

LDA 토픽모델링을 활용한 역대 대통령 취임사에 내재된 정책기조 분석

박 태 정*

서울사이버대학교 사회복지전공

본 연구는 역대 대통령 취임사에 내재한 정책 기조의 공통 주제를 찾는 것을 목적으로 하였다. 초대 이후 20대에 이르는 대통령들의 취임사 전체를 분석 대상으로 적용하여 파이선 프로 그래밍에 기반한 TF, TF-IDF 분석, 단어네트워크 분석, LDA 토픽모델링을 실시하였다. 분석 결과, 대통령의 언어에서 민족성과 역사성이 반복되어 나타났고, 경제, 문화, 통일과 국민의 삶과 직결된 생활에 대한 강조가 두드러져 있었다. 핵심 연구 문제인 정책 기조의 공통주제에 대한 LDA 토픽모델링 분석 결과 세계 경제와 국가발전, 자유민주주의와 개혁, 민족 역사성과 정치, 시민 생활과 복지로 확인되었다. 이러한 정책 기조는 국민 욕구로부터 발현되고 있어서, 역대 정부별 정책 성과 평가에 적용하여 후속 연구가 이어져야 함을 제안하였다.

주요어: 대통령 취임사, LDA 토픽모델링, 단어네트워크분석, 정책기조, 파이선

* 주저자: 박태정/서울사이버대학교 사회복지대학 교수/서울시 강북구 솔매로 49길
/Tel: 02-944-5019/ E-mail: jfkj01@hanmail.net

I. 서론

국민이 행복한 국가를 만들기 위한 노력은 특정 정부의 전유물이 아니다. 국민의 인간다운 삶과 기본권을 국가가 제도로서 보장하는 게 복지국가라고 할 때 이를 실행하는 수단은 정책이다. 정부의 정책 가치와 대통령이 지향하는 가치는 정책에 반영되어 실시되며, 이는 국민의 삶에 대해 크고 작은 영향을 미친다. 그런 만큼 「헌법」 제66조에서 정한 국가 원수를 선출하는 대통령 선출과정에서도 이념과 정파는 다르지만, 모든 후보자가 한 목소리로 자신의 정책 기조와 가치가 한국 사회와 국민에게 가장 잘 부합한다고 주장하는 모습을 종종 볼 수 있다.

일제 강점기를 지나 한국 정부가 수립된 이후 지금까지 여·야를 넘나드는 정권교체가 있었지만, 새로운 정부가 시작되는 때 시기마다 이른바 ‘정책기조’는 등장했다. 그렇다면 역대 정부들의 정책가치는 과연 무엇을 지향하였을까? 가령 ‘가장 중요한 정책기조는 무엇이었을까?’라는 질문 역시 이러한 맥락에서 나타날 수 있는 궁금증이다. 이 연구는 그 실마리를 대통령 취임사로부터 풀어내 보고자 한다.

대통령이 정책을 입안하고 실행하는 데 미치는 영향은 매우 크고 그 결과는 국민의 삶과 직결되어 있다. 대통령들의 정치적 가치 지향성은 국내·외 각종 행사에서 발표되는 연설을 통해 가장 보편적이고 명확하게 드러나는 경향이 있다(백자성·윤지원, 2022). 정부 정책 방향이 제시되는 통로는 다양하지만, 무엇보다 대통령 당사자의 연설은 가장 강력한 수단이며 그 가운데 취임사는 최대한의 파급 효과가 있다(박준형 외, 2017). 그러므로 취임사는 국가 비전을 제시하고 대통령의 정치철학, 정책 기조와 방향을 국민에게 전달할 수 있는 대표적인 수단으로 여겨진다. 취임사는 대통령의 통치이념과 새 정부 국정운영의 방향이 드러난 공적 자료일 뿐만 아니라 나라와 국민을 위해 미래에 어떠한 일을 할 것인지를 밝히는 일종의 정책공약이다(윤우제 외,

2011). 따라서 대통령 취임사는 대통령 개인의 성향을 이해하는 데에 그치지 않고, 해당 시대의 정치적 인 상황과 국가의 방향성을 파악하는 데에도 유용하다.

물론 역사적으로 역대 대통령들이 추진한 정책의 방향성이 일관적이지만은 않았으며, 대통령의 이념적 경향 및 리더십 스타일에 따라 여러 가지 방향으로 표출되는 것은 분명하다(김혁, 2014). 하지만 매 정부 출범을 앞두고 대통령이 구상하는 국가와 정책에 대한 비전은 취임사에 고스란히 반영된다는 점에서, 축적된 취임사로부터 확인되는 주제들은 단지 대통령 개인의 발화 의도라는 차원을 넘어 시대를 아우르는 고민과 욕구를 확인하는 것에도 적용할 수 있다.

이러한 점에 착안하여, 이 논문에서는 한국 정부가 수립된 후 20대 대통령으로 이어져 오는 동안 발표된 취임사에 드러난 공통 정책 기조를 찾아보고자 한다.

그간 대통령선거 혹은 각 정부의 성격에 관한 연구는 비교적 활발하게 진행됐다. 최근 발표된 연구만 하더라도 선거와 관련된 프레임 분석하거나 흔히 표심으로 표현되는 유권자 투표 성향, 투표율 분석과 연결되는 연구(이수범·박희영, 2022; 구분상, 2021; 이준한·황원재, 2021; 김성연, 2022), 지역과 진영을 주제로 한 연구(가상준, 2022; 김재한, 2014) 등을 들 수 있다. 그러나 선거를 통해 출범한 정부의 가치 표상이나 정책기조에 관한 연구는 상대적으로 적은 편이다. 후보자, 여야정당 간 복지정책과 공약을 비교하거나(김현정, 2022), 국정 목표 전반에 대한 분석(전일욱, 2021), 국민의 복지 인식 변화를 살펴본 연구(박종구·이보라, 2015) 등의 연구들이 발표되고 있지만, 시계열적 관점에서 역대 정부를 관통하는 정책 기조가 무엇이며 이것이 시사하는 바를 분석한 연구 역시 미미하다.

그러므로 이번 연구는 초대 대통령 이후 지금까지 13명의 대통령 취임사에 내재된 정책 기조 분석

을 처음으로 시도한다는 점에서 의미가 있다. 나아가 이를 통해 도출된 주제들은 한국 사회가 안고 있는 고민과 욕구를 이해하는 데 도움이 되며, 그와 관련된 정책평가 등을 주제로 하는 후속 연구로도 이어질 수 있을 것이다

II. 이론적 배경

1. 정책기조와 대통령 취임사

정책기조란 정책의 방향, 내용, 성격, 과정 등을 규정해 주는 사고 정향, 철학, 가치 등을 포괄하는 개념이다. 우리나라는 대통령제를 채택하고 있는 만큼 각 정부에 따라 설정되는 정책 기조는 특정한 방향으로 고정된다기보다 해당 임기에 제약받기 쉽다는 점에서 정부별로 차이가 나타난다(임주영·박형준, 2015). 정책의 정통성과 지속성 유지라는 점에서 정책기조가 지나치게 자주 바뀌는 것은 바람직하지 않겠으나 정부별로 조금씩 달라질 것임은 분명하다.

이렇게 볼 때 국가 운영에 대한 정책 기조 형성의 주체는 그러한 기조를 정책의 형성, 집행, 평가에 반영하고자 하는 정책결정권자이다. 따라서 대통령을 중심으로 한 정부 조직이 변화될 때 정책 기조의 형성 주체는 대통령 당선인이며 대통령직 인수과정에서 각 장관 내정자와 함께 정책 기조가 만들어진다(박정택, 2007).

이렇게 수립된 정책 기조가 가장 먼저 발화되는 시점이 취임사이다. 당선 후 첫 공식 연설인 취임사는 국가와 국민, 사회구조를 둘러싼 역사적 인식, 국내 및 국제정치 상황과 관련한 서사에 기초하여 임기를 통해 달성하려는 정책의 의도와 목표, 가치를 제시하는 역할을 한다. 또한 취임사는 국가의 정책 방향이자 구성원들에게 비전을 제시하는 최고 리더의 말이며 시대와 역사인식을 반영한 실행과제

를 피력하는 복합적 기능이 있다(임순미, 2009).

취임사에 드러난 정책방향, 특징 등을 다룬 선행 연구를 보면 우선 전일욱(2022)은 대통령 취임사를 통해 진보정부와 보수정부의 국정목표를 비교하였는데 보수정부는 경제목표를 우선시한 반면 진보정부는 사회목표를 강조한 것으로 제시하였다. 강태완·김은정(2004)은 박정희 대통령부터 노무현 대통령까지 취임연설문을 분석하여 그 수사적 특징과 역할 규정을 살펴본 바 있다. 3공화국을 대상으로 5대부터 7대까지 대통령 취임사를 분석하여 제5대에서는 정치목표를 제6대에서 경제목표, 제7대의 경우 사회·후생목표로 그 강조점이 변화되었음을 보여준 연구도 수행되었다(전일욱, 이병관, 2021). 임순미(2009)는 김대중 대통령과 이명박 대통령의 취임사 분석을 통해 개인적 가치 프레임과 시대적 배경에 대한 해석을 통해 정치적 상징 이미지를 찾는 한편 대국민 설득과정의 공통점과 차이점을 제시하기도 하였다. 한국과 미국 대통령 취임사에 대한 비교분석도 시도되었는데 임성호(2018)는 버락 오바마, 도널드 트럼프, 문재인 대통령을 대상으로 취임사에서 추출된 지도자 이미지와 그 속에 담긴 가치의 강조점, 공통점, 차이점 등을 분석한 바 있다.

한편 해외에서는 미국을 중심으로 대통령 취임사를 대상으로 다양한 분석이 시도되고 있는데 Morel(2015)은 링컨 대통령의 두 번째 재취임 당시 연설에서 남북전쟁의 승전과 노예해방이라는 중요 사건을 국가 재통합이라는 비전을 제시하는 데 활용하였음을 제시하기도 하였다. 같은 연설에 대해 Burt(2015)는 북부와 남부 모두가 미국 남북전쟁의 폭력과 노예 문제에 대해 책임이 있다는 점에 대해 취임사를 통해 집합적, 도덕적 책임을 강조한 점이 두드러진다는 결과를 도출하였다. 레이건 대통령 취임사를 분석하여 경제 위기의 해결책으로 작은 정부가 제시되었고 자유주의의 진보적 가치를 수용하는 가치 체계와 이른바 개별화된 아메리칸드림에 대한 강력한 서사가 나타났다고 제시한 연구도 수

행되었다(Rowland & Jones, 2016). Vitale(2015)은 2007년에 아르헨티나 대통령으로 취임한 크리스티나 페르난데스 데 키르히너 대통령이 자신의 정치 지도력을 합법화하기 위해 취임 연설을 어떻게 활용하였는지에 대해 분석하기도 했다.

대통령제를 본격적으로 시행한 지 근 백 년이 채 되지 않는 만큼 대통령 취임사를 비롯한 연설문을 분석하는 국내 연구들은 미국 등의 국가와 비교하여 아직 초기 단계다. 앞에서 설명하였듯이, 2000년대 이후부터 서서히 시도되고 있는데 대체로 특정 시기 혹은 진보·보수 정부의 비교, 정치적 이념이 다른 대통령 간의 취임사 특징 분석 등이 주로 진행되었고 대통령 취임사의 정치적 수사 변화, 가치관 등을 확인하는 연구들이 일부 수행되었다.

이러한 점에서 본 연구는 초대부터 현직에 이르는 전체 대통령의 취임사를 분석한다는 점에서 시계열적 흐름 전체를 대상으로 하는 차이점이 있고, 핵심 국정 가치라는 구체적인 주제로 분석을 시도

한다는 점에서 의미가 있다.

2. 역대 대통령별 재임기간 및 주요 특징

우리나라는 1948년 7월 20일, 간접선거를 통해 초대 대통령으로 당선된 이승만 대통령 이래 2023년, 현재까지 75년간 총 20대에 걸쳐 13명의 대통령이 국정을 담당하였다. 이 시기에 근대화, 산업화, 민주화 과정을 차례로 거쳐 오면서 미처 다 해결하지 못한 의제들과 함께 시대변화에 따른 신사회문제까지 오늘날 같은 시간대로 발생하고 있다.

과거와 현재의 사회문제들이 긴밀하게 엮여 있다는 점에서 역대 대통령 재임 기간에 발생한 정치, 사회, 경제 분야 주요 사건들은 한국 사회를 설명하고 이를 통해 발생하게 되는 국민 욕구를 이해하는데에도 도움이 된다. 우리나라는 13명의 대통령이 20대에 걸쳐 정부 운영을 맡아왔는데 재임 기간별 주요 특징을 정리하면 다음 <표 1>과 같다.

<표 1> 역대 대통령별 재임기간 및 주요 특징

	대통령	임기	취임식일자	선거방식	재임기간	비고
1	이승만	4년	1948. 7. 24.	간접선거	1948. 7. ~1960. 4.	<ul style="list-style-type: none"> • 한국전쟁발발 • 재임(12년)
2		4년	1952. 8. 15.	직접선거		
3		4년	1956. 8. 15.	직접선거		
4	윤보선	4년	1960. 8. 13.	간접선거	1960. 8. ~1962. 3.	<ul style="list-style-type: none"> • 3·15부정선거(1960)에 의해 1960. 8. 12. 재선거 당선 • 5·16군사정변(1961) 으로 1962. 3. 22. 하야
5	박정희	4년	1963. 12. 17.	직접선거	1963. 12. ~1979. 10.	<ul style="list-style-type: none"> • 1979. 10. 26. 사거(死去) (재임16년)
6		4년	1967. 7. 1.	직접선거		
7		4년	1971. 7. 1.	직접선거		
8		4년	1972. 12. 27.	간접선거		
9		6년	1978. 12. 27.	간접선거		
10	최규하	6년	1979. 12. 21.	간접선거	1979. 12. ~1980. 8.	<ul style="list-style-type: none"> • 권한대행
11	전두환	6년	1980. 9. 1.	간접선거	1980. 9. ~1988. 2.	<ul style="list-style-type: none"> • 민주화운동 • 6·29 민주화 선언 (1987)
12		7년	1981. 3. 3.	간접선거		

	대통령	임기	취임식일자	선거방식	재임기간	비고
13	노태우	5년	1988. 2. 25.	직접선거	1988. 2.~1993. 2.	• 직선제 당선
14	김영삼	5년	1993. 2. 25.	직접선거	1993. 2.~1998. 2.	• 금융실명제(1993) • 1995년 전국동시지방선거 등 민선지방자치제 실시 • IMF긴급구제금융조치(1997)
15	김대중	5년	1998. 2. 25.	직접선거	1998. 2.~2003. 2.	• 대북포용정책 • 노벨평화상수상 (2000) • IMF 차입금 조기상환(2001)
16	노무현	5년	2003. 2. 25.	직접선거	2003. 2.~2008. 2.	• 헌정 최초 탄핵소추 및 기각 (2004)
17	이명박	5년	2008. 2. 25.	직접선거	2008. 2.~2013. 2.	• 글로벌 금융위기(2008) • 4대강 사업(2015)
18	박근혜	5년	2013. 2. 25.	직접선거	2013. 2~2017. 3.	• 2017. 2. 탄핵 (재임 4년)
19	문재인	5년	2017. 5. 10.	직접선거	2017. 5.~2022. 5	• COVID-19 발생(2020)
20	윤석열	5년	2022. 5. 10.	직접선거	2022. 5.~2027. 5.	

위 표 1에서 지적한 것과 같이 초대 정부부터 나타난 부정선거 등과 같은 문제는 결국 우리나라의 정치와 민주화를 오랫동안 가로막은 방해물이 되었다. 광복 이후 온전한 국권 회복이 제대로 시작되기도 전 발발했던 한국전쟁의 결과로 맞게 된 분단국가 상황은 통일과 대북정책이 등장한 원인이자 배경이다. 뿐만 아니라 이러한 분단 현실은 우리나라를 위협하는 지정학적인 위협 요인으로 자리 잡았다. 독재·군사정권에 대한 민주화운동은 시민정신으로 되찾은 민주주의로 이어져 지금까지 이어지고 있다. 그러나 여전히 청산되지 못한 친일 잔재, 독재·군사정권이 남긴 구습과 폐단, 가치 갈등에 의한 정치 분열 현상은 성숙한 자유민주주의 국가로 성장하는 데 걸림돌이 되고 있다.

한편 1990년대 이후 겪게 된 외환위기는 그 이전 단계에서 해결하지 못한 사회문제와 맞물려 단일국가 차원을 넘어서는 신사회위험을 초래하였다. 이러한 변화는 우리나라 경제·산업 전반에 걸쳐 시장친화적인 구조조정을 불러왔지만, 취약계층과 사회적 약자들이 겪게 된 각종 문제는 친시장적 조정

결과의 한계를 고스란히 보여주었다. 형식적으로는 외환위기로부터 시작된 경제문제를 조기상환으로 해결한 것처럼 보일지라도, 그로부터 촉발된 경제와 복지 그리고 사회의 지속 가능한 발전을 둘러싼 부조화와 불균형의 상황은 여전히 해결해야 하는 숙제로 남아 있다.

III. 연구방법

1. 자료수집

역대 대통령 취임사에 내재된 정책 기조 변화 양상을 분석하기 위한 이 연구에서는 초대 이래 20대까지 이르는 취임사 전체를 대통령기록관에서 확보하였다. 대통령기록관에는 취임사 이외 각종 연설문, 영상 및 문헌자료 등이 제공되고 있는데 온라인으로 공개된 대통령기록관 사이트의 기록컬렉션 사이트에서 취임사에 해당하는 기록물을 모두 선별, 추출한 후 자료 수집을 진행하였다. 이 가운데 중복

으로 게시된 내용을 제외하고 총 20건의 취임사 전문(全文)을 분석 대상으로 하였다.

2. 분석방법

개인과 집단이 어울려 살아가는 사회에서는 서로 다른 행위자들에 의한 텍스트 데이터가 끊임없이 만들어진다. 이러한 데이터는 주로 사적 공간 내 상호작용에서 생성되지만, 공적 영역에서도 다양한 형태의 텍스트 데이터로 생산되고 있다. 최근에는 정보통신 및 데이터 처리 기술의 발전에 따라 공공부문의 텍스트 데이터가 디지털화되어 빅데이터로 구축되는 추세이다.

이러한 배경에서 공공 빅 텍스트 데이터를 분석하는 데 알고리즘 기술을 활용하는 텍스트 마이닝이 주목받고 있다. 텍스트 마이닝은 자연어 처리 기술을 적용하여 대규모 텍스트 데이터로부터 의미 있는 정보와 지식을 추출하는 기술을 의미한다. 텍스트 마이닝은 정보 추출(information retrieval), 감성분석(sentiment analysis), 주제 발굴(topic modeling) 등에서 적용된다(이상국, 2022: 68).

과거에는 텍스트, 이른바 비계량적 자료에 대해 내용분석을 적용해왔다. 내용분석은 기록물, 매체, 발화 녹취록, 연설문 등과 같은 자료의 내용에 대해 주로 질적 방법을 적용하여 그 속에 담긴 메시지를 분류하여 의미나 특성을 파악하는 방식인데 이를 통해 의미 분석, 시간적 추이에 따른 변화 경향, 차이점 등을 파악해 왔다.

그러나 내용분석에 대해 외적 타당성의 제한, 개념적 모호성, 연구자 분석에 대한 과도한 의존, 편향적인 자료수집 가능성과 데이터 처리량의 한계와 같은 문제가 발생하는 점이 종종 지적됐다(박준형 외, 2017).

내용분석 뿐만 아니라 질적 연구에서 자주 적용하는 현상학적 해석 역시 연구하고자 하는 행위자의 경험을 밝히기 위해서는 개인의 체험을 끌어낼

수 있는 질문을 통해 분석되어야 한다. 그런 만큼 참여자들의 구체적인 진술과 독특한 경험을 토대로 현상의 본질적인 의미를 밝히는 데는 의미가 있겠으나 공적 영역에서 이미 생산된 텍스트를 분석하는 것에는 적절하지 못하다(김해숙 외, 2018).

이러한 점에서 텍스트 마이닝은 수집된 자연어 데이터에 대해 계량적으로 더 정교한 결과를 도출하고 간결하게 표현할 수 있는 장점이 있다. 나아가 정성적 분석에만 의존할 때 발생하기 쉬운 연구자의 인지 편향에 따른 정보 왜곡 현상도 해소할 수 있으므로 본 연구의 분석 방법으로 채택하였다.

3. 분석절차 및 분석도구

취임사와 같은 비정형 텍스트 데이터를 분석하기 위해서는 전처리과정(pre-processing)을 반드시 실시해야 한다. 본 연구에서는 먼저 수집된 취임사에 대해 인용표기, 따옴표 등 불필요한 문장부호 제거, 한자병기 단어 정리, 오타자 검수, 띄어쓰기 및 동의어 정리 등을 차례로 실시하였다.

그 이후 연구 목적과 거리가 먼 불용어를 제거하였는데 이때 사용자 불용어 수집사전 방식을 적용하였다. 불용어는 연구대상 자료 출현빈도는 높지만 의미·맥락에는 영향을 미치지 않는 단어인 바 ‘그리고’, ‘그래서’, ‘그러나’ 등과 같은 접속사, ‘또’, ‘또한’, ‘더불어’ 등과 같이 문장 연결에 적용되는 부사, ‘이’, ‘그’, ‘저’ 등의 관형사, ‘것’, ‘등’, ‘뿐’ 등의 의존 명사들을 제거하였다.

불용어처리 및 형태소 분석이 모두 마무리된 후, 최종 전처리된 자료를 바탕으로 파이선의 LDA 토픽 모델링 관련 패키지 및 라이브러리를 활용하여 분석을 시행하였다.

먼저 TF(Term Frequency, 단어 빈도, 이하 TF) 분석과 TF-IDF(Term Frequency-Inverse Document Frequency, 단어 빈도-역문서빈도, 이하 TF-IDF) 분석을 수행하였다. TF 분석결과는 문서에서 해당

단어가 몇 번 출현했는지 빈도를 나타내는 것인데, 많이 등장하는 단어가 가장 유용하다고 판단하는 것은 일반용어의 과잉 표상과 같은 문제가 발생할 수 있으므로 그러한 단어의 중요도를 낮추어 볼 필요가 있다. 한편 TF-IDF 분석결과는 단어 중요도의 가중치로서 특정 단어가 문서 내에서 얼마나 중요한지를 나타내는 통계적 수치이므로, 두 방법 모두를 적용하여 비교·분석하였다.

두 번째, 단어 네트워크(Word-Network)분석을 수행하였다. 이 방법은 자연어 처리 기법 가운데 하나로써 자연어 데이터에서 도출된 주요 주제어 간의 중요도나 관련성을 파악하고 시각화하는 데 유용하다. 이를 위해 문서를 단어로 분리하고, 중복되는 단어를 제거하여 단어 집합을 만들고, 문서에서 단어 간의 연결 관계와 중요도를 확인하였다.

이때 연결 중심성 지수와 매개중심성 지수를 적용하였다. 연결 중심성 지수는 텍스트에서 특정 단어가 다른 단어들과 얼마나 많은 관계를 맺고 있는지 보여준다. 매개중심성지수 분석은 각 단어가 다른 단어 간의 중개자 역할을 하는지 측정하여 매개성이 높은 단어를 찾을 수 있게 한다(Borgatti, 2005). 본 논문에서는 두 가지 지수를 동시에 적용함으로써 단어네트워크 설명력을 향상하고자 한다.

마지막으로 토픽 응집도 분석을 수행하여 LDA 토픽 모델링(잠재 디리클레 할당, Latent Dirichlet Allocation)에 적용될 토픽 수를 결정한 후 최종 분석 결과를 시각화하였다. 아울러 시각화를 위해 Sievert and Shirley(2014)가 개발한 LDAvis도구를 적용하였다. LDAvis는 웹 기반 토픽 모델링 시각화 도구로서, 주제와 단어의 관계를 전반적으로 살펴볼 수 있으며, 각 토픽과 토픽 내 단어를 중요도에 따라 순위화하여 해당 문서 집합에서 주요한 토픽 및 단어를 쉽게 파악할 수 있는 특징이 있다.

그뿐만 아니라 개별 토픽의 속성도 동시에 볼 수 있어서 이해가 용이하고 토픽과 용어 사이의 관계를 통해 학문적 통찰을 갖기에 유용하다. 이러한 점

을 고려하여 본 논문에서도 시각화 결과를 제시하는 데 LDAvis를 적용하였다.

IV. 연구결과

1. TF 및 TF-IDF 분석결과

역대 대통령 취임사 데이터의 TF, TF-IDF 분석한 결과는 아래 <표 2>와 같다. TF 분석을 통해 가장 많이 등장한 키워드를 빈도순으로 보면 세계, 민족, 평화, 역사, 자유 등의 순으로 드러났는데, 특정 분야, 정책영역과 연결되는 구체적인 단어보다는 추상적이고 광범위한 가치를 표현하는 단어가 빈번하게 사용된 것으로 볼 수 있다. 정책분야의 경우 경제, 문화, 통일 영역이 상위 10개 이내에 포함된 것을 알 수 있다.

한편 단순 등장 빈도가 갖는 한계를 보완하기 위해 특정 단어의 중요성을 수치로 나타낸 가중치(term-weighting) 값인 TF-IDF 결과의 경우, 민족, 역사, 세계, 발전, 정치의 순으로 상위 5개 값이 나타났으며, 그 외 존경, 헌법이 새롭게 포함되어 있다.

분단 이후 민주화 이전까지 자유민주주의적 법질서와 가치에 입각한 정권교체보다는 군부독재 등 민주주의 가치를 훼손하면서 집권·통치했던 역사적 사실에서 보면, 이러한 키워드는 민주화 이전 시기의 경우 집권 대통령 개개인들이 해당 정권의 정당성을 그럴듯하게 포장하기 위한 수사적 차원에서 사용했을 것으로도 보인다. 민족과 역사성이 두드러져 보이는 것은 정부와 정권의 정통성을 확보하려는 목적이 대통령 취임사의 공통적인 서사구조로 반영된 것으로 생각해 볼 수 있다.

TF 분석에서 상위 5순위 이내에 포함된 자유의 경우 TF-IDF 결과에서는 9순위로 밀려져 있었고 TF 분석에서 상위 20순위 이내 포함되지 않았던 평화는 오히려 8순위로 측정되었다. 통일 단어 역시

<표 2> TF 및 TF-IDF 분석결과

TF			순위	TF-IDF		
키워드	빈도	점수		키워드	빈도	점수
세계	127	0.1284	1	민족	125	0.0400
민족	125	0.1264	2	역사	97	0.0386
평화	102	0.1031	3	세계	127	0.0351
역사	97	0.0981	4	발전	94	0.0332
자유	95	0.0961	5	정치	90	0.0317
발전	94	0.0950	6	존경	36	0.0308
정치	98	0.0910	7	경제	84	0.0306
경제	84	0.0849	8	평화	102	0.0302
문화	75	0.0758	9	자유	95	0.0278
통일	70	0.0708	10	민주주의	67	0.0234
민주주의	67	0.0677	11	조국	54	0.0228
조국	54	0.0546	12	문화	75	0.0221
협력	52	0.0526	13	동포	39	0.0214
민주	50	0.0506	14	통일	70	0.0210
북한	48	0.0485	15	북한	48	0.0196
국제	45	0.0455	16	헌법	37	0.0192
안정	42	0.0425	17	책임	42	0.0182
책임	42	0.0425	18	국제	45	0.0171
정신	41	0.0415	19	협력	52	0.0168
동포	39	0.0394	20	생활	38	0.0168

TF-IDF 순위상 그 중요성이 TF 등장 빈도와 비교하여 상대적으로 낮다.

요컨대 민족성과 역사성을 바탕으로 하는 서사구조 바탕으로 각 정부의 정통성을 강조한 것이 엿보이고, 자유, 경제, 민주주의 등의 단어가 비중 있게 다뤄져 왔음을 알 수 있다.

2. 단어연결망 분석결과

취임사에 나타난 주요 단어간의 연결망(Word-Network)을 분석하기 위해 연결중심성(degree centrality) 지수와 매개 중심성지수를 각각 적용하여 분석한 결과는 아래 <표 3>과 같다.

먼저 연결중심성지수에 따른 네트워크를 중심으로 살펴보면, 국제(0.058), 경제(0.055), 국정(0.055), 개인(0.053), 교육(0.053)의 순으로 나타났다. 특히 연결중심성 네트워크에서 분석결과 상위 10단 단어 가운데 국제를 제외한 나머지 단어들은 TF 빈도의 상위 20순위 내에 들어가 있지 않았다.

연결중심성 분석결과에서 나타난 주요 10개 주제는 국제, 경제, 국정, 개인, 개혁, 교육, 가치, 관계, 갈등, 건설 등이다. 이는 역대 취임사에서 가장 중심적인 역할을 담당한 단어이다. 특히 국제, 경제, 국정은 다른 주제에 비해 더 높은 연결 수와 중심성을 보이고 있는데, 이는 정부 수립 이후 20대에 이르기까지 한국의 국제정책, 경제정책이 각 분야와

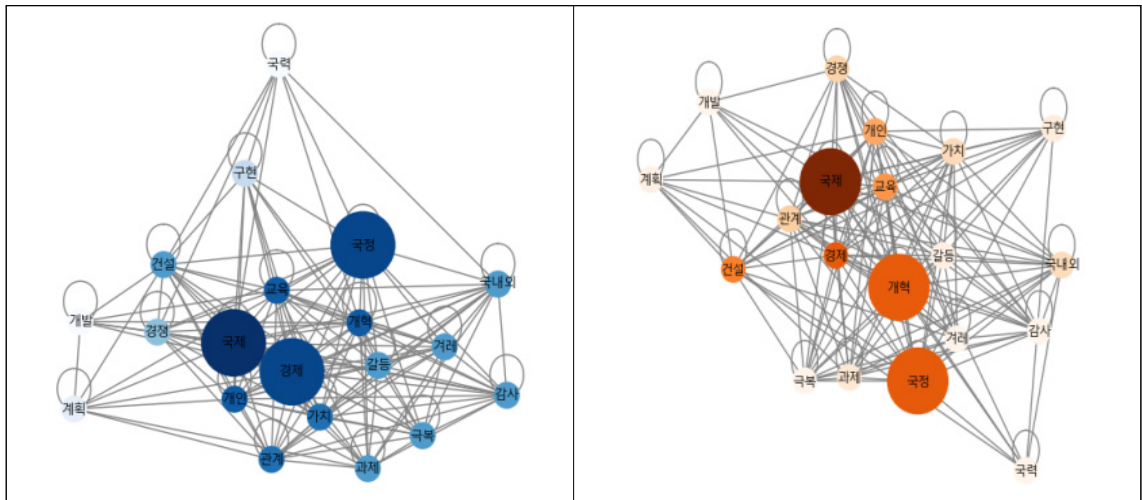
긴밀하게 연결되어 있음을 알 수 있다.

일제 강점기, 한국전쟁 등의 역사적 경험은, 한국은 국제 정세변화에 따라 국가의 존립이 좌우될 수 있다는 인식이 영향을 미친 것으로 보인다. 이는 진보·보수 정부를 떠나 국제정세 대응과 경제정책이 모든 국정운영 과정에서 중심을 차지해야 한다는 점에 대해 역대 대통령 인식이 일치되어 있음을 보여준다.

한편 매개 중심성은 네트워크상에서 다른 주제어들이 상호작용할 때 중개자 역할을 하는 주제어를

확인할 수 있는 지표로서 매개중심성이 높을수록 다른 단어 간의 상호작용에서 중요한 역할을 하는 것으로 해석할 수 있다. 매개중심성지수 분석결과 국제(0.045), 국정(0.031), 개혁(0.031), 경제(0.031), 건설(0.026)의 순서로 나타났다. 이러한 주제어들이 다른 주제어와의 연결관계에서 매개(중개) 요인으로 작용하는 것을 알 수 있다. 특징적인 것은 국제와 국정 주제어의 뒤를 이어 “개혁”이 상위 3순위에 있다는 점이다.

<표 3> 단어네트워크 분석결과



<연결중심성 네트워크>

<매개중심성 네트워크>

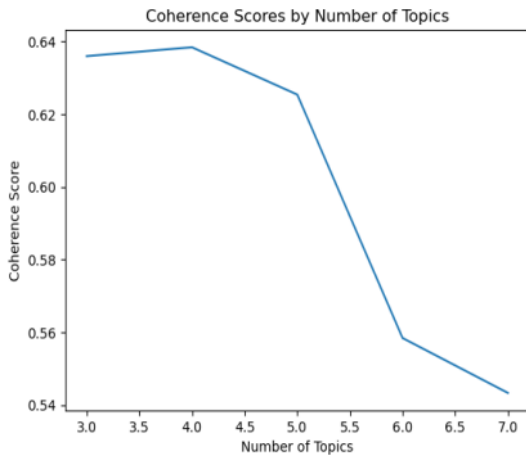
연결중심성			순위	매개중심성		
단어	지수	TF		단어	지수	TF
국제	0.058	45	1	국제	0.047	45
경제	0.055	84	2	국정	0.031	18
국정	0.055	18	3	개혁	0.031	33
개인	0.053	22	4	경제	0.031	84
개혁	0.053	33	5	건설	0.026	34
교육	0.053	34	6	교육	0.022	34
가치	0.050	26	7	개인	0.019	22
관계	0.050	22	8	관계	0.013	22
갈등	0.044	21	9	경쟁	0.012	18
건설	0.044	34	10	가치	0.010	26

이는 곧 우리나라가 당면해 있는 전체 의제를 매개하는 중심 가치는 모든 영역에서의 변화와 쇄신, 즉 개혁이라는 것으로 해석할 수 있다.

위 <표 3>에서도 확인할 수 있는 것처럼, 대부분의 주제어를 개혁이 매개하고 있는 것은 한국 정부와 정책은 변화되어야 하며, 국정 전반에 걸쳐 내재한 구습과 폐단을 줄이고 이를 위한 혁신적인 정책 조치가 필요하다는 것을 의미한다.

3. LDA 토픽모델링 분석 결과

LDA 토픽모델링을 실시하기 위해 토픽응집성 지수 분석을 선행하였다. 지수 분석결과는 최적의 토픽 개수를 결정하는데 유용한 기준이기 때문에, 해당 지표값을 참고하여 최종 토픽수를 결정하였다. 분석결과는 다음과 그림과 같다.



<그림 1> 응집성 지수 분석결과

분석결과 토픽 개수는 4개에서 가장 좋은 결과(응집성 지수: 0.6384)를 보여주었다. 이를 기준으로 LDA 토픽모델링 분석을 한 바 다음 표 4와 같다. 토픽별 주제어는 개별 주제의 상위 구성비를 차지한 단어를 고려하되, 다른 영역에 나타나지 않거나 상대적으로 높은 순위의 단어를 확인하여 그 맥락

을 연결하는 말로 명명하였다.

<표 4> 토픽모델링 결과

	주제 1	주제 2	주제 3	주제 4
	세계경제와 국가발전	자유민주주의와 개혁	민족역사성과 정치	시민생활과 복지
비율	32.2%	27.6%	20.8%	19.5%
1	세계	민주주의	역사	민족
2	경제	자유	정치	문화
3	평화	평화	민족	역사
4	민족	세계	발전	세계
5	발전	발전	경제	통일
6	자유	정치	평화	복지
7	통일	존경	세계	책임
8	조국	북한	통일	민주
9	문화	개혁	협력	생활
10	민주	교육	정책	시민

주제 1에서 확인된 정책기조는 가장 높은 비중을 차지한 세계경제와 국가발전이다. 대통령과 정부가 바뀔 때마다 다양한 정책의제가 나타나지만, 역대 정부 핵심 정책기조는 주제 1에서와 같이 세계경제 상황과 국가 발전 방안 수립에 대한 것이었다. 분단 국가라는 지정학적 상황, 세계 경제 구조 변화에 따라 민감한 영향을 받아왔던 한국 사회의 근현대사적 특징이 반영된 것으로도 볼 수 있다. 이는 대통령들의 국가의제 인식에도 영향을 주었는데 우리나라는 외부 경제 요인 변화에 대해 다른 국가에 비해 더 크게 영향을 받기 때문에, 세계 경제 상황에 기민하게 대처해야 한다는 정책기조가 그동안 두드러졌던 것으로 보인다.

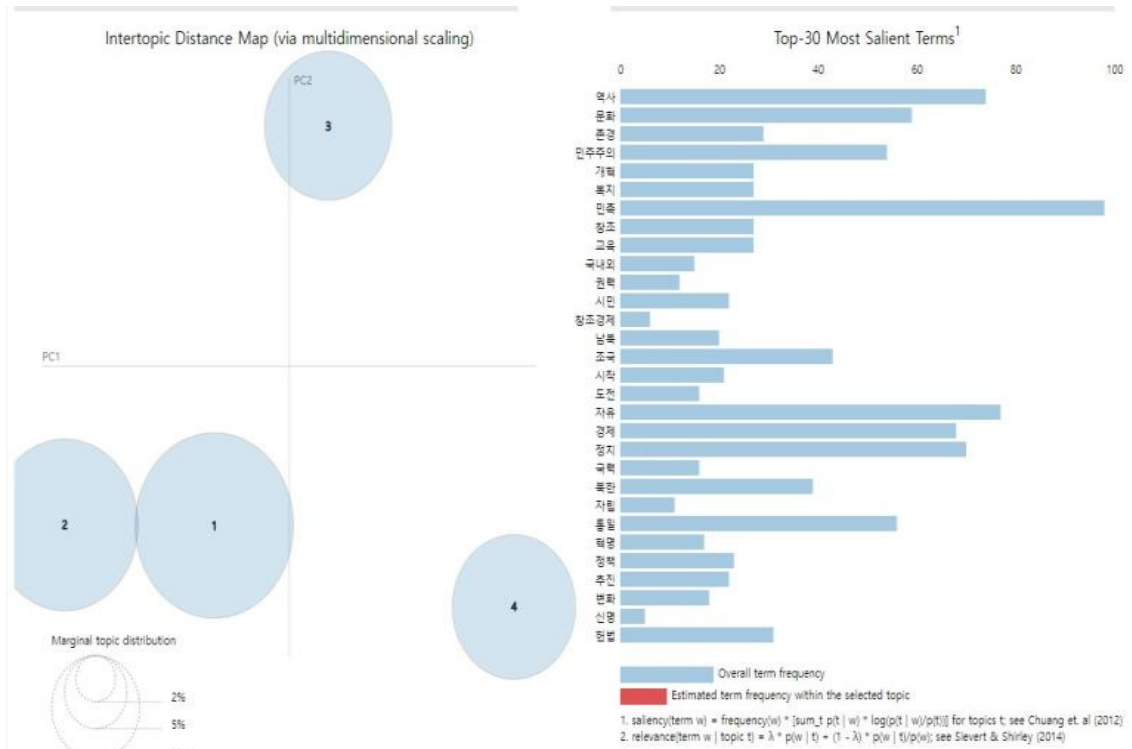
주제 2는 자유민주주의와 개혁에 대한 정책기조이다. 이 주제에서는 민주주의, 자유가 가장 높은 비중을 차지하였는데, 전술한 매개중심성에서 강조된 단어인 개혁 역시 본 주제에 유일하게 포함되어

있었다. 한국의 민주주의는 독재·군사정권의 억압과 이에 맞서 대항해 온 민주화·시민 운동에서 비롯되었음은 잘 알려진 사실이다. 역대 대통령과 정부가 바뀌는 동안 수많은 폐단과 구습이 반복되었을지라도, 국민은 자유민주주의를 포기하지 않았다. 민주화 이후 보다 성숙한 민주주의 사회를 이루려는 국민 욕구는 더 커지고 있다. 이러한 맥락에서 한국 민주주의의 성장과 개혁과 관련한 정책기조가 비교적 꾸준히 다뤄졌던 것을 주제 2를 통해 알 수 있다.

주제 3은 민족역사성과 정치이다. 반만년을 이어 오는 민족 고유의 역사성에 대한 강조는 특정 정부 혹은 대통령의 정책 기조가 갖는 의미 그 이상이다. 비록 각 대통령의 역사 인식과 이념적 접근방식에서 차이는 있었을지라도 민족 역사성을 지키고 정

치발전을 목표로 둔 정책을 어떻게 만들어 갈 것인가에 대한 고민은 계속됐다고 말할 수 있다.

마지막으로 주제 4의 경우, 다른 주제와 달리, 문화, 복지, 책임, 생활, 시민 등 개별 국민의 구체적인 생활세계와 연결되는 지점이 등장한다. 상대적으로 다른 주제에 비해 차지하는 비중은 제일 낮지만, 역대 대통령 모두 국가 책임에 의한 복지국가 관련 정책기조를 저마다의 관점에서 접근해 온 것으로 생각해 볼 수 있다. 이 결과를 LDAvis 분석을 통해 시각화해 보면 아래 그림 2와 같이 4개의 토픽 간 중첩은 거의 일어나지 않았고 주제 1의 비중이 가장 높았다. 주제 간 비중 차이는 크지 않은 바 특정 주제로 쏠렸다고보다는 추출된 4가지 주제가 비교적 고르게 강조되고 있는 확인할 수 있다.



<그림 2> LDAvis 분석 결과

V. 결론

대통령은 국가정책의 형성과 의사결정 과정에 막대한 영향을 미치며, 각 대통령의 중심 가치가 무엇인가에 따라 정책 우선순위가 결정되고 그에 따라 정책이 수립되어왔다. 이에 본 연구에서는 역대 대통령의 취임사에 내재된 정책기조를 구성하는 주제를 찾아내고자 하였다.

연구분석 결과에서도 드러났듯이, 정책 분야의 경우, 정치, 문화, 안보, 복지, 교육 등 다양한 영역이 나타났지만, 그 가운데 경제영역이 가장 강조되고 있었다. 특기할 만한 것은 연결중심성과 매개중심성 모두 '경제' 키워드가 상위를 차지한 점이다. 매개중심성 결과로 보면 경제 분야 등 여러 정책 분야에서 개혁이 주요한 매개어로 나타났다. 토픽모델링 분석결과, 세계 경제와 국가발전, 자유민주주의와 개혁, 민족 역사성과 정치, 시민 생활과 복지가 우리나라 근현대 정책기조를 구성하는 공통 주제임을 확인하였다.

이러한 주제들이 주는 의미는 한국 사회의 쟁점들과 맞닿아 있다. 한국전쟁, 군부독재, 민주화 투쟁, 국정농단에 따른 탄핵에 이르기까지 우리 사회의 외피(外皮)는 안정적인 성장이 담보된 국가의 성과 거리가 멀었다. 또한 시민사회의 성장과 시민의식의 형성 속도 역시 압축적인 근·현대화 과정에서 사회변화 속도와 반드시 정비례하지 않았던 경험을 안고 있다. 정경유착으로 인한 자유민주주의 시장경제 질서 왜곡, 경제발전 수준과 국민의 삶의 질 향상 정도가 균형을 맞추지 못해 발생했던 각종 문제를 경험해온 국민이 정치·경제의 민주화를 원하는 것은 어쩌면 당연한 결과이다.

국민의 궁극적인 바람과 지향점은 자유민주주의에 터 잡은 시민으로서의 인간다운 삶이다. 그러므로 대통령이 바뀌어 온 모든 과정에서 국민은 이러한 욕구를 해소할 수 있는 정책을 원해왔고, 역대 대통령 역시 수준과 정도의 차이는 있지만 이를 인

지하였으므로 취임사의 정책기조로 발화되었다.

달리 설명하면 국민은 이러한 4가지 주제로서의 공적 가치를 개인의 삶이라는 사적자치 영역에서 떼어놓은 적이 없었다는 설명도 가능하다. 그래서 특정 대통령, 정부가 아닌 20대에 이르는 전체 정부와 역대 대통령의 취임사로부터 이러한 주제가 공통으로 추출된 것이다.

대통령의 취임사가 국민 욕구와 사회의제를 반영하는 언어라고 하였을 때 이번 분석을 통해 드러난 4가지의 주제는 특정 정부 혹은 시·공간성을 초월하고 있는 정책기조라는 점에서 의미가 있다. 국민이 원하는 정책에 가장 근접해 있는 의제라고도 말할 수 있다.

현재 한국 사회는 국내·외에 걸친 다양한 사회변화와 문제에 직면하고 있다. 이러한 시기일수록 정책 대상자인 국민 당사자들의 욕구에 기반하여 정책이 수립되어야 한다. 바로 이 점에서 역대 대통령 취임사를 통해 확인한 4가지 주제는 정부수립 이후 우리 사회를 지배하고 있는 근본 욕구이다. 역대 대통령 취임사에서 찾은 4대 주제는 현재진행형으로 전개되고 있는 우리 사회의 욕구이며 과제, 그 자체인 셈이다.

이번 연구는 역대 대통령 연설을 통해 추출된 공통 정책 기조, 즉 주제를 확인하는 데 초점을 두었다. 시계열적인 흐름을 관통하여 나타나는 중심 주제를 확인한 만큼, 이러한 주제들이 실제로 정책으로 실체화되었는지도 분석할 필요가 있다. 토픽모델링에 의한 주제 도출 연구의 특성상 이를 다루지 못한 것은 연구 범위와 분석 방법에서 비롯되는 한계이다. 동시에 후속 연구 주제로 확장되어야 할 필요성과 가능성을 보여주는 것이기도 하다. 이를 바탕으로 역대 대통령 혹은 정부 정책에 대한 공과(功過) 평가 등의 연구가 이어진다면 국민의 욕구에 기반한 지속 가능한 정책 개발에도 활용될 수 있을 것이다.

참고문헌

1. 국내

가상준(2023). 2022 대통령 선거 투표참여와 정서적 양극화: 부정적 투표(negative voting)를 중심으로. 서울: 경희대학교 인류사회재건연구원 연구보고서.

강태완, 김은정(2004). 역대 대통령의 연설문에 나타난 수사적 특징과 역할규정. 한국스피치커뮤니케이션학회 학술대회 자료집. 67-106.

구본상(2021). 성차별 인식은 대선 후보 선택에 영향을 주는가?: 제19대 대통령선거 분석. 한국정당학회보. 20(2): 39-71.

김성연(2022). “근접성 모형, 방향성 모형, 그리고 한국 유권자들의 이념 투표: 2010년대 이후 대통령 선거 분석 결과.” 미래정치연구. 12(2): 5-28.

김승욱, 남현주(2022). 의존을 포용하는 평등의 이상과 복지국가 - 롤즈의 정치이론에 대한 키테이의 의존비판에 관한 연구. 한국사회복지학. 74(4): 41-63.

김재한(2014). 18대 총선에서 나타난 선거지역주의. 한국지역지리학회지. 20(3): 315-333.

김해숙, 임남열, 이혜진(2018). “살인사건 피해 유가족의 심리적 경험에 대한 현상학적 연구”. 한국심리학회지: 문화 및 사회문제. 24(2):301-335.

김학용, 김학봉(2014). 조선왕조실록 네트워크의 동적 변화 분석. 한국콘텐츠학회논문지, 14(9): 529-537.

김 혁(2014). 대통령의 리더십과 정책의제설정양태에 대한 연구. 한국정치연구, 23(2), 77-102.

김현정(2022). 대통령 후보자 선출과정에서 여야 양당의 복지정책 논쟁에 관한 분석. 한국과 세계. 4(2).

박정택(2007). 대통령직 인수와 정책기조의 형성 - 문민정부 이후 대통령직 인수 사례의 분석을 중심으로-. 한국정책학회보 한국정책학회. 16(4):

1-30.

박종구, 이보라(2015). 18대 대선 전·후의 국민 복지인식 변화에 관한 연구. 한국행정논집. 27(2): 377-397.

박준형, 유호선, 김태영, 한희정, 오효정(2017). 대통령 연설기록 내용분석을 통한 역대 대통령의 중심가치 연구. 한국기록관리학회지. 17(2): 57-78.

백자성, 윤지원(2022). 역대 대통령 연설문을 통해 추론한 한국 전략문화의 속성. 국방연구. 65(4): 1-27.

윤우재, 서인석, 이동규(2011). 문민정부 이후 역대 대통령 미래비전 리더십 연구: 취임사에 나타난 국가의제 및 연설(言說: statement)을 중심으로. 한국행정연구. 20(1): 37-66.

이수범, 박희영(2022). 제20대 대통령선거의 표심 분석: 신문기사의 의미연결망을 중심으로. 선거연구. 17: 65-98.

이준한, 황원재(2021). 12시면 최종투표율을 알 수 있다: “선거 당시 시간대별 투표율 분석과 예측 모델.” 선거연구. 15:5-34.

임성호(2018). 버락 오바마, 도널드 트럼프, 문재인 대통령의 취임사 속에 담긴 지도자 이미지 및 가치 비교 연구. 한국소통학회 2018 봄철 정기학술대회 발표집, 93-99.

임순미(2009). 정치리더의 가치프레임: 김대중 대통령과 이명박 대통령의 취임사 비교. 국가정책연구, 23(4), 35-67.

임주영, 박형준(2015). 직무특성과 환경변화에 따른 대통령의 국정가치 변화분석 -이명박 정부의 사례 중심으로-. 한국정책학회보 한국정책학회. 24(2): 149-183.

전일욱(2021). 제8·9대 대통령 취임사에 나타난 국정목표. 한국행정사학회지. 53: 55-79.

전일욱(2022). 진보정부와 보수정부의 국정목표 비교연구: 노무현·이명박 대통령 취임사를 중심으로. 한국행정사학회지. 56: 45-63.

전일욱, 이병관(2021). 제3공화국의 국정 목표에 관한 연구: 제5·6·7대 대통령 취임사를 중심으로. 한국행정사학회지. 51: 31-58.

2. 외국

Borgatti, S. P. (2005). Centrality and network flow. *Social networks*, 27(1), 55-71.

Burt, J. (2015). Collective Guilt in Lincoln's Second Inaugural Address. *American Political Thought*, 4(3), 467 - 488.

<https://doi.org/10.1086/682042>

Morel, L. (2015). Of Justice and Mercy in Abraham Lincoln's Second Inaugural Address. *American Political Thought*, 4(3), 455 - 466.

Rowland, R. C., & Jones, J. M. (2016). Reagan's Strategy for the Cold War and the Evil Empire Address. *Rhetoric and Public Affairs*, 19(3), 427 - 464.

Sievert, C., & Shirley, K. (2014). LDAvis: A method for visualizing and interpreting topics. *In Proceedings of the workshop on interactive language learning, visualization, and interfaces*. pp. 63-70.

Vitale, M. A. (2015). Legitimizing Leadership: Argentine President Cristina Fernández de Kirchner's 2007 Inaugural Address. *Rhetoric Society Quarterly*, 45(3), 250

투고일자: 2023. 5. 15.

심사일자: 2023. 6. 11.

게재확정일자: 2023. 6. 16.

Analyzing the Policy Direction in the Presidential Inaugural Addresses Using LDA Topic Modeling

TaeJeong Park

Department of Social Welfare, Seoul Cyber University

This study investigated the underlying policy direction encapsulated in the presidential inaugural addresses delivered by the 13 presidents who have led Korea since its establishment, focusing on those who assumed office in their 20s. The study applied various text analysis techniques, including Python programming-based Term Frequency (TF) and Term Frequency – Inverse Document Frequency (TF-IDF) analysis, word network analysis, and Latent Dirichlet Allocation (LDA) topic modeling to analyze the complete set of inaugural addresses. The study results indicated that the presidential speeches emphasized national and historical identity, the economy, culture, unification, and people's everyday lives. Regarding policy direction, the LDA topic modeling analysis revealed four primary themes: the global economy and national development, liberal democracy and reform, national identity and politics, and citizen's life and welfare. These policy orientations are derived from the desires and demands of the citizens, highlighting their importance in shaping government policies. Therefore, research should apply these findings to evaluate the policy performance of past administrations to facilitate subsequent studies.

Key words: Inaugural speeches, LDA topic modeling, Word-Network analysis, Policy orientation, Python