

<http://dx.doi.org/10.17703/JCCT.2023.9.6.309>

JCCT 2023-11-38

텍스트 마이닝(Text mining) 기법을 활용한 『제1차 조선교육령』과 『조선교육요람』(1913, 1920)의 영어 번역본 분석

Analysis on the English Translation of *The First Chosen Educational Ordinance, Manual of Education of Koreans (1913), and Manual of Education in Chosen 1920 (1920) Using Text Mining Analytics*

탁진영*, 곽은주**, 진실로***손민주****, 김동미*****

Jinyoung Tak*, Eunjoo Kwak, Silo Chin***, Minjoo Shon****, Dongmie Kim*******

요약 본 논문의 목표는 1911년에 공포된 『제1차 조선교육령』, 1913년에 발행된 『조선교육요람』과 1920년에 발행된 『조선교육요람(1920)』의 영어 텍스트를 비교 분석하고, 이를 통해 조선총독부의 조선 통제 수단으로 활용된 식민지 교육정책을 추적하는 것이다. 본 논문은 조선총독부의 식민지 교육정책의 전체적인 흐름과 그 배경을 파악하기 위해 기존의 단편적이고 질적인 역사서 연구와는 달리 통시적·양적 연구를 시도하였다. 이를 위해 상위 50개 단어 빈도 순위와 워드 클라우드(Word Cloud)와 CONCOR(CONvergence of iteration CORrelation)의 텍스트 마이닝 기법을 활용하여 다음과 같은 결과를 도출하였다. 첫째, 1911년 『제1차 조선교육령』은 ‘전체적인 규정’이나 ‘교육과정 및 운영’에 초점을 두었고 둘째, 1913년의 『조선교육요람』은 ‘교육 방식 및 매체’와 전반적인 ‘교육 내용’에 대한 지침을 싣고 있었다. 그리고 마지막으로 1920년 『조선교육요람(1920)』은 교육의 구체적인 ‘교육 실행 및 교육의 주체’에 관한 내용을 담고 있었다.

주요어 : 조선교육요람, 텍스트 마이닝, 조선교육령, 조선총독부, 워드 클라우드

Abstract The purpose of this paper is to investigate how Japan tried to dominate Chosen through educational policies by analyzing three official English texts published by the Japanese Government-General of Korea: *the First Chosen Educational Ordinance* declared in 1911, *the Manual of Education of Koreans*(1913), and *the Manual of Education in Chosen 1920*(1920). In order to pursue this purpose, the present study carried a corpus-based diachronic analysis, rather than a qualitative analysis. Facilitating text analytics such as Word Cloud and CONCOR, this paper derived the following results: First, *the first Chosen Educational Ordinance*(1911) includes overall educational regulations, curriculum, and operations of schools. Second, *the Manual of Education of Koreans*(1913) contains the educational medium and contents on how to educate. Finally, it can be proposed that *the Manual of Education in Chosen 1920*(1920) contains specific implementation of education and the subject of education.

Key words : Manual of Education of Koreans, Manual of Education in Chosen 1920, Text Analytics, the First Chosen Educational Ordinance, Japanese Government-General of Korea, Word Cloud.

*, 세종대학교 영어영문학과 직위 교수(주저자)

** , 세종대학교 영어영문학과 직위 교수(공동저자)

***, 세종대학교 영어영문학과 직위 초빙교수(공동저자)

**** 세종대학교 영어영문학과 직위 석사생 연구원(공동저자)

***** 대신대학교 상담영어학부 직위 조교수(교신저자)

접수일: 2023년 9월 15일, 수정완료일: 2023년 10월 10일

게재확정일: 2023년 11월 1일

Received: September 15, 2023 / Revised: October 10, 2023

Accepted: November 1, 2023

*****Corresponding Author: belle-kim75@hanmail.net

Dept. English language & Literature, Daeshin University
Korea

I. 서론

일본은 일제강점기 동안 조선의 역사 의식과 민족의식 말살을 위해 조선인에게 창씨개명[1], 신사참배, 그리고 황국신민서사시 암송 등의 정책을 실시하였다. 이러한 일본의 조선에 대한 식민지화는 문서상으로 ‘치밀’하고 ‘교활’하게 진행되었다. 즉, 조선에 대한 식민지 정당화를 대외적으로 공식화하기 위해 한국어와 일본어로 자료를 제작하였을 뿐만 아니라 서구 열방의 언어인 영어로 문서를 제작하고 유통시켰다[2]. 이는 다양한 영문 자료를 통해 국제여론을 호의적으로 조성하고 식민통치 역량을 과시하려는 일본의 의도적인 전략에 기인한다. 가령, 일본어판 『시정연보』(1906-1942)는 조선에 대한 행정자료와 일본 본토에 제출한 제반 보고서의 기초자료뿐만 아니라 국내외 선전을 위한 홍보자료로도 활용되었다. 일본은 영문 『시정연보』(1907-1939)(*Annual Report on Reforms and Progress in Chosen*)를 1년에서 2년, 3년 단위로 출간하였으며 한국 체류 서양 선교사들과 외교관들 및 영미권 거주 학자들, 해외통신원들 및 유력인사들과 대령도서관에 무료 배포하였다[3]. 따라서 『시정연보』 영문판은 일본의 ‘성공적’ 식민통치에 호의적 국제여론 조성을 위한 중요 자료로 활용되었다.

한편, 일본은 앞서 언급한 식민지 정책뿐만 아니라 한 차원 더 면밀하게 식민지화를 강행하였다. 예를 들어, 일본은 조선의 민족의식과 정체성 절단을 위한 방편으로 ‘교육정책’을 수정 시행하여 점진적으로 조선을 우민화(愚民化)[4]하고 동화(同化)[5]되게 하여 일본에 존속 시키고자 하였다. 일본은 이러한 교육정책을 시행하기 위해 1890년에 ‘교육에 관한 칙어’를 기반으로 1911년, 1922년, 1938, 1943년에 『조선교육령』을 4차례 공포하였다. 또한, 『조선교육령』의 포고를 위해 교육령 세부 지침을 담은 『조선교육요람』을 발행하였고, 『조선교육요람』 일본어판을 1913년, 1915년, 1919년, 1929년에, 영문판을 1913년과 1920년에 발행하였다[6].

본 논문의 목적은 텍스트마이닝 기법을 활용하여 『제1차 조선교육령』과 1913년의 『조선교육요람』 그리고 1920년의 『조선교육요람(1920)』의 영문판을 비교 분석하고 이를 통해 일제강점기 조선총독부의 조선 식민지 지배 전략을 추적하는 것이다.

II. 선행연구 검토

1. 『조선교육령』과 『조선교육요람』

『조선교육령』은 일제강점기 동안 일본이 조선의 식민지 교육정책을 어떻게 시행하였는지 관찰할 수 있는 중요한 자료이다. 일본은 1911년, 1922년, 1938년에 이어 1943년까지 총 4차례에 걸쳐 조선의 교육 및 교과 과정을 개편하였고, 조선에 철저히 식민지화하기 위한 수단으로 활용하였다.

총 4차례에 걸친 『조선교육령』의 시기와 내용을 살펴보면 다음과 같다. 1911년에 공포된 『제1차 조선교육령』에는 조선의 교육을 수준이 낮은 저급한 교육으로써 실업·기술교육 등으로 대체하는 정책들이 포함되어 있다. 또한, 일본어 수업을 늘리고, 조선어 수업 시간을 줄였으며, 보통교육에 해당하는 수업연한도 4년으로 단축시켰다[7]. 『제1차 조선교육령』의 공포된 영향권으로 1913년, 1915년, 1919년의 『조선교육요람』이 제작되었다. 본 논고에서 논하고자 하는 『조선교육요람』 영어 번역본은 각 1913년과 1920년에 제작되었다.

1919년 3·1운동의 직접적인 영향으로 고안된 『제2차 조선교육령』(1922)의 주요 내용은 문화식민지 정책의 가시적인 포고였다. 그 유희정책의 일례로 지난 『제1차 조선교육령』의 공포 이후 4년으로 단축된 수업연한이 일본과 동일하게 6년으로 변경되었다. 하지만, 입학 조건 등에서는 한국인을 차별하는 즉 여전히 불평등한 동화(同化)정책이 작용하였다.

『제3차 조선교육령』은 1938년에 공포되었으며, 중일전쟁 등 앞으로의 국제 정세에 대응하기 위한 교육정책을 담고 있다. 또한, 학교의 명칭에도 변화가 있었는데, 보통학교를 심상소학교로 고등보통학교를 중학교로, 여자고등보통학교를 고등여학교로 변경하였다.

마지막으로 1943년에 공포된 『제4차 조선교육령』의 특징은 군부에 의한 철저한 교육 통제가 이루어졌다는 것이다. 가령, 전시체제를 위한 학생동원 강화정책을 세웠으며, 이에 따라 교과서도 침략 정책에 어떻게 하면 대응할 수 있는지에 맞도록 편찬되었다.

위에 기술된 바와 같이 현재까지 『조선교육령』은 다수 연구자에 의해 수행되었으나 이와 비교하여 『조선교육요람』의 연구는 양과 질적인 면에서의 거의 전무하다. 최근 들어 몇몇 학자들이 『조선

교육요람』의 연구 가치의 중요성에 이목을 집중하고 관련 연구를 진행하고 있다. 그 사례로 진실로(2022)는 영문판 『조선교육요람』(1913), 『조선교육요람(1920)』(1920)을 어휘·면, 항목 수 그리고 어휘 사용 양상 등을 비교하여 해당 텍스트의 역사적 의미를 도출해 냈다. 이 분석 결과에 의하면, 일본의 조선 교육의 궁극적인 목적은 식민지 교육정책을 통해 조선이 이전보다 더 발전하고 있음을 서구 열강에 선전하기 위함이고, 이보다 더 중요한 것은 일본에 필요한 노동력을 제공할 수 있는 단순하고 기초적이며 실무 중심의 교육을 실시하는 것이었다(진실로, 2022:161)[8].

탁진영 외(2023)[9]는 1913년과 1920년 조선총독부 발행 『조선교육요람』의 영문판 분석을 기반으로 『제1차 조선교육령』이 공포된 1911년부터 『제2차 조선교육령』이 공포된 1922년 사이의 식민지 교육정책을 관찰하였다. 그 결론을 살펴보면, 일본은 교육기관의 운영 주체에 따라 차별적인 교육정책을 펼쳤으며, 다양한 유형의 교육기관을 통해 조선인에게 저급하고 차별적인 교육정책을 시행한 것을 확인할 수 있었다(탁진영 외, 2023:167). 지금까지 제1차에서부터 4차까지의 『조선교육령』의 내용 변화에 대해 살펴보았다.

본 논고는 1911년에 공포된 『제1차 조선교육령』과 『조선교육요람』(1913, 1920)의 번역본을 비교하는 것이다. 따라서 『제1차 조선교육령』에서 언급된 내용을 근거로 『조선교육요람』(1913)과 『조선교육요람(1920)』(1920)의 차이점을 관찰하도록 하겠다.

2. 역사서 텍스트마이닝에 관하여

현대에 들어서 인문학이 ‘디지털 인문학’으로 불리며 컴퓨터 발달로 인해 새로운 연구 방법으로 출현하였다. 특히, 국내의 역사서가 텍스트마이닝 연구로 인해 그동안에 진행되어왔던 시도들과는 달리 ‘디지털화’ 또는 ‘디지털화된 자료’의 데이터베이스 구축을 기반으로 데이터마이닝 기술 개발 등 새로운 방법의 연구가 시작되었다. 이를 대표할 수 있는 사례로 심재권(2020)을 언급할 수 있는데, 심재권(2020)[10]은 이순신이 직접 작성한 「난중일기」 중 임진년 일기와 임진왜란이 발발하자 그 당시의 상

황, 조선수군의 출전, 왜군의 정황, 전투 결과 등을 보고한 장계 등을 분석하였다. 이러한 연구를 위해 임진년(1592)을 크게 3기로 구분하여 각 시기에 따라 텍스트 마이닝과 연결망 기법을 활용하여 분석하였다.

라미경(2022)[11]은 사회과학 전반에 걸쳐 확산되어 있는 일본군 ‘위안부’에 대한 논의 중 국내 문헌 연구를 통하여 기존 연구의 동향을 분석하고 현재 진행되고 있는 일본군 ‘위안부’ 연구의 실태와 쟁점을 비판적으로 검토했다. 이러한 연구를 위해 모두 3분기로 나누어 텍스트 마이닝하고 분석하여 주장을 제시하였다.

마지막으로 최근의 연구로 서호준(2021)[12]을 언급할 수 있다. 서호준(2021)은 4차 산업혁명과 역사학 간의 연결고리를 찾고자 해방 이후부터 2010년대까지 약 75년에 이르는 고대사의 연구 경향을 크게 7개 시기로 구분하여 텍스트 마이닝과 소셜 네트워크 분석을 시도하였다. 이러한 결과를 바탕으로, 서호준(2021)은 논저의 수가 크게 증가하고 있는 최근으로 올수록 그 정확성이 높아지는 현상이 관찰된다고 주장하였다. 그러나 이러한 연구에도 불구하고, 텍스트 네트워크 분석은 기본적으로 키워드 분석을 위한 것이므로 문장 단위 분석이나 고차원적인 개념분석에는 약점을 갖고 있음을 밝히며 역사서를 연구할 때는 면밀하고 신중한 검토가 연구에서는 이루어져야 한다고 주장하였다.

III. 빅데이터 분석

본 연구는 조선총독부가 발행한 『제1차 조선교육령』(1911)과 1913년의 『조선교육요람』, 1920년의 『조선교육요람(1920)』 영문 번역본을 분석 대상으로 한다.

이를 위해 국립중앙박물관 아카이브를 통하여 해당 자료를 pdf 형태의 텍스트를 수집한 후, 텍스트 마이닝 [8](text mining)용 텍스트 형태로 변환하기 위해 pdf 파일을 OCR 프로그램 중 하나인 text scanner를 활용하여 Word 파일로 전환하였다. 이후, 분석의 정확성을 높이기 위해 원문과의 대조를 통해 OCR 과정 가운데 발생한 오류를 수정하였다.

그 이후 파이썬(Phython)의 기본 내장 함수와 리(re) 모듈의 정규 표현식을 이용해 텍스트의 영어 대문

자와 소문자를 소문자로 통일하고, 특수문자와 숫자 등을 제거하는 전처리 과정을 거쳤다. 특히 본 연구는 텍스트 내에서 ‘조항’의 의미로 쓰인 ‘article’의 약자인 ‘art.’와 ‘기술’의 의미로 쓰인 ‘art’가 특수문자의 처리 후 동일한 어휘인 ‘art’로 분석이 되어 ‘art.’를 ‘article’로 변환하였다.

다음으로 파이썬의 대표적인 자연어처리 라이브러리인 엔엘티케이(NLTK)를 활용하여 토큰화(tokenization)와 품사 태깅(POS tagging), 표제어 추출(lemmatization), 대명사와 전치사 등의 불용어(stopwords) 제거 과정을 거친 후 본 연구에서 분석할 핵심어를 추출하였다.

IV. 실험 및 분석

1. 핵심어 빈도 분석 결과

각 텍스트에서 추출한 출현 빈도 상위 50개의 핵심어는 다음 표 1과 같다.

표 1. 상위 50개 단어의 빈도 순위
Table 1. Top Fifty Word Frequency

순위	1911년 『조선교육령』	1913년 『조선교육요람』	1920년 『조선교육요람』
1	school	school	school
2	article	common	education
3	teach	educational	common
4	pupil	korean	give
5	give	give	korean
6	subject	government	high
7	education	education	teacher
8	study	high	government
9	high	teacher	educational
10	hour	chosen	chosen
11	year	public	special
12	language	condition	public
13	general	general	year
14	common	year	japanese
15	make	book	book
16	concern	private	condition
17	book	industrial	time
18	chosen	subject	industrial
19	national	concern	study
20	aim	time	general
21	governor	course	make
22	special	people	establish
23	allot	study	student
24	girl	ordinance	language
25	teacher	show	people
26	weak	text	regulation
27	regulation	work	private

28	per	teach	teach
29	moral	also	course
30	teaching	local	subject
31	instruction	language	girl
32	matter	girl	pupil
33	ordinance	establish	seoul
34	knowledge	make	number
35	industrial	japanese	also
36	private	student	text
37	time	regulation	national
38	course	good	use
39	condition	other	ordinance
40	use	object	primary
41	government	well	other
42	table	take	object
43	industry	special	show
44	art	number	good
45	other	authority	institute
46	one	use	result
47	necessary	national	concern
48	order	knowledge	knowledge
49	number	aim	work
50	obtain	instruction	local

먼저 각 텍스트에 자주 등장한 핵심어를 비교해보면, 세 텍스트 모두 ‘school’의 빈도수가 가장 높았다. 특히, 1911년 『조선교육령』에서는 ‘article’과 ‘teach’와 같은 어휘들이 많이 언급되었고, 1913년과 1920년 『조선교육요람』에서는 ‘common’, ‘educational’ 그리고 ‘education’과 같은 단어들이 빈번하게 출현하였다. 이는 세 텍스트 모두 조선의 교육기관에 관련한 내용을 담고 있으며, 그중에서도 『조선교육요람』은 교육기관의 유형과 교육의 체계 등과 같은 세부 사항들에 대해 주로 다루고 있으며, 『조선교육령』은 주로 교육 전반적인 조항에 대해 제시하고 있다는 것을 보여준다.

2. 워드 클라우드(Word Cloud)

본 논문은 어휘 빈도수 기반의 워드 클라우드(Word Cloud)[9]를 활용하여 조선총독부가 발행한 1911년 『조선교육령』, 1913년도 『조선교육요람』 그리고 1920년 『조선교육요람(1920)』에 나타나는 어휘 양상의 흐름을 비교 분석하였다.

다음의 <그림 1>은 1911년 『조선교육령』과 1913년 『조선교육요람』 그리고 1920년 『조선교육요람(1920)』의 어휘 빈도수를 분석하여 이를 워드 클라우드(Word Cloud)로 시각화한 자료이다.

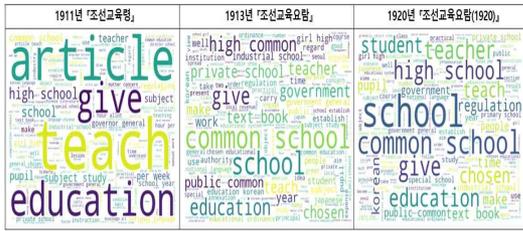


그림 1. 1911년 『조선교육령』, 1913년 『조선교육요람』과 1920년 『조선교육요람(1920)』의 워드 클라우드 비교표
 Figure 1. Word Cloud of *the First Chosen Educational Ordinance*, *Manual of Education of Koreans(1913)* and *Manual of Education in Chosen 1920(1920)*

위의 워드 클라우드(Word Cloud) 분석을 통해 본 연구는 1911년 『조선교육령』, 1913년 『조선교육요람』, 1920년 『조선교육요람(1920)』의 텍스트에서 ‘school’, ‘education’, ‘give’ 등이 공통적으로 빈도수가 높은 어휘로 분석됨을 확인할 수 있었다. 그러나 자세히 살펴보면 1911년 『조선교육령』에서는 ‘조항’을 나타내는 ‘article’과 ‘가르치다’의 ‘teach’의 출현 빈도가 높았고, 1913년 『조선교육요람』의 워드 클라우드에서는 다른 영어 텍스트에서 두드러지게 나타나지 않는 ‘text book’, ‘use’가 비교적 크게 시각화되었다는 것을 알 수 있었다. 또한, 1920년 『조선교육요람(1920)』에서는 다른 텍스트에 비해 ‘student’와 ‘pupil’의 빈도수가 높게 분석되었다. 이러한 양상은 1911년의 『조선교육령』에서는 조선총독부가 ‘교육(education)’에 대해 각 ‘조항(article)’별로 어떻게 ‘가르쳐야(teach)’ 할지에 관한 전체적인 교육의 틀을 공포하는 목적을 가진 텍스트임을 뒷받침해 준다. 또한, 1913년의 『조선교육요람』에는 ‘무엇으로’ 가르쳐야 하는가, 즉 ‘교육 매체’에 주안점을 두었고, 이는 ‘text book’의 출현 빈도로 설명될 수 있다. 그리고 1920년의 『조선교육요람(1920)』에는 ‘누가’, ‘누구에게’, ‘무슨 과목을 가르쳐야 할지’에 관한 것으로 ‘korean’, ‘chosen’, ‘pupil’, ‘student’ 등이 1913년보다 더 두드러지게 나타났다. 1920년에 이러한 단어들 이 굵게 표시된 이유는 한국어와 조선어의 수업을 대폭 줄이고 교원 임용에 있어서 한국인을 배제하고 일본인을 임용하는 경우가 늘었기 때문[13]이라고 추적할 수 있다.

다음의 <그림 2>는 바이그램 워드 클라우드(Bigram Word Cloud)로써 ‘공기어’(co-occurrence word)의 출현을 나타낸 것이다. ‘바이그램 워드 클라우드’라는 단어의 의미는 전체적인 텍스트 안에서 두 단어가 동시에 출현

하는 현상을 워드 클라우드로 나타낸 것이다.

아래의 <그림 2>는 1911년 『조선교육령』, 1913년 『조선교육요람』, 1920년 『조선교육요람(1920)』에서 두 단어가 동시에 출현하는 빈도수의 수를 워드 클라우드로 나타낸 것이다. 가령, 가장 크고 굵게 나타난 공기어는 전체 텍스트 안에서 출현한 빈도가 가장 많은 단어라고 예측할 수 있다. 아래 <그림 2>에서와 같이 1911년의 『조선교육령』에서는 다른 두 비교 텍스트에 나타나지 않는 ‘governor_general’, ‘school_year’, ‘subject_study’, ‘per_week’ 등이 빈번하게 출현했다는 것을 알 수 있다.

또한, 1913년 『조선교육요람』에는 ‘knowledge_art’, ‘education_korean’ 등의 출현이 이 다른 두 비교 대상 텍스트에서 출현하는 것보다 두드러지게 나타난다.

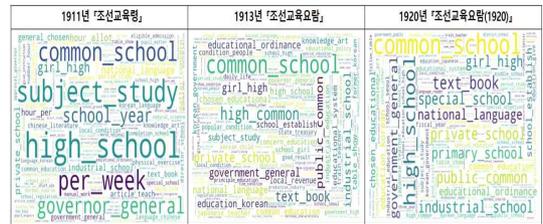


그림 2. 1911년 『조선교육령』, 1913년 『조선교육요람』과 1920년 『조선교육요람(1920)』의 Bigram 워드 클라우드(Word Cloud) 비교표
 Figure 2. Bigram Word Cloud of *the First Chosen Educational Ordinance*, *Manual of Education of Koreans(1913)* and *Manual of Education in Chosen 1920(1920)*

위의 <그림 2>의 1920년 『조선교육요람(1920)』에서는 다른 텍스트에 비해 학교의 종류와 연관되는 어휘인 ‘common_school’, ‘special_school’, ‘primary_school’, ‘industrial_school’, ‘high_school’, ‘national_school’, ‘private_school’ 등의 출현이 눈에 띄게 높다.

본 연구는 이에 대한 분석으로 1911년 『조선교육령』의 ‘school_year’, ‘governor_general’, ‘per_week’ 등과 같은 어휘의 높은 출현 빈도는 1911년의 『조선교육령』이 1913년 『조선교육요람』 그리고 1920년 『조선교육요람(1920)』의 발판[14]이 되었다고 추론한다. 다시 말해, ‘교육정책의 방법’ 또는 조선총독부가 언어 정책에 있어서 조선어나 ‘일본어(모국어)’에 대한 시간 배율과 학교의 수업연한을 개괄적으로 기획한 시기[15]이기 때

문에 워드 클라우드에 위와 같이 두드러지게 나타났다고 볼 수 있다. 또한 1913년 『조선교육요람』에는 조선 교육의 문제점을 지적하며 고급 문학이나 학문 등의 ‘지능적 과목’을 조선인들에게 가르치기보다는 ‘단순한 기능’ 교육하고자 했던 일본의 교육정책[16] 의도가 드러나게 되고 이는 ‘knowledge_art’의 높은 빈도에서 유추해볼 수 있다. 그리고 1920년의 『조선교육요람(1920)』에는 다양한 학교 종류가 출현하였는데, 이는 1920년에 조선총독부가 사립학교에 수업에 대한 규제를 강화하여 조선인들이 보통학교 또는 기술학교[17]로 많이 다니게 하는 정책을 실시하였기 때문이다. 또한, 이러한 학교들에서는 고등 학문보다 단순 기능적 과목을 더 많이 편성하였는데 이것은 일본이 조선인에게 우민화(愚民化)정책을 시행하기 위한 교육정책의 한 방편이라는 것을 예측할 수 있다.

지금까지 워드 클라우드(Word Cloud)라고 불리는 핵심 단어 시각화를 통해 1911년의 『조선교육령』, 1913년 『조선교육요람』과 1920년 『조선교육요람(1920)』 내의 어휘 출현 빈도수를 비교하여 보았다.

3. CONCOR 분석

본 연구는 어휘 출현 빈도에 기반한 워드 클라우드(Word Cloud) 외에 다양한 텍스트 마이닝 분석 방법을 조사하고, 해당 텍스트 내에서 주제가 유사한 어휘를 찾아내어 군집화하는 CONCOR(CONvergence of iteration CORrelation) 분석을 시도하였다. CONCOR 분석이란 구조적 등위성 측정 방법을 통해 행위자들 간의 복잡한 연결 관계에서 일정한 패턴을 찾은 후 그 패턴이 유사한 것들끼리 블록화하는 분석기법이다[18].

본 연구는 텍스트 내에서 유사한 연결구조를 가진 어휘들을 군집화하기 위해 우선, 1911년의 『조선교육령』, 1913년 『조선교육요람』, 1920년 『조선교육요람(1920)』의 텍스트에서 가장 많이 출현하는 상위 50개 어휘를 추출하여 분석하였다.

우선, 차례대로 1911년의 『조선교육령』과 1913년의 『조선교육요람』을 살펴본 다음 1920년의 『조선교육요람(1920)』을 관찰하였다.

아래의 <그림 3>은 1911년 『조선교육령』의 상위 빈도수 50개 어휘의 주제별 유사성에 대한 CONCOR 분석 결과이다.

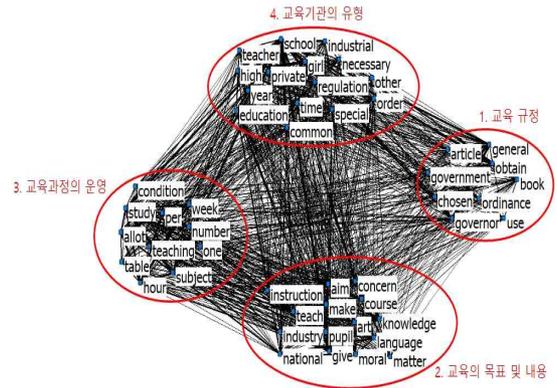


그림 3. 1911년 『조선교육령』의 CONCOR 분석
Figure 3. CONCOR Analysis of the First Chosen Educational Ordinance,

위의 <그림 3>에서 나타나듯이, 1911년의 『조선교육령』은 총 네 개의 주제(‘교육 규정’, ‘교육의 목표 및 내용’, ‘교육과정의 운영’ 및 ‘교육기관의 유형’)로 분석된다. 맨 오른쪽 시계방향의 첫 번째 동그라미에 포함되어 있는 ‘article’, ‘government’, ‘ordinance’ 등의 어휘는 ‘교육 규정’이라는 주제로 묶여질 수 있다. 그리고 두 번째 동그라미에서는 ‘instruction’, ‘aim’, ‘teach’는 ‘교육 목표’와 ‘knowledge’, ‘art’, ‘language’는 ‘교육 내용’을 담는 어휘로 분석될 수 있다. 이어서 세 번째 동그라미에는 ‘교육과정의 운영’을 나타내는 ‘hour’, ‘week’, ‘per’, ‘allot’의 어휘가 군집을 이루고 있고, 이는 ‘수업 운영’과 연관된 환경이나 조건을 지칭하는 어휘로 분석될 수 있다. 결론적으로, 1911년 『조선교육령』은 교육에 관한 내용 및 기관에 관한 ‘기초적인 규정’의 토대를 마련하고 이를 공포하는 것을 목표로 발행된 자료이다.

아래의 <그림 4>는 1913년 『조선교육요람』의 출현 빈도수 상위 50개 어휘의 주제별 상관관계를 분석한 결과를 시각화한 자료이다.

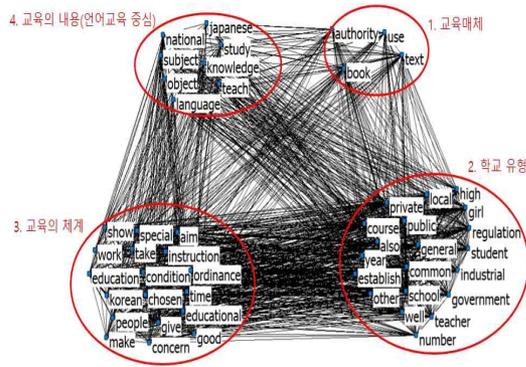


그림 4 1913년 『조선교육요람』의 CONCOR 분석
 Figure 4. CONCOR Analysis of *Manual of Education of Koreans*(1913)

위의 <그림 4>을 통해 1913년 『조선교육요람』도 1911년 『조선교육령』과 동일하게 네 개의 주제로 군집화됨을 알 수 있다. 그러나 1913년 『조선교육요람』의 주제는 ‘교육 매체,’ ‘학교 유형,’ ‘교육의 체계’와 ‘교육의 내용’으로 분석되어, 1911년 『조선교육요람』에는 없었던 ‘교육 매체’가 새로운 주제로 나타났음이 확인되었다.

어휘들의 주제별 묶음을 세부적으로 관찰하면, 첫 번째 주제 묶음은 ‘text’, ‘book’, ‘use’ 등으로 ‘교육 매체’에 관한 것이며, 두 번째 동그라미는 ‘industrial’, ‘common’, ‘private’, ‘local’ 등의 ‘학교 유형’이고 세 번째 묶음은 ‘aim’, ‘time’, ‘concern’, ‘condition’ 등의 ‘교육의 체계’라고 볼 수 있다. 그리고 마지막으로, ‘language’, ‘japanese’, ‘study’, ‘knowledge’, ‘subject’ 등의 어휘를 포함한 묶음은 ‘교육의 내용’ 특히, ‘언어 교육’과 관련된 주제이다. 이와 같은 주제별 묶음이 무엇을 의미하는지에 대해 1920년 『조선교육요람(1920)』의 주제별 CONCOR 분석과 비교하여 시간적 어휘변화의 추이를 살펴보도록 하겠다.

아래의 <그림 5>는 1920년 『조선교육요람(1920)』에 출현한 어휘 중 상위 50개 어휘의 주제별 유사성을 분석한 결과이다.

<그림 5>에서 알 수 있듯이, 1920년 『조선교육요람 1920』은 ‘교육 대상,’ ‘교육 내용,’ ‘교육기관 유형’과 ‘교육 체계’의 네 가지 주제로 군집화되는 텍스트로 분석되었다. 특히 1920년 『조선교육요람(1920)』의 주제 중 ‘교육 대상’은 다른 텍스트에서는 주제로 분류되지 않았던 주제이다. 즉 1913년 『조선교육요람』에서 ‘교육 매체’로 분석되었던 ‘text’, ‘book’의 어휘가 1920년 『조선교

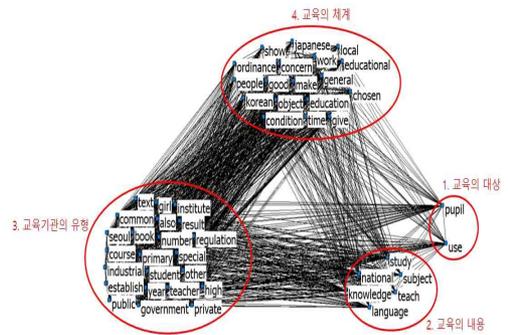


그림 5. 1920년 『조선교육요람(1920)』의 CONCOR 분석
 Figure 5. CONCOR Analysis of *the Manual of Education in Chosen 1920*(1920)

육요람』에서는 개별적으로 군집화되는 대신 ‘교육기관 유형’에 묶여서 출현하였고, 대신 1920년 『조선교육요람』에서는 새로운 어휘인 ‘pupil’이 ‘교육 대상’으로 군집화되어 출현하였음을 알 수 있다.

지금까지 1911년의 『조선교육령』, 1913년 『조선교육요람』과 1920년의 『조선교육요람(1920)』을 텍스트 마이닝 기법을 활용하여 분석하였다. 그리고 7년에 해당하는 시간의 격차를 두고 일본이 조선총독부를 통하여 조선의 교육정책을 어떻게 바꾸려 했는지 어휘의 주제별 변화 추이를 살펴보았다.

V. 결론

‘디지털 인문학’이라는 의미는 과거의 인문학을 연구하는 방법과는 다르게 학제 간 연구의 토대로 인문학과 디지털의 복합적인 연구를 의미한다.

본 연구는 역사적 의미를 지닌 사료를 디지털과 접목하여 비교 분석함으로써 역사적 의미를 되새기고 추적하는 것을 목적으로 한다. 인문학에서 연구된 사료의 의미가 디지털로 연구되었을 때도 동일한 의미를 지니는지 파악하기 위해 각 텍스트로부터 출현 빈도 상위 50개 핵심어를 도출하였고, 워드 클라우드와 바이그램 워드 클라우드를 통해 텍스트에 나타난 특징적인 어휘 양상을 분석하였다. 그 결과, 각 사료에서 나타나는 시대별 어휘 특성들로 일본의 조선에 대한 식민 교육정책 방향성을 인지할 수 있었다. 즉, 1911년 『조선교육령』은 1913년 『조선교육요람』과 1920년 『조선교육요람(1920)』의 교육정책을 어떻게 이끌어 가야 할지에 대한

기반에 관한 내용이었고, 1913년의 『조선교육요람』이 『조선교육령』의 취지 및 의의에 대한 기초작업의 ‘교육 행정 시스템’에 초점을 맞추고 있다면 1920년의 『조선교육요람(1920)』은 ‘누가’ ‘누구를’ ‘무엇으로’ ‘어떻게’ 가르쳐야 할지’에 관한 내용으로써 한 걸음 더 세부화된 ‘교육 주체와 대상’의 주제를 다루고 있다는 것이다.

더 나아가 본 연구는 CONCOR 분석 방법을 활용하여 영문 텍스트의 어휘를 주제별로 군집화하여 각 텍스트의 특징을 도출하였다. 그 결과, 1911년 『조선교육령』, 1913년 『조선교육요람』 그리고 1920년의 『조선교육요람(1920)』에서 나타나는 어휘의 변화가 워드 클라우드와 마이그램 워드 클라우드와 다소 유사한 어휘 패턴을 나타낸다는 것을 알 수 있었으며, CONCOR 분석을 통해서 워드 클라우드에서 분석할 수 없었던 텍스트별 주요 주제를 알아낼 수 있었다.

다시 말해, 1913년의 『조선교육요람』에서는 ‘교육의 매체’ 및 ‘교육 내용’이 주로 나타났으며, 1920년 『조선교육요람(1920)』에서는 ‘교육의 대상’, ‘학교 기관의 유형’이 주로 나타나는 성향이 있었다. 마지막 1911년의 『조선교육령』은 ‘교육과정의 운영’, 및 ‘교육의 목표 및 내용’에 관한 것으로써, 1913년 『조선교육요람』, 1920년 『조선교육요람(1920)』의 토대가 된다는 것을 살펴볼 수 있었다.

본 연구의 1911년 『조선교육령』, 1913년 『조선교육요람』, 1920년 『조선교육요람(1920)』의 텍스트 마이닝 분석 결과는 그동안의 질적 연구에서 찾아볼 수 없었던 어휘 연결관계를 전체적으로 조망하여 조선총독부의 교육정책을 통시적으로 분석하는데 기여하였다. 즉, 기존의 질적 연구에서는 『제1차 조선교육령』에서 공포한 교육정책에 대한 설명과 분석이 대부분이었고, 『제1차 조선교육령』 시기 동안 교육정책에 변화가 있었는지 그리고 교육정책이 실질적으로 어떻게 진행되었는지 대한 언급은 없었다. 그러나 본 연구는 『제1차 조선교육령』 시기에 발행된 1913년과 1920년 『조선교육요람』의 텍스트마이닝 분석을 통하여 그 텍스트들 사이의 연계성을 추론할 수 있었다. 결론적으로, 본 연구는 역사서의 어휘에 대한 코퍼스 분석을 통하여 질적 연구에서 찾아볼 수 없는 ‘교육정책’에 관한 통시적 역사의 흐름을 일목요연하게 읽을 수 있었다.

References

- [1] K. M. Koo, “The Name-Changing Program and the Response of Koreans under the Japanese Colonial Government,” *Korean Journal of International Relations*, Vol. 45, No. 4, pp. 31-53, December, 2005. UCI: G704-000367.2005.45.4.005
- [2] B. Y. Park, “International Society’s Opinion regarding the English Annual Report On Reforms And Progress in Chosen(1907-39) and Overseas Propaganda of the Japanese Colonial Government in Korea,” *The journal of Korean Studies*, Vol. 57, pp. 177-208, February, 2022. DOI: 10.25232/ku.2022.57.177
- [3] D. M. Kim, and J. Y. Tak, “Education Policies of the Early Japanese Colonial Period: Through Text-Mining Analysis on English Translation,” *The Journal of Interpretation and Translation Education*, Vol. 20, No. 4, pp. 29-49, December, 2022. DOI: 10.23903/kaited.2022.20.4.002
- [4] J. H. Park, and S. J. Kim, “Japanese Education of Making the Korean the Ignorant and History textbook - Focusing on textbook for Primary Schools compiled by the Japanese Government - General of Korea in the period of “1 Cha Chosun Education Law,” *The Korea Journal Japanese Education*, Vol. 1, pp.183-196, December, 2016. DOI: 10.18631/jalali.2012.56.019
- [5] S. W. Oh, “A Study on the Change of the Colonial Education Policy during Colonial Rule of Japan - Focusing on Morality-Cultivation Text Book Published by Japanese-Government General of Chosun,” *Journal of the society of Japanese Language and Literature, Japanology*, No. 27, pp. 273-292, December 2002. UCI: G704-000518.2005.27.009
- [6] W. B. Song, “An Analysis of Articles on the Korean Peninsula Revealed from the History Textbook of Japanese Government General of Korea (First Joseon Educational Ordinance) - Focused on Ancient Times Part-,” *Journal of Japanese History*, No. 39, pp. 231-258, January 2018. DOI:10.16957/sa.93.201802.231
- [7] J. Y. Tak, and M. J. Shon, “Analysis of Educational Institutions in *Manual of Education of Koreans(1913)* and *Manual of Education in Chosen 1920(1920)* Published by Japanese Government-General of Chosen: with Word Frequency Analysis and Distributed

- Representation Model,” *Journal of Humanities*, Vol. 1, pp. 167-199, August 2023.
- [8] S. L. Chin, “Analysis of Japanese Colonial Education Policy hidden in English Texts: Comparison between *Manual of Education of Koreans* (1913) and *Manual of Education in Chosen(1920)*,” *The Journal of Interpretation and Translation Education*, Vol. 20, No. 3, pp. 159-178, November, 2022. DOI: 10.23903/kaited.2022.20.3.008.
- [9] J. Y. Tak, and M. J. Shon, “Analysis of Educational Institutions in *Manual of Education of Koreans(1913)* and *Manual of Education in Chosen 1920(1920)* Published by Japanese Government-General of Chosen: with Word Frequency Analysis and Distributed Representation Model,” *Journal of Humanities*, Vol. 1, pp. 167-199, August 2023.
- [10] J. K. Shim, “Analysis of Yi Sun-shin’s Activities and Situation in the Year of the Imjin(1592) by using Text-Mining and Network,” *Journal of Association for Korean Public Administration History*, Vol. 49, pp. 139-169, August 2020.
- [11] M. K. Ra, “Current Status Analysis and Task of Domestic Research on Japanese Military ‘Sexual Slavery’ - Focusing on the Text Mining Technique,” *Journal of Patriots and Veterans Affairs in the Republic of Korea*, Vol. 21, No. 1, 9-36, March 2020, DOI:10.24004/tafp.2022.21.1.001
- [12] H. J. Seo, “Big Data and Research Trends of Korean Ancient History,” *Daegusahak*, Vol. 144, pp. 1-71, DOI:10.17751/DHR.144.1
- [13] M. S. Kang, “A Historical Study on the Enactment of the Second Chosen Educational Ordinance in the Japanese Colonial Period,” *The Korean Journal of Educational Ideas*, Vol. 23, No. 3, pp. 25-53, December 2009. DOI:10.17283/jkedi.2009.23.3.25
- [14] M. S. Kang, “A study on the student labor mobilization in the late Japanese colonial period,” *The Korean journal of history of education*, Vol. 30, No. 2, pp. 1-23, October 2008. DOI: 10.15704/kjhe.30.2.200810.1
- [15] M. S. Kang, “A Study on the Enactment of the First Chosen Educational Ordinance and Reform of School System in the Japanese Colonial Period,” *The Korean journal of history of education*, Vol. 31, No. 1, pp. 7-34, March 2009. DOI: 10.15704/kjhe.31.1.200904.7
- [16] H. J. Kim, “The Educational Policy and the Publication of Textbook by the Chosun Government-General,” *Studies on History Education*, Vol. 9, pp. 295-329, June 2009. DOI: 10.16976/kahe.2009..9.295
- [17] M. S. Kang, “A Study on the Systematization of School in the Japanese Colonial Period -Focused on the Enactment of the Second Chosen Educational Ordinance”, *The Korean journal of history of education*, Vol. 32, No 1, pp. 1-23, April 2004. DOI: 10.15704/kjhe.32.1.2010 04.1
- [18] J. R. Lee, “Digital humanities, new ways of text analysis when language changes to numbers,” *Concept and communication*, No. 22, pp.321-329, December 2018, DOI:10.15797/concom.2018..22.010

※ 이 논문은 2021년 대한민국 교육부와 한국
연구재단 일반공동연구지원사업의 지원을
받아 수행된 연구임. (NRF-2021S1A5A2A0
3062359)