

대한민국 대통령의 취임사 연구* **

-텍스트 마이닝에서의 특성값을 중심으로-

조 용 립***

목 차

1. 시작하는 말
2. 연구 내용 및 방법
3. 취임사 어휘 분석 결과
4. 끝맺는 말

1. 시작하는 말

2022년 3월 9일은 제20대 대통령 선거가 있었으며, 5월 10일에 제20대 신임 대통령이 취임하게 된다.

대통령은 국무총리 및 국무위원의 임명권을 가지며(제69조), 국회에서 제출한 법률안을 거부할 권한을 가진다(제40조). 대통령은 또한 긴급한 경우에 법률의 효력을 가진 명령을 발하거나, 재정상 필요한 처분을 할 수 있다(제57조). 대통령의 임기는 4년으로, 1회 중임이 가능하다(제55조).

위 인용문은 대한민국 헌법 초안이다. 현행 대한민국의 헌법과도 유사한 내용으로 대한민국의 대통령은 강력한 권한과 책임을 지니며, 지금까지 9번의

* 이 논문은 2019년 대한민국 교육부와 한국연구재단의 인문사회분야 신진연구자지원 사업의 지원을 받아 수행된 연구임(NRF-2019S1A5A8034379)

** 이 논문은 2021년 한글학회 연합 가을 학술대회에서 발표한 내용을 곁고 다듬은 것이다. 학술대회에서 도움말을 주신 선생님들과 익명의 논문 심사자들에게 감사사를 드린다.

*** 목원대학교 기초교양학부 조교수

개헌을 통해 현재의 대통령제 중심의 헌법을 유지하고 있다.¹⁾

선서의 내용²⁾을 요약하면 대통령은 직무상 ‘헌법 수호 의무’, ‘국가의 독립·보전 의무’, ‘직무 수행 의무’, ‘겸직 금지 의무’, ‘평화 통일 노력 의무’를 지게 된다. 즉 대통령의 발언은 고도의 정치 행위이기도 하지만 직무상 의무를 성실히 수행하는 역할을 하며, 사적 담화가 아닌 헌법과 기타 법률 등 권리와 의무를 다하는 행위로서 기록되며, 그 발언은 실행까지 담고 있어 상당한 파급력을 가지고 있다.

1) 연구 목적

대통령 연설문 중 취임사는 국가적 당면 과제와 사회적 지향점으로 기록된 역사적 가치가 매우 크다. 따라서 대통령 연설문을 텍스트 마이닝 방식으로 살펴 시대적 변화에 따라 국가 운영 방식의 변화를 살필 수 있다는 점에서 그 의의와 효용 가치가 있다.

또한 대통령 연설문 중 취임사는 대통령으로서 첫 공식 발언으로 그 중요도가 높고, 법률 적합성, 정치적 가치관, 개인적 신념 등을 담고 있다. 결국 원고의 충실성, 주제의 명확성, 내용의 전달성 등에서 중요성이 있다고 볼 수 있다.

이 연구에서는 1948년 이후부터 대한민국 역대 대통령 취임사를 텍스트 마이닝, Big kinds, Google 트렌드 분석을 활용하여 ‘어휘의 빈도수 및 특성’, ‘어휘의 속성 파악’, 명사 추출 ‘워드 클라우드’ 방식으로 분석하고자 한다.

2) 선행 연구 검토

텍스트 트렌드 연구는 기존의 전문가 의견, 설문 조사 등과 같은 전통적인 방법에서 벗어나 해당 분야에서 생산되는 텍스트 자체를 분석함으로써, 데이

1) 대한민국임시정부 건국강령 초안은 1941년 11월 28일 임시정부 국무회의에서 수정을 거쳐 “3·1 운동을 통해 대한민국을 건립한 독립정신을 계승한다.”라고 규정하고 있다. 임시정부에서 작성한 헌법 초안을 바탕으로 그 역사적 정통성을 이어받은 대한민국은 1948년 7월 17일 제헌 헌법을 제정하였다.(문화재청고시 제2018-182호, 관보 제19410 호, 2018.12.31.)

2) “나는 헌법을 준수하고 국가를 보위하며 조국의 평화적 통일과 국민의 자유와 복리의 증진 및 민족문화의 창달에 노력하여 대통령으로서의 직책을 성실히 수행할 것을 국민 앞에 엄숙히 선서합니다.”(헌법 제69조)

터 기반의 객관적인 결과를 탐구하는 데에 의의가 있다.³⁾ 지금까지 대통령 연설문과 관련한 연구를 아래와 같이 요약할 수 있다.

<표 1> 대통령 연설문 관련 연구 목록

분야	내 용
사회 과학 분야	강태완 외(2004) 역대, 수사적 특징, 역할 규정 조수곤 외(2015) 미국 대통령 연설문, 텍스트 마이닝 트렌드 분석 홍주현 외(2014) 대통령 후보자의 연설문 네트워크 분석 곽송연(2018) 민주주의 담론 변화 연구 정의룡(2018) 대통령 지시사항에 대한 의미 연결망 분석
정치 분야	이귀혜(2007) 설득 수사학 이서영(2011) 설득적 스피치의 언어적 메시지 구현 방법 연구 권향원 외(2013) 언어적 상징 전략 박영환(2013) 레토릭과 지도도 연구 황창호 외(2015) 역대 대통령 연설문의 유형과 특징 비교
과학 분야	박영재 외(2017) Word2Vec을 이용한 연설문 네트워크 분석(물리)
국어 국문학 분야	김현국(2001) 연설문의 문체 연구 이정복(2003) 연설문의 경어법 분석 이창수(2008) 노무현, 연설문, 정체성 연구 김병홍(2017) 대통령 취임사의 언어 특성 분석 조원형(2017) 대통령 과거사 관련 연설 텍스트 언어학적 분석 김재희(2018) 사회언어학적 의미 연구

지금까지 대통령 연설문에 관한 연구는 사회과학, 정치 분야에 국한되었거나, 국어국문학 분야에서 문체, 높임법, 언어 특성에 관한 연구가 진행되었다. 또한 대통령 연설문 중에서 일부 대통령을 연구하거나 유형을 제한한 연구가 진행됐다.

2. 연구 방법 및 설계

1) 연구 대상

- 3) 텍스트 데이터로부터 의미 있는 정보의 추출 및 분석을 위한 텍스트 마이닝(Text Mining, Chakraborty et al., 2013)은 문서 요약(Kim et al., 2013), 정보검색(Pai et al., 2013), 감성분석(Liu, 2012), 트렌드 분석(Hung and Zhang, 2012), 군집분석(Aggarwal and Zhai, 2012) 및 분류분석(Chen and Chen, 2011) 등 다양한 주제의 연구로 심화하는 현상을 보인다(조수곤 외, 2015: 453).

이 연구에서는 1948년 초대 이승만 대통령부터 2017년 제19대 문재인 대통령까지 71년간, 취임사 19편을 대상으로 한다.⁴⁾

<표 2> 대통령별 기록물 대상 자료

구분	대통령	재임 기간	재직 개월	연설문 자료
1~3대	이승만	1948. 7~1960. 4	11년 10개월	998건
4대	윤보선	1960. 8~1962. 3	1년 7개월	3건
5~9대	박정희	1962. 3~1979.10	15년 10개월	1,270건
10대	최규하	1979.10~1980. 8	8개월	58건
11~12대	전두환	1980. 9~1988. 2	7년 5개월	602건
13대	노태우	1988. 2~1993. 2	5년	601건
14대	김영삼	1993. 2~1998. 2	5년	728건
15대	김대중	1998. 2~2003. 2	5년	822건
16대	노무현	2003. 2~2008. 2	5년	780건
17대	이명박	2008. 2~2013. 2	5년	819건
18대	박근혜	2013. 2~2017.3	4년	- ⁵⁾
19대	문재인	2017. 5~2022.5	5년	-
계				6,681건

대통령 취임사는 대통령으로서의 첫 임무를 시작할 때 국민에게 자신의 취임 목표를 밝히는 것으로 대통령의 국정운영 철학과 미래를 살펴볼 수 있는 소중한 자료이다. 이 연구는 행정부 수반의 기록 중에서 취임사의 어휘 분석을 통해 키워드를 파악하고자 한다.⁶⁾

4) 이 연구의 분석 대상 자료는 ‘행정안전부 국가기록원 대통령기록관’에서 제공하고 있다.(<https://www.pa.go.kr/research/contents/speech/index.jsp>)

이 연구에서는 대통령 연설문 6,681건의 일부인 취임사 19편만을 대상으로 하며, 다른 연설문은 차기 연구에서 진행하기로 한다.

5) 18대 ~ 19대 대통령의 연설문 자료는 대통령기록관에 자료 이전이 완료되는 대로 다음 연구에서 더 밝히기로 한다.

6) 취임사 어휘 분석의 기준은 다음과 같다.

가. 어휘는 명사를 추출한다. 단, 복합명사는 문맥을 해치지 않기 위해 단어어, 파생어의 구분 없이 분석하였다.

나. 의미 파악이 어려운 의존명사, 대명사는 제외한다. 단, ‘저, 제, 여러분, 자기, 우리, 본인’ 등 자신을 지칭할 때 예외로 한다.

다. 접두사 및 접미사는 단어어로 처리한다.

한편, 취임사에 나타난 빈도수 제시는 지면의 한계로 상위 10~20개 내외로 제시한다.

특히 이 연구의 중심은 특성값을 통해 이루어진다. 특성값은 특성추출에 기반해 계산된 중요도 값이다. 특성추출이란 자료에 등장하는 명사 중에서 해당 자료에서 중요하다고 판단하여 추출한 키워드이다. 이때 중요도는 ‘텍스트 랭크(Text Rank)’ 알고리즘으로 자동 판단된다. 텍스트 랭크 알고리즘은 단어 간 연결망을 그려서 중심이 되는 단어를 찾는 알고리즘이다. 특정 문서에서 같이 사용된 단어 간에 연결망을 그리고, 그 연결망에서 다른 단어들과 많이 연결될수록 중요한 단어라고 판단된다.⁷⁾

2) 텍스트 마이닝 연구 방법

본 연구의 방법으로는 텍스트 마이닝 기법을 활용한다. 텍스트 마이닝은 텍스트 데이터를 대상으로 정보를 추출하는 과정으로 수치 데이터의 정보를 추출하는 데이터 마이닝과는 차이가 있다. 흔히 문서 색인으로 정의할 수 있는데, 정보검색에서 사용되는 텍스트를 구조적인 데이터로 표현하는 기법을 이용한다. 정보검색에서는 사용자에게 더욱 신속하고 정확한 텍스트를 조회하기 위한 텍스트 마이닝은 자연어 처리 분야에서 문장 단위의 처리 기법을 이용하여 의미를 분석하는 것이다(조태호, 2001:77).

지금까지 텍스트 마이닝은 외국어를 중심으로 이루어져 왔다. 한국어는 형태소 분석의 난해함으로 완벽한 문장을 분석하기 어려웠다. 한국어를 분석할 수 있는 분석기의 개발로 한국어를 분석하는 텍스트 마이닝 연구가 가능해졌다. 최근에는 여러 도구의 개발로 원문 자료를 바로 형태소 분석, 주제어 추출 기법으로 바로 연구할 수 있다.⁸⁾

대통령 취임사 자료를 컴퓨터가 이해할 수 있는 txt 또는 excel 파일로 전환해야 프로그램에서 인식할 수 있다. 또한 전처리 과정을 거쳐, ‘기호, 구두점, 불용어’ 등을 소거하여야 프로그램에서 처리할 수 있다. 본 연구에 사용되는 프로그램은 Knowledge Matrix Plus(KM+, KISTI(2016)),⁹⁾ BigKinds(한국언론진

7) 특성값과 특성추출에 대한 더욱 자세한 내용은 한국언론진흥재단, 빅카인즈를 참조할 수 있다.(<https://www.bigkinds.or.kr/v2/analysis/featureExtraction.do>).

8) 조용림, 「텍스트 마이닝을 활용한 글쓰기 자료 분석」, 『문화와 융합』42-5, 한국문화융합학회, 2020. 280-281쪽.

9) KM+는 Knowledge Matrix Plus(KM+) ver.0.80 for supporting Scientometric Network Analysis, Department of Scientometric Research, Korea Institute of

흥재단)¹⁰⁾ 구글 워드 클라우드¹¹⁾를 사용하였다. 따라서 이 연구에서는 ‘KM+’, ‘빅카인즈’의 형태소 분석을 사용하여 상호 텍스트 검증을 하고 ‘어휘 빈도’를 파악한다. ‘어휘 빈도’와 함께 핵심 키워드를 추출하고, 구글 ‘워드 클라우드’를 사용하여 키워드 이미지를 제시하고, ‘빅카인즈’의 특성값을 통해 그 의미를 분석하고자 한다.

3. 대통령별 취임사 어휘 분석

1) 1대 ~ 4대(이승만, 윤보선) 대통령의 취임사 어휘

<표 3> 1대 ~ 4대(이승만, 윤보선) 대통령의 취임사 어휘

순위	1대	빈도	2대	빈도	3대	빈도	4대	빈도
1	나	9	세계	15	농민	15	경제	11
2	나라	7	자유	14	나라	12	국민	11
3	자리	7	나라	13	가지	9	외교	9
4	말	6	사람	9	경제	9	정치	9
5	사람	6	전쟁	9	사람	9	자유	6
6	동포	4	목적	8	생산	9	정부	6
7	눈물	4	문제	8	우리나라	8	행정	6
8	진행	4	원수	8	공산당	7	국가	5
9	일	4	주의	8	성공	6	정책	5
10	대우	4	민주	7	국민	5	책임	5
11			승전	7	진전	5	독재	4
12							민주	4
13							4월혁명	4

Science and Technology Information으로 KISTI(한국과학기술정보연구원) 과학계량연구실에서 제공된 것이다.

10) 한국언론진흥재단 ‘빅카인즈’(<https://www.bigkinds.or.kr>)에서 제공하는 ‘형태소 분석’, ‘특성추출’ 분석기를 이용한다.

11) 구글 클라우드(<https://cloud.google.com/>)를 이용한다.



1대~4대 대통령의 취임사 어휘의 빈도수이다. 1대~3대는 이승만 대통령, 4대는 윤보선의 대통령의 취임사이다. 1대 이승만 대통령은 ‘나’, 2대에는 ‘세계’, 3대에는 ‘농민’이 가장 많이 사용된 것으로 나타났다. 4대 윤보선 대통령은 ‘경제’와 ‘국민’이 가장 높은 빈도를 보인다. 워드 클라우드에서 그 내용을 시각화한 것이다.

<표 4> 1대~4대 대통령 취임사의 특성값

	1대 키워드	특성값	2대 키워드	특성값	3대 키워드	특성값	4대 키워드	특성값
1	눈물	0.00846	문제	0.0055	우리나라	0.01370	외교 정책	0.01453
2	대한민국	0.00755	우리나라	0.0051	농민	0.00879	국민 경제	0.01083
3	국회의원	0.00663	사람	0.0050	국민	0.00874	정권 외교	0.01046
4	대우	0.00544	원수	0.0038	사람	0.00853	정치적 자유	0.00848
5	우방	0.00506	전쟁	0.0036	일본	0.00809	경제적 자유	0.00827
6	애국남여	0.00502	세계	0.0035	한국	0.00755	제2공화국	0.00709
7	영광	0.00500	목적	0.0033	미국	0.00670	경제적 위기	0.00704
8	친선	0.00494	승전	0.0029	공산당	0.00656	책임감	0.00652
9	신성	0.00474	자유	0.0028	경제	0.00647	국민	0.00615
10	의장	0.00435	결심	0.0027	국제	0.00543	외교 활동	0.00599

11	협약의 진행	0.0043	제국주의	0.0027	생산력	0.00541	동북아	0.00557
12	과거 40년	0.00421	직책	0.0026	환산율	0.00495	국가 운명	0.00554
13	부의장	0.00421	실패	0.0026	동포	0.00493	민주주의	0.00538
14	공산당	0.00419	통일	0.0026	쌀값	0.00492	사월혁명	0.00524
15	자리	0.00418	민중	0.0024	농림부	0.00487		

<표 4>는 1대 ~ 4대 대통령 취임사의 ‘특성값’이다.¹²⁾ 1대 취임사의 특성값에서는 ‘눈물’ 키워드가 높게 나타났다. ‘눈물’은 총 4회 나타나는데¹³⁾ 임시정부를 잇고 대한민국의 초대 대통령으로서 건국 상황에서 힘겨웠던 상황을 ‘눈물’이라는 단어를 통해 제시하여, 그 특성값이 높게 나타난 결과로 보인다. 특성값이 높은 상위 키워드를 살펴보면, ‘대한민국, 애국남여, 국회의원, 의장, 부의장’에서 국내 정치와 관련한 키워드가 등장하며, ‘대우, 우방, 애국남여, 친선, 협의, 공산당’ 등 다른 나라와의 관계를 중시한 키워드가 특성값이 높게 나타났다.

2대 취임사의 특성값에서는 ‘문제’ 키워드가 높게 나타났다. ‘문제’는 총 8회 쓰였는데, 한국전쟁과 관련한 부분에서 3회, 국내 정치와 세계 정세와 관련한 부분에서 5회가 나타난다. 즉, 전쟁 상황에서 2대 대통령에 취임하게 되면서 여러 상황을 ‘문제’로 인식하여 특성값이 높게 나타난 것으로 보인다. 특성값이 높은 상위 키워드에는 ‘원수, 전쟁, 세계, 목적, 승전, 자유, 통일’ 등 한국전쟁과 관련한 키워드가 등장하고 있다.

3대 취임사의 특성값에서는 ‘우리나라’ 키워드가 높게 나타났다. ‘우리나라’는 총 8회 나타나는데, ‘국민에게 불공평’, ‘공산당’, ‘난동과 혁명운동’, ‘내정’, ‘주권’, ‘교육’, ‘생산력’, ‘농민’과 관련하여 국가의 중대한 의제를 제시하는 부분에서 ‘우리나라’가 사용되고 있다. 특성값이 높은 상위 키워드에는 ‘농민, 국민, 사람, 경제, 생산력, 쌀값, 농림부’ 등 한국전쟁 직후에 전쟁으로 인한 먹거리의 중요성과 궁핍한 삶을 이겨내기 위한 키워드가 나타났으며, ‘일본, 한

12) 이하의 특성값은 소수점 이하 4~5자리 이상의 값만 제시하기로 한다.

13) 기쁨이 極하면 우습으로 변하여 눈물이 된다는 것을 <중략> 나의 致賀하는 남여동포가 모다 눈물을 씻으며 <중략> 축전 오는 것을 보면 모다 눈물을 금하기 어렵다 <중략> 木石肝膽이 아닌만치 뼈에 맺히는 눈물을 금하기 어렵습니다.<1대 대통령 취임사에서>

국, 미국, 국제, 환산율, 동포’ 등 전쟁 이후의 국제 정치가 중요하다는 사실이 특성값에 나타나고 있다.

4대 취임사의 특성값에서는 ‘외교 정책’ 키워드가 높게 나타났다. 4대 윤보선 대통령은 4.19 혁명 후 대통령에 당선되어 의원내각제의 제2공화국의 대통령이 되었다. 이전 정부의 비민주적 태도를 지적하면서 주변 강대국과의 외교에 중점을 두고 평화를 위한 활동을 다짐하면서 ‘정권 외교, 외교 활동’ 등의 키워드가 함께 나타나고 있다. 상위 키워드에는 ‘국민 경제, 경제적 위기’ 등의 피해를 입은 경제에 대한 경각심을 높이고자 하는 키워드가 나타나고 있으며, ‘자유, 책임감, 민주주의, 사월혁명’ 등의 부정과 부패를 딛고 일어서고자 하는 의지를 나타내는 키워드가 상위 특성값에 나타났다. 그러나 1960년 8월에 취임한 뒤 이듬해 5.16 군사쿠데타를 맞아 약 1년간 대통령직을 유지하였으나 사실상 권한이 없는 대통령으로 결국 1962년 3월 대통령직에서 하야하여 취임사를 실천할 시간이 없었던 것으로 보인다.

2) 5대 ~ 9대(박정희) 대통령의 시대별 취임사 어휘

<표 5> 5대 ~ 9대 대통령 취임사의 어휘 빈도

순위	5대	빈도	6대	빈도	7대	빈도	8대	빈도	9대	빈도
1	사회	25	민주	11	평화	12	민족	20	국민	22
2	정치	23	사회	9	국민	11	유신	17	민족	17
3	국민	16	경제	8	사회	8	국민	15	사회	12
4	민족	16	국민	8	조국	8	번영	15	자주	10
5	민주	14	민족	8	통일	8	역사	13	고도 산업	9
6	공화국	9	민주주의	8	민족	7	조국	12	세계	8
7	역사	9	통일	8	민주	7	통일	11	역사	8
8	혁명	9	공업	7	발전	6	국력	7	평화	7
9	대혁신	8	조국	7	국력	5	평화	7	국가	6
10	조국	8	건설	6	근대화	5			국력	6
11	혁신	8	근대화	6	세계	5			바탕	6
12	협조	8	동포	6					통일	6



<표 5>와 <그림 5 ~ 9>는 5대 ~ 9대 대통령의 취임사 어휘 빈도수이다. 5대에는 ‘사회’, 6대에는 ‘민주’, 7대에는 ‘평화’, 8대에는 ‘민족’, 9대에는 ‘국민’이 가장 많이 사용된 것으로 나타났다. 특히 자기 자신을 나타내는 대명사로 ‘본인’이 2회 사용되면서 처음 등장한다.

박정희 대통령의 재임 기간은 15년 10개월로 역대 대통령 중 재임 기간이 가장 길다. 1961년 5.16 군사쿠데타로 의원내각제 체제의 제2공화국을 전복하였다. 이후 군사혁명위원회 의장, 내각 수반, 국가재건최고회의 의장, 대통령 권한대행을 맡아 대통령제로 개헌을 하였고, 1963년 12월 17일 제5대 대통령에 취임하며 제3공화국이 출범하게 되었다. 1967년 7월 1일 제6대 대통령 재임하였고, 1969년 3선 개헌안을 통과시켜 1971년 7월 1일 제7대 대통령에 취

임하였다.¹⁴⁾ 1972년 10월 유신 헌법을 선포하여 6년씩 제한 없는 간선 대통령 제에서 1972년 12월 27일 8대 대통령에 취임하면서 4공화국이 출범하였다. 이후 1978년 12월 27일 역시 간선으로 9대 대통령에 취임하게 되었다.

<표 6> 5대 ~ 7대 대통령 취임사의 특성값

	5대 키워드	특성값	6대 키워드	특성값	7대 키워드	특성값
1	공화국	0.00683	민주주의	0.01075	통일 조국	0.01144
2	정치 사회적	0.00677	공산주의	0.00608	평화 통일	0.00967
3	민주주의	0.00606	근대화	0.00602	통일 국력	0.00943
4	국민	0.00538	아시아	0.00514	한국 국민	0.00801
5	정치적 대립	0.00429	공업입국	0.00406	민주 발전	0.00676
6	공화국 건설	0.00427	경제발전	0.00390	국민 생활	0.00661
7	민주주의 전도	0.00342	민주사회	0.00379	국가 발전	0.00611
8	정치적 역량	0.00339	대한민국	0.00377	문화 한국	0.00609
9	정치적 자주	0.00335	창조력	0.00365	자유 민주	0.00566
10	영광	0.00314	전근대적 요소	0.00349	세계 평화	0.00550
11	정치 사회적 불안정	0.00295	복지사회	0.00347	조국 근대화	0.00532
12	공화국 탄생	0.00285	역사적 대업	0.00336	인류 사회	0.00519
13	정신적 혁명	0.00256	경제적 자립	0.00307	자유 이념	0.00512
14	민주적 역량	0.00253	경제건설	0.00298	생활 이념	0.00473
15	반세기	0.00241	통일조국	0.00295	민주 사회	0.00465

5대 취임사의 특성값에서는 ‘공화국’이 가장 높게 나타났다. 의원내각제에서 대통령제로 전환된 것이 특성값을 높인 것으로 보인다. 이후에 나타나는 ‘정치 사회, 민주주의, 정치적 대립, 공화국 건설’ 등도 특성값이 높게 나타나는데 대통령제로의 전환 어휘로 볼 수 있다. 이 외에 ‘국민, 영광, 정신적 혁명, 반세기’ 등의 어휘가 보이는데 5.16 군사쿠데타를 군사혁명이라 칭하며, ‘국민 개개인의 정신적 혁명’을 함께 요구하고 있는 것으로 보인다. 특이한 어휘로 ‘복지’ 어휘가 대통령 취임사에 처음 등장한다.

6대 취임사의 특성값에서는 ‘민주주의’가 가장 높게 나타났다. ‘공산주의,

14) 당시 헌법에서는 4년씩 3선까지 할 수 있었기에 7대 대통령의 임기는 1975년 6월 30일까지였지만 10월 유신으로 1년 5개월 만에 막을 내렸다.

민주사회, 역사적 대업, 통일조국’ 등 북한과의 사상적 대결, 국내 정치의 안정을 바라는 어휘가 주요 등장한 것으로 보인다. 이후 ‘근대화, 공업입국, 경제발전, 창조력, 경제적 자립, 경제 건설’ 등 경제 성장을 위한 경제개발 2차 5개년 계획을 주창하여 특성값이 높게 나온 것으로 볼 수 있다. 이 특성값을 반영하는 “우리의 적은 빈곤과 부정·부패와 공산주의입니다. 나는 이것을 우리의 삼대 공적으로 생각합니다.”라는 문장을 통해 6대 취임사의 핵심을 알 수 있다.

7대 취임사의 특성값에는 ‘통일 조국’이 가장 높게 나타났다. 이와 함께 ‘평화 통일, 통일 국력, 세계 평화, 자유 이념, 평화 통일 제의’ 등 남북문제 관련 어휘가 특성값이 높게 나타났다. 주요 어휘로 ‘민주 발전, 국민 생활, 국가 발전, 생활 이념’, ‘문화 한국, 인류 사회’ 등의 어휘가 나타나면서 국민 생활 안정 등에 집중하는 모습을 보이는 것으로 보인다.

<표 7> 8대 ~ 9대 대통령 취임사의 특성값

	8대 키워드	특성값	9대 키워드	특성값
1	민족 통일	0.01043	국민 생활	0.00836
2	통일 조국	0.00911	민족 문화	0.00691
3	민족 문화	0.00743	통일 조국	0.00673
4	유신 질서	0.00711	고도 산업 사회	0.00661
5	국력 배양	0.00702	국가 발전	0.00614
6	역사 발전	0.00693	민족사	0.00590
7	유신 이념	0.00688	평화 통일	0.00584
8	사회 보장 제도	0.00612	민족 중흥	0.00531
9	사회 복지	0.00587	국민 정신	0.00513
10	정치 제도	0.00576	역사 창조	0.00509
11	정신 문화	0.00543	국민 정신과 사회 기강	0.00481
12	국가 발전	0.00538	국제 경제	0.00479
13	유신 과업	0.00508	민족 웅비	0.00473
14	사회 기풍	0.00506	자립 경제	0.00429
15	새마을 정신	0.00466	방위 산업	0.00423

8대 취임사의 특성값에는 ‘민족 통일’이 가장 높게 나타났다. 분단 27년 동안 단절된 남북문제에 대한 언급이 ‘통일 조국’ 어휘와 함께 특성값이 높게 나타난 것으로 보인다. 이후 ‘유신 질서, 유신 이념, 유신 과업’의 어휘는 1972년

10월 유신 이후에 새로운 헌법 질서를 주창하고 그 정신을 펼치고자 하는 의지를 보인 어휘가 특성값이 높게 나타난 것으로 보인다. ‘사회 보장 제도, 사회 복지, 새마을 정신’ 등 경제발전을 확인한 다음 5대 취임사에 처음 등장한 ‘복지’를 내세움으로써 그 자신감을 ‘복지’와 ‘새마을’이라는 어휘에 담고 있으며, 국민의 정신 개혁을 이루려고 하는 의지를 취임사에서 제시하고 있다.

9대 취임사의 특성값에는 ‘국민 생활’이 가장 높게 나타났다. 이전부터 선언했던 경제 개발 계획에 따라 경공업 중심에서 중화학 공업 국가를 선언했던 내용들이 국민의 노력으로 경제에 대한 자신감을 얻게 된 결과로 보인다. ‘고도 산업 사회, 국가 발전, 국민 정신(과 사회 기강), 자립 경제’ 등이 ‘국민 생활’과 연관되어 나타나고 있다. 더불어 ‘통일 조국, 평화 통일, 민족 중흥, 민족 웅비’ 등의 어휘는 ‘통일’을 중시하려는 의지의 표현으로 앞선 취임사에서부터 나타났던 어휘들이 지속해서 보인다.

3) 10대 ~ 13대(최규하, 전두환, 노태우) 대통령의 취임사 어휘

<표 8> 10대 ~ 13대 대통령 취임사의 어휘 빈도

순위	10대 키워드	빈도수	11대 키워드	빈도수	12대 키워드	빈도수	13대 키워드	빈도수
1	국민	36	국민	30	본인	31	국민	48
2	헌법	21	민주	22	국민	29	시대	23
3	국가	21	사회	21	민족	19	민족	20
4	경제	19	정치	21	시대	17	민주	20
5	안정	18	정부	16	역사	17	사람	17
6	본인	17	국가	14	국가	17	나라	13
7	사회	15	발전	12	해방	11	세계	12
8	발전	12	복지	12	생활	10	국가	12
9	정부	11	본인	11	정부	10	통일	11
10	정치적	8	경제	11	북한	10	정부	11
11	국제	8	민주주의	10	사회	9	사회	10
12	개정	8					민 의	10



<표 8>과 <그림 10 ~ 13>은 10대 ~ 13대 대통령의 취임사 어휘 빈도수이다. 10대 ~ 11대에는 ‘국민’, 12대에는 ‘본인’, 13대에는 다시 ‘국민’이 빈도수가 높은 어휘로 나타났다.

<표 9> 10대 ~ 13대 대통령 취임사의 특성값

순위	10대 키워드	특성값	11대 키워드	특성값	12대 키워드	특성값	13대 키워드	특성값
1	본인	0.00738	민주주의	0.00654	본인	0.00427	민주주의 시대	0.00965
2	국민 각계각층	0.00457	유지 발전	0.00418	북한주민	0.00408	민주주의	0.00806
3	나라 경제	0.00435	가치관	0.00360	국민	0.00384	국가 발전	0.00515
4	정치적 발전	0.00408	새마을 운동	0.00350	민주주의	0.00375	국민	0.00459
5	국가적 현실	0.00329	대통령 자신	0.00329	국가안보	0.00373	정직	0.00388
6	사회적 변화	0.00293	민족복지 국가 건설	0.00318	식민통치 36년	0.00368	민족자존	0.00356
7	헌법개정	0.00280	자주국방	0.00289	제5공화국 헌법	0.00338	서울올림픽	0.00292
8	국가적 난구	0.00254	근로자	0.00288	권력남용	0.00331	국민화합	0.00292
9	현행 헌법	0.00246	정치풍토	0.00266	물질적 빈곤	0.00290	동아시아	0.00287

10	국민생활	0.00225	평화적 해결	0.00253	평화통일	0.00283	20세기	0.00280
11	사회적 불안	0.00224	경제발전	0.00251	구헌법	0.00282	민족통일	0.00279
12	군사적 조치	0.00218	국제경쟁력	0.00250	민족사	0.00275	고도성장	0.00267
13	국민적 합의	0.00215	평화적 정권 이양	0.00245	북한당국	0.00264	보통사람들	0.00264
14	주변 정세	0.00212	제11대 대통령	0.00232	정부 고위관리	0.00262	행복한 사회	0.00262
15	북한공산 집단	0.00211	북한공산 집단	0.00226	일본	0.00262	진취적 사회	0.00256

10대 취임사 특성값은 ‘본인’으로 나타났다. ‘10.26’사태 후 2달 여만의 제4 공화국 헌법에 따라 1979년 12월 6일 대통령으로 취임하였고, 12월 21일에 발표된 취임사에서 ‘정치, 국가, 사회, 헌법, 불안, 합의’ 등 ‘위기관리정부’의 책임을 제시하고 있다. 대통령에 오를 수 있었던 이유도 당시 헌법에 ‘대통령 궐위 시 3개월 이내 후임자를 선출한다.’라는 내용에 따라 차기 정권에 권력을 이양하는 바를 천명하여¹⁵⁾ 대통령에 선출될 수 있었고, 재임 8개월 만인 1980년 8월 16일에 대통령직에서 사임하였다.

11대 취임사 특성값은 ‘민주주의’로 나타났다. 11대 대통령 취임 전 그는 1979년 12.12 군사 반란, 1980년 5.17 내란으로 권력을 장악한 후 5.18 광주민주화운동을 무력 진압하였다. 이후 통일주체국민회의에서 치러진 대통령 선거에서 당선되어 1980년 9월 1일 대통령에 취임하였다. 11대 취임사에서 ‘민주주의’가 특성값이 높게 나타났으며, 1. 우리 정치풍토에 맞는 민주주의 2. 진정한 복지사회 이룩, 3. 정의로운 사회 구현, 4. 교육혁신과 문화창달로 국민정신 개조를 국정 목표로 삼아 정치(평화적 정권 이양 등), 경제(새마을 운동, 유지발전 등), 남북문제(자주국방, 평화적 해결, 북한 공산집단 등)와 관련한 어휘들이 특성값이 높게 나타났다.

12대 취임사에서 특성값이 높게 나타난 어휘는 ‘본인’이다. 5대 취임사에서 처음 등장한 ‘본인’은 대통령이 자신을 나타내는 대명사로 자연스럽게 자리하

15) 1979년 11월 6일, 유신 헌법에 따라 새 대통령 선출 및 새 대통령의 헌법을 개정한다는 ‘시국에 관한 담화’를 발표하였다. 서중석, 『한국현대사 60년』, 역사비평사, 2007, 159쪽.

고 있다. 7년 단임제의 새 헌법을 통과시킨 후 11대 대통령직을 사임하고 제5 공화국을 출범시켰다. ‘본인’ 외에도 남북문제 관련한 ‘북한주민, 북한당국, 평화통일, 국가안보’ 등의 어휘, ‘민주주의, 제5공화국헌법, 권력남용, 구헌법’과 같은 정치 관련 어휘 등이 특성값이 높게 나타났다.

13대 취임사의 특성값은 ‘민주주의 시대’가 가장 높게 나타났다. 12대 대통령 임기 말 대통령 직선제 개헌을 요구하는 국민의 요구가 있었지만 4.13 호헌조치에 국민의 시위는 더욱 악화하였고, 결국 6월 항쟁이 일어나게 되었다. 6.10항쟁 후 대선후보였던 노태우 후보가 6.29선언으로 대통령 직선제 개헌을 받아들였고, 1988년 2월 25일 제13대 대통령으로 취임하였다. 취임 전후의 상황에서 ‘민주주의’라는 시대정신이 취임사에 반영된 것으로 보인다.

또한 ‘민주주의, 국민, 국회화합’ 등의 어휘가 특성값이 높게 나타나고 있으며, ‘보통사람들, 행복한 사회’를 중심으로 90년대를 맞이하는 희망을 취임사에서 보여주고 있는 것으로 보인다. 다음으로 ‘국가발전, 서울올림픽, 동아시아, 20세기, 고도성장’ 등의 어휘에서 경제발전의 의지를 나타내는 어휘가 주요 어휘로 등장한다.

4) 14대 ~ 16대(김영삼, 김대중, 노무현) 대통령의 취임사 어휘

<표 10> 14대 ~ 16대 대통령 취임사의 어휘 빈도

순위	14대 키워드	빈도수	15대 키워드	빈도수	16대 키워드	빈도수
1	국민	28	국민	39	국민	21
2	민족	15	경제	26	동북아	18
3	사회	12	정부	25	평화	17
4	정부	12	기업	20	시대	15
5	신한국	11	남북	14	세계	14
6	시대	10	나라	12	사회	13
7	세계	9	사회	12	한반도	12
8	정치	8	혁명	12	북한	10
9	사람	7	민주	11	동북아 시대	9
10	자유	6	협력	10	발전	9
11	통일	6	극복	10	경제	8
12	창조	6	문화	10		
13	민주	6				



14대 ~ 16대 취임사 어휘 빈도수이다. 여기에서는 모두 ‘국민’이 빈도수가 가장 높은 어휘로 나타났다. 특히 14대 ‘문민정부’, 15대 ‘국민의 정부’, 16대 ‘참여 정부’로 공식 명칭을 정한 바 있다.

<표 11> 14대 ~ 16대 대통령 취임사의 특성값

	14대 키워드	특성값	15대 키워드	특성값	16대 키워드	특성값
1	신한국	0.01275	21세기	0.00593	동북아 시대	0.01206
2	신한국 창조	0.00830	민주주의	0.00569	세계 경제	0.00768
3	부정부패	0.00585	중소기업	0.00396	21세기	0.00492
4	공동체	0.00578	21세기 외교	0.00367	지식정보화 시대	0.00437
5	민주주의	0.00568	시장경제	0.00358	외환위기	0.00425
6	근로자	0.00519	애국심	0.00350	한미동맹	0.00420
7	국가기강	0.00428	세계화 시대	0.00331	유럽연합	0.00398
8	신바람	0.00419	5대 혁명	0.00324	지속적 성장	0.00372
9	한민족 시대	0.00412	정권교체	0.00323	핵무기	0.00353
10	신제품	0.00390	정보화 시대	0.00316	정보화	0.00353
11	해외 동포	0.00389	경제발전	0.00314	국민	0.00326
12	봄날	0.00383	교류협력	0.00310	창의력	0.00324

13	한라산	0.0382	남북기본합의서	0.00304	중국	0.00324
14	남북한 동포	0.0382	대기업	0.00301	고이즈미	0.00321
15	국회의사당	0.0382	남북문제	0.00300	대외 경제	0.00319

14대 취임사의 어휘에는 ‘신한국, 신한국 창조’가 높게 나타났다. 약 30여 년간 군인 출신의 대통령이 집권한 것과 다르게 취임사 내에서도 ‘문민 정부’를 강조하고 있는데, 그와 결을 같이하면서 ‘신한국 창조’를 주창하고 있다. 이전 정부의 ‘부정부패’를 청산하고, ‘신바람, 공동체, 민주주의, 근로자’를 통해 새로운 대한민국 건설을 이룩하려는 내용을 담고 있다. 또한 ‘한민족 시대, 봄날, 한라산, 남북한 동포’ 등의 어휘 등은 남북 평화와 통일의 가치를 포함하는 것으로 보인다.

15대 취임사의 어휘에는 ‘21세기’가 특성값이 높게 나타났다. 취임 전 IMF 외환위기와 역사상 처음으로 여야 정권교체 정권이라는 특수성을 지니고 있다. 15대 취임사 초반에 “정부수립 50년 만에 처음 이루어진 여야 간 정권교체, 국민의 정부, 인간혁명, 농업혁명, 도시혁명, 사상혁명, 산업혁명의 5대 혁명”, “정보화 혁명, 세계경제시대로의 전환”으로 시작하고 있다. ‘21세기’와 연관된 어휘들로 인터넷 등 정보화 혁명을 강조하는 정보통신이 강한 국가를 취임사에 담고 있다. “6.25 이후 최대의 국난이라고 할 수 있는 외환위기”를 언급하면서 ‘중소기업, 시장경제, 경제발전, 대기업’ 등의 어휘를 통해 외환위기 극복과 시장경제 발전 등의 어휘가 특성값이 높게 나타났으며, ‘21세기 외교, 세계화 시대, 남북기본합의서, 남북문제’ 등이 남북문제 해결을 위한 어휘들로 등장한다. 이러한 의지가 2000년에 분단 이후 최초 남북정상회담이 성사되는 결과를 낳았다고 볼 수 있다.

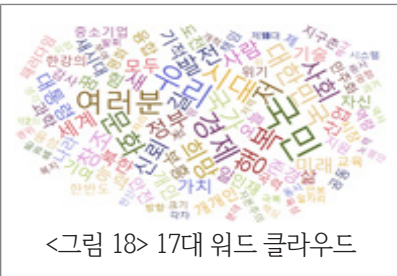
16대 취임사의 어휘에는 ‘동북아 시대’가 특성값이 높게 나타났다. ‘21세기, 한미동맹, 유럽연합, 핵무기, 중국, 고이즈미’의 어휘들과 함께 국정 목표 중 하나인 ‘평화와 번영의 동북아 시대’를 이루기 위한 외교 안보 등 한반도 평화 체제를 전 정부에서부터 연속하여 구축하려는 의지로 보인다. 더불어 ‘동북아 시대’를 열기 위해 ‘원칙과 신뢰, 공정과 투명, 대화와 타협, 분권과 자율’이라는 국정운영의 좌표를 제시하였다. 이후 ‘세계 경제, 외환위기, 지속적 성장, 대외 경제’ 등의 어휘가 등장하면서 경제적인 관점의 내용이 중요하게 다루어

졌다. 특히 취임사 초반에 ‘대구지하철 참사’를 언급하면서 ‘안전 사회를 만들도록 최선을’을 다한다는 언급으로 취임사를 시작하고 있다.

5) 17대 ~ 19대(이명박, 박근혜, 문재인) 대통령의 취임사 어휘

<표 12> 17대 ~ 19대 대통령 취임사의 어휘 빈도

순위	17대 키워드	빈도수	18대 키워드	빈도수	19대 키워드	빈도수
1	국민	30	국민	58	대통령	33
2	나라	23	시대	23	국민	24
3	사회	22	행복	21	대한민국	11
4	대한민국	17	문화	20	나라	10
5	세계	17	경제	20	역사	7
6	정부	15	대한민국	12	선거	6
7	문화	15	사회	12	정치	6
8	기업	14	국가	12	세상	5
9	선진	14	희망	10	시작	5
10	국가	13	개인	10	존경	5
11	경제	11	창조	10		
12	존경	11				



<표 12> 와 <그림 17 ~ 19>는 17대 ~ 19대 취임사 어휘 빈도수이다. 17대 ~ 18대에서도 ‘국민’이 빈도수가 가장 높은 어휘로 나타났으며, 19대 취임사에서 는 ‘대통령’이 가장 높은 빈도를 보인다.

<표 13> 17대 ~ 19대 대통령 취임사의 특성값

	17대 키워드	특성값	18대 키워드	특성값	19대 키워드	특성값
1	대한민국 국민	0.00968	국민행복 시대	0.01655	대한민국 대통령	0.02653
2	대한민국	0.00858	대한민국	0.01302	대한민국	0.02061
3	대한민국 선진화	0.00622	국가 발전	0.01031	대통령 선거	0.01880
4	대한민국 대통령	0.00589	창조경제	0.00970	광화문 대통령 시대	0.01510
5	국가 발전	0.00488	지구촌	0.00674	한반도 평화	0.00720
6	일자리	0.00428	글로벌 경제	0.00586	권위적 대통령	0.00652
7	문화 수준	0.00355	경제부흥	0.00519	사드	0.00626
8	국제사회	0.00334	문화융성	0.00510	전임 대통령	0.00626
9	과학기술	0.00329	행복한 나라	0.00503	현직 대통령	0.00619
10	평화 통일	0.00329	중소기업	0.00474	동반자	0.00599
11	지구 환경 변화	0.00316	과학기술	0.00440	동북아 평화	0.00583
12	경쟁력	0.00283	한강	0.00417	사드 문제 해결	0.00555
13	생활 형편	0.00269	국제사회	0.00392	정치 발전	0.00553
14	동맹관계	0.00267	일자리	0.00391	세대 간 갈등	0.00525
15	글로벌 외교	0.00251	글로벌 경제 위기	0.00388	비정규직	0.00505

17대 취임사의 어휘에는 ‘대한민국 국민’이 특성값이 높게 나타났다. 더불어 ‘대한민국, 대한민국 선진화, 대한민국 대통령, 국가 발전, 일자리, 경쟁력, 생활 형편’ 등 국정 목표로 제시한 내용 중 ‘경제’와 관련한 어휘가 특성값이 높게 나타났다. 유사한 어휘도 보이는데 ‘문화 수준, 과학기술, 지구 환경 변화’ 등의 어휘에서도 문화 및 환경 산업을 통해 경제 성장 동력을 만들자는 의지를 보이면서 ‘경제’에 초점을 두는 취임사 내용으로 볼 수 있다.

18대 취임사의 어휘에는 ‘국민행복 시대’가 특성값이 높게 나타났다. ‘대한민국, 국가 발전, 창조경제, 경제부흥, 문화융성, 행복한 나라, 중소기업, 한강, 일자리, 글로벌 경제 위기’ 등이 경제를 발전시켜 ‘국민행복 시대’로 나아가고자 하는 의지를 보이는 어휘가 특성값이 높게 나타나고 있다.

19대 취임사의 어휘에는 ‘대한민국 대통령’이 특성값이 높게 나타났다. 이 어휘는 ‘대한민국, 대통령 선거, 광화문 대통령 시대, 권위적 대통령, 전임 대통령, 현직 대통령, 동반자, 정치 발전, 세대 간 갈등’ 등 특성값이 높은 다른 어휘들과 관계가 깊다. 전임 대통령이 탄핵으로 임기가 끝난 상태에서 19대 대통령에 취임하게 되면서 ‘국민의 나라, 정의로운 대한민국’을 구호로 5대 국정 목표를 발표했다. ‘국민이 주인인 정부’와 관련이 깊으며, ‘소통으로 통합하는 광화문 대통령’에 초점을 둔 취임사의 내용이 특성값이 높게 나타났다. 또한 권력의 분산을 위해 ‘광화문 대통령시대, 권위적 대통령’의 어휘가 등장하며, 분열과 갈등을 해소하기 위해 ‘세대 간 갈등, 비정규직’이 특성값이 높게 나타난 것으로 보인다. 취임사 초반에 “승자와 패자가 없는, 동반자, 정치 혼란, 통합의 국가”를 천명하였고, ‘한반도 평화, 사드, 동북아 평화, 사드 문제 해결’ 등의 어휘를 통해 평화와 번영의 한반도를 만들겠다는 국정전략과 관계 깊은 어휘들이 사용되었다. 또 많은 이들에게 “기회는 평등하고, 과정은 공정하며, 결과는 정의로울 것”이라는 취임사의 핵심 문장이 등장한다.

4. 끝맺는 말

지금까지 이 연구에서는 대한민국 역대 대통령의 취임사 어휘를 특성값 중심으로 파악하였다. 형태소 분석을 통해 단어를 정제한 후 어휘 빈도수와 특성값을 찾아내고, 워드 클라우드로 시각화하였다. 기존 어휘 빈도수를 통한 분석이 아닌 특성값과 워드 클라우드를 통해 특성을 대표하는 객관적 주제어를 추출할 수 있었다. 그 특성값을 제시하면 다음과 같으며, 시대별로 요약하는 것으로 결론을 대신하도록 한다.

<표 14> 대통령 취임사의 상위 어휘 빈도 및 특성값

	빈도	특성값
1대	나	눈물
2대	세계	문제
3대	농민	우리나라
4대	경제, 국민	외교 정책
5대	사회	공화국
6대	민주	민주주의

7대	평화	통일 조국
8대	민족	민족 통일
9대	국민	국민 생활
10대	국민	본인
11대	국민	민주주의
12대	본인	본인
13대	국민	민주주의 시대
14대	국민	신한국
15대	국민	21세기
16대	국민	동북아 시대
17대	국민	대한민국 국민
18대	국민	국민행복 시대
19대	대통령	대한민국 대통령

정부수립부터 1950년대까지는 이승만 대통령이 집권하던 시기이다. 1대 취임사의 특성값에서는 ‘눈물’ 키워드가 높게 나타났다. 임시정부를 잇고 대한민국의 초대 대통령으로서 건국 상황에서 힘겨웠던 상황이 ‘눈물’ 여회를 통해 특성값이 높게 나타난 것으로 보인다. 2대 취임사의 특성값에서는 ‘문제’ 키워드가 높게 나타났는데, 한국전쟁 관련, 국내 정치와 세계정세와 관련한 부분에서 ‘문제’가 쓰이며, 특성값이 높게 나타난 것으로 보인다. 3대 취임사의 특성값에서는 ‘우리나라’ 키워드가 높게 나타났다. ‘국민에게 불공평’, ‘공산당’, ‘난동과 혁명운동’, ‘내정’, ‘주권’, ‘교육’, ‘생산력’, ‘농민’과 관련하여 국가의 중대한 의제를 제시하는 부분에서 ‘우리나라’가 사용되고 있다.

1960 ~ 1970년대에는 윤보선, 박정희 대통령이 집권하던 시기이다. 4대 취임사의 특성값에서는 ‘외교 정책’ 키워드가 높게 나타났다. 4.19 혁명 후 대통령에 당선되어 의원내각제의 제2공화국의 대통령이 되었는데, 이전 정부의 비민주적 태도와 주변 강대국과의 외교에 중점을 두고 평화를 위한 활동을 다짐하면서 ‘정권 외교, 외교 활동’ 등의 키워드가 특성값이 높게 나타나고 있다. 5대 취임사의 특성값에서는 ‘공화국’이 가장 높게 나타났다. 의원내각제에서 대통령제로 전환된 것이 특성값을 높인 것으로 보인다. 6대 취임사의 특성값에서는 ‘민주주의’가 가장 높게 나타났다. ‘공산주의, 민주사회, 역사적 대업, 통일

조국’ 등 북한과의 사상적 대결, 국내 정치의 안정을 바라는 어휘가 중요하게 등장한 것으로 보인다. 7대 취임사의 특성값에는 ‘통일 조국’이 가장 높게 나타났다. 이와 함께 ‘평화 통일, 통일 국력, 세계 평화, 자유 이념, 평화 통일 제의’ 등 정치 사상적인 어휘가 특성값이 높게 나타났다. 8대 취임사의 특성값에는 ‘민족 통일’이 가장 높게 나타났다. 더불어 분단 27년 동안 단절된 남북문제에 대한 언급이 ‘통일 조국’ 어휘와 함께 특성값이 높게 나타난 것으로 보인다. 9대 취임사의 특성값에는 ‘국민 생활’이 가장 높게 나타났다. 이전부터 선언했던 경제 개발 계획에 따라 경공업 중심에서 중화학 공업 국가를 선언했던 내용들이 국민의 노력으로 경제에 대한 자신감을 얻게 된 결과로 보인다.

1970년대 말 ~ 1980년대까지는 최규하, 전두환, 노태우 대통령이 집권하던 시기였다. 10대 취임사 특성값은 ‘본인’으로 나타났다. ‘위기관리정부’의 책임을 제시하면서 차기 정권에 권력을 이양하는 취임사로 보아 주체성을 강조한 것으로 보인다. 11대 취임사에서는 ‘민주주의’가 특성값이 높게 나타났다. 12대 취임사 특성값은 ‘본인’으로 나타났다. 5대 취임사에서 처음 등장한 ‘본인’은 대통령이 자신을 나타내는 대명사로 자연스럽게 자리하였는데, 이후에는 나타나지 않고 있다. 13대 취임사의 특성값에는 ‘민주주의 시대’가 가장 높게 나타났다. 국민의 대통령 직선제 개헌을 요구하였고, 결국에는 대통령 직선제 개헌을 받아들인 상황에서 ‘민주주의’라는 국민의 정신이 취임사에 반영된 것으로 보인다.

1990 ~ 2000년대까지는 김영삼, 김대중, 노무현 대통령이 집권하던 시기였다. 14대 취임사의 어휘에는 ‘신한국, 신한국 창조’가 높게 나타났다. 새로운 문명정부, 새로운 시대를 강조하면서 ‘신한국 창조’가 특성값이 높게 나타났다. 15대 취임사의 어휘에는 ‘21세기’가 특성값이 높게 나타났다. “정부수립 50년 만에 처음 이루어진 여야 정권교체, 국민의 정부, 인간혁명, 농업혁명, 도시혁명, 사상혁명, 산업혁명의 5대 혁명”, “정보화 혁명, 세계경제시대로의 전환”을 제시하면 새로운 세기를 맞이하는 의지로 ‘21세기’가 특성값이 높게 나타난 것으로 보인다. 16대 취임사의 어휘에는 ‘동북아 시대’가 특성값이 높게 나타났다. ‘21세기, 한미동맹, 유럽연합, 핵무기, 중국, 고이즈미’의 어휘들이 ‘평화와 번영의 동북아 시대’를 이루고 남북 한반도 평화 체제를 구축하고자 하는 의지

가 특성값이 높게 나타난 것으로 보인다.

2000년대 후반부터 현재까지는 이명박, 박근혜, 문재인 대통령이 집권하던 시기였다. 17대 취임사의 어휘에는 ‘대한민국 국민’이 특성값이 높게 나타났다. 그 중 ‘대한민국 선진화’ 등 ‘경제’와 관련한 어휘가 특성값이 높게 나타났다. 18대 취임사의 어휘에는 ‘국민행복 시대’가 특성값이 높게 나타났다. ‘창조 경제, 경제부흥, 문화융성, 행복한 나라’ 등 경제를 발전시켜 ‘국민행복 시대’로 나아가고자 하는 의지를 취임사에 담아 특성값이 높게 나타난 것으로 보인다. 19대 취임사의 어휘에는 ‘대한민국 대통령’이 특성값이 높게 나타났다. 촛불 혁명으로 전직 대통령이 탄핵당하면서 갑작스럽게 대통령에 취임하였다. 따라서 ‘대통령 선거, 광화문 대통령 시대, 권위적 대통령, 전임 대통령, 현직 대통령, 동반자, 정치 발전, 세대 간 갈등’ 등 ‘국민이 주인인 정부’, ‘소통으로 통합하는 광화문 대통령’에 초점을 둔 취임사의 내용으로 특성값이 높게 나타난 것으로 보인다.

자연어는 인공어와 달리 컴퓨터가 직접 처리하기 곤란하여 이를 구조적인 데이터로 처리할 필요가 있다. 따라서 표본 텍스트에 단어의 가중치를 문서 내의 출현 빈도에 부여하고 이를 근거로 일부의 단어를 특성값 키워드로 추출할 수 있다. 텍스트 마이닝의 정보 필터링을 통해 취임사 자료를 입력하고, 주제어 관련 연관성에 따른 특성값의 결과를 얻을 수 있어서 단순 어휘 빈도수보다 비정형 텍스트의 질적 향상을 도모할 수 있었다.

■ 참고문헌

- 강태완, 김은정, 「역대 대통령의 연설문에 나타난 수사적 특징과 역할규정」, 『사회과학연구』30권 2호, 2004, 53-89쪽.
- 박송연, 「1980년대 이후 한국 대통령 연설문에 나타난 민주주의 담론 변화 연구」, 『OUGHTOPIA』Vol.33 No.2, 경희대 인류사회재건연구원, 2018, 35-74쪽.
- 권향원, 최도림, 「대통령의 언어적 상징 전략에 대한 연구」, 『정부학연구』19권 3호, 고려대학교 정부학연구소, 2013, 285-320쪽.
- 김병홍, 「대통령 취임사의 언어 특성 분석」, 『우리말연구』51집, 우리말학회, 2017, 161-185쪽.
- 김재희, 「대통령 연설문에 나타난 사회언어학적 의미 연구」, 『텍스트언어학』 44권, 한국텍스트언어학회, 2018, 1-33쪽.
- 김현국, 「연설문의 문체 연구 - 대통령 취임사를 중심으로」, 『청람어문교육』 23권, 청람어문교육학회, 2001, 243-297쪽.
- 길호현, 「텍스트마이닝을 위한 한국어 불용어 목록 연구」, 『우리말글』 제78집, 우리말글학회, 2018, 1-25쪽.
- 박영재, 김영빈, 정선영, 김영진, 손승우, 「Word2Vec을 이용한 한국 대통령 연설문 네트워크 분석」, 『새물리』Vol.67 No.5, 한국물리학회, 2017, 569-574쪽.
- 박영환, 「대통령의 레토릭과 한국 대통령의 지도도」, 『한국정치연구』 22권 3호, 서울대학교 한국정치연구소, 2013, 63-94쪽.
- 박준형, 유호선, 김태영, 한희정, 오효정, 「대통령 연설기록 내용분석을 통한 역대 대통령의 중심가치 연구」, 『한국기록관리학회지』17-2, 한국기록관리학회, 2017, 57-78쪽.
- 박진우, 「9/11 이후 부시 대통령 연설문 연구」, 인천대대학원 석론, 2007.
- 서진수, 『R라뷰 - R을 활용한 데이터 분석 입문편』, 더알음, 2015.
- 이귀혜, 「한국 대통령들의 설득 수사학」, 『한국소통학보』제8호, 한국소통학회, 2007, 276-312쪽.
- 이서영, 「설득적 스피치의 언어적 메시지 구현 방법에 관한 연구」, 『한국소통학회 가을철 정기학술대회 발표집』, 한국소통학회, 2011, 125-144쪽.

- 이슬기, 박영민, 「쓰기 수행 수준에 따른 중학생 논설문의 텍스트 시각화 분석」, 『학습자중심교과교육연구』17-15, 학습자중심교과교육연구, 2017, 401-422쪽.
- 이시연, 「버락 오바마 대통령 연설문의 코퍼스 분석」, 충남대 대학원 석론, 2012.
- 이정복, 「대통령 연설문의 경어법 분석」, 『배달말』33권, 배달말학회, 2003, 213-237쪽.
- 이창수, 「노무현 전 대통령의 연설문에 나타난 정체성 연구」, 『담화와 인지』제15권 제2호, 담화인지언어학회, 2008, 143-161쪽.
- 전영환, 「연설문 텍스트마이닝을 통한 미국 연방공개시장위원회 주요 인사들의 통화정책 성향 분석」, 서울대 대학원 석론, 2014.
- 정승훈, 「비판적 담화분석을 통한 대통령의 현충일 추념사 언어표현 연구」, 고려대 대학원 석론, 2020.
- 정의룡, 「대통령 지시사항에 대한 의미연결망 분석」, 『The Journal of the Convergence on Culture Technology(ICCT)』Vol4 No.1, 국제문화기술진흥원, 2018, 129-137쪽.
- 조수근, 조재희, 김성범, 「텍스트마이닝을 활용한 미국 대통령 취임 연설문의 트렌드 연구」, 『대한산업공학회지』41-5, 대한산업공학회, 2015, 453-460쪽.
- 조원형, 「대통령의 과거사 관련 연설에 대한 텍스트언어학적 분석」, 『수사학』제29집, 한국수사학회, 2017, 217-245쪽.
- 조용림, 「텍스트 마이닝을 활용한 글쓰기 자료 분석」, 『문화와 융합』 42권 5호, 한국문화융합학회, 2020, 279-307쪽.
- 조태호, 「텍스트 마이닝의 개념과 응용」, 『지식정보인프라』 5, 한국과학기술정보연구원, 2001, 76-85쪽.
- 홍주현, 윤해진, 「대통령 선거 후보자의 연설문 네트워크 분석」, 『한국콘텐츠학회논문지』제14권 제9호, 한국콘텐츠학회, 2014, 24-44쪽.
- 황창호, 이혁우, 임동완, 「역대 대통령의 연설문의 유형과 특징에 대한 비교 연구」, 『한국공공관리학보』제29권 제4호, 한국공공관리학회, 2015, 27-50쪽.
- Aggarwal, C. C. and Zhai, C., “Mining text data”, Springer, (2012).
- Chakraborty, G., Pagolu, M., and Garla, S., “Text Mining ad Analysis

: Practical Methods, Examples, and Case Studies Using SAS", SAS Institute, (2013).

Chen, Y. T. and Chen, M. C., "Using chi-square statistics to measure similarities for text categorization", Expert systems with applications, 38, (2011), pp.3085-3090.

Hu, X. and Liu, H., "Text analytics in social media", Mining text data, (2012), pp.385-414.

■ 국문초록

자연어는 인공어와 달리 컴퓨터가 직접 처리하기 곤란하여 이를 구조적인 데이터로 처리할 필요가 있다. 따라서 표본 텍스트에 단어의 가중치를 문서 내의 출현 빈도에 가중치를 부여하고 이를 근거로 일부의 단어를 키워드로 선택한다.

텍스트 마이닝의 정보 필터링을 통해 취임사 자료를 입력하고, 주제어 관련 연관성에 따른 특성값의 결과를 얻을 수 있어서 단순 어휘 빈도수보다 비정형 텍스트의 질적 향상을 도모할 수 있었다.

지금까지 이 연구에서는 대한민국 역대 대통령의 취임사 어휘를 분석하였다. 형태소 분석을 통해 단어를 정제한 후 어휘 빈도수와 특성값을 찾아내고, 워드 클라우드로 시각화하였다. 이 연구는 기존 어휘 분석의 단편적 분석이 아닌 특성값과 워드 클라우드를 통해 특성을 대표하는 객관적 주제어를 추출할 수 있었다.

주제어 ● 대통령, 취임사, 텍스트 마이닝, 형태소 분석, 어휘 빈도, 워드 클라우드.

■ Abstract

A Study on the Inaugural Address of the President of Korea : Focusing on feature values in Textmining

Joe, Yong Lim

Unlike artificial languages, natural languages are difficult for computers to process directly. Therefore, it is necessary to process it as structured data. In the end, by giving weight to the sample text, the frequency of appearance in the document was weighted.

The inaugural address data was entered through information filtering of text mining, and it was possible to obtain the result of the characteristic value according to the relation to the main word.

This study analyzed the vocabulary of the inaugural speeches of past presidents of the Republic of Korea. After refining words through morpheme analysis, we found vocabulary frequency and characteristic values, and the data was visualized in wordcloud. This study did not perform a fragmentary analysis of the existing lexical analysis. This study presented feature values and word clouds. Therefore, it was possible to extract objective keywords representing the characteristics of the keywords.

Keyword • President, Inaugural Address, Text Mining, Stemming, Vocabulary, Frequency, Wordcloud.