

신문기사 텍스트마이닝으로 본 제 1 공화국 시기 ‘主權’의 담론과 의미 *

The Discourse and Meaning of ‘Sovereignty’ in the First Republic of Korea through Text Mining Analysis of Newspaper Articles

안범영(Ahn, BeomYoung)**

 0009-0005-3762-5105

허 수(Hur, Soo)***

 0009-0000-2056-7133

목차

1. 머리말
2. 연구 방법
3. 제 1 공화국 시기 ‘主權’의 담론
 - 3.1 담론의 해석 및 분류
 - 3.2 담론의 시기별 변동
4. 제 1 공화국 시기 ‘主權’의 의미
 - 4.1 전기(1948~1955 년)
 - 4.2 후기(1956~1960 년)
5. 맺음말

초록

이 글은 대한민국 제 1 공화국 시기에 발행된 경향신문, 동아일보, 조선일보의 주권 관련 기사 859 건을 토대로 텍스트마이닝 방법론을 활용하여 당대 주권 개념의 담론과 의미, 그리고 그 변동을 분석한 연구이다. 머신러닝을 접목한 자연어처리 기법인 토픽 모델링을 통해 거시적 차원에서 당대 ‘주권’의 담론을 파악하였으며, 단어 연결망을 통해 미시적 차원에서 ‘주권’의 의미를 확인하였다. 먼저 ‘주권’과 관련하여 제 1 공화국 시기는 1948~1955 년(전기), 1956~1960 년(후기)으로 구분될 수 있었다. 담론을 분석한 결과, 대내적 담론 4 개(개헌, 정치, 사회상, 선거)와 대외적 담론 4 개(해양주권, 세계-냉전, 세계-독립, 분단)가 도출되었다. 그 초점이 전기에는 대외적 담론에 있었으나 후기에 들어 대내적 담론으로 이동하였으며, 특히 후기의 변화는 선거의 담론에 집중되었다. 의미를 분석한 결과, 먼저 전기와 후기 모두 국민주권으로의 인식과 주권의 결함에 대한 인식이 드러났다. 전기의 경우 ‘주권’은 이양과 회복의 대상이면서도, 특히 영역의 측면에서 침해와 제약의 대상이었다. 다만 후기의 경우 ‘주권’은 옹호와 투쟁의 대상이었으며, 그 행사 수단으로 선거가 제시되어 주권 개념에 능동성이 부여되었다.

주제어: 주권, 토픽 모델링, 단어 연결망, 제 1 공화국, 디지털 역사학

*이 글은 제 1 저자의 학사학위논문으로, 2024 년도 서울대학교 미래 기초학문분야 기반조성사업의 재원으로 학부 인문 아너스 프로그램의 지원을 받아 수행된 결과물을 수정 및 보완한 것이다. 논문 작성에는 서울대학교 역사학부 김경숙 교수와 이민용 조교수, 한국학중앙연구원 한국학대학원 인문정보학 전공 김병준 조교수 및 서울대학교 국사학과 대학원생 김진규, 허현주 씨의 도움을 받았다.

**제 1 저자: 서울대학교 역사학부, 학부생, intazero@snu.ac.kr

***교신저자: 서울대학교 역사학부, 교수, crctaper@snu.ac.kr

Abstract

This paper analyzes the discourse and meaning of 'Sovereignty' in the First Republic of Korea by employing text mining methodology, based on 859 sovereignty-related articles from Kyunghyang Shinmun, Dong-A Ilbo, and Chosun Ilbo published during the era. Topic modeling - the Natural Language Processing(NLP) technology that integrates Machine Learning - was used to uncover the discourse related to the concept of sovereignty at a macro level, and co-word network analysis was employed to identify the meaning at a micro level. First, the period of the First Republic was divided into two phases: the former period(1948-1955) and the latter period(1956-1960). The discourse analysis revealed four domestic discourses(Constitutional Amendment, Politics, Social Situation, and Election) and four international discourses(Maritime Sovereignty, World-Cold War, World-Independence, and the Division of Korea). The focus of the discourse shifted from international(the former) to domestic(the latter), with a particular emphasis on the discourse of election. The meaning analysis, at first indicated that both periods reflect an awareness of popular sovereignty and a recognition of deficiencies in its exercise. During the former period, 'Sovereignty' was perceived to be transferred and restored, but to be infringed and restricted especially in terms of territory. In the latter, 'Sovereignty' became a matter of defense and struggle with election being presented as the means of its exercise, thereby imparting proactivity to the concept.

Keywords: Sovereignty, Topic Modeling, Co-word Network, the First Republic of Korea, Digital History

1. 머리말

1948년 헌법(제헌헌법)에 따르면 대한민국은 민주공화국이며(제1조), 그 주권과 권력은 국민으로부터 나온다(제2조). 두 조항은 9번의 헌법 개정 및 제정 과정에서 단 한 번도 삭제되거나 후순위로 밀려난 바가 없었던 것들로, 현행 헌법 제1조에서도 확인해 볼 수 있다.¹ 대한민국이 국민주권에 기반을 둔 민주공화국이라는 대내외적 선언은 1919년 임시정부 수립, 1948년 8월 15일 정부 수립 이후 지금에 이르기까지 절대 불변하였던 것이었다. 또 대한민국은 1948년 12월 12일 제3차 유엔 총회에서 그 정부가 한반도 내 유일 합법 정부로 승인되었다. 비록 5·10 총선거가 가능했던 지역(38선 이남)에 한정된다는 제약이 존재하였으나, 대한민국은 한반도 내 유일한 주권국가로서 그 정통성을 인정받을 수 있게 되었다.²

주권(主權, sovereignty)은 대내적으로는 최고의 정치권력을, 대외적으로는 타 권력체로부터의 독립과 평등을 의미한다. 대내적 최고성과 대외적 독립성으로 함축되는 이 개념은 고대 로마에서 기원하여 중세 서양을 거치며 그 필요성이 대두되기 시작하였고, 근대에 장 보댕(Jean Bodin)의 《국가론》을 통해 대내적 차원에서의 주권 개념이 처음으로 정립되었다. 이후 30년 전쟁과 베스트팔렌 조약으로 촉발된 국제법의 성장과 함께 대외적 차원에서의 개념이 점차 정립되었다.³

근대 서양에서 형성된 주권 개념이 한반도에 수용되고 전개된 것은 개화기 이후였다. 주권 개념이 막 전래되었을 때 조선의 지식인들은 전통적 사대관계와 배치되었던 이 개념에 보수적으로 접근하였다. 이후 주권 개념이 주로 대외적 차원에서 논의되는 동시에 여러 주권론이 경합하였으며, 대한제국 수립을 거치면서 주권에 관한 논의는 대외적으로는 청(淸)의 예속에서 탈피한 자주 독립국가로의 전환으로, 대내적으로는 군주주권론으로 귀결되었다. 20세기에 들어 주권 개념은 법학 교과서, 잡지, 신문 등을 통해 지속적으로 보급되고 전파되었다. 또 국권을 상실해가면서 군주주권론이 우세한 가운데에서도 국민주권의 인식과 민족적 주권 인식이 점차 부상하였다. 이러한 양상은 국권 상실 이후에도 계속 진행되어 1917년 〈대동단결선언〉에서 그러한 인식이 명문화되는 성과가 나타났으며, 1919년 3·1운동 이후 수립된 대한민국임시정부의 《대한민국임시헌법》(1919.9.11.)의 주권 법정화(法定化)에도 영향을 미쳤다. 이와 같은 흐름을 토대로 대한민국임시정부는 다섯 차례 헌법을 개정하면서 망명정부였던 당시의 현실로 인해 발생한 주권적 측면에서의 한계를 인식하면서도, 지속적으로 대한민국이 민주공화국임을 천명하였다.⁴

상술하였듯이 해방 이후 대한민국 정부 수립을 통해 국민주권이 대내외적으로 회복·확립되었으나, 그럼에도 당시 대한민국의 주권은 일말의 불완전성을 내재하고 있었다. 먼저 영토적 측면에서 대한민국은 조선민주주의인민공화국(‘북한’)의 존재로 인해 헌법상 영토주권과 실질적 영토주권이 괴리되어 있었다.⁵ 한반도가 분단되면서 국민의 왕래도 단절되었으므로, 이러한 대한민국의 불완전주권성은 영토적 분단과 더불어 인적 분단에서도 기인하였다.⁶ 이와 더불어 샌프란시스코 강화조약이 체결되기 전까지 한국과 일본은 국제법적으로 완전히 분리되지 않은 상태에 있기도 했다. 미국은 연합국최고사령관 각서(SCAPIN) 699호를 통해 한반도에 대한 일본의 통치권 및 행정권이 상실되었다는 사실을 일본에 통보하였다. 이로써 한반도에 대한 일본의 주권은 통치적 차원에서 상실되었다. 그러나 미국은 선례(先例)에 따라 이러한 주권과는 별개로 한반도에 대한 영토적 주권은 별도의 조약을 체결하기 전까지 일본이 보유하고 있다고 보았다. 이로 인해 대일강화조약이 체결되기 이전까지 한반도에 대한 일본의 영토주권은 여전히 법적으로 잔존하고 있었다.⁷

그러나 적어도 명목상으로 대한민국 국민이 주권을 담지하고 있었다는 사실은 부인할 수 없다. 1948년 정부 수립을 통해 명목상 온전해진 국민주권은 초기에 어떻게 인식되었을까. 제1공화국의 주권 인식에 대한 실마리를 찾기 위해 먼저 금보운의 연구를 주목할 수 있다. 금보운은 1957년 전후에 주목하여 6·25전쟁 이후에도 지속되었던 주한미군과 유엔군의 광범위한 공간적 권한에 대한민국 정부와 민간인들이 가졌던 반응과 주권 인식을 밝혔다. 이 연구에 따르면 민·관은 그들의 권한이 주권자인 자신들보다 우위에 있다는 사실을 확인하였다. 또 분단 체제하에서 외국군의 군사력을 사용하며 주권 영역 내에서 권한을 공유하는 ‘제한적 주권’을 가진 탓에 정부는 공간적 권한을 일부 회수하기 위해 노력하면서도 그들의 주둔 자체를 반대하지는 못했다.⁸ 다음으로 한모니카는 미군정이 대한민국 정부에 정권을 이양하는 과정을 검토하면서, 1948년 8월 24일 체결된 ‘대한민국 대통령과 주한 미군사령관간에 체결된 과도기에 시행될 잠정적 군사안전에 관한 행정협정’(한미과도군사협정)에 대하여 당시 정부는 주권 침해와 관련하여 협정이 끼칠 수 있는 과장과 위험성을 인지하면서도 이를 묵과한 채 비밀리에 협정을 체결하였다는 사실을 제시하였다.⁹

시각을 더욱 확장하여 1987년 이전 대한민국에서의 주권 인식을 분석한 연구를 살펴보자면 다음과 같다. 허은은 1960~1980년대 학생운동의 변화를 추적하여 당시 대학생들 비롯한 지식인들의 주권 개념에 대한 인식이 한국의 절대적 주권 확보에서 남북한과 미국이 연결된 ‘상호의존적 주권’ 인식과 전 민족적 자주권 확보로 전환되었음을 밝히면서, 여기에 미국의 존재가 끼친 영향을 규명하였다.¹⁰ 정무용은 1961년에 체결된 한미경제기술원조협정에 관한 정치·사회적 갈등을 분석하면서, 조항을 둘러싼 주권 침해 논란과 이에 대한 정부의 반박과 대응을 포괄하여 협정이 조인, 비준되는 과정을 설명하였다. 아울러 반대파를 향한 정부의 이데올로기적 공세로 논쟁의 축이 변화하는 양상을 보여주면서도, 그 이면에는 경제적 자립과 그 실천법을 두고 대립하는 측면도 있었다는 사실을 밝혔다.¹¹崔航航은 한반도 인근 바다의 해양주권에 대하여 박정희 정부가 국가주도 경제개발을 위한 에너지 자원 확보의 맥락에서 제1공화국의 인식을 계승하는 동시에 국제공약을 취사선택하면서 인식했다는 사실을 밝혔다. 그리고 당시의 분단체제와 국제관계하에서 그러한 인식의 한계가 드러나는 과정을 소개하였다.¹²

이렇게 제1공화국, 더 나아가 1987년 이전 대한민국에서의 주권 인식은 지금까지 한미관계와 주한미군과 같은 특정 주제, 혹은 대학생과 같은 인식 주체와 결부되어 분석되었다. 그러나 선행연구는 그 수도 많지 않을뿐더러, 허은의 연구를 제외하면 주권 개념 그 자체가 조망된 적이 없다는 한계를 지닌다. 허은은 주권 개념에 대한 인식을 통시적으로 추적하고 있기는 하지만, 연구는 제2공화국 시기에서 본격적으로 출발한다는 점에서 그 이전 시기에 대해서는 여전히 시계열적

관점에서 분석된 바가 없다.

현대와 달리 정부 수립 이전 근대 한국에서의 인식을 분석한 연구는 상대적으로 다수가 존재한다. 이러한 연구들은 주권 개념의 수용 혹은 주권의 인식 주체에 주목한 경우, 오영달과 같이 주권을 바라보는 특정 시각을 실제 역사에 투사하여 사회과학적 연구를 수행한 경우로 나눌 수 있다.¹³ 한국근대사 속 주권 개념의 수용과 변용에 대한 연구는 현재 많이 진척되어 있다. 그러나 대부분 질적 방법론을 채택하여 양적 방법론을 통한 분석이 본격적으로 시도된 적은 드물었다는 점에서 여전히 연구의 공백이 발생하고 있으며, 이는 현대사 영역에서의 연구 역시 동일하다.

이에 본고는 토픽 모델링, 단어 연결망과 같은 디지털 인문학 방법론을 활용하여 양적 방법론의 측면에서 대한민국 제1공화국 시기 ‘주권’의 담론과 의미를 통시적으로 분석한다. 이 작업은 제1공화국 시기에 발행된 《경향신문》, 《동아일보》, 《조선일보》 기사 데이터 분석을 통해 진행된다. 이 글이 해명하고자 하는 과제는 대한민국임시정부의 헌법에서 기원하여 1948년 정부 수립을 통해 비로소 체감될 수 있게 된 국민주권이 초기 대한민국에서 어떠한 맥락에서 인식되었으며, 또 그러한 인식이 어떻게 변화했는지를 규명한다는 점에서 의미가 있다.

따라서 본고의 구성은 다음과 같다. 2장에서는 《경향신문》, 《동아일보》, 《조선일보》를 주목한 이유, ‘주권’을 분석하기 위해 기사 데이터에 적용한 기준, 그렇게 선별된 데이터의 기술통계량과 함께 토픽 모델링과 단어 연결망 기법으로 결과를 도출하기 위해 거쳤던 과정을 상세히 소개한다. 이와 더불어 통시적 분석에 필요한 시기 구분 결과에 대해 설명한다. 3장에서는 토픽 모델링으로 도출된 토픽을 소개하고, 그렇게 도출된 당대 ‘주권’의 담론과 그 변화에 대해 설명하면서 거시적 차원에서 개념에 접근한다. 4장에서는 ‘주권’의 의미와 그 변화를 알 수 있는 시기별 단어 연결망을 제시한 뒤, 그것들이 가지는 의미를 해석함으로써 미시적 차원에서 개념에 접근한다.

2. 연구 방법

연구의 진행 과정을 본격적으로 설명하기에 앞서 본고에서 사용하는 주요 방법론인 토픽 모델링과 단어 연결망에 대해 개관할 필요가 있다. 토픽 모델링은 대량의 텍스트에서 잠재된 토픽을 찾는 통계적 방법론을 말한다. 여기서 토픽은 특정 주제로 분류될 가능성이 높은 단어들의 집합이다. 토픽 모델링은 인공지능의 하위 개념인 기계학습(Machine Learning)의 범주에 포함되며, 비지도학습을 통해 구현되어 텍스트에 대한 인간의 개입 없이 문서를 학습하여 토픽과 그를 구성하는 단어들을 산출할 수 있다. 토픽 모델링의 대표적인 방법론은 잠재 디리클레 할당(Latent Dirichlet Allocation)이다. 이 모델은 문서들이 다수의 잠재된 토픽을 함유하고 있다는 가정하에서 분석을 시행한다. 그렇기에 LDA 모델을 통해 문서에 대한 각 토픽의 비율을 알 수 있기도 하다. 아울러 LDA를 확장한 토픽 모델링 방법론이 존재하는데, 예를 들어 저자, 연도와 같은 부가적인 정보, 즉 메타데이터(metadata)를 추가로 고려한 디리클레 다항 회귀분석(Dirichlet Multinomial Regression)과 같은 모델이 존재한다.¹⁴ 본고에서는 이러한 토픽 모델링 방법론 중 DMR 모델을 사용하며, 그 메타데이터를 발행연도로 하여 토픽 분석에 발행연도를 추가로 고려한다.

단어 연결망은 유사도 등에 기반하여 단어들 간의 연결 관계를 시각화한 그래프이다. 연결망은 기본적으로 개체에 상응하는 노드(Node), 그리고 개체와 개체 간의 연결 관계를 표시하는 링크(Link)로 구성된다. 노드의 경우 대체로 연결망의 중심에 있을수록, 그리고 다른 노드와의 연결 관계가 많을수록 중요하다고 취급된다.¹⁵ 또한 노드의 크기, 링크의 두께를 통해 단어의 중요도와

연결 관계의 강약을 파악할 수 있고, 군집화(Clustering)를 통해 노드들의 집단을 확인해 볼 수도 있다. 아울러 단어 이외에도 저자, 토픽과 같은 데이터를 노드로 삼아 연결망을 형성할 수 있으며, 토픽-단어 2원 연결망과 같이 노드의 종류를 다양화할 수도 있다.

본고는 이 방법론들을 바탕으로 신문 기사를 사료로 활용하여 제1공화국 시기 ‘주권’의 담론과 의미를 논의한다. 이러한 분석에서 《경향신문》, 《동아일보》, 《조선일보》 등의 신문 기사는 어떤 사료적 가치를 가질까. 1950년대에 식자층이 급증하면서 신문의 대사회적 영향력이 증가하였고, 세 신문은 《서울신문》, 《한국일보》와 더불어 당대에 인기 있는 신문들이었다.¹⁶ 1954년 10월 말 일간신문의 총 발행 부수는 대략 50만 부였는데, 이중 《동아일보》는 8만 부, 《조선일보》는 6만 부, 《경향신문》은 4만 3천 부가 발행되었다. 1955년 말에는 전체 198만 부 중 석간 《동아일보》가 17만 6000부, 석간 《경향신문》이 10만 부, 조간 《조선일보》가 8만 부를 발행하였다. 1959년에는 《동아일보》가 35만 부, 《경향신문》이 20만 부, 《조선일보》가 10만 부를 발행하였다.¹⁷ 발행 부수에서 알 수 있듯이, 세 일간지는 발행 부수의 양적 측면에서 높은 순위에 있었다.

아울러 당시의 신문 매체들은 정론지(政論紙)의 성격을 매우 강하게 띠고 있었다. 또 신문사의 운영을 광고 수입에 크게 의존하지 않고 있었기에, 권력으로부터 상대적으로 자유로울 수 있었다. 그로 인해 언론 매체에 대한 정권의 탄압이 가해져도 《경향신문》, 《동아일보》, 《조선일보》는 ‘야당지’의 기능을 수행하면서 정권의 행위를 비판할 수 있었다.¹⁸ 실제로 1952년 1월부터 1953년 3월까지의 기간 동안 정부 시책에 대한 찬성과 비판 논평은 《조선일보》의 경우 100건-173건, 《경향신문》은 73건-151건, 《동아일보》는 12건-191건을 기록했다.¹⁹ 이렇게 《경향신문》, 《동아일보》, 《조선일보》는 정권의 권력에 굴복하지 않고 당시의 사건과 현안을 포착하고 논의할 수 있었다. 주권(主權)에 대한 논의 역시 그러하였으며, 이는 기사 텍스트에 남아있는 다수의 증거를 통해 확인된다.

현재 세 신문은 모두 디지털화된 동시에 데이터베이스화까지 완료되어 있다. 본고에 활용된 《경향신문》 기사는 <네이버 뉴스 라이브러리>에서, 《조선일보》와 《동아일보》 기사는 각각 <조선 뉴스 라이브러리 100>, <동아디지털아카이브>에서 수집되었다.²⁰ 분석의 기본 도구로 프로그래밍 언어 ‘Python’을 사용하였으며, 개발 환경으로 ‘Jupyter Notebook’과 ‘Google Colaboratory’를 사용하였다. 단어 연결망 작성은 ‘NetMiner 4.5.1.d’를²¹ 사용했다.

본고는 신문기사 선별을 다음 기준에 따라 진행하였다. 먼저 기사 데이터셋에서 기사 제목, 제목 및 본문, 발행일자, 발행연도, 신문사에 대한 열(column)만을 남겼다. 이후 대한민국 정부가 수립된 1948년 8월 15일부터 3차 개헌으로 제2공화국이 수립되기 전날인 1960년 6월 14일까지의 기사 중 주권, 主權이 포함된 기사만을 분리하였다. 그러나 주권은 主權과 더불어 株券, 舟券으로도 이해될 수 있었고, 또 자주권(自主權), 거주권(居住權), 영주권(永住權), 종주권(宗主權), 국민주권(國民主權), 주권재민(主權在民)과 같이 다양한 문자열 혹은 단어가 결부되어 사용되는 경우를 다수 발견할 수 있었다. 이에 기사들을 살펴면서 主權과 아무 관련이 없는 기사들을 제외하였다.

다음으로 기사의 제목에 주권(主權) 혹은 그를 포함하는 단어가 1번 이상 언급되거나, 그러한 단어가 본문에 3번 이상 언급된 기사만을 선별하였다. 본문에 단어가 1~2번만 언급된 경우, 실제 주제와 상관없이 단순히 단어를 언급하였을 가능성을 배제하기 어려웠다. 또 대량의 데이터를 다루는 상황에서 주권과 상관이 없으나 미처 발견하지 못했던 기사들을 제외하는 효과도 있었다. 이러한 이유에서 상술한 기준에 부합하는 기사가 주권을 글의 주된 요소로 다루고 있다고 간주하였으며, 그러한 기사들만을 사료로 선별하였다. ‘主權(주권)’과 같이 한자어와 그 독음이 같이 나열된 경우에는, 1번만 언급된 것으로 간주하였다.

이와 더불어 주권(主權)이 단체 이름에 들어가는 경우를 고려하였다. 예를 들어 1950년대 중후반부터 장택상을 위원장으로 하는 ‘국민주권옹호투쟁위원회’를 다룬 기사들을 다수 확인할 수 있었다. 이러한 단체들의 경우 그 이름에 主權이 포함된다는 점에서 당시 사회에서 단어가 어떻게 인식되고 수용되었는지를 보여주는 좋은 지표라 할 수 있다. 다만 단체들이 언급된 모든 기사가 주권을 논하고 있다고 단언하기는 어려웠다. 그렇기에 단체가 언급된 기사의 경우, 주권(主權)이 그 이름에서만 언급된 기사는 모두 제외하였다.

표 1. 주권 관련 기사 수 (연도별, 신문사별)

연도 \ 신문사	경향신문	동아일보	조선일보	합계
1948	10	12	11	33
1949	26	41	24	91
1950	4	15	3	22
1951	0	12	2	14
1952	27	31	12	70
1953	19	26	19	64
1954	57	43	44	144
1955	16	21	25	62
1956	21	36	33	90
1957	23	28	18	69
1958	30	39	18	87
1959	14	26	22	62
1960	3	24	24	51
합계	250	354	255	859

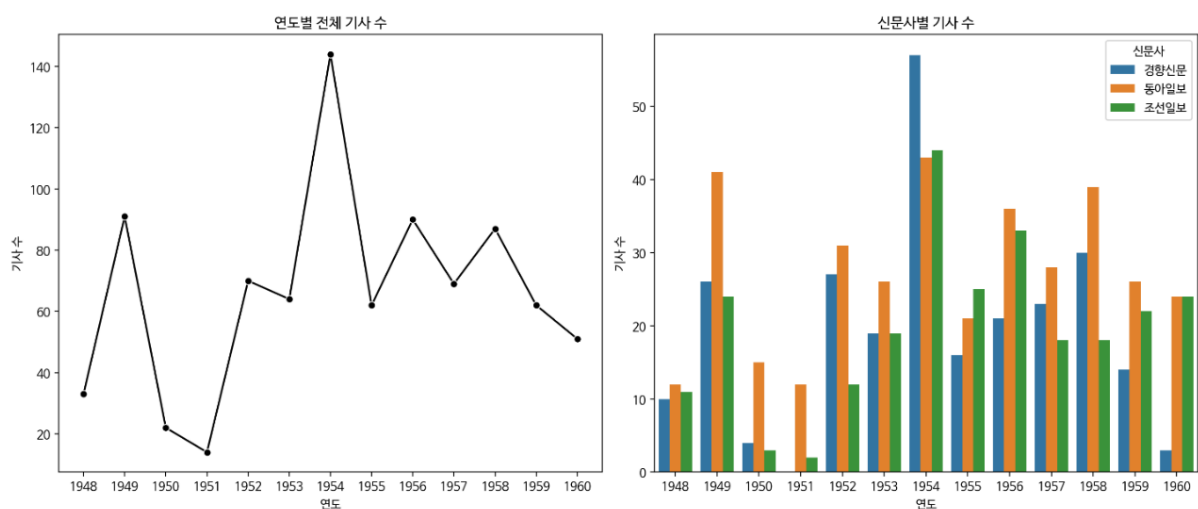


그림 1. 주권 관련 기사 수 그래프 (연도별, 신문사별)

이러한 기준을 통하여 모두 859건의 기사를 선별하였다. 연도별, 신문사별 기사 수와 그래프는 〈표 1〉, 〈그림 1〉과 같다. 나아가 데이터의 ‘제목 및 본문’ 열에 대해 ‘Python’의 ‘hanja’ 라이브러리를²² 통해 국한문혼용체로 작성된 원문을 순수한 한글 문장으로 변환하고 ‘kiwi’ 라이브러리를²³ 사용하여 형태소 분석을 진행한 뒤, 일반 명사와 고유 명사를 중심으로 단어를 선별하면서 1차적으로 제목 및 본문에 대한 코퍼스 구축을 완료하였다. 이후 분석 오류 수정, 용어의 통일 혹은 분리, 불용어 제거의 과정을 거쳐 코퍼스를 직접 정돈하였다.

다음으로 토픽 모델링을 수행하고 단어 연결망을 형성하는 데 필요한 공기어, 즉 주권 개념과 같이 나타나는 단어의 문맥 범위를 설정하였다. 토픽 모델링의 경우 상술한 바와 같이 거시적 차원에서 주권 개념의 담론을 살펴보는 데 목적이 있으므로, 화제적 차원에서 ‘주권’을 분석하기 위해 공기어의 문맥을 넓게 설정할 필요가 있었다.²⁴ 따라서 토픽 모델링에서의 문맥은 기사 전체로 설정하였다. 환언하면 하나의 기사가 하나의 문맥, 즉 문서에 해당하므로 토픽 모델링에는 총 859개의 문서가 투입되었다. 이와 달리 단어 연결망은 주권 개념의 의미에 집중하는 데 목적이 있으므로, 공기어의 문맥을 좁게 설정하였다. 이에 단어 연결망에서의 문맥은 앞서 구축한 코퍼스를 기반으로²⁵ ‘주권’, ‘주권자’, ‘주권선’, ‘자주권’, ‘주권론’, ‘주권성’ 등 ‘주권’과 그를 포함하는 단어, 그리고 그 주변의 단어 4개를 문맥으로 설정하였다.²⁶ 앞서 살펴본 주권 관련 기사 내에서 ‘주권’ 관련 단어의 개수가 둘 이상 되는 경우도 많았으므로, 단어 연결망 분석에는 총 3,028개의 문서가 투입되었다.

이렇게 분석의 재료가 되는 데이터를 준비한 뒤, ‘tomotopy’ 라이브러리를²⁷ 사용하여 토픽 모델링을 실시하였다. 본격적인 분석에 앞서 데이터로부터 산출될 토픽의 개수를 지정할 필요가 있었다. 토픽의 개수를 지정하는 것은 전적으로 연구자의 몫이지만, 분석의 성능을 평가할 수 있는 지표를 통해 그 개수를 결정하는 것이 좋다.²⁸

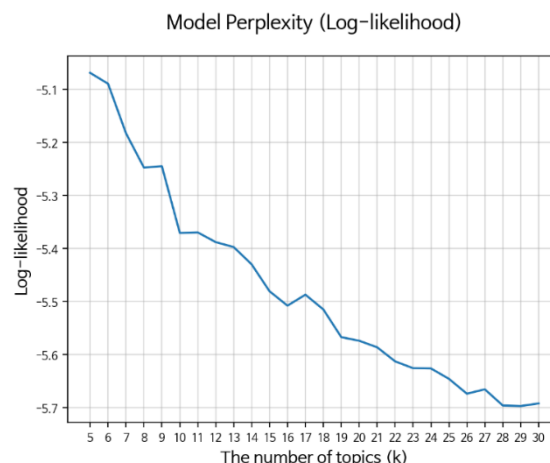


그림 2. 혼잡도(Perplexity) 그래프

본고는 〈그림 2〉와 같이 토픽의 개수에 따른 혼잡도(Perplexity)를 지표로 하여 분석의 성능을 파악하였다. 혼잡도는 언어 모델링(Language Modeling)에서 사용되는 지표로, 그 점수가 낮을수록 분석의 성능이 좋다는 것을 의미한다.²⁹ 그러나 이와 같은 지표들을 통해 도출되는 분석 성능이 토픽에 대한 이해(interpretation)와 반드시 직결되는 것은 아니며,³⁰ 인간 실험을 통해서도 증명된 바가 있다.³¹ 혼잡도는 토픽의 개수에 증가함에 따라 지속적으로 감소하지만, 방금 살펴본 바와 같이 많은 수의 토픽이 반드시 좋은 결과를 이끌어내는 것은 아니다. 따라서 적절한 지점에서 토픽의 개수를 결정해야 한다. 그중에는 혼잡도의 경향이 감소에서 증가로 변화하기 직전의 지점을

주목하는 방법이 있다. 본고는 이 방법을 채택했으며, <그림 2>의 그래프에 따라 $k=8, k=10, k=16, k=26$ 과 같은 후보를 선정하였다. 이후 ‘주권’ 관련 기사를 분석하기에 도출될 수 있는 토픽이 비교적 한정된다는 점을 고려하면서 최종적으로 8개의 토픽을 선정하였다.

토픽의 개수를 결정한 뒤에는 각 토픽을 구성하는 단어를 30개로 설정하고 토픽을 도출하였다. 아울러 각 토픽 내 여타 단어들과 이질적인 단어들을 제거하기 위해서, 토픽 단어들의 유사도를 점별 상호정보량(PMI)을 통해 계산하여 토픽을 보정하였다.³²

단어 연결망은 다음과 같은 과정을 거쳤다.

- ① 시기별 고빈도 100개 단어 Node-Cut → ② 단어 연결망 1차 작성 → ③ Link-Cut
→ ④ Component 분석 → ⑤ Community 분석 및 단어 연결망 완성

①은 ‘Python’에서, 나머지는 ‘NetMiner’를 통해 수행하였다. ①은 중요 단어들의 연결 관계를 간결하게 시각화하기 위한 기초 작업으로, 후술할 각 시기의 고빈도 100개 단어에 관한 정보만을 간추린 파일을 작성하였다. 이후 ②의 과정으로 코사인 유사도(Cosine Similarity)³³를 통해 단어 간의 연결관계를 계산하면서 연결망을 1차적으로 작성하였다. 다만 ①을 거쳐도 노드와 그에 따른 링크가 너무 많아 연결망이 번잡해지기에, 중요한 단어와 연결관계를 간결히 파악하기 위해 ③에서 일정 유사도(weight) 미만인 링크를 삭제하였다.³⁴ ③까지 수행하면 연결망이 한결 간결해지지만 아무 노드와 연결되어 있지 않은 고립 노드, 연결관계가 희박한 노드들이 발생한다. 이를 삭제하기 위해 ④의 과정으로 Component 분석을 실시하였다. 마지막으로 ⑤에서 Community 분석을 실시하여 노드들을 군집화했다.³⁵ 그리고 연결 중심성에 따라 노드의 크기를,³⁶ 소속 클러스터에 따라 색상과 모양을 결정하고, 유사도를 기준으로 링크의 두께를 조정하여 연결망을 완성하였다.

표 2. 연도별 상관관계

연도	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960
1948	1	0.811314	0.734659	0.958514	0.353064	0.894726	0.756177	0.546024	0.427708	0.423308	0.163085	0.784708	-0.06204
1949	0.811314	1	0.62946	0.820834	0.091776	0.635117	0.658995	0.658643	0.255834	0.202724	-0.06471	0.507835	-0.25216
1950	0.734659	0.62946	1	0.822561	0.351109	0.648186	0.637414	0.369783	0.633627	0.669319	0.431976	0.907061	0.244604
1951	0.958514	0.820834	0.822561	1	0.312535	0.809372	0.825446	0.659227	0.595921	0.567408	0.263549	0.818153	0.071629
1952	0.353064	0.091776	0.351109	0.312535	1	0.689045	-0.03623	0.187747	-0.10955	-0.10444	-0.3028	0.609629	-0.23652
1953	0.894726	0.635117	0.648186	0.809372	0.689045	1	0.481539	0.449725	0.181008	0.19732	-0.02932	0.829218	-0.19426
1954	0.756177	0.658995	0.637414	0.825446	-0.03623	0.481539	1	0.698881	0.510312	0.487095	0.220302	0.489837	-0.03801
1955	0.546024	0.658643	0.369783	0.659227	0.187747	0.449725	0.698881	1	0.198693	0.089801	-0.21479	0.31365	-0.30385
1956	0.427708	0.255834	0.633627	0.595921	-0.10955	0.181008	0.510312	0.198693	1	0.982597	0.853021	0.581055	0.813017
1957	0.423308	0.202724	0.669319	0.567408	-0.10444	0.19732	0.487095	0.089801	0.982597	1	0.913221	0.619034	0.834482
1958	0.163085	-0.06471	0.431976	0.263549	-0.3028	-0.02932	0.220302	-0.21479	0.853021	0.913221	1	0.400734	0.923828
1959	0.784708	0.507835	0.907061	0.818153	0.609629	0.829218	0.489837	0.31365	0.581055	0.619034	0.400734	1	0.267498
1960	-0.06204	-0.25216	0.244604	0.071629	-0.23652	-0.19426	-0.03801	-0.30385	0.813017	0.834482	0.923828	0.267498	1

마지막으로 시기 구분은 연도별 각 토픽의 비율에 따라 계산한 상관관계에 따른 결과를, 단어 기반의 계층적 군집화(Hierarchical Clustering)에 따른 결과를 통해 보완, 확정했다. DMR 모델을 통한 토픽 모델링은 동일한 코퍼스의 내용 변화를 반영하기 때문에, 이를 바탕으로 연도별 유사도를 상관관계수로 측정하여 시기를 구분할 수 있다.³⁷ 본고는 앞서 서술했듯이 DMR 모델에 대한 메타데이터를 기사의 발행연도로 설정하였기에 연도별 각 토픽의 비율을 산출할 수 있었다. 이를 기반으로 피어슨 상관관계수(Pearson Correlation Coefficient)³⁸를 활용하여 연도별 상관관계를 계산할 수 있는데, 그 결과는 <표 2>와 같다.

이때 서로 양의 상관관계를 강하게 보이는 연도들을 확인하기 위해 상관계수가 0.6 이상인 부분을 음영 처리하여 시기 구분 결과를 시각화하였다. 이를 통해 1948~1955년을 전기(前期), 1956~1960년을 후기(後期)로 하는 결과를 얻을 수 있었다. 다만 기준이 되는 상관계수의 값을 필자가 자의로 결정하였고, 또 각 시기 내에서 강한 양의 상관관계를 보이지 않는 연도들이 존재한다는 점에서 보완이 필요했다. 따라서 본고는 ‘scikit-learn’ 라이브러리를³⁹ 사용하여 토픽 모델링에서 투입한 859개 문서를 토대로 TF-IDF 값⁴⁰ 합산 상위 50개 단어에 대한 문서-단어 행렬(DTM)을 만들고, 다시 연도-단어 행렬로 변환하였다. 이를 기반으로 연도별 상관관계를 코사인 유사도로 계산한 뒤, ‘SciPy’ 라이브러리를⁴¹ 통해 계층적 군집화 분석을 실시하였다.

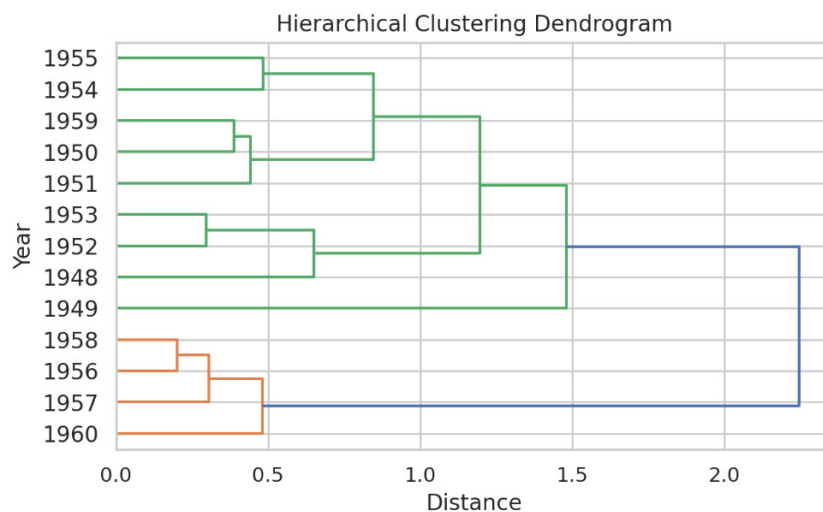


그림 3. 계층적 군집화에 따른 연도 덴드로그램(Dendrogram)

〈그림 3〉에 따른 결과 역시 1948~1955년을 전기로, 1956~1960년을 후기로 하여 제1공화국을 두 시기로 구분하고자 하는 본고의 결정을 뒷받침한다. 다만 이 그래프에 따르면 1959년은 1956, 1957, 1958, 1960년과 다른 집단에 존재하고 있다. 그러나 〈표 2〉를 참고하였을 때 1959년이 집단의 나머지 연도들과 완전히 유리되어 있지는 않은 것으로 파악되고, 본고는 제1공화국 시기에서의 통시적 변화를 추적하는 것을 주요 목적으로 두는 만큼 시기 구분의 측면에서 1956~1960년을 후기로 그대로 두어도 된다고 판단했다. 따라서 본고에서는 대한민국 제1공화국의 시기를 1948~1955년, 1956~1960년으로 이분하여 분석을 진행하였다.⁴²

3. 제 1 공화국 시기 ‘主權’의 담론

3.1 담론의 해석 및 분류

2장에서 살펴본 바와 같이, 당시 ‘主權’의 담론과 그 변화를 알기 위해 토픽 모델링 분석을 실시하여 8개의 토픽을 도출하였다. 〈표 3〉은 각 토픽과 그를 구성하는 단어들에 대한 정보가 나열되어 있는 표이다. 토픽의 이름은 필자가 직접 명명하였다.

표 3. 토픽 관련 정보

토픽	단어
T0 (30) (개헌)	헌법 국회 대통령 국민 의원 규정 개헌 개정 국가 정부 법률 주권 개헌안 자유당 국민투표 제도 조항 행정부 이승만 제한 사항 책임 국회의원 국민투표제 제정 법 중대 반대 선거 입법
T1 (29) (해양주권)	일본 한국 정부 선언 해양 주권 국제 인접 주권선 수역 자원 침범 보호 어업 국가 태도 국제법 어선 영토 영해 평화선 자유 인정 관계 공해 한일회담 동경 대륙붕 조치
T2 (30) (세계-냉전)	미국 서독 소련 주권 프랑스 조약 영국 독일 회의 중국 협정 회담 회복 유럽 대통령 방위 비준 재무장 국가 완전 합의 부여 자유 NATO 평화 가입 요구 EDC 운하 동독
T3 (30) (정치)	국민 정치 국가 민주 민주주의 자유 나라 주권 정당 민족 사회 발전 행사 선거 권리 책임 정신 공무원 인민 권력 정책 법 보장 조직 기본 기관 헌법 행정 주권자 운동
T4 (30) (사회상)	주권 언론 사건 신문 법 대회 학생 자유 보도 시민 불법 대한민국 기자 수호 경찰 국가보안법 구속 무효 운동 반대 전국 공산당 결의 진보당 범죄 침해 정신 평화 체포 당국
T5 (29) (선거)	선거 의원 자유당 야당 민주당 국민 주권 투쟁 경찰 투표 표 국회 등록 부통령 데모 여당 위원 방해 여야 유권자 사건 사태 지방 당선 개표 행사 옹호 위원회 원내
T6 (29) (세계-독립)	인도네시아 네덜란드 주권 공화국 정부 이양 연방 법령 국가 규정 관계 군정 경제 독립 합중국 기관 협정 조치 행정 군사 영국 금융 군 해결 법 협조 대륙 임명 재정
T7 (30) (분단)	한국 유엔 정부 미국 대한민국 통일 주권 세계 평화 국제 원조 독립 공산 민족 북한 총회 결의 위원회 경제 자유 소련 수립 국가 국내 전쟁 남북 공산주의 협정 반대 해결

먼저 T0은 ‘헌법’, ‘국회’, ‘개헌’, ‘개헌안’과 같은 단어들을 감안하였을 때 개헌과 관련된 토픽이라 할 수 있다. 제1공화국 시기에는 대통령 직선제가 주요 골자였던 1952년 개헌(발췌 개헌), 순수 대통령제 확립과 초대 대통령에 대한 중임 제한 철폐가 주요 특징이었던 1954년 개헌(사사 오입 개헌)이 실시되었다. 이와 더불어 개헌안에 대한 논쟁이 시기 내내 지속되었다. 제헌 이래 2년이 채 지나지 않았던 1950년 1월에 대통령중심제에서 내각제로의 변경을 주요 골자로 하는 개헌안이 제출되고 공방이 벌어진 바가 있었다. 제3대 대선이 있었던 1956년 이후에는 장면의 부통령 당선으로 위기감을 느낀 자유당이 부통령제를 없애고 권력을 영속화하기 위해 내각제 개헌을 추진했다. 개헌이 실패하자 자유당은 정·부통령 조항만을 수정하는 개헌을 목표하기도 했다. 또 이 과정에서 민의원의원선거법, 국가보안법, 지방자치법 등의 개정이 이루어져 민주당의 동의 아래 진보당이 축출되고 선거와 언론의 자유가 축소되었으며, 지방자치제가 유명무실화되었다.⁴³

T1은 ‘해양’, ‘주권’, ‘주권선’, ‘어업’, ‘국제법’과 같은 단어들을 통해 대한민국의 해양주권과 관련된 토픽이라고 보았다. 이 토픽은 일본과 상당히 연관되어 있다. 이승만 정부는 일본의 어업 활동 수역을 제한한 ‘맥아더 라인’의 폐지, 그리고 한국 근해에서 빈번히 조업하여 한국 수산업을 위협한다고 인식되었던 일본 어선들에 대응하고자 했다. 이를 위해 국제법적 검토를 거쳐 1952년 ‘대한민국 인접 해양의 주권에 대한 대통령의 선언’을 선포하였다. 이로써 한반도 인근 바다에 이른바 ‘평화선’을 긋고, 선을 침범하는 일본 어선을 나포하였다. 어선의 나포는 한일 양국에서 정치적 이슈로 부상하였으며, 한일 간의 어업 협상은 1951년부터 시작된 한일회담에서 가장 주요한 의제 중 하나가 되었다.⁴⁴

T2는 ‘미국’, ‘서독’, ‘소련’, ‘프랑스’, ‘영국’, ‘NATO’, ‘EDC’와 같은 단어를 고려하였을 때 세계, 그리고 냉전과 관련 있는 토픽이었다. 아울러 T2는 특히 초기 냉전에서의 서독의 행보와 연관이 있다. 제2차 세계대전 이후 미국·영국·프랑스·소련에게 점령되었던 독일은 냉전이 시작되면서 서독과 동독으로 분리되었다. 서독은 처음에 서유럽 국가들로부터 재무장에 대한 견제를 받았으나 점차 냉전의 파트너로서의 위상이 대두되었다. 이에 서독은 1955년 주권을 회복하면서 재무장을 실시하고, 프랑스 의회가 비준하지 않아 무위로 돌아간 유럽안보공동체(EDC) 대신 북대서양조약기구(NATO)에 가입하였다. 이를 통해 냉전이 유럽이 과거의 적대 관계를 극복하고 통합하는 데 원동력을 제공했다는 사실을 알 수 있으며,⁴⁵ 당시의 한국 언론은 주권과 관련하여 이러한 서독의 변신에 크게 관심을 기울였다.

T3는 ‘국민’, ‘정치’, ‘민주주의’와 같은 단어들을 통해 정치와 관련된 토픽이라고 생각했다. T3 토픽의 비중이 매우 높았던 기사로 민주정치에 대하여 논하는 글들을 다수 발견할 수 있었다. T4는 ‘대회’, ‘언론’, ‘국가보안법’, ‘진보당’과 같은 단어들로 보아 당시의 사회상(社會相)과 연관된 토픽이라고 보았다. 당시 사회의 모습을 살펴보는 기사와 더불어 1958년 진보당 사건에 대한 기사, 주권 수호를 주장하기 위한 목적을 가진 여러 쟁기대회에 대한 기사들을 확인할 수 있었다.

T5는 ‘선거’, ‘자유당’, ‘민주당’, ‘투표’, ‘개표’와 같은 단어들로 말미암아 선거와 관련된 토픽임을 알 수 있었다. 이승만이 초대 대통령, 이시영이 초대 부통령으로 당선된 이래 제1공화국에서 선거는 정치에서 가장 중요한 요소였다. 초기에는 대다수의 국회의원이 무소속이었으며, 이승만을 지지하는 의원은 적었다. 1952년에 최초로 시행되었던 도·시·읍·면의회 선거는 이승만에 충성하는 세력을 대거 양성하여 당시 이승만과 대립하였던 국회의원들을 압박하는 수단으로 작용하기도 했다. 6·25전쟁 이후에도 변화는 계속되었다. 1954년 제3대 총선을 기점으로 정당 정치가 확립되었다. 또 개헌을 통해 권력을 영속화하려는 자유당의 행보에 야권이 반발하면서 1955년 민주당이 창당된 뒤로 선거는 기본적으로 자유당, 민주당 보수 양당이 경쟁하는 구도가 되었다. 이 과정에서 진보당은 ‘북진통일론’에 대항하여 ‘평화통일론’을 주장하는 제3의 길을 걸으며 조봉암이 대통령 후보로서 1956년 대선에서 2등을 차지하는 등 돌풍을 일으켰다. 그러나 ‘용공’에 극히 민감했던 사회 분위기 속에서 진보당 사건 등을 통해 경쟁에서 축출되었다. 진보당이 선전한 1956년 대선은 장면이 부통령으로 당선되었다는 점에서 자유당에 대한 민심의 이반이 드러나는 계기이기도 했다. 한편 1950년대 후반에 들어 이승만의 사당(私黨)으로서의 성격을 더욱 강하게 가지게 된 자유당은 국가권력을 활용하여 보수 야당인 민주당에도 심대한 정치적 탄압을 가하였다. 제1공화국의 마지막 대선이었던 1960년 대선을 앞두고서는 부정선거 음모 폭로, 3·15 부정선거와 같은 중대 이슈가 발생하였고, 이는 4·19 혁명을 통한 정권의 몰락으로 이어졌다.⁴⁶

T6는 ‘인도네시아’, ‘네덜란드’, ‘주권’, ‘이양’, ‘독립’과 같은 단어를 근거로 독립의 측면에서 세계와 관련되는 토픽이라고 판단하였다. 이는 네덜란드에 대한 인도네시아의 독립과 특히 관련이

있었다. 수백 년 동안 네덜란드의 식민지였던 인도네시아는 제2차 세계대전 종전 이후 4년 동안 독립을 위한 혁명기에 있었다. 네덜란드는 그 여왕을 중심으로 하는 네덜란드-인도네시아 연방공화국을 계획하면서도, 인도네시아의 독립 세력을 무력으로 진압하기를 기도하였다. 전쟁이 계속되는 가운데 미국과 유엔이 개입하면서 1949년 헤이그 회의를 통해 인도네시아의 독립 문제가 논의되었고, 이를 통해 인도네시아는 이리안 지역을 제외한 모든 지역에 대한 주권을 이양받으며 독립을 성취할 수 있었다.⁴⁷ 당시 대한민국의 신문들은 주권과 관련하여 이러한 인도네시아의 독립과 네덜란드의 주권 이양에 대해 크게 관심을 기울인 바 있다.

마지막으로 T7은 ‘유엔’, ‘통일’, ‘전쟁’, ‘남북’과 같은 단어들을 통해 한반도의 분단과 관련이 있는 토픽이라고 간주했다. 해방 이후 냉전 체제가 수립되면서 한반도는 38선을 통해 양단(兩斷)되었다. 6·25전쟁 이후로는 군사분계선과 북방한계선(NLL)을 통해 분단되어 현재에 이르고 있다. 담론은 헌법상 대한민국의 영토이지만 사실상 외국이었던 북한, 그리고 한반도의 통일과도 연결되었다. 이승만 정부는 전쟁 이래 휴전 반대와 북진통일을 주창했고, 이는 이승만의 국부로서의 위상을 회복하게 하면서 제1공화국의 정치 구도에 영향을 주었다. 상술한 바와 같이 진보당은 북진통일론에 대항하여 평화통일론을 주장하기도 했다.⁴⁸

정리하자면 T0(개헌), T3(정치), T4(사회상), T5(선거)는 ‘주권’의 대내적 담론으로, T1(해양주권), T2(세계-냉전), T6(세계-독립), T7(분단)은 대외적 담론으로 분류할 수 있다.

3.2 담론의 시기별 변동

그렇다면 이러한 토픽 즉 담론들은 연도별, 시기별로 어떻게 변동하였을까. <그림 4>는 연도별 토픽의 변화를 담은 그래프이며, <그림 5>는 시기별 토픽의 변화를 담은 그래프이다.

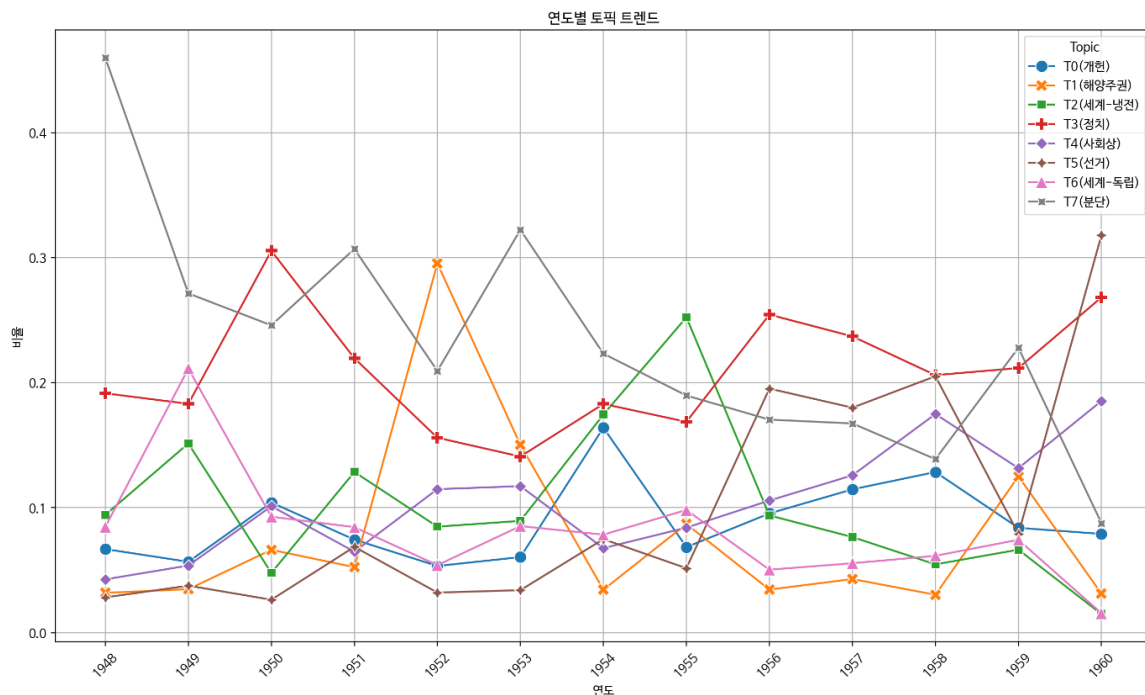


그림 4. 연도별 토픽 트렌드

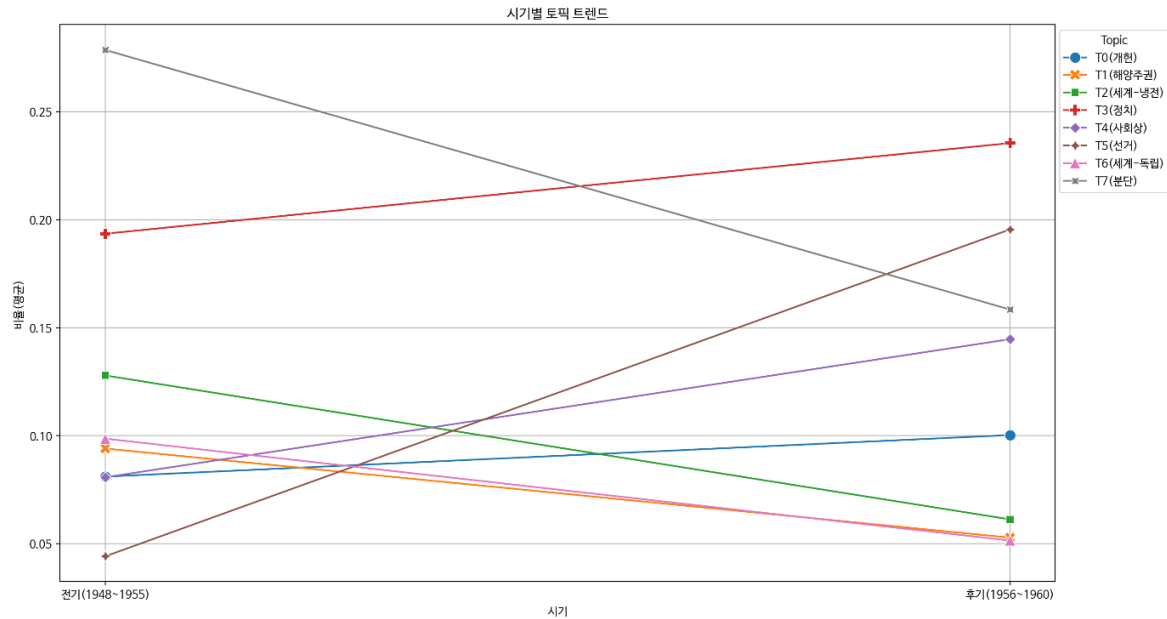


그림 5. 시기별 토픽 트렌드

〈그림 5〉의 경우 각 시기별로 연도별 비율의 평균을 계산하여 그래프를 그린 것이다.⁴⁹ 두 그래프에서 주목할 수 있는 변화는 다음과 같다. 먼저 전기와 후기를 비교하였을 때 T0(개헌), T3(정치), T4(사회상), T5(선거)와 같이 대내적 담론에 속하는 토픽의 비율은 모두 증가한 반면, T1(해양주권), T2(세계-냉전), T6(세계-독립), T7(분단)과 같이 대외적 담론에 속하는 토픽의 비율은 모두 감소하였다. 전자에서 특히 주목되는 토픽은 T5이다. 전기에서 T5의 비율은 여타 토픽과 비교하였을 때 가장 낮지만, 후기에서는 T3 다음으로 높은 모습을 보여준다. 이러한 T5 비율의 급격한 증가는 〈그림 4〉에 따르면 1956년을 기점으로 나타난다. 상술한 바와 같이 1956년에는 대선이 있었고, 이 선거는 이승만과 자유당에 대한 민심의 변화가 나타나는 선거였다. 또 1959년에 들어 그 비율이 감소하면서도, 이듬해인 1960년에 약 3배 증가하는 모습을 보이기도 한다.

표 4. 시기별 토픽의 순위

시기	토픽 순위
전기	T7 - T3 - T2 - T6 - T1 - T0 - T4 - T5
후기	T3 - T5 - T7 - T4 - T0 - T2 - T1 - T6

이와 더불어 〈표 4〉를 통해 시기별 토픽의 순위를 살펴볼 수 있다. 전기에서 대외적 담론에 속하는 토픽(빨강으로 표기)은 전반적으로 상위권에, 대내적 담론에 속하는 토픽(파랑으로 표기)은 T3를 제외하면 모두 하위권에 있으나, 후기에 들어 이러한 관계는 역전된다. 또 전기와 후기의 순위를 비교하였을 때 대내적 담론에 해당하는 토픽의 순위는 모두 증가했으며, 대외적 담론에 해당하는 토픽의 순위는 모두 감소했다. 이러한 순위의 변동은 전기에서 후기로 이행하면서 ‘주권’에 대한 대외적 담론의 비중은 감소하고 대내적 담론의 비중은 증가하였다는 사실을 뒷받침하면서도, 대외적 담론에서 대내적 담론으로 초점이 이동하였다는 사실을 보여준다.

종합하자면 제1공화국 시기 ‘주권’의 담론은 개헌, 정치, 사회상, 선거와 같은 대내적 담론과 해양, 세계, 분단과 같은 대외적 담론으로 이분될 수 있다. ‘주권’ 담론의 초점은 전기에는 대외적 담론에 맞추어져 있었으나, 후기에 들어 대내적 담론으로 이동하였다. 이는 대외적 담론의 비중 감소와 대내적 담론의 비중 증가로 이루어졌으며, 특히 선거의 담론에 관하여 변화가 집중되었다.

4. 제 1 공화국 시기 ‘主權’의 의미

4.1 전기(1948~1955 년)

앞서 토픽의 해석 및 분류, 그리고 시기별 변동을 분석하면서 거시적 차원에서 ‘주권’이 어떻게 이해되었는지를 살펴보았다. 이와 함께 제1공화국 시기 ‘주권’의 의미와 그 변화가 어떠하였는지를 살펴보기 위해서는 단어 혹은 개념에 대한 미시적 차원에서의 분석이 필요하다.

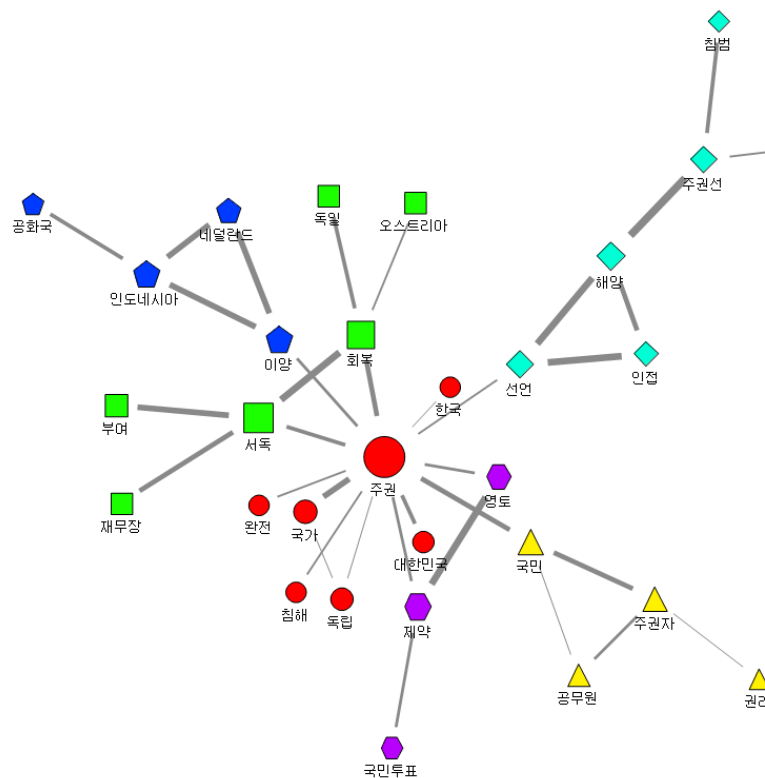


그림 6. 전기(1948~1955 년) 단어 연결망

표 5. 전기 단어 연결망 생성 정보

전기(1948~1955)	Link-Cut	Component	Fusion Level
	$w \geq 0.203$	≥ 3 (Node: 100 → 30)	9(6 Clusters)

〈그림 6〉은 전기 단어 연결망이며, 〈표 5〉에 그 생성 정보가 나열되어 있다. 〈그림 6〉에서 다음과 같은 점들을 주목할 수 있다. 먼저 3장에서 보았던 토픽들을 연결망의 클러스터를 통해 확인해 볼 수 있다. ‘해양’, ‘주권선’, ‘선언’과 같은 단어들을 통해 해양주권 클러스터(◆)는 T1과 연결되며, ‘서독’, ‘재무장’, ‘회복’ 등을 통해 세계-냉전 클러스터(■)는 T2와 연결된다. 또 정치 클러스터(▲)는 ‘국민’, ‘주권자’, ‘공무원’, ‘권리’를 통해 T3와, 세계-독립 클러스터(●)는 ‘네덜란드’, ‘인도네시아’, ‘이양’ 등을 통해 T6와 연관된다. 이렇게 클러스터들을 통해 대외적 담론들과의 연결 관계가 대두되고 있다.

이와 더불어 ‘주권’과 연결된 단어들에서도 전기의 연결망이 가지는 의의를 발견할 수 있다. 주권 클러스터(●)는 ‘주권’과 더불어 ‘한국’, ‘대한민국’과 같은 단어들이 존재한다는 점에서 대한민국의 주권에 관한 군집이라 생각된다. 여기서 ‘주권’은 단어 ‘국민’과 연결되어 있고, 다시 ‘국민’은 정치 클러스터의 ‘주권자’와 연결되어 있다. 이를 통해 제1공화국에서 ‘주권’은 국민주권의 맥락에서 인지되었다는 사실을 확인할 수 있다. 주권은 또한 ‘이양’과 ‘회복’의 대상이면서도, ‘침해’되고 ‘제약’되는 것이기도 했다. 주권 클러스터 안에 ‘침해’가 존재하는 점, 그리고 ‘주권’이 ‘제약’과 ‘회복’, ‘이양’과 연결되는 점을 감안했을 때 대한민국의 주권은 미군정으로부터 이양되어 회복되었으나 그럼에도 침해되고 제약되어 불완전한 측면이 존재하는 개념이었다. ‘주권’ - ‘영토’ - ‘제약’이 삼각관계(clique)를 형성한 것을 볼 때 그러한 불완전성은 특히 영토와 관련되어 인식되었던 것으로 풀이되며, 이는 전재성, 구선영, 그리고 나가사와 유코가 선행연구에서 지적하였던 바와 합치된다. 비슷한 맥락에서 해양주권 클러스터에서 나타나는 ‘해양’ - ‘주권선’ - ‘침범’의 연결관계를 통해 한국의 해양주권에 대한 침해와 제약의 인식도 확인해 볼 수 있다.

즉 이 시기 ‘주권’은 회복되고 이양되어 국민이 소유하게 된 권리였지만, 특히 영역의 측면에서 불완전성이 내재되고 있었다.

4.2 후기(1956~1960년)

담론에서 시기별 변동이 나타났듯이, 단어적 차원에서의 ‘주권’ 인식에도 변화가 있었다. 〈그림 7〉은 후기 단어 연결망이며, 〈표 6〉에 그 생성 정보가 제시되어 있다. 전기와 마찬가지로 후기 단어 연결망에서도 토픽과의 연결점을 찾아볼 수 있다. 해양주권 클러스터(●)를 통해 T1의 존재를 여전히 확인할 수 있는 가운데 전기의 T3가 T5로 대체되었고, 분단 클러스터(◆)를 통해 T7의 존재를 짐작해 볼 수 있다. 이를 통해 후기의 단어 연결망에서 대외적 담론을 대신하여 대내적 담론이 대두되지는 않았으며, ‘주권’의 사용은 3장에서 분석했던 담론의 변동과 연관되지는 않는다는 사실을 알 수 있다.

그럼에도 후기 단어 연결망은 주권이 ‘옹호’해야 할 대상이며, 또 이를 위해서는 ‘투쟁’해야 하는 것으로 그 인식이 변화하였다는 사실을 보여준다. 전기 연결망에도 존재하는 ‘주권’ - ‘국민’ - ‘주권자’의 연결관계로 말미암아 후기에 ‘주권’은 국민이 소유하는 권리라는 국민주권의 이념으로 인식되었다. 이와 함께 전기에서의 ‘이양’, ‘회복’, ‘침해’, ‘제약’이 사라진 대신, ‘투쟁’과 ‘옹호’가 새로이 등장하였다. ‘주권’이 ‘행사’, 그리고 ‘선거’와 연결되어 있는 점을 고려한다면 당시 국민이 가지는 주권을 행사할 수 있는 주요 방법으로 선거가 생각되었고, 그렇게 선거를 통해 행사된 주권은 곧 옹호해야 할 대상이자 투쟁으로 지켜내야 할 대상이었다.

이러한 변화는 2장에서 언급한 바가 있었던, 1956년 7월 야당계 국회의원 72인의 연합체로서

결성된 국민주권옹호투쟁위원회의 등장과 무관하지 않다. 이 위원회는 1956년 8월의 지방선거를 앞두고 경찰 등의 공권력을 동원하여 공정한 선거에 지장을 초래했던 자유당에 대항하기 위해 결성된 단체이다. 더 큰 맥락에서 이승만과 자유당의 행적 전반을 비판하고자 했던 위원회는 민주 정치와 공명선거를 촉구하고자 했다. 이를 위해 10가지 측면에서 정권의 실정을 비판, 경고하고 이를 시정할 것을 촉구하는 결의문을 작성하거나, 의원직 사퇴를 통해 그 뜻을 관철하고자 했다. 또 1957년 5월 25일 장충단에서 있었던 시국강연회를 방해한 투석, 방화 행위와 이를 사전에 단속하지 않았던 경찰을 비판하기도 했다.⁵⁰ 이후 위원회는 여전히 정치를 과행으로 이끌었던 여당의 행위를 제어하지 못하면서 총선을 앞두고 있었던 1958년 3월 24일에 해체되었다.⁵¹

정부 수립을 통해 적어도 명목상으로 온전히 국민에게 부여되었던 주권이 무결(無缺)하였다면, 굳이 주권에 대해 옹호하고 투쟁할 필요는 없었을 것이다. 즉 1956년 이후 자유당 집권의 영속화를 위한 정치적 탄압으로 인해 국민의 주권이 위태로운 상황에서 전기와 달리 ‘주권’에 능동성이 부여된 것이었다.

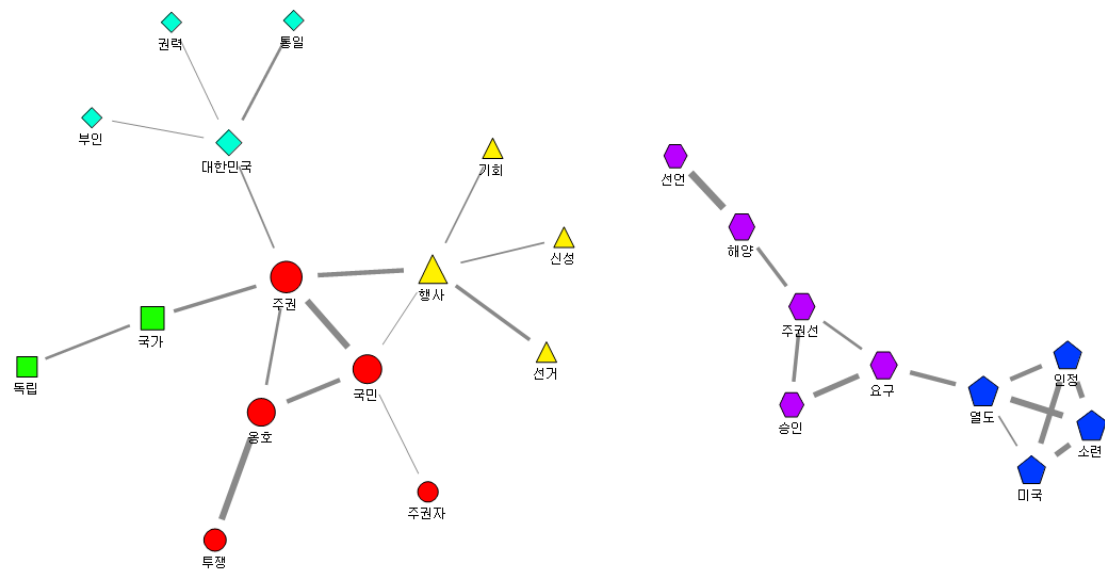


그림 7. 후기(1956~1960년) 단어 연결망

표 6. 후기 단어 연결망 생성 정보

후기(1956~1960)	Link-Cut	Component	Fusion Level
	$w \geq 0.209$	≥ 4 (Node: 100 → 24)	6(6 Clusters)

5. 맺음말

본고에서는 텍스트마이닝 방법론을 통해 대한민국 제1공화국 시기 발행된 《경향신문》, 《동아일보》, 《조선일보》 기사를 기반으로 당대 ‘주권’의 담론과 의미를 통시적 관점에서 규명하였다. 먼저 ‘주권’과 관련하여 제1공화국 시기는 1948~1955년(전기)과 1956~1960년(후기)으로 구분되었다. 이러한 시기 구분을 기반으로 토픽 모델링을 통해 ‘주권’ 담론의 양상과 그 변동을 분석하면서 거시적 차원에서 주권 개념에 접근하였으며, 단어 연결망을 통해 ‘주권’의 의미와 그 변화를 살펴보면서 미시적 차원에서 개념에 접근하였다. 토픽 모델링 결과 8개의 담론이 도출되었으며, 이는 대내적 담론과 대외적 담론으로 분류되었다. 대외적 담론의 비중 감소와 대내적 담론의 비중 증가로 대외적 담론에 있었던 전기의 초점은 후기에 대내적 담론으로 이동하였으며, 변화는 특히 선거의 담론에 집중되었다. 단어 연결망을 통해 본 ‘주권’의 의미와 그 변동은 담론의 분석에서 나타난 결과와는 독립되었다. ‘주권’의 사용과 관련하여 전기와 후기 모두 대외적 담론과의 연결관계를 찾을 수 있는 가운데, 국민주권의 맥락에서 주권이 인식되었다. 전기의 경우 ‘주권’은 이양과 회복의 대상인 동시에 침해와 제약의 대상으로 이해되었으며, 특히 영역의 측면에서 불완전성이 내재되었다. 후기의 경우 ‘주권’의 행사 수단으로 선거가 제시되었다. 또 ‘주권’은 옹호의 대상이자 투쟁으로 수호하여야 할 대상으로 변화하였다. 이는 후기에 등장한 국민주권옹호 투쟁위원회와 무관하지 않으며, 주권의 결함에 대한 전기의 인식을 공유하면서도 전기와는 달리 주권 개념에 능동성이 부여되었다는 것을 의미하였다.

이렇게 ‘주권’의 능동성은 제1공화국에 종언을 고한 4·19 혁명이 시민의 혁명이 되는 데 지대한 영향을 끼쳤으리라 생각한다. 주권은 과연 혁명 이후 무결해졌을까. 1920년부터 2019년까지 발행된 모든 《동아일보》 기사 데이터를 담은 코퍼스에서 웹 분석 도구를 통해 ‘주권’의 공기어 상위 30개 목록을 검색한 결과,⁵² 1960년대부터 현재까지 모든 시기에서 ‘침해’의 등장을 확인할 수 있었다. 이것이 대한민국의 주권에 무엇을 시사하는지는 후속 연구를 통해 보완하여야 하겠으나, 최소한 주권의 무결성 혹은 안전성에 대한 논란은 이후에도 신문에서 계속 회자되었다.

이와 더불어 ‘주권’의 유의어에 대해 고민할 필요가 있을 것이다. 본고는 ‘주권’과 그를 포함한 단어들에 한하여 연구를 진행하였지만, ‘주권’의 대표적인 유의어로 ‘국권(國權)’이 있다. 1957년 완간된 《조선말 큰사전》에 따르면 국권은 주권과 통치권을 아우르는 나라의 권력을 의미하는 단어로,⁵³ 주권과 거의 동일한 뜻을 가진다. 국권의 용례에 대하여, 1957년 10월 17일의 국회본회의에서 김달호 의원은 주한미군의 총격 사건에 대한 한국의 미온적 반응을 비판할 때 국권을 주권과 사실상 동일한 의미로 사용한 바 있다.⁵⁴ 연구 대상이었던 《경향신문》, 《동아일보》, 《조선일보》 데이터에서도 국권을 언급한 기사들이 다수 발견되기도 했다. 당시의 신문기사 등에서 활용된 주권과 국권 개념의 유사도를 양적 방법론을 활용하여 검토한 뒤 두 개념을 함께 분석한다면, 당대 주권 개념의 담론과 의미에 대해 더욱 풍부한 결과를 기대할 수 있을 것이다. 이렇게 유의어를 고려한 분석은 미래의 연구과제로 남겨두고자 한다.

¹ 구 《대한민국헌법》 제 1 조, 제 2 조(1952.7.7. 헌법 제 2 호로 개정되기 이전의 것). 현행 헌법에는 다음과 같이 규정되어 있다: 제 1 조 ① 대한민국은 민주공화국이다. ② 대한민국의 주권은 국민에게 있고 모든 권력은 국민으로부터 나온다.

² 제 3 차 유엔 총회(1948). “The problem of the independence of Korea (A/RES/195(III))”. 25. <https://digitallibrary.un.org/record/210026>; 박태균(2018). “정전협정과 종전선언”. <역사비평>. 177. <http://dx.doi.org/10.38080/crh.2018.05.123.171>

³ 박상섭(2008). <국가·주권>. 소화. 183-220. <http://lod.nl.go.kr/resource/KMO200901235>

- ⁴ 정용화(2004). “근대 한국의 주권개념의 수용과 적용”. <세계정치>. <https://www.riss.kr/link?id=A75510922>; 신옥희(2004). “근대 한국의 주권 개념”. <세계정치>. <https://www.riss.kr/link?id=A75510618>; 김현철(2005). “개화기 『만국공법』의 전래와 서구 근대주권국가의 인식 - 1880년대 개화파의 주권 개념의 수용을 중심으로 -”. <한국학>. <https://www.riss.kr/link?id=A79866058>; 全鍾杙(2006). “近代主權概念의 受容과 展開 - 1876년부터 1900년까지 開化知識人의 개혁론을 중심으로”. 서울대학교 대학원 박사학위논문. <https://doi.org/10.23170/snu.000000048604.11032.0000667>; 박상섭(2008). <국가·주권>. 소화. <http://lod.nl.go.kr/resource/KMO200901235>; 오영달(2008). “대한제국기 이승만의 서구 인권 및 주권론 수용: - 그의 『독립정신』에 나타난 정치사상을 중심으로”. <韓國民族文化>. <https://www.riss.kr/link?id=A75058018>; 강상규(2010). “동아시아 문명권에서 ‘주권’과 ‘국제’ 개념의 탄생 - 『만국공법』의 판본 비교와 번역”. <중국학보>. <https://www.riss.kr/link?id=A60303736>; 김동택(2017). “근대 한국 주권 개념 연구”. <21세기정치학회보>. <http://dx.doi.org/10.17937/topsr.27.2.201706.27>; 전종익(2017). “대한민국임시정부 이전 정치체제 구상 - 1910년대 군주제와 공화제를 중심으로”. <법사학연구>. <http://dx.doi.org/10.31778/lawhis.56.201710.219>; 왕신(2021). “한민족의 주권(主權) 인식 연구 - 개항부터 『대한민국 임시헌법』(1911.9.11.) 제정까지 -”. 한국학중앙연구원 한국학대학원 박사학위논문. <https://www.riss.kr/link?id=T15930878>; 왕신(2021). “한말 법학 교과서에 나타난 주권 개념 연구: 『헌법(憲法)』, 『법학통론(法學通論)』, 『평시 국제공법(平時國際公法)』을 중심으로”. <역사문제연구>. <http://doi.org/10.36432/CSMKH.46.202111.5>; 왕신(2022). “한일합병 이후 근대 주권 개념의 변용 및 주권 용어의 법정화(法定化)”. <한국학>. <http://doi.org/10.25024/ksq.45.3.202209.179>; 오영달(2022). “대한민국의 건국 시기에 대한 논쟁의 재조명: 이중 주권론의 시각”. <한국동양정치사상사연구>. <http://dx.doi.org/10.35161/rkapt.2022.03.21.1.145>
- ⁵ 구선영(2019). “국제관계 형성의 기점으로서의 베스트팔렌 신화 재고: 각국의 역사적 문맥에 배태된 다양한 주권 개념 형성에 관한 시론”. <국제정치논총>. 44. <http://doi.org/10.14731/kjir.2019.06.59.2.7>
- ⁶ 전재성(2017). “동북아의 불완전한 주권국가들과 복잡적 무정부상태”. <세계정치>. 98. <https://www.riss.kr/link?id=A105354062>
- ⁷ 나가사와 유코(2012). “일본 패전 후의 한반도 잔여주권(殘餘主權)과 한일 ‘분리’ - 신탁통치안 및 대일강화조약의 ‘한국포기’ 조항을 중심으로(1945~1952)”. <아세아연구>. <https://www.riss.kr/link?id=A99553344>
- ⁸ 금보운(2021). “이승만 정권기 민·관의 주한미군 재산반환 요청과 주권 인식”. <국학연구>. <http://doi.org/10.36093/ks.2021.46.013>
- ⁹ 한모니까(2013). “1948년 대한민국 정부 수립과 주한미군의 정권 이양 과정 및 의미”. <동방학지>. 306. <https://www.riss.kr/link?id=A103538878>
- ¹⁰ 허은(2006). “한국 학생운동의 ‘주권’ 인식 변화와 반미자주화 운동”. <지역과 전망>. <https://www.riss.kr/link?id=A87034514>
- ¹¹ 정무용(2015). “1961년 한미경제기술원조협정을 둘러싼 정치·사회적 갈등”. <인문과학연구>. <https://www.riss.kr/link?id=A104223485>
- ¹² 崔航航(2021). “박정희 정부의 해양주권 인식과 동아시아 ‘자주’ 외교(1965~1974) - ‘한일대륙봉협정’ 체결 과정을 중심으로”. 고려대학교 대학원 석사학위논문. <https://doi.org/10.23186/korea.000000235007.11009.0001178>
- ¹³ 근대 한국에서의 주권 인식을 다룬 주요 연구에 대해서는 미주 4의 목록을 참조할 것.
- ¹⁴ Blei, David(2012). “Probabilistic Topic Models.” *Communications of the ACM*. <http://doi.org/10.1145/2133806.2133826>
- ¹⁵ 노드의 중심성(Centrality)을 평가하는 기준으로 링크의 개수를 통해 인접 노드와의 연결성을 측정하는 연결 중심성(Degree Centrality), 얼마나 중심에 근접한지를 측정하는 근접 중심성(Closeness Centrality), 노드 집단을 서로 연결하는 핵심 고리인지를 보는 매개 중심성(Betweenness Centrality), 얼마나 중요한 노드들과 연결되어 있는지를 보는 위세 중심성(Eigenvector Centrality)이 있다.
- ¹⁶ 김동윤(2007). “1950년대 신문소설의 위상”. <대중서사연구>. 9-11. <http://dx.doi.org/10.18856/jpn.2007.17.001>
- ¹⁷ 韓元永(1999). <韓國現代新聞連載小說研究>. 國學資料院. 62. <http://lod.nl.go.kr/resource/KMO200000966>; 강준만(2019). <한국 언론사: 한성순보에서 유튜브까지>. 인물과사상사. 222, 240. <http://lod.nl.go.kr/resource/KMO201911499>
- ¹⁸ 강준만(2019). <한국 언론사: 한성순보에서 유튜브까지>. 인물과사상사. 207-247. <http://lod.nl.go.kr/resource/KMO201911499>
- ¹⁹ 韓元永(1999). <韓國現代新聞連載小說研究>. 國學資料院. 60.

<http://lod.nl.go.kr/resource/KMO200000966>

²⁰ 보다 구체적으로 서술하자면, 연구는 아카이브들을 통해 구축된 빅데이터에서 시작되었다. 《경향신문》의 경우, 앞서 서술한 세 신문에서 1945 년 11 월부터 1960 년 6 월까지 발행한 주권 관련 기사로 구성된 데이터셋(Dataset)에서 필요한 데이터를 추출하였다. 《조선일보》와 《동아일보》는 1920~1999 년 발행된 전체 기사의 데이터셋을 활용하였다. 전자의 데이터셋은 김진규(서울대 국사학과 박사수료)가 제작하였으며, 후자의 데이터셋은 김병준(한국학중앙연구원 한국학대학원 인문정보학 전공 조교수)이 구축하였다. 아울러 전자는 직접, 후자는 교신저자로부터 제공받았다. 다만 신문들이 가지는 논조와 성향으로 인해 객관적 시각에서 당시의 시대상을 살피는 데 지장이 있을 수 있다. 실제로 당시의 시대상을 가장 잘 반영하고 있는 신문은 《서울신문》과 《한국일보》라고 일반적으로 평가되고 있지만, 데이터 확보에 어려움이 있어 두 신문에 대한 분석은 차후의 과제로 남겨두기로 한다.

²¹ Cyram. “NetMiner 4.5.1.d”. <https://www.netminer.com>

²² 변수민. “hanja”. <https://github.com/suminb/hanja>

²³ 이민철(2024). “Kiwi: 통계적 언어 모델과 Skip-Bigram 을 이용한 한국어 형태소 분석기 구현”. <디지털인문학>. <http://doi.org/10.23287/KJDH.2024.1.1.6>

²⁴ 공기어(共起語)는 특정 문맥에서 어떠한 단어와 유의미하게 같이 출현하는 단어를 말한다. 공기어의 문맥은 연구 목적에 따라 그 범위가 다르게 설정될 수 있다. 예를 들어 단어의 공기 관계를 정밀히 포착하고자 할 경우 공기어의 문맥을 3~4 어절 정도로 좁힐 수 있다. 단어의 의미 포착에서 벗어나 화제적 차원에서 분석하고자 할 경우 문맥 범위를 넓혀 문단 단위로 설정할 수도 있다(김일환(2019). “인문학을 위한 신문 빅 데이터와 텍스트 마이닝”. <語文論集>. 49-50. <http://dx.doi.org/10.15565/jll.2019.06.78.41>). 아울러 문맥의 설정이 반드시 어절 기준으로 국한되는 것은 아니며, 문장, 음절과 같은 다양한 기준을 채택할 수 있다. 예를 들어 이윤원은 20 세기 초반 한국 사회가 근대적 혁명 개념을 수용하는 과정 속에서 드러났던 ‘혁명’의 발화 맥락과 그 변화를 규명하기 위해, 당대의 언론 매체 데이터에서 공백을 제외한 ‘혁명’의 앞뒤 각 20 음절을 문맥으로 간주하여 그 범위 안에서 명사와 고유명사를 위주로 공기어를 선별한 바 있다(이윤원(2023). “20 세기 초반 한국 사회의 ‘혁명’ 개념 수용 과정 - 언론 매체의 공기어 연결망 분석을 중심으로 -”. 서울대학교 대학원 석사학위논문. 22-23. <https://www.riss.kr/link?id=T16748783>).

²⁵ 데이터셋의 기사 제목 및 본문은 제대로 문장화되어 있지 않아 어절이나 문장을 기준으로 문맥을 설정하여 공기어를 선별하는 것은 무리가 있었다. 상술한 코퍼스는 제목과 본문에 사용된 모든 단어가 각 기사별로 정리되어 있었기에, 이를 기반으로 문맥을 설정했다.

²⁶ 기본적으로 전후(前後) 두 단어를 포함하였다. 다만 앞 혹은 뒤에 있는 단어의 개수가 충분치 않을 경우, 다른 방향에서 단어를 더 수집하였다. 예를 들어 어떤 한 문맥(문서)에서 ‘주권’ 앞에 단어가 존재하지 않았을 경우, ‘주권’ 바로 뒤에 있는 단어 4 개를 수집하였다. 이렇게 하나의 문서당 ‘주권’ 관련 단어를 포함하여 5 개의 단어가 존재하도록 하였다. 기사의 제목 및 본문이 극도로 짧아 전체 단어 수가 5 개 이하인 경우에는, 기사의 단어를 모두 수집하여 하나의 문서로 간주하였다.

²⁷ 이민철. “tomotopy”. <http://doi.org/10.5281/zenodo.6868418>

²⁸ 許洙, 金惠珍, 鄭瑜炅(2022). “대한제국기 ‘집단적 주체’의 의미망 - 『황성신문』과 『대한매일신보』의 사설 기사를 중심으로”. <대동문화연구>. 257. <http://doi.org/10.18219/ddmh..119.202209.245>

²⁹ Blei, David; Ng, Andrew; Jordan, Michael(2003). “Latent Dirichlet Allocation.” *Journal of Machine Learning Research*. 1008. <https://dl.acm.org/doi/10.5555/944919.944937>

³⁰ Blei, David(2012). “Probabilistic Topic Models.” *Communications of the ACM*. 83. <http://doi.org/10.1145/2133806.2133826>

³¹ Chang, Jonathan; Boyd-Graber, Jordan; Gerrish, Sean; et al. (2009). “Reading Tea Leaves: How Humans Interpret Topic Models.” *Advances in Neural Information Processing Systems*. <https://dl.acm.org/doi/10.5555/2984093.2984126>

³² 許洙, 金惠珍, 鄭瑜炅(2022). “대한제국기 ‘집단적 주체’의 의미망 - 『황성신문』과 『대한매일신보』의 사설 기사를 중심으로”. <대동문화연구>. 257. <http://doi.org/10.18219/ddmh..119.202209.245>

³³ 코사인 유사도는 두 벡터(vector) 간의 코사인 각도를 통해 서로의 유사도를 계산하는 방법이다. -1~1 의 값을 가지며, 1 에 가까울수록 높은 유사성을 가진다. 각(θ)을 통해 그 관계를 계산하기에 객체가 되는 문서의 길이로부터 비교적 자유롭다는 특징이 있다(유원준, 안상준. “05-01 코사인 유사도(Cosine Similarity) - 딥 러닝을 이용한 자연어 처리 입문”. <https://wikidocs.net/24603>). 문서뿐만 아니라 공기어 간의 관계를 파악할 수도 있는데, 자세한 방법은 이수상(2014). “언어 네트워크 분석 방법을 활용한 학술논문의 내용분석”. <정보관리학회지>. 57. <http://doi.org/10.3743/KOSIM.2014.31.4.049> 를 참조할 것.

³⁴ 유사도를 내림차순으로 정렬하여 순위에 따른 그래프를 그린 뒤, Elbow Point 후보 범위 중 $w = 0.20$ 주변을 선정하여 유사도가 급격히 하락하기 직전의 지점을 적절히 선택하였다.

- ³⁵ Component 는 연결의 끊김이 없는 노드들의 집단을 말하며, Community 는 노드끼리 응집된 정도에 따라 구분되는 Component 의 하부단위를 뜻한다(허수(2016). “네트워크분석을 통해 본 1980년대 ‘민중’ - 『동아일보』의 용례를 중심으로 -”. <개념과 소통>. 70. <http://dx.doi.org/10.15797/concom.2016..18.002>).
- ³⁶ 각 노드의 연결 중심성은 0~1 사이의 값으로 정규화되어 있어 적절한 표현을 위해 특정 범위 내의 값으로 다시 설정할 필요가 있었다. 노드의 크기를 연결 중심성에 비례하게 조정하기 위해, 각 노드의 연결 중심성에 200 을 곱하고 20 을 더한 뒤 소수 첫째 자리에서 반올림하였다. 일반화된 식은 다음을 참조할 것. Adler, Jonas. “How do you de-normalise?”. <https://stackoverflow.com/questions/45257856/how-do-you-de-normalise>
- ³⁷ 許洙, 金惠珍, 鄭瑜炅(2022). “대한제국기 ‘집단적 주체’의 의미망 - 『황성신문』과 『대한매일신보』의 사설 기사를 중심으로”. <대동문화연구>. 260. <http://doi.org/10.18219/ddmh..119.202209.245>
- ³⁸ 피어슨 상관계수는 두 변수 간의 선형(linear) 상관관계를 나타내는 숫자이며, -1~1 의 값을 가진다. 1 에 가까울수록 양의 상관관계를 강하게 나타내며, -1 에 가까울수록 음의 상관관계를 강하게 나타낸다. 0 혹은 그에 근접한 값은 두 변수 간의 상관관계가 존재하지 않는다는 것을 의미한다.
- ³⁹ Pedregosa, Fabian; Varoquaux, Gaël; Gramfort, Alexandre; et al. (2011). “Scikit-learn: Machine Learning in Python.” *Journal of Machine Learning Research*. <http://dl.acm.org/doi/10.5555/1953048.2078195>
- ⁴⁰ TF-IDF 에서 TF(Term Frequency)는 문서 내 단어의 빈도를 나타내는 값이다. IDF(Inverse Document Frequency)는 기본적으로 (단어가 등장하는 문서 수 ÷ 전체 문서 수)의 역수에 상용로그를 취한 값이다. TF-IDF 는 TF 와 IDF 를 곱한 것으로, 점수가 높을수록 전체 문서 중 특정 문서에서만 등장하는 중요한 단어라는 의미를 가진다. 한편 IDF 의 계산 불능 문제 등을 방지하기 위한 여러 변형 공식이 존재하고 있는데, 이 글에서는 ‘scikit-learn’ 라이브러리에서 제공하는 공식을 사용하였다.
- ⁴¹ Virtanen, Pauli; Gommers, Ralf; Oliphant, Travis E.; et al. (2020). “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python.” *Nature Methods*. <http://doi.org/10.1038/s41592-019-0686-2>
- ⁴² 이 글에 관한 연구데이터는 <https://github.com/intazero/Sovereignty-1stROK> 에서 확인할 수 있다.
- ⁴³ 서희경(2020). <한국헌정사, 1948-1987>. 포럼. 131-348. <http://lod.nl.go.kr/resource/KMO202040246>
- ⁴⁴ 조윤수(2008). “‘평화선’과 한일 어업 협상 - 이승만 정권기의 해양질서를 둘러싼 한일간의 마찰 -”. <일본연구논총>. <http://dx.doi.org/10.35368/kjjs.2008..28.008>
- ⁴⁵ 황영주, 이승근(2005). “초기 유럽통합과정과 냉전의 영향 - NATO 의 성립과 EDC 의 실패를 중심으로 -”. <대한정치학회보>. <https://www.riss.kr/link?id=A82606386>
- ⁴⁶ 홍석률, 박태균, 정창현(2018). <한국현대사 2: 경제성장과 민주주의, 그리고 통일의 과제>. 푸른역사. 19-40. <https://www.riss.kr/link?id=M14926680>; 서희경(2020). <한국헌정사, 1948-1987>. 포럼. 348. <http://lod.nl.go.kr/resource/KMO202040246>
- ⁴⁷ 양승윤(2010). <인도네시아사>. 한국외국어대학교출판부. 359-402. <http://lod.nl.go.kr/resource/KMO201115830>
- ⁴⁸ 홍석률, 박태균, 정창현(2018). <한국현대사 2: 경제성장과 민주주의, 그리고 통일의 과제>. 푸른역사. 27-29, 32. <https://www.riss.kr/link?id=M14926680>
- ⁴⁹ 許洙, 金惠珍, 鄭瑜炅(2022) 262 쪽의 방법을 참조하였다.
- ⁵⁰ 民主黨 편(1957). <鬭爭의 足跡>. 民主黨. 78-81, 108-110, 124-126, 134-135. <https://lod.nl.go.kr/page/CNTS-00121177292>
- ⁵¹ 동아일보. “『不當한 一方的散會』主權鬭委 閉院強行에 非難聲明 (1958.3.26.)”. <https://newslibrary.naver.com/viewer/index.naver?articleId=1958032600209101011&editNo=1&printCount=1&publishDate=1958-03-26&officeId=00020&pageNo=1&printNo=10947&publishType=00010>
- ⁵² 최재웅, 이도길(2014). “물결 21 코퍼스: 공개 웹 자원 및 활용 도구”. <민족문화연구>. <http://dx.doi.org/10.17948/kcs.2014..64.3>; 분석 도구는 고려대학교 민족문화연구원 디지털인문학센터에서 제공하는 《동아일보 코퍼스 분석 도구》(<https://corpus.korea.ac.kr/donga>)를 활용하였다.
- ⁵³ 한글학회(1957). <조선말 큰사전>. 을유문화사. 419, 2785. <https://hangeul.or.kr/homepage/custom/koreanbig>; 사전에서 제시하는 주권과 국권의 정의는 다음과 같다. ‘주권(主權) [이] ① 가장 중요한 권리. ② (법) 나라를 구성하는 요소인 가장 높고 독립한 권력. 국권(國權) [이] (정) 나라의 권력. 주권(主權)과 통치권.’
- ⁵⁴ 국회도서관. “제 3 대 국회 제 26 회 제 17 차 국회본회의 1957 년 10 월 17 일(목)”. <https://dataset.nanet.go.kr>; 관련 발언은 다음과 같다. ‘뿐만 아니라 오늘날 우리 국민은 이와 같은 외국인 미군 범죄자가 우리의 애국적인 국민을 함부로 살상하는 데 대해서 우리 경찰 내지 검찰국에서 즉각 체포하기를…… 뿐만 아니라 체포해서 처분하기를 희망하고 있는 것입니다. 이 점에 대해서 과거 재정협정이라는 것이 있어 가지고 공무집행에 한해서 행해진 문제 기타 무어라고 무어라고 조건을 들어

가지고 현행범 내지 또 살인자에 대해서도 우리 국권은 체포할 수 없다 하는 식의 협정이 되어 있는 것입니다. (중략) 우리는 우리의 국권의 행사에 대해서 여기에 와 있는 미군인들이 우리 국법의 준수에 협조해야 하는 것입니다. 협조해야 할 주체가 어디에 있는 것입니까? 우리나라 주권행사에 대해서 우리를 원조해 줄려고 온 그 군대는 전체적으로 우리 국법의 수행에 대해서 협조해 줄 의무가 있는 것입니다. (중략) 뿐만 아니라 어떻게 해야지 우리가 약소국가로서 독립…… 완전독립을 지향하는 마당에서 우리의 이 국권 우리의 이 주권의 행사를 좀 더 국내외적으로 보장할 수 있느냐 하는 것을 우리가 깊이 생각해야 하는 것입니다.’ (밑줄은 필자 표기)

참고문헌

- 자료

경향신문, 동아일보, 조선일보

국회도서관. “제 3 대 국회 제 26 회 제 17 차 국회본회의 1957 년 10 월 17 일(목)”. <https://dataset.nanet.go.kr>

대한민국헌법

民主黨 편(1957). <鬭爭의 足跡>. 民主黨. <https://lod.nl.go.kr/page/CNTS-00121177292>

제 3 차 유엔 총회(1948). “The problem of the independence of Korea (A/RES/195(III))”.
<https://digitallibrary.un.org/record/210026>

한글학회(1957). <조선말 큰사전>. 을유문화사. <https://hangeul.or.kr/homepage/custom/koreanbig>

- 단행본

강준만(2019). <한국 언론사: 한성순보에서 유튜브까지>. 인물과사상사.
<http://lod.nl.go.kr/resource/KMO201911499>

박상섭(2008). <국가·주권>. 소화. <http://lod.nl.go.kr/resource/KMO200901235>

서희경(2020). <한국헌정사, 1948-1987>. 포럼. <http://lod.nl.go.kr/resource/KMO202040246>

양승윤(2010). <인도네시아사>. 한국외국어대학교출판부. <http://lod.nl.go.kr/resource/KMO201115830>

韓元永(1999). <韓國現代新聞連載小說研究>. 國學資料院. <http://lod.nl.go.kr/resource/KMO200000966>

홍석률, 박태균, 정창현(2018). <한국현대사 2: 경제성장과 민주주의, 그리고 통일의 과제>. 푸른역사.
<https://www.riss.kr/link?id=M14926680>

- 논문

강상규(2010). “동아시아 문명권에서 ‘주권’ 과 ‘국제’ 개념의 탄생 - 『만국공법』의 판본 비교와 번역”.
<중국학보>. <https://www.riss.kr/link?id=A60303736>

구선영(2019). “국제관계 형성의 기점으로서의 베스트팔렌 신화 재고: 각국의 역사적 문맥에 배태된 다양한 주권 개념 형성에 관한 시론”. <국제정치논총>. <http://doi.org/10.14731/kjir.2019.06.59.2.7>

김보운(2021). “이승만 정권기 민·관의 주한미군 재산반환 요청과 주권 인식”. <국학연구>. <http://doi.org/10.36093/ks.2021..46.013>

김동윤(2007). “1950년대 신문소설의 위상”. <대중서사연구>. <http://dx.doi.org/10.18856/jpn.2007..17.001>

김동택(2017). “근대 한국 주권 개념 연구”. <21세기정치학회보>. <http://dx.doi.org/10.17937/topsr.27.2.201706.27>

김일환(2019). “인문학을 위한 신문 빅 데이터와 텍스트 마이닝”. <語文論集>. <http://dx.doi.org/10.15565/jll.2019.06.78.41>

김현철(2005). “개화기 『만국공법』의 전래와 서구 근대주권국가의 인식 - 1880년대 개화과의 주권 개념의 수용을 중심으로 -”. <한국학>. <https://www.riss.kr/link?id=A79866058>

나가사와 유코(2012). “일본 패전 후의 한반도 잔여주권(殘餘主權)과 한일 ‘분리’ - 신탁통치안 및 대일강화조약의 ‘한국포기’ 조항을 중심으로(1945~1952)”. <아세아연구>. <https://www.riss.kr/link?id=A99553344>

박태균(2018). “정전협정과 종전선언”. <역사비평>. <http://dx.doi.org/10.38080/crh.2018.05.123.171>

신옥희(2004). “근대 한국의 주권 개념”. <세계정치>. <https://www.riss.kr/link?id=A75510618>

오영달(2008). “대한제국기 이승만의 서구 인권 및 주권론 수용: - 그의 『독립정신』에 나타난 정치사상을 중심으로”. <韓國民族文化>. <https://www.riss.kr/link?id=A75058018>

오영달(2022). “대한민국의 건국 시기에 대한 논쟁의 재조명: 이중 주권론의 시각”. <한국동양정치사상사연구>. <http://dx.doi.org/10.35161/rkapt.2022.03.21.1.145>

왕신(2021). “한민족의 주권(主權) 인식 연구 - 개항부터 『대한민국 임시헌법』(1911.9.11.) 제정까지 -”. 한국학중앙연구원 한국학대학원 박사학위논문. <https://www.riss.kr/link?id=T15930878>

왕신(2021). “한말 법학 교과서에 나타난 주권 개념 연구: 『헌법(憲法)』, 『법학통론(法學通論)』, 『평시 국제공법(平時國際公法)』을 중심으로”. <역사문제연구>. <http://doi.org/10.36432/CSMKH.46.202111.5>

왕신(2022). “한일합병 이후 근대 주권 개념의 변용 및 주권 용어의 법정화(法定化)”. <한국학>. <http://doi.org/10.25024/ksq.45.3.202209.179>

이민철(2024). “Kiwi: 통계적 언어 모델과 Skip-Bigram 을 이용한 한국어 형태소 분석기 구현”. <디지털인문학>. <http://doi.org/10.23287/KJDH.2024.1.1.6>

이수상(2014). “언어 네트워크 분석 방법을 활용한 학술논문의 내용분석”. <정보관리학회지>. <http://doi.org/10.3743/KOSIM.2014.31.4.049>

이운원(2023). “20 세기 초반 한국 사회의 ‘혁명’ 개념 수용 과정 - 언론 매체의 공기에 연결망 분석을 중심으로 -”. 서울대학교 대학원 석사학위논문. <https://www.riss.kr/link?id=T16748783>

전재성(2017). “동북아의 불완전한 주권국가들과 복합적 무정부상태”. <세계정치>. <https://www.riss.kr/link?id=A105354062>

全鍾杙(2006). “近代主權概念의 受容과 展開 - 1876 년부터 1900 년까지 開化知識人의 개혁론을 중심으로”. 서울대학교 대학원 박사학위논문. <https://doi.org/10.23170/snu.000000048604.11032.0000667>

전종익(2017). “대한민국임시정부 이전 정치체제 구상 - 1910 년대 군주제와 공화제를 중심으로”. <법사학연구>. <http://dx.doi.org/10.31778/lawhis..56.201710.219>

정무용(2015). “1961 년 한미경제기술원조협정을 둘러싼 정치·사회적 갈등”. <인문과학연구>. <https://www.riss.kr/link?id=A104223485>

정용화(2004). “근대 한국의 주권개념의 수용과 적용”. <세계정치>. <https://www.riss.kr/link?id=A75510922>

조윤수(2008). “‘평화선’과 한일 어업 협상 - 이승만 정권기의 해양질서를 둘러싼 한일간의 마찰 -”. <일본연구논총>. <http://dx.doi.org/10.35368/kjjs.2008..28.008>

최재웅, 이도길(2014). “물결 21 코퍼스: 공개 웹 자원 및 활용 도구”. <민족문화연구>. <http://dx.doi.org/10.17948/kcs.2014..64.3>

崔航航(2021). “박정희 정부의 해양주권 인식과 동아시아 ‘자주’ 외교(1965~1974) - ‘한일대륙붕협정’ 체결 과정을 중심으로”. 고려대학교 대학원 석사학위논문. <https://doi.org/10.23186/korea.000000235007.11009.0001178>

한모니까(2013). “1948 년 대한민국 정부 수립과 주한미군의 정권 이양 과정 및 의미”. <동방학지>. <https://www.riss.kr/link?id=A103538878>

허수(2016). “네트워크분석을 통해 본 1980 년대 ‘민중’ - 『동아일보』의 용례를 중심으로 -”. <개념과 소통>. <http://dx.doi.org/10.15797/concom.2016..18.002>

許洙, 金惠珍, 鄭瑜炅(2022). “대한제국기 ‘집단적 주체’의 의미망 - 『황성신문』과 『대한매일신보』의 사설 기사를 중심으로”. <대동문화연구>. <http://doi.org/10.18219/ddmh..119.202209.245>

허은(2006). “한국 학생운동의 ‘주권’ 인식 변화와 반미자주화 운동”. <기억과 전망>. <https://www.riss.kr/link?id=A87034514>

황영주, 이승근(2005). “초기 유럽통합과정과 냉전의 영향 - NATO 의 성립과 EDC 의 실패를 중심으로 -”. <대한정치학회보>. <https://www.riss.kr/link?id=A82606386>

Blei, David; Ng, Andrew; Jordan, Michael(2003). "Latent Dirichlet Allocation." *Journal of Machine Learning Research*.
<https://dl.acm.org/doi/10.5555/944919.944937>

Blei, David(2012). "Probabilistic Topic Models." *Communications of the ACM*. <http://doi.org/10.1145/2133806.2133826>

Chang, Jonathan; Boyd-Graber, Jordan; Gerrish, Sean; et al. (2009). "Reading Tea Leaves: How Humans Interpret Topic Models." *Advances in Neural Information Processing Systems*. <https://dl.acm.org/doi/10.5555/2984093.2984126>

Pedregosa, Fabian; Varoquaux, Gaël; Gramfort, Alexandre; et al. (2011). "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research*. <http://dl.acm.org/doi/10.5555/1953048.2078195>

Virtanen, Pauli; Gommers, Ralf; Oliphant, Travis E.; et al. (2020). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python." *Nature Methods*. <http://doi.org/10.1038/s41592-019-0686-2>

- 웹 자원

변수민. "hanja". <https://github.com/suminb/hanja>

유원준, 안상준. "05-01 코사인 유사도(Cosine Similarity) - 딥 러닝을 이용한 자연어 처리 입문".
<https://wikidocs.net/24603>

이민철. "tomotopy". <http://doi.org/10.5281/zenodo.6868418>

Adler, Jonas. "How do you de-normalise?". <https://stackoverflow.com/questions/45257856/how-do-you-de-normalise>

Cyram. "NetMiner 4.5.1.d". <https://www.netminer.com>