐한 담론에 대한 BERT 기반 클러스터링 결과

<그림3>은 중국어 SNS 데이터를 대상으로 일본어와 동일한 분석 절차(BERT 임베딩, UMAP 차원 축소, HDBSCAN 클러스터링 적용)를 적용한 결과로, 각 문서가 2차원 공간상에서 의미적으로 유사한 문서들과 함께 하나의 군집으로 시각화된 모습을 보여준다. 또한, 일본어와 동일한 개수의 주요 토픽을 도출하였는데, 아래 <그림4>은 중국어 혐한 담론에서 도출된 각 토픽 클러스터별 상위 키워드를 제시한 것이다.



<그림4> 중국어 토픽별 주요 키워드

이를 통해 해당 담론의 주제적 특성과 중국내 혐한 감정의 경향성을 유추할 수 있는 것이다.

3.3 토픽 클러스터링과 담론분석

BERT 기반 토픽 모델링을 통해 도출된 클러스터들을 통해 각기 다른 주제적 특성과 정서적 맥락을 파악할 수 있다. 본 연구에서는 이들 클러스터를 종합적으로 분석하여 일본과 중국의 혐한 담론이 어떠한 구조로 형성되고 있는지를 다층적으로 살펴보았다.

<표2> 주요 일본어 토픽

순위	토픽1	토픽2	토픽3	토픽4	토픽5	토픽6	토픽7	토픽8	토픽9	토픽10
1	韓国	ネットウヨ	日本	メディア	ドラ	中国	日本	風潮	日本	韓国
2	日本	自分	日本人	報道	寝る	製品	韓国	親しむ	韓国	乞食
3	日本人	保守	良い	韓国	韓国	自民党	日本人	老若男女	日本人	コカコーラ
4	中国	陰謀	歴史	日本	ネットウヨ	中国人	関係	コンテンツ	日本語	好き
5	反日	出る	無い	ニュース	花王	維新	国民	身近	良い	ルール
6	朝鮮	好き	外国	大谷	フジテレビ	侵略	叩く	今時	お互い	彼女
7	北朝鮮	のに	自分	反日	コン	関係	大統領	問う	分かる	ペプシコーラ
8	嫌い	良い	差別	テレビ	藤木	まさに	反日	主流	教える	凄い
9	多い	反日	文化	新聞	本家	買う	断絶	実際	最下位	破る
10	問題	行く	来る	大統領	かま	食料	左翼	影響	嫌う	買う

<표2>의 일본어 코퍼스에서 도출된 총 10개의 토픽을 의미적 유사성과 주제적 응집도에 따라 세 가지 주요 군집으로 재분류할 수 있다(아래 <표3>).

<표3> 주요 일본어 토픽 재분류

구분	관련 토픽	주요 내용	주요 키워드	특징 요약
군집 1: 정치·외교 중심 담론	1, 6, 7, 9	한일 관계, 정치, 외교, 무역 이슈 중심의 비판 담론	韓国(한국), 日本(일본), 中国(중국), 自民党(자민당), 維新(유신), 製品(제품), 関係(관계), 国民(국민)	우의 시각에서의 정치화된 혐한 감정, 실질적 행동(불매 등) 반영
군집 2: 감정적 혐오 담론	2, 5, 10	강한 정서적 반감 및 감정 중심의 혐한 표현	ネットウヨ(넷우익), 嫌い(싫어함), 断絶(단절), 破る(깨뜨리다), 買う(사다), ペプシコーラ(펩시콜라), 彼女(여자친구)	극우 담론 주도, 소비 보이콧, 혐오 정서의 직접적 표출
군집 3: 문화·일상 혼합 담론	3, 4, 8	미디어, 콘텐츠, 스포츠 등 일상적 비교와 폄하 담론	メディア(미디어), 報道(보도), フジテレビ(후지TV), コンテンツ(콘텐츠), 教える(가르치다), 実際(실제), 最下位(최하위)	일상 속 우월감 표현, 은연중의 혐한 감정

각 토픽은 정치·외교, 감정적 혐오, 문화·일상이라는 세 가지 클러스터(군집)으로 재구성할 수 있다. 군집 1은 ‘정치·외교 중심의 비판 담론’으로 ‘한국(韓国)’, ‘일본(日本)’, ‘중국(中国)’, ‘자민당(自民)’,

‘유신(維新)’, ‘제품(製品)’, ‘국민(國民)’ 등 정치, 외교, 무역 등 거시적 이슈와 관련한 단어들이 집중되어 있다. 특히 ‘자민당(自民黨)’이나 ‘유신(維新)’과 같은 보수 정치 세력에 대한 언급은 이 담론이 일본 내 우익 정치 성향과 밀접하게 연결되어 있음을 시사한다. 또한 ‘한국(韓國)’과 ‘제품(製品)’이라는 단어들이 연속적으로 나열된 것에서 혐한 감정이 한·일 간 무역 분쟁이나 불매운동 등 실질적 행동으로 이어지고 있음을 알 수 있다. 이러한 단어 구성은 일본 내 보수주의적 외교 인식이 혐한 정서를 정치화하는 데 영향을 미치고 있어, 혐한 담론이 단순한 편견을 넘어 정치·외교적 정당화의 논리로 전개되고 있음을 보여준다.

군집 2는 ‘감정적 혐오 및 비하 중심 담론’으로 ‘넷우익(ネトウヨ)’, ‘단절(斷絶)’, ‘싫어함(嫌い)’, ‘깨뜨리다(破る)’, ‘여자친구(彼女)’, ‘사다(買う)’, ‘펍시콜라(ペプシコーラ)’ 등의 단어들이 포함되어 있다. 해당 군집은 이성적 비판보다는 직접적 혐오와 반감을 표현하는 감정 중심의 혐한 담론이 나타난다. 특히 ‘넷우익(ネトウヨ)’이라는 단어는 일본 내 온라인 극우 커뮤니티와 연계된 사용자들의 혐한 담론이 구조화되어 있음을 보여주는 단서로 작용하고 있다. 또한, 특정 소비재 브랜드에 대한 언급은 한국 또는 친한 성향을 가진 브랜드에 대해 부정적인 태도나 보이콧 움직임이 나타나고 있음을 의미한다. 이는 감정 중심의 혐오 담론이 즉각적이고 강도 높은 실질적 소비행태를 유도할 수 있으므로, 사회적 전파력과 파급력이 크다는 점에서 특별한 관찰이 필요하다.

군집 3은 ‘문화·일상 혼합 담론’으로 ‘미디어(メディア)’, ‘보도(報道)’, ‘후지TV(フジテレビ)’, ‘콘텐츠(コンテンツ)’, ‘가르치다(教える)’, ‘실제(実際)’, ‘최하위(最下位)’ 등 일상 생활과 문화 콘텐츠 소비 경험을 중심으로 구성되어 있다. 이 담론에서는 노골적인 혐한 감정보다는, 미디어의 보도 방식, 콘텐츠 소비 경험, 한·일 간 스포츠 경쟁 등 일상적 경험 속에서 은연중에 표현되는 냉소와 부정적 인식이 드러나는 것이 특징적이다. 이러한 비교 우위와 냉소적 언어를 통해 드러나는 혐한 담론은 표면적으로는 감정의 강도가 낮아 보일 수 있으나, 사회 전반에 은연중에 스며드는 우월주의 정서가 나타나는 방식으로 작동할 수 있는 것이다.

<표4> 주요 중국어 토픽

순위	토픽1	토픽2	토픽3	토픽4	토픽5	토픽6	토픽7	토픽8	토픽9	토픽10
1	中国	女权	棒子	中国	日本	美国	韩致	烦妈	台湾	微博
2	韩国	韩女	仇日	人民	历史	战争	韩云圪	韩国	台湾人	韩娱
3	日本	女性	我要	华人	中日	越战	知道	看到	大陆	粉丝
4	美国	女人	今天	共产党	经济	韩国	韩复	讨论	韩战	天天
5	就是	生产率	嫌仇	批评	反日	国家	韩正	中秋	问题	碰瓷
6	中共	男性	真的	中华	仇日	反战	香菇	喜欢	独立	活疹
7	仇日	男人	我会	国家	反中	北约	甜橙	韩国	蒋介石	一些
8	文化	韩男	这么	反贪	中日韩	联合国	韩三骗	棒子	收复	全文
9	不是	极端	支那	教育	就是	二战	是不是	唯一	不是	真的
10	这个	支女	椅子	西方	道歉	伊拉克	谢罪	泡菜	美国	社交

중국의 혐한 담론 10개의 토픽(<표4>)도 의미적 유사성과 주제적 응집도에 따라 다음 네 가지 주요 군집으로 재분류할 수 있다.

<표5> 주요 중국어 토픽 재분류

구분	관련 토픽	주요 내용	주요 키워드	내용 요약
군집 1: 정치·외교적 대립 및 체제 비판	1, 4, 6, 9	한중 외교 갈등, 반미·반한 정서	中国(중국), 韩国(한국), 美国(미국), 台湾(대만), 人民(인민), 国家(국가), 战争(전쟁), 北约(나토), 联合国(유엔), 韩战(한국전), 独立(독립), 中共 (중국공산당)	한국의 역사 교육, 한국의 외교 노선과 미국·대만과의 협력에 대한 비판 및 중국 중심 체제 정당화
군집 2: 젠더 및 반페미니즘 담론	2, 7	반페미니즘, 성 별 갈등 조롱	韩女(한국 여성), 女性(여성), 女权 (여성권리), 男性(남성), 支女(페미 지지자), 极端(극단), 韩致/韩正(한 국 정치인), 韩三蹦(한국 비하어), 谢 罪(사죄), 是不是(그렇지 않다)	한국의 젠더 이슈를 조롱하며 협한 정서를 젠더담론에 접 목, 급진 페미니즘에 대한 반 감 표출
군집 3: 역사·이념 갈등	3, 5	역사 왜곡, 이념 갈등 비판	仇日(반일), 中日(중일), 历史(역사), 反日(반일), 反中(반중), 经济(경제), 教育(교육), 西方(서방), 中日韩(한 중일), 道歉(사죄)	드라마 등 한류 콘텐츠에 나 타나는 한국의 반일 감정을 비판하며, 역사 인식 및 교육 을 문제 삼음. 서구 추종 경향 에 대한 불만
군집 4: 문화 소비 기반 혐 한 담론	8, 10	한류 피로감, 문 화적 비하	棒子(한국인 비하), 泡菜(김치), 粉 丝(팬), 微博(웨이보), 韩娱(한류), 天天(매일), 活泼(활기찬), 社交 (SNS), 破碎(깨지다), 喜欢(좋아하 다), 讨厌(싫어하다)	한국 문화 콘텐츠 소비와 함 께 나타나는 피로, 거부감, 감 정적 혐한

군집 1은 중국의 외교 및 지정학적 관점에서 한국의 외교노선을 비판하는 담론으로 구성되어 있다. ‘台湾(대만)’, ‘美国(미국)’과의 관계에서 한국이 ‘中国(중국)’에 비우호적인 태도를 취하고 있다는 인식을 엿볼 수 있다. ‘北约(나토)’, ‘联合国(유엔)’, ‘独立(독립)’ 등의 키워드들은 한국이 서방 중심의 국제 질서에 편승하고 있다는 비판적 시야를 반영하며, 특히 ‘反战(반전)’이나 ‘战斗(전투)’와 같은 단어는 한국이 미국 중심의 군사 전략에 참여하는 것을 중국 내에서 부정적으로 인식하고 있다는 부정적 평가로 연결된다. 이러한 담론은 한국에 대한 직접적 비난보다는 한국의 외교노선이 중국과의 한·중관계에서 갈등 구조가 부각되고 있다는 비판적 인식이 작용하고 있는 것이다.

군집 2는 한국 여성에 대한 비난과 동시에 극단적 페미니즘에 대한 거부 반응이 나타나는 담론이다. ‘韩女(한국 여성)’, ‘女权(페미니즘)’, ‘支女(페미 지지자)’ 등의 키워드는 한국 사회의 젠더 갈등을 중국인들이 조롱하거나 공격하는 주요 대상으로 삼고 있음을 보여준다. 특히 ‘韩三蹦(한국 비하어)’이라는 신조어는 ‘한국 여성들이 연애, 결혼, 출산을 거부한다’는 중국 내 편견을 반영하며, 한국의 급진적 여성주의에 대한 거부감을 표현하는 데 사용된다. 이는 중국 내 보수적인 성 역할 규범과 충돌하며, 한국을 중국 정서 기준으로 볼 때 ‘부도덕한 국가’로 타자화하여 인식하는 젠더 기반 혐오 감정과 연결된다.

군집 3은 한·중·일 역사 갈등 속에서 한국의 역사 인식을 문제 삼는 담론이다. 핵심 키워드인 ‘仇日(반일)’, ‘反日(반일)’, ‘中日(중일)’, ‘中日韩(중국-일본-한국)’, ‘历史(역사)’ 등은 한국이 일본에 대해 지속적인 반일 정서를 표출하면서도, 실제 정치·경제적 차원에서는 일본 및 서방 국가들과 협력하거나 그들과 유사한 태도를 보인다는 중국인의 인식이 반영되어 있다. 즉, ‘道歉(사죄)’, ‘教育(교육)’, ‘西方(서방)’ 등의 키워드는 한국이 일본의 과거사에 대해 반복적으로 사과를 요구하면서도, 정작 서구 혹은 일

본과의 관계 속에서는 타협적인 태도를 취한다는 비판으로 이어진다. 이는 중국 내에서 한국의 역사 인식이 일관성을 결여하고 정치적으로 선택적이라는 인식으로 귀결된다.

군집 4는 중국 내 한국 대중문화 소비 환경에 대해 ‘호감’과 ‘혐오’가 동시에 교차하는 양가적 정서를 반영하는 담론이다. 주요 키워드인 ‘棒子(한국인 비하 표현)’, ‘泡菜(김치)’, ‘韩娱(한류)’, ‘微博(웨이보)’, ‘社交(SNS)’, ‘粉丝(팬)’, ‘活泼(활기찬)’, ‘破碎(깨지다)’, ‘天天(매일)’ 등을 통해, 중국 이용자들이 한국 문화, 특히 한류 콘텐츠를 일상적으로 소비하면서도, 동시에 국가 정체성·민족 감정과 충돌할 때 비하적 표현이나 정서적 거부 반응을 드러내는 양상을 보여준다. 한국 대중문화 소비와 관련한 주요 키워드들과 함께 ‘喜欢(좋아하다)’와 ‘讨厌(싫어하다)’가 동시에 자주 등장하는 것은 한류에 대한 감정이 이중적이라는 것을 보여주며, 이는 팬덤과 일반 대중 사이에서 한국 문화에 대한 사회적 거리감과 정체성 혼란이 공존하고 있음을 의미한다.

4. 결론

본 연구는 인공지능 언어모델인 BERT를 활용하여 일본과 중국의 온라인 공간에서 나타나는 혐한 담론 구조와 정서적 특징을 비교 분석하였다. 트위터, 웨이보, 바이두 등 양국을 대표하는 SNS 플랫폼에서 수집한 약 1만 4천 건의 자연언어 데이터를 기반으로, 각국 이용자들이 어떠한 방식으로 혐한 감정을 구성하고 표현하는지를 계량적으로 탐색한 것이다.

분석 결과, 일본의 혐한 담론은 ‘정치·외교적 한·일관계’, ‘감정적 혐오론’, ‘문화적 비교’라는 세 가지 주제로 분화되어 나타났다. 정치·외교적 한·일관계 담론은 보수 정당과 외교 갈등을 중심으로, 한국에 대한 체계적 비판을 담고 있고, 감정적 혐오론과 관련된 담론에서는 네티즌과 소비 보이콧이 결합된 혐오 표현이 강하게 나타났다. 반면, 문화와 일상이 혼합된 담론은 한국 콘텐츠나 미디어에 대한 평가 속에서 은근한 우월주의와 냉소가 드러났다.

중국의 혐한 담론은 ‘정치·외교적 비판’, ‘젠더 갈등과 반페미니즘’, ‘역사·이념 대립’ 및 ‘문화 괴로움’ 등 네 가지 축으로 구성되었다. 특히 중국의 경우 혐한 감정이 젠더 이슈나 외교 갈등과 복합적으로 얽혀 나타났는데, 여성혐오와 민족주의 정서가 함께 작동하는 사례도 다수 관찰되었다. 또한 한류 콘텐츠에 대한 소비 괴로움과 문화적 거리감이 혐한 정서로 전이되는 양상도 발견되었다.

이와 같은 분석 결과는 일본과 중국의 혐한 담론이 표면적으로는 유사한 양상을 보이지만, 그 내부 구조는 각국의 사회·문화적 맥락에 따라 서로 다른 방식으로 형성된다는 점을 시사한다. 특히 본 연구의 인공지능 기반의 토픽모델링은 단어 사이의 의미적 관계와 문맥적 흐름을 반영함으로써, 혐한 감정이 어떻게 구조화되어 전개되고 있는지를 보다 입체적으로 파악할 수 있다. 이러한 분석은 일·중 언어 학습자가 단어 수준의 언어 의미의 이해를 뛰어넘어, 실제 온라인 담론에서 사용되는 표현의 감정적·사회문화적 함의를 파악하는 데 실질적인 도움이 될 수 있다. 나아가, 혐오 표현에 담긴 화용적 전략, 감정 구조, 담화 맥락 등을 분석하고 학습함으로써, 학습자는 비판적 언어 이해 능력과 더불어 다문화적 감수성을 함께 함양할 수 있다. 이는 궁극적으로 실질적인 의사소통 역량의 향상은 물론, 일본 및 중국 사회에 대한 심층적인 문화적 이해로도 이어질 수 있을 것이다.

◀ 참고문헌 ▶

- 김선욱 · Yang, Kiduk(2022) 「LDA와 BERTopic을 이용한 토픽모델링의 증강과 확장 기법 연구」 『정보관리학회지』 39(3), 한국정보관리학회, pp.99-132
- 김혜진(2022) 「텍스트마이닝 기법을 이용한 한국 사회의 혐오 양상 분석」 『한국문헌정보학회지』 56(4), 한국문헌정보학회, pp.431-453
- 박광현 · 나승훈 · 신종훈 · 김영길(2019) 「BERT를 이용한 한국어 자연어처리: 개체명 인식, 감성분석, 의존 파싱, 의미역 결정」 『한국정보과학회 학술발표논문집』, 한국정보과학회, pp.584-586
- 박명희 · 최은봉(2019) 「일본사회 혐한의 확산: 자정의 담론구조와 한일관계의 부침」 『일본연구논총』 50, 현대일본학회, pp.6-26
- 박재언 · 김지호 · 이홍철(2022) 「BERT 기반의 사전 학습 언어 모델을 이용한 한국어 문서 추출 요약 베이스라인 설계」 『한국정보기술학회논문지』 20(6), 한국정보기술학회, pp.19-32
- 이희진(2017) 「한류콘텐츠 이용정도가 중국인의 협한정서에 미치는 영향: 한류호감 한국인에 대한 긍정적 인식의 이차매개효과 검증을 중심으로」 『한국콘텐츠학회논문지』 17(10), 한국콘텐츠학회, pp.394-405
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K.(2019) “BERT: Pre-training of deep bidirectional transformers for language understanding”, Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), pp.4171 - 4186
- Grootendorst, M.(2022) “BERTopic: Neural topic modeling with a class-based TF-IDF procedure”, arXiv preprint, arXiv:2203.05794, <https://arxiv.org/pdf/2203.05794.pdf>
- Lee, C. H., Lee, Y. J., & Lee, D. H.(2020) “A study of fine tuning pre-trained Korean BERT for question answering performance development”, Journal of Information Technology Services, 19(5), Information Technology Services, pp.83-91
- Liu, Y., & Lapata, M.(2019) “Text summarization with pretrained encoders”, Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), The Association for Computational Linguistics, pp.3730-3740
- McInnes, L., Healy, J., & Melville, J.(2018) “UMAP: Uniform Manifold Approximation and Projection for dimension reduction”, arXiv preprint, arXiv:1802.03426, <https://arxiv.org/pdf/1802.03426.pdf>, pp.1-63
- Pires, T., Schlinger, E., & Garrette, D.(2019) “How multilingual is multilingual BERT?”, Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp.4996-5001

<Abstract>

A Comparative Analysis of Anti-Korean Sentiment in Japanese and Chinese Online Discourse Using AI-Based Methods

This study conducts a comparative analysis of anti-Korean discourse and its emotional characteristics in online spaces in Japan and China, utilizing a BERT-based AI language model. Drawing upon approximately 14,000 posts collected from major social media platforms such as Twitter, Weibo, and Baidu, the study quantitatively explores how users in each country construct and express anti-Korean sentiment.

The analysis reveals that Japanese discourse centers around three main themes: political and diplomatic Korea-Japan relations, emotional hatred, and cultural comparison. Discourses related to political relations focus on conservative parties and diplomatic tensions, articulating structured criticism of Korea. Meanwhile, expressions of emotional aversion are often tied to net-rightist rhetoric and boycott movements. Cultural discourses, on the other hand, exhibit subtle superiority and cynicism in the evaluations of Korean media and content.

Chinese anti-Korean discourse is structured around four major themes: political and diplomatic criticism, gender conflict and anti-feminism, historical and ideological confrontation, and cultural fatigue. In particular, the Chinese context demonstrates a complex entanglement of anti-Korean sentiment with gender issues and diplomatic conflicts, often blending misogyny with nationalist emotions. Additionally, consumption fatigue and cultural distance regarding Korean pop culture contribute to the spread of anti-Korean sentiment.

These findings indicate that while Japanese and Chinese anti-Korean discourse may appear similar on the surface, their underlying structures are shaped by distinct socio-cultural contexts. The AI-based topic modeling applied in this study enables a more nuanced understanding by capturing the semantic relationships and contextual dynamics of language use. Such insights can support language learners in moving beyond surface-level vocabulary acquisition toward a deeper comprehension of the emotional and socio-cultural nuances in online discourse. Furthermore, by analyzing the pragmatic strategies, affective structures, and discursive contexts embedded in hate expressions, learners can cultivate both critical language awareness and intercultural sensitivity—ultimately enhancing communicative competence and fostering deeper cultural understanding of Japanese and Chinese societies.

Key Words : AI (Artificial Intelligence), topic modeling, BERT, Anti-Korean sentiment, discourse, analysis

◆ 李俊瑞(Lee, Jun-Seo)

- 소속 : 성결대학교 글로벌물류학과 교수
- E-mail : jslee@sungkyul.ac.kr

논문투고일	2025. 05. 06.
심사개시일	2025. 05. 11.
심사완료일	2025. 05. 26.
게재확정	2025. 06. 04.