

특집 논문

김일성·김정일·김정은 시대의
북한 고고학 비교*

: 『조선고고연구』 ‘멀리서 읽기’

고일홍 (서울대)

국문요약

『조선고고연구』는 사회과학원 고고학연구소에서 1986년부터 계간으로 발행해 온 북한 고고학의 대표 학술지로, 북한 고고학의 역사를 추적할 수 있는 가장 중요한 일차자료라고 할 수 있다. 본 논문에서는 이러한 『조선고고연구』에 수록된 글들을 분석하여 김일성·김정일·김정은 체제의 북한 고고학 활용 전략이 어떻게 변화하였는지를 추적하였다. 분석 대상 텍스트가 방대한 관계로, 이번 분석을 위해 ‘디지털 인문학’의 ‘멀리서 읽기’ 방법론을 도입하였다. 본 분석에서 사용된 ‘멀리서 읽기’ 기법은 총 3가지였다. 우선, 김일성·김정일·김정은 시대에 게재된 모든 논문의 제목으로 구성된 말뭉치에 대한 분석을 진행하였다. 각 시대의 주요어를 워드 클라우드로 시각화하였고, 토픽 모델링을 통해 대표 주제 5개 및 연관 주요어를 제시하였다. 그 결과 김일성·김정일 시대를 거치면서 ‘발전’의 중요도가 낮아지고, ‘평양’의 중요도가 높아졌음이 확인되었는데, 이는 김정일 시대 ‘조선민족제일주의’의 영향을 받은 결과라고 판단하였다. 또한 김정은 시대에 들어와서는 ‘왕릉’, ‘개성’, ‘문화유산’의

DOI URL: <https://doi.org/10.35369/jpus.15.1.202306.55>

* 이 논문은 2020년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2020S1A6A3A02065553). 또한 이 논문은 2022년도 서울대학교 통일평화 연구원의 재원으로 통일·평화기반구축사업의 지원을 받아 수행된 결과물임.

중요도가 높아지고 '만월대'가 새롭게 주요어로 등장하였음을 확인하였는데, 이는 '문화유산보호법'의 채택 및 '사회주의기업 책임관리제'의 영향을 받은 결과라고 판단하였다. 한편, 『조선고고연구』 수록된 '교시적' 글의 제목과 가장 먼저 등장하는 교시의 내용에 대해 주요어 분석과 시맨틱 네트워크 분석을 실시하였다. 그 결과, 김정일·김정은 시대의 교시 중 가장 중요한 주요어가 '역사'에서 '민족'으로 전환되었고, 또한 김정은 시대 교시의 시맨틱 네트워크에서 '문화유산'의 중심성이 '민족' 다음으로 높게 나왔음을 확인할 수 있었다. 이는 김정은 시대 '민족' 담론의 변화, 그리고 북한 정권이 고고학 성과를 바탕으로 추구하고자 하는 목적의 변화를 반영한 결과라고 판단하였다.

주제어: 북한 고고학, 『조선고고연구』, 김정은 시대, 멀리서 읽기, 디지털 인문학

I. 들어가며

고고학과 민족주의(nationalism)는 19세기 후반부터는 불가분의 관계를 맺어 왔다. 둘의 관계는 매우 복잡하고 다양하며, 긍정과 부정이 공존하기 때문에 기계적으로 비판만 해서는 안 된다는 것이 고고학계의 입장이다.¹⁾ 예를 들어, 20세기 전반에는 민족주의의 영향으로 사회진화론적 보편사에 매몰되어 있던 고고학자들이 지역사에 눈을 돌리게 되었고, 식민지화를 경험했던 집단들에게 이러한 지역사의 발견은 - 모든 인류가 존엄하다는 생각과 결부된 경우에 한해서는 - '문명화'의 탈을 쓰고 자행되었던 식민주의적 역사 왜곡과 인종주의에 대한 저항의 기반을 마련해 주었다.²⁾ 그러나 고고학이 이처럼 민족주의로부터 자유로울 수 없음을 인

1) 예를 들어, Trigger, B.G. "Alternative Archaeologies" in Historical Perspective," in J. Habu, C. Fawcett, J.M. Matsunaga, eds., *Evaluating Multiple Narratives* (New York, NY.: Springer, 2008).

2) Kohl, P.L., Fawcett, C. "Archaeology in the service of the state: theoretical considerations" in P.L. Kohl, Fawcett, C., eds., *Nationalism, Politics and the Practice*

정할지라도, 그것의 과도한 영향은 반드시 경계해야 한다. 그것을 잘 보여주는 것이 1990년대 이후 북한 고고학계가 진행한 단군릉의 발굴과 대동강문화론의 제기이다.

주지하다시피 북한 고고학계는 이미 1960년대 이래 주체사상의 영향을 받고 있었다. 제4차 노동당대회(1961년)를 계기로 주체사상과 유일사상체계가 확립되고 지식인에 대한 학문적 통제가 강화되면서 주체사상이 고고학계에 유입되었다. 또한 제5차 노동당대회(1970)를 통해 당 규약에 주체사상이 추가되면서 1970년대부터는 김일성 어록이 고고학 논문에 직접 인용되기 시작했다.³⁾ 이 당시까지만 하더라도 고고학의 주된 목적은 고조선에 대한 교육강화 등 인민의 계몽이었고, 따라서 1960-70년대까지는 북한의 ‘주체사상 고고학’이 민족문화의 우수성과 계승성을 강조하는 ‘과도한 민족주의 고고학’ 정도였다고 평가할 수 있다.⁴⁾ 그러나 1989년에 발표된 김정일의 ‘조선민족제일주의’ 담화를 계기로 1990년대부터 고고학이 본격적인 체제 선전의 도구로 전환되면서 북한 고고학은 새로운 국면을 맞이하게 되었다. 특히 기원전 10세기로 비정되었던 고조선의 기원 연대가 단군릉의 발굴로 기원전 3000년기 초까지로 올라가면서, 그동안 유적·유물에 대한 실증적 연구의 결과로 축적되었던 기존 편년안이 한순간에 무너졌다. 그리고 그 이후 진행된 고고학 자료의 연대에 대한 해석은 정치색이 짙은 새로운 선사시대 편년안을 재생하는 도구로 전락하였다. 이렇게 되면서 북한에서 출간된 고고학 텍스트의 학문적 의미는 많이 약해졌지만, 그 대신 김정일 시대, 김정은 시대의 정치수사를 반영하는 자료로써 북한학 연구 분야에서의 중요성은 커졌다고

of Archaeology (Cambridge: Cambridge University Press, 1995), pp. 4-5.

3) 이기성, 「북한학계 선사고고학의 최근 연구 동향」 『통일고고학을 위한 연구현황과 과제 진단』 한국고고학회 2016년 학술회의 (2016년 8월 23일), 17~19쪽.

4) 위의 글, 25쪽.

볼 수 있다.

이에 본고에서는 1986년부터 2019년까지 매해 4호씩 발간된 북한 고고학의 대표 학술지 『조선고고연구』의 텍스트를 분석하여 김일성·김정일·김정은 체제의 북한 고고학 활용 전략이 어떻게 변화하였는지는 추적하고자 한다. 현재까지 북한의 고고학사를 통시적으로 조망하거나 『조선고고연구』에 대한 분석을 통해 북한 고고학의 연구 경향을 추적하는 작업은 간헐적으로 이루어진 바 있다.⁵⁾ 그러나 본고와 같이 고고학이 아닌 정치 담론의 관점에서 북한 고고학의 성과를 분석한 시도는 거의 없었던 것으로 판단된다.

한편, 본고에서는 ‘멀리서 읽기(distant reading)’ 방법을 도입하여 『조선고고연구』의 텍스트에 대한 분석을 진행하고자 한다. ‘멀리서 읽기’란 방대한 분량의 텍스트를 꼼꼼하게 살펴보고 분석하기보다는 디지털 기술을 활용하여 주요어의 통시적 변화양상을 추적하거나 주요어 간의 의미 관계를 확인하는 등 텍스트의 ‘구조’를 파악하는 방법론을 의미한다. 이 방법론은 문학비평에서 처음으로 시도된 이후 다양한 학문분야에서 연구사를 추적하는 데 활발히 사용되고 있다.⁶⁾ 이에 본고에서는 『조선고

⁵⁾ 남한 고고학자들에 의한 북한 고고학사 연구는 1990년을 전후에 시작되었는데, 2015년까지의 연구는 이기성의 2015년 논문에 잘 정리되어 있다(이기성, 「북한 선사고고학의 패러다임- 학사적 검토를 중심으로-」 『고고학』 제14권 3호(2015), 9~10쪽, 27~28쪽). 또한 2015년부터 2017년까지 출간되거나 이기성의 2015년 논문에서 누락된 북한 고고학사 선행연구는 고일홍의 2018년 논문에 정리되어 있다(고일홍, 「北韓 考古學界 最近 現況 檢討 - 『조선고고연구』에 대한 內容分析을 中心으로 -」 『湖西考古學』(2018), 103쪽). 2018년 이후에 출간된 북한 고고학사 관련 연구는 다음과 같다: 이형우, 「정치·사회적 맥락에 따른 북한 구석기 연구 변화」 『문화재』 제53권 1호(2020); 강현숙, 「북한의 고구려 고고학 조사연구의 성과와 과제」 『문화재』 제53권 1호(2020); 윤상덕, 「북한의 신라 고고학 연구 현황과 특징」 『문화재』 제53권 2호(2020); 이기성, 「북한 청동기시대 고고학 연구 경향의 변화와 쟁점」 『문화재』 제53권 3호(2020).

⁶⁾ Underwood, T. "Distant Reading and Recent Intellectual History," in Gold, M.K. &

고연구』에 수록된 총 1657편의 글 제목을 대상으로 주요어 및 주요 주제의 통시적 변화양상을 3개의 분기로 나누어(김일성 시대, 김정일 시대, 김정은 시대) 추적하겠다. 또한 『조선고고연구』에 수록된 ‘교시적’ 성격의 글 35편에 포함된 교시⁷⁾의 텍스트를 분석하고 워드 클라우드나 시맨틱 네트워크 등의 형태로 그 경향성을 시각화하겠다. 이처럼 북한 고고학의 텍스트와 디지털 인문학의 방법론을 접목함으로써 이 분야에서 생성된 연구의 결과가 폭넓은 연구자층에 의해 활용될 수 있도록 하는 것이 본고의 궁극적인 목적이다.

II. 『조선고고연구』와 ‘멀리서 읽기’ 방법론의 중요성

1. 『조선고고연구』의 가치

그동안 북한에서는 고고학 분야의 연구 성과물이 다양한 학술지들을 통해 소개되었으나,⁸⁾ 이들 중 창간 이래 명칭 변경 없이 지속적으로 정기 출간되고 있는 학술지는 『조선고고연구』가 유일하다.⁹⁾ 북한 사회과

Klein, L.F. eds., *Debates in the Digital Humanities 2016* (Minneapolis, MN: University of Minnesota Press, 2016).

- 7) 본고에서는 ‘교시’를 북한의 최고 지도자가 내리는 지침 등을 의미하는 포괄적인 개념으로 이해하고 사용하고 있다. 즉, 김정일의 ‘말씀’ 등도 ‘교시’의 범주에 넣고 있다.
- 8) 『문화유물』, 『문화유산』, 『고고민속』, 『민족문화유산』, 『歷史諸問題』, 『역사과학』, 『조선건축』 등.
- 9) 국립문화재연구소, 『북한정기간행물- 고고학문헌목록』 (2016), 2쪽.

학원 고고학연구소에서 1986년부터 계간으로 발행해온 『조선고고연구』에는 다양한 종류의 글(논문, 발굴 및 조사보고, 강좌, 상식, 학계 소식 등)이 수록되어 있어서, 북한 고고학의 현주소에 대한 균형 잡힌 이해를 제공하고 있다.¹⁰⁾

『조선고고연구』가 이렇듯 북한 고고학계를 대변하는 중요 학술지이다 보니, 남한의 국립문화재연구소에서는 1986년부터 2016년까지 『조선고고연구』에 게재된 모든 글에 대한 서너 문단 분량의 요약과 한 문단 분량의 해제를 수록한 해제집 두 권을 출간하기도 하였다.¹¹⁾ 1986년부터 2019년까지, 33년이라는 기간 동안 수록된 총 1657편의 글들은 북한 고고학의 역사를 추적할 수 있는 가장 중요한 일차자료라고 할 수 있다. 따라서 『조선고고연구』는 최근 북한 고고학의 연구 경향 및 그 통시적 변화 양상을 파악하는 데 적극적으로 활용되기도 했다.¹²⁾

그러나 『조선고고연구』의 가치는 고고학 분야에만 한정되어 있는 것은 아니다. 이 학술지에는 사회과학원 고고학연구소뿐만 아니라, 김일성 종합대학, 중앙력사박물관, 조선민속박물관, 문화보존연구소, 건설건재대학, 묘향산력사박물관, 해주력사박물관, 개성력사박물관, 평안남도 력사박물관, 고려박물관, 함경북도 명승지 및 문화유적관리소 등에 소속된 연구자 330여명의 글이 수록되어 있는 만큼, 북한의 연구 생태계를¹³⁾ 파

10) 고일홍, 「北韓 考古學界 最近 現況 檢討 - 『조선고고연구』에 대한 內容分析을 中心으로 -,」, 105쪽.

11) 국립문화재연구소, 『조선고고연구 해제집 1』(2017a); 국립문화재연구소, 『조선고고연구 해제집 2』(2017b).

12) 예를 들어, 고일홍, 「북한의 발해 연구사 서술을 위한 새로운 접근 - 북한 학술지 목록화 사업 결과물을 활용한 내용분석」 『인문논총』 제77권 4호 (2020); 양시은, 「최근 북한 고고학계의 고구려 연구 동향: 『조선고고연구』를 중심으로」 『고구려발해연구』 제56권 (2016).

13) ‘연구 생태계’란 자연 생태계의 유지 및 발전의 원리를 사회과학적 해석에 적용한 개념이다. 이 개념을 도입하면 생태계 내부에서 연구가 생산되고 활용되는

약하는 데에도 매우 중요한 자료이기도 하다.

또한 『조선고고연구』에는 그 제목에서 ‘김일성’, ‘김정일’, ‘김정은’, ‘당’, ‘령도’ 등의 키워드가 등장하는 소위 ‘교시적’인 성격의 글이 총 35편 수록되어 있다. 이러한 교시적 글의 연도별 수록 양상이나 필진에 대한 연구는 이미 진행된 바 있는데,¹⁴⁾ 전체 글 편수의 2% 정도에 불과한 이러한 교시적 글의 존재만 가지고 『조선고고연구』의 학문적 의의를 무비판적으로 폄하하는 것은 옳지 않다는 의견이 제시되었다.

『조선고고연구』에 수록된 교시적 성격의 글은 고고학자에게는 당혹스럽게 다가오겠지만, 북한학 연구에서는 귀중한 자료로 활용될 수 있다. 왜냐하면 이 35편의 글은 김일성·김정일·김정은 시대에 모두 걸쳐 분포하고 있어서, 시대별 교시 내용의 특징과 그 변화 양상을 확인할 수 있기 때문이다. 이는 특히 북한의 정치 수사, 담론 등을 다루는 연구에 보탬이 될 수 있으며, 바로 이러한 이유로 『조선고고연구』는 고고학뿐만 아니라 북한학 분야에서도 관심을 가질 만한 자료라고 평가할 수 있다.

2. ‘멀리서 읽기’ 방법론의 가능성

『조선고고연구』가 가치가 있는 자료임은 분명하나, 굳이 이 학술지에 수록된 1657편의 글을 세밀하게 분석할 만큼의 가치가 있는 자료인지에 대해서는 회의가 든다. 우선, 앞서 언급했듯이, 2016년까지 출간된 글들

과정, 유효기간이 지난 연구가 소멸하고 또 다른 연구로 대체되는 과정 등을 살펴보고, 연구자가 소속된 연구 집단이나 집단과 개인 연구자 간의 관계망에 주목하게 된다. 박초희·이성숙, 「연구 생태계 관점에서 본 국내 행정학 분야의 구조적 특성 분석 - 공저자, 인용, 키워드 네트워크 중심으로 -」 『한국비블리아학회지』 제31권 1호 (2020), 214쪽.

14) 고일홍, 「北韓 考古學界 最近 現況 檢討 - 『조선고고연구』에 대한 內容分析을 中心으로 -」, 120~123쪽.

에 대해서는 이미 요약과 해제 작업이 진행된 바 있다. 이 해제집은 북한 고고학 전반의 본질을 예리하게 파악하고 날카롭게 비판하는 내용을 담고 있지는 않다. 『조선고고연구』에 대한 그러한 심도 있는 분석은 일부 세부전공분야에 한해서만 이루어지고 있을 뿐이며,¹⁵⁾ 『조선고고연구』에 수록된 글 전체에 대해 단독 연구자나 하나의 연구집단이 심도 있는 분석을 진행하기를 기대하는 것은 다소 무리라고 판단된다. 그렇다면 『조선고고연구』 자료 전체를 조망하면서 『조선고고연구 해제집』에서 제공되는 것 이상의 시사점을 얻을 수 있는 현실적인 방법은 없는가? 필자는 그 방법을 ‘멀리서 읽기’에서 찾았다.

‘멀리서 읽기’는 모레티가¹⁶⁾ 2000년 제창한 개념으로, 몇 개의 특정한 텍스트만을 읽지 않고, 거대한 양의 데이터를 축적하고 분석함으로써 문학을 이해하고자 하는 방법을 의미한다. 문학비평가인 모레티는 지금까지 나온 모든 문학사(文學史) 연구가 과거에 출간되었던 모든 텍스트가 아닌 몇몇 정전(正典) 텍스트를 중심으로 구축된 현실을 비판하며 ‘멀리서 읽기’라는 혁신적인 방법론의 필요성을 제안했던 것이다.¹⁷⁾ 이 새로운 방법론은 ‘디지털 인문학’으로의 전환을 주도하였으며, 이후 다양한

15) 예를 들어, 양시은, 「최근 북한 고고학계의 고구려 연구 동향: 『조선고고연구』를 중심으로」.

16) Moretti, F., “Conjectures on World Literature,” *New Left Review*, no. 1. (2000), <<https://newleftreview.org/issues/ii1/articles/franco-moretti-conjectures-on-world-literature>>

17) 모레티는 기존과는 다른 문학사 연구를 주장하기 위해 ‘멀리서 읽기’ 방법론을 제창하였다. 예를 들어, 1700년부터 2000년 사이에 영국, 일본, 스페인, 이탈리아, 나이지리아에서 출간된 모든 소설의 5년 단위의 출간 빈도에 대한 그래프를 생성하여 세계 문학사를 광범위한 시각에서 조망하였다. 또한 1710년부터 1850년 사이에 출간된 모든 영국 소설의 5년 단위의 출간 빈도에 대한 그래프를 생성하여 영국 소설의 성장 과정을 추적하였다. Moretti, F., 2005., *Graphs, Maps, Trees: Abstract Models for a Literary History* (London: Verso, 2005), pp. 5-8.

연구 분야에서 활용되고 있다.

일각에서는 ‘멀리서 읽기’가 더 객관적인 접근이라고 보기도 하지만, 이러한 입장은 많은 비판을 받기도 하였다.¹⁸⁾ 필자 역시 ‘멀리서 읽기’에서 연구자의 주관이 배제되어 있다고 생각하지는 않는다. 아래에서 다시 언급하겠지만, ‘멀리서 읽기’는 ‘텍스트 마이닝’에 기반하고 있는데, 텍스트 마이닝의 선행 작업인 ‘정규화’와 ‘토큰화’ 작업에는 연구자의 판단이 개입한다. 또한 데이터 클리닝 작업을 거친 텍스트를 대상으로 자연어 처리를 진행하고 ‘워드 클라우드’를 이용해서 주요어를 시각화하는 과정에서도 연구자는 표출되는 주요어의 개수를 지정해야 한다. 텍스트를 구성하는 단어 간의 관계를 나타낸 ‘시맨틱 네트워크’를 생성할 때도 분석 대상 범위를 지정해야 하고, 유의미한 단어 간 링크의 출현 빈도를 나타내는 값도 정해야 한다. 이처럼 ‘멀리서 읽기’도 연구자의 판단이 분석의 최종 결과물에 큰 영향을 끼치는 방법론임은 분명하다. 하지만 만약 데이터 세트와 연구자가 분석 과정에 대한 정보가 공개된다면 다른 연구자는 해당 분석을 재현할 수 있고, 또한 변수의 조정을 통해 새로운 해석을 도모할 수 있다.

‘멀리서 읽기’ 방법이 지원해주는 이와 같은 ‘재현성’은 북한 자료에 대한 남한 연구자의 분석이 그나마 객관성을 담보하고 있다는 평가를 받을 수 있게 하는 근거를 제공해준다. 일례로 필자는 앞서 ‘멀리서 읽기’라고 명시하지는 않았지만, 이 방법을 적용하여 기존의 남한 연구자들이 제시했던 것과는 다른 시각에서 북한의 발해 연구사를 서술한 바 있다.¹⁹⁾ 그 이유는 남북 연구자 모두가 검토하고 의견을 나눌 수 있는 ‘데이터 기반’

18) 예를 들어, Ascari, M. “The Dangers of Distant Reading: Reassessing Moretti’s Approach to Literary Genres,” *Genre*, vol. 47, no. 1 (2014).

19) 고일홍, 「북한의 발해 연구사 서술을 위한 새로운 접근 - 북한 학술지 목록화 사업 결과물을 활용한 내용분석」.

연구사를 제시하고자 했기 때문이었다.

따라서 본고에서는 이렇듯 다양한 학문 분야에서 그 유용성이 충분히 입증된 바 있는 ‘멀리서 읽기’ 방법론을 활용하여 『조선고고연구』의 특징을 통시적으로 살펴보고자 한다. 구체적으로는 ‘워드 클라우드’를 생성하고 ‘토픽 모델링’을 진행할 예정이다. 이와 같은 방식의 연구는 흔히 논문의 제목과 초록으로 구성된 텍스트를 대상으로 진행된다. 그런데 『조선고고연구』에는 (북한의 학술지가 흔히 그러하듯이) 별도의 초록이 제시되어 있지 않다. 따라서 여기에서는 이 학술지에 수록된 글들의 ‘제목’만을 가지고 분석을 진행할 예정이다. 참고로 글의 제목만으로도 충분히 유의미한 분석을 진행할 수 있다는 사실은 모레티가 1740년부터 1850년까지 110년에 해당되는 기간 동안 출판된 7000권의 소설 제목에 대하여 ‘멀리서 읽기’를 함으로써 새로운 영국 소설사를 제시했던 사례를 통해서 알 수 있다.²⁰⁾ 또한 논문 초록이나 전문이 아니라 제목 데이터만을 가지고 시맨틱 분석이나 지식 그래프를 생성하는 것이 충분히 유의미한 결과를 제공할 수 있음을 입증하는 연구 결과도 발표된 바 있다.²¹⁾

물론 『조선고고연구』에 수록된 텍스트의 전문을 분석한다면 더욱 심도 있는 시사점을 확보할 수 있겠지만, 디지털화된 텍스트가 존재하지 않아 현재로서는 불가능하다. 다만 이 글에서는 35편의 교시적인 글에서 직접 인용된 김일성, 김정일, 김정은의 교시를 분석하는 것은 중요하다고 판단되어, 그 텍스트에 한해서는 필자가 이를 디지털화하여 시맨틱 네트워크 분석 등을 실시하였다.

²⁰⁾ Moretti, F. “Style, Inc. Reflections on Seven Thousand Titles (British Novels, 1740–1850),” *Critical Inquiry*, vol. 36, no. 1 (2009).

²¹⁾ Galke, L., Mai, F., Schelten, A., Brunsch, D., Scherp, A. “Using Titles vs. Full-text as Source for Automated Semantic Document Annotation,” *K-CAP 2017*, December 4th–6th, 2017, Austin, Texas, United States (2017).

III. 『조선고고연구』 텍스트의 분석

1. 『조선고고연구』 제목 데이터 세트의 특징

『조선고고연구』에는 ‘논문’ 이외에도 북한 고고학계의 관심사에 대한 파악을 가능하게 하는 다양한 종류의 글들이 수록되어 있다²²⁾. 각 종류의 글에서 중요하게 다루고 있는 주제나 글의 종류별로 확인되는 통시적 양상을 파악하는 것도 흥미로운 연구 주제가 될 수 있겠으나, 이는 본고에서 주요하게 다룰 내용이 아니라고 판단된다. 따라서 『조선고고연구』에 대한 분석을 위해 구축한 데이터 세트에서는 이러한 구분을 반영하지 않았다.

오히려 본고의 목적을 위해서는 『조선고고연구』의 글들을 시기별로 구분하고, 주제나 표현의 통시적 변화양상을 파악하는 것이 더 중요하다고 판단되었다. 북한 고고학의 분기설정에 대해서는 여러 선행연구에서 이미 다루어진 바 있다.²³⁾ 그러나 이러한 기존의 분기설정안들은 고고학사의 관점에서 이루어진 것이고, 또한 『조선고고연구』가 출간되기 시작한 1986년 이후의 시기의 대해서는 세부적으로 구분하지 않고 있다. 따라서 본 분석을 위해서는 북한 정치 수사학 연구(특히, 고고학 분야와 관련이 있는 ‘민족’ 등에 대한 인식 연구)에서 일반적으로 사용되고 있는 분기설정 단위인 ‘김정일 시대’와 ‘김정은 시대’를 사용하였다.²⁴⁾

²²⁾ 논문 1062편; 발굴보고, 발굴소식, 발굴 및 답사소식, 발굴 및 조사보고, 조사 및 발굴보고, 조사발굴기, 조사발굴보고, 소개자료, 유적소개, 유물소개, 유적소개, 유적발굴기 252편; 자료, 자료 및 소개, 자료소개, 자료와 소개, 자료와 연구, 소개자료 99편; 강좌 137편; 학계소식, 소식 43편, 상식 28편; 교시적 글 35편; 강령 1편.

²³⁾ 이기성, 「북한 청동기시대 고고학 연구 경향의 변화와 쟁점」, 185-187쪽 참고.

본고에서 ‘김정일 시대’는 1990년부터 2011년까지 해당된다. 1990년은 비록 김일성이 생존해 있었으나, 김정일이 1990년 5월 최고인민회의 제 9기 1차회의에서 국방위원회 제1부위원장으로 선출되어 당뿐만 아니라 군에서도 공식적으로 2인자의 지위를 확보하였기 때문에 ‘김정일 공식승계 단계’의 시작으로 대체로 이해되고 있다.²⁵⁾ 또한 향후 북한 고고학의 방향에 지대한 영향을 끼친 김정일의 ‘조선민족제일주의’ 담화가 발표된 것이 1989년 12월이었으므로 1990년을 ‘김정일 시대’의 시작연대로 삼고, 『조선고고연구』가 출간되기 시작한 1986년부터 1989년까지를 본고에서는 ‘김일성 시대’로 규정하였다. 한편, 김정은 시대의 시작을 2012년으로 설정한 이유는, 그가 후계자로 내정되고 우상화 과정이 체계적으로 진행되기 시작한 것은 2009년이었으나, 『조선고고연구』에 수록된 교시적인 글에서 김정은의 교시가 인용되기 시작한 것은 2012년부터이기 때문이다. 이상과 같은 시기 구분에 근거하여 분류 및 분석된 『조선고고연구』 수록 글은 ‘김일성 시대(1986-1989)’ 총 152편, ‘김정일 시대(1990-2011)’ 총 993편, ‘김정은 시대(2012-2019)’ 총 510편이다.

한편, 『조선고고연구』에는 교시적 내용을 담고 있는 글이 총 35편 수록되어 있는데, ‘김일성 시대’에 1편, ‘김정일 시대’에 22편, ‘김정은 시대’에 12편이 해당된다. 참고로 『조선고고연구』의 모든 글에는 김일성·김정일·김정은의 교시가 포함되어 있으나, 대체로 3번째 문단부터는 고고학 관련 내용이 나오고 주체사상 등과 관련된 수사적인 표현들은 거의 등장하지 않는다. 반면 글 제목에서부터 아예 ‘김일성’, ‘김정일’, ‘김정은’, ‘당’, ‘령도’와 같은 표현이 등장하고 항상 목차의 맨 앞에 배치되어 있는

24) 예를 들어, 임사라·양문수, 「김정일 시대와 김정은 시대의 경제개혁 조치 비교 연구」 『현대북한연구』 제25권 1호 (2022).

25) 예를 들어, 최진옥, 『김정일의 당권장악과정 연구』 (서울: 民族統一研究院, 1996), 4쪽.

일부 글의 경우에는 텍스트 전반에 걸쳐 수사적 표현이 지속적으로 등장한다.

따라서 교시적 내용을 담고 있는 이 글들에 한해서는 글의 초반에 등장하는 교시 문구에 대한 분석을 진행하겠다. 물론 논문 전체에 대한 텍스트 분석을 진행할 수도 있겠지만, 아직은 『조선고고연구』 텍스트 전문에 대한 디지털화가 진행되어 있지 않다. 해당 자료의 디지털화는 개별 연구자가 단독으로 할 수 있는 성격의 작업이 아닌 만큼, 본고에서는 교시적 문구에 대해서만 텍스트 분석을 하는 ‘파일럿 스터디’ 성격의 분석에 치중하도록 하겠다.

2. 『조선고고연구』 제목의 주요어와 그 변화 양상

텍스트 분석에 앞서 진행해야 하는 전처리 작업은 크게 두 가지인데, 하나는 ‘정규화(normalization)’ 작업이고, 또 하나는 ‘토큰화(tokenization)’ 작업이다.²⁶⁾ 정규화 작업은 텍스트의 언어에 따라 그 과정이 다른데, 가령 영어 텍스트의 경우에는 대문자·소문자 구분을 없애고, 문장에서 가장 많이 사용되는 ‘the’, ‘and’, ‘it’과 같은 단어들을 제거해야 한다. 한국어의 경우에는 조사나 문장부호 등을 제거하는 기본적인 클리닝 작업 이후, 연구 목적에 따라 텍스트의 구조를 파악하는 데 방해되는 단어를 제거하는 추가적인 클리닝 작업이 이루어져야 한다. 본 분석의 대상 텍스트가 고고학 논문의 제목이다 보니 논문의 주제·필자·게재 시기와 상관없이 반복적으로 등장하는 단어들이 있는데, 시간과 관련 있는 ‘시대’, ‘시기’, 공간과 관련 있는 ‘지역’, ‘구역’, 그리고 논문 제목에서 흔히 사용되는 단어인 ‘검토’, ‘고찰’ 등이 이에 해당된다. 이와 같은 성격의 단어들

²⁶⁾ Graham, S., Milligan, I., Weingart, S. *Exploring Big Historical Data: The Historian's Macroscope* (London: Imperial College Press, 2016), pp. 60-67.

은 ‘기피어’로 설정하여 분석을 진행하였다.

‘정규화’ 작업의 또 다른 중요한 과정으로는 ‘표제어(lemma)’ 지정과 ‘어간 추출(stemming)’이 있다. ‘어간 추출’은 다양한 형태의 단어를 단일화하는 작업으로, 과거형·현재형·미래형이 다르게 표현되는 언어들에서는 중요한 작업이 되겠지만, 한국에서는 ‘표제어’ 지정으로 대체될 수 있는 작업이다. 표제어 지정은 연구 목적에 따라 동일한 범주로 간주될 수 있는 단어들을 묶고 그 대표 단어를 설정하여 단어의 빈도를 계산할 때 수치가 분산되는 것을 막아준다. 예를 들어, 본 분석에서는 ‘기와’, ‘막새’, ‘암기와’, ‘수기와’, ‘암기와막새’, ‘수기와막새’의 표제어로 ‘기와’를 설정하고, ‘고구려’, ‘고구려적 성격’, ‘고구려 유적 유물’, ‘고구려 사람’, ‘고구려사람’의 표제어로 ‘고구려’를 설정하였다(부록11 참고).

‘토큰화 작업’은 텍스트를 토큰(token)이라 불리는 단위로 나누는 작업인데, 한국어는 기본적으로 형태소 단위로 토큰화가 이루어진다. 그러나 종종 연구 목적에 따라 복수의 형태소 단위를 토큰으로 지정할 필요가 있다. 예를 들어, 본 분석에서는 ‘번개무늬그릇’ 등과 같은 북한 고고학의 개념어나, ‘락랑’과 같이 북한 표기법으로 인해 프로그램이 하나의 고유명사 형태로서 인식하지 못한 경우에 그것을 하나의 토큰으로 지정하였고(‘지정어’ 설정), ‘-동’, ‘-리’와 같은 지명의 경우에도 행정구역 단위가 포함되도록 지정어를 설정하였다.

이와 같은 전처리 과정을 진행한 텍스트를 대상으로 우선 주요어를 ‘워드 클라우드’의 형태로 시각화하였다.²⁷⁾ 본 분석에서는 용어의 상대적 중요도를 ‘단순 출현빈도’가 아닌 ‘TF-IDF 지수’²⁸⁾를 통해 측정하는 Net-

²⁷⁾ 본 분석에서는 NetMiner 4.0 소프트웨어를 이용해서 명사를 추출하여 워드 클라우드를 생성하였다.

²⁸⁾ TF-IDF는 단어의 빈도와 역 문서 빈도를 사용하여 해당 텍스트 내에서 단어의 중요성의 정도를 판단하는 데 사용되는 통계적 지표이다.

Miner4.0 소프트웨어의 틀을 이용해서 워드 클라우드를 생성하였다. 이때 워드 클라우드를 통해 표출될 주요어 개수의 선정은 두 가지 전략에 입각해서 이루어졌다. 우선 김일성 시대, 김정은 시대, 김정일 시대의 말뭉치(corpus)를 비교하기 위해서는 동일한 개수의 주요어가 표출되는 워드 클라우드를 생성할 필요가 있었다. 그런데 김일성 시대 말뭉치의 경우에는 151개의 제목으로만 이루어져 있어서, 유의미한 패턴을 확인하기 위해서는 소량의 주요어를 표출하였다.²⁹⁾ 이렇게 생성된 김일성, 김정은, 김정일 시대 말뭉치의 20개 주요어가 각각 표출된 워드 클라우드는 <그림 1>과 같다.

<그림 1> 김일성, 김정일, 김정은 시대 말뭉치(corpus)의 20개 주요어가 각각 표출된 워드 클라우드



²⁹⁾ 분석에 사용되는 다양한 워드 클라우드 생성 프로그램들을 보면 일반적으로 50-200개 표출이 기본으로 되어 있다. 표출되는 수가 너무 많으면 잘 보이지 않아 문제가 될 수 있으나, 표출되는 단어의 수가 적은 것은 문제가 되지 않는다.

김일성·김정일 시대 주요어를 비교해보면 김일성 시대 워드 클라우드에서 보이는 ‘동굴’, ‘돌칸흙무덤’, ‘벽화’, ‘개성’, ‘중세’, ‘일본’, ‘발전’이 김정일 시대 워드 클라우드에서는 ‘고대’, ‘질그릇’, ‘사람’, ‘기와’, ‘가마터’, ‘고인돌’, ‘청동기’으로 대체되었음을 알 수 있다. 또한 시간이 지나면서 ‘평양과 ‘신석기’의 중요도가 높아졌음을 알 수 있다. 이러한 변화는 김정일 시대부터 평양 지역에서 중요한 신석기시대 및 청동기시대 유적이 발견되고, 그 자료를 기반으로 ‘대동강문화론’이 제창되었던 결과로 이해할 수 있다.

김정일·김정은 시대의 주요어를 비교해보면 대부분의 주요어는 유지되고, ‘가마터’, ‘벽화무덤’, ‘고려’가 ‘집자리’, ‘벽화’, ‘조선사람’으로 대체되었음을 알 수 있다. ‘가마터’ → ‘집자리’에 대해서는 역시 2000년대 이후로 평양의 표대 유적과 리천리 유적에서 신석기시대 및 청동기시대 주거지가 지속적으로 발굴되면서 고대 집자리 연구가 집중적으로 이루어진 것의 결과로 이해할 수 있다.

세 시대에 걸친 주요어의 통시적 변화 양상을 보았을 때 가장 눈에 띄는 패턴은 ‘우리 나라’의 중요도가 점점 낮아지고, 그 대신 ‘문화’의 중요도가 점점 높아지는 현상이다. 이와 같은 경향성의 함의에 대해서는 고찰에서 살펴보도록 하겠다.

한편, 김정일 시대와 김정은 시대의 말뭉치는 그 규모가 커서, 각각의 시기 동안 중요하게 다루어진 주제들과 그것들 간의 관계에 대한 심도 있는 분석이 가능하다. 따라서 본 분석에서는 각각의 말뭉치를 대상으로 보다 많은 주요어가 표출된 워드 클라우드를 생성하고, 연이어 토픽 모델링 분석을 실시하였다.

우선, 김정일 시대의 주요어 100개를 표출시킨 워드 클라우드를 생성한 결과, ‘청동기’(시대)의 중요도나 ‘신석기’(시대)의 중요도보다 낮은 것으로 시각화되었는데, 그 이유는 이 시대에 집중적으로 발굴조사 및 연

구된 평양의 청동기시대 ‘리천리유적’이나 청동기시대의 대표적인 문화인 ‘팽이형토기문화’, ‘좁은숫단검’(문화) 등이 ‘청동기’(시대)와는 별도의 주요어로 인식되었기 때문임을 확인할 수 있었다. 이에 청동기시대 관련 주요어들을 ‘청동기’라는 표제어에 포함시킨 또 다른 워드 클라우드를 생성하였다. ‘청동기’ 표제어 보정 이전과 이후 생성된 워드 클라우드를 <그림 2>에 제시되어 있다. ‘청동기’ 표제를 보정한 결과, 워드 클라우드에서는 ‘청동기’의 주요도가 고구려 다음으로 ‘무덤’과 유사한 정도의 중요도를 보이고 있음을 확인할 수 있었다 .

<그림 2> 김정일 시대 말뭉치 ‘청동기’ 보정 이전 및 이후 생성 워드 클라우드 (주요어 100개 표출)



① ‘청동기’ 보정 이전



② ‘청동기’ 보정 이후

이와 같은 중요어를 가진 김정일 시대의 말뭉치로부터 유추될 수 있는 연구 주제들을 파악하기 위해 토픽 모델링 기법을 활용하였다. ‘토픽 모델링’(topic modeling)이란 텍스트 마이닝 기법을 활용한 통계적 추론을 통해 연구자의 주관이 개입되는 것을 최소화하면서 연구 주제를 포착하는 방법을 의미한다.³⁰⁾ 토픽 모델링 알고리즘은 확률 모델 알고리즘으로, 말뭉치 단어들의 통계적 분포에 대한 분석을 통해 복수의 잠재적 주제들을 추론하는 것을 가능하게 해준다.³¹⁾ 본 분석에서는 토픽 모델링 기법 중에 PLSI(Probabilistic latent semantic indexing)을 발전시킨 잠재 디리클레 할당(LDA) 알고리즘을 활용한 NetMiner4.0 소프트웨어의 토픽 모델링 툴을 이용하였다.

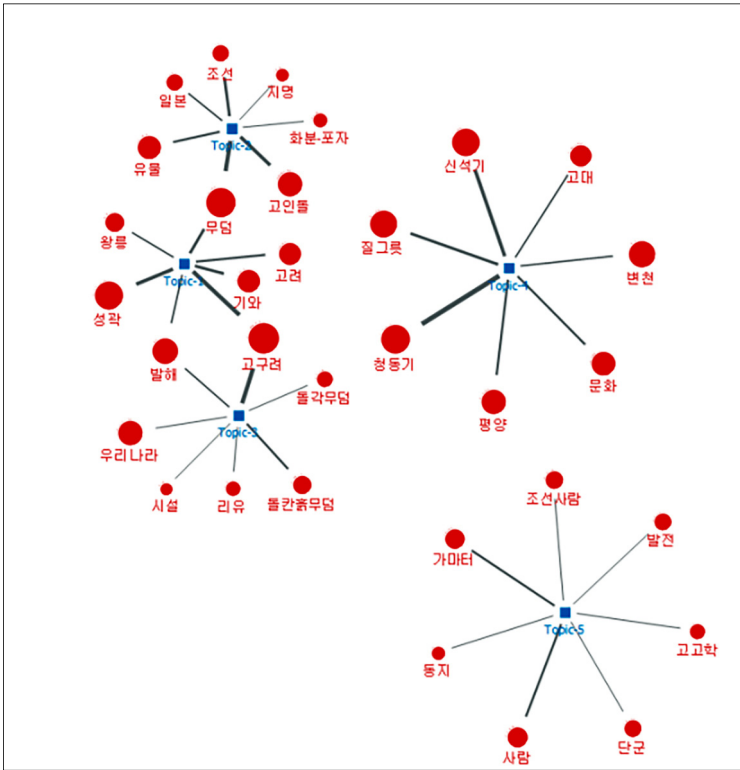
청동기시대 관련 주요어를 다시 ‘청동기’라는 표제어에 포함시킨 김정일 시대 말뭉치를 대상으로 토픽 모델링을 통해 총 5개의 토픽을 유추한 결과가 <그림 3>이다. 이 중 ‘토픽-1’, ‘토픽-2’, ‘토픽-3’은 모두 북한 고고학의 핵심 관심 사항인 선사시대(‘고인돌’)에서 고구려, 발해, 고려시대까지 이어지는 조선 무덤의 실증적 연구에 관한 주제 및 관련 주요어임을 알 수 있다. 토픽-2에서 ‘일본’이 연관 주요어로 등장하는 이유는, 한반도에서 기원한(북한에서는 ‘조선’이라는 표현을 사용한다) 무덤 형식의 일본으로의 전파가 이 시기 북한 무덤 연구의 중요한 관심사 중 하나였기 때문이다. ‘토픽-4’는 김정일 시대에 본격적으로 진행된 ‘대동강문화론’ 연구와 관련된 주제임을 한 눈에 볼 수 있다. 북한은 평양 지역이 예로부터 세계 문명의 중심지 중 하나였음을 주장하기 위해 김정일 시대에 들어와 신석기시대 토기 제작기술의 우월성이나 단군조선의 등장 기반이

30) 박중도, 「토픽모델링을 활용한 다문화 연구의 이슈 추적 연구」 『한국문헌정보학회지』 제53권 3호 (2019).

31) 정영주·김혜진, 「토픽모델링을 활용한 학교도서관 연구동향 분석」 『한국도서관·정보학회지』 제51권 3호 (2020).

된 청동기시대 문화에 대한 연구를 많이 진행해 왔음이 잘 알려져 있다. 그리고 ‘토픽-5’는 고고학 자료를 통해 드러난 단군을 조상으로 둔 조선 민족의 우월성 담론과 관련이 있음을 유추할 수 있다.

〈그림 3〉 김일성 시대 ‘청동기’ 보정 말뭉치에 대한 토픽 모델링 분석 결과
(5개 주제 네트워크 생성)



김정은 시대의 말뭉치 분량은 김일성 시대 그것의 절반 정도가 되는 관계로, 시험적으로 워드 클라우드에서도 김일성 시대에 표출된 것의 절

반에 해당되는 주요어 50개를 시각화하였다. 그러나 워드 클라우드를 비교할 때에는 말뭉치의 규모를 고려하기 보다는, 비슷한 형태의 워드 클라우드를 비교하는 것이 원칙인 관계로,³²⁾ 김정은 시대의 말뭉치에 대해서도 앞선 시대와 마찬가지로 주요어 100개를 표출한 ‘청동기’ 표제어 보정 이전의 워드 클라우드와 보정 이후 생성된 워드 클라우드를 생성하였다. 그 결과는 〈그림 4〉에 제시되어 있다.

‘청동기’ 표제어 보정 이전의 워드 클라우드를 보면, 김정일 시대에 비해 ‘팽이형토기문화’나 ‘표대유적’의 중요도가 낮아지고, ‘리천리유적’의 중요도가 높아진 것을 확인할 수 있다. 이는 리천리 유적에 대한 발굴조사의 대부분이 비록 김정은 시대 이전에 진행되었으나, 그 성과를 정리한 『리천리유적에 대한 연구』³³⁾가 2014년에 출간된 이후 이 유적에 대한 연구를 중심으로 대동강 유역의 청동기시대 문화에 대한 복원이 이루어지고 있기 때문이다.

한편, 김정은 시대에 들어와 ‘통일거리’가 새롭게 주요어로 등장한 사실이 흥미롭다. 통일거리 건설현장에서 약 2,200여 기의 낙랑 고분이 발굴된 것은 1990년부터 1992년 사이의 기간이다. 이 당시의 조사 내용은 김정일 시대 동안에는 단행본 보고서의 형태로 출간되었고, 『조선고고연구』를 통해 두 차례만 소개되었을 뿐이다. 하지만 김정은 시대에 들어서서 평양의 개발로 지속적으로 발견되고 있는 낙랑 고분의 발굴조사 내용이 보고서가 아닌 논문의 형태로 발신되고 있어서, 그 결과로 ‘통일거리’가 김정은 시대의 주요어로 등장한 것이다.

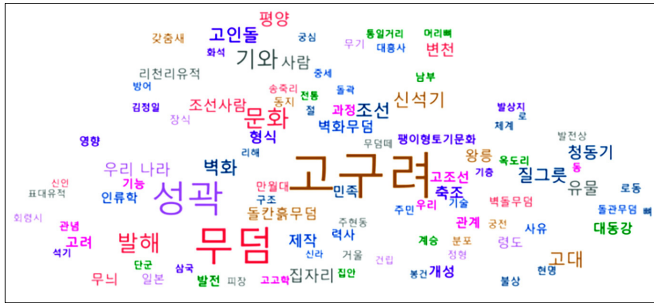
김정은 시대에 들어와 ‘왕릉’, ‘개성’, ‘문화’의 중요도가 높아지고 또한

32) Castellà, Q, Sutton, C., “Word Storms: Multiples of Word Clouds for Visual Comparison of Documents,” WWW’14, April 7-11, 2014, Seoul, Korea (2014).

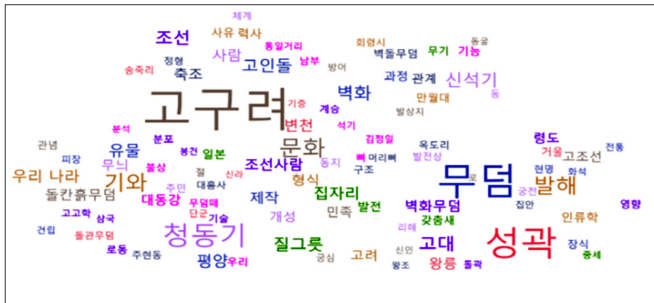
33) 차달만, 『리천리유적에 대한 연구(조선사회과학학술집 478 고고학편)』 (평양: 사회과학원, 2014).

‘만월대’가 새롭게 주요어로 등장한 사실은 매우 흥미롭다. 이는 남북발 굴조사단이 2007년부터 2015년까지 총 7차에 걸쳐 개성의 만월대 유적을 공동으로 조사한 성과에 따른 자연스러운 결과로 이해할 수도 있겠다. 그러나 또 한편으로는, 2013년에 ‘개성 역사유적 지구’가 세계문화유산으로 지정된 사실을 통해 알 수 있듯이 김정은 정권에서 역사문화콘텐츠의 활용과 사업화에 관심을 가진 결과로도 이해할 수 있다. 이 부분에 대해서는 고찰에서 추가적으로 검토하도록 하겠다.

〈그림 4〉 김정은 시대 말문치 ‘청동기’ 보정 이전 및 이후 생성 워드 클라우드 (주요어 100개 표출)



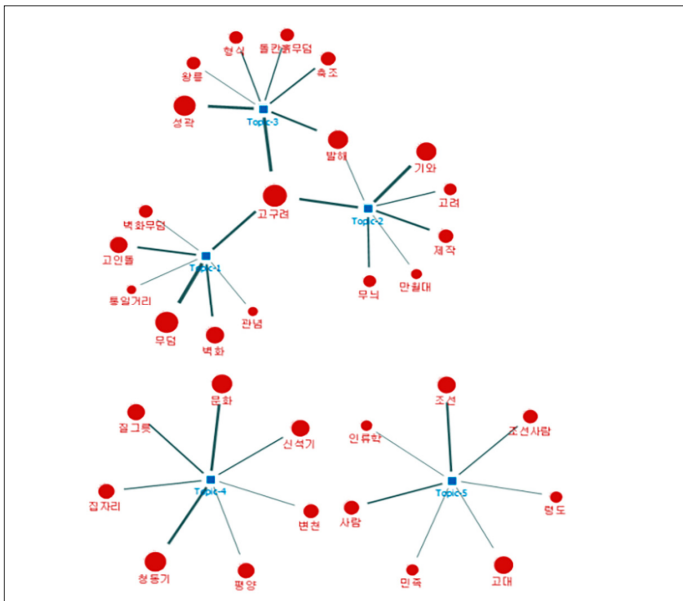
① ‘청동기’ 보정 이전



② ‘청동기’ 보정 이후

다음으로 ‘청동기’ 표제어를 보정한 김정은 시대 말뭉치를 대상으로 토픽 모델링을 한 결과가 <그림 5>이다. 유추한 총 5개의 주제 중 서로 연결된 ‘토픽-1’, ‘토픽-2’, ‘토픽-3’은 평양을 중심으로 확인된 청동기시대, 낙랑, 고구려 무덤(토픽-1), 역사시대(고구려, 발해, 고려)의 문화전통(토픽-2), 그리고 고구려-발해의 고고학 자료(토픽-3)에 관한 것임을 알 수 있다. 또한 ‘토픽-4’와 ‘토픽-5’는 각각 김정일 시대의 결과에서와 마찬가지로 ‘대동강문화론’ 연구와 관련된 주제와 조선 민족의 우월성 주장과 관련이 있음을 유추할 수 있다. 다만 흥미로운 것은 ‘토픽-5’의 경우, 김정일 시대와 달리 ‘단군’이 연관 주요어로 등장하지 않고, 그 대신 ‘령도’가 보인다는 사실이다.

<그림 5> 김정은 시대 ‘청동기’ 보정 말뭉치에 대한 토픽 모델링 분석 결과
(5개 주제 네트워크 생성)



3. 『조선고고연구』 수록 교시의 주요어와 시맨틱 네트워크

『조선고고연구』에 수록된 교시적 성격의 글은 고고학자에게는 관심의 대상이 되지 않으나, 북한 정치학의 수사나 담론에 주목하는 연구자들에게는 흥미로운 자료가 될 수 있다. 따라서 본 장에서는 교시적 성격인 글의 ‘제목’으로 구성된 말뭉치와 그러한 글의 초반에 늘 동일한 형식으로 등장하는 교시의 ‘내용’으로 구성된 말뭉치를 구축하였다. 이러한 두 종류의 말뭉치를 대상으로 김정일 시대와 김정은 시대의 담론·수사를 비교하였다.³⁴⁾ 참고로 교시적 글의 제목이나 교시의 내용에 대해서는 모든 형태소를 파악하는 것이 중요하다고 판단되어 단어에서 ‘조사’를 삭제하는 것 이외의 정규화 작업은 진행하지 않았다.

또한 비교의 방법으로 우선 워드 클라우드를 이용하였으나, 토픽 모델링은 진행하지 않았다. 이는 말뭉치의 크기가 토픽 모델링을 진행하여 유의미한 결과를 얻기에는 너무 작다고 판단했기 때문이다. 그 대신 교시의 내용에 대해서 시맨틱 네트워크를 다시금 생성하였다. 『조선고고연구』 수록 글의 제목에는 적용하지 않았던 시맨틱 네트워크 분석을 교시의 내용에 대해서 진행한 이유는 텍스트로부터 ‘개념’들을 추출하고 그 개념들 간의 ‘인접성 관계’를 노드와 링크로 이루어진 네트워크로 시각화하는 방법이 정치 담론을 규정하는 틀(structuring framework)의 구조를 시각화하는 데 용이한 도구로 여겨지고 있기 때문이다.³⁵⁾ 본고에서는 비정형 텍스트 데이터로부터 자동으로 노드를 추출해주는 NetMiner4.0 소

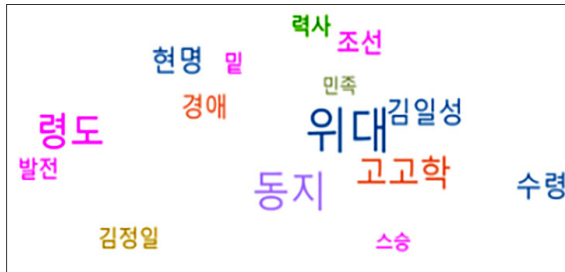
34) 김일성 시대에 해당되는 교시적인 글은 한 편에 불과하여 비교분석 대상으로 삼지 않았다.

35) 예를 들어, Baden, C. *Communication, Contextualization, & Cognition: Patterns & Processes of Frames' Influence on Peoples Interpretations of the EU Constitution* (Delft: Eburon Academic Publishers, 2010).

소프트웨어의 시맨틱 네트워크 툴을 이용해서 시맨틱 네트워크를 생성하였다.

우선 김정일 시대와 김정은 시대의 교시적 성격의 글 제목을 비교한 워드 클라우드를 보면 <그림 6>과 같다. 둘을 비교해 보았을 때 가장 눈에 띄는 차이는 김정일 시대에는 ‘김일성’과 더불어 ‘김정일’이 제목에 등장한 것과 달리, 김정은 시대에는 ‘김정일’만 등장하고, ‘김정은’은 등장하지 않았다는 사실이다. 또한 김정은 시대의 제목 주요어로 ‘민족’은 사라지고, 그대신 ‘발굴’, ‘유물’, ‘유적’, ‘사업’, ‘연구’와 같은 용어들이 새롭게 등장한다는 사실인데, 이는 김정은 정권이 고고학의 도구화를 통해 무엇을 추구하고 있는지에 대한 중요한 시사점을 제공한다.

<그림 6> 김정일 시대, 김정은 시대 교시적 성격의 글 제목의 모든 주요어를 시각화한 워드 클라우드



① 김정일 시대



② 김정은 시대

이러한 경향성은 『조선고고연구』에 수록된 교시적인 글의 교시 내용에도 반영되어 있다. 교시 내용 말뭉치를 대상으로 워드 클라우드를 생성하고 비교한 워드 클라우드가 <그림 7>인데, 이를 보면 의미있는 변화들이 확인된다. 우선, ‘민족’이 ‘역사’를 대신하여 가장 중요한 주요어가 되었다. 또한 ‘유산’, ‘문화유산’, ‘복구’와 같은 새로운 주요어들이 등장하였는데, 이 역시 문화유산을 적극적으로 활용하고자 하는 김정은 정권의 전략을 반영하는 것으로 판단된다.

<그림 7> 김정일 시대, 김정은 시대 교시적 성격의 글 초반 교시 내용의 모든 주요어를 시각화한 워드 클라우드



① 김정일 시대



② 김정은 시대

한편, 『조선고고연구』에 수록된 교시적인 글의 교시 내용을 대상으로 시맨틱 네트워크도 생성하였는데, 이때 단어 간 네트워크 생성을 위한 ‘윈도우 크기(window size)’는 ‘5’로 설정하였고, 유의미하다고 판단되는 단어 간 링크의 출현 빈도 값은 ‘3’으로 설정하였다.³⁶⁾ 즉, 본 분석에서는 연속된 5개의 단어를 분석의 단위로 삼고, 그 단위에서 두 단어가 3번 이상 공반되면 두 단어 사이에 연결성이 있는 것으로 파악하였다.

이와 같은 과정을 통해 생성한 시맨틱 네트워크가 <그림 8>에 제시되어 있다. 김정일 시대의 교시 내용을 대상으로 한 네트워크에서는 ‘력사’가 가장 중요한 용어일 뿐만 아니라, 중심성도 가장 높은 단어임을 확인할 수 있다. 이는 김정일 시대에는 단군릉의 발굴과 결부되어 그간 왜곡되었던 소위 ‘우리나라의 역사를 바로 세우는 것이 중요한 화두였기 때문인 것으로 판단된다. 그와는 달리 김정은 시대에 들어와서는 ‘민족’이 가장 중요한 용어이자 중심성도 가장 높은 단어로 정착되었음을 알 수 있다. 또한 워드 클라우드에서는 ‘력사’가 ‘문화유산’보다 중요도가 더 높은 것으로 시각화 되었으나, 시맨틱 네트워크에서는 ‘문화유산’의 중심성이 훨씬 더 높다는 사실을 알 수 있다.

또한 김정은 시대의 교시에서는 ‘위대한’, ‘김일성’, ‘김정일’이 등장하지만 이것들이 다른 개념들과는 밀접하게 연결되어 있지 않음이 포착된다. 실제로 김정은 시대의 교시 내용을 보면 일종의 양극화가 확인된다. 우선 김정은 시대의 교시적 글에 포함된 김정은의 교시 중 김정일, 김정일의 업적을 찬양하는 두 종류의 교시가 반복되어 등장한다. 그러나 이와는 달리, 반복되지 않고 실제적인 내용을 담은 교시도 등장한다. 반면, 김정일 시대의 경우에는 교시적 글에서 교시의 반복은 확인되지 않는다.

³⁶⁾ 시맨틱 네트워크 분석에서 ‘윈도우 값’이나 ‘출현 빈도 임계치’의 원리는 고일홍, 「북한의 발해 연구사 서술을 위한 새로운 접근 - 북한 학술지 목록화 사업 결과물을 활용한 내용분석」, 219~220쪽 등을 참고할 수 있다.

IV. 김일성 · 김정일 · 김정은 시대의 고고학에 대한 고찰

김일성 · 김정일 시대의 주요어 상위 20개를 표출한 워드 클라우드 (<그림 1> 참고)에서 포착되는 가장 큰 변화는 김일성 시대의 ‘발전’이 김정일 시대에는 사라진 점, 그리고 ‘평양의 중요성이 높아진 점이다. 이러한 현상은 1989년에 김정일이 제시한 조선민족제일주의의 영향을 반영하는 것으로 이해할 수 있다. 우선 ‘발전’은 맑스-레닌주의에 의한 유물사관과 관련이 있는 표현인데, 역사학에서의 주체사상이 본격화된 이후에도 지속적으로 사용되었다. 그러나 조선민족제일주의는, 비록 초기 민족주의의 애국주의적 전통을 계승했다고 할 수는 있으나, 사실은 수령, 당, 사상을 비롯한 체제 수호 의지가 그 본질이라고 할 수 있다.³⁷⁾ 따라서 이러한 조선민족제일주의의 영향으로 김정일 시대의 역사적 담론에서는 ‘발전’ 개념이 약화된 것이라고 평가할 수 있다.

또한, 잘 알려져 있듯이 북한 사회 내에서 조선민족제일주의 기반 민족주의를 고조시킨 것은 1993년 단군릉의 발굴이었고, 그것을 보다 체계화된 형태로 완성시킨 것이 ‘대동강문화론’이다. 주지하다시피 ‘대동강문화론’의 핵심 장소는 ‘평양’이다. 과거에 평양이 한민족의 삶에 중심적 역할을 해 왔듯이, 앞으로도 그럴 것이라는 자기예언적 희망을 담고 있는 것이³⁸⁾ ‘대동강문화론’의 핵심 내용 중 하나이다. 따라서 김정일 시대에 들어와 ‘평양’의 중요도가 높아진 것은 이러한 맥락에서 이해할 수 있다. 이러한 경향을 가장 잘 대변하는 것이 『조선고고연구』 1994-1호에 게재

37) 김동후, 「북한 역사학의 민족주의적 특징과 고대사 인식체계」 『북한학연구』 제18권 2호 (2022), 132쪽.

38) 위의 글, 133쪽.

된 「평양은 고대문화의 중심지」라는 제목의 논문이다.³⁹⁾ 『조선고고연구』 1994-1호는 북한 고고학에서 실로 중대한 전환을 보여주고 있는데, 이는 「단군릉 발굴정형에 대하여」, 「단군릉에서 나온 사람뼈의 인류학적특징에 대하여」, 「강동군 단군릉에서 발굴된 사람뼈에 대한 절대년대 측정결과에 대하여」가 각각 첫 번째에서 세 번째 논문으로 연이어 수록되고 있기 때문이다.⁴⁰⁾ 이 논문들의 뒤를 이어 네 번째 논문으로 수록된 것이 바로 「평양은 고대문화의 중심지」인데, 논문의 서두에서는 “평양시 강동군에서 단군릉이 발굴됨으로써 우리 나라가 반만년의 유구한 력사와 찬란한 문화를 가진 동방의 첫 선진문명국가였다는 것이 확증되었다”라고 밝힌 다음에,⁴¹⁾ 평양이 고대 문화의 중심지인 근거를 4가지 제시하고, 또한 이러한 평양의 고대 문화를 주변 지역들이 계승하였다고 주장하고 있다.

조선민족제일주의로 촉발된 이러한 김정일 시대 고고학의 방향성은 토픽 모델링의 결과에서도 확인되는데(〈그림 3〉 참고), ‘평양’이 관련 주요어로 등장하는 [토픽-4]와 ‘단군’이 관련 주요어로 등장하는 [토픽-5]에 주목할 수 있다. [토픽-4]는 평양이라는 장소의 중요성을 재생산하는 담론과 관련이 있는 주제로 이해된다. 관련 중요어로 ‘신석기’, ‘청동기’가 등장하는 이유는, 김정일 시대의 상위 100개 주요어를 표출한 워드 클라우드에서 등장하는 평양의 ‘표대유적’과 ‘리천리유적’에서 중요한 신석기 시대 및 청동기시대 유구와 유물이 발굴되어, 관련 연구가 평양의 중요성을 재생산하는 데 결정적인 역할을 했기 때문이다. [토픽-5]는 단군을

39) 석광준, 「평양은 고대문화의 중심지」, 『조선고고연구』 1994-1 (1994).

40) 박진옥, 「단군릉 발굴정형에 대하여」, 『조선고고연구』 1994-1 (1994); 장우진, 「단군릉에서 나온 사람뼈의 인류학적특징에 대하여」, 『조선고고연구』 1994-1 (1994); 김교경·전영수, 「강동군 단군릉에서 발굴된 사람뼈에 대한 절대년대 측정결과에 대하여」, 『조선고고연구』 1994-1 (1994).

41) 석광준, 「평양은 고대문화의 중심지」, 17쪽.

뿌리로 둔 조선사람의 우월성을 재생산하는 주제로 이해할 수 있다. 북한의 초기 민족주의가 소극적·방어적 측면에서 각 민족의 독자성과 자주성을 긍정하는 성격을 가졌던 것과 달리, ‘대동강문화론’으로 인해 북한의 후기 민족주의에서는 사회주의 민족의 우수성과 정통성이 적극적으로 강조되었다고 이해되고 있는데,⁴²⁾ [토픽-5]가 이러한 입장을 대변하는 것으로 판단된다.

2. 『조선고고연구』에 반영된 김정은 시대의 문화유산정책 및 관광정책

김일성·김정일 시대의 교시적인 글에 수록된 교시의 내용을 분석한 결과(〈그림 7〉 참고), 포착되는 가장 큰 변화 중 하나는 ‘밭굴’, ‘유물’, ‘유적’, ‘사업’, ‘연구’와 같은 용어들의 등장이다. 이러한 변화는 앞서 언급했듯이 2012년에 ‘문화유산보호법’이 채택된 영향으로 이해할 수 있다. 왜냐하면 이때부터 기존의 ‘문화유물보호법’에서는 ‘문화유물’이라고 칭했던 용어가 모두 ‘문화유산’으로 변경되었기 때문이다.⁴³⁾ 김정은은 ‘문화유산보호법’을 제정하고 2년 후인 2014년에 10월 24일 노동당 중앙위원회 책임일꾼(고위간부)들과 「민족유산보호사업은 우리 민족의 역사와 전통을 빛내는 애국사업이다」라는 내용의 담화를 나누었는데, 이는 이후 북한 내에서의 문화유산 정책에 직접적인 영향을 미친 것으로 평가되고 있다.⁴⁴⁾ 이 담화에서는 민족문화유산 보호와 대외홍보를 위한 남북, 국제 교류의 중요성이 강조되었고, 고고학과 관련이 있는 내용으로는 ‘민

42) 김동후, 「북한 역사학의 민족주의적 특징과 고대사 인식체계」, 132~133쪽.

43) 김유진, 「북한의 통치이념 변화에 따른 문화유산정책 연구」(한국전통문화대학교 문화유산학과 문화재법정책전공 석사학위논문, 2016), 86~87쪽.

44) 정창현, 「김정은시대 북한의 문화유산정책 변화와 남북교류」 『통일인문학』 제 77권 (2019), 379쪽.

죽문화유산 보존관리와 관련해 역사유적유물의 원상 보존관리, ‘민족유산보호사업의 과학화’, ‘역사박물관 등 ‘역사교양거점’을 활용한 교양사업’, ‘박물관·유적·명승지 참관·견학 활성화’ 등이 거론되었다.⁴⁵⁾ <그림 8>의 오른쪽에 제시된 김정은 시대의 교시 내용 시맨틱 네트워크에서 중심성이 두 번째로 높은 노드로 ‘문화유산’이 등장하고, 또한 이 노드와 연결된 왼쪽 아래의 클러스터에서 ‘복구’, ‘개선’, ‘유적’, ‘고증’과 같은 노드들이 분포해 있는 양상은 이러한 2014년 담화의 영향을 반영하는 것으로 볼 수 있다.

한편, 김일성 시대에 들어와 ‘만월대’가 주요어로 정착하게 된 것은 (<그림 4> 참고) ‘사회주의기업 책임관리제’의 도입과 연결해서 생각해 볼 수 있다. ‘사회주의기업 책임관리제’의 도입으로 북한의 기관들은 실질적인 수익을 내야 하는 상황에 직면하게 되었는데, 이렇듯 ‘실리’가 중요해진 시대적 배경 속에서 문화유산보호를 위한 예산 확보 차원에서 ‘개성 역사유적 지구’를 세계문화유산으로 등재한 것이라고 해석되기도 한다.⁴⁶⁾

현재 북한 내에서 만월대가 속해있는 세계문화유산 ‘개성 역사유적 지구’가 중요하게 여겨질 수밖에 없는 이유는 관광지를 경제적 이익을 가져다주는 상업적 공간으로만이 아니라, 국제적인 교류와 상호협조의 공간으로도 인식하고 있기 때문이다.⁴⁷⁾ 즉, 관광지를 국제화의 도구로도 접근하고 있는데, ‘세계문화유산’만큼이나 국제적인 교류와 상호협조에 적합한 장(場)도 없을 것이다.

45) 정창현, 「김정은시대 북한의 문화유산정책 변화와 남북교류」, 379쪽.

46) 위의 글, 381쪽.

47) 강채연, 「김정은 시대 관광산업의 국제화전략과 협력의 선택적 이중구조」 『한국동북아논총』 제25권 2호 (2020), 113쪽.

3. 『조선고고연구』에 반영된 북한 이데올로기적 담론의 변화

김일성·김정일·김정은 시대를 거치면서 확인되는 가장 큰 변화 중 하나는 ‘우리 나라’라는 표현의 중요성이 줄어든다는 사실이다. 김일성 시대에는 그 중요성이 ‘고구려’의 뒤를 이어 ‘무덤’과 비슷하였으며(그림 1) 참고), 김정일 시대에는 그 중요성이 줄어들었지만 그래도 [토픽-3]의 연관 주요어로(그림-3) 여전히 그 존재감을 유지했다. 그러나 김정은 시대에 들어와서는 그 중요성이 매우 약화되었고, 토픽 모델링 분석 결과에서도 더 이상 등장하지 않게되었다. 이러한 추세가 가지고 있는 함의를 파악하기 위해 북한 정권의 담론에 관한 선행연구를 참고하고자 하였으나, ‘우리 나라’에 관한 논의는 찾아볼 수 없었다. 따라서 토픽 모델링 분석 결과에서 ‘우리 나라’가 사라지는 대신 새롭게 연관 주요어로 등장한 ‘민족’에 주목하였다(그림 5) 참고).

그런데 주지하다시피 피상적으로만 보면 이러한 경향은 김정은 시대의 ‘민족’ 담론에 대한 선행연구와 상충되며, ‘조선민족제일주의’에서 ‘김정일애국주의’로의 전환 과정에 관한 내용과도 맞지 않는다. 즉, 2012년에 ‘김정일애국주의’가 확산되면서 북한에서는 ‘민족’보다는 ‘국가’를 앞세운 애국주의 담론이 제창된 것으로 알려져 있다.⁴⁸⁾ 다만, 2018년부터 북한에서는 ‘민족’이 다시 전면에 강조되는 양상도 관찰되고, 또한 ‘민족’ 담론이 특히 ‘문화유산’ 관련해서 부각되었다고 하니,⁴⁹⁾ 이러한 담론의 흐름이 김정은 시대의 『조선고고연구』 글 제목 말뭉치는 물론, 교시적 글에 포함된 교시 내용에서도 포착되는 ‘민족’의 주요어로서의 위상에 대한

48) 김보민, 「북한 민족제일주의 담론의 변화 - 조선민족제일주의에서 우리 국가제일주의로」 『현대북한연구』 제24권 1호 (2021).

49) 전미영, 「김정은시대 북한 민족주의: 담론·문화·정책」 『북한학보』 제43권 1호 (2018).

설명이 될 수 있을 것 같다.

이밖에 『조선고고연구』에 대한 분석을 통해 확인할 수 있었던 또 하나의 담론적 변화는 김정일 시대 글의 제목에서는 ‘김일성’과 ‘김정일’ 둘 다 주요어였던 것과 달리, 김정은 시대의 논문에서는 오로지 ‘김정일’만이 주요어로 등장한다는 사실이다. 이는 현재 김정은이 김정일 생전에 북한 혁명의 ‘유일한’ 혁명사상으로 규정됐던 ‘김일성의 혁명사상’(김일성주의)이 아닌 ‘김정일의 혁명사상’(김정일주의)을 정식화하고 체계화하고 있는 사실과 관련이 있어 보인다.⁵⁰⁾

V. 마치며

본고에서는 북한 고고학의 대표적인 일차자료인 『조선고고연구』를 대상으로 여러가지 디지털 기법을 이용하여 ‘멀리서 읽기’를 진행하였다. 필자는 북한에서 생산된 텍스트야말로 ‘멀리서 읽기’에 매우 적합하다고 생각한다. 왜냐하면 관련 텍스트의 분량이 방대한 것에 비해 그 자료에 관심이 있는 연구자가 (최소한 고고학 분야에서는) 많다고는 할 수 없고, 한정된 풀의 연구자가 무한하지 않은 시간을 투자하기에는 텍스트의 내용을 읽는 과정이 즐겁거나 생산적이라고 말하기 어렵기 때문이다. ‘멀리서 읽기’ 방법은 데이터를 수합하고 클리닝하는 과정이 지난하지만, 한번 구축된 데이터 세트는 메타 데이터만 제공된다면 다른 연구자와 공유되어 새로운 연구를 촉발할 수도 있다.

⁵⁰⁾ 김진환, 「김정은 시대 지배이데올로기의 특징과 전망: ‘김일성주의’에서 ‘김일성-김정일주의’로」 『북한연구학회보』 제17권 2호 (2013).

물론 ‘멀리서 읽기’가 제 기능을 발휘하기 위해서는 텍스트의 내용을 분석하고 맥락을 파악하는 ‘가까이 읽기(close reading)’도 함께 진행되어야 한다. 본고의 경우, 『조선고고연구』에 대한 ‘가까이 읽기’가 충분하지 않았던 측면이 있다. 이는 우선 ‘멀리서 읽기’를 통해 확보된 북한 고고학 관련 시사점에 대해 ‘가까이 읽기’를 통해 파고드는 작업이 본 학술지의 성격에 맞지 않다고 판단했기 때문이다. 또한 ‘멀리서 읽기’를 통해 확보된 북한의 정치 수사나 담론에 관한 시사점을 ‘가까이 읽기’를 통해 심층적으로 파악하는 것은 필자의 역량 밖의 일이었기 때문이다. 본 연구는 무엇보다도 『조선고고연구』에 대한 ‘멀리서 읽기’가 북한학 연구에서도 기여할 수 있는 바가 있음을 보여주기 위한 파일럿 스터디이다. 따라서 본고를 통해 북한학 연구자들이 『조선고고연구』에 관심을 갖고 다양한 관점에서 이 텍스트를 분석하여 본 연구의 부족한 지점들을 보충해줄기를 기대하고 있다. 최근 북한학 분야에서 『로동신문』 등에 대한 텍스트 분석이 활발히 이루어지기 시작했다는 사실을 감안하면⁵¹⁾ 이는 충분히 가능하다고 생각된다. 또한 이 파일럿 스터디는 『조선고고연구』 텍스트와 ‘디지털 인문학’의 시너지 효과를 충분히 설득력 있게 제시하고 있다고 판단되며, 앞으로 『조선고고연구』 전체 텍스트의 디지털화를 향한 첫 걸음이 되기를 희망하는 바이다.

■ 접수: 2023년 5월 11일 / 심사: 2023년 5월 30일 / 게재 확정: 2023년 6월 7일

51) 예를 들어, 박철수, 「텍스트마이닝을 활용한 북한 지도자의 신년사 및 연설문 트렌드 연구」 『Journal of Information Technology Applications & Management』 26(3) (2019); 임수진, 『북한사회의 집단적 자기화-타자와 담론 연구 - 『로동신문』을 중심으로』 서울대학교 사회학과 박사학위논문 (2021).

【참고문헌】

- 국립문화재연구소, 『조선고고연구 해제집 1』 대전: 국립문화재연구소, 2017a.
- 국립문화재연구소, 『조선고고연구 해제집 2』 대전: 국립문화재연구소, 2017b.
- 차달만, 『리천리유적에 대한 연구(조선사회과학학술집 478 고고학편)』, 평양: 사회과학출판원, 2014.
- 최진욱, 『김정일의 당권장악과정 연구』, 서울: 民族統一研究院, 1996.
- Baden, C. *Communication, Contextualization, & Cognition: Patterns & Processes of Frames' Influence on Peoples Interpretations of the EU Constitution*, Delft: Eburon Academic Publishers, 2010.
- Graham, S., Milligan, I., Weingart, S. *Exploring Big Historical Data: The Historian's Macroscopic*, London: Imperial College Press, 2016.
- Moretti, F. *Graphs, Maps, Trees: Abstract Models for a Literary History*, London: Verso, 2005.
- 강채연, 「김정은 시대 관광산업의 국제화전략과 협력의 선택적 이중구조」 『한국동북아논총』 제25권 2호, 2020, 101~125쪽.
- 강현숙, 「북한의 고구려 고고학 조사연구의 성과와 과제」 『문화재』 제53권 1호, 2020, 106~125쪽.
- 고일홍, 「北韓 考古學界 最近 現況 檢討 - 『조선고고연구』에 대한 內容分析을 中心으로 -」 『湖西考古學』 제41호, 2018, 102~131쪽.
- 고일홍, 「북한의 발해 연구사 서술을 위한 새로운 접근 - 북한 학술지 목록화 사업 결과물을 활용한 내용분석」 『인문논총』 제77권 4호, 2020, 193~237쪽.
- 김교경 · 전영수, 「강동군 단군릉에서 발굴된 사람뼈에 대한 절대년대 측정결과에 대하여」 『조선고고연구』 제1994년 1호, 1994, 11~16쪽.
- 김동후, 「북한 역사학의 민족주의적 특징과 고대사 인식체계」 『북한학연구』 제18권 2호, 2022, 121~159쪽.
- 김보민, 「북한 민족제일주의 담론의 변화 - 조선민족제일주의에서 우리 국가제일주의로」, 『현대북한연구』 제24권 1호, 2021, 139-177.
- 김진환, 「김정은 시대 지배이데올로기의 특징과 전망: '김일성주의'에서 '김일성-김

- 정일주의로」『북한연구학회보』 제17권 2호, 2013, 29~56쪽.
- 박중도, 「토픽모델링을 활용한 다문화 연구의 이슈 추적 연구」『한국문헌정보학회지』 제53권 3호, 2019, 273~289쪽.
- 박진욱, 「단군릉 발굴정형에 대하여」『조선고고연구』 제1994년 1호, 1994, 2~6쪽.
- 박철수, 「텍스트마이닝을 활용한 북한 지도자의 신년사 및 연설문 트렌드 연구」『Journal of Information Technology Applications & Management』 제26권 3호, 2019, 43~59쪽.
- 박초희·이성숙, 「연구 생태계 관점에서 본 국내 행정학 분야의 구조적 특성 분석 - 공저자, 인용, 키워드 네트워크 중심으로 -」『한국비블리아학회지』 제31권 1호, 2020, 213~235쪽.
- 석광준, 「평양은 고대문화의 중심지」『조선고고연구』 제1994년 1호, 1994, 17~20쪽.
- 양시은, 「최근 북한 고고학계의 고구려 연구 동향: 『조선고고연구』를 중심으로」『고구려발해연구』 제56호, 2006, 85~105쪽.
- 윤상덕, 「북한의 신라 고고학 연구 현황과 특징」『문화재』 제53권 2호, 2020, 270~285쪽.
- 이기성, 「북한학계 선사고고학의 최근 연구 동향」『통일고고학을 위한 연구현황과 과제 진단』, 한국고고학회 2016년 학술회의 (2016년 8월 23일), 3~29쪽.
- 이기성, 「북한 청동기시대 고고학 연구 경향의 변화와 쟁점」『문화재』 제53권 3호, 2020, 184~201쪽.
- 이형우, 「정치·사회적 맥락에 따른 북한 구석기 연구 변화」『문화재』 제53권 1호, 2020, 126~149쪽.
- 임사라·양문수, 「김정일 시대와 김정은 시대의 경제개혁 조치 비교 연구」『현대북한연구』 제25권 1호, 2022, 49~88쪽.
- 장우진, 「단군릉에서 나온 사람뼈의 인류학적특징에 대하여」『조선고고연구』 제1994년 1호, 1994, 7~10쪽.
- 전미영, 「김정은시대 북한 민족주의: 담론·문화·정책」『북한학보』 제43권 1호, 2018, 218~249쪽.
- 정영주·김혜진, 「토픽모델링을 활용한 학교도서관 연구동향 분석」『한국도서관·정보학회지』 제51권 3호, 2020, 101~121쪽.
- 정창현, 「김정은시대 북한의 문화유산정책 변화와 남북교류」『통일인문학』 제77호, 2019, 363~395쪽.

- Ascari, M., "The Dangers of Distant Reading: Reassessing Moretti's Approach to Literary Genres," *Genre*, vol. 47, no.1(2014), pp. 1-19.
- Castellà, Q, Sutton, C., "Word Storms: Multiples of Word Clouds for Visual Comparison of Documents," *WWW'14, April 7-11, 2014, Seoul, Korea*, 2014.
- Galke, L., Mai, F., Schelten, A., Brunsch, D., Scherp, A. "Using Titles vs. Full-text as Source for Automated Semantic Document Annotation," *K-CAP 2017, December 4th-6th, 2017, Austin, Texas, United States*, 2017.
- Kohl, P.L., Fawcett, C. "Archaeology in the service of the state: theoretical considerations" in P.L. Kohl, Fawcett, C., eds., *Nationalism, Politics and the Practice of Archaeology*, Cambridge: Cambridge University Press, 1995, pp. 3-20.
- Moretti, F., "Conjectures on World Literature," *New Left Review* 1, 2000. (<<https://newleftreview.org/issues/ii1/articles/franco-moretti-conjectures-on-world-literature>>)
- Moretti, F. "Style, Inc. Reflections on Seven Thousand Titles (British Novels, 1740-1850)," *Critical Inquiry*, vol. 36, no. 1 (2009), pp. 134-158.
- Trigger, B.G., "Alternative Archaeologies" in Historical Perspective," in J. Habu, C. Fawcett, J.M. Matsunaga, eds., *Evaluating Multiple Narratives*, New York, NY.: Springer, 2008, pp. 187-196.
- Underwood, T. "Distant Reading and Recent Intellectual History," in Gold, M.K., Klein, L.F. eds., *Debates in the Digital Humanities 2016*, Minneapolis, MN: University of Minnesota Press, 2016, pp. 530-533.
- 김유진, 「북한의 통치이념 변화에 따른 문화유산정책 연구」 한국전통문화대학교 문화유산학과 문화재법정책전공 석사학위논문, 2016.
- 임수진, 『북한사회의 집단적 자기화-타자와 담론 연구 - 『로동신문』을 중심으로』 서울대학교 사회학과 박사학위논문, 2021.

부록

『조선고고연구』 수록 글 제목 말뭉치 데이터 클리닝에 사용된
‘동이어’ 목록 (괄호 안 단어가 표제어임)

(비파형단검), 비파형단검문화, 비파형
(좁은숫단검), 좁은숫단검문화, 초기좁은숫단검문화, 좁은숫단검관계
(구석기), 구석기시대
(신석기), 신석기시대
(청동기), 청동기시대, 청동
고조선, 고조선관계
(벽돌무덤), 벽돌칸무덤
(고구려), 고구려적성격, 고구려 유적 유물, 고구려 사람, 고구려사람
(행렬도), 행렬
(질그릇), 그릇
(리조), 리조왕조, 리조시기
(고려), 고려왕조, 고려시기
(성곽), 성, 산성
(대동강), 대동강류역
(침촌형), 침촌리, 침촌
(양릉),릉
(해선시 답동), 답동, 해선리
(절), 절유적
(법종), 신라종, 고려종, 리조종
(기와), 막새, 암기와, 수기와, 암기와막새, 수기와막새
(팽이형토기문화), 팽이형토기, 팽이, 팽이그릇집자리
(돌각무덤), 돌각담무덤
(령도), 령도밀
(표대유적), 표대, 표대부락
(리천리유적), 리천리
(미송리류형문화), 미송리류형
(안학궁), 안학, 안학궁터
(고인돌), 고인돌떼
(남연해주), 남연해주일대
(우리 나라), 우리나라
(단군), 단군릉, 단군조선
(방사성탄소년대측정), 방사성탄소, 탄소년대, 탄소년대측정
(락랑), 락랑유적
(화분-포자), 화분, 포자
(벽화), 벽화고분

Comparative Study of North Korean Archaeology
in the Kim Il Sung, Kim Jong Il, and Kim Jong Un Eras
: A ‘Distant Reading’ of *Joseon Gogo Yeongu*

Ko, Ilhong (Seoul National University Asia Center)

Abstract

Joseon Gogo Yeongu is the representative academic journal on North Korean archeology, published quarterly by the Institute of Archeology of the Academy of Social Sciences since 1986, and can be said to be the most important primary source for tracing the history of North Korean archeology. By analyzing the titles of the texts included in *Joseon Gogo Yeongu*, this paper examines the changing ways in which the regimes of Kim Il Sung, Kim Jong Il and Kim Jong Un have utilized archaeology. Due to the vast nature of the text to be analyzed, the ‘distant reading’ method of ‘digital humanities’ was adopted. A total of three different ‘distant reading’ techniques were used in this analysis. First of all, text analysis was conducted on all papers published in the Kim Il Sung, Kim Jong Il, and Kim Jong Un Eras, and the key words of each era were visualized using ‘word clouds.’ 5 representative themes and related key words were also identified through ‘topic modeling.’ As a result, it was confirmed that the importance of ‘our country’ gradually decreased over the three eras, and the importance of ‘culture’ gradually increased instead. In addition, it was confirmed that in the Kim Jong Un era, the importance of ‘royal tomb,’ ‘gaeseong,’ and ‘culture’ increased and ‘Manwoldae’ appeared as a new key word. This

result stems from the fact that the superiority of the Joseon people and the Daedonggang Culture theory continued to be important research topics in North Korean archaeology even in the Kim Jong Un era. On the other hand, the government also had great interest in the development of North Korea's cultural heritage through archaeology. Keyword analysis and semantic network analysis were also conducted on the titles of the 'doctrine' articles included in *Joseon Gogo Yeongu*. It was found that the most important term of the previous era, 'minjok', disappeared in the titles during the Kim Jong Un era. However, analysis on the contents of the teachings of the 'doctrine' articles revealed that the most important keyword had changed from 'history' to 'minjok.' The centrality of 'cultural heritage' in the semantic network was found to be high, following 'minjok.' This was judged to be the result of a new understanding of the concept of 'minjok' in the Kim Jong Un era, as well as changes in North Korean regime's pursuit as to what is to be achieved from the archaeological findings.

Keywords: North Korean archaeology, *Joseon Gogo Yeongu*, Kim Jong Un era, distant reading, digital humanities

고일홍 (Ko, Ilhong)

서울대학교 아시아연구소 HK연구교수로 재직 중이다. 한국의 선사·고대 이주와 교류에 관심이 있는 고고학자로, 고대 동아시아 세계의 교역 네트워크와 환동해 지역 교류사에 대한 연구를 수행하며 지역 네트워크들의 연결성과 아시아 세계의 재생산에 주목하고 있다. 서울대학교 통일평화연구원의 지원을 받아 북한고고학에 대한 연구도 수행하고 있으며, 디지털 인문학과 역사학에 대해서도 고민하고 있다.