

연구노트

선거여론조사 뉴스의 주요 주제*

김지범** · 장덕현*** · 김솔이****

선거여론조사가 난립하고 응답률이 감소하여 국민의 불신을 초래하고 있다. 이를 해결하고자 정부, 정치권, 조사기관이 선거여론조사 제도의 개선을 추진해 왔으나, 여전히 상황이 나아지지 않고 있다. 선거여론조사에 대한 법률적 정의는 없지만, 선거여론조사로 간주되지 않는 예외 조항이 있다. 이러한 이유로 인해, 국민들은 선거여론조사와 비선거여론조사를 쉽게 구분하기가 어렵다. 따라서, 만약 국민이 선거여론조사의 중요성을 비선거여론조사보다 높게 평가하더라도, 두 조사를 구별하지 못한다면 선거여론조사에 대해 특별히 관심을 가지거나 참여하려는 동기가 생기지 않을 수 있다. 본 연구는 간접적으로나마 국민이 직관적으로 선거여론조사와 비선거여론조사를 구분할 수 있는지 알고자 '선거여론조사' 뉴스를 검색해 분석했다. 분석의 결과, 수시로 실시되는 선거여론조사에는 선거와 관련된 명확한 주제뿐만 아니라 북한 문제나 정치 스캔들처럼 선거와 직접적인 관련이 없다고 여겨질 수 있는 주제들도 혼재되어 있는 것으로 나타났다. 향후, 중앙선거여론조사심의위원회가 선거여론조사의 기간, 문항 수, 내용 등에 관한 구체적인 기준을 제시한다면 선거여론조사의 문제점을 개선하고 전반적인 조사환경에 긍정적인 변화를 가져오는 첫걸음이 되리라 기대한다.

주제어: 선거여론조사, 중앙선거여론조사심의위원회, 선거 뉴스

* 본 논문은 2020년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구이다 (NRF-2020S1A3A2A03096777). 본 연구의 토대가 된 기초연구는 중앙선거여론조사심의위원회 연구용역 결과보고서(2023.10)이다.

** 성균관대학교 사회학과 교수(jbk7000@gmail.com), 제1저자.

*** 한국갤럽 연구위원(dhjang@gallup.co.kr), 제2저자.

**** 성균관대학교 서베이리서치센터 선임연구원(sorikim0313@gmail.com), 교신저자.

I. 연구의 배경

1987년 한국궐립의 대선 예측조사가 조사업계의 확산계기가 된(박무익 2002) 이후, 정당의 기능과 역할이 무색하게 2002년 민주당 대선 후보 국민참여경선, 노무현과 정몽준 후보의 단일화, 그리고 한나라당 이명박과 박근혜 대선 후보 경선, 20대 민주당과 국민의 힘 대선 후보 경선 등 후보 공천과 경선은 선거여론조사와 떼어놓을 수 없는 상황이 되었다(이동운 2020; 김현성 2021: 313-325). 선거여론조사가 무소불위의 영향력을 행사하고 있음에도 불구하고 조사 양산, 저비용 조사방법, 낮은 응답률, 미디어의 선거조사 보도 행태가 선거여론조사에 대한 불신을 초래하고 있다.

해외의 선거여론조사는 주로 선거운동기간 동안 실시되며, 후보자와 선거 쟁점에 관한 주제를 다루는 조사를 의미한다(Traugott & Lavrakas 2008: 11). 반면, 한국에서는 선거여론조사가 일 년 내내 실시되고 있다. 중앙선거여론조사심의위원회(이하 중앙여심위)(2023: 10)에 따르면, 선거여론조사는 ‘선거에 관한 여론조사’로, 법률상 명확한 정의는 존재하지 않는 것으로 보고하고 있다. 중앙여심위의 웹사이트의 첫 번째 FAQ는 “선거여론조사는 무엇인가요?”라는 질문에 “당선인을 예상하게 하는 정당 또는 후보자에 대한 지지도 조사를 포함하여 당해 선거를 동기로 하거나 빌미로 하는 등 당해 선거와 관련 있는 여론조사는 모두 선거여론조사라고 할 수 있습니다. * 예: 선거공약정책에 대한지지·반대에 관한 여론조사 등”(중앙여심위, 2024.10.16. 접속)으로 언급된다. 예로, 뉴스 미디어를 통해 수시로 접하는 대통령 지지도 혹은 직무 수행 평가는 정당 지지도 문항이 포함되어 조사되었기에 선거여론조사로 간주된다. 또한, 공직선거법 제8조의8(선거여론조사심의위원회)에서는 다음과 같이 선거여론조사에 해당되지 않는 예외 규정을 명확히 밝히고 있다.

1. 정당이 그 대표자 등 당직자를 선출하기 위하여 실시하는 여론조사
2. 후보자(후보자가 되려는 사람을 포함한다)의 성명이나 정당(창당준비위원회를 포함한다)의 명칭을 나타내지 아니하고 정책·공약 개발을 위하여 실시하는 여론조사

3. 국회의원 및 지방의회의원이 의정활동과 관련하여 실시하는 여론조사. 다만 제60조의2제1항에 따른 해당 선거의 예비 후보자 등록 신청 개시일부터 선거일까지 실시하는 여론조사는 제외한다.
4. 정치, 선거 등 분야에서 순수한 학술·연구 목적으로 실시하는 여론조사
5. 단체 등이 의사결정을 위하여 그 구성원만을 대상으로 실시하는 여론조사

선거여론조사가 과도하게 실시되고 있지만, 응답률은 점점 떨어지고 있다. 지난 1년간 정당 및 후보 지지도 조사의 피조사 경험률은 2007년 28%에서 2017년 44%까지 증가하였다(김지범 외 2019). 2023년 1월 1일부터 2024년 10월 31일까지 중앙선거심위에 등록된 2,652건의 선거여론조사를 살펴보면, 조사기관에서 전체 약 1억 9천만 명에게 접촉을 시도하였다. 이중 비적격 사례(결번/사업체번호/팩스/대상지역 아님/할당초과 등)인 약 4천7백만 명을 제외한 약 1억 명은 접촉에 실패(통화 중/부재 중/접촉 안 됨 등)했고, 약 4천2백만 명은 수신 후 거절 및 이탈했으며, 약 2백만 명이 조사에 참여했다. 즉, 조사기관의 수많은 조사 요청에도 불구하고 조사대상자의 거절로 인해 아주 소수의 조사 참여자만을 대상으로 선거여론조사가 진행된다. 또 다른 사례로, 한국갤럽 데일리오피니언 응답률(AAPOR COOP3 협조율)의 추이를 보면 2016년 연평균 21.5%로 가장 높았고, 2018년부터 2021년까지 15% 전후를 유지하다가, 2023년은 10% 미만으로 하락하였다. 조사대상자는 한정되어 있는데 조사 요청이 너무 많아 조사 협력자가 고갈될 수 있는 공유자원(common pool resource) 문제로 귀결되고 있다(Leeper 2019).

선거여론조사의 난립은 조사방법과 결부되어 있다. 조사기관은 선거조사를 수행할 때 대부분 전화조사방식을 이용하는데, 특히 전화면접보다는 ARS(Automatic Response System, 전화자동응답시스템)가 더 많이 사용되고 있다. <표 1>을 보면 중앙선거여론조사심의위원회에 등록된 2,652건의 선거여론조사 중 61%가 ARS조사, 36%는 전화면접조사, 2%는 ARS와 전화면접 혼합 조사, 그리고 1%는 인터넷, 스마트폰 앱 조사였다. ARS 조사가 더 빈번하게 활용되는 이유는 저비용으로 자료를 신속하게 수집할 수 있기 때문이다(이갑윤 외 2023: 51). 그러나, 전화면접조사에 비해 응답률이 낮을 뿐만 아니라 대표성의 문제를 안고 있다. 예를 들어, 주요 정당의 지지율을 과대평가하고 무당파 비율을 낮추는 경향이 확인된다(최종호 외 2022).

<표 1> 선거여론조사 조사방법

조사방법	빈도	%
무선 ARS	935	35.3
무선전화면접	896	33.8
무선 ARS, 유선 ARS	687	25.9
무선전화면접, 유선전화면접	45	1.7
무선 ARS, 유선전화면접	20	0.8
인터넷조사	13	0.5
무선 ARS, 무선전화면접	12	0.5
스마트폰앱조사	10	0.4
인터넷조사, 스마트폰앱조사	10	0.4
무선 ARS, 유선 ARS, 무선전화면접	9	0.3
무선 ARS, 유선 ARS, 무선전화면접, 유선전화면접	6	0.2
무선전화면접, 인터넷조사	4	0.2
무선 ARS, 무선전화면접, 유선전화면접	3	0.1
유선 ARS	2	0.1
합계	2,652	100

기간: 2023년 1월 1일 ~ 2024년 10월 31일.

출처: 중앙선거여론조사심의위원회.

이렇게 만들어진 선거여론조사 결과가 뉴스 미디어를 통해 거의 매일 노출되고 있지만, 기자들의 여론조사에 대한 이해 부족으로 인한 ‘몰상식’ 기사와 경마식 보도의 문제점이 지속적으로 제기되고 있다(신창운 2014). 하나의 예로, 표본오차 범위 내의 대통령 지지율 변화를 상승 혹은 하락으로 묘사하는 보도 행태를 들 수 있다. 더불어, 언론인들이 자율적으로 제정하여 2016년 12월 8일부터 실시된 선거여론조사보도준칙이 잘 지켜지지 않고 있는 점과 보도에 앞서 언론사가 짧은 기간에 적은 예산으로 다수의 조사를 진행하는 문제도 지적되고 있다(한규섭 외 2021).

현재 증폭되고 있는 선거여론조사의 위기 상황에 대응하여 정부, 정치권, 조사가

관들은 각각 구체적이고 실질적인 방안을 마련하며 규제 제도를 추가하거나 기준을 강화하고 있다. 2014년 중앙여심위가 출범한 이후, 선거여론조사 기준 마련, 정보의 투명공개, 불법 선거여론조사기관 통제, 선거여론조사기관 등록제 도입, 가상번호 사용 등 다양한 제도적 개선이 시행되었다. 최근에는 국회의원 14인(장제원 의원 대표 발의, 2022.11.23.)이 응답률 5% 미만 공표 금지를 포함한 공직선거법 일부개정법률안을 발의했다. 또한 한국조사협회(KORA)는 선거여론조사의 신뢰성을 제고하기 위해 ‘정치선거 전화 여론조사기준’을 수립하여, 전국 단위 조사에서 휴대전화 가상번호를 사용할 경우에는 최소 10%, RDD(Random Digit Dialing, 전화번호 임의 걸기) 방식의 경우에는 최소 7%의 응답률 기준을 설정하는 결정을 내렸다(KBS 뉴스 2023).

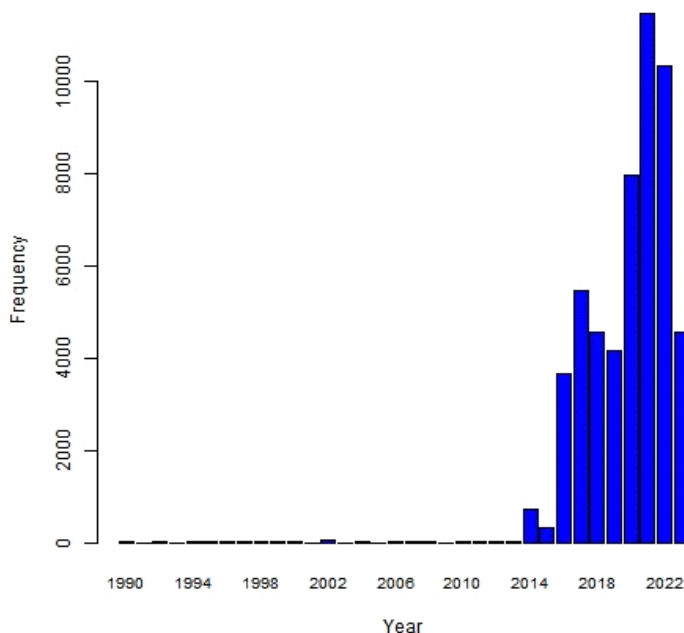
지금까지 선거여론조사 위기에 관한 대응 방안은 선거여론조사 자체에만 초점이 맞춰져 왔다. 비록 선거여론조사 제도 개선이 분명히 의미 있는 대책이라고 할지라도, 이러한 논의가 이루어지는 전반적인 조사환경에 대한 고려는 부족한 상황이다. 특히 비선거여론조사는 관리와 통제가 충분하지 않으며, 선거여론조사의 정의도 엄밀하지 않고 예외 규정이 존재하여 국민이 선거여론조사와 비선거여론조사의 차이를 명확히 인식할 수 있는지를 확신하기 어렵다. 예를 들어, 중앙여심위에 등록된 조사에 북한 관련 질문이 포함되면 선거여론조사로 취급되지만(예, 뉴스1 2024; 강원일보 2024), 동일한 질문이 등록되지 않은 조사에서는 선거여론조사로 간주되지 않는다. 또 다른 사례로, ‘국민 76% 정몽규·홍명보 사퇴해야’라는 기사(헤럴드경제 2024)도 중앙여심위에 등록된 선거여론조사에 기반하고 있다. 이는 중앙여심위가 정당이나 후보 지지도가 포함된 조사를 선거여론조사로 간주하기 때문이며, 결국 특정 문항이 선거여론조사에 해당하는지는 정당 지지도 등 관련 문항의 포함 여부에 따라 결정된다는 것을 보여준다. 국민이 비록 선거여론조사가 비선거여론조사보다 더 중요한 기능을 수행한다고 믿더라도, 두 조사의 차이를 분명하게 인식하지 못한다면 특별히 선거조사에 더 우호적으로 참여할 동기가 없을 것 같다.

이러한 문제의식에서 본 연구는 뉴스 미디어에서 다루는 선거여론조사 관련 기사를 분석하여, 선거여론조사에 대한 정부 규정이 국민의 인식과 괴리가 있는지를 간접적으로 탐색하고자 한다. 선거여론조사 기사 주제가 국민이 쉽게 선거여론조사와 연관 짓지 못하는 것이라면, 현행 제도를 철저히 보완하거나 새로운 법과 정책의 도입이 필요함을 시사한다.

II. 자료와 분석

1. 선거여론조사 온라인 기사

‘선거여론조사’ 키워드를 포함한 온라인 기사는 최근 들어 급격하게 증가하고 있다. 빅카인즈에서 수집한 연도별 기사 수(<그림 1>)를 살펴보면, 선거여론조사 관련 기사는 1990년부터 2013년까지 연 100건 미만이었으나 선거여론조사심의위원회가 출범한 2014년에 733건, 선거여론조사 등록이 상시화된 2016년에 3,500건으로 대폭 증가했다. 제도 도입이 ‘선거여론조사’ 용어의 사용과 직접 관련된 것으로 보인다. 2014년 이전에는 여론조사, 선거조사, 정치조사 등 다양한 용어로 선거여론조사를 규정하였다.



<그림 1> 연도별 선거여론조사 온라인 기사 수

기사 수가 가장 많았던 연도는 2021년으로, 총 11,481건이었다. 2021년은 제20대 대통령 선거를 앞둔 해로, 유권자들의 후보자와 정당에 대한 지지도를 반영한 온라인 기사들이 대량으로 생성된 것으로 보인다. 이러한 경향은 선거 이후에도 지속되어, 현재까지 상당수의 선거여론조사 관련 기사가 보도되는 추세이다.

2. 토픽 모델링 방법

1) 자료의 출처 및 범위

본 연구에서는 국내 주요 신문사에서 발행된 ‘선거여론조사’ 온라인 뉴스 기사를 대상으로 했다. 온라인 뉴스 기사는 특정 시기의 국내 온라인 뉴스를 검색하고 분석이 가능한 플랫폼인 빅카인즈(<https://www.bigkinds.or.kr/>)를 통해 수집했다. 수집한 자료는 데이터 제공 초기 연도인 1990년부터 2023년 9월 9일 사이에 작성된 54,642개의 기사이다.

2) 자료 수집 방법

기사 수집은 빅카인즈의 상세 검색 기능 중 ‘키워드를 중심으로 한 검색’에서 ‘선거여론조사’를 키워드로 지정하여 수행했다. 빅카인즈에서 검색한 기사는 각 신문의 웹페이지를 직접 방문하여 전문을 크롤링할 필요 없이 정리된 형태의 자료를 엑셀 파일로 바로 다운로드할 수 있었다.

3) 데이터 전처리

수집한 자료는 본격적인 분석에 앞서 전처리 과정을 거쳤다. 각 과정은 중복기사 제거, 명사 추출, 유의어 대체 및 불용어 제거 순으로 이루어졌다.

(1) 기사 제거

전체 기사에서 유효한 기사만 선택하기 위해 두 가지 과정을 거쳤다. 먼저, 빅카인즈에서 제공한 엑셀 자료의 ‘분석 제외 여부’ 칼럼에서 ‘예외’인 기사를 삭제했다. 예외로 분류된 기사들은 주로 주요 행정직 인사, 부고, 동정, 또는 사진 기사로 본 연구의 취지에 적합하지 않았다. 다음으로, 일자, 기고자, 제목, 본문이 일치하는 경우 중복기사로 간주하여 하나의 기사만 분석에 포함하고 중복되는 기사를 삭제하여 총 53,690개 기사를 분석에 이용하였다.

(2) 명사 추출

수집한 엑셀 파일에는 ‘일자’, ‘언론사’, ‘기고자’, ‘제목’, ‘각종 분류’, ‘인물’, ‘위치’, ‘기관’, ‘키워드’, ‘본문’, ‘URL’ 등 다양한 상세 정보가 포함되어 있다. ‘본문’은 뉴스 기사의 전문이 아닌 상위 일부만 포함하고 있다는 점과 일부 과거 기사는 URL로도 원문 확인이 불가능하다는 점에서 본 연구의 분석에는 ‘키워드’를 대신 활용하였다. 키워드는 기사 본문, 제목, 사진 제목 등에서 명사를 추출한 결과이다. 다만 추출된 키워드를 텍스트 분석에 바로 활용하기에는 한계가 있어 분석 전 전처리 과정을 거쳤다.

(3) 전처리

자료 전처리는 크게 다섯 단계로 진행하였다. 첫 번째로, 검색 키워드를 제외하였다. 기사 수집에 사용한 키워드인 ‘선거여론조사’는 모든 문서에 포함되어 있어 의미가 없으므로 삭제하였다. 두 번째로, 수집된 키워드 중 오·탈자로 의심되는 단어는 원문과 대조하여 검토하였다. 가령, ‘코로’를 ‘코로나’, ‘황교’를 ‘황교안’, ‘조사서’를 ‘조사’로 수정하였다. 세 번째로, 외래 약어, 한자, 숫자를 한국어로 변환하는 표준화 작업을 수행하였다. ‘LH’는 ‘한국토지주택공사’, ‘shy’는 ‘샤이’, ‘檢’은 ‘검찰’, ‘尹’은 ‘윤석열’, ‘文’은 ‘문재인’, ‘95%’는 ‘구십오프로’, ‘20대’는 ‘이십대’ 등으로 대체했다. 네 번째로, 유의어 및 동의어를 통일하였다. 예를 들어, ‘코로나바이러스’, ‘코로나19’, ‘코로나-19’, ‘COVID’, ‘COVID19’, ‘COVID-19’는 ‘코로나’로 대체하였다. 다섯 번째로, 불용어를 처리하였다. 출현 빈도는 높으나 분석적 가치가 낮은 단어들을 제거하고, 특수기호(% , ? , - 등)와 영문도 함께 제거하였다. ‘사진’, ‘제목’, ‘화면’, ‘캡처’, ‘세부’, ‘홈페이지’, ‘참조’ 등이 그 예이다.

4) 분석방법: 구조적 토픽 모델링

구조적 토픽 모델링(Structural Topic Modeling, 이하 STM)은 텍스트 마이닝 방법의 하나로, 텍스트 데이터에서 주제를 추출할 때 해당 텍스트의 메타데이터를 동시에 고려하는 기법이다. 메타데이터(metadata)는 텍스트의 외부적 특성을 나타내는 정보로, 본 연구의 분석자료인 온라인 기사의 경우 발행 날짜, 저자, 신문사 등이 이에 해당한다. 따라서 STM은 텍스트 내용만을 분석하는 일반적인 토픽 모델링(예: 잠재적 디리클레 할당(LDA))보다 정교한 분석이 가능하며, 메타데이터를 활용한 다양한 해석을 도출할 수 있는 기법이다.

‘선거여론조사’와 관련된 이슈들은 시간에 따라 중요성이 달라질 수 있다. 따라서 STM을 통해 주요 토픽이 무엇인지, 그리고 이러한 토픽들이 시간에 따라 어떻게 변화하는지를 추적하고자 했다. 이를 위해 먼저 전체 뉴스 기사를 분류할 수 있는 적정 토픽 개수를 정하고, 각 토픽에 분류된 단어와 대표 기사를 살펴봄으로써 토픽명을 부여하였다. 다음으로 전체 문서에서 각 토픽이 차지하는 비중과 토픽들의 연도별 발현 비중을 도출하였다.

3. 선거여론조사 토픽과 연도별 발현

1) 선거여론조사 토픽

<표 2>는 STM으로 추출한 토픽의 이름, 주요 단어, 비중, 문서 수를 정리한 것이다. 적절한 토픽 수를 결정하기 위해 의미적 일관성(coherence)과 배타성(exclusivity)을 지표로 활용하였다. 분석의 결과, 토픽이 15~20개일 때 두 지표의 균형이 가장 적절한 것으로 나타났으며, 이 범위 내에서 각 토픽의 주요 단어와 대표 원문을 검토하여 최적의 토픽 수를 15개로 결정하였다.

각 토픽의 이름은 해당 토픽 내 주요 단어들을 고려하여 일차적으로 선정하였다. 이후 각 토픽을 가장 잘 대표하는 기사의 원문을 검사하여 선정한 이름의 적합성을 검토하여 최종적으로 확정하였다. 각 토픽의 주요 단어들은 해당 토픽과 밀접한 연관성을 가지는 단어들로, 토픽의 특성과 주요 내용을 반영하였다. 최빈 단어는 STM 분석결과의 ‘Highest Probability’로 해당 토픽에 자주 등장하는 단어를 지칭하며, 독점 단어는 ‘Frequency and Exclusivity (FREX)’로 토픽 고유의 특징을 잘 나타내는 단어를 의미한다. 비중은 토픽이 전체 문서에서 차지하는 비율을 나타내며, 문서 수는 각 기사에서 가장 높은 비중을 차지하는 토픽을 해당 기사에 할당한 후 토픽별로 합산한 값이다.

토픽 분석결과를 살펴보면, 추출된 토픽의 과반수는 정치적 동향 및 지지율 관련 내용이었다. 전체 문서 내에서 가장 큰 비중을 차지하는 토픽은 ‘대통령 지지율 동향(토픽1)’으로, ‘연속’, ‘하락’, ‘상승’과 같은 지지율 변화 관련 단어들과 ‘포인트’라는 변동 정도를 나타내는 단어가 ‘대비’, ‘지난주’ 등의 시간 표현과 함께 등장하여 지지율의 동향을 구체적으로 보여주었다. 다음으로 비중이 높은 토픽은 ‘대통령 직무 수행 평가(토픽2)’로서 대통령의 국정 수행에 대한 국민의 긍정적 및 부정적 평가를 다루었다. 다만 조사회사별로 용어 사용이 달라, 리얼미터의 ‘대통령 지지

도’와 한국갤럽의 ‘대통령 직무평가’는 유사하거나 동일한 개념을 측정한 것으로 볼 수 있다.

대통령 관련 토픽 외에도 다양한 정치인과 정당의 지지율 관련 토픽들이 확인되었다. 여기에는 대선 주자들 간의 지지율 비교를 포함한 ‘대선 후보 지지율과 단일화(토픽3)’, 정치인 및 대선 주자들의 선호도 조사로 구성된 ‘주요 정치인 선호도(토픽4)’, 정당별 지지율과 그에 따른 민심을 보여준 ‘정치적 성향과 민심(토픽6)’이 포함되었다. 또한 정당에 대한 지지도와 응답자의 성향을 나타낸 ‘정당 지지도와 투표(토픽10)’, 정당별 지지율과 지도자들의 선호도로 이루어진 ‘정치 주요 인물과 정치 현안(토픽13)’, 그리고 선거구별 총선 경쟁과 후보자 지지율을 다룬 ‘지역선거와 정치동향(토픽14)’이 확인되었다.

한편 정치적 상황이나 이슈에 초점을 맞춘 토픽들도 있었다. ‘정치 스캔들과 법적 대응(토픽7)’은 ‘대통령’, ‘장관’, ‘의혹’과 같은 최빈 단어와 ‘징계’, ‘청문회’, ‘영장’ 등의 독점 단어를 통해 정치 스캔들에 초점을 맞추었다. ‘정치적 의사소통과 여론(토픽8)’은 ‘이야기’, ‘생각’, ‘발언’ 등의 최빈 단어와 ‘박지훈’, ‘황보선’ 등의 독점 단어로 정치인들의 공개적인 의사소통과 여론의 경향성을 보여주었다. ‘부산시 관련 정치 이슈와 보궐선거(토픽15)’는 ‘부산’을 중심으로 한 시장 선거와 ‘박형준’, ‘김영춘’과 같은 인물들, ‘보궐선거’, ‘가덕신공항’ 등의 키워드를 통해 지역 발전과 정치적 상황을 다루었다.

다음으로, 정부 정책 및 사회적 이슈 관련 토픽도 추출되었다. ‘정부 정책과 사회 문제(토픽12)’는 ‘코로나’ 대응과 ‘부동산’ 정책 등과 같은 주요 정부 정책에 대한 국민의 반응, 그리고 ‘청년’ 지원, ‘일자리’ 창출 등 사회경제적 정책에 대한 대중의 의견을 담고 있었다. ‘북한에 대한 여론(토픽9)’은 ‘북한’과 관련된 사안이나 ‘남북 정상회담’과 같은 중대 이슈에 대한 ‘찬성’ 및 ‘반대’ 의견, ‘평화’와 ‘사면’ 등의 주제에 대한 국민의 ‘공감’ 정도를 보여주었다.

그리고 선거 및 여론조사 방법론 관련 토픽도 확인되었다. ‘선거법 준수와 위반(토픽11)’은 ‘선관위’에 의해 공표된 ‘후보자’의 ‘공직선거법’ 위반 사례와 이에 대한 ‘위원회’의 조치를 포함하였고, ‘혐의’와 ‘고발’, 그리고 그에 따른 ‘위반’ 행위의 내용을 담고 있었다. ‘여론조사 방법과 신뢰도(토픽5)’는 ‘응답’, ‘무선’, ‘유선’ 등의 최빈 단어와 ‘면접’, ‘가중’, ‘표집’ 등의 독점 단어들을 통해 여론조사 방식의 다양성과 기술적인 측면을 보여주었으며, ‘95%’ 신뢰 수준과 ‘지역별’, ‘연령별’ 등의 단어들이 주요하게 나타났다.

<표 2> 선거여론조사 뉴스의 토픽별 주요 단어와 비중

토픽	토픽명	최빈 단어	독점 단어	비중	문서
1	대통령 지지율 동향	포인트, 지지율, 하락, 상승, 기록, 리얼미터, 대통령, 민주당, 집계, 결과, 전주, 지난주, 국민, 대비, 지도도	연속, 하락, 전주, 주간, 상승, 대비, 지난주, 하락세, 소폭, 포인트, 리얼미터, 경신, 리얼, 미터, 집계	0.164	9,698
2	대통령 직무 수행 평가	평가, 대통령, 부정, 긍정, 한국갤럽, 국정, 수행, 포인트, 결과, 지지율, 직무, 국민, 윤석열, 지지, 지도도	긍정, 부정, 직무, 평가, 수행, 부정률, 외교, 직무수행, 평가자, 전반적, 미흡, 독단적, 긍정률, 취임, 한국갤럽	0.108	5,870
3	대선 후보 지지율과 단일화	후보, 대선, 포인트, 지지율, 국민, 이재명, 안철수, 윤석열, 지지, 결과, 단일, 오차범위, 격차, 대결, 단일화	단일, 단일화, 후보, 대선후보, 김동연, 심상정, 김은혜, 다자, 토론, 물결, 다자대결, 정권교체, 대선, 선대위, 안철수	0.095	6,257
4	주요 정치인 선호도	의원, 대표, 지사, 총장, 이재명, 이낙연, 대선, 윤석열, 주자, 국민, 결과, 지지, 포인트, 선호도, 홍준표	이낙연, 총장, 선호도, 검찰총장, 대권, 최재형, 지사, 정세균, 감사원장, 주자, 원희룡, 추미애, 김기현, 총리, 당심	0.087	6,247
5	여론조사 방법과 신뢰도	응답, 전화, 무선, 방식, 응답률, 유선, 신뢰, 수준, 지지, 표본오차, 구십오프로, 지역별, 연령, 대상, 지역	무선, 유선, 자동, 면접, 가중, 임의, 표집, 전화, 생성, 추출, 방식, 통계, 부여, 응답률, 유선전화	0.070	2,023
6	정치적 성향과 민심	지지율, 보수, 민주당, 총선, 대선, 정치, 지지, 대통령, 정권, 여당, 야당, 민심, 상황, 중도, 표심	현상, 중도, 전문가들, 샤이, 세대, 시대정신, 소장, 표심, 변수, 투표율, 보수, 정권, 엄경영, 교체, 명지대	0.063	2,468
7	정치 스캔들과 법적 대응	대통령, 대표, 민주당, 장관, 국민, 국회, 의원, 검찰, 수사, 정부, 청와대, 의혹, 원내, 인사, 조국	징계, 청문회, 체포, 김남국, 대통령실, 비대위, 부결, 코인, 정지, 윤리위, 조국, 해임, 강행, 검찰개혁, 영장	0.060	3,123
8	정치적 의사소통과 여론	이야기, 생각, 국민, 대통령, 정치, 대표, 사람, 의원, 말씀, 상황, 발언, 윤석열, 정도, 입장, 자신	말씀, 박지훈, 이야기, 황보선, 배종찬, 김준일, 라디오, 공격, 출연, 본인, 이동형, 개인적	0.048	2,141

<표 2> 선거여론조사 뉴스의 토픽별 주요 단어와 비중 (계속)

토픽	토픽명	최빈 단어	독점 단어	비중	문서
9	북한에 대한 여론	반대, 찬성, 의견, 응답, 대결, 결과, 지지, 응답자, 북한, 국민, 전국, 구십오프로, 대통령, 리얼미터, 한국당	찬성, 찬반, 반대, 남북, 사면, 한반도, 김정은, 북한, 의견, 남북정상회담, 공감, 올림픽, 회담, 북미, 평화	0.047	2,187
10	정당 지지도와 투표	지지, 민주당, 정당, 투표, 국민, 정의당, 비례, 지지도, 결과, 더불어민주당, 응답자, 응답, 시장, 기록, 여성	의향, 비례, 미래한국당, 더불어민주당, 시민당, 무등일보, 중부일보, 데일리리서치, 호감, 작위추출, 투표, 남성, 제선거구, 기타, 비례정당	0.045	2,204
11	선거법 준수와 위반	결과, 공표, 보도, 후보자, 선관위, 등록, 정당, 위반, 위원회, 혐의, 기관, 투표, 고발, 행위, 왜곡	선관위, 공직선거법, 여심위, 과태료, 선거법, 위반, 왜곡, 벌금, 공표, 신고, 행위, 혐의, 고발, 불법, 중앙선관위	0.045	2,522
12	정부 정책과 사회 문제	정책, 정부, 코로나, 부동산, 경제, 공약, 청년, 지원, 추진, 대책, 국민, 강조, 일자리, 사회, 지급	주택, 백신, 방역, 지원금, 공급, 접종, 산업, 집값, 재난, 소상공인, 확진자, 성장, 공공, 사업, 부동산	0.044	1,922
13	정치 주요 인물과 정치 현안	대표, 문재인, 한국당, 안철수, 호남, 지사, 새누리당, 대선, 정당, 지지율, 홍준표, 더불어민주당, 시장, 유승민, 국민의당	안희정, 반기문, 김무성, 바른정당, 손학규, 권한, 유엔, 호남, 대행, 최순실, 탄핵, 황교안, 남경필, 사무총장, 친박	0.044	2,284
14	지역 선거와 정치 동향	후보, 지역, 의원, 새누리당, 더불어민주당, 총선, 결과, 선거구, 공천, 대구, 교육감, 예비, 지지율, 지역구, 당선	원주, 수성갑, 춘천, 종로, 창원성산, 교육감, 강기윤, 동작, 조희연, 순천, 무소속, 여영국, 노관규, 예비후보, 이인선	0.040	2,768
15	부산시 관련 정치 이슈와 보궐선거	시장, 민주당, 서울, 부산, 지역, 국민, 의원, 출마, 경남, 보궐, 위원장, 지방, 경선	박형준, 김영춘, 보궐, 가덕도, 박영선, 오거돈, 부산시장, 보선, 서병수, 보궐선거, 광역, 가덕신공항, 이연주, 김태호, 부산시	0.040	1,976

앞의 15개 토픽을 단순화하여, 선거여론조사 토픽과 비선거여론조사 토픽으로 구분할 수 있다. 선거여론조사 토픽은 대선 후보 지지율과 단일화(토픽3), 주요 정치인 선호도(토픽4), 여론조사 방법과 신뢰도(토픽5), 정치적 성향과 민심(토픽6), 정당 지지도와 투표(토픽10), 선거법 준수와 위반(토픽11), 지역선거와 정치 동향(토픽14), 부산시 관련 정치 이슈와 보궐선거(토픽15)로 약 49%를 차지하였다. 즉, 뉴스 기사에서 선거와 관련된 주제와 비선거와 관련된 여론조사 주제가 비슷한 비중으로 보도되었다고 볼 수 있다.

2) 선거여론조사 토픽 연도별 발현 비중

<그림 2>는 선거여론조사 관련 온라인 기사에서 나타난 토픽들의 연도별 비중 변화를 시각화한 결과이다. 뉴스 기사를 수집한 전체 기간 중 선거여론조사 제도 개선으로 온라인 기사 수가 증가한 2014년 이후의 결과를 제시하였다.

대통령 활동과 선거에 따라 토픽들의 비중이 변화가 있었다. 대통령 지지율(토픽1)과 대통령 직무수행 평가(토픽2)는 하나의 대통령 관련 토픽으로 볼 수 있었으며, 2019년처럼 전국 단위 선거가 없던 해에도 대통령과 정부 활동에 대한 기사가 꾸준히 보도되었다. 주요 정치인에 대한 관심(토픽4)도 지속적으로 높은 비중을 차지했는데, 특히 2016년 대통령 국정농단·탄핵 사태나 2021년 대선 주자 경쟁과 같은 정치적 변화가 있을 때 더욱 증가하였다. 대선과 관련된 토픽3(대선 후보 지지율과 단일화)은 대선 주기와 맞물려 시기별 변동이 컸던 반면, 여론조사 자체에 대한 관심(토픽5: 여론조사 방법과 신뢰도)은 비교적 안정적인 비중을 유지하여, 여론조사의 방법론과 신뢰성에 대한 지속적인 논의가 이루어졌음을 알 수 있다.

정치적 이슈와 관련된 토픽들(토픽6: 정치적 성향과 민심, 토픽7: 정치 스캔들과 법적 대응, 토픽8: 정치적 의사소통과 여론)의 비중 변화는 한국 정치의 역동성을 반영했다. 특히 토픽7의 경우, 특정 시기에 비중이 급증하는 모습을 보이는데, 이는 주요 정치 스캔들이 발생했을 때 여론의 관심이 집중되는 현상을 반영했다. 관련된 예를 보면, 2019년 조국 전 법무부 장관 관련 의혹, 2022년 대선 후보들의 각종 의혹 제기, 그리고 2023년 여·야 주요 정치인들을 둘러싼 수사와 논란을 들 수 있다.

북한 관련 여론(토픽9)과 지역 선거 및 정치 동향(토픽14), 그리고 부산시 관련 정치 이슈(토픽15)의 비중 변화는 국내·외 정세와 지역 정치의 중요성이 시기에 따라 달라짐을 보여주었다. 예를 들어, 북한 관련 여론의 경우 남북정상회담이 열린 2018, 2019년에 비중이 높았다. 한편 선거법 준수와 위반(토픽11)에 대한 관심은

2014년 이후 급격히 감소했는데, 이는 선거 관련 제도가 안정화되면서 관련 논란이 줄어들었음을 의미할 수 있다.



<그림 2> 선거여론조사 뉴스 토픽 연도별 발현 비중 추이

이를 통해 선거여론조사 관련 토픽의 발현 비중 변화를 살펴볼 수 있었다. 정치인과 선거에 직접 관련된 주제들이 지속적으로 높은 비중을 차지했으며, 선거와 직접적인 관련성이 적은 토픽들도 꾸준히 일정 비중을 유지했음을 확인할 수 있었다. 이는 결과적으로 일반 국민들이 선거여론조사와 일반여론조사를 명확히 구분하기 어려울 수 있음을 시사한다.

III. 결론

본 연구는 빅카인즈에서 1990년부터 2023년 9월 9일까지 선거여론조사 관련 뉴스가 담고 있는 내용을 토픽 모델링을 통하여 살펴보았다. 수시로 실시되는 선거여론조사에는 선거와 관련된 명확한 주제뿐만 아니라 북한 문제나 정치 스캔들처럼 선거와 직접적인 관련이 없다고 여겨질 수 있는 주제들도 혼재되어 있었다. 또한, 선거여론조사와 비선거여론조사의 비중도 거의 차이가 없었고, 시간의 흐름에도 이러한 경향은 크게 변하지 않는 것으로 관찰되었다. 즉, 중앙여심위의 선거여론조사 규정과 실제로 여론조사에 참여하는 국민이 느끼는 선거여론조사의 의미 사이에 간극이 있을 수 있음을 시사한다.

물론, 실제로 국민들에게 위의 주제를 통해 선거여론조사와 비선거여론조사의 차이를 분간할 수 있는지 직접 물어보지 못했다는 점은 본 연구의 분명한 제한점이다. 향후 연구에서는 선거여론조사 요청을 받은 국민이 갖는 선거여론조사 이해도와 관점을 파악할 필요가 있다.

다양한 질문을 포함한 많은 조사가 상시 진행되는 선거여론조사로 등록되어 있다는 점에서 국민들이 선거여론조사와 비선거여론조사를 구별하지 못한다면, 선거여론조사만을 개선시키려는 조치는 효과적이지 않을 가능성이 있다. 또한, 선거여론조사가 직면한 조사의 난립, 낮은 응답률, 불신 등의 문제가 선거여론조사로만 국한되는 것이 아니라는 점에서, 전반적으로 조사기후(survey climate)를 개선시키려는 노력이 병행되어야 한다는 점도 선거여론조사의 개선을 위해 고려해야 할 사항이다. 이러한 점을 전제로 중앙여심위, 정치권, 뉴스 미디어, 그리고 연구자에게 몇 가지 제도 개선 방안을 제안한다.

첫째, 중앙여심위는 선거여론조사의 포괄적 정의를 다시 돌아보아야 할 때이다. 앞에서 언급했듯이, 비선거여론조사를 제외한 선거여론조사만으로도 조사 과잉 상태인 바, 2년간 조사기관의 전화 수신 후 거절 및 이탈만 4천 2백만 명이다. 지금까지 선관위는 선거여론조사 규제 범위를 확장하는 방향으로 제도를 개선해 왔지만(예, 선거여론조사기관 설립 요건 등), 여전히 선거여론조사가 직면한 문제들은 해소되지 않았다. 이제는 미온적인 대응을 넘어 새로운 해결 방안을 찾아야 할 때이다. 무엇보다도, 선거여론조사에 대한 명확하고 합의된 정의를 먼저 확립할 필요가 있다. 현재 언론에서 매일 다뤄지는 대통령 직무 수행 평가, 국정 지지율, 정당 지지율, 현안 선호도 등을 선거여론조사 범위에서 제외하고, 일반적인 내용도 축소해 문항 수를 제한하는 방안을 검토할 수 있다. 더불어 선거여론조사를 선거운동 기간으로 한정하는 것도 하나의 해결책이 될 수 있다.

둘째, 정부 통계의 외부 조사환경을 향상시키는 방법으로 소셜마케팅이 제안된 바 있듯이(Lorenc et al. 2013), 선거여론조사가 선거운동 기간에만 수행된다면 효과적인 홍보를 통해 다른 조사와 차별화를 이루고 국민들의 인식 변화를 유도할 가능성도 있다. 예를 들어, 모든 선거여론조사의 경우 면접원이 조사의 시작부터 “이 조사는 선거여론조사입니다”를 명확히 밝히고 진행하는 방식을 도입할 수 있다. 현재 뉴스에서 여론조사 결과를 보도한 후 중앙여심위를 참조하라고 안내하는 것처럼, 조사 시작 단계에서 중앙여심위가 관리하는 선거여론조사라는 점을 각인시키는 방법이다. 현재 뉴스 미디어에서 중앙여심위를 참조하라는 언급은 조사결과가 신뢰성이 있다는 점을 강조하려고 하는 것일 뿐 하락하는 응답률을 개선하는 데는 효과적이지 않다.

셋째, 선거여론조사의 이해관계자인 정치권, 뉴스 미디어, 조사업계, 그리고 중앙여심위가 모두 참여하여 조사환경을 바꿀 수 있는 실질적인 대안을 심각하게 논의해야 한다. 정치권, 뉴스 미디어, 연구자, 조사회사 등 선거여론조사를 의뢰하는 기관들이 공동 조사를 실시하여 조사대상자에게 끊임없이 조사 요청이 가지 않도록 하는 방법도 공유자원 문제를 해결하는 한 방안이 될 수 있다(Leeper 2019; 박인서·백영민 2021). 특히, 정치권의 후보 단일화나 공천 과정에서 여론조사에 의존하는 관행은 선거여론조사가 비판과 논란의 중심에 서게 만드는 주요 원인 중 하나이다. 선거여론조사 문제를 해결하기 위해 가장 중요한 시작점은 선거여론조사에 대해 방관적 태도를 보여온 정치권이 그 활용 여부와 범위에 대해 새로운 결단을 내리는데 있다고 해도 과언이 아니다.

중앙여심위는 선거여론조사 개선을 위해 더욱 적극적인 역할을 해야 한다. 선거여론조사라는 명목으로 이루어지는 잦은 접촉 시도와 이에 따른 거절 사례를 보면, 국민들이 선거여론조사에 상당한 피로감을 겪고 있다는 점을 확인할 수 있다. 이러한 과도한 선거여론조사는 조사환경 전반에 부정적인 영향을 미치며, 결과적으로 선거여론조사의 질적 저하를 초래하는 악순환으로 이어진다고 볼 수 있다. 난립하고 남용되는 선거여론조사를 개선하기 위해, 보다 구체적이고 명확한 기준을 설정해야 한다.

참고문헌

- 김지범·장덕현·김솔이·강정환. 2019. “조사동향: 여론조사에 대한 조사.” 《조사연구》 20(1): 131-164.
- 김현성. 2021. 《선거로 읽는 한국 정치사》. 웅진지식하우스.
- 박무익. 2002. “조사동향: 선거여론조사의 역사와 의의.” 《조사연구》 3(1): 91-118.
- 박인서·백영민. 2021. “선거기간 설문조사의 응답률 변화 연구.” 《언론정보연구》 58(3): 5-45.
- 신창운. 2014. 《위기의 여론조사》. 리북.
- 이갑윤·이지호·이현우. 2023. 《한국의 여론 조사》. ㈜푸른길.
- 이동윤. 2020. “한국의 정당공천제도: 정당의 국회의원 후보는 누가 결정하는가?” 《정치정보연구》 23(2): 265-290.
- 중앙선거여론조사심의위원회 FAQ. 2024.
<https://www.nesdc.go.kr/portal/bbs/B0000046/list.do>
- 최종호·김태균·백다예·왕정하·한강욱. 2022. “조사방법 차이가 정치적 태도 조사결과에 미치는 영향: 중앙여심위 선거여론조사 통합자료 분석.” 《한국정당학회보》 21(3): 5-40.
- 한규섭·김미나·김춘석·박종희·이내영. 2021. “불신받는 대선 여론조사-난립하는 군소 여론조사 업체들이 대선 좌우하는 현실 진단.” 《관혼저널》 63(4): 67-111.
- 헤럴드경제. 2024.10.02. “국민 76% 정몽규·홍명보 사퇴해야”.
- KBS 뉴스. 2023.10.22. 한국조사협회 “여론조사는 전화면접만 실시...응답률 최소 10% 넘어야”.
- 뉴스1. 2024.10.25. “국민 73%, 북러 군사협력 '핵실험 수준' 위협으로 느껴[갤럽]”
- 강원일보. 2024.10.24. “[도민 의식 조사]도민 4명 중 1명만 ‘남북 통일’ 바란다”

- Leeper, T.J. 2019. “Where Have the Respondents Gone? Perhaps We Ate Them All.” *Public Opinion Quarterly* 83(S1): 280-288.
- Lorenc, B., G. Loosveldt, M.H. Mulry, and D. Wrighte. 2013. “Understanding and Improving the External Survey Environment of Official Statistics.” *Survey Methods: Insights from the Field* 161: 1-14.
- Traugott, M.W. and P.J. Lavrakas. 2008. *The Voter's Guide to Election Polls*. (4th edition). Lanham, MD: The Rowman & Littlefield Publishers.

<접수 2024.10.18; 수정 2024.11.06; 게재확정 2024.11.24>

Main Topics in Election Poll News

Jibum Kim

(Sungkyunkwan University)

Deok Hyun Jang

(Korea Gallup Poll Co., Ltd.)

Sori Kim

(Sungkyunkwan University)

The proliferation of election polls and the decline in response rates have led to public distrust of election polls. To tackle these challenges, the government, political entities, and polling organizations have tried to reform the election polling system, yet the problems persist. While there is no legal definition of an election poll, there are exceptions where certain polls are not considered as such. As a result, it seems difficult for the public to clearly distinguish between election and non-election polls. Therefore, even if people recognize the importance of election polls over non-election polls, they are unlikely to feel a particular motivation to participate if they cannot differentiate between them. This study indirectly assesses whether people can intuitively distinguish between election and non-election polls by analyzing ‘election poll’ news articles. The analysis revealed that election polls, conducted frequently, include not only election-related topics but also subjects like attitudes toward North Korea or political scandals, which may seem unrelated to elections. Moving forward, if the National Election Poll Deliberation Committee sets specific standards for the duration, number of questions, and content of election polls, it could serve as a starting point for resolving the difficulties surrounding election polling and bring positive changes to the overall polling environment.

Key words: election polls, National Election Survey Deliberation Commission (NESDC), election news