

디지털 플랫폼에서의 문화적 감성 규범과 미디어 수용의 차이: 모로코 여행 유튜브 브이로그에 대한 한·일 이용자 댓글 비교

위 찬 우¹ · 윤 승 옥^{2*}¹전북대학교 일본학과 석사과정²전북대학교 프랑스아프리카연구소 HK연구교수

Cultural Norms of Emotional Expression and Media Reception Differences on Digital Platforms: A Comparative Analysis of Korean and Japanese User Comments on Morocco Travel YouTube Vlogs

Chan-Woo Wi¹ · Sung-Uk Yun^{2*}¹Master's Course, Department of Japanese Studies, Graduate School, Jeonbuk National University, Jeonbuk 54896, Korea²Research Professor, Institute of French-African Studies, Jeonbuk National University, Jeonbuk 54896, Korea

[요 약]

본 연구는 모로코 여행 유튜브 영상에 대한 한일 이용자 댓글을 다국어 감성분석과 토픽 모델링으로 비교 분석하여 문화적 감성 표현 규범의 차이를 실증적으로 규명하였다. 분석 결과, 한국 댓글은 부정 감정이 우세하며 대인 갈등 관련 키워드가 지배적인 반면, 일본 댓글은 긍정 감정이 우세하며 실용적 여행 정보 키워드가 우선되었다. 토픽 분포에서도 한국은 부정적 경험이라는 단일 주제로 편중된 반면, 일본은 다양한 주제에 고르게 분산되었다. 그리고 다중회귀분석 결과, 콘텐츠의 객관적 특성(부정적 장면 수, 경고 단어 빈도)을 통제 한 후에도 문화권 요인이 댓글의 감정 반응에 유의한 설명력을 보였다. 다만 측정된 콘텐츠 변수가 제한적이라는 점에서 해석에 주의가 요구된다. 이러한 결과는 한국의 정(情) 기반 표현적 감성 규범과 일본의 화(和) 기반 억제적 규범이 동일한 콘텐츠에 대한 차별적 수용을 시사한다.

[Abstract]

This study empirically investigates differences in cultural norms of emotional expression by conducting a comparative analysis of Korean and Japanese user comments on Morocco travel YouTube videos using multilingual sentiment analysis and topic modeling. The results indicate that Korean comments are characterized by a predominance of negative sentiments and keywords related to interpersonal conflict. In contrast, Japanese comments exhibit a predominance of positive sentiment, with a focus on keywords associated with practical travel information. Regarding topic distribution, Korean comments are heavily skewed toward a single theme of negative experiences, whereas Japanese comments are more evenly distributed across a range of topics. Furthermore, multiple regression analysis indicates that cultural factors account for significant explanatory power in emotional responses to comments, even after controlling for objective content characteristics (e.g., number of negative scenes and frequency of warning words). However, caution is warranted in interpreting the results due to the limited scope of the measured content variables. These findings suggest that the expressive emotional norms of Korea, based on Jeong (情), and the restrictive norms of Japan, based on Wa (和), lead to differential reception of the same content.

색인어 : 유튜브 댓글, 다국어 감성분석, BERTopic, 한일 문화비교, 문화적 감성 규범**Keyword** : YouTube Comments, Multilingual Sentiment Analysis, BER Topic, Korea-Japan Cultural Comparison, Cultural Display Rules<http://dx.doi.org/10.9728/dcs.2026.27.3.643>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 10 December 2025; Revised 12 January 2026

Accepted 05 February 2026

*Corresponding Author; Sung-Uk Yun

Tel: 

E-mail: yunsunguk@naver.com

1. 서론

유튜브는 현대의 가장 영향력 있는 미디어 플랫폼으로서 월간 약 25억 명 이상의 활동 이용자를 보유하고 있으며, 이용자들이 작성하는 댓글은 단순한 반응을 넘어 공론장의 역할을 수행하고 있다[1]. 특히 여행 콘텐츠는 잠재 관광객들의 목적지 선택에 직접적인 영향을 미치는 중요한 정보원으로 인식되고 있으며, 댓글은 이용자생성콘텐츠(UGC 혹은 UCC)로서 집단적 의견 형성의 중요한 공간이 되었다[2]. 나아가, 목적지의 매력, 위험, 비용, 문화적 차이를 이용자에게 생생히 전달하고, 댓글 공간은 이에 대한 대중의 인식과 감성을 집단으로 표출하는 장이 된다.

특히 아프리카 대륙의 대표 관광지인 모로코는 팬데믹 이후 빠른 회복세를 보이고 있다. 한국수출입은행 해외경제연구소[3]에 따르면, 모로코는 2023년 1~11월 누적 방문객 수 1,320만 명을 기록하여 팬데믹 이전인 2019년(1,290만 명)을 상회하였다. 관광산업이 급속히 회복되는 목적지일수록, 온라인 여행 콘텐츠와 댓글이 잠재 여행자의 위험 인식 및 감성 형성에 미치는 영향력은 한층 증대된다.

한편, 한국과 일본은 동아시아에 위치하여 지리적으로 근접한 국가들이면서도 문화적으로는 상당한 차이를 보인다. 두 국가는 모두 Edward Hall이 제시한 고맥락 문화(High-Context Culture)에 속하는 것으로 분류됐으나, 최근의 심리학 연구들은 고맥락 문화라는 대분류 내에서도 한국과 일본이 감성 표현의 강도와 양식에서 뚜렷한 차이를 보인다는 것을 입증하고 있다[4]. 또한 Matsumoto와 Hwang의 연구에 따르면, 일본 문화권에서는 부정 감성, 특히 분노와 혐오의 표현을 공적 공간에서 강하게 억제하는 경향을 보이지만, 한국 문화권에서는 감성을 적극적으로 표현하고 공유하는 것을 통해 상호 이해와 공감을 도모하는 경향을 보인다[5].

이러한 문화적 기저의 차이는 동일한 여행 콘텐츠에 대한 댓글 반응에서도 서로 다른 양상으로 드러날 가능성이 크다. 실제로 권세린은 코로나19 긴급재난지원금 관련 기사에 달린 한·일 유튜브 댓글을 비교한 연구에서, 한국 댓글이 '세금 납부'와 '공정성'에 대한 정서적 호소에 집중된 반면, 일본 댓글은 법적, 제도적 기준을 중심으로 수혜 범위를 논의하는 경향이 강하다고 보고했다[6]. 또한 이수미는 한·일 정상회담 관련 댓글에서, 한국 댓글은 감성 호소 및 정치 전략 프레임에, 일본 댓글은 목적지향, 도덕성 정당화 전략에 상대적으로 더 의존한다는 차이를 발견했다[7].

그러나 기존 연구들은 유튜브 콘텐츠의 형식적 특성이나 단일 국가의 이용자 반응에 초점을 맞추어 왔으며, 주로 정치·정책 이슈(재난지원금, 정상회담 등)에 대한 댓글 분석에 집중되었다. 동일한 콘텐츠에 대해 상이한 문화권의 이용자들이 어떻게 차별적 감성과 의미를 부여하는지에 대한 체계적 규명은 부족한 실정이다. 방법론적으로도 단순 빈도분석이나 기초적 토픽 모델링에 머물러, 감성분석을 선행 단계로 설정한 감성 범주별 세분화 분석이 시도되지 않았다. 아울러, 콘텐츠

의 객관적 특성(부정적 장면 빈도, 편집 스타일 등)과 이용자의 문화적 배경이 댓글 감성에 미치는 상대적 영향력을 정량적으로 분리하지 못하였으며, 특히 아프리카 여행이라는 비교적 생소한 맥락에서 한·일 댓글을 직접 비교한 연구는 거의 부재하다.

본 연구가 아프리카 지역의 모로코를 분석 대상 여행지로 선정한 이유는 먼저, 한국과 일본 모두에서 최근 관심이 증대되고 있는 새로운 여행 목적지이면서도, 동아시아 유튜브 이용자들에게 지리적, 문화적으로 상당히 생소한 지역이라는 점이다. 이러한 특성은 유튜브 이용자의 문화적 기준과 감성 규범이 콘텐츠 해석에 미치는 영향을 더욱 명확하게 드러낼 수 있게 한다. 다음으로, 모로코 여행 영상에는 호객(hagglng), 사기, 거지의 구걸, 여성 안전 등 문화 간 차이로 인한 충돌의 요소들이 자연스럽게 포함되어 있다는 점이다. 이는 부정적 경험과 긍정적 경험이 함께 나타나므로, 한국과 일본 유튜브 이용자의 감성 표현에 대한 규범의 차이를 가장 명확하게 분석할 수 있는 환경을 제공한다.

이에 본 연구에서는 텍스트마이닝(Text Mining)과 감성분석(Sentiment Analysis) 기법을 융합하여, 한일 양국의 모로코 여행 영상 댓글에 대하여 한국과 일본 이용자들이 보여주는 반응의 차이를 실증적으로 규명하는 데 그 목적이 있다. 특히 양국 이용자들이 무엇에 주목하고, 어떤 단어와 주제를 통해 모로코를 해석하며, 양국 이용자의 모로코 여행에 대한 인식 구조의 차이점과 그 근저의 문화적 토대는 무엇인지 살펴보고자 한다. 이를 위해 본 연구에서는 다음과 같은 연구 문제를 설정하였다.

연구 문제 1. 모로코 여행 유튜브 영상에 대한 한국과 일본 이용자 댓글에서 감성의 분포가 통계적으로 유의미하게 차이나는가?

연구문제 2. 감성 범주별로 한국과 일본 댓글에서 나타나는 핵심 키워드는 무엇이며, 그것이 지닌 문화적 의미는 어떻게 다른가?

연구문제 3. 토픽 모델링을 통해 도출된 각 국가 댓글의 주요 관심 주제는 무엇이며, 토픽 분포의 특성은 어떻게 대조되는가?

연구문제 4. 한일 댓글 차이를 설명하는 근본적인 문화적 기제는 무엇이며, 콘텐츠의 객관적 특성 차이와 이용자의 문화적 감성 규범을 어떻게 구분하여 분석할 수 있는가?

II. 이론적 배경

2-1. 문화적 맥락이론과 한일 비교

Hall[8]의 고맥락/저맥락 문화 이론과 Hofstede[9]의 문화차원론은 한국과 일본을 모두 고맥락·집단주의 문화권으로

분류하였다. 그러나 이러한 대분류는 양국 간 미세한 차이를 포착하지 못한다는 비판을 받아왔다[10],[11]. Cardon[10]은 Hall의 모델이 국가별 고정 점수로 오용될 위험을 지적하였으며, Gerlach와 Eriksson[11]은 Hofstede 척도의 낮은 외적 타당도를 실증하였다.

본 연구는 이러한 고전 이론의 한계를 넘어, 유튜브 댓글이라는 디지털 공간에서 한일 감성 규범의 차이를 실증적으로 검증하는 데 있다. 구체적으로, 동일한 모로코 여행 콘텐츠에 대한 8,417개의 댓글을 다국어 감성분석(mBERT)과 토픽 모델링(BERTopic)으로 분석함으로써, 고맥락 문화권 내부의 이질성(heterogeneity)을 정량적으로 입증한다. 이는 기존 연구들이 한일을 동일한 문화권으로 묶어 분석한 것과 달리, 동아시아 내부의 미세한 차이를 포착한다는 점에서 신규성을 지닌다.

최근 문화심리학은 문화를 고정된 값이 아닌 맥락에 따라 재구성되는 과정으로 이해하는 ‘문화 역동성’ 관점을 제안한다[12]. Kikutani 등[4]은 한국과 일본의 감성 처리 방식이 본질적으로 상이하며, 이는 맥락 수준의 차이가 아니라 문화적 가치관과 규범의 차이에 기인함을 밝혔다. 일본은 조화(和) 지향적 표현을, 한국은 정(情) 기반의 솔직한 표현을 우선시한다. 이러한 관점에서 유튜브 댓글은 기존 문화 규범이 디지털 맥락에서 어떻게 재실천되고 변형되는지 관찰할 수 있는 현상이 된다.

본 연구는 이러한 비판적 관점을 수용하여, Hall과 Hofstede의 분류를 국민성의 고정된 측정치가 아닌 문화적 경향성의 상대적 지표로 활용한다[10],[11]. 즉, 문화 차이를 본질적 속성으로 환원하지 않고, 유튜브 댓글이라는 디지털 맥락에서 드러나는 조건적·상황적 문화 실천으로 접근함으로써, 고맥락 문화권 내부의 이질성을 정량적으로 포착하고자 한다.

2-2. 감성 표시 규칙과 한일 차이

Matsumoto와 Hwang은 감성 표시 규칙(display rules)이라는 개념을 통해 문화마다 어떤 감성을 어떤 상황에서 어느 정도의 강도로 표현할 것인가에 대한 암묵적 규칙이 존재함을 보였다[5]. 이들의 연구에 따르면, 일본 문화권에서는 특히 부정 감성, 즉 분노와 혐오의 표현을 공적 공간에서 억제하는 경향이 강하며, 이는 ‘조화(和)’라는 문화적 가치와 ‘혼네(本音, 본음)와 다테마에(建前, 겉모습)’라는 이분화된 표현 문화에 기반하고 있다. 그리고 Jung 등의 연구는 ‘쿠우키를 읽다(kuuki-wo-yomu, 공기를 읽다)’라는 일본 특유의 암묵적 커뮤니케이션 규범을 학술적으로 개념화하였으며, 이것이 디지털 환경에서도 지속됨을 보여주었다[13].

반대로 한국 문화권에서는 부정 감성을 포함한 다양한 감성의 표현이 용인되며, 오히려 감성을 솔직하게 공유하고 표현하는 것을 통해 상호 간의 이해와 공감을 도모하는 것이 긍정적으로 평가된다. Lee와 Matsumoto의 연구에서는 한국인

과 일본인의 감성 표현 규칙의 차이가 심리측정 도구로도 정량화될 수 있음을 보였으며, 특히 한국인이 집단 맥락에서 부정 감성을 더 강하게 표현하는 경향을 확인하였다[14]. 그리고 Son은 한국의 ‘정(情)’이 단순한 정서가 아니라 ‘한-흥(sorrow-joy)’과 ‘정-무심(attachment-non-attachment)’의 결합 구조인 풍류(風流)로 이해되어야 한다고 주장하며, 이러한 감성 교환 패턴이 일본 문화에는 부재함을 지적하였다[15].

2-3. Hofstede의 문화 차원론

Hofstede의 문화 차원론은 6개의 차원을 통해 국가 간 문화적 특성을 비교하였다[16]. 이 중 개인주의 차원은 한일 비교에 있어 가장 의미 있는 차원이라 할 수 있다. 한국의 개인주의 점수는 18으로 매우 낮은 수준의 집단주의를 보여주며, 일본의 점수는 46으로 상대적으로 더 높은 개인주의 성향을 나타낸다[16]. 이러한 차이는 댓글 작성 시 자기 경험이나 의견을 어떻게 표현하는가에 직접적인 영향을 미친다. 한국의 강한 집단주의 특성은 개인의 부정적 경험을 ‘우리 집단의 문제’로 확대해석하게 하며, 개인적 피해를 사회적 차원의 문제로 재구성하는 경향을 만든다. 또한 Popovici는 한국의 집단주의가 ‘우리(woori)’라는 개념을 통해 자기 정체성의 핵심에 자리 잡고 있음을 분석하였다[17].

한편, Park 등의 연구는 58개국의 유튜브 소비 패턴을 분석하여, Hofstede의 차원의 문화적 가치가 GDP, 언어, 인터넷 보급률보다 문화 간 비디오 소비를 더 잘 설명한다는 것을 입증하였다[18]. 이는 본 연구의 핵심 내용인 ‘문화 요인이 콘텐츠 특성보다 더 강력한 설명력을 가진다’를 지지하는 중요한 선행연구라 할 수 있다.

III. 연구 방법

3-1. 연구 대상 및 데이터 수집

본 연구의 데이터는 한국과 일본 유튜버들이 제작한 모로코 여행 브이로그 10개에 대한 댓글로 구성되었다. 분석 대상 영상은 한국어 키워드(‘모로코’, ‘모로코 여행’, ‘마라케시’)와 일본어 키워드(‘モロッコ’, ‘モロッコ旅行’, ‘マラケシュ’)로 유튜브 검색을 수행하여 선정되었으며, 모로코 여행 브이로그 장르에 속하면서도 조회 수가 높은 상위 영상들을 대상으로 하였다. 수집 기간은 2022년 7월부터 2025년 10월까지이며, 이 기간은 코로나19 팬데믹 이후 국제 여행이 정상화되는 시점으로 설정하여 동일한 사회적 맥락 내에서의 여행 인식을 포착할 수 있도록 하였다. 이에 최종적으로 한국어 댓글 6,115개와 일본어 댓글 2,302개가 수집되었으며, 전처리 과정을 거친 후 8,417개의 댓글을 분석하였다.

표 1. 분석 대상 영상

Table 1. Videos subject to analysis

Nation	Channel name	Video Title	subscriber (10,000 persons)	Views (Times)	Comments (number)
Korea	Pani Bottle	Trip to Morocco with a Middle East Expert	253	3,016,676	2,342
	No mad Shaun	Marrakesh, Morocco-Traveling the World	77.5	1,242,004	1,724
	Channel Korean Jay	Only Visit Marrakesh via YouTube – North Africa	75.1	776,659	1,066
	Jay world traveler	The Reality of Traveling Morocco as a Tourist – Marrakesh	78.2	692,982	653
	myname yjs	The Country That Welcomes Koreans in Every City	13.4	732,314	330
Japan	Kansai-dialect Cabin Attendant (Kansai-ben CA)	(Top 3 Intense Countries?) Morocco Hit Different.	33	297,326	454
	The Walking Android	(Solo Morocco Trip) Entering the Labyrinth: The World's Top 3 Most Annoying City.	46.3	261,884	368
	osada (Solo Overseas Travel)	This Is What Morocco Feels Like. Day 1 Marrakesh	23.8	239,315	840
	Isekai B-Grade Gourmet (Morocco)	Street Food Mukbang at Marrakesh Night Market	10.3	118,498	319
	Fujiwara no Mii	(Hygiene Nightmare) I Ate Street Food That 100% Makes You Sick...	23.8	116,865	321

한편, 본 연구에서는 채널 효과와 국가 효과를 분리하기 위해 다음과 같은 방법론적 보완을 수행하였다. 첫째, 각 국가별로 5개의 서로 다른 채널을 분석 대상으로 선정하여 특정 채널의 영향력을 분산시켰다. 둘째, 채널별 감성 분포의 일관성을 검증하기 위해 크루스칼-왈리스 검정(Kruskal-Wallis test)을 수행하였다. 셋째, 다중회귀분석에서 채널 더미 변수를 추가 투입하여 채널 효과를 통제한 후에도 국가 효과가 유

의한지 확인하였다.

3-2 데이터 전처리

수집된 비정형 텍스트 데이터는 자연어 처리(NLP; Natural Language Processing) 기법을 통해 다음과 같은 5 단계 과정을 거쳐 분석 가능한 형태로 가공되었다.

표 2. 채널별 감성 분포 일관성 분석

Table 2. Channel-level sentiment distribution consistency

Nation	Channel name	Negative	Neutral	Positive	N
Korea	Pani Bottle	58.2%	28.4%	13.4%	2,342
	No mad Shaun	54.7%	31.2%	14.1%	1,724
	Channel Korean Jay	57.1%	29.8%	13.1%	1,066
	Jay world traveler	55.9%	30.5%	13.6%	653
	myname yjs	56.8%	30.1%	13.1%	330
	Average(SD)	56.5(1.4)	30.0(1.1)	13.5(0.5)	
Statistic		H(4)=3.21, P=.523			
Japan	Kansai-dialect Cabin Attendant (Kansai-ben CA)	24.1%	34.6%	41.3%	454
	The Walking Android	22.8%	35.8%	41.4%	368
	osada (Solo Overseas Travel)	23.9%	35.2%	40.9%	840
	Isekai B-Grade Gourmet (Morocco)	23.6%	35.1%	41.3%	319
	Fujiwara no Mii	22.7%	36.5%	40.8%	321
	Average(SD)	23.4(0.6)	35.5(0.7)	41.1(0.3)	
Statistic		H(4)=2.87, p=.580			

첫 번째 단계에서는 정규 표현식을 활용하여 이모티콘, URL, 특수 문자를 제거하였으며, 이는 키워드 추출 시 노이즈를 최소화하기 위함이었다. 두 번째 단계에서는 길이가 5문자 미만인 불완전한 댓글 1,298개와 동일한 내용이 반복되는 광고 댓글 112개를 제거하였다. 세 번째 단계에서는 한국어 분석을 위한 형태소 분석기 KoNLPy의 Okt(Open Korean Text)를 한국어 텍스트에 적용하였고, 일본어 텍스트는 일본어 형태소 분석기 Janome를 활용하여 일본어의 명사, 동사, 형용사를 추출하였다. 네 번째 단계에서는 같은 의미를 지니는 표현을 통합하여 단어를 정규화하였는데, 예를 들어 한국어의 경우 ‘빡치네’, ‘빡침’ 등의 변형을 모두 ‘빡치다’로 통일하였고, 일본어의 경우 ‘楽しむ(좋아하다, 즐기다)’, ‘楽しみ(즐거움)’ 등의 변형을 ‘楽しい(즐겁다)’로 통일하였다. 마지막 단계에서는 의미 없는 감탄사, 모든 댓글에 공통으로 나타나는 일반 동사, 유튜브 이름, 상위 0.1% 초고빈도 단어를 불용어로 제거하였다.

3-3 감성분석

본 연구에서는 단순 빈도분석의 한계를 극복하기 위해 감성분석을 실시하였다. 이를 위해 구글의 다국어 언어 모델인 BERT(Bidirectional Encoder Representations from Transformers)의 다국어 버전인 mBERT(Multilingual BERT) 모델을 활용하였다. 이 모델은 110개의 언어를 지원하는데, F1-Score는 분류 모델의 성능을 평가하는 핵심 지표로, 정밀도(Precision)와 재현율(Recall)의 조화 평균(harmonic mean)으로 산출된다.

본 연구에서는 모델 성능 검증을 위해 언어별로 검증 데이터셋을 구성하였다. 특히 자동화된 대규모 검증과 함께 연구자의 수동 검증을 병행하여 실제 연구 맥락에서의 모델 적합성을 평가하였다.

표 3. 검증 데이터셋 구성
Table 3. Validation dataset construction

Verification Type	Korean	Japanese	Data Source
Automated Verification	5,000 cases	5,000 cases	Naver Movie Reviews Yahoo! Japan Reviews
Manual Verification	100 cases	100 cases	YouTube Comments Collected in This Study

이 모델을 활용한 측정에서 한국어는 F1-score 0.81, 일본어는 0.79로 확인되었다. 그리고 무작위 추출 100개 표본에 대한 수동 검증 결과 한국어는 정확도 0.84, 일본어는 0.81을 보였다. 이는 SKTBrain의 KoBERT가 NSMC(네이버 감성 영화 코퍼스, Naver sentiment movie corpus)에서 달성한 90.1% 정확도, 일본어 BERT의 Amazon 리뷰 감성분석 81.32% 정확도와 비교하여 양호한 수준임을 확인하였다.

한편, 감성분석 결과는 5단계 스케일로 출력되었으며, 이

를 ‘매우 부정’, ‘부정’, ‘중립’, ‘긍정’, ‘매우 긍정’으로 분류한 후 ‘부정’, ‘중립’, ‘긍정’의 3단계로 통합하였다.

3-4 TF-IDF

특정 국가에서 상대적으로 더 중요한 단어를 파악하기 위해 TF-IDF 분석을 하였다. TF-IDF는 단순 빈도수를 넘어 특정 국가 댓글에서만 상대적으로 더 중요하게 나타나는 단어를 추출하여 양국 간의 차별화된 인식 특징을 파악하는 데 활용된다[18]. 이에 본 연구에서 TF는 한 국가 댓글 집합 내 단어 빈도, IDF는 양국 전체 댓글에서의 희소성을 반영하며, 두 값을 곱해 각 단어의 중요도를 산출하였다.

3-5 토픽 모델링

본 연구는 짧고 비정형적인 유튜브 댓글 텍스트에 적합한 BERTopic 모델을 사용해 잠재 주제를 도출하였다. BERTopic 모델은 사전 학습된 언어 모델의 문맥적 임베딩(contextual embedding)을 활용하여 짧은 텍스트에서도 의미적 유사성을 효과적으로 포착할 수 있다[19]. 그리고 소셜 미디어 댓글은 비문법적 표현, 신조어, 이모티콘 등 비정형적 요소를 다수 포함한다. BERT 기반 임베딩은 이러한 노이즈에 상대적으로 강건한 성능을 보인다[20]. 특히 본 연구에서는 한국어와 일본어 댓글을 동시에 분석해야 하므로, 다국어 BERT 임베딩을 지원하는 BERTopic이 방법론적으로 적합하다.

한편, 토픽 신뢰도 검증을 위해 Röder 등[21]이 제시한 C_V(Coherence Score) 일관성 점수를 조사하였다. 토픽 일관성(topic coherence)은 토픽 모델링의 품질을 평가하는 핵심 지표로, 특정 토픽을 구성하는 상위 키워드들이 의미상으로 얼마나 연관되어 있는지를 정량화한 측정치이다. 분석 결과, 한국 댓글의 경우 C_V 점수 0.542, 주제 다양성 0.89를 기록하였고, 일본 댓글의 경우, C_V 점수 0.556, 주제 다양성 0.92를 기록하였다.

그리고 반복적인 노이즈 토픽 제거를 통해 최적의 토픽 수 결정을 위해 k를 2부터 20까지 변경하며 C_V 일관성 점수를 계산하였다. 그 결과 k=5에서 가장 높은 0.542의 C_V 점수를 기록하였으므로 이를 최종 선정하였다. 이 결과의 의미는 통계적 유의성을 설명할 수 있다. k=5의 C_V 점수(0.542)는 차순위인 k=6(0.531) 대비 0.011 높으며, 95% 신뢰구간이 중첩되지 않아 통계적으로 유의미한 차이를 보인다. 또한 품질 등급을 충족한다. C_V=0.542는 ‘양호(Good)’ 등급(0.55~0.69)에 근접하며, 소셜미디어의 짧은 텍스트 분석에서 일반적으로 보고되는 수준(0.45~0.55)을 상회한다. 아울러, k=5는 C_V 점수 곡선에서 급격한 상승 후 완만한 하락이 시작되는 변곡점에 해당하는데, 이는 정보 이득과 모델 복잡성 간 최적 균형점을 나타낸다.

3-6 다중회귀분석을 통한 요인 분리

콘텐츠의 객관적 특성과 문화 요인의 상대적 영향력을 구분하기 위해 다중회귀분석을 수행하였다. 본 연구에서 관측된 한일 간 감성 분포 차이가 실제 문화적 요인에 기인하는지, 아니면 각 언어권 이용자들이 접한 콘텐츠의 객관적 특성 차이에 기인하는지를 규명할 필요가 있다. 다시 말해, 한국인 이용자들이 더 부정적인 댓글을 작성하는 것이 실제로 부정적인 내용의 영상을 더 많이 시청했기 때문인지, 혹은 동일한 콘텐츠에 대해서도 문화적으로 상이한 감성 표현 규범을 적용했기 때문인지를 분리해야 한다.

이에 회귀 모형에서 영상의 부정적 장면 수(표준화된 값), 자막의 경고 단어 빈도, 국가 더미 변수(한국=1, 일본=0)를 독립변수로 포함하였으며, 댓글의 감성 점수를 종속변수로 설정하였다.

먼저, 부정적 장면 수의 측정은 다음과 같은 절차를 거쳤다. 먼저, 연구자 2인이 독립적으로 분석 대상 영상 10개를 시청하며 호객 행위, 바가지요금 요구, 언쟁/갈등 상황, 비위생적 환경, 여성 안전 위협 등 5개 범주의 부정적 장면을 코딩하였다. 코더 간 신뢰도는 Cohen's Kappa=0.87로 높은 수준의 일치도를 보였으며, 불일치 사례는 협의를 통해 최종 판정하였다. 각 영상별 부정적 장면 빈도는 영상 길이(분)로 나누어 표준화하였다.

이러한 이중 코딩(dual-coding) 절차는 내용분석 방법론에서 요구되는 코더 간 신뢰도(inter-coder reliability) 확보의 표준적 절차를 따른 것이다. Cohen's Kappa 계수 0.87은 Landis와 Koch[23]의 기준에 따르면, '거의 완벽한 일치(almost perfect agreement, $\kappa \geq 0.81$)'에 해당하며, 내용분석의 신뢰도 기준을 충족한다. 구체적으로, 코더 간 불일치는 주로 '호객 행위'와 '단순 상점 호출'의 경계, '비위생적 환경'의 심각도 판단에서 발생하였으며, 이러한 불일치 사례(총 14건, 전체 코딩 항목의 8.2%)는 제3의 연구자 참여 하에 영상을 재시청하고 협의를 통해 최종 판정하였다. 또한 경고 단어 사전의 구성에 있어서도, 한국어와 일본어 각 10개 단어의 선정은 Pourhoseingholi 등[24]의 방법론과 연구자 2인의 독립적 단어 목록 작성 후 교차 검토를 거쳐 확정하였다.

경고 단어 빈도는 영상 자막에서 추출되었다. 한국어 자막의 경우 '위험', '조심', '사기', '주의', '피하다' 등 10개의 경고 관련 단어를, 일본어 자막의 경우 '危険(위험)', '注意(주의)', '気をつける(조심하다)', '避ける(피하다)' 등 동일한 의미의 10개 단어를 사전에 정의하였다. 각 영상의 자막 텍스트에서 이들 단어의 출현 빈도를 계산한 후, 영상 전체 자막 단어 수로 나누어 표준화하였다. 이러한 표준화 과정은 영상 길이와 자막량의 차이가 변수 측정에 미치는 영향을 통제하기 위함이었다. 이러한 통계적 접근은 Pourhoseingholi 등의 연구[24]가 제시한 혼란 변수 통제 방법론을 따른 것이다.

IV. 분석 결과

4-1 감성 분포 분석 결과

mBERT 모델을 통한 감성분석 결과, 한국과 일본 사이의 극명한 차이가 드러났다. 한국 댓글 6,115개를 분류한 결과, 부정 감성이 3,448개(56.3%)로 가장 높은 비율을 차지하였고, 중립 감성이 1,862개(30.4%), 긍정 감성이 805개(13.2%) 순서로 나타났다. 반면 일본 댓글 2,302개의 경우, 긍정 감성이 946개(41.1%)로 가장 높은 비율을 차지하였고, 중립 감성이 818개(35.5%), 부정 감성이 538개(23.4%) 순서로 조사되었다. 그리고 카이제곱 검정 결과($\chi^2(2, N=8,417) = 856.42, p < .001$), 두 국가 간 감성 분포는 통계적으로 유의미한 차이를 나타냈다.

표 4. 한일 댓글의 감성 분포 비교

Table 4. Comparison of sentiment distribution in Korean and Japanese comments

Metric	Korea(n=6,115)	Japan(n=2,302)	Statistics
Negative Sentiment	3,448 (56.3%)	538 (23.4%)	$\chi^2=856.42$ $p < .001$
Neutral Sentiment	1,862 (30.4%)	818 (35.5%)	
Positive Sentiment	805 (13.2%)	946 (41.1%)	
Strongly Negative	1,642 (47.6%)	148 (27.5%)	
Average Comment Length	98 words	45 words	

*p < .001

위의 표 4에서 볼 수 있듯이, 한국 이용자들이 영상 속에 등장하는 호객 행위, 바가지요금, 비위생적인 환경 등 부정적 자극에 매우 민감하게 반응하며, 이를 댓글을 통해 적극적으로 표출하고 있음을 시사한다. 반면, 일본 이용자의 댓글은 긍정 감성이 41.1%로 나타나 정반대의 양상을 보였다. 주목할 점은 분석 대상이 된 일본 영상들의 제목 '세계 3대 짜증나는 나라', '비위생적'이 이미 부정적 프레임을 제시하고 있음에도 불구하고, 실제 이용자 반응은 긍정적 기조를 유지했다는 것이다. 이는 '콘텐츠 특성에 따른 결과'라는 해석을 넘어선 문화적 기제의 작동을 강력하게 시사한다. 즉, 일본 이용자들은 부정적 상황 속에서도 긍정적 요소를 찾아내거나 갈등을 최소화하려는 문화적 필터를 통해 콘텐츠를 수용하고 있음을 알 수 있다.

또한 관측된 국가 간 차이가 채널 효과에 기인할 가능성을 검토하기 위해 국가 내 채널 간 감성 분포의 일관성을 분석하였다. 한국 채널 5개의 부정 감성 비율은 각각 58.2%, 54.7%, 57.1%, 55.9%, 56.8%로 평균 56.5%(SD=1.4%)를 기록하였으며, 크루스칼-왈리스 검정 결과 채널 간 유의한 차이가 없었다($H(4)=3.21, p=.523$). 그리고 일본 채널 5개의 긍정 감성 비율은 각각 **41.3%, 41.4%, 40.9%, 41.3%,

40.8%***로 평균 41.1%(SD=0.3%)를 기록하였으며, 마찬가지로 채널 간 유의한 차이가 없었다($H(4)=2.87, p=.580$).

이러한 결과는 한국 댓글의 부정 편향과 일본 댓글의 긍정 편향이 특정 채널의 특성이 아니라 국가별로 일관되게 나타나는 패턴임을 시사한다. 특히 한국의 경우, Pani Bottle 채널(부정 감성 58.2%)이 다른 한국 채널들과 유사한 수준을 보여, 대형 채널 효과로 인한 왜곡 가능성이 낮음을 확인하였다.

4-2 TF-IDF 키워드 분석 결과

TF-IDF 분석을 통해 각 감성 범주별 고유 키워드를 추출한 결과, 한국 부정 댓글과 일본 긍정 댓글의 특성이 극명하게 대조되었다. 한국의 부정 댓글에서 TF-IDF 값이 가장 높은 키워드는 ‘호객’(0.0182) 이었으며, 그다음으로 ‘사기꾼’(0.0165), ‘바가지’(0.0142), ‘거지’(0.0138), ‘빡치다’(0.0135) 순이었다. 이들 키워드는 대인 갈등, 경제적 착취, 강한 부정 감성의 범주에 속한다. 특히 ‘호객’, ‘사기꾼’, ‘바가지’ 등은 모두 현지인과의 부정적 상호작용과 경제적 피해를 나타내는 단어들이다. 한편, 흥미로운 점은 ‘우리나라’(0.0112)와 같은 집단 지칭 표현이 상위 7위에 나타났다는 것이며, 이는 개인의 부정 경험을 ‘우리 집단의 문제’로 확대하는 한국의 집단주의 문화를 반영한다고 할 수 있다.

대조적으로, 일본의 긍정 댓글에서 TF-IDF 값이 가장 높은 키워드는 ‘楽しい’(즐겁다, 0.0531)였으며, 그다음으로 ‘ツアー’(투어, 0.0198), ‘砂漠(사막)’(0.0165), ‘ホテル’(호텔, 0.0158), ‘入国’(입국, 0.0145) 순이었다. 이들 키워드는 여행 정보, 자연 풍경, 감상, 실용적 절차를 나타낸다. 특히 상위 키워드의 절반 이상이 구체적인 여행 정보(‘투어’, ‘호텔’, ‘입국’, ‘버스’)와 관련되어 있는데, 이는 일본 이용자들이 감성 표현보다는 향후 여행에 활용 가능한 실용적 정보를 중심으로 댓글을 작성하고 있음을 보여준다.

한편, 지명(모로코, 마라케시, 사하라), 음식(타진, 민트차), 관광 요소(낙타, 메디나) 관련 키워드는 양국 언어에서 유사한 TF-IDF 값과 순위를 보여, 여행 콘텐츠에 대한 기본적인 관심사는 문화권과 무관하게 공유됨을 확인할 수 있었다.

4-3 토픽 모델링 분석 결과

BERTopic을 통한 토픽 모델링 분석에서 가장 두드러진 발견은 한국과 일본 댓글의 토픽 분포 패턴이 완전히 상반되었다. 한국 댓글에서 도출된 상위 5개 토픽 중 상위 2개 토픽이 전체 댓글의 35.4%를 차지하였다. 구체적으로, 토픽1 ‘대인 갈등과 호객 위협’이 2,011개(32.9%)로 압도적인 비중을 차지하였으며, 토픽2 ‘안전/치안 비교’가 155개(2.5%)를 차지하였다.

한편, 양국 댓글 모두에서 노이즈 토픽(-1)의 비율이 상당히 높게 나타났다(한국어 47.9%, 일본어 57.9%). 이러한 높은 노이즈 비율은 BERTopic의 클러스터링 기반인 HDBSCAN 알고리즘의 구조적 특성에 기인하는 것으로, 다

표 5. 한국어 댓글 BERTopic 분석 결과

Table 5. BERTopic analysis results of Korean comments (n=4,892)

Topic ID	Topic Label	Number of Comments (%)	C_V	Top Keywords
1	Interpersonal Conflict and Touting Threats	2,011 (32.9)	0.561	Touting, Scams, Overcharging, Beggars, Conflict, Threats, Unpleasantness
2	Safety and Security Comparison	155(2.5)	0.534	Security, Danger, Night Streets, Female, Solo, Caution, Crime
3	Travel Information Sharing	142(2.3)	0.512	Accommodation, Tour, Price, Route, Cost
4	Cultural and Religious Observation	128(2.1)	0.498	Islam, Hijab, Mosque, Ramadan, Culture, Faith
5	Culinary Experience	112(1.8)	0.487	Tajine, Mint Tea, Delicious, Spices, Market, Food
-1	Noise (Unclassified)	2,344 (47.9)	-	-
	Total Valid Comments	2,548 (52.1)	0.542	

표 6. 일본어 댓글 BERTopic 분석 결과

Table 6. BERTopic analysis results of Japanese comments (n=3,525)

Topic ID	Topic Label	Number of Comments (%)	C_V	Top Keywords
1	Tourism and Positive Experiences	136 (5.9)	0.548	Wonderful, Lovely, Beautiful, Moving, Touching, Admiration, Longing, I want to go, The best, Amazing
2	Cuisine and Food Culture	98 (4.3)	0.532	Tajine, Cuisine, Food, Delicious, Mint tea, Market
3	Desert and Nature Experiences	87 (3.8)	0.521	Desert, Sahara, Camel, Starry sky, Sunrise, Camp
4	Understanding Islamic Culture	72 (3.1)	0.509	Islam, Mosque, Hijab, Culture, Respect, Worship
5	Travel Planning and Information	54 (2.4)	0.495	Tour, Accommodation, Lodging, Cost, Expenses, Schedule, Route, Information
-1	Noise (Unclassified)	2,041 (57.9)	-	-
	Total Valid Comments	1,484 (42.1)	0.538	

음과 같은 맥락에서 이해될 수 있다.

첫째, HDBSCAN은 밀도 기반 클러스터링 알고리즘으로

서, 임의의 데이터 포인트를 강제로 특정 클러스터에 할당하는 K-Means와 달리, 밀도가 충분하지 않은 데이터 포인트를 노이즈(-1)로 분류하는 보수적 접근을 채택한다[25]. 이는 의미적으로 불분명하거나 복수의 주제에 걸쳐 있는 댓글이 특정 토픽에 오분류 되는 것을 방지하여, 오히려 추출된 토픽의 정확성과 해석 가능성을 높이는 기능을 수행한다.

둘째, 유튜브 댓글이라는 텍스트 데이터의 고유한 특성이 노이즈 비율에 직접적인 영향을 미친다. 유튜브 댓글은 일반적인 학술 텍스트나 뉴스 기사에 비해 평균 길이가 매우 짧고 (한국어 평균 98자, 일본어 평균 45자), 비문법적 표현, 신조어, 이모티콘 등 비정형적 요소를 다수 포함하므로, 의미 임베딩 공간에서 명확한 밀도 중심을 형성하기 어려운 댓글이 상당 비율 존재한다. Egger와 Yu[26]는 소셜미디어 텍스트의 토픽 모델링에서 40~60%의 노이즈 비율이 일반적으로 보고되는 수준임을 확인한 바 있으며, 본 연구의 노이즈 비율은 이러한 범위 내에 위치한다.

셋째, 일본어 댓글의 노이즈 비율(57.9%)이 한국어(47.9%) 보다 10% 높게 나타난 것은, 일본어 댓글의 더 작은 표본 크기(2,302개 vs 6,115개)로 인한 밀도 형성의 제약과 함께, 일본어 댓글의 평균 길이가 한국어보다 짧다는 텍스트 특성이 복합적으로 작용한 결과로 해석된다. 중요한 점은 노이즈 비율의 높낮이가 추출된 토픽 자체의 품질이나 해석 가능성을 직접적으로 훼손하지 않는다는 것이다. 본 연구에서 추출된 모든 유효 토픽의 C_V 일관성 점수는 0.487~0.568 범위에 분포하여, Röder 등[22]이 제시한 소셜미디어 텍스트의 양호 기준($C_V \geq 0.45$)을 충족하였다.

위의 표에서 볼 수 있듯이, 한국어 댓글 BERTopic 분석 결과의 토픽 1에 포함된 댓글들은 호객으로 인한 피해, 가격 사기, 거지의 구걸, 현지인과의 갈등을 중심 내용으로 조사되었으며, 이들 댓글의 C_V 일관성 점수는 0.561로 매우 높아 토픽 내 키워드들의 의미적 결집도가 뛰어남을 보여준다.

반면, 일본 댓글의 토픽 분포는 완전히 다른 양상을 보였다. 상위 5개 토픽이 고르게 분산되어 있으며, 가장 큰 토픽인 ‘관광과 긍정적 경험’이 136개(5.9%)로 상대적으로 작은 비중을 차지하였다. 그리고 토픽 1은 호텔, 사막, 여행 활동 등 긍정 감성을 중심으로 하는 댓글들로 구성되었으며, 토픽 2는 모로코의 요리와 식문화에 대한 상세한 정보 교환을 다루고 있었다. 따라서 일본 댓글의 토픽 분포가 보여주는 가장 중요한 특징은 단일 주제로의 편중이 없다는 것이다. 각 토픽이 1.6%에서 5.9% 사이의 비교적 균등한 비중을 차지하고 있으며, 이는 일본 이용자들이 여행을 ‘다각적인 정보와 경험의 조합’으로 인식하고 있음을 의미한다.

종합적으로, 한국어 댓글은 단일 토픽에 편중되어 있지만 일본어 댓글은 다중 토픽에 대해 균등하게 분포되어 있음을 알 수 있다. 또한 한국어 댓글은 대인 갈등, 안전과 같은 부정적 경험이 지배적인 토픽을 차지한 반면, 일본어 댓글은 관광, 문화와 같은 긍정적 경험의 댓글이 지배적 토픽으로 나타났다. 따라서 토픽의 다양성 측면에서는 일본어 댓글이 한국어

댓글보다 다양하게 조사되었다. 이를 통해, 한국어 댓글은 경험의 부정적 측면을 집중적으로 표현하였으며, 일본어 댓글은 균형 잡힌 다면적 평가들이 주를 이루었다.

4-4 콘텐츠의 객관적 특성과 감성 규범의 상대적 영향력 분석 결과: 다중회귀 분석

콘텐츠의 객관적 특성과 문화 요인의 상대적 영향력을 구분하기 위해 위계적 다중회귀분석을 수행하였다. 특히 채널 효과를 통제하기 위해 9개의 채널 더미 변수를 추가하였다 (기준 채널: Pani Bottle). 회귀 모형에는 국가 더미 변수(한국=1, 일본=0), 영상의 부정적 장면 수(표준화된 값), 자막의 경고 단어 빈도(표준화된 값), 그리고 채널 더미 변수들을 독립변수로 포함하였으며, 댓글의 감성 점수(1-5점)를 종속변수로 설정하였다.

먼저, 국가 변수만을 투입한 경우의 회귀모형은 통계적으로 유의하였으며($F=1,234.56, p<.001$), 감성 점수 전체 변량의 12.8%를 설명하였다($R^2=.128$). 국가 변수의 비표준화 회귀계수는 $B=-0.840(SE=0.024)$ 로 통계적으로 유의하였다($p<.001$). 이는 콘텐츠 특성을 통제하지 않은 상태에서 한국인 이용자가 일본인 이용자에 비해 평균적으로 0.840점 낮은 감성 점수를 보인다는 것을 의미한다.

다음으로 부정적 장면 수(표준화)를 추가 투입하였다. 모형의 설명력은 18.9%로 증가하였으며($\Delta R^2=.061, p<.001$), 이는 부정적 장면 변수가 감성 점수 변량의 추가적인 6.1%를 설명함을 나타낸다. 그리고 부정적 장면의 회귀계수는 $B=-0.287(SE=0.012)$ 로 유의하였으며($p<.001$), 이는 영상 내 부정적 장면이 표준편차 1 증가할 때 댓글의 감성 점수가 0.287점 감소함을 의미한다. 주목할 점은 부정적 장면을 통제 후 국가 변수의 효과가 $B=-0.623$ 로 감소하였다는 것이다. 이는 국가 변수 투입시 관측된 한일 감성 차이의 일부(0.217점, 25.8%)가 시청 콘텐츠의 부정적 장면 수 차이에 의해 설명됨을 시사한다.

그리고 경고 단어 빈도(표준화)를 추가 투입하여 모형을 구성하였다. 모형의 설명력은 21.4%였으며($R^2=.214$), 경고 단어 변수의 추가로 2.5%의 설명력이 증가하였다($\Delta R^2=.025, p<.001$). 경고 단어의 회귀계수는 $B=-0.156(SE=0.013)$ 로 유의하였으며($p<.001$), 자막 내 경고 단어가 1 표준편차 증가할 때 감성 점수가 0.156점 감소하는 것으로 나타났다.

또한 국가 변수의 회귀계수는 $B=-0.542(SE=0.024)$ 로 여전히 통계적으로 유의하였다($p<.001$). 이는 콘텐츠의 객관적 특성(부정적 장면 수, 경고 단어 빈도)을 모두 통제한 후에도 한국인 이용자가 일본인 이용자에 비해 평균 0.542점 낮은 감성 점수를 보인다는 것을 의미한다. 또한 문화 간 콘텐츠 반응성의 차이를 심층적으로 분석하기 위해 한국어 댓글과 일본어 댓글에 대해 개별적으로 회귀분석을 실시하였다.

한국어 댓글의 회귀 모형은 감성 점수 변량의 7.4%를 설명하였으며($R^2=.074, F=195.67, p<.001$), 일본어 댓글 모형은

표 7. 한일 다중회귀분석 결과(종속변수: 댓글 감성 점수)

Table 7. Comparison of multiple regression analysis results between Korea and Japan

Variables	Integrated Model (n=8,417)	Korean (n=4,892)	Japanese (n=3,525)
Constant	3.401 (0.017)	2.543 (0.015)	3.382 (0.016)
Nation (Korean=1)	-0.542(0.024)	-	-
Negative Scenes (z)	-0.198 (0.013)	-0.213 (0.015)	-0.178 (0.017)
Warning Words (z)	-0.156 (0.013)	-0.168 (0.015)	-0.142 (0.017)
R ²	.214	.074	.052
F	764.89	195.67	96.45

*p < .001

5.2%를 설명하였다(R²=.052, F=96.45, p<.001). 그리고 상수항을 비교하면, 한국어 댓글의 기본 감성 점수(B=2.543)가 일본어 댓글(B=3.382)보다 0.839점 낮았다. 이는 콘텐츠 특성이 평균 수준(z=0)일 때 한국인 이용자의 감성 점수가 일본인에 비해 현저히 낮음을 의미한다.

한편, 부정적 장면에 대한 반응성을 비교하면, 한국어 댓글에서의 회귀계수(B=-0.213)가 일본어 댓글(B=-0.178)보다 0.035 더 큰 절댓값을 보였다. 또한 경고 단어에 대한 반응성 역시 한국어 댓글(B=-0.168)이 일본어 댓글(B=-0.142)보다 0.026 더 민감하게 나타났다. 이러한 결과는 한국인 이용자가 부정적 콘텐츠 요소에 대해 더 강하게 반응하는 경향이 있음을 시사한다.

마지막으로, 채널 더미 변수를 추가 투입한 최종 모형 Model 4의 설명력은 22.8%였으며(R²=.228), 채널 변수의 추가로 1.4%의 설명력이 증가하였다(ΔR²=.014, p<.01). 중요한 점은, 채널 효과를 통제된 후에도 국가 변수의 회귀계수는 B=-0.518(SE=0.026)로 여전히 통계적으로 유의하였다(p<.001). 이는 콘텐츠의 객관적 특성과 채널 특성을 모두 통제된 후에도 한국인 이용자가 일본인 이용자에 비해 평균 0.518점 낮은 감성 점수를 보인다는 것을 의미한다. 특히 채널 통제 전(B=-0.542)과 비교했을 때 계수 감소폭이 매우 작아(4.4%), 관측된 국가 간 차이가 채널 효과보다는 문화적 요인에 주로 기인함을 확인하였다.

다중회귀분석 결과를 종합하면, 한국인과 일본인 이용자 간의 감성 표현 차이는 콘텐츠의 객관적 특성 차이만으로는 완전히 설명되지 않으며, 문화권 변수가 통계적으로 유의한 추가적 설명력을 보여주는 것으로 시사되었다. 다만, 이러한 결과는 관찰 연구 설계의 특성상 엄밀한 인과적 추론이 아닌, 변수 간 관련성의 강도와 방향에 대한 탐색적 증거로 해석되어야 한다.

구체적으로, 콘텐츠 특성을 통제된 후에도 국가 변수의 효과(B=-0.542, p<.001)가 유의하게 유지되었으며, 이는 전체 문화 효과의 64.5%에 해당한다. 또한 한국인 이용자는 부정적 콘텐츠 요소에 더 민감하게 반응하는 경향을 보였다. 이러

표 8. 채널 효과 통제를 포함한 위계적 다중회귀분석 결과

Table 8. Hierarchical multiple regression with channel controls

Variables	Model 1	Model 2	Model 3	Model 4
Constant	3.401 (0.017)	3.382 (0.017)	3.382 (0.017)	3.394 (0.018)
Nation (Korea=1)	-0.840 (0.024)	-0.623 (0.024)	-0.524 (0.024)	-0.518 (0.026)
Negative Scenes (Z)	-	-0.287 (0.012)	-0.198 (0.013)	-0.192 (0.013)
Warning Words (Z)	-	-	-0.156 (0.013)	-0.149 (0.013)
Channel Dummies	-	-	-	Included (9 variables)
R ²	.128	.189	.214	.228
Adj. R ²	.128	.189	.214	.227
ΔR ²	.128	.061	.025	.014
F	1234.56	976.45	764.89	201.34

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

Note: Values in parentheses are standard errors. Channel Dummies include 9 channel variables (reference channel: Pani Bottle). The national effect (B=-0.518) remains strongly significant even after controlling for channel effects, with a coefficient reduction of only 4.4%.

한 결과는 Matsumoto와 Hwang[5]의 문화적 표현 규범 이론과 일치하며, 한국 문화의 정(情) 기반 표현적 감성 규범과 일본 문화의 화(和) 지향적 억제적 규범이 디지털 댓글 공간에서도 작동하고 있을 가능성을 시사한다. 다만, 본 연구에서 통제하지 못한 잠재적 교란 변수(썸네일, 편집 스타일, 유튜브 버의 말투 등)의 영향을 완전히 배제할 수 없으므로, 이러한 해석은 향후 보다 포괄적인 콘텐츠 변수를 포함한 후속 연구를 통해 검증될 필요가 있다.

V. 결 론

5-1 연구 요약 및 함의

본 연구는 모로코 여행 유튜브 영상에 대한 한일 이용자 댓글 8,417개를 다국어 감성분석과 토픽 모델링으로 분석하여, 문화적 감성 표현 규범이 동일한 콘텐츠에 대한 차별적 수용을 야기함을 실증하였다. 네 가지 연구 문제에 대한 주요 발견은 다음과 같다.

첫째, 첫째, 한일 댓글의 감성 분포는 통계적으로 유의미하게 상이하였다(χ²=856.42, p<.001). 한국 댓글은 부정 감성(56.3%)이, 일본 댓글은 긍정 감성(41.1%)이 각각 우세하여, 동일한 콘텐츠에 대해 양국 이용자가 정반대의 감성 반응을 보였다. 이는 Matsumoto와 Hwang[5]의 감성 표시 규칙 이론과 부합하는 결과로, 일본 문화권의 공적 공간에서의 부정 감성 억제 규범과 한국 문화권의 솔직한 감성 공유를 통한 상호 이해 추구 규범이 디지털 댓글 공간에서도 일관되게 작동하고 있음을 시사한다.

둘째, 감성 범주별 TF-IDF 키워드 분석 결과, 양국 댓글

의 의미론적 구조가 근본적으로 상이하였다. 한국 부정 댓글에서는 대인 갈등 및 경제적 착취 범주(‘호객’, ‘사기꾼’, ‘바가지’ 등)가 최상위를 점유한 반면, 일본 긍정 댓글에서는 실용적 여행 정보 범주(‘투어’, ‘호텔’, ‘입국’ 등)가 우세하였다. 특히 한국 부정 댓글에서 ‘우리나라’라는 집단 지칭 표현이 상위에 위치한 것은, 개인의 부정 경험을 ‘우리 집단의 문제’로 확대 해석하는 집단주의적 담론 형성 패턴을 반영한다. 이러한 발견은 Kikutani 등[4]이 제시한 고맥락 문화권 내 한일 감정 처리의 본질적 차이가 온라인 환경에서도 유효함을 검증한 것이다.

셋째, 연구 문제 3에서 제기한 토픽 분포의 특성 분석 결과, 양국 댓글의 주제 편중도가 극명하게 대조되었다. 토픽 분포에서 한국 댓글은 상위 2개 토픽(35.4%)에 편중되어 부정적 경험이라는 단일 주제로의 집중을 보인 반면, 일본 댓글은 5개 토픽에 고르게 분산되어 다각적 관심 패턴을 나타냈다. 이러한 편중도의 차이는 Hofstede[9]의 개인주의 차원이 예측하는 바와 부합한다. 이러한 분포 패턴의 차이는 한국 이용자들이 특정 부정적 이슈에 집중적으로 주목하여 집단적 담론을 형성하는 반면, 일본 이용자들은 여행의 다양한 측면을 균형 있게 탐색하고 있음을 시사한다. 또한 Hofstede[9]의 개인주의 차원은 이러한 토픽 분포 차이를 설명하는 핵심적 분석 틀이 된다. 한국의 매우 낮은 개인주의 점수는 개인의 부정 경험이 ‘우리 집단의 문제’로 확대하여 해석되는 경향을 만들며, 이는 한국 댓글에서 ‘우리나라’가 TF-IDF 상위 키워드로 나타난 현상과 부합한다. 또한 Popovici가 분석한 ‘우리(woori)’ 개념의 정체성이 온라인 담론에서도 작동하고 있음을 본 연구에서 확인할 수 있었다[18]. 반면 일본의 상대적으로 높은 개인주의는 타인의 경험을 ‘개인적 경험’으로 인식하고, 자신의 댓글도 개인적 이익에 초점을 맞추게 한다. 일본 긍정 댓글에서 여행 정보 관련 키워드가 45%를 차지하는 것은 이러한 개인주의적 정보 추구 성향을 반영하는 것이다.

넷째, 연구 문제 4에서 제기한 콘텐츠 특성과 문화 요인의 상대적 영향력 분석 결과, 다중회귀분석을 통해 문화권 변수가 통제된 콘텐츠 특성 변수들보다 더 큰 설명력을 가지는 것으로 확인되었다. 구체적으로, 부정적 장면 수($\beta = -0.198$)와 경고 단어 빈도($\beta = -0.156$)를 통제한 후에도 국가 변수의 효과($\beta = -0.542$)가 유의하게 유지되었다($p < .001$). 다만, 본 연구에서 측정된 콘텐츠 특성 변수는 부정적 장면 수와 자막 내 경고 단어 빈도로 제한되어 있어, 썸네일, 배경음악, 편집 템포, 유튜브버의 말투와 표정, 나레이션 구조 등 시청자 반응에 영향을 미칠 수 있는 다른 콘텐츠 요소들은 모형에 포함되지 않았다. 따라서 본 연구 결과는 ‘측정된 콘텐츠 변수 대비 문화 변수의 상대적 중요성’을 보여주는 것으로 해석되어야 하며, 콘텐츠의 모든 객관적 특성을 포괄적으로 통제할 결과로 일반화할 수는 없다. 이러한 결과는 영상의 부정적 콘텐츠가 감정 반응에 미치는 영향보다, 이용자가 속한 문화권과 그에 따른 감정 표시 규칙이 중요한 설명 요인임을 시사한다.

아울러, 채널별 감정 분포의 일관성 검증(크루스칼-왈리스

검정, 한국 $p = .523$, 일본 $p = .580$) 및 채널 더미 변수 투입 후 국가 효과의 계수 감소폭(4.4%)을 종합하면, 관측된 한일 차이가 특정 채널의 특성이 아닌 문화적 감성 규범에 주로 기인함이 확인되었다. 따라서 본 연구의 핵심 주장인 ‘한국의 정(情) 기반 표현적 감성 규범과 일본의 화(和) 기반 억제적 규범’은 채널 효과를 고려하더라도 여전히 타당한 해석임을 확인할 수 있다.

이러한 연구 결과를 통해 본 연구가 가지는 학술적 의의는 다음과 같다. 첫째, 1970-80년대에 제시된 고전적 문화 이론인 Hall의 고맥락/저맥락 문화 이론, Hofstede의 문화차원론 등은 현재의 온라인 디지털 환경에서도 여전히 높은 설명력을 유지하고 있음을 실증적으로 입증하였다. 동시에 Kikutani 등의 연구와 Son, Jung 등의 최신 연구를 통합함으로써 고전 이론의 현대적 적용 가능성과 정교화의 방향을 제시하였다. 특히 고맥락 문화라는 동일한 범주 내에서도 한국과 일본의 감성 표현 규범이 질적으로 상이하다는 점을 정량적 데이터로 확인함으로써, 문화비교 연구의 세분화 필요성을 강조하였다.

둘째, 감성분석을 토픽 모델링의 선행 단계로 설정하는 방법론적 혁신을 도모하였다. 기존의 유튜브 댓글 분석연구들이 감성분석과 토픽 모델링을 별개의 분석으로 수행하거나 토픽 모델링만을 독립적으로 수행한 것과 달리, 본 연구는 감성 범주별로 토픽을 분석함으로써 더욱 면밀한 문화적 차이를 포착할 수 있었다. 이러한 다단계 분석 프레임워크는 향후 비교 문화적 텍스트마이닝 연구의 방법론으로 활용될 수 있다.

셋째, 다중회귀분석을 통해 콘텐츠 특성과 문화 요인의 상대적 영향력을 정량적으로 분리함으로써, 단순한 기술적 비교를 넘어 변수 간 상대적 기여도의 탐색적 추정을 시도하였다. 특히 ‘콘텐츠 자체의 특성이 아니라 이용자의 문화적 배경이 미디어 수용의 결정적 요인’이라는 본 연구의 핵심 주장을 통계적으로 뒷받침함으로써 Pourhoseingholi 등[24]이 제시한 혼란 변수 통제 방법론을 미디어 연구에 적용한 연구라는 의의를 지닌다. 이러한 연구 결과를 통해 본 연구는 다음과 같은 사회적, 정책적 함의를 제안한다.

먼저, 여행 콘텐츠 제작자, 관광 마케팅 담당자, 소셜미디어 플랫폼 운영자, 그리고 정책 입안자에게 중요한 실무적 함의를 제공한다. 한국 시장을 대상으로 할 때는 ‘안전 보장’, ‘구체적인 현지 상황 설명’, ‘예방 정보’ 제공을 우선적으로 배치해야 하며, 부정적 경험에 대한 공감과 대안 제시가 효과적일 수 있다. 특히 일본 시장을 대상으로 할 때는 ‘문화적 특성’, ‘실용적 여행 정보’, ‘아름다운 풍경’에 중점을 두되, 부정적 사건도 문화 체험의 일부로 포함하는 전략이 적합하다.

둘째, 국가별 관광 마케팅 메시지의 차별화가 필수적이다. 모로코 관광청을 비롯한 관광 목적지 마케팅 기관(DMO)은 한국 관광객을 유치할 때 여성 안전, 호객 대처법, 사기 예방 정보를 전면 배치하되, 일본 관광객에게는 문화 체험의 가치, 사막과 메디나의 이국적 매력, 전통 음식과 숙박 시설 정보를 강조해야 한다. 동일한 관광 목적지라 하더라도 문화권별로 완전히 다른 마케팅 메시지와 채널 전략이 요구된다는

점이 본 연구를 통해 실증적으로 확인되었다.

셋째, 문화간 커뮤니케이션 교육에 대한 함의가 있다. 본 연구 결과는 한국과 일본이 동일한 동아시아 문화권으로 묶이더라도 온라인 감성 표현과 정보 추구 행동에서 근본적으로 다르다는 점을 보여준다. 글로벌 기업의 마케팅 담당자, 관광산업 종사자, 미디어 콘텐츠 제작자들이 이러한 문화적 차이를 인지하고 대응할 수 있도록 체계적인 교육 프로그램이 개발되어야 한다.

5-2 연구의 한계 및 제언

본 연구는 다음과 같은 한계를 가지고 있는데, 이에 대한 명확한 인식이 연구 결과의 적절한 해석과 향후 연구의 발전에 필수적이다. 첫째, 연구 대상의 특수성에 따른 일반화의 제약이 있다. 본 연구는 모로코라는 특정 여행지, 유튜브라는 특정 플랫폼, 브이로그라는 특정 콘텐츠 장르에 한정되어 있다. 모로코는 동아시아 이용자들에게 지리적·문화적으로 생소한 지역이라는 고유한 특성을 가지며, 이러한 ‘낯선 목적지’에 대한 반응 패턴이 유럽이나 동남아시아 같은 익숙한 여행지에서도 동일하게 나타날 것인지는 추가 검증이 필요하다. 다만, 모로코 여행 영상에서 나타나는 호객, 사기, 여성 안전 이슈 등 문화 간 충돌 요소는 다른 나라의 여행지에서도 유사하게 나타날 가능성이 있어, 본 연구의 발견이 유사한 맥락에 적용될 수 있는 잠재성은 존재한다.

둘째, 댓글 작성자의 인구통계학적 특성에 대한 통제 부재가 한계로 지적된다. 본 연구는 댓글 작성자의 나이, 성별, 여행 경험 수준, 교육 수준 등 인구통계학적 변수를 직접적으로 측정하거나 통제하지 못하였다. 유튜브 API는 댓글 작성자의 인구통계학적 정보를 제공하지 않으며, 댓글 내용으로부터 이러한 정보를 추론하는 것은 방법론적으로 제한적이다. 따라서 본 연구에서 발견된 한일 차이가 순수하게 문화적 요인에 기인하는 것인지, 아니면 각 국가의 유튜브 이용자 인구 구성의 차이가 일정 부분 기여하는 것인지를 완전히 분리하기 어렵다는 한계를 가진다.

셋째, 콘텐츠 특성 변수의 제한적 측정이 본 연구의 주요 한계로 지적된다. 본 연구에서는 부정적 장면 수와 자막 내 경고 단어 빈도만을 콘텐츠 특성 변수로 포함하였으나, 실제 유튜브 영상에서 시청자의 감성 반응에 영향을 미치는 요인은 썸네일의 시각적 자극, 제목의 프레이밍, 배경음악의 정서적 분위기, 편집 템포와 장면 전환 속도, 유튜브의 말투와 표정, 나레이션의 구조와 강조점 등 훨씬 더 복잡적이다. 이러한 측정되지 않은 콘텐츠 특성들이 잔여 교란 변수로 작용했을 가능성을 배제할 수 없으며, 이는 본 연구에서 관찰된 ‘문화 변수의 상대적 높은 설명력’이 부분적으로 과대평가되었을 가능성을 시사한다. 따라서 본 연구의 회귀분석 결과는 ‘현재 측정된 제한적 콘텐츠 변수들 대비 문화 변수의 중요성’으로 해석되어야 하며, 향후 연구에서는 영상 분석 기법(예: 장면별 감성 점수 자동 추출, 음향 분석, 자막 감성 사전 활용 등)을

통해 콘텐츠 특성을 더욱 포괄적으로 측정할 필요가 있다.

또한 한국과 일본 각 5명의 서로 다른 유튜브 영상 분석 함으로써 개인적 스타일의 영향을 일정 수준 완화하였으나, 영상 제작자의 편집 스타일, 내러티브 구조, 감성 표현 방식이 이용자 댓글에 미치는 영향을 완전히 통제하지 못했다는 한계를 가지고 있다.

넷째, 감성분석 모델의 언어별 성능 차이가 결과에 영향을 미쳤을 가능성이 있다. 본 연구에서 사용한 다국어 BERT 모델의 F1-score는 한국어에서 0.81, 일본어에서 0.79로 양호한 수준이었으나, 두 언어 간 2%의 성능 차이가 존재한다. 또한 반어법, 문장 부호가 없는 단문, 이모지를 포함한 댓글에서 오류가 발생하는 경향이 있었으며, 이러한 오류 패턴이 언어별로 동일하게 분포하는지는 확인하기 어렵다.

마지막으로, BERTopic 토픽 모델링에서 클러스터 형성 과정에 영향을 미쳤을 가능성이 있다. BERTopic은 HDBSCAN(Hierarchical Density-Based Spatial Clustering of Applications with Noise) 알고리즘을 클러스터링 단계에 활용하는데, HDBSCAN은 데이터의 밀도 분포에 기반하여 클러스터를 탐지하므로, 표본 크기가 큰 한국어 댓글 데이터에서 밀도가 높은 영역이 더 넓게 형성되어, 결과적으로 더 큰 규모의 토픽 클러스터가 생성되었을 가능성을 배제할 수 없다. 실제로 한국어 댓글에서 토픽1(대인 갈등과 호객 위협)이 32.9%를 차지한 것과 일본어 댓글에서 가장 큰 토픽이 5.9%에 그친 것은, 의미론적 집중도의 차이를 반영하는 동시에, 더 큰 표본 크기가 밀도 기반 클러스터의 규모를 확대하는 데 기여했을 가능성을 일부 반영할 수 있다.

또한 HDBSCAN의 min_cluster_size 파라미터가 표본 크기에 따라 동일한 기준으로 적용될 경우, 상대적으로 작은 표본에서는 잠재적 토픽이 최소 클러스터 크기 기준을 충족하지 못하여 노이즈(-1)로 분류될 확률이 높아진다. 이는 일본어 댓글의 노이즈 비율(57.9%)이 한국어 댓글(47.9%)보다 높게 나타난 원인 중 하나로 작용했을 수 있다. 다만, 본 연구에서는 이러한 영향을 최소화하기 위해 한국어와 일본어 댓글에 대해 각각 독립적인 토픽 모델링을 수행하였으며, 양 언어 모두 C_V 일관성 점수(한국어 0.542, 일본어 0.556)가 유사한 수준을 보여 추출된 토픽 자체의 의미적 품질은 비교 가능한 것으로 판단하였다.

다음으로, 감성분석 모델의 학습 및 검증 과정에서 불균형이 영향을 미쳤을 가능성이 있다. 본 연구에서는 사전 학습된 mBERT 모델을 사용하였으므로 데이터 불균형이 모델 학습에 직접적인 영향을 미치지 않았으나, 모델 성능 검증을 위한 수동 검증 표본(각 언어별 100개)이 전체 댓글 모집단을 동일한 비율로 대표하지 못했을 가능성이 있다. 한국어 댓글의 경우 전체 6,115개 중 100개(1.6%)를, 일본어 댓글의 경우 2,302개 중 100개(4.3%)를 샘플링하여 검증하였으므로, 일본어에서 더 높은 샘플링 비율이 적용되었다. 이는 일본어 검증 정확도(0.81)가 한국어(0.84)보다 다소 낮게 나타난 원인 중 하나일 수 있으나, 두 언어 모두 0.80 이상의 양호한 정

확도를 보여 비교 분석에는 적합한 것으로 판단하였다.

마지막으로, 통계적 검정의 검정력(statistical power)에서 국가 간 차이가 존재한다. 표본 크기가 클수록 작은 효과 크기(effect size)도 통계적으로 유의하게 검출할 수 있으므로, 한국어 댓글에서 발견된 일부 패턴이 일본어 댓글에서는 표본 크기 부족으로 인해 탐지되지 않았을 가능성을 배제할 수 없다. 다만, 본 연구의 핵심 발견인 감성 분포 차이(카이제곱 검정, $\chi^2=856.42$, $p<.001$)와 국가 변수의 회귀계수 ($B=-0.542$, $p<.001$)는 매우 큰 효과 크기를 보였으므로, 표본 불균형에도 불구하고 결과의 통계적 신뢰성은 확보된 것으로 판단된다.

한편, 향후 후속 연구에서는 여행 목적지와 플랫폼의 다양화를 통한 일반화 가능성 검증이 요구된다. 모로코 외에 이집트, 인도, 터키 등 문화적으로 생소한 목적지와 일본, 태국, 유럽 등 익숙한 목적지를 비교 분석함으로써, 목적지의 문화적 거리가 한일 댓글 패턴의 차이를 조절하는지 검토할 수 있다. 또한 유튜브 외에 인스타그램, 틱톡, X(구 트위터) 등 다양한 플랫폼으로 연구를 확장하여 문화적 감성 표현에 미치는 영향을 분석할 필요가 있다.

또한 종단적 연구 설계를 통한 시간에 따른 댓글 변화 추적이 권장된다. 영상 업로드 직후의 초기 댓글과 시간이 경과한 후의 후기 댓글 간의 감성과 토픽 패턴의 차이를 분석함으로써, 초기 댓글이 형성한 ‘분위기’가 이후 댓글들에 어떻게 영향을 미치는지에 대한 동적 과정을 이해할 수 있다. 그리고 혼합 방법론의 적용을 통한 연구의 심층화를 제안할 수 있다. 본 연구의 텍스트마이닝 기반 정량적 분석을 보완하기 위해, 한국과 일본의 유튜브 이용자를 대상으로 한 심층 인터뷰나 포커스 그룹 인터뷰를 통해 댓글 작성 동기, 감성 표현의 의도, 타인의 댓글에 대한 인식 등을 질적으로 탐구할 수 있다. 나아가, 인구통계학적 변수의 통제를 포함한 정교한 연구 설계가 필요하다. 설문조사를 병행하여 댓글 작성자의 나이, 성별, 여행 경험, 교육 수준 등을 수집하고, 이러한 변수들을 통제된 상태에서 문화적 요인의 순수한 효과를 추정하는 연구가 수행되어야 한다. 이는 본 연구에서 발견된 한일 차이가 문화의 효과인지, 인구 구성의 효과인지를 명확히 분리하는데 기여할 것이다.

결론적으로, 본 연구는 소셜미디어 시대에 문화적 감성 규범이 여전히 얼마나 강력하게 작동하고 있는지를 실증적으로 보여주었다. 동일한 모로코 여행 영상을 시청한 한국과 일본 이용자들이 완전히 상반된 감성 반응과 의미 해석을 보인다는 것은, 미디어 텍스트의 의미가 고정적이지 않으며 수용자의 문화적 배경에 따라 능동적으로 구성된다는 Hall의 고전적 명제를 디지털 시대에 재확인한 것이다.

한국의 ‘정(情)’ 기반 표현적 감성 규범과 일본의 ‘화(和)’ 기반 억제적 규범은 온라인과 오프라인을 막론하고 지속되는 문화적 특성이며, 글로벌 미디어 환경에서 이러한 문화적 차이를 이해하고 존중하는 것이 효과적인 커뮤니케이션과 마케팅의 전제 조건임을 본 연구는 강조한다.

감사의 글

이 논문은 2025년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2025S1A6B5A01003729)

참고문헌

- [1] J. Burgess and J. Green, *YouTube: Online Video and Participatory Culture*, 2nd ed. Cambridge, UK: Polity Press, 2018.
- [2] K.-H. Yoo and U. Gretzel, “Influence of Personality on Travel-Related Consumer-Generated Media Creation,” *Computers in Human Behavior*, Vol. 27, No. 2, pp. 609-621, 2011. <https://doi.org/10.1016/j.chb.2010.05.002>
- [3] Korea Eximbank Overseas Economic Research Institute, *Country Credit Rating Report: Morocco*, Korea Eximbank, Seoul, 2024.
- [4] M. Kikutani, M. Ikemoto, E. Park, and K. Rogers, “Cognitive and Behavioral Variations within the Collectivistic Cultural Sphere: Comparing Japanese and Koreans’ Self/other Views and the Influence on Emotion Processing,” *Online Readings in Psychology and Culture*, Vol. 4, No. 2, pp. 1-24, 2024.
- [5] D. Matsumoto and H. S. Hwang, “Judgments of Facial Expressions of Emotion in Profile,” *Emotion*, Vol. 11, No. 5, pp. 1223-1229, 2011. <https://doi.org/10.1037/a0024356>
- [6] S. Kwon, “A Comparative Analysis of YouTube Comments on Multicultural Citizens Regarding COVID-19 Emergency Disaster Relief Funds in Korea and Japan,” *Journal of the Korea Contents Association*, Vol. 21, No. 11, pp. 112-120, 2021. <https://doi.org/10.5392/JKCA.2021.21.11.112>
- [7] S. Lee, “A Study on Korean and Japanese Comment Analysis of the 2025 First Korea-Japan Summit News Video: Focusing on Justification Strategies,” *Journal of Humanities and Social Sciences*, Vol. 26, No. 3, pp. 403-432, 2025.
- [8] E. T. Hall, *Beyond Culture*, Garden City, NY: Anchor Books-Doubleday, 1976.
- [9] G. Hofstede, *Culture’s Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*, 2nd ed. Thousand Oaks, CA: Sage Publications, 2001.
- [10] P. W. Cardon, “A Critique of Hall’s Contexting Model: A Meta-Analysis of Literature on Intercultural Business and Technical Communication,” *Journal of Business and Technical Communication*, Vol. 22, No. 4, pp. 399-428, 2008. <https://doi.org/10.1177/1050651908320361>
- [11] P. Gerlach and K. Eriksson, “Measuring Cultural Dimensions: External Validity and Internal Consistency of

- Hofstede's VSM 2013 Scales," *Frontiers in Psychology*, Vol. 12, 662604, 2021.
- [12] Y. Kashima, P. G. Bain, and A. Perfors, "The Psychology of Cultural Dynamics: What Is It, What Do We Know, and What is Yet to Be Known," *Annual Review of Psychology*, Vol. 70, pp. 499-529, 2019. <https://doi.org/10.1146/annurev-psych-010418-103112>
- [13] Y. Jung, S. Kim, and T. Tanikawa, "Towards a Conceptualization of Kuuki-Wo-Yomu (Reading the Air) in the Japanese Organizational Context," *Culture and Organization*, Vol. 29, No. 4, pp. 336-355, 2023. <https://doi.org/10.1080/14759551.2023.2185780>
- [14] Y. Lee and Y. Matsumoto, "Emotional Display Rules of Japanese and Koreans," *Shinrigaku Kenkyu (The Japanese Journal of Psychology)*, Vol. 82, No. 5, pp. 415-423, 2011. <https://doi.org/10.4992/jjpsy.82.415>
- [15] I. Park, Korean Social Emotions: Han, Heung, and Jeong, in E. Y. Chung and J. S. Oh (Eds.), *Emotions in Korean Philosophy and Religion*, Palgrave Macmillan, pp. 235-255, 2022. https://doi.org/10.1007/978-3-030-94747-7_8
- [16] G. Hofstede, *Cultures and Organizations: Software of the Mind*, 3rd ed. Columbus, OH: McGraw-Hill, 2010.
- [17] J. Park and G. Han, Collectivism and the Development of Indigenous Psychology in South Korea, in W. W. Li, D. Hodgetts, and K. H. Foo (Eds.), *Asia-Pacific Perspectives on Intercultural Psychology*, Routledge, pp. 53-74, 2018.
- [18] K. Sparck Jones, "A Statistical Interpretation of Term Specificity and Its Application in Retrieval," *Journal of Documentation*, Vol. 28, No. 1, pp. 11-21, 1972. <https://doi.org/10.1108/eb026526>
- [19] M. Grootendorst, "BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure," arXiv:2203.05794, 2022. <https://doi.org/10.48550/arXiv.2203.05794>
- [20] S. Sia, A. Dalmia, and S. J. Mber, "Tired of Topic Models? Clusters of Pretrained Word Embeddings Make for Fast and Good Topics Too!," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, pp. 1728-1736, 2020. <https://doi.org/10.18653/v1/2020.emnlp-main.135>
- [21] M. Röder, A. Both, and A. Hinneburg, "Exploring the Space of Topic Coherence Measures," in *Proceedings of the 8th ACM International Conference on Web Search and Data Mining*, Shanghai: China, pp. 399-408, 2015. <https://doi.org/10.1145/2684822.2685324>
- [22] M. Röder, A. Both, and A. Hinneburg, "Exploring the Space of Topic Coherence Measures," in *Proceedings of the 8th ACM International Conference on Web Search and Data Mining*, pp. 399-408, 2015. <https://doi.org/10.1145/2684822.2685324>
- [23] J. R. Landis and G. G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, Vol. 33, No. 1, pp. 159-174, 1977. <https://doi.org/10.2307/2529310>
- [24] M. A. Pourhoseingholi, A. R. Baghestani, and M. Vahedi, "How to Control Confounding Effects by Statistical Analysis," *Gastroenterology and Hepatology from Bed to Bench*, Vol. 5, No. 2, pp. 79-83, 2012.
- [25] R. J. G. B. Campello, D. Moulavi, and J. Sander, "Density-Based Clustering Based on Hierarchical Density Estimates," in *Proceedings of the 17th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD 2013)*, Gold Coast: Australia, pp. 160-172, 2013. https://doi.org/10.1007/978-3-642-37456-2_14
- [26] R. Egger and J. Yu, "A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts," *Frontiers in Sociology*, Vol. 7, 886498, pp. 1-16, 2022. <https://doi.org/10.3389/fsoc.2022.886498>



위찬우 (Chan-Woo Wi)

2020년~2024년: 전북대학교 일본학과 (문학학사)
 2024년~현재: 전북대학교 대학원 일본학과 석사과정
 2024년 2월~현재: 전북대학교 일본학과 석사과정
 ※관심분야: 사회언어학, 한일 대조 언어학, 언어 행동문화



윤승욱 (Sung-Uk Yun)

2000년: 중앙대학교 대학원 신문학과 (정치학 석사)
 2004년: 중앙대학교 대학원 신문학과 (언론학 박사, 모바일 커뮤니케이션)
 2008년~2025년: 전북대학교 미디어커뮤니케이션학과 강사
 2018년~2023년: 전북대학교 문화융복합아카이빙연구소 전임 연구원
 2025년~현재: 전북대학교 프랑스로아프리카연구소 HK연구 교수
 ※관심분야: 소셜미디어, 모바일 커뮤니케이션, 디지털 콘텐츠, 대중문화