



토픽모델링을 통한 여성체육학 분야의 연구동향 분석: 한국여성체육학회지를 중심으로

노미영*(국민대학교, 조교수)

국문초록

본 연구의 목적은 토픽모델링을 활용하여 한국여성체육학회지 연구 동향을 분석하고 시간이 지남에 따라 연구주체의 동향이 어떻게 변화하는지 규명하는데 있다. 이를 위해 한국여성체육학회지 창간호인 1987년 1권 1호부터 2021년 35권 4호까지 총 1,277편의 논문을 대상으로 자료를 수집하였다. 수집된 자료는 토픽모델링을 분석하기 위해 토큰화를 통해 명사단어를 추출하고, Stopwords를 사용해 불필요한 단어를 제거하였으며, 동일한 의미의 단어들은 하나의 공통단어로 변환하는 등의 텍스트 전처리 과정을 수행하였다. 이후 Python 3.8 프로그램과 SPSS 22.0을 사용하여 키워드 빈도분석, 단순회귀분석, 토픽모델링의 LDA 알고리즘을 활용하여 토픽을 산출하였다. 연구결과는 다음과 같다. 첫째, 키워드 빈도분석결과, 운동(287회), 여성(207회), 스포츠(161회), 참여(141회), 노인(138회), 무용(104회), 선수(96회), 여가(82회), 건강(38회) 등의 빈도순으로 나타났다. 둘째, 토픽모델링 분석결과, 8개의 주요 연구주체(토픽)가 도출되었고, ‘운동경험’(18.8%), ‘선수 대상 연구’(16.0%), ‘여가활동 참여’(14.4%), ‘노인 체육 및 건강’(12.7%), ‘만족도 관련 연구’(11.6%), ‘여성 스포츠 관련 연구’(9.6%), ‘무용 관련 연구’(8.9%), ‘생리적 변인 연구’(8.0%) 순으로 비중 있게 다루어졌다. 셋째, 시기별 개별 토픽의 연구 동향을 살펴본 결과, Topic 7 ‘무용 관련 연구’는 Cold 토픽, topic 8 ‘생리적 변인’은 Hot 토픽으로 나타나 시간에 따른 연구주체의 동향을 확인할 수 있었다. 이러한 결과는 여성 체육학이 나아가야 할 학문적 발전에 중요한 기초자료로 활용될 수 있을 것이다.

한글주요어 : 토픽모델링, LDA 알고리즘, 한국여성체육학회지, 연구 동향

* 노미영, 국민대학교, E-mail : rohmi0828@kookmin.ac.kr

I. 서론

여성 체육학이란 개개의 여성 또는 다른 사람들(예: 동성과 이성) 간의 상호관계 속에서 발생하는 스포츠의 사회문화적인 현상을 학제적 접근방법을 통해 탐구하는 융합학문으로 정의 내릴 수 있다(유정애, 2014). 한국여성체육학회는 1954년 출범을 시작으로 우리나라 여성 체육학의 학문적 발전과 전문 여성 체육인 양성 및 그 위상을 정립하는데 선도적인 역할을 담당하고 있다(한국여성체육학회, 2014). 특히 한국여성체육학회의 학술지인 한국여성체육학회지는 1987년 창간호를 시작으로 지금까지 총 101권의 학술지를 발행하며 양적인 성장뿐만 아니라 여성 체육 관련 연구의 산실로 자리매김하고 있다.

하지만, 한국여성체육학회는 오랜 역사를 가지고 있음에도 불구하고 여성 체육학이라는 학문의 정체성을 확립하기 위한 지속적인 담론이 부족한 실정이며 이에 본 학회가 더욱 발전해 나가기 위해서는 연구의 지속적인 탐색 및 연구 동향 분석이 필요하다고 하였다(유정애, 2014). 이러한 맥락에서 많은 학자들은 특정 분야의 연구 동향을 파악하는 것은 현시점까지의 연구 성과를 점검하고 그것에 기초하여 미래의 연구과제를 제시하는데 중요한 작업이라고 하였다(변기용, 이석열, 김수홍, 2010; Lee, Wu, & Tsai, 2009; Lin, Lin, & Tsai, 2014). 따라서 한국여성체육학의 학문적 정체성과 미래 과제를 탐색하기 위해서 관련 연구 동향 분석이 필수적이라 할 수 있다.

체육학 분야의 연구 동향에 대한 분석은 2000년대부터 하위 분야별로 체육측정평가(신승윤, 2015; 이용국, 박재현, 2017), 스포츠사회학(박

소영, 이근모, 2007; 장익영, 이원미, 2019), 특수체육학(윤지운, 2017; 이경옥, 하지연, 2000; 정경환, 원영신, 구송광, 2015) 등 체육학 분야의 발전을 위해 지속적으로 이루어지고 있다.

구체적으로 여성 체육에 대한 연구 동향을 분석한 기존연구를 살펴보면, 대부분 내용분석(content analysis)이 주를 이루고 있다(김수현, 2010; 김원정, 홍애령, 2017; 김인형, 이근모, 2005). 선행연구를 살펴보면, 대부분 연구자의 특정 관심 주제에 관한 내용으로, 내용분석의 주요 키워드는 여성 엘리트 스포츠(김인형, 이근모, 2005), 여성의 여가 제약(2010, 김수현), 여학생의 신체활동(김원정, 홍애령, 2017) 등 여성과 스포츠라는 특정 주제에 관한 분석이 주를 이루고 있어 여성 체육에 대한 거시적인 연구 동향을 파악하는데 한계가 있다. 또한, 전통적으로 널리 사용되고 있는 내용분석은 연구방법론적 측면에서 연구주제, 연구대상 및 방법 등을 체계적으로 분석하여 전반적인 연구의 흐름을 파악할 수 있다는 장점이 있지만, 분석과정에서 연구자의 주관적 판단을 배제할 수 없어 외적 타당성을 확보하기 어렵다는 문제가 있다(Danowski, 1993; Rourke & Anderson, 2004).

최근 들어 이러한 내용분석의 연구방법론적 한계를 극복하고 연구 동향을 보다 객관적으로 분석하기 위해 컴퓨터 프로그램을 활용한 토픽모델링(topic modeling) 기법의 연구들이 증가하고 있다. 토픽모델링은 비정형화된 텍스트 내 방대한 문헌 데이터에서 숨겨진 핵심 토픽(주제)을 추출하는 알고리즘 기반 텍스트마이닝(text mining)의 기법 중 하나이다(Blei, 2012). 토픽모델링 방법론

을 적용한 연구 동향 분석 연구는 2010년 초기 문헌정보학 연구 분야를 시작으로, 방대한 데이터를 처리하는데 시간적 제약이 많은 내용분석 방법이 갖는 단점을 보완하고 보다 객관적인 분석이 가능하여 현재는 다양한 연구영역으로 확대되어 사용되고 있다(최성철, 박한우, 2020).

토픽모델링은 텍스트 자료에서 숨겨진 연구의 주제를 추출하는 방식의 확률모델 알고리즘(Blei, 2012; Blei & Laerty, 2006)으로, 이 기법이 사용된 초기에 Deerwester 등(1990)이 제안한 LSI(Latent Semantic Indexing) 알고리즘과 Hofmann(1999)이 제안한 pLSA(Probabilistic Latent Semantic Analysis) 알고리즘, 그리고 Blei, Ng와 Jordan(2003)에 의해 제시된 LDA(Latent Dirichlet Allocation) 알고리즘이 있다. LSI는 문서와 단어 간 매트릭스(term)를 통해 숨겨진 의미를 추출하는 방식으로 토픽모델링 기법이 소개된 초기에 가장 기본이 되는 알고리즘이었지만, 계산방식에서 SVD(Singular Value Decomposition)의 특성상 새로운 문서에 대한 업데이트가 어렵다는 단점이 있다. 이후 이러한 단점을 개선한 LDA(Blei, Ng, & Jordan, 2003) 알고리즘은 현재까지 가장 널리 사용되고 있다.

LDA 알고리즘은 각 문서 내 존재하는 다수의 토픽들 간의 확률분포를 바탕으로 특정 토픽의 중요도를 알 수 있으며 토픽들을 독립적으로 분류하고 일관성 있는 주제를 추출하는데 효과적이라는 장점을 가지고 있다(Blei, Ng, & Jordan, 2003). 또한, 확률적 모델인 LDA는 LSI와 비교해서 모듈성(modularity)과 확장성(extensibility) 측면에서

우수하여 LSI가 할 수 없는 더 복잡한 모델에 사용하기에 더 적합한 알고리즘이다(Blei, Ng, & Jordan, 2003). 실제로 최성철과 박한우(2020)는 국내 사회과학 분야를 중심으로 토픽모델링 연구 동향을 살펴본 결과, 90% 이상이 LDA를 사용하고 있다고 보고하였다. 이에 본 연구에서는 토픽모델링에 보다 더 적합한 알고리즘으로 학계에서 가장 널리 활용되고 있는 LDA 알고리즘을 적용할 것이다.

최근 체육학 분야에서도 토픽모델링을 적용한 한국체육측정평가학회지 연구 동향 분석(이용국, 박재현, 2017), 토픽모델링을 적용한 한국스포츠사회학회지의 연구주제 분석(장익영, 이원미, 2019) 등 해당 학문 분야의 연구특성과 동향을 파악하기 위해 토픽모델링 기법을 사용하고 있다. 이에 반해 한국여성체육학회지는 양적, 질적으로 성장하였음이 입증되었지만, 한국여성체육학회지의 연구특성 및 연구 동향을 심층적으로 파악한 연구는 미비한 실정이다. 김지훈, 원영신, 김수현(2016)의 연구에서 한국여성체육학회지를 대상으로 핵심어 연결망 분석을 통해 여성 체육 관련 주요 키워드와 거시적인 지식구조를 살펴볼 수 있었지만, 네트워크 연구분석의 특성상 미시적인 측면에서의 연구주제별 특성을 탐색할 수 없다는 한계점이 존재한다.

따라서 본 연구에서는 한국여성체육학의 학문적 정체성과 여성 체육의 미래 과제를 탐색하기 위하여 한국여성체육학회지 창간호인 1987년부터 2021년까지 등재된 논문을 대상으로 토픽모델링을 적용하여 연구 동향을 분석하고자 한다. 또한, 시간이 지남에 따라 연구주제의 동향이 어떻게 변화하는지 규명하고자 한다. 이러한 시도는 현재까

지의 한국여성체육학회지의 연구 동향을 탐색하여 미래의 연구과제를 제시하고 여성 체육학의 학문적 발전을 위한 기초자료로 의의가 있을 것이다.

II. 연구방법

1. 분석대상 및 자료수집

본 연구는 한국여성체육학회지 창간호인 1987년 1권 1호부터 2021년 35권 4호에 게재된 총 1,277편의 논문을 분석대상으로 선정하였다. 자료 수집을 위해 디비피아(DBpia) 홈페이지(<https://www.dbpia.co.kr>)에서 학술지 검색 메뉴를 활용하여 창간호부터 현재까지 한국여성체육학회지에 수록된 모든 논문의 연도, 권, 저자명, 논문 제목 등을 웹 크롤링(crawling)을 통해 수집하였다. 최종 분석대상의 연도별 빈도수와 출간 동향에 대한 결과는 <그림 1>과 같다.

2. 분석절차

본 연구는 여성 체육학의 연구 동향을 분석하기 위해 한국여성체육학회지에 등재된 논문 제목 데이터를 활용하였다. 분석대상은 1987년 창간호부터 2021년 12월 31일까지 총 1,277편의 논문을 대상으로 하였다. 이후 토픽모델링 분석에 적합한 형태로 변환하기 위해 수집된 자료를 기초로 텍스트 전처리(text preprocessing) 과정을 수행하였다.

텍스트 전처리 과정은 Python 프로그램을 활용하여 아래 순차적 단계에 따라 진행하였다. 1단계는 수집된 자료를 단어 단위로 분리하는 토큰화(tokenization) 작업을 통해 명사단어를 추출하였다. 2단계는 불필요한 단어(예: 연구, 영향, 분석 등)를 제거하기 위한 작업으로 Stopwords를 사용하였다. 3단계는 동일한 의미이지만 다른 방식으로 표현된 단어들은 하나의 공통단어(예: 여성과 여자 → 여성, 노인과 노년 → 노인 등)로 변환하였다. 마지막으로 전체 자료에서 5회 이상 반복적으로 관찰

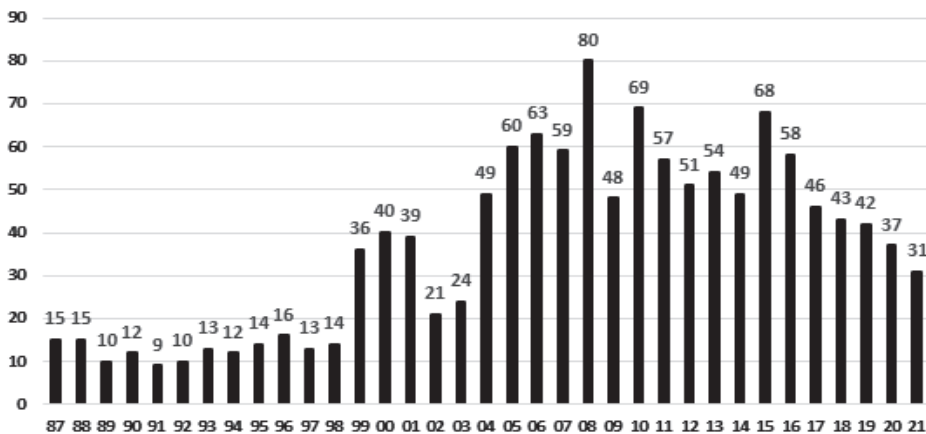


그림 1. 연도별 한국여성체육학회지 출간 동향

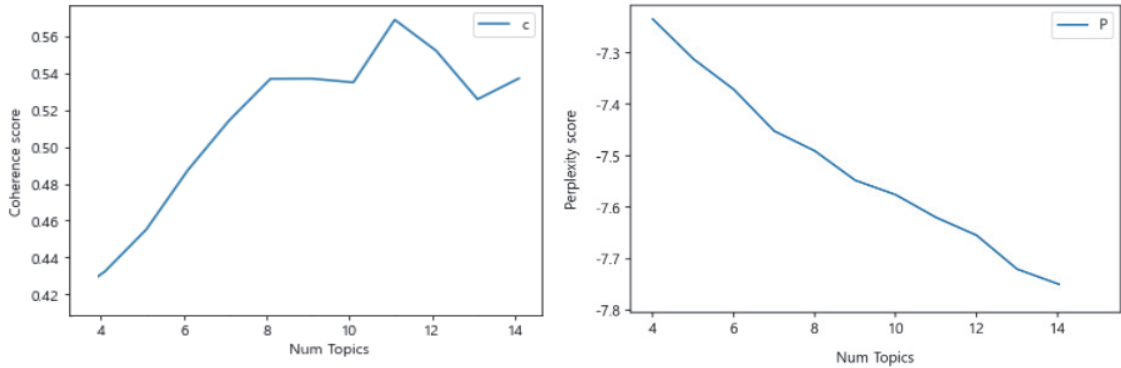


그림 2. 토픽 개수의 따른 응집도(왼쪽)와 혼란도(오른쪽) 그래프

되는 단어는 하나의 단어(예: 신체와 활동 → 신체_활동, 생활과 체육 → 생활_체육 등)로 변환하였다.

텍스트 전처리 과정을 모두 마친 데이터를 기초로 키워드 빈도분석을 실시하였다. 이후 토픽모델링을 통해 8개의 토픽을 산출하였다. 본 연구에서는 토픽모델링 분석결과의 타당성을 확보하기 위해 객관적인 두 지표인 응집도(coherence)와 혼란도(perplexity)에 기초하여 토픽 수를 산정하였다(그림 2). 가장 적합한 토픽의 수는 응집도의 그래프에서 기울기가 0 또는 부적으로 변화되기 전 가장 높은 응집도 점수에서의 토픽의 수로 선정할 수 있다(Prabbakaran, 2018). 또한, 응집도는 높을수록(Newman et al., 2010) 혼란도는 반대로 값이 낮을수록 좋은 성능을 지녔다고 평가한다(Griffiths & Steyvers, 2004). 이에 따라 응집도 점수는 토픽의 수가 8개 또는 11개일 때 가장 적당하다는 것을 확인할 수 있었으며 혼란도는 토픽의 수가 증가할수록 감소하는 경향이 나타났기 때문에 두 지표의 결과와 산출된 토픽들의 해석가능성을 고려하여 토픽의 수는 최종적으로

8개로 판단하였다.

구체적인 토픽별 해석 과정은 토픽별 상위 키워드와 관련 논문들의 초록을 참고하여 전문가 회의(체육학 박사 2명, 체육측정평가 박사 1명)를 통해 문맥을 분석한 후 토픽별 주요 연구주체를 정리하였다. 마지막으로 시간에 따라 개별 토픽들에 대한 연구동향이 어떻게 변화하고 있는지 규명하기 위해 연도별 토픽점유율을 산출하여 독립변수는 연도로, 종속변수는 연도별 토픽점유율을 사용하여 단순회귀분석을 실시하였다. 분석을 통해 산출된 회귀분석의 회귀계수 값을 이용하여 토픽의 연도별 추이를 판단하였으며(Griffiths & Steyvers, 2004), 통계적으로 유의한 회귀계수 값이 양수(+)이면 Hot 토픽(상승 연구주제), 음수(-)이면 Cold 토픽(하향 연구주제), 그리고 통계적으로 유의하지 않은 값은 Neutral(중립) 토픽으로 해석하였다(김창식, 최수정, 박기영, 2017). 전체적인 분석절차는 다음 <그림 3>과 같다.

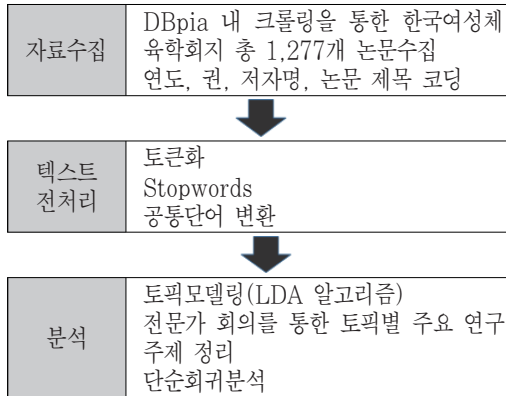


그림 3. 자료수집 및 분석절차

3. 자료처리방법

본 연구의 목적은 토픽모델링을 활용하여 한국여성체육학회지에 게재된 연구주제의 동향을 분석하는 것이다. 토픽모델링은 텍스트 자료 내 핵심어를 바탕으로 연구의 주제(topic)를 추출하는 확률모델 알고리즘으로, 특정 연구 동향을 분석하는데 매우 유용한 도구이다(Blei, 2012). 이에 연구의 목적을 달성하기 위해 먼저 수집된 자료에서 나타나는 키워드를 가지고 빈도분석을 통해 빈도수 상위 30개의 단어를 추출하였다. 이후 토픽모델링의 LDA 알고리즘을 활용하여 8개의 토픽을 산출하였으며, 토픽별 키워드는 10개로 정리하였다. 마지막으로 시간에 따라 연구 동향이 어떻게 변화하고 있는지 규명

표 1. 자료처리방법에 사용된 Python 패키지

분석방법	패키지
자료수집	BeautifulSoup, Reqeusts, Selenium
텍스트 전처리	KoNLPy
빈도분석	NLTK
토픽모델링	gensim

하기 위해 단순회귀분석을 실시하였다. 자료처리 및 분석을 위해 Python 3.8 프로그램과 SPSS 22.0을 사용하였으며, 모든 통계적 유의수준은 .05로 설정하였다. 자료처리방법에 사용된 Python 패키지는 다음 <표 1>에 제시하였다.

Ⅲ. 연구결과

1. 키워드 분석결과

한국여성체육학회지의 연구주제를 대상으로 키워드 분석을 하기 위해 텍스트 전처리 과정을 모두 마친 최종자료를 기반으로 빈도분석을 실시하였다. <표 2>는 빈도수 상위 30개의 단어와 그 빈도를 제시하였다. 분석결과, 총 9193개의 단어가 나타났으며 '운동'이라는 단어가 287회로 가장 많이 나타났다. 그리고 키워드 빈도를 순차적으로 나열하면 '여성' 207회, '스포츠' 161회, '참여' 141회, '노인' 138회, '무용' 104회, '만족' 95회, '여가' 82회, '프로그램'

표 2. 빈도수 상위 30개 단어

단어(n)	단어(n)	단어(n)
1 운동(287)	11 여가(82)	21 지도자(47)
2 여성(207)	12 사회(74)	22 유형(46)
3 스포츠(161)	13 효과(67)	23 변화(46)
4 참여(141)	14 프로그램(65)	24 몰입(46)
5 노인(138)	15 체력(61)	25 학생(45)
6 무용(104)	16 신체(57)	26 경험(41)
7 선수(96)	17 비만(54)	27 신체_활동(39)
8 만족(95)	18 행동(53)	28 특성(38)
9 체육(87)	19 요인(51)	29 건강(38)
10 활동(87)	20 대학생(49)	30 인식(37)
심리(36회),	댄스_스포츠(34회),	콜_밀도(24회),
청소년(23회),	보행(21회),	교양_체육(9회),
진로(9회) 등		



그림 4. 키워드의 워드 클라우드(word cloud)

65회, '지도자' 47회, '신체_활동' 39회, '건강' 38회 등으로 나타났다. 기타 추출된 키워드는 심리, 댄스_스포츠, 골 밀도, 청소년, 교양 체육 등으로 나타났다. 그림 4는 추출된 키워드(상위 250개)를 기초로 워드 클라우드(word cloud)로 도식화한 그림이다.

2. 토픽모델링 분석결과

한국여성체육학회지의 연구주제를 구체적으로 파악하기 위해 토픽모델링의 LDA 알고리즘을 적용해 분석한 결과, 총 8개의 토픽이 도출되었으며 토픽별 비중이 높은 주요 키워드 순으로 정리하여 <표 3>에 제시하였다. 토픽별 연구 주제명은 일반적으로 상위 키워드 1과 2를 중심으로 정하며 연구자의 식견에 따라 달라질 수 있다(오정심, 2020). 이에 본 연구에서는 상위 키워드와 관련 논문의 초록을 참고하여 전문가 회의를 통해 내용을 분석하여 토픽별 주요 연구주제를 정리하였다. 그리고 <표 4>는 개별 토픽과 관련된 주요 논문 목록을 제시한 것이며 문서와 토픽 간 행렬에 의한 확률값인 감마값이 높은 논문을 연도순으로 제시하였다.

분석결과를 살펴보면, Topic 1 '운동 경험' 이 18.8%로 가장 높았으며 Topic 2 '선수 대상 연구'

표 3. 토픽모델링 분석결과

Topic	퍼센트	주요 키워드				
Topic 1 (운동 경험)	18.8%	운동(.113)	여성(.085)	경험(.029)	학생(.024)	신체_활동(.023)
Topic 2 (선수대상 연구)	16.0%	비만(.022)	노인(.018)	성인(.018)	변화(.018)	복합_운동(.016)
Topic 3 (여가활동 참여)	14.4%	선수(.052)	프로그램(.044)	행동(.043)	효과(.040)	체력(.035)
Topic 4 (노인체육 및 건강)	12.7%	운동(.027)	심리(.025)	요인(.023)	체육교사(.020)	유형(.019)
Topic 5 (만족도 관련 연구)	11.6%	참여(.104)	활동(.044)	몰입(.040)	여가(.033)	대학생(.030)
Topic 6 (여성 스포츠 관련 연구)	9.6%	생활_체육(.027)	사회(.025)	스트레스(.024)	운동(.021)	이용(.018)
Topic 7 (무용 관련 연구)	8.9%	노인(.086)	체육(.053)	교육(.026)	문화(.026)	건강(.023)
Topic 8 (생리적 변인)	8.0%	청소년(.018)	적응(.018)	생활(.015)	인지(.014)	현황(.014)
		만족(.070)	지도자(.033)	의도(.031)	인식(.030)	태권도(.030)
		대학(.025)	적응(.024)	수업(.024)	기능(.022)	직무(.020)
		스포츠(.145)	여성(.028)	학교(.021)	개발(.020)	참여_동기(.016)
		댄스(.014)	보행(.014)	관점(.014)	이미지(.013)	자본(.012)
		무용(.046)	자기(.025)	진로(.023)	수준(.023)	선수(.022)
		중심(.020)	고등학교(.018)	학습(.018)	리더십(.018)	대학(.013)
		신체(.051)	불안(.031)	혈중(.024)	기술(.022)	능력(.020)
		여학생(.020)	골 밀도(.016)	섭취(.016)	성과(.016)	대사_증후군(.015)

표 4. Topic과 관련된 주요 논문 목록

Topic (논문수)	논문제목(게재연도)	Topic (논문수)	논문제목(게재연도)
Topic 1 (220)	직장체육 활동참여 실태조사 연구(1992)	Topic 5 (201)	생활체육 참가 주부의 가족 스트레스와 생활만족에 관한 연구(2000)
	대학생의 스포츠활동 참여실태에 관한 연구(1996)		체육수업 반 구성형태에 따른 여자 고등학생들의 체육수업 관심도 및 만족도 조사연구(2000)
	여성노인의 운동참여와 삶의 질에 관한 연구(2005)		중학생이 인식하는 체육교사의 변혁적 리더십과 체육수업만족에 관한 연구(2008)
	성인 여성의 아쿠아로빅 운동 참여정도와 자기관리가 운동몰입, 자아존중감에 미치는 영향(2009)		골프갤러리의 개인적 특성에 따른 만족도 분석(2008)
	운동참가여부에 따른 중년여성의 건강신념이 수면의 질과 행복지수에 미치는 영향(2013)		스포츠동아리 참여 대학생들의 라이프스타일이 자아탄력성과 생활만족에 미치는 영향(2015)
Topic 2 (109)	국가대표 복싱선수의 등속성 근력과 무산소성 능력간의 상관관계(2006)	Topic 6 (120)	여성의 스포츠 참여에 대한 사회적 배경(1987)
	고등학교 태권도 선수의 선수 만족이 자기실현에 미치는 영향(2006)		신문기사를 통해 본 여성 스포츠(2009)
	스포츠지도자의 슈퍼리더십이 선수들의 셀프리더십과 개인학습에 미치는 영향(2011)		라인댄스 참가 중년여성의 신체 이미지와 사회적 지지 및 삶의 기대와의 관계(2010)
	합의적 질적연구 방법을 적용한 고등학교 학생선수의 진로교육 개입 성과 탐색(2016)		여성의 평등권, 건강권, 행복권 추구를 위한 스포츠 서비스 환경개선 전략(2015)
	20대 은퇴선수를 위한 진로집단상담 프로그램 개발: 아틀러의 상담이론을 중심으로(2021)		여성, 스포츠 그리고 미디어(2016)
Topic 3 (158)	도시 주부의 여가 생활 실태(1993)	Topic 7 (92)	자유주의 페미니즘적 맥락에서 본 최승희의 삶과 무용의 재조명(1997)
	직장인의 여가인지에 관한 연구(1994)		강령탈춤의 기본 춤사위를 응용한 정신지체아의 무용치료프로그램에 영향을 미치는 요인에 관한 연구(1999)
	노인의 개인별 특성에 따른 여가만족도 분석(2008)		무용이 정적 자세 조절에 미치는 영향(2007)
	노인의 스포츠라이프 스타일이 여가 만족도에 미치는 영향(2009)		무용치료 사례연구(2008)
	여성의 여가제약 및 여가제약협상에 관한 연구동향 분석(2010)		한국무용이 무산소성 역치 수준에 미치는 영향(2008)
Topic 4 (245)	노인들의 신체활동 참여의 지속성과 재사회화에 관한 연구(2003)	Topic 8 (132)	여자 스포츠선수의 월경이상과 골밀도에 관한 연구(1993)
	노인들의 사회인구학적 요인, 건강체력, 일상생활 및 운동능력과의 관련성(2008)		여대생의 골밀도 분석 연구(2007)
	수중운동이 노년기 여성의 신체구성, 골밀도 및 건강체력에 미치는 영향(2009)		중년여성들에 있어 신체조성, 최대산소섭취량 및 근력과 골밀도와의 상관관계(2000)
	리듬운동 테라피프로그램이 경증치매 노인의 인지기능, 신체구성 및 건강관련 체력에 미치는 영향(2013)		대학 동계 스키 수업 중 넘어짐에 따른 뇌진탕 발생률에 관한 연구(2006)
	걷기 동아리 참여 노인들의 커뮤니케이션이 운동 관여도와 운동지속의도에 미치는 영향(2019)		10주간의 요가프로그램이 유방절제술 환자의 신체조성과 관절의 가동성에 미치는 효과(2006)

가 16.0%, Topic 3 '여가활동 참여'가 14.4%로 위 3개의 토픽이 비중 있게 다뤄지고 있는 연구영역으로 나타났다. 그 다음으로 노인 체육 및 건강

(12.7%), 만족도 관련 연구(11.6%), 여성 스포츠 관련 연구(9.6%), 무용 관련 연구(8.9%), 생리적 변인(8.0%) 순으로 나타났다.

토픽별 상위 10개의 주요 키워드 산출결과, Topic 1은 운동, 여성, 경험, 학생, 신체_활동, 비만, 노인, 성인, 변화, 복합_운동의 단어가 주요 키워드로 나타났다. Topic 2의 주요 키워드는 선수, 프로그램, 행동, 효과, 체력, 운동, 심리, 요인, 체육 교사, 유형으로 나타났다. Topic 3은 참여, 활동, 몰입, 여가, 대학생, 생활_체육, 사회, 스트레스, 운동, 이용이 주요 키워드로 나타났다. Topic 4는 노인, 체육, 교육, 문화, 건강, 청소년, 적응, 생활, 인지, 현황이 주요 키워드로 나타났으며, Topic 5는 만족, 지도자, 의도, 인식, 태권도, 대학, 적응, 수업, 기능, 직무가 주요 키워드로 나타났다. Topic 6의 주요 키워드는 스포츠, 여대생, 학교, 개발, 참여_동기, 댄스, 보행, 관점, 이미지, 자본으로 나타났으며, Topic 7은 무용, 자기, 진로, 수준, 선수, 중심, 고등학교, 학습, 리더십, 대학으로 나타났다. 마지막으로 Topic 8은 8.0%로 가장 작은 비중을 차지하며 주요 키워드는 신체, 불안, 혈중, 기술, 능력, 여학생, 골_밀도, 섭취, 성과, 대사_증후군으로 나타났다.

3. Hot 토픽과 Cold 토픽

한국여성체육학회지의 연구주제를 대상으로 시간이 지남에 따라 연구주제의 동향이 어떻게 변화하는지 규명하기 위해 토픽모델링을 통해 도출된 토픽들의 연도별 토픽점유율을 산출하였다. 이후 독립변수를 연도로, 종속변수는 연도별 토픽점유율을 기준으로 단순회귀분석을 실시하였다. 회귀모형에 대한 적합성은 잔차분석(residual analysis)을 통해 검토하였다(김진환, 김영준, 2021). 잔차의

독립성은 Durbin-Watson을 실시하였으며 Shapiro-wilk 정규성 검정을 실시하였다. 분석결과, Durbin-Watson 값은 1.874~1.975로 2 이하로 나타났으며 모두 정규성 가정이 성립되었다($p < .05$). 그리고 분석결과, 통계적으로 유의한 회귀계수 값이 양수(+)이면 Hot 토픽(상승 토픽),

음수(-)이면 Cold 토픽(하향 토픽), 그리고 통계적으로 유의하지 않으면 Neutral(중립) 토픽으로 해석한다. 이러한 기준을 기초로 회귀분석의 결과는 <표 5>에 제시하였으며, 분석결과 1개의 Hot 토픽과 1개의 Cold 토픽이 도출되었다. Topic 7 무용 관련 연구는 통계적으로 유의한 음의 회귀계수 값으로 Cold 토픽에 해당하여 연구주제의 비중이 감소하는 경향을 보였다. 반면에 Topic 8 생리적 변인은 통계적으로 유의한 양의 회귀계수 값으로 Hot 토픽에 해당하여 연구주제의 비중이 증가하는 경향을 보였다. 그리고 도출된 Hot 토픽과 Cold 토픽의 시기별 추이를 살펴보기 위해 총 5개의 시기(1: 1987년~1999년, 2: 2000년~2004년, 3: 2005년~2009년, 4: 2010년~2014년,

표 5. 단순회귀분석 결과

Topic	β	p	Shapiro- p	Durbin- Watson	Hot /Cold
Topic 1	.010	.726	.000	1.880	-
Topic 2	-.027	.338	.000	1.946	-
Topic 3	.014	.612	.000	1.944	-
Topic 4	.027	.330	.000	1.874	-
Topic 5	-.021	.443	.000	1.964	-
Topic 6	-.011	.686	.000	1.962	-
Topic 7	-.062	.026	.000	1.973	Cold
Topic 8	.056	.047	.000	1.975	Hot

Neutral Topic = -

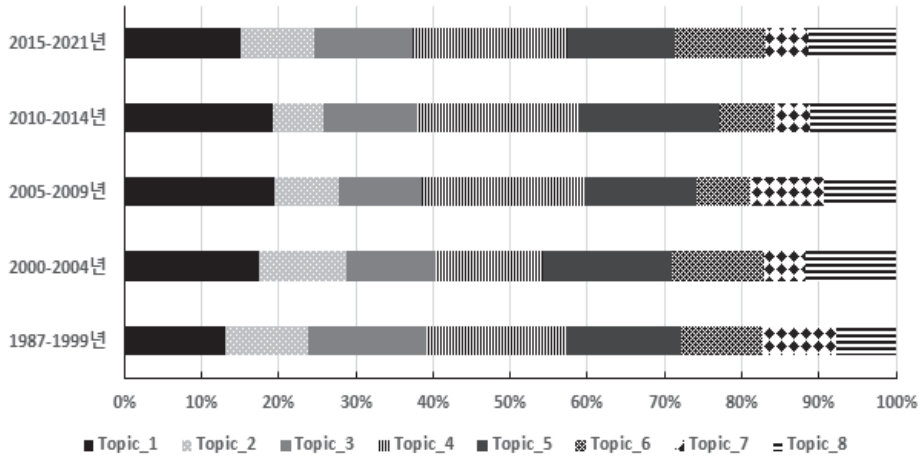


그림 5. 시기에 따른 토픽별 점유율

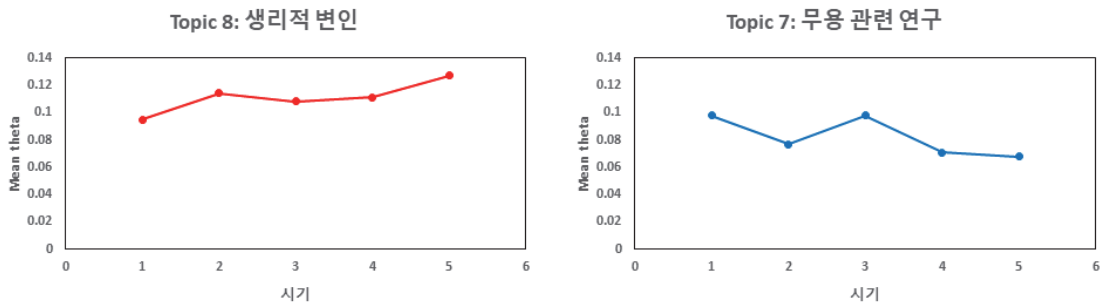


그림 6. Hot 토픽(왼쪽)과 Cold 토픽(오른쪽)의 변화 추이

5: 2015년~2021년)로 구분하여 시기별 토픽점유율을 산출하였으며 그 결과를 <그림 5>에 제시하였다. 그리고 통계적으로 유의미하게 나타난 Hot 토픽과 Cold 토픽의 그래프는 <그림 6>과 같다.

IV. 논의

많은 학자들은 특정 분야의 연구 동향을 파악하는 것은 과거의 연구 성과를 점검하고 앞으로의 연

구과제를 제시하는데 중요한 작업이라고 주장하였다(Lee, Wu, & Tsai, 2009; Lin, Lin, & Tsai, 2014). 지금까지 여성체육에 대한 연구동향 분석은 대부분 내용분석이 주를 이루고 있다(김수현, 2010; 김원정, 홍애령, 2017). 하지만 전통적으로 널리 사용되고 있는 내용분석은 연구방법론적 측면에서 연구주제, 연구대상 및 방법 등을 체계적으로 분석하여 전반적인 연구 동향을 파악할 수 있지만, 분석과정에서 연구자의 주관적 판단을 배제할 수 없어 외적 타당성을 확보하기 어렵다는 문제

가 있다(Rourke & Anderson, 2004).

따라서 본 연구에서는 연구 동향을 보다 객관적으로 분석하기 위해 토픽모델링(LDA 알고리즘)을 적용하여 한국여성체육학회지의 연구 동향을 분석하고자 하였다. 또한, 시간이 지남에 따라 연구주제의 동향이 어떻게 변화하는지 규명하고자 하였다. 분석결과, 8개의 토픽이 도출되었으며 시간에 따라 상승하는 연구주제로 1개의 Hot 토픽, 하향하는 연구주제로 1개의 Cold 토픽을 확인할 수 있었다. 이러한 결과를 토대로 토픽별로 다음과 같은 논의를 도출하였다.

Topic 1은 운동, 여성, 경험, 학생, 신체_활동, 비판, 노인, 성인, 변화, 복합_운동의 단어가 주요 키워드로 나타났으며 이러한 내용으로 보아 운동경험에 대한 연구주제로 확인할 수 있었다. Topic 1의 주요 키워드와 이를 포함하는 관련 대표 논문에 비추어볼 때, Topic 1은 성인 여성과 노인을 주 대상으로 운동참여 경험에 대한 실태조사를 파악하고 운동참여를 통한 긍정적인 변화(예: 운동몰입, 자아존중감, 건강신념, 삶의 질 등)를 살펴보는 연구의 흐름을 드러내 보인다고 할 수 있다. 이러한 연구주제의 흐름은 한국여성체육학회(2014)에서 내용분석을 통해 연구 동향을 분석한 결과, 운동에 따른 신체적 변화에 대한 연구가 전체 연구 중 최고의 빈도(91편, 10%)로 나타났다는 연구결과를 부분적으로 지지하는 결과로, 중요한 연구주제로 관련 연구가 지속되고 있는 것을 확인할 수 있었다.

Topic 2의 주요 키워드는 선수, 프로그램, 행동, 효과, 체력, 운동, 심리, 요인, 체육교사로 나타나 선수 대상 연구를 주로 다루고 있음을 확인할 수 있

었다. Topic 2의 주요 키워드와 이를 포함하는 관련 대표 논문에 비추어볼 때, 선수를 중심으로 체력과 심리요인에 대한 연구 및 선수들의 행동에 영향을 미치는 지도자 영향요인 등 폭넓게 연구가 이루어지고 있었다. 이러한 결과는 김지훈, 원영신, 김수현(2016)의 연구에서 한국여성체육학회지에 게재된 여성 체육 관련 연구의 핵심어 연결망 분석을 한 결과, 선수는 연결중심성이 높은 상위 핵심어로 나타나 이를 뒷받침해주고 있다. 또한, Topic 2에 포함하는 대표 논문들을 내용분석한 결과, 2016년부터 현재까지 총 26편의 논문 중 선수의 진로 탐색을 위한 연구는 3편(8.7%) 보고되었지만 이러한 결과는 토픽모델링 기법만을 통해서 확인하기 어려운 결과였다. 본 연구는 게재된 논문의 제목만을 키워드로 하여 토픽모델링 분석을 실시하였기 때문에 토픽에 해당하는 주요 키워드를 산출하는데 있어 한계가 있는 것으로 사료된다. 따라서 추후에는 토픽모델링과 함께 기존의 내용분석을 활용하여 심층적인 연구 동향 분석이 진행될 필요가 있다.

Topic 3은 다양한 여가활동 참여에 대한 내용이 주요한 연구주제로 포착되었다. 주요 키워드로는 참여, 활동, 여가, 생활_체육 등이 상위 키워드로 나타났다. Topic 3의 주요 키워드와 이를 포함하는 관련 대표 논문에 비추어볼 때, 여가 인식 및 여가 참여에 따른 다양한 측면에서의 긍정적 효과를 다루는 연구 흐름을 파악할 수 있었다. 그리고 시기별 내용분석을 통해 2000년 전후로 전반기에는 여성의 여가 생활 실태 및 여가 참여요인 탐색 연구(김준희, 1996; 박계순, 1993)가 주를 이루었으며 이후 여가만족에 영향을 미치는 연구 탐색 연구(김영

숙, 2003; 박장근, 장경호, 임란희, 2009)의 흐름이 나타나고 있음을 확인할 수 있었다. 지금까지 토픽모델링을 적용해 한국여성체육학회지의 연구 동향을 분석한 기존 선행연구가 없어 본 연구결과를 직접적으로 비교하기는 어렵지만 최창환과 김혜련(2019)이 네트워크 기반 키워드 분석을 적용하여 여성 신체활동 연구의 지식구조를 분석한 결과, 여가 및 사회 연구영역의 지식구조를 형성하고 있다는 연구결과와 유사한 결과였다. 또한, 최창환과 김혜련(2019)의 연구에서 여성체육학 연구와 관련하여 다수의 학술지 중 한국여성체육학회가 출판현황에서 상위 학술지로 나타났다. 따라서 본 연구에서 여성의 여가활동 및 참여에 대한 내용을 주요한 연구 주제로 다루고 있음을 알 수 있었다.

Topic 4는 노인, 체육, 교육, 문화, 건강, 청소년, 적응, 생활, 인지, 현황이 주요 키워드로 나타났다. Topic 4의 주요 키워드와 관련 대표 논문에 비추어볼 때, Topic 4는 노인을 주요 대상으로 여러 측면에서의 건강에 관한 연구 주제로 가늠해 볼 수 있었다. 이러한 결과는 2000년 이후 고령화 시대에 대한 이슈가 부각되면서 노인 대상 건강 관련 연구가 증가하고 있다는 연구결과(김지훈, 원영신, 김수현, 2016; 한국여성체육학회, 2014)를 지지하는 것이다. 이러한 연구 흐름에 편승하여 이원미와 권구명(2019)은 토픽모델링을 적용하여 노인 체육만을 대상으로 연구 동향을 분석한 결과, 노인과 관련하여 여가와 건강이 주요 토픽으로 도출되어 유사한 결과를 보였다. 게다가 Topic 4를 대표하는 관련 논문을 살펴봄으로써, 운동이 노인의 인지적 기능에 미치는 영향에 대한 연구(김대권, 양

은혜, 2013; 이청무, 이난희, 2012) 등 2010년 이후부터 인지기능에 대한 연구가 새롭게 나타남을 알 수 있었다. 이와 같이 초고령화 시대로 가는 시점에서 노인체육 및 건강에 대한 연구주제는 지속적인 관심 영역이 될 것으로 사료된다.

Topic 5는 주요 키워드와 관련 대표 논문에 비추어 만족도와 관련된 연구에 대한 내용을 주요 연구 주제로 다루고 있었다. 주요 키워드로는 만족, 지도자, 의도, 대학, 적응, 수업, 기능, 직무 순으로 나타났다. 구체적으로 Topic 5와 관련한 연구를 살펴보면, 체육수업 만족(김경숙, 김나정, 2000; 정성우, 표내숙, 김명철, 2008)과 생활 만족(김준희, 박선영, 2000; 김효림, 이미연, 전익기, 2015) 등에 대한 연구가 꾸준히 이루어 있어 여러 측면에서의 만족도에 관한 연구주제를 파악할 수 있었다.

Topic 6의 주요 상위 키워드는 스포츠, 여성, 학교, 참여_동기, 댄스 등의 단어가 주요 키워드로 나타났다. 주요 키워드와 이를 포함하는 관련 대표 논문에 비추어볼 때, 젠더적 관점에서 여성과 스포츠에 대한 연구주제를 확인할 수 있었다. 게다가 관련 연구로 라인댄스 참가 중년여성의 신체 이미지와 사회적 지지 및 삶의 기대와의 관계 연구(문행자, 2010)와 같이 남성보다 여성에게 더 중요시되는 신체적 이미지 관련 연구로도 범위가 확대되고 있음을 알 수 있었다.

Topic 7의 주요 키워드는 무용, 자기, 진로, 수준, 선수, 고등학교, 리더십 순으로 나타났다. 주요 키워드와 관련 대표 논문에 비추어볼 때, Topic 7의 주요 연구주제는 무용과 관련된 연구로 가늠할 수 있었다. 특히 시간에 따라 연구주제의 상승과

하락을 살펴보기 위해 회귀계수 값을 이용하여 분석한 결과, Topic 7은 통계적으로 유의한 음의 회귀계수 값으로 Cold 토픽에 해당하여 연구주제의 비중이 감소하는 경향을 보이고 있었다. 이러한 결과를 내용분석을 통해 연구 동향을 분석한 결과(한국여성체육학회, 2014)와 분석방법의 차이로 직접적으로 비교하기는 제한적이지만 무용과 관련한 연구가 2010년 이후 다양한 스포츠 종목으로 확대되어 나타났다는 연구결과를 부분적으로 지지하는 결과였다. 본 연구에서도 2010년 이후로 다양한 융합 연구의 등장으로 무용과 관련된 연구는 감소하는 것으로 보이며 또한, 무용 관련 연구들이 더 전문화된 무용 관련 학회지(예: 한국무용교육학회지, 한국무용학회지, 한국무용과학회지 등)로 투고되면서 감소하였을 가능성도 있을 것이다. 이에 추후 연구에서는 관련 학회지의 연구동향 분석결과를 비교하는 등의 더욱 객관적 지표를 가지고 이러한 현상의 원인을 다각도로 분석할 필요가 있다.

그리고 김지훈, 원영신, 김수현(2016)의 연구에서 한국여성체육학회지를 대상으로 여성체육 관련 핵심어 연결망 분석을 살펴본 결과, 무용 관련 연구 동향을 단순히 2개(1987~1999, 2000~2015)의 시기로 구분하여 주요 키워드로만 결과를 해석하고 있었다. 이에 본 연구에서는 처음으로 현재까지 한국여성체육학회지에 투고된 논문을 대상으로 토픽모델링을 활용하여 여성 체육 관련 연구주제를 분류하였고 시간의 흐름에 따른 연구 동향을 통계적 기법을 활용하여 회귀계수 값을 이용하여 살펴봄으로써, 시간에 따른 연구주제의 변화 추이를 과학적으로 입증할 수 있었음에 연구의 의미가 있다.

Topic 8의 주요 키워드는 신체, 불안, 혈중, 여학생, 골_밀도 등으로 나타났다. 주요 키워드와 관련 대표 논문에 비추어볼 때, Topic 8은 광범위하게 생리적 변인과 관련한 내용을 주요 연구주체로 다루고 있었다. 또한, 생리적 변인과 관련한 주제는 통계적으로 유의한 양의 회귀계수 값으로 Hot 토픽에 해당하여 연구주제의 비중이 증가하는 경향을 보였다. 그리고 Topic 8과 관련된 대표 논문들을 살펴보면, 생리적 변인과 관련하여 신체 조성, 골밀도, 뇌진탕 등 매우 광범위하게 측정되고 있음을 알 수 있었다. 한국은 학생들의 비만과 체력 저하를 방지하기 위해 2007년 학생건강체력평가 시스템(PAPS)을 개발(교육인적자원부, 2007)하였으며 이 시기를 기점으로 국민들의 건강한 삶을 위한 건강 관련 체력의 측정 및 평가가 활성화되었다. 즉 이러한 시대적 흐름에 편승하여 여성 체육학 분야의 연구도 체력과 관련된 생리적 요인을 측정하는 자연과학 분야의 비중이 증가하는 추세를 보이는 것으로 사료된다. 따라서 본 연구결과를 토대로 앞으로 여성 체육 관련 연구의 폭넓은 융합 연구가 활발하게 진행될 것으로 예측할 수 있다.

V. 결론 및 제언

본 연구는 한국여성체육학회지를 대상으로 토픽모델링을 적용하여 토픽별 연구주제 및 시간에 따른 연구 동향을 규명하였다. 연구결과를 요약하면 다음과 같다. 첫째, 키워드 빈도분석을 통해 빈번히 출현하는 단어와 그 빈도를 살펴볼 수 있었다. 둘째, 토픽

모델링을 활용하여 8개의 토픽을 도출할 수 있었다. 셋째, 연도별 토픽점유율을 기준으로 단순회귀분석을 실시한 결과, Topic 7 '무용 관련 연구'는 Cold 토픽, Topic 8 '생리적 변화'는 Hot 토픽으로 나타나 시간에 따른 연구주제의 동향을 확인할 수 있었다. 본 연구의 결과는 여성 체육학의 학문의 정체성 수립과 앞으로의 여성 체육학이 나아가야 할 학문적 발전에 중요한 기초자료로 활용될 수 있을 것이다.

본 연구의 결과를 종합해보면, 한국여성체육학회를 대상으로 토픽모델링을 적용하여 8개의 토픽별 연구주제 및 시간에 따른 연구 동향을 파악할 수 있었다. 게다가 내용분석의 한계를 인식하고 비정량적 데이터를 정량화시키는 텍스트 기반의 연구 동향 분석법을 새롭게 소개하고 적용하였다는 점에서 중

요한 의의가 있다. 다만 본 연구의 결과는 시대에 따른 논문 제출형태의 차이로 인해 논문 제목 자료만을 활용하여 연구주제를 산출하였다는 점, 토픽모델링의 특성상 소수의 키워드로 형성된 연구주제를 파악하기 어렵다는 점에서 한계가 있다. 물론 본 연구에서는 이러한 한계점을 인식하고 토픽모델링을 통해 도출된 토픽들의 연구주제는 질적 분석에 기초한 전문가 회의를 통해 관련 논문들을 토대로 연구주제를 정리하려 노력하였다. 따라서 추후 연구는 토픽모델링과 함께 기존의 내용분석을 활용하여 심층적인 연구 동향 분석이 진행될 필요가 있을 것이다. 또한, 연구목적, 연구방법, 연구결과 및 논의의 내용을 포함한 초록의 내용을 기초로 한 토픽모델링을 적용하여 살펴볼 필요가 있을 것이다.

참고문헌

- 고대선, 강익원, 이금룡(2008). 노인들의 사회인구학적 요인, 건강체력, 일상생활 및 운동능력과 관련된 성. **한국여성체육학회지**, 22(2), 215-228.
- 고재욱, 김창우(2006). 대학 동계 스키 수업 중 넘어짐에 따른 뇌진탕 발생률에 관한 연구. **한국여성체육학회지**, 20(5), 147-156.
- 고정화(1992). 직장체육활동참여실태조사연구: 인천지역공단을 중심으로. **한국여성체육학회지**, 6, 72-81.
- 교육인적자원부(2007b). **학생건강체력평가시스템 구축**. 교육인적자원부.
- 김경숙, 김나정(2000). 체육수업 반 구성형태에 따른 여자고등학생들의 체육수업 관심도 및 만족도 조사연구. **한국여성체육학회지**, 14(2), 27-37.
- 김동현, 장덕선(2021). 20 대 은퇴선수를 위한 진로집단상담 프로그램 개발: 아들러의 상담이론을 중심으로. **한국여성체육학회지**, 35(1), 69-84.
- 김대권, 양은혜(2013). 리듬운동 테라피프로그램이 경증치매 노인의 인지기능, 신체구성 및 건강 관련 체력에 미치는 영향. **한국여성체육학회지**, 27(4), 115-127.
- 김명주(1994). 직장인의 여가인지에 관한 연구. **한국여성체육학회지**, 8, 39-48.
- 김미숙, 양점홍(2006). 10 주간의 요가프로그램이 유방절제술 환자의 신체조성과 견관절의 가동성에 미치는 효과. **한국여성체육학회지**, 20(3), 13-24.
- 김병훈, 김소연(2019). 걷기 동아리 참여 노인들의 커뮤니케이션이 운동 관여도와 운동지속의도에 미치는 영향. **한국여성체육학회지**, 33(2), 109-122.
- 김수현(2010). 여성의 여가제약 및 여가제약협상에 관한 연구동향 분석. **한국여성체육학회지**, 24(2), 167-178.
- 김영숙(2003). 시니어 에어로빅스 참가 노인의 여가만족에 관한 연구. **한국여성체육학회지**, 17(1), 33-46.
- 김원정, 홍애령(2017). 국내 여학생 신체활동의 연구동향 분석을 통한 체육 활성화 방안 탐색. **한국여성체육학회지**, 31(3), 73-93.
- 김인형, 이근모(2005). 스포츠사회학: 여성 엘리트 스포츠의 사회학적 연구동향 분석. **한국체육학회지**, 44(6), 213-223.
- 김준희(1996). 여가사회화 참여에 기여하는 요인. **한국여성체육학회지**, 10, 83-94.
- 김준희, 박선영(2000). 생활체육 참가 주부의 가족 스트레스와 생활만족에 관한 연구. **한국여성체육학회지**, 14(2), 39-53.
- 김진환, 김영준(2021). 토픽모델링을 이용한 영업 분야 연구 동향 분석. **마케팅연구**, 36(1), 99-120.
- 김창식, 최수정, 곽기영(2017). 토픽모델링과 시계열회귀분석을 활용한 정보시스템분야 연구 동향 분석. **한국디지털콘텐츠학회 논문지**,

- 18(6), 1143-1150.
- 김효림, 이미연, 전익기(2015). 스포츠동아리 참여 대학생들의 라이프스타일이 자아탄력성과 생활만족에 미치는 영향. **한국여성체육학회지**, 29(4), 223-237.
- 남인수(2009). 성인여성의 아쿠아로빅 운동 참여정도 와 자기관리가 운동몰입, 자아존중감에 미치는 영향. **한국여성체육학회지**, 23(2), 61-74.
- 문행자(2010). 라인댄스 참가 중년여성의 신체 이미지와 사회적 지지 및 삶의 기대와의 관계. **한국여성체육학회지**, 24(2), 155-165.
- 민경은, 조미혜(2008). 무용치료 사례연구: 발달장애 아동들을 중심으로. **한국여성체육학회지**, 22(3), 59-70.
- 박계순(1993). 도시 주부의 여가 생활 실태: 울산시 주부를 대상으로. **한국여성체육학회지**, 7, 22-47.
- 박소영, 이근모(2007). 인문, 사회과학편: 한국스포츠사회학회지 연구동향 분석: 1993-2005. **한국체육학회지**, 46(1), 259-268.
- 박장근, 김진영, 차선동(2008). 골프갤러리의 개인적 특성에 따른 만족도 분석. **한국여성체육학회지**, 22(5), 129-141.
- 박장근, 장경호, 임란희(2009). 노인의 스포츠라이프 스타일이 여가 만족도에 미치는 영향. **한국여성체육학회지**, 23(1), 85-95.
- 박장근, 차선동, 임란희(2008). 노인의 개인별 특성에 따른 여가만족도 분석. **한국여성체육학회지**, 22(3), 95-104.
- 변기용, 이석열, 김수홍(2010). 한국 고등교육 정책 연구의 동향 분석: 연구주제 및 연구자를 중심으로. **아시아교육연구**, 11, 121-145.
- 신승윤(2015). 체육측정평가의 연구 동향 (1998~2014). **체육과학연구**, 26(2), 205-218.
- 오정심(2020). 빅데이터 토픽모델링 및 네트워크 분석을 통한 문화콘텐츠학 지식구조 연구. **문화정책논총**, 34(2), 35-70.
- 유미희(1997). 자유주의 페미니즘적 맥락에서 본 최승희의 삶과 무용의 재조명. **한국여성체육학회지**, 11, 115-128.
- 유상건(2016). 여성, 스포츠 그리고 미디어. **한국여성체육학회지**, 30(4), 39-53.
- 유정애(2014). 한국여성체육학회의 학문적 정체성과 미래 과제 탐색. **한국여성체육학회지**, 28(4), 77-90.
- 윤지운(2017). 지식네트워크 분석을 활용한 특수체육의 연구동향 분석. **한국특수체육학회지**, 25(1), 121-133.
- 이경옥(2007). 여대생의 골밀도 분석. **한국여성체육학회지**, 21(5), 1-12.
- 이경옥, 하지연(2000). 특수체육의 연구동향에 관한 분석. **한국특수체육학회지**, 8(2), 17-34.
- 이용국, 박재현(2017). 토픽모델링을 적용한 한국체육측정평가학회지 연구동향 분석. **한국체육측정평가학회지**, 19(2), 11-22.
- 이원미, 권구명(2019). 토픽모델링을 적용한 노인체육의 연구주제: 국내 인문사회과학 분야를 중심으로. **한국체육학회지**, 58(6), 253-262.
- 이청무, 이난희(2012). 복합운동이 여성노인의 신경영양성인자 및 인지기능에 미치는 영향. **한국여성체육학회지**, 26(1), 173-189.

- 이천희(2006). 고등학교 태권도 선수의 선수만족이 자기실현에 미치는 영향. **한국여성체육학회지**, 20(5), 29-40.
- 이청무(1993). 여자 스포츠선수의 월경이상과 골밀도에 관한 연구. **한국여성체육학회지**, 7, 105-112.
- 이현정, 광정현(2009). 신문기사를 통해 본 여성스포츠: 1991년~2008년까지 신문기사를 중심으로. **한국여성체육학회지**, 23(4), 159-172.
- 임변장(1987). 여성의 스포츠 참여에 대한 사회적 배경. **한국여성체육학회지**, 1, 23-26.
- 임호남(2009). 수중운동이 노년기여성의 신체구성, 골밀도 및 건강체력에 미치는 영향. **한국여성체육학회지**, 23(3), 39-48.
- 장익영, 이원미(2019). 토픽모델링을 적용한 한국스포츠사회학회지의 연구주제 분석. **한국스포츠사회학회지**, 32(2), 20-35.
- 정미라(2007). 무용이 정적 자세 조절에 미치는 영향. **한국여성체육학회지**, 21(3), 65-74.
- 정경환, 원영신, 구승광(2015). 특수체육 연구동향 분석: 2011-2014. **한국특수체육학회지**, 23(1), 79-90.
- 정성우, 표내숙, 김명철(2008). 중학생이 인식하는 체육교사의 변혁적 리더십과 체육수업만족에 관한 연구. **한국여성체육학회지**, 22(2), 77-94.
- 정수진, 손태열(2008). 한국무용이 무산소성 역치 수준에 미치는 영향. **한국여성체육학회지**, 22(1), 113-122.
- 정효정(2016). 텍스트 마이닝을 이용한 혁신 분야의 국외 연구 동향 분석. **기술혁신연구**, 24(4), 249-275.
- 조성립, 조현익, 소영호(2013). 운동참가여부에 따른 중년여성의 건강신념이 수면의 질과 행복 지수에 미치는 영향. **한국여성체육학회지**, 27(1), 45-61.
- 조은혜(2003). 노인들의 신체활동 참여의 지속성과 재사회화에 관한 연구. **한국여성체육학회지**, 17(1), 99-114.
- 진행미(1996). 대학생의 스포츠활동 참여실태에 관한 연구. **한국여성체육학회지**, 10, 77-82.
- 최미자(1999). 강령달춤의 기본춤사위를 응용한 정신지체아의 무용치료프로그램에 영향을 미치는 요인에 관한 연구. **한국여성체육학회지**, 13(1), 155-168.
- 최성애(2011). 스포츠지도자의 슈퍼리더십이 선수들의 셀프리더십과 개인학습에 미치는 영향. **한국여성체육학회지**, 25(1), 157-169.
- 최성철, 박한우(2020). 토픽모델링 연구동향 분석: 공학과 사회과학 분야 Kci 등재지를 중심으로. **한국자료분석학회지**, 22(2), 815-826.
- 한시완(2016). 합의적 질적연구 방법을 적용한 고등학교 학생선수의 진로교육 개입 성과 탐색. **한국여성체육학회지**, 30(2), 17-36.
- 허현미(2015). 여성의 평등권, 건강권, 행복권 추구를 위한 스포츠 서비스 환경개선 전략. **한국여성체육학회지**, 29(3), 249-272.
- 홍양자, 이승민, 조영희, 강승애(2005). 여성노인의 운동참여와 삶의 질에 관한 연구. **한국여성체육학회지**, 19(3), 97-105.
- 한국여성체육학회(2014). **한국여성체육학회 60주년 기념집**. 서울: 효일문화사.
- Berry, T. R., Strachan, S. M., & Verkooijen, K. T. (2014). The relationship between exercise schema and identity. *International Journal*

- of Sport and Exercise Psychology, 12, 49-63.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.
- Blei, D. M., & Lafferty, J. D. (2006). Dynamic topic models. *Paper presented at the Proceedings of the 23rd International Conference on Machine Learning*, 113-120.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3, 993-1022.
- Danowski, J. A. (1993). Network analysis of message content. *Progress in communication sciences*, 12, 198-221.
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391-407.
- Hofmann, T. (1999). Probabilistic latent semantic indexing. *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 50-57.
- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences of the United States of America*, 101 Suppl 1, 5228-5235.
- Lee, M., Wu, Y., & Tsai, C. (2009). Research trends in science education from 2003 to 2007: A content analysis of publications in selected journals. *International Journal of Science Education*, 31(15), 1999-2020.
- Lin, T., Lin, T., & Tsai, C. (2014). Research trends in science education from 2008 to 2012: A systematic content analysis of publications in selected journals. *International Journal of Science Education*, 36(8), 1346-1372.
- Newman, D., Lau, J. H., Grieser, K., & Baldwin, T. (2010, June). Automatic evaluation of topic coherence. *In Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 100-108). Association for Computational Linguistics.
- Prabhakaran, Selva (2018). Topic Modeling with Gensim (Python). Available: <https://www.machinelearningplus.com/nlp/topic-modeling-gensim-python/>.
- Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development*, 52(1), 5.

Analysis of Research Trends in The Journal of Korean Association of Physical Education and Sport for Girls and Women Using Topic Modeling

Miyoung Roh(Kookmin University, Assistant Professor)

ABSTRACT

The purpose of the study was to analyze research trends of the journal of Korean Association of physical education and sport for girls and women using topic modeling and to identify how research trends change over time. For the purpose of this study, a total of 1,277 articles were collected from the first issue vol. 1(1) in 1987 to vol. 35(4) in 2021 of the journal of Korean Association of physical education and sport for girls and women. To analyze by topic modeling, collected data were subjected to text preprocessing, such as extracting noun words through tokenization, removing unnecessary words using Stopwords, and converting words with the same meaning into a single common word. Then, topics were extracted using LDA(Latent Dirichlet Allocation) algorithm of topic modeling and keyword frequency analysis, correlation analysis using Python 3.8 program and SPSS 22.0. The results are as follow: First, frequently used words were exercise(287 times), women(207 times), sports(161 times), participation(141 times), elderly(138 times), dance(104 times), athletes(96 times), leisure(82 times) and health (38 times). Second, 8 topics extracted from the data include Topic 1 'Exercise experience'(18.8%), Topic 2 'Study on athletes'(16.0%), Topic 3 'Participation in leisure activities'(14.4%), Topic 4 'Physical education and health for the elderly'(12.7%), Topic 5 'Study on satisfaction'(11.6%), Topic 6 'Women and Sport'(9.6%), Topic 7 'Study on dance'(8.9%), Topic 8 'Study on physiological variables'(8.0%). Third, we found that Topic 7 'dance-related research' was a Cold Topic and Topic 8 'physiological variables related-research' was a Hot Topic. These results can be used as important basic data for the academic development of women's sports studies.

Key words : topic modeling, LDL algorithm, korean association of physical education sport for girls and women, research trend

논문 접수일 : 2022. 2. 16

논문 승인일 : 2022. 3. 18

논문 게재일 : 2022. 3. 31