

토픽모델링을 활용한 여군에 대한 언론 보도 연구

조선웅*

초 록

본 연구에서는 여군에 대한 언론 보도 구조 파악을 위해 1990년부터 2022년까지의 여군에 관한 동아일보와 중앙일보, 한겨레와 경향신문 기사를 분석하였다. 빅카인즈를 통해 자료를 수집하였고, 파이썬을 활용하여 토픽모델링 등의 분석 방법을 적용하였다. 분석 결과, 보수와 진보 언론사는 여군 관련 보도에 있어서 다소 차이를 보였으나 그러한 차이는 시간이 지나면서 축소되었고, 최근으로 올수록 양 언론사 모두 성폭력 관련 이슈를 중요하게 보도하고 있었다. 이러한 연구 결과를 바탕으로 몇 가지 함의를 도출하여 제시하였다.

주제어 : 여군, 뉴스, 토픽모델링, 동적 토픽모델링

* 서울대학교 사회학과 박사수로 (skingsuper@gmail.com)

I. 서론

한국에서 여군에 대한 언론 보도 및 연구가 증가한 것을 보면 많은 사람들이 여군에 대해 관심을 갖고 있는 것 같다¹⁾. 1950년 여자의용군 500여 명을 시초로 하는 한국 여군은 2020년 1만 3천 명을 넘어섰다. 이는 전체 간부 대비 7.4% 정도인데, 2022년에는 그 비율이 9%로 증가하였다. 국방부는 2027년까지 여군 비율 15.3% 달성 계획을 가지고 있어서, 앞으로도 여군 규모가 확대될 것으로 보인다(국방부, 2020: 123; 국방부, 2023: 287)²⁾. 또한 2019년에 여군 최초 소장 진급자가 나오는 등 상위 계급으로의 진출도 활발히 이루어지고 있다. 남성 병역자원 한계, 여권 신장, 국방과학기술 발달 등의 상황을 고려하면(김경순, 2011), 향후 여군 증가 및 역할 확대 추세가 지속될 것 같다.

이런 맥락에서 여군에 대한 다양한 연구들이 이루어지고 있는데, 본 연구에서는 그동안 주목하지 않았던 여군 관련 뉴스를 분석해보려 한다. 쉽게 접할 수 있는 텍스트 자료인 뉴스에 대해서 통상 질적인 분석이 이루어져 왔으나, 최근 분석 방법의 발달에 힘입어 보다 더 면밀한 분석이 가능해졌다. 요즈음 널리 활용되어 영향력이 증가한 SNS는 다소 개인의 경험에 초점을 맞춘 반면, 뉴스는 대체로 사회적 관심을 반영하고 있다는 측면에서 중요성이 있다(Iyengar, 1994; Rapp et al., 2013).

뉴스가 사실을 전달해야 한다는 데에는 이견이 없을 것이다. 정확한 정보 전달이 뉴스의 가장 기본적인 기능 중 하나이기 때문이다. 하지만 정보 전달이 단순한 사실의 전달에 그치는 경우는 드물다. 왜냐하면 많은 이슈들이 사실 전달뿐만 아니라 가치 판단을 요구하는 영역을 포함하고 있기 때문이다. 이처럼 관점이 요구되는 이슈들에 대해서 각 언론사는 그들이 갖고 있는 정체성이나 가치관, 처한 환경 등에 따라 다른 목소리를 낸다(박윤정 외, 2023; Baron, 2006). 이를 이분법적으로 구분하는 것은 개별 언론사의 입장을 단순화할 뿐만 아니라 언론사 간의 차이를 간과하는 단점이 있지만, 이러한 이분법

1) 빅카인즈 전국일간지 대상 여군 검색 기사는 1990년 59건, 2000년 163건, 2010년 240건, 2020년 486건이다. RISS에서 제목에 여군이 포함된 학술논문은 1990년-1999년 4건, 2000년-2009년 15건, 2010년-2019년 51건이다.

2) 칸터(Kanter, 1993)에 따르면 조직 내 여성 비율이 약 15%보다 낮게 되면, 개인 특성이 여성이라는(열세적) 지위에 묻히게 된다. 하지만 여성 비율이 15%를 넘어서서 35% 정도에 이르게 되면 여성끼리 연대할 수 있을 뿐만 아니라 조직 문화에 영향을 미치는 것이 가능하다. 한국 여군이 국방부의 계획대로 2027년경 15% 비율을 넘어서게 된다면, '토큰(tokens)'의 지위에서(의미 있는) '소수자(minority)'의 지위로 거듭나게 될 것이라 기대할 수 있다.

이 통용되는 것이 엄연한 현실이기도 하다. 이러한 이분법은 흔히 ‘보수’와 ‘진보’로 불린다.

특히, 사람들은 본인의 성향에 따라 특정 언론사의 기사에 더 많이 노출되는 경향이 있다. 즉, 사람들은 특정 이슈에 대해 본인의 생각과 방향이 유사한 뉴스를 들으며 본인의 성향을 더 강화해 나가는 것이다. 심지어 개인이 가지고 있는 가치관과 일치하지 않은 정보를 접할 때 그러한 정보에 대해 신뢰하지 않는 경향을 보인다(김영옥, 2009; Gentzkow & Shapiro, 2006). 그 결과, 성향이 다른 사람들 간 생각 차이는 더 깊어질 수 있고, 심할 경우 이는 갈등으로 표출될 수 있다.

한국의 보수와 진보 언론사 역시 많은 사안들에 대해 다른 목소리를 낸다(김영옥, 2011; 이재경, 2008). 안보 영역을 예로 들자면, 보수의 경우 상대적으로 북한을 대결 상대로 보고 한미동맹을 강조하는 경향이 있다. 반면 진보의 경우 상대적으로 북한을 대화 상대로 보고, 한미동맹에 대해서도 다소 유연한 경향을 취하기도 한다(김경희·노기영, 2011; 손영준, 2004; 신동호, 2020). 젠더 영역의 경우 보수는 상대적으로 남성의 입장에 서는 경향이 있으며, 진보는 상대적으로 여성주의적 시각을 견지하는 경향이 있다(이수범·송민호, 2020; 홍지아, 2022). 예를 들어 보수 언론은 대체로 여성주의 시각을 축소하거나 부정적으로 바라볼 뿐만 아니라 젠더 갈등의 원인보다는 갈등 그 자체를 보도하는 경향이 있다(홍지아, 2017).

그동안 여군 이슈에 대해 한국의 보수와 진보 언론사가 어떠한 입장을 보이는지 분석한 선행 연구는 찾아보기 힘들다³⁾. 하지만 앞서 살펴본 내용을 토대로 한다면 다음과 같이 예상해 볼 수 있다. 보수 언론사는 상대적으로 여성이 입대하여 군 조직의 일원으로서 어떠한 쓰임을 받고, 그 결과 군 전투력에 어떻게 기여할 수 있는지에 집중할 것으로 보인다. 여성이 군 조직에 진출하는 과정에서 발생하는 여성 개인의 긴장과 갈등 그리고 부정적인 요소보다는 군 조직의 관점에 초점을 맞추어 여군의 기능적인 측면에 주목할 것으로 예상된다. 진보 언론사는 상대적으로 군 조직의 입장보다는 여군 개인의 입장에서 사회화 과정 간 겪는 어려움에 주목할 것으로 예상된다. 또한, 여성이 군에 진출함으로써 얻게 되는 장점보다 이 과정에서 발생하는 부작용에 더 초점을 맞출 것으로 보인다.

그런데 그동안 여군이 큰 후퇴 없이 양적으로 성장하고 질적으로도 발전해

3) 여군 관련 뉴스를 분석한 연구로는 정인정 외(2022)가 있다. 그런데 이는 시기적으로 2014-2019년으로 국한되고, 보수와 진보의 차이에 대해 논하지 않고 있으며, 시기에 따른 토픽의 변화를 분석하지 않고 있어 본고와 다소 차이가 있다.

왔다는 점을 고려하면 보수와 진보 언론사의 여군 관련 뉴스의 차이가 크게 두드러지지 않을 수도 있다. 앞서 언급했다시피 약 500여 명에 불과했던 여군은 현재 1만 3천 명을 넘어섰고, 소장 진급자가 나왔으며 군의 다양한 병과 및 직위에 진출했다. 앞으로도 특별한 구조적 변화가 전제되지 않는 한 여군이 양적으로 확대되고 그들의 역할이 늘어날 것이라고 볼 수 있다. 따라서 이러한 여군 발전의 당위성과 경향성을 고려한다면 의외로 보수와 진보 언론사의 여군 관련 보도가 큰 차이를 보이지 않을 수도 있다.

하지만 표면적으로는 여군 발전 과정이 순탄하게 이루어진 것 같지만, 그 이면을 살펴보면 다양한 긴장과 갈등이 수반되었다는 점도 고려해야 한다. 여군 발전은 거저 주어진 것이 아니라 가족을 포함한 각계각층의 노력과 희생에 기반하여 이루어졌다(조선웅, 2015). 군 조직 내부에서는 여군과 남군의 긴장과 갈등이 존재했고, 이러한 과정 속에서 많은 여군들의 좌절이 있었다(김재은·김지현, 2021).

이러한 배경 속에 본고에서는 그동안 여군 관련 뉴스가 어떻게 보도되었는지 살펴보려 한다. 특히, 보수와 진보 언론의 여군 관련 보도가 어떠한 공통점과 차이점을 갖는지 분석하고 그 함의를 찾으려 한다. 또한, 시간에 따라 여군 관련 뉴스 보도가 어떻게 변화되었는지를 분석할 것이다. 뉴스와 우리의 인식 및 행동 등이 상호 영향을 주고받는다라는 점을 감안한다면, 본고의 작업은 향후 군 조직 내 여군의 진출 및 적응, 역할 수행 등에 유용한 함의를 제공할 수 있을 것이다.

II. 선행 연구 검토: 여성의 군 조직 진출과 갈등

군 조직은 오랜 시간 동안, 거의 모든 지역에 걸쳐서 남성들의 영역이었다. 군 조직의 성별 구성을 근거로 들면, 비교적 최근까지도 남성은 군 조직의 약 97%를 차지하고 있다(Goldstein, 2001: 10). 공동체의 존립이 위태로운 상황에서는 군에서 여성의 비율이 높아지기도 하였지만, 전쟁이 끝나면 대체로 이들은 다시 군에서 자리를 잃고 떠나야 했다(Carreiras, 2006: 5-12).

한국의 경우도 국제 사회의 경향에서 크게 벗어나지 않는다. 한국전쟁 발발로 국가에 큰 위기가 닥치자 여자의용군이 조직되어 여성이 군 조직에 참여하게 되었다. 1950년 9월, 모집된 500명 중 491명이 훈련을 수료하고 여자의용군 1기로 임관하였다. 임관 후 전방 부대에 배치받게 된 여군은 문서 연락,

필서, 통역, 정훈, 위문 등의 임무를 담당했다. 1951년 8월 이후 전선이 다소 안정되자 당국에서는 여군을 전선의 후방에 배치하였고, 이에 반발하여 전방 지역에서 계속 근무하기를 원했던 여군들은 전역하게 되었다. 전역하지 않은 여군들은 계속해서 군에 남아 임무를 수행했지만, 주로 후방에서 남군을 보조하는 임무에 국한될 뿐이었다(민경자, 2008).

이후 군 조직에서 주변적인 여군의 지위는 약 40년 동안 큰 변화가 없었다. 먼저 여군의 규모는 1989년에도 755명에 불과했다(국방부, 2011: 296). 여자의용군이 약 500명으로 출발했다는 것을 상기하면 그동안 여군 규모에 거의 변화가 없었다고 볼 수 있다.

그뿐만 아니라 이 시기 동안 여군들은 남군들처럼 병과 선택의 자유를 가질 수 없었다. 왜냐하면 여군은 ‘여군병과’에 소속되어 보조적인 역할만을 수행할 수 있었기 때문이다. 자노비츠(Janowitz, 2017[1961])에 따르면 군 조직에서는 필수 경력에 대한 믿음이 강하고, 이에 따라 엘리트 군인이 되기 위해서는 필수 경력을 성실하게 이행해야 한다. 이러한 필수 경력은 전투 임무를 주로 하는 지휘관과 참모직을 번갈아 하는 것을 의미한다. 그런데 여군은 제도적으로 전투 병과에 대한 접근이 차단되어 있기 때문에 상위직으로 진출하는 것이 불가능했다.

또한 이 시기 동안 여군은 여성으로서의 기본권도 상당 부분 침해당했다. 짧은 예외 기간을 제외하면 기혼자는 군에 지원할 수 없었고, 미혼으로 입대한 여군은 결혼과 출산이 금지되기도 하였다. 일부 예외 기간이 있긴 하지만 대체로 하사관은 결혼과 출산이 모두 금지되었고, 장교의 경우는 결혼은 허용하되 출산은 금지하는 기조가 지속되었다(국방부, 2011; 민경자, 2008).

하지만 이러한 분위기는 1980년대 후반부터 바뀌게 된다. 1988년경에는 대통령 지시에 따라 여군 역할 확대에 관한 연구가 각 군에서 이루어졌고, 1988년부터 여군 장교와 하사관의 출산이 허용되었으며(국방부, 2011; 민경자, 2008: 344), 여군병과가 해체되어 1990년부터는 여군들이 군의 다양한 병과로 진출할 수 있게 되었다. 여군병과 해체 이후 여군 신규 임용 규모를 늘리면서 1999년에는 2,085명의 여군이 군에 복무하고 있었다(국방부, 2011: 502).

그런데 이 과정에서 여군의 복무를 촉진하기 위한 제도의 정비는 군 내부에서 주도적으로 이루어지기보다 군 외부의 영향으로부터 비롯되는 경우들이 있었다. 예를 들어, 군 내 불임휴직제도가 2010년도에 신설될 것은 여군의 필요성에 의해서 선제적으로 만들어진 것이 아니라, 국가·지방 공무원에게 먼저 적용되었던 것이 군대로 확대된 것이다(국방부훈령 제1279호). 오히려 군 내

부에서는 1999년 여군담당관의 건의에 의해 보건휴가가 폐지되는 등 모성보호제도에 다소 소극적인 입장이 었보인다(국방부, 2011: 507).

이러한 갈등에도 불구하고 구조적인 변화가 이루어진 1990년 이후 대체로 여군 복무를 촉진하기 위한 제도의 정비가 지속적으로 이루어졌다. 1990년 여군병과가 해체되었고, 1997년부터 1999년까지 공사, 육사, 해사에 최초의 여생도가 입학하게 되었다. 여군병과 해체는 여군이 군의 다양한 병과로 진출하는 것을 가능하게 하였고, 각 군 사관학교 입학 허용은 여성이 남성과 동일한 양성교육을 받게 하였다. 여자의용군교육대를 모체로 하여 그동안 여군만을 양성했던 여군학교도 2002년에 문을 닫게 되었다. 양성과정, 병과 선택 등에서 군 조직의 주류와 분리되어 있던 여군이 어느 정도 군 조직에 통합되었다고 볼 수 있다.

2000년대에는 여군의 군 조직 내 진출을 보여주는 상징적인 일들이 다수 발생했다. 각 군 사관학교를 졸업한 여군 장교들이 그동안 여성에게 허용되지 않았던 직위에 최초로 진출하게 된 것이다. 또한 이 시기에는 여군 최초의 전투 병과 준장 진급자가 나오게 되었다.

2010년대로 접어들면서는 그동안의 제도적 통합에도 불구하고 일부 남아있던 여군에 대한 배제가 대부분 폐지되었다. 2011년에 최초 여 학군단이 설치되었고, 2014년에는 육군 삼사관학교가 여생도 입학을 허용하게 되었다. 또한 여군병과 폐지 이후에도 여군을 받아들이지 않았던 기갑, 포병 등의 병과가 여군에게 개방됨으로써, 2014년을 기점으로 육군의 모든 병과가 여군을 받아들여지게 되었다. 그동안 여군이 전투 직위를 맡더라도 보병 부대 지휘관(자)이 아닌 신병 부대 지휘관(자)에 임명되는 등 남녀가 차별적인 경력관리모형을 적용받았지만, 2018년 관련 법령이 개정되면서 남녀 동일 경력관리모형을 적용하게 되었다.

그런데 이러한 제도의 개선이 꼭 여군의 복무 환경 개선 및 군 조직으로의 통합을 의미하지는 않을 수 있다. 종합적인 여건이 성숙되지 않은 상황에서 여군을 위한 제도가 정비될 경우 남군은 이를 불공정하다고 느끼며 상대적 박탈감을 호소하기도 하고, 그 결과 제도적인 차별의 개선에도 불구하고 보이지 않는 배제와 차별이 지속될 수 있다는 점에 주목해야 한다(Yoder, Adams, & Prince, 1983: 334).

특히, 이 과정에서 신체는 중요한 의미를 갖는다. 애초에 여성의 군 조직 참여를 가로막던 핵심적인 명분은 상대적으로 '부족한' 신체 능력이었다. 과학기술의 발달 및 군사 전략의 변화로 더 이상 신체 능력이 과거와 같은 효용을 갖기 어려움에도 불구하고 여전히 신체 능력은 여군을 이등 군인으로 속박하

는 효과를 지니기도 한다. 때로는 신체 능력 그 자체에 의해서 여군은 열등한 군인이 되기도 하지만 남성의 신체는 우월하고 여성의 신체는 열등하다는 오래된 관념이 문화로 굳어져서 여군을 옥죄기도 한다(김엘리, 2012: 164-167; Centeno & Enriquez, 2016: 13-17). 신체 능력 그 자체는 여군이 노력을 통해서 증명해 낼 수 있는 영역이지만 두 번째의 경우에는 쉽게 극복될 수 있는 것이 아니다.

그런데 중요한 점 중 하나는 남군이 여군을 온전한 동료로 여기지 않고, 성적인 대상으로 보는 경향이 있다는 것이다. 신체를 매개로 한 여군에 대한 차별적 인식은 비교적 양성평등한 군대문화를 가지고 있는 미군에서도 관찰된다. 미국의 경우 남군이 여군을 바라보는 관점은 흔히 slut-bitch의 이분법으로 이해된다. 이는 여군이 성적으로 접근 가능하면 slut이라 칭하고, 그렇지 않으면 bitch로 명명하는 것이다. 즉, 남군이 여군을 바라보는 관점은 군인적인 요소가 전혀 없이 오로지 그들의 성적인 신체에 집중되어 있는 것이다(King, 2015).

한국의 경우도 녹록치 않다. 한국은 가부장제와 남아선호사상이 뿌리 깊은데, 이는 상대적으로 여성의 권리가 취약하다는 것을 의미한다. 또한 여전히 군 병력 충원의 상당 부분은 남성 징병에 의존하고 있기 때문에, 여군 간부의 증가에도 불구하고 여전히 군 조직 구성원의 대부분은 남성이고, 남성 중심적인 문화가 만연해 있다(김엘리, 2012; 조선웅, 2019). 이러한 가운데 비교적 최근 들어 여군 복무 편의를 위한 제도들이 빠르게 정비되고 있기 때문에 제도적 차별의 개선에도 불구하고 여군에 대한 은근한 차별과 괴롭힘이 발생할 가능성이 높다(류영숙, 2002).

이러한 경우 여군 신체는 은근한 차별과 괴롭힘을 수행할 최적의 대상이 될 수 있다. 또한 여전히 여군이 하위 계급에 과다대표되고 있다는 점을 고려하면(국회뉴스ON, 2019.10.10.), 군 조직의 위계서열이 신체를 매개로 한 차별 및 괴롭힘과 결합하여 부정적인 효과를 배가할 수 있다(김수미·이새롬·한민경, 2023; 피우진, 2017). 또한, 군 조직의 특수성, 이로부터 기인한 군사법원 시스템은 군 내 성폭력 사건을 더 악화시키는 원인으로 지목되기도 한다(이동임, 2021). 언론 보도에 따르면 군 내 성폭력 사고는 2017년 102건에서 2021년 999건으로 급격히 증가했다고 하여(여성신문, 2022), 이러한 우려가 괜한 것이 아님을 알 수 있다.

이처럼 한편으로는 여군을 위한 제도가 정비되고 여군의 규모가 확대되었으며 상위 계급으로의 진출이 이루어졌으나, 다른 한편으로는 여전히 여군이 군 조직에 통합되고 역량을 발휘하는 데 어려운 환경이 지속되고 있는 것으로 보인다.

Ⅲ. 자료와 방법론

본 연구에서는 여군을 검색어로 하여 수집된 언론 기사에 대해 분석하였다. 수집 기간은 1990년부터 2022년까지이다. 언론사는 ‘보수’와 ‘진보’를 각각 2개씩 선정하였다. 보수 언론사로는 동아일보와 중앙일보를, 진보 언론사로는 한겨레와 경향신문을 대상으로 하였다⁴⁾. 자료 수집은 여러 언론사의 뉴스를 제공하는 빅데이터를 활용하였다. 빅데이터에서 제공하는 자료의 형태는 본문에서 명사를 추출한 키워드이다.

본문에 여군이라는 단어가 포함된 기사가 모두 검색되다보니, 적절하지 않은 기사들을 제외하는 작업이 선행되어야 했다. 여씨 성을 가진 남성을 대상으로 한 기사, 여군이 언급되기는 했지만 정작 기사는 연예 관련 내용인 경우, 여군 대상 포토 뉴스여서 내용이 없는 경우, 중복 기사 등 분석에 적절하지 않다고 판단되는 기사들을 제외하였다. 특히 중앙일보의 경우 MBC 예능 프로그램 ‘진짜 사나이’에 출연했던 여성 연예인에 관한 중복 기사들이 상당히 많았는데, 이러한 기사는 여군과 거의 관련 없는 것들이어서 제외시켰다. 그 결과 동아일보와 중앙일보는 최초 2,820개의 기사에서 1,092건을 제외하고 1,728건을 분석에 활용하였다. 한겨레와 경향신문의 경우는 최초 1,876건에서 279건을 제외한 후 1,597건을 분석에 활용하였다. 검색 키워드인 ‘여군’, 신문명을 비롯한 몇몇 불필요 단어들과 1글자 단어는 불용어 처리하였다.

분석의 주요 방법은 토픽모델링을 활용하여 여군 관련 뉴스를 분석하는 것이다. LDA(Latent Dirichlet Allocation) 토픽 모델링은 문서가 복수 개의 토픽으로 구성되어 있고, 각 토픽은 단어의 집합으로 이루어졌다는 가정에 기반한다. 이때 토픽은 단어나 문서처럼 관찰 가능한 것이 아니라 숨겨진 것인데, 이를 파악하면 대규모 텍스트 자료를 효과적으로 이해할 수 있다는 장점이 있다(Blei, 2012; Blei, Ng, & Jordan, 2003).

그런데 문서에 시간 정보가 있는 경우에는 시간의 흐름에 따라 토픽의 비중이 어떻게 변화하였는지를 확인할 수 있다. 이는 시간 정보가 무시되는 LDA 토픽 모델링의 단점을 보완할 수 있는 방법이다. 먼저, 특정 시간 동안 만들어진 토픽 분포의 평균을 계산해서 토픽 분포의 변화를 확인하는 방법이 가능하다. 그런데 이처럼 토픽 트렌드를 통해 시간에 따른 토픽의 변화를 확인하는 것은 토픽 분포의 변화에도 불구하고 토픽 그 자체는 바뀌지 않는다는 가정에

4) 조선일보의 경우 자료수집이 원활하지 않아 분석에서 제외하였다. 그럼에도 불구하고 보수 2곳, 진보 2곳의 언론사를 통해 자료 수집을 하고 있기 때문에 전반적인 경향성을 파악하여 분석하는 데 문제가 없을 것으로 생각된다.

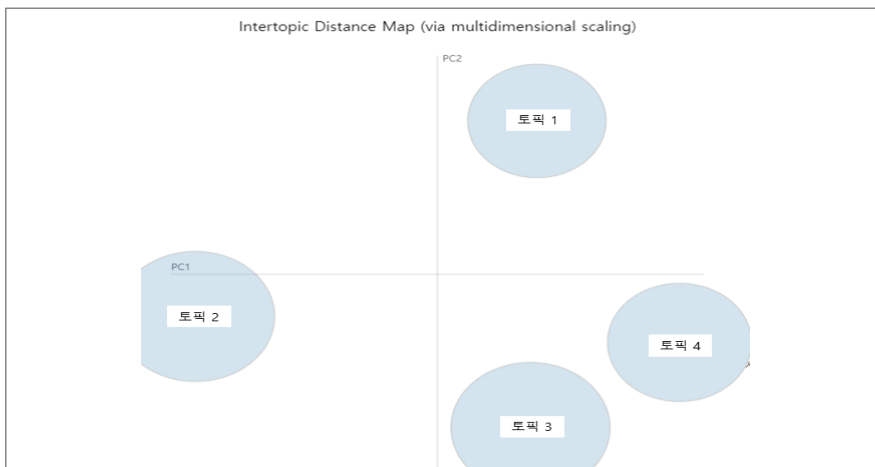
기반한다. 따라서 보다 더 현실적인 가정은 시간이 지나면서 토픽의 내용이 바뀔 수 있다는 점이고, 이에 기반한 방법이 바로 동적 토픽 모델링(Dynamic Topic Modeling)이다(Blei & Lafferty, 2006). 본고에서는 시간에 따른 토픽의 변화를 확인하기 위해 동적 토픽 모델링 방법을 활용하였고, 시간은 1년 단위로 설정하였다.

이상의 분석은 파이썬을 활용하여 진행하였고, 토픽 모델링 분석과 시각화에는 gensim 패키지 등을 이용하였다.

IV. 결과

1. 동아일보와 중앙일보에 대한 분석

먼저, 토픽 개수를 정하기 위해 gensim 패키지를 활용해 토픽 개수 2개부터 40개까지에 대해 응집도(coherence)와 혼란도(perplexity)를 확인하였다. 하지만 응집도는 대체로 토픽 개수가 적을수록 높은 반면 혼란도는 토픽 개수가 많을수록 낮게 나와서 두 기준을 모두 만족하는 토픽 개수를 정할 수 없었다. 따라서 해석 가능성을 고려해 응집도에 더 초점을 맞추되 각 토픽별로 중복되지 않고 잘 구분되는지를 감안하여 토픽 개수를 4개로 정하였다. [그림 1]은 pyLDAvis를 활용하여 각 토픽별 거리와 분포를 시각화 한 것이다.



[그림 1] 보수 언론 토픽의 분포와 토픽 간 거리 시각화

토픽 개수 4개로 LDA 토픽모델링을 수행한 결과는 <표 1>과 같다. 첫 번째 토픽은 ‘여군 장교 훈련’으로 명명하였다. 각 군의 여군 장교들이 각 직위에서 훈련하며 임무를 수행하는 것과 관련된 내용이다. 과거에는 주로 행정이나 보조적인 일을 도맡아 했던 여군이 군의 핵심 영역이라고 할 수 있는 전투 직위에 처음 보직되고, 이러한 내용이 ‘최초’라는 표현과 함께 보도되었다. 그러한 측면에서 각종 항공기의 ‘조종사’는 여군이 주변부에서 탈피하여 군의 핵심 영역으로 진입하는 상징성을 더 극적으로 보여주는 단어로 보인다. 이처럼 여군의 진출에 대한 언론의 관심은 주로 장교에게 집중되었고 부사관은 다소 소외된 것으로 보인다.

두 번째 토픽은 ‘성폭력 사건’으로 명명하였다. ‘사건’, ‘성추행’, ‘피해자’, ‘조사’, ‘성폭력’ 등의 키워드를 통해 어떤 내용의 뉴스들이 두 번째 토픽으로 분류되었는지를 선명하게 보여준다. 눈에 띄는 점 중 하나는 ‘하사’로 이는 성폭력 사건의 피해자가 주로 하사를 포함한 하위 계급이라는 점을 시사한다.

<표 1> 보수 언론 여군 뉴스 토픽 모델링 결과

구분	토픽 1	토픽 2	토픽 3	토픽 4
토픽명	여군 장교 훈련	성폭력 사건	국가와 전쟁	여군 지원 및 복무
키워드	1 여성	사건	여성	여성
	2 장교	육군	대통령	한국
	3 육군	부대	북한	복무
	4 해군	성추행	미국	북한
	5 조종사	국방부	미군	지원
	6 공군	피해자	사람	군인
	7 훈련	하사	부대	남성
	8 대위	수사	한국	훈련
	9 선발	조사	훈련	후보
	10 전투	성폭력	영화	교육
전체 토큰 대비 각 토픽 비중	21.2%	27.8%	28%	23%

세 번째 토픽은 ‘국가와 전쟁’으로 명명하였다. 주요 키워드로 북한, 미국, 한국의 국가명이 눈에 띈다. 특정 국가의 군 관련 활동이나 국가 간 관계 속에서 여군이 언급된 뉴스들이 세 번째 토픽으로 분류되었다.

네 번째 토픽은 ‘여군 지원 및 복무’로 명명하였다. 특히 여성의 경우 병역의 의무가 없는 상황에서 자발적으로 지원하여 직업으로서 복무하게 되는데,

이와 관련된 뉴스들이 이 토픽으로 분류된 것으로 보인다.

〈표 2〉는 동적 토픽 모델링을 수행하여 보수 언론 첫 번째 토픽의 키워드 변화를 제시한 것이다. 먼저, ‘여군 장교 훈련’에 있어서 상대적으로 육군에 가장 많은 관심이 쏠렸던 것을 알 수 있다. 이는 육군의 규모가 타군보다 커서 그만큼 보도할 내용이 많은 점에 기인할 수 있다. 2000년대 중반까지는 해군보다 공군에 대한 관심이 더 높게 나타났으나 그 이후에는 해군이 공군을 제치고 주요 키워드로 부상하는데, 이는 여군의 잠수함 최초 임무에 대한 기대감 등이 반영된 것으로 보인다.

다음으로, 여성(여자 포함) 키워드가 점점 후순위로 밀린 점에 주목할 필요가 있다. 이와 더불어 2010년까지 1-2순위 키워드에 머물렀던 장교라는 단어가 후순위로 밀렸음을 알 수 있다. 즉, 남성 장교가 당연시되던 상황에서 여성 장교의 출현 자체가 많은 관심을 받았다가 점차 시간이 지나면서 그보다는 이들이 수행하는 훈련으로 관심의 초점이 바뀐 것이다.

〈표 2〉 보수 언론 첫 번째 토픽 키워드 변화

연도	키워드				
	1	2	3	4	5
1990년	장교	훈련	육군	여성	여자
1991년		육군	훈련		
1992년			여성	공군	
1993년					여성
1994년					
1995년					
1996년					
1997년		여성	육군		
1998년					
1999년					
2000년					
2001년					
2002년					
2003년					
2004년	육군			장교	훈련
2005년		여성	훈련		
2006년			해군		
2007년					
2008년					

연도	키워드					
	1	2	3	4	5	
2009년	장교	육군	훈련	여성		
2010년	훈련	장교	육군	해군	여성	
2011년		해군	여성	장교	육군	
2012년				육군	장교	
2013년		여성	해군			
2014년		해군	육군	여성		장교
2015년		육군	해군			
2016년						
2017년						
2018년						
2019년	육군	훈련	공군	부대		
2020년		해군		훈련	장교	
2021년			장교		공군	
2022년						

〈표 3〉 보수 언론 두 번째 토픽 키워드 변화

연도	키워드						
	1	2	3	4	5		
1990년	사건	피고인	육군	장교	국방부		
1991년							
1992년							
1993년							
1994년							
1995년	피고인	사건	성추행	성추행	사단		
1996년							
1997년	사건	피고인			성추행	육군	성희롱
1998년							
1999년		성추행			장교		사단
2000년							
2001년	성추행	사건	육군	육군			
2002년							
2003년							
2004년	사건	성추행		육군	장교	사단	
2005년							
2006년							
2007년							

연도	키워드				
	1	2	3	4	5
2008년		육군	성추행		
2009년				부대	
2010년			부대	성추행	혐의
2011년					
2012년			성추행	부대	사단
2013년					
2014년			부대	성추행	혐의
2015년					
2016년		부대	육군	육군	피해자
2017년					
2018년			하사	성추행	전역
2019년					
2020년		하사	부대	중사	피해자
2021년					
2022년					

〈표 3〉은 동적 토픽 모델링을 수행하여 보수 언론 두 번째 토픽의 키워드 변화를 제시한 것이다. 먼저, ‘육군’, ‘사단’이 꾸준히 중요 키워드로 도출된 점이 눈에 띈다. 이는 사단장이 성폭력 사건의 가해자인 경우가 종종 있었고, 그에 대한 여파가 상당했기 때문인 것으로 보인다. 표에서 제시하진 못했지만 사단장 계급을 뜻하는 ‘소장’이란 단어가 1990년부터 2003년까지 10번째 키워드 안에 등장하고 있다.

다음으로 ‘사건’이란 키워드와 ‘성추행’이란 키워드를 살펴볼 필요가 있다. 5번째 키워드 안에 ‘성폭력’이란 단어를 찾아볼 수 없고, 10번째 키워드 안에서 ‘성폭행’이란 단어도 찾을 수 없다. 해당 기간 동안 여군에 대한 성폭행 사건이 없었던 것은 아니다. 대체로 ‘성추행 사건’이란 표현이 가장 많이 사용된 것으로 보인다.

마지막으로 1998년까지 ‘피고인’ 키워드가 중요하게 등장하다가 사라진 점, 2018년부터 ‘피해자’ 키워드가 5번째 안에 등장한 것에 주목할 필요가 있다. 보수 언론의 ‘성폭력 사건’에 대한 뉴스가 과거에는 가해자인 ‘피고인’에게 초점을 더 맞추었다면, 최근에는 피해자인 여군에게 더 초점을 맞추고 있는 것으로 보인다.

〈표 4〉는 동적 토픽 모델링을 수행하여 보수 언론 세 번째 토픽의 키워드 변화를 제시한 것이다. 먼저, 1990년부터 2000년 중반까지 이라크가 주요 키

워드로 도출된 것은 걸프 전쟁, 이라크 전쟁, 아부그라이브 교도소 가혹행위 사건 등의 영향으로 보인다. ‘북한’ 키워드는 이라크에 대한 관심이 집중되었을 때를 제외하고는 상위 키워드로 계속 등장하고 있음을 알 수 있다. 또한 전 기간에 걸쳐서 ‘미국’ 키워드 역시 중요하게 등장하고 있는데, 특히 이라크와 전쟁을 수행하는 기간에는 미국뿐만 아니라 ‘미군’도 주요 키워드로 도출되고 있다. 키워드를 상위 10개까지 확대했을 경우에는 최근으로 올수록 ‘중국’과 ‘우크라이나’ 키워드도 등장하고 있다. 즉, 국제사회에서 발생하는 주요 전쟁과 여군, 그리고 한국을 중심으로 한 국제관계 관련 여군 뉴스들이 세 번째 토픽으로 분류되었다고 볼 수 있다.

〈표 5〉는 동적 토픽 모델링을 수행하여 보수 언론 네 번째 토픽의 키워드 변화를 제시한 것이다. 네 번째 토픽명은 ‘여군 지원 및 복무’인데, 시간이 흐르면서 토픽의 내용이 변화하고 있음을 볼 수 있다. 1990년대와 2000년대 초 중반까지는 여성이 군에 지원해서 시험을 치르고 선발되어 복무하는 내용들이 이 토픽의 주제로 보인다. 주요 키워드를 10개까지 확대하였을 경우 ‘취업’이라는 키워드도 도출되고 있어, 여성이 군을 직업 중 하나로서 바라보고 접근하는 경향을 포착할 수 있다.

〈표 4〉 보수 언론 세 번째 토픽 키워드 변화

연도	키워드				
	1	2	3	4	5
1990년	이라크	미국	사람	대통령	북한
1991년					
1992년		사람	미국		
1993년					
1994년	사람	미국	이라크	북한	대통령
1995년	미국	사람	북한	이라크	
1996년	사람	북한	미국	미군	
1997년		미국	북한		한국
1998년					대통령
1999년		북한	사람	미국	사람
2000년	미국		미군		
2001년		미군		북한	
2002년	이라크		미국	북한	사람
2003년					영화
2004년				대통령	북한

연도	키워드					
	1	2	3	4	5	
2005년	북한			북한	대통령	
2006년		미군		이라크	한국	
2007년		미국	미군	한국	대통령	
2008년		한국	미국	미군	전쟁	
2009년				전쟁	미군	사람
2010년						
2011년				대통령		사람
2012년						사진
2013년		대통령	한국	미국	중국	
2014년						
2015년			미국	한국	사람	
2016년						
2017년	대통령	북한	미국	한국	사람	
2018년				사람	여성	
2019년	북한	대통령	여성	미국	사람	
2020년	대통령	북한		사람	미국	
2021년						
2022년						

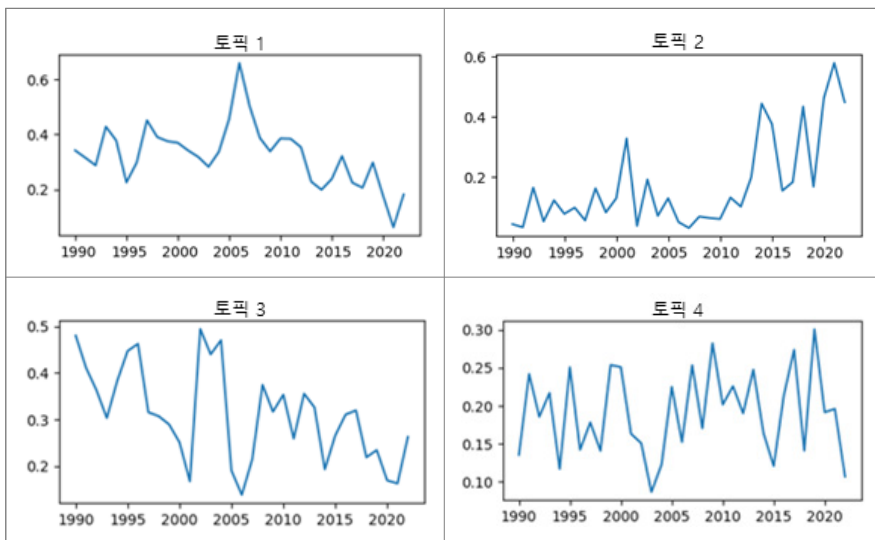
2010년대 중반부터는 ‘국방’ 키워드와 ‘병력’ 키워드가 등장하고 있는데, 이는 복무 제도의 주요 내용이 더 이상 여성이 군에 지원하여 복무하는 그 자체에만 있지 않음을 보여준다. 키워드를 10개까지 살펴보면 점점 국가, 국방부, 군대 등의 키워드가 나타나는데, 이는 남성 병역 자원의 한계와 맞물려 여군 증가의 당위성 부여, 여군 병력의 활용 등에 관한 주제의 뉴스들이 보도되었기 때문인 것으로 보인다.

〈표 5〉 보수 언론 네 번째 토픽 키워드 변화

연도	키워드					
	1	2	3	4	5	
1990년	여성	사회	복무	시험	제도	
1991년					남성	
1992년			남성	남성	시험	
1993년				복무		
1994년				복무		
1995년					시험	
1996년		남성	사회			
1997년			대학			
1998년						
1999년			사회			
2000년			대학		지원	
2001년			대학			
2002년			지원		후보	
2003년						
2004년		지원	남성			
2005년			후보	채용		
2006년			남성		복무	
2007년		복무	지원	남성	국방	
2008년			국방	의원		
2009년		국방	의원	복무	국방부	
2010년			복무	국방부	남성	
2011년		남성		남성	국방부	
2012년				국방	병력	
2013년				병력	국방	
2014년						
2015년						
2016년						
2017년						
2018년						
2019년						
2020년						
2021년						
2022년						

[그림 2]는 시간에 따른 보수 언론 여군 관련 뉴스 각 토픽 비중 변화를 시각화 한 것이다. 먼저 가장 눈에 띄는 것은 두 번째 토픽인 ‘성폭력 사건’의 비중이 시간이 흐름에 따라 상당히 증가한 것이다. 특히 2010년 이후로 증가세가 확연해져서 최근에는 가장 높은 토픽 비중을 보이고 있다.

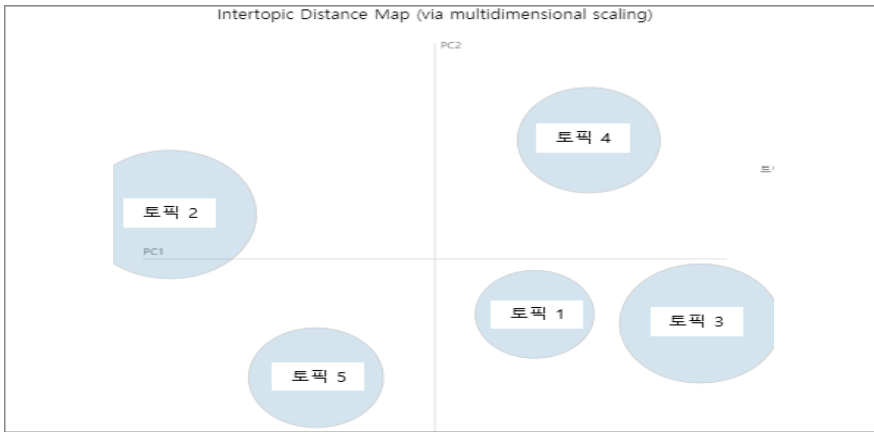
다음으로 다른 토픽의 비중은 대체로 큰 변화가 없거나 감소하는 경향을 확인할 수 있다. 첫 번째 토픽인 ‘여군 장교 훈련’의 경우 2005년에서 2010년 사이 토픽 비중이 상당히 높았음을 알 수 있는데, 이 시기는 사관학교에 최초로 입학한 여성들이 임관 후 진급을 거듭하며 상징적인 ‘최초 진출’을 연이어 하던 때이다. 비중이 가장 높을 때를 기준으로 보자면 여군 뉴스의 전반적인 흐름이 ‘여군 장교 훈련’에서 ‘성폭력 사건’ 토픽으로 변화하고 있음을 알 수 있다.



[그림 2] 보수 언론 여군 관련 뉴스 각 토픽 비중의 변화

2. 한겨레와 경향신문에 대한 분석

한겨레와 경향신문에 대해서도 앞선 경우와 마찬가지로 gensim 패키지를 활용해 토픽 개수 2개부터 40개까지에 대해 응집도와 혼란도를 확인하였으나, 두 기준을 모두 만족하는 토픽 개수를 도출할 수 없었다. 따라서 보수 언론 뉴스 분석에 적용했던 기준처럼 응집도에 초점을 맞추고, 각 토픽별로 중복되지 않도록 하여 토픽 개수를 5개로 정하였다. [그림 3]은 pyLDAvis를 활용하여 각 토픽별 거리와 분포를 시각화 한 것이다.



[그림 3] 진보 언론 토픽의 분포와 토픽 간 거리 시각화

〈표 6〉 진보 언론 여군 뉴스 토픽 모델링 결과

구분		토픽 1	토픽 2	토픽 3	토픽 4	토픽 5
토픽명		여군 지원	성폭력 사건	국가와 전쟁	여군과 사회	여군 복무
키워드	1	대학	사건	여성	하사	국방부
	2	인권	피해자	남성	사람	의원
	3	여성	성폭력	한국	전역	육군
	4	학교	부대	사회	군인	부대
	5	지원	해군	군대	사건	국방
	6	국방부	성추행	여자	생각	사건
	7	사건	육군	전쟁	트랜스젠더	장관
	8	교육	장교	군인	복무	공군
	9	육군	공군	남자	대통령	진급
	10	장교	피해	후보	사회	인사
전체 토큰 대비 각 토픽 비중		13%	27.7%	23.8%	18.7%	16.7%

〈표 6〉은 진보 언론 여군 뉴스에 대해 토픽 개수 5개로 LDA 토픽 모델링을 수행한 결과이다. 첫 번째 토픽은 ‘여군 지원’으로 명명하였다. 대학을 졸업하고 군에 지원하는 것과 관련된 뉴스들이 이 토픽으로 분류되었다. 두 번째 토픽은 ‘성폭력 사건’으로 가장 많은 비중을 차지하고 있다. 특히 ‘해군’과 ‘공군’이 주요 키워드로 도출되어서 보수 언론의 ‘성폭력 사건’ 토픽 키워드와 다소 차이를 보이고 있다.

세 번째 토픽은 ‘국가와 전쟁’으로 명명하였다. ‘한국’, ‘군대’ 등의 키워드가 도출되었고, 키워드의 범위를 조금 넓힐 경우 ‘미국’, ‘국가’ 등의 키워드가 나타났다. 상대적으로 보수 언론의 ‘국가와 전쟁’ 토픽에서 다른 국가명이 상위 키워드로 출현하였던 것에 반해, 진보 언론의 ‘국가와 전쟁’ 토픽은 ‘한국’, ‘사회’ 등이 주요 키워드로 도출되는 것으로 보아 국내 문제에 좀 더 초점을 맞추고 있는 것으로 보인다.

네 번째 토픽은 ‘여군과 사회’로 명명하였다. 다른 토픽들이 대체로 군 조직 구성원으로서의 ‘군인’에 초점을 맞추고 있는 것에 반해 이 토픽에서는 군 조직과 사회의 경계에서 발생하는 일들을 조망한 것으로 보인다. 변희수 하사 사건도 이 토픽에 포함이 된다. 마지막 다섯 번째 토픽은 ‘여군 복무’로 명명하였다. 각 군 및 부대에서 복무하고 진급하는 것과 관련된 뉴스들이 이 토픽으로 분류된 것으로 보인다.

〈표 7〉은 동적 토픽 모델링을 수행하여 진보 언론 첫 번째 토픽의 키워드 변화를 제시한 것이다. 이 토픽은 여성이 군에 지원하여 군인이 되는 것과 관련된 것인데, 시간이 지남에 따라 ‘여성’이라는 키워드가 점점 뒤로 밀리고 있는 것을 확인할 수 있다. 그와는 대조적으로 ‘국방부’나 ‘복무’와 같은 키워드들은 대체로 시간이 지남에 따라 중요도가 올라가고 있는데, 이는 여군 모집에 있어 ‘여성’이라는 젠더가 강조되던 과거의 경향이 벌어지고 있기 때문인 것으로 보인다.

〈표 7〉 진보 언론 첫 번째 토픽 키워드 변화

연도	키워드				
	1	2	3	4	5
1990년	여성	지원	모집	인력	대학
1991년					
1992년					
1993년					
1994년					
1995년			대학	인력	
1996년					
1997년					
1998년					
1999년					
2000년		대학	모집	국방부	
2001년					
2002년					
2003년					
2004년					
2005년			국방부	모집	
2006년				국방	
2007년					
2008년					
2009년					
2010년		지원	국방		
2011년				복무	
2012년					
2013년					
2014년		지원	대학	국방	
2015년			복무	국방	
2016년	국방		복무		
2017년	복무		제도		
2018년	여성	복무	제도	정부	
2019년		정부	복무	제도	
2020년		위원회			
2021년		위원회			여성
2022년	인권위				

〈표 8〉은 동적 토픽 모델링을 수행하여 진보 언론 두 번째 토픽의 키워드 변화를 제시한 것이다. 2010년대 중반부터 ‘성폭행’, ‘성폭력’, ‘피해자’ 등의 키워드가 도출되고 있음을 알 수 있다. 이를 통해 유추해보면 점점 시간이 지남에 따라 다양한 유형의 성폭력에 대한 보도가 이루어지고 있고, 피해자에게 초점을 맞추고 있는 것으로 보인다.

〈표 8〉 진보 언론 두 번째 토픽 키워드 변화

연도	키워드				
	1	2	3	4	5
1990년	사건	성희롱	성추행	육군	장교
1991년					
1992년					
1993년					
1994년					
1995년					
1996년					
1997년					
1998년					
1999년					
2000년	성추행	사건	장교		성희롱
2001년					
2002년					
2003년					
2004년	사건	성추행	육군	장교	혐의
2005년					부대
2006년					대위
2007년		대위	육군	장교	대위
2008년				성추행	부대
2009년				부대	성추행
2010년		육군	대위		
2011년					
2012년		부대			
2013년		육군	부대		
2014년		성추행	육군	부대	대위
2015년		육군	부대	성추행	성폭행
2016년		성폭력	피해자	부대	육군
2017년				조사	부대
2018년				부대	조사
2019년				수사	성추행
2020년					
2021년		공군			
2022년		사망			

〈표 9〉는 진보 언론 세 번째 토픽의 키워드 변화를 제시한 것이다. 미국과 이라크 간 전쟁 등이 주로 보도되던 것에서 점차 한국의 군대와 사회에 관한 것으로 토픽의 내용이 변화하였음을 알 수 있다.

〈표 9〉 진보 언론 세 번째 토픽 키워드 변화

연도	키워드				
	1	2	3	4	5
1990년	여성	이라크	미군	전쟁	미국
1991년					
1992년					
1993년					
1994년					
1995년		미군	이라크	미국	전쟁
1996년					
1997년			미국	미국	전쟁
1998년					
1999년		전쟁		미군	사회
2000년					
2001년			미군	전쟁	이라크
2002년					
2003년		미국	미국	이라크	전쟁
2004년					
2005년		미국	이라크	미군	
2006년					
2007년			미군	이라크	
2008년		전쟁			이라크
2009년			한국	미군	
2010년		사회			사진
2011년			남성		
2012년		한국		미국	남성
2013년			사회		
2014년		남성		여자	
2015년			한국		군대
2016년		사회		남성	
2017년			군대		한국
2018년		사회		남성	
2019년			군대		한국
2020년		사회		남성	
2021년			군대		한국
2022년		사회		한국	
2023년			군대		남성

〈표 10〉 진보 언론 네 번째 토픽 키워드 변화

연도	키워드						
	1	2	3	4	5		
1990년	투표	사람	여자	영화	생각		
1991년							
1992년							
1993년	사람	투표	군인	자신	복무		
1994년							
1995년		여자				영화	생각
1996년							
1997년							
1998년		영화				여자	군인
1999년							
2000년						생각	
2001년		영화					군인
2002년						생각	
2003년							
2004년							
2005년							
2006년							
2007년							
2008년							
2009년							
2010년							
2011년							
2012년							
2013년							
2014년							
2015년							
2016년							
2017년							
2018년	전역	하사	사람	트랜스젠더			
2019년			군인				
2020년	하사	전역	트랜스젠더	군인			
2021년							
2022년							

〈표 10〉은 동적 토픽 모델링을 수행하여 진보 언론 네 번째 토픽의 키워드 변화를 제시한 것이다. 과거 영화나 방송 등에서 조망된 여군 관련 내용들과 관련된 토픽에서 점차 ‘전역’, ‘엄마’ 등의 키워드가 도출되는 등 사적인 측면과 관련한 내용으로 변화하였음을 알 수 있다. 또한 최근으로 오면서 트랜스젠더의 군 복무 문제와 변희수 하사 관련 문제로 토픽의 내용이 변화하고 있다.

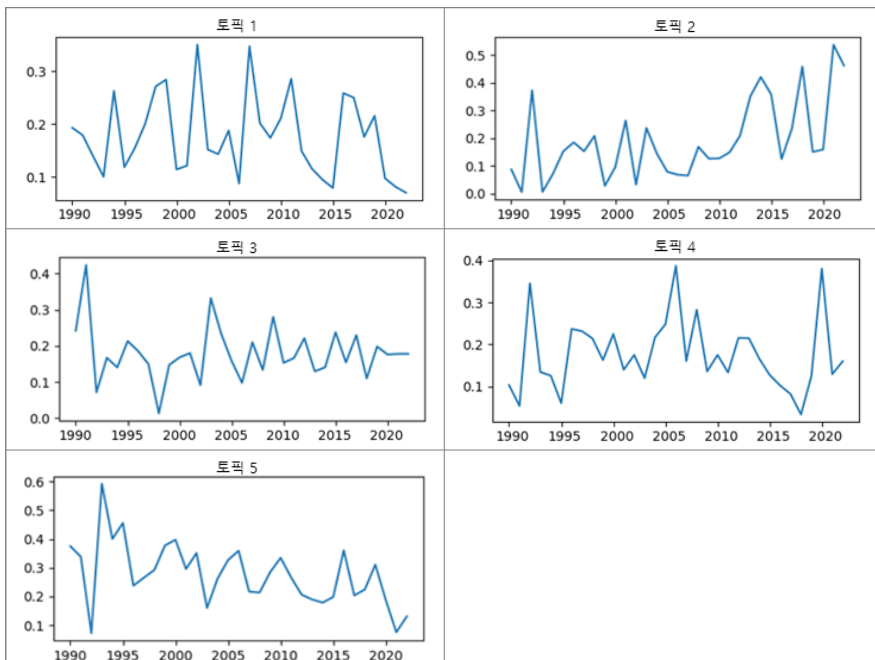
〈표 11〉 진보 언론 다섯 번째 토픽 키워드 변화

연도	키워드				
	1	2	3	4	5
1990년	장교	부대	훈련	공군	육군
1991년				공군	공군
1992년			육군		
1993년			여성	여성	
1994년		훈련			부대
1995년			훈련	훈련	
1996년		부대			육군
1997년			여성	여성	
1998년		육군			부대
1999년			공군	공군	
2000년		여성			부대
2001년			육군	육군	
2002년		육군			여성
2003년			부대	부대	
2004년		장교			해군
2005년	육군			공군	공군
2006년	부대	육군			
2007년		장교	장교		
2008년	해군			부대	
2009년		공군	공군		
2010년	해군			육군	장교
2011년		부대	부대		
2012년				해군	장교
2013년		장교	장교		
2014년	육군			부대	해군
2015년		장교	부대		
2016년			장교	장교	
2017년		장교			
2018년				부대	장교

연도	키워드				
	1	2	3	4	5
2019년		해군	부대	공군	장교
2020년			공군	부대	
2021년				장교	부대
2022년					

〈표 11〉은 동적 토픽 모델링을 수행하여 진보 언론 다섯 번째 토픽의 키워드 변화를 제시한 것이다. 여군 복무와 관련한 다섯 번째 토픽은 시간에 따른 큰 변화를 파악하기가 힘들어 보인다.

[그림 4]는 시간에 따른 진보 언론 여군 관련 뉴스 각 토픽의 비중 변화를 보여준다. ‘성폭력 사건’으로 명명한 두 번째 토픽을 제외한 다른 모든 토픽이 큰 변화가 없거나 대체로 토픽 비중이 감소하는 경향을 보이고 있다. 이는 보수 언론 여군 관련 뉴스의 토픽 비중 변화에서도 동일하게 관찰된 경향이다. 특히 보수 언론 여군 관련 뉴스에서도 그랬던 것처럼 ‘성폭력 사건’ 토픽의 비중이 상당히 높아서 토픽 비중만 기준으로 했을 경우 두 번째 토픽과 나머지로 나누어도 될 정도이다.



[그림 4] 진보 언론 여군 관련 뉴스 각 토픽 비중의 변화

3. 보수와 진보 언론의 여군 관련 뉴스 비교

보수와 진보 언론의 여군 관련 뉴스를 분석한 결과 몇 가지 차이가 도출되었다. 예상했던 것과 같이 보수 언론사는 군 조직 내 여군 활용에 더 초점을 맞춘 경향이 나타났는데, 특히 2000년대 중반에는 여군 장교의 진출을 집중 조명하였다. 여군 지원 및 복무와 관련한 토픽에서도 최근으로 오면서 병역 자원으로서의 여군에 초점을 맞추고 있다. 대체로 군 당국의 입장이 반영된 토픽의 뉴스들이 보도되고 있는 것이다.

이와는 대조적으로 진보 언론은 군 당국의 입장보다는 여군 개인의 입장을 강조한 토픽의 뉴스가 보도된 것으로 보인다. 특히 군 조직 구성원으로서의 여군이 아니라 좀 더 사적인 측면의 여군에 초점을 맞춘 토픽이 도출되었을 뿐만 아니라, 군 조직 내에서 여군이 겪는 피해와 관련된 내용들이 더 두드러지게 보도되고 있음을 알 수 있다.

그런데 여군 관련 뉴스의 보수와 진보 차이는 주로 과거에 더 두드러졌다. 즉, 언론의 정파성에 따른 효과가 과거에 강하게 나타났고, 최근으로 오면서 그러한 효과가 감소하고 있는 것이다. 동적 토픽 모델링을 통해서 살펴 보았듯이, 최근으로 오면서 ‘성폭력 사건’을 제외한 다른 토픽들은 그 비중이 큰 변화가 없거나 감소하는 경향을 뚜렷이 보였다. 이는 보수와 진보 언론사 모두에서 동일하게 나타났으며, 결국 토픽의 비중만 기준으로 하면 ‘성폭력 사건’ 토픽과 나머지 토픽으로 분류해도 무방할 정도가 되었다. 물론 ‘성폭력 토픽’에서도 보수와 진보 언론사의 접근이 다소 차이가 나는 것은 사실이다. 상대적으로 보수 언론사는 피고인에, 진보 언론사는 피해자에 초점을 맞추었다. 하지만 이러한 차이도 주로 과거에 두드러졌고, 최근으로 오면서 그 차이가 축소되었다. 즉, 토픽 모델링을 통해서 살펴본 바에 따르면, 보수와 진보 언론의 여군 관련 보도는 시간이 지나면서 차이가 감소하였고, 성폭력 사건에 대한 집중 조명으로 수렴하게 되었다.

V. 논의

본 연구에서는 여군 관련 뉴스의 특징을 포착해내기 위해 동아일보, 중앙일보, 한겨레, 경향신문의 여군 기사를 분석하였다. 빅카인즈에서 1990년부터 2022년까지의 자료를 수집하였고, 파이썬을 통해 LDA 토픽 모델링과 동적

토픽 모델링 방법을 활용하였다. 그 결과 보수와 진보 언론 간 여군 관련 보도의 차이점을 도출하였고, 시간에 따라 여군 관련 뉴스의 토픽이 변화하는 것을 포착하였다.

연구 결과를 바탕으로 논의할 점을 제시하면 다음과 같다. 먼저, 현 시점에서 보수와 진보 언론의 여군 관련 뉴스가 성폭력 사건으로 수렴되는 것은 한편으로는 당연해 보인다. 그만큼 군 조직 내 성폭력 문제가 심각하고 많은 국민들이 이에 충격을 받으며 분노하고 있다는 것을 시사한다. 이러한 이유는 여러 가지가 있겠지만, 성폭력 문제는 ‘범죄’라는 사실에 새삼 주목할 필요가 있다. 언론의 정파성이 개입할 여지가 있는 이슈들은 대체로 가치 판단이 필요한 영역이다. 그런데 성폭력 문제는 이러한 정파성이 끼어들 여지가 없는 질 나쁜 범죄일 뿐이다. 군 조직의 존재 이유가 국가와 국민의 안전 보장에 있다는 점을 고려하면, 전우애로 똘똘 뭉쳐야 할 조직 구성원 간 성폭력 범죄는 더 무겁게 다가온다. 더군다나 특수한 목적을 위해서 필요한 강력한 위계서열이 영동하게 작동하고 있다는 데에 많은 사람들이 분노할 수밖에 없다.

현재와 같은 수준으로 군 조직 내 성폭력 사건이 발생한다면, 그동안 여군의 규모를 늘리고 그들의 진출을 독려하며 제도를 정비했던 것이 큰 의미를 갖기 어려워 보인다. 양적인 확대, 상징적인 진출, 가시적인 제도의 개선에만 매몰될 것이 아니라 실제 복무 환경 및 동료들의 인식을 제고하려는 노력이 그 어느 때보다 필요해 보인다.

이를 개선하려는 군 조직의 자체적인 노력에도 상황이 그렇게 호전되지 못한 점과 위계 서열 및 폐쇄성 등 군 조직의 독특한 특징이 문제를 일으키는 주요 원인으로 지목되는 점을 고려한다면, 군 조직과 민간 사회와의 접점을 늘리고 군 조직의 사회 적합성을 향상시키는 방법이 유용할 수도 있을 것 같다.

다음으로 방법론적인 측면에서 토픽 모델링의 토픽 개수를 결정하는 문제에 대해 생각해 볼 필요가 있다. 토픽 모델링은 인간의 주관성을 배제하고 객관적인 방법으로 대량의 텍스트 자료를 파악하기 위한 방법으로, 그 결과는 토픽 개수를 어떻게 결정하느냐에 따라 다소 차이를 보인다. 따라서 그동안 토픽 개수를 결정하는 다양한 방법들이 제시되어 왔다. 그런데 본 연구에서는 비교적 일반적으로 사용되는 응집도와 혼란도 모두를 만족시키는 토픽 개수를 도출할 수 없었고, 이에 따라 응집도를 기준으로 하여 토픽 모델링을 수행하고 분석을 진행하였다. 이러한 문제가 비단 필자만의 문제는 아닐 것이다. 따라서 텍스트 자료의 성격과 연구 목적에 따라 토픽 문제의 결정에 있어 연구자의 세심한 관심이 필요할 것으로 판단된다.

마지막으로 토픽 모델링의 한계에 대해서도 고려할 필요가 있다. 앞서 언급

했듯이 토픽 모델링은 대량의 문서를 효율적으로 이해할 수 있다는 큰 장점을 가지고 있다. 하지만 보다 섬세한 접근과 이해가 필요한 경우에는 토픽 모델링의 방법이 다소 부적절할 수도 있기 때문에, 토픽 모델링을 활용한 연구 결과의 해석에 주의를 기울일 필요가 있다. 보수 언론과 진보 언론의 여군에 대한 보도에서 ‘성폭력 사건’ 토픽이 중요하게 대두된 것은 맞지만, 본고의 분석을 통해서 그 구체적인 면면과 미묘한 차이를 들여다보기 쉽지 않다. 통상 보수 언론과 진보 언론의 성폭력 관련 보도 양태가 달랐다는 점을 고려한다면, 여군 보도에 있어 ‘성폭력 사건’ 토픽의 대두에도 불구하고 보수 언론과 진보 언론 간 미묘한 차이가 존재할 수 있다. 이 지점에 대해서는 추후 질적인 방법을 통한 보완이 요구된다.

본 연구는 보수와 진보 각각 2개 언론사를 선정하여 총 4개의 언론사에 대해 분석하였기 때문에 결과 해석 및 일반화에 다소 주의가 필요할 수 있다. 그럼에도 불구하고 여군 관련 뉴스를 토픽 모델링 방법을 통해서 분석하고, 특히 동적 토픽 모델링을 통해 시간에 따른 토픽의 변화를 드러냈다는 점에서 그 의의가 있다.

참고문헌

- 국방 모성보호제도 운영 훈령(국방부훈령 제1279호, 시행 2010.10.28.)
- 국방부(2011). *여군60년사*. 서울: 국방부.
- _____(2020). *2020 국방백서*. 서울: 국방부.
- _____(2023). *2022 국방백서*. 서울: 국방부.
- 국회뉴스ON(2019.10.10.). “[2019국감] 김중로의원 “장성급 여군 5명에 불과
...최고 계급 17년째 ★”. <https://naon.go.kr/content/html/2019/10/10/d3808e6d-208e-480e-ae4d-987e804af478.html>에서 2023.9.8.
인출.
- 김경순(2011). “국방환경의 변화와 여군의 역할 확대”. *민족연구*. 제45호. 113-137.
- 김경희·노기영(2011). “한국 신문사의 이념과 북한 보도방식에 대한 연구”. *한국
언론학보*. 제55권 제1호. 361-387.
- 김수미·이새롬·한민경(2023). “군 내 성희롱 피해 원인 연구”. *한국형사정책학회*.
제34권 제4호. 117-144.
- 김영옥(2009). “독자가 본 한국 언론의 정파성”. *미디어인사이트*. 제3호. 1-30.
- _____(2011). “한국 언론의 정파성과 사회적 소통의 위기”. *한국언론학회 심포지
움 및 세미나*. 제5호. 107-136.
- 김엘리(2012). “초남성 공간에서 여성의 군인되기 경험”. *한국여성학*. 제28권 제3
호. 145-180.
- 김재은·김지현(2021). “여군의 군생활 경험과 적응 과정: 포커스그룹 인터뷰를 중
심으로”. *한국심리학회지 : 상담 및 심리치료*. 제33권 제1호. 507-538.
- 류영숙(2002). *여성장교의 경험으로 본 한국군대의 젠더정치*. 연세대학교 사회학
과 석사학위논문.
- 민경자(2008). “여군의 창설과 발전”. *군사*. 제68호. 325-366.
- 박윤정·이세영·이은성·금희조(2023). “ccLDA와 의미네트워크를 기반으로 하는
뉴스 의제의 정파성 분석”. *한국콘텐츠학회논문지*. 제23권 제1호. 48-60.
- 손영준(2004). “미디어 이용이 보수 진보적 의견에 미치는 영향”. *한국언론학보*.
제48권 제2호. 240-266.
- 신동호(2020). “텍스트마이닝을 이용한 보수와 진보 언론의 ‘국가안보’ 프레임 분
석”. *군사논단*. 제104호. 141-166.
- 여성신문(2022.02.18.). “군 성폭력 4년간 10배 폭증...성추행 가장 많아”. <https://www.womennews.co.kr/news/articleView.html?idxno=220363>에서

2023.9.8. 인출.

- 이동임(2021). “군 성폭력범죄 대응의 문제점과 개선방안에 관한 연구”. *안전문화연구*. 제13호. 59-79.
- 이수범·송민호(2020). “디지털 성범죄 사건 보도에 대한 비판적 담론분석: n번방 사건을 중심으로”. *언론정보연구*. 제57권 제4호. 150-195.
- 이재경(2008). “한국의 저널리즘과 사회갈등: 갈등유발형 저널리즘을 극복하려면”. *커뮤니케이션 이론*. 제4권 제2호. 48-70.
- 정인정·이지수·조은영·장정운(2022). “워드 임베딩을 활용한 ‘여군’ 이미지 분석: 2014-2019년 기사 말뭉치를 중심으로”. *한국군사학논집*. 제78권 제1호. 323-348.
- 조선웅(2015). “육군 ‘군인 부부’ 여군의 일과 가정 양립에 관한 연구”. *한국군사학논집*. 제71권 제3호. 33-58.
- _____(2019). “여성의 군대 지원 동기에 관한 연구: 육군 여군 장교를 중심으로”. *여성연구*. 제100권 제1호. 97-124.
- 파우진(2017). *여군은 초콜릿을 좋아하지 않는다*. 서울 : 삼인.
- 홍지아(2017). “젠더화된 폭력에 대한 뉴스 보도”. *한국언론정보학보*. 제83호. 186-218.
- _____(2022). “젠더갈등은 어떻게 우리 사회의 주요 담론이 되었는가?: 보수언론의 젠더갈등 기사 분석을 중심으로”. *미디어, 젠더 & 문화*. 제37권 제2호. 99-155.
- Baron, D. P.(2006). Persistent media bias. *Journal of Public Economics*, 90(1-2), 1-36.
- Blei, D. M.(2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.
- Blei, D. M., & Lafferty, J. D.(2006). Dynamic topic models. *Proceedings of the 23rd international conference on Machine learning*.
- Blei, D. M., Ng, A. Y., & Jordan, M. I.(2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Carreiras, H.(2006). *Gender and the Military: Women in the Armed Forces of Western Democracies*. London and New York: Routledge.
- Centeno, M. A. & Enriquez, E.(2016). *War & Society*. Malden: Polity.
- Gentzkow, M. & Shapiro, J. M.(2006). Media bias and reputation. *Journal of Political Economy*, 114(2), 280-316.

- Goldstein, J. S.(2001). *War and Gender: How Gender Shapes the War System and Vice Versa*. New York: Cambridge University Press.
- Iyengar, S.(1994). *Is Anyone Responsible?: How Television Frames Political Issues*. Chicago: The University of Chicago Press.
- Janowitz, M.(2017[1961]). *The Professional Soldier : A Social and Political Portrait*. New York: Free Press.
- Kanter, R. M.(1993). *Men and Women of the Corporation*. New York: BasicBooks.
- King, A.(2015). The female combat soldier. *European Journal of International Relations*, 22(1), 1-22.
- Rapp, A., Beitelspacher, L. S., Grewal, D., & Hughes, D. E.(2013). Understanding social media effects across seller, retailer, and consumer interactions. *Journal of the Academy of Marketing Science*, 41(5), 547-566.
- Yoder, J. D., Adams, J., & Prince, H. T.(1983). The Price of a Token. *Journal of Political and Military Sociology*, 11(2), 325-337.

Abstract

A Study on Media Coverage of Female Soldiers in the Military Using Topic Modeling

Sunwoong Cho

This study analyzed news articles on female soldiers in South Korean newspapers Dong-A Ilbo, JoongAng Ilbo, Hankyoreh, and Kyunghyang Shinmun from 1990 to 2022 to understand the structure of media coverage of women in the military. The data was collected through BigKinds, and analytical methods such as topic modeling were applied using Python. As a result, conservative press and liberal media showed some differences in reporting on female soldiers, but such differences narrowed over time, and as of late, both media have emphasized on topics related to sexual assault. Based on these findings, several implications were drawn and presented.

Keywords : Female Soldiers, News, Topic Modeling, Dynamic Topic Modeling.

* PhD Candidate in Department of Sociology, Seoul National University.