

신문 기사 속의 직업명사와 관련어*

정유진 · 강범모**

(고려대학교)

Chung, Eugene & Kang, Beom-mo. 2012. Occupation Nouns and Related Words in Newspaper Articles. *The Linguistic Association of Korea Journal*. 20(2). 145-170. In this paper, we have examined the aspects of occupation nouns and related words in news articles. Our investigation is based on the [Trends 21] Corpus (Kim et al., 2010; Kang & Kim, 2011), which is a morphologically annotated 400 million word corpus of Koreans newspaper texts. It is a collection of Korean newspaper texts covering most of the topics in print. The corpus is composed of four major Korean national newspapers issued from the year 2000 to 2009. We adopt a network analysis using co-occurring words with target words, i.e., occupation nouns. The extraction of co-occurring words is based on the statistical information (t-score) about the co-occurrences of words. The co-occurring words are regarded as not only semantically related but also mentally related to target words; thus we call the co-occurring words in a given context related words. A network is a graph-based structure made up of nodes and links. We adopt co-occurrence network analysis, which views lexical relationships in terms of nodes (representing individual words within the network) and connections or links (which represent relationships between the words). Pajek (Nooy, Mrvar & Vladimir, 2005) is used to visualize the relationship between the occupation nouns and their related terms. In addition, we have identified peoples' interest in social, cultural and economic issues that are portrayed in Korean newspaper articles.

주제어(Key Words): 신문 기사(newspaper articles), [물결 21] 코퍼스([Trends 21] Corpus), 공기어(co-occurring words), 관련어(related words), 공기 네트워크(co-occurrence network), 직업명사(occupation nouns)

* 이 논문은 2007년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (KRF-2007-361-AL0013).

** 제1저자: 정유진, 교신저자: 강범모

1. 서론

이 연구에서 우리는 2000년부터 2009년 사이에 발간된 조선일보, 중앙일보, 동아일보, 한겨레신문에서 나타나는 직업명사와 그 관련어를 추출하여 관련어 네트워크(co-occurrence network)를 제시한다. 연구대상이 되는 직업명사와 관련어의 양상을 알아보기 위해서 사회 연결망 분석의 네트워크 분석 방법을 이용한다. 명사의 관련어는 그 명사의 뜻을 분명하게 할 뿐만 아니라 사회의 관심을 반영하기도 한다(강범모·김홍규, 2011)¹⁾. 이를 기반으로 신문 기사에서 빈번히 언급되는 직업이 어떻게 대중이나 언론 매체의 관심을 반영하고 기술되는지 살펴보고 직업에 따른 관련어 양상을 분석한다. 특히 상위어, 하위어와 같이 계층관계에 있는 단어가 관련어를 통해 어떻게 그 의미가 두드러지게 되는지를 밝히고 문서의 주제에 따라 어떤 관련어와 함께 사용되는지 그 특징을 살펴보고자 한다.

신문이나 방송과 같은 언론 매체는 어떤 대상에 대한 대중의 관심이나 이해에 중요한 역할을 한다(강범모·김홍규, 2011). 이런 역할의 매개체가 되는 것이 언어이며 구성 성분인 단어의 의미를 제대로 알기 위해서는 그 단어의 사전적인 의미 외에 단어가 사용되는 문맥을 살펴보면 더 정확한 뜻을 알 수 있다(Ostler & Atkins, 1991; Lyons, 1995; Copestake & Briscoe, 1995). 특히 문맥 내에서 같이 나타나는 단어(co-occurring word)는 그 단어의 의미를 구체화하는데 도움을 준다. 어떤 단어와 의미적으로 관련 되어있는 단어가 관련되어 있지 않은 단어들보다 더 자주 동일 문맥 내에서 출현할 가능성이 높기 때문이다. Charles & Miller(1989)의 연구에 따르면 ‘big-little’이나 ‘large-small’과 같은 반의어 쌍들은 우연히 한 문장에 나타날 것으로 기대되는 빈도보다 더 자주 한 문장에서 사용된다고 한다. 즉 관련이 있는 단어들은 한 문장 내에서 자주 함께 사용된다는 것이다(Justeson & Katz, 1990).

이 연구에서 신문 기사를 연구 자료로 삼은 것처럼 신문이나 블로그에 나타나는 단어를 통해 사용 양상을 살펴보는 연구는 국내외 여러 기관과 연구자들에 의해 시행이 되어오고 있다(Saga et al., 2008; Michel et al., 2010; 김홍규 외, 2010; 다음커뮤니케이션, 2010; 정유진·강범모 2011). 다음 절에서는 이 연구의 언어자료로 사용된 [물결 21] 코퍼스에 대한 소개와 연구 대상이 되는 직업명사에 대해 언급한다. 3절에서는 직업명사와 관련어의 네트워크 분석을 시도하며 직업명사의 분류에 따른 상위어와 하위어의 관계를 관련어로 알아보고자 한다. 결론에서 이 연구에 대한 의의와 앞으로의 연구 과제에 대해 언급함으로써 논문을 정리 하도록 한다.

1) 퍼스(Firth, 1957)는 단어의 의미를 파악하는 여러 가지 방법 중 하나는 연구하고자 하는 단어와 같이 나타나는 단어들로 알 수 있다고 제시한다. 전통적인 방법에 따라 개념적 정의에 따른 단어 의미의 파악도 가능하지만, 단어가 사용되는 환경에서 의미의 일면을 파악할 수도 있다. 예를 들어, 사람들은 모르는 단어의 의미를 그 용법을 보고 확실히 알기도 한다.

2. 분석 대상과 방법

2.1. [물결 21] 코퍼스(2000-2009)

[물결 21] 코퍼스(김홍규 외, 2010; 강범모·김홍규, 2011)는 고려대학교 민족문화연구원이 한국을 대표하는 주요 4대 일간지의 2000년 이후 신문 텍스트 전체를 가공하여 언어 자원 데이터베이스로 구축하고 있는 자료이다. 이 연구의 자료가 된 코퍼스는 2000년부터 2009년까지 10년 동안 발간된 조선일보, 동아일보, 중앙일보, 한겨레신문에서 제공받은 약 200만 개의 신문 기사로 구성되어 있다. 이는 각 신문사가 매년 평균 약 5만 건의 기사를 싣고 있는 셈이다. 어절 수로는 약 4억 어절 이상이며 이는 British National Corpus의 약 4배, 즉 장편소설 약 8천권에 해당하는 분량이다.

이밖에 1년 동안 한 신문에 게재된 어절 수의 평균은 약 1천만 어절이며, 신문에 게재되는 1일 평균 기사 수는 약 156건이다. 명사만 논하자면 50만개 이상의 일반 명사가 실렸으며 이는 고려대 한국어 대사전에 수록된 일반 명사 수의 약 2배이다. 연도별 그리고 주제별(면종별) 기사 수와 어절 수는 다음과 같다.

표 1. [물결 21] 코퍼스에 포함된 기사 수와 어절 수 통계 (강범모·김홍규, 2011)

연도	기사 수	어절 수
2000년	231,361	39,591,435
2001년	220,109	39,001,720
2002년	214,941	41,572,150
2003년	205,635	42,283,054
2004년	200,260	39,971,067
2005년	180,413	37,416,116
2006년	162,055	36,947,492
2007년	181,493	42,179,864
2008년	176,697	41,773,224
2009년	165,050	42,787,812
전체	1,938,014	403,523,934

표 2. 주제별 기사 수와 어절 수 통계(강범모 · 김홍규, 2011)

주제	기사 수	어절 수
정치	194,056	40,949,310
국제	147,175	28,703,955
경제	352,886	69,700,427
사회	276,103	53,562,058,
문화	256,566	69,547,683
스포츠	154,543	25,917,307
과학	22,195	5,176,213
사실	34,543	8,750,631
오피니언	127,967	39,403,975
지역	31,562	10,878,509
기획	280,624	40,578,452
사람들	59,791	10,355,414
전체	1,938,014	403,523,934

주제에 따른 기사 수와 어절 수에서도 차이가 있는데 각 신문사가 가장 많은 지면을 할당하는 내용은 '경제 > 문화 > 사회 > 기획 > 정치 > ... > 사실 > 지역 > 과학'의 순서로 나타난다. [물결 21] 코퍼스는 다양한 주제와 많은 수의 기사를 포함하고 있으므로 이를 바탕으로 직업명사와 관련어가 신문에서 어떻게 기술되며 대중의 관심을 반영하는가를 고찰하는 기본 자료로 사용하기에 적합하다.

2.2. 직업명사

이 연구에서는 신문에 나타나는 직업에 대한 견해, 사회상, 그리고 사람들의 관심을 알아보기 위하여 직업명사를 대상으로 그 관련어들을 살펴보고자 한다. 현실에서는 수많은 직업을 가진 사람들이 사회 여러 분야에서 영향력을 발휘하며 살아가고 있다. 직업의 여러 종류와 관련어들을 통해서 그 직업의 성격과 그 직업이 우리 사회에 어떤 역할을 하며 사람들은 직업에 대해 어떤 관심을 갖고 있는지 고찰한다. 그리고 직업과 관련어를 살펴봄으로써 사회 변화의 흐름을 읽어볼 수 있을 것으로 기대한다. 한 예로 농업, 어업, 임산업에 관련된 종사자는 지난 반세기 동안 급격하게 감소하여 1955년 79.5%에서 1996년 16.1%로 축소되었는데(장홍근 외, 2006) 이러한 사실은 연구 대상이 되는 코퍼스에서 이 직종에 관련된 직업명사의 빈도가 현저히 적음을 통해서도 나타난다.

직업 자체에 대해 생각해 보면, 직업이란 생계를 유지하기 위하여 자신의 적성과 능력에 따라 일정한 기간 동안 종사하는 일(표준국어대사전, 2000)이며 이는 경제적으로 보상이 되는 활동이자 개인의 능력을 발휘하는 자기 실현의 활동이기도 하다. 김동규(2004)에 의하면

직업은 현대사회에서 사람들을 진정한 사회 구성원이 되게 하는 사회적 연대의 원천이라고 하였다. Rothman(1998)은 직업이 재화를 생산하거나 서비스를 제공하는 활동의 사회적, 기술적 측면으로 이루어지며 직업의 역할은 기술적인 차원과 더불어 사회적 차원의 의미가 결합된 것이라고 주장한다. 모든 직업의 역할에는 독특한 역사와 함께 사회문화적, 경제적, 정치적 맥락이 존재한다. 어떤 직업에 종사하는 사람은 일의 성격을 규정하는 작업 집단의 구성원으로 기능하며 또한 직업에는 직업 집단의 규범과 가치 및 그 집단의 신념체계 등이 존재한다(유홍준, 2005). 이와 같이 직업은 현대 사회의 핵심 구성 요소이기에 직업을 나타내는 명사와 관련어를 기반으로 직업의 역할에 대한 여러 맥락을 보고자 한다.

연구 대상이 되는 단어는 「한국표준직업분류」(통계청, 2010)²⁾의 국제표준직업분류를 근거로 대분류 10개(표3 참조)에 속하는 명사들 중 일부를 선정하였다. 하나의 직업은 직무상 유사성을 갖고 있는 여러 직무의 묶음이다³⁾.

표 3. 한국표준직업분류(통계청, 2010)

대분류	주요직업
고위임직원 및 관리자	대통령, 국무총리, 장·차관, 국회의원, 지방의회 의원, 정당노조, 경제단체, 기업의 최고 책임자, 행정기관, 기업, 단체의부서장
전문가 및 관련종사자	자연과학 및 공학 전문가, 의학 분야 전문가, 경제 분야 전문가, 교수, 교사, 법률 전문가, 정보 전문가, 사회과학 전문가, 예술가, 종교 관련 종사자
기술공 및 준전문가	각 분야 전문가와 유사한 업무 분야에서 전문가를 보조 보완하는 업무
사무종사자	사무용기기 조작원, 경리사무원, 금융사무원, 자재·생산·운송 관련 사무원, 도서·우편 및 관련 사무원, 금전출납원, 고객관련 서비스 사무원 등

- 2) 직업분류는 우리나라에 존재하는 직무가 어떤 수준에서든지 분류에 포괄되어야 한다는 포괄성의 원칙과 동일하거나 유사한 직무는 같은 단위 직업으로 분류되어야 한다는 배타성의 원칙에 근거하여 이루어졌다. 통계청에서 제시한 한국표준직업분류는 주어진 직무의 업무와 과업을 수행하는 능력인 직능을 근거로 편제되며, 직능수준과 직능유형을 고려하고 있다. 직능수준은 정규교육, 직업훈련, 직업경험 그리고 선천적 능력과 사회 문화적 환경 등에 의해 결정된다. 직능유형은 직무수행에 요구되는 지식의 분야, 사용하는 도구 및 장비, 투입되는 원재료, 생산된 재화나 서비스의 종류와 관련된다.
- 3) 한국표준직업분류에는 직무와 직업을 다음과 같이 정의 내리고 있다. 직무(job)란 “생산 활동에 종사하는 개별 종사자에 의하여 계속적으로 수행되었거나 또는 수행되도록 설정, 교육, 훈련된 업무”로 직업분류의 가장 기본적인 개념이며, 직업(occupation)은 “동일한 형태의 일의 집합”이다.

서비스종사자 및 판매종사자	항공기·여객선·열차·버스 등 관련시설의 서비스 종사자, 여행 안내원, 가사 근로자, 조리사, 이·미용사, 장의사, 소방관, 경찰관, 교도관, 모델, 점원, 판촉원, 판매원 등
농업과 어업 숙련 근로자	곡물·특용작물·채소 재배자, 과수·원예작물 재배자, 낙농 및 가축생산자, 동물 조련사, 임업 종사자, 어업근로자, 수렵원
기능원 및 관련 기능 종사자	단순한 기계조작이 아닌 기능과 숙련이 요구되는 작업, 광원, 조각원, 목공, 타일원, 미장원, 배관원, 용접원, 제관원, 가구제조원, 모피 가공원, 염색원
장치와 기계 조립원 및 종사자	산업용 기기 조작원(굴삭기, 재단기, 선풍기, 용광로, 주조, 유리성형, 식각기, 절삭기, 재봉기, 합판 제조기 등), 인쇄와 제본기 조작원, 열차·자동차 등의 운전 조작원
단순 노무직 근로자	신문 판매원, 행사, 방문 판매원, 구두 미화원, 세차원, 아파트 관리인, 가사보조원, 청소원, 호텔 객실종사원, 세탁인, 배달원, 수위, 경비원, 집표원, 광산인부, 건설인부, 환경미화원, 하역원
군인	군인, 군사, 요원

모든 직업명사를 고찰하기에는 다양한 직업이 있으므로 여기서는 위에 제시된 직업 부류 중 고위임직원 및 관리자, 전문가 및 관련 종사자, 사무종사자, 서비스 종사자 집단에 속하는 명사를 살펴보고자 한다. 특히 해당 부류 집단에 속하는 명사들 가운데 국립국어연구원에서 발표한 외국인을 위한 한국어 학습용 기본 어휘 6,000 단어(국립국어연구원, 2003) 중 직업을 지칭하는 명사를 실제 연구 대상으로 선정한다. 연구대상이 되는 직업명사는 상위어에 해당하는 직업 부류 명사와 하위어에 해당하는 직업명사를 선택하여 상위어/하위어의 관련어가 해당 명사를 이해하는데 어떻게 뜻을 구체화하는지 논의한다. 대상어는 고위임직원 및 관리자에 해당하는 ‘정치인, 대통령, 국회의원, 국무총리’와 전문가 및 관련 종사자에 속하는 ‘법조인, 변호사, 판사, 교원, 강사, 교수, 연구원, 교사, 선생님’, 사무종사자에 해당하는 ‘사무직, 회사원, 직장인, 은행원’ 그리고 서비스 종사자인 ‘공무원, 경찰관, 소방관’이다. 대상명사는 ‘정치인, 법조인, 교원, 사무직, 공무원’과 같은 상위 개념어와 각각의 하위어인 ‘대통령, 국무총리, 국회의원’, ‘변호사, 판사’, ‘경찰관, 소방관’, ‘강사, 교수, 교사, 선생님’, ‘회사원, 직장인, 은행원’으로 구분할 수 있다⁴⁾.

[물결 21] 코퍼스에 나타나는 연구 대상어의 빈도는 다음과 같다.

4) 이 논문에서 통계청 자료에 나온 수많은 직업명사를 모두 고려할 수는 없으므로 빈도와 기본 어휘 자료를 바탕으로 일부 직업명사 중 몇 분야에서 몇 단어를 선택하였다.

표 4. 대상어의 빈도 현황

대상어	절대빈도	상대빈도 ⁵⁾	대상어	절대빈도	상대빈도
정치인	49,624	1,339	교원	58,450	1,023
대통령	762,331	20,485	강사	31,993	862
국무총리	18,856	509	교수	426,323	11,559
국회의원	43,321	1,169	교사	145,468	3,941
법조인	5,046	136	선생님	41,150	1,113
변호사	77,256	2,086	사무직	2,527	68
판사	41,693	1,126	회사원	15,022	405
공무원	117,310	3,164	직장인	22,166	598
경찰관	16,610	448	은행원	2,614	70
소방관	4,092	110			

연구 대상으로 선정된 직업 가운데 신문 기사에서 가장 많이 언급되는 대상은 ‘대통령’이며 대부류에 속하는 사무직 및 그 하위의 ‘은행원’은 상대적으로 적은 빈도를 보인다. ‘대통령’ 이외에도 ‘교수, 교사, 공무원, 변호사’가 상위 빈도를 차지하고 있다.

2.3. 분석방법: 공기 네트워크

네트워크 분석은 그래프 이론을 이용하여 개체 간의 관계를 분석하는 정량적인 방법론이다. 이런 분석은 개체 간의 연결 관계나 발생 빈도와 주체의 흐름에 대한 정보 그리고 개체의 유형에 대한 정보를 분석할 수 있는 기반을 제공한다. 다시 말해, 분석 대상이 되는 요소를 그래프 이론, 통계확률과 시뮬레이션 등을 이용하여 그 구조를 분석하는 방법이다(Scott, 2000). 대표적인 예로, 시각화된 사회 연결망 분석(social network analysis), 의미 연결망(semantic network), 공기 네트워크(co-occurrence network)가 있다. 이 외에 네트워크 분석을 이용하여 정치인들 간의 상호 연결 관계나 자연과학 학술지에서 사용되는 키워드들의 연관도를 파악하는 것에 초점을 둔 동시단어 분석(co-word analysis)이 시도되기도 하였다(Callon et al., 1991; Bhattacharya & Basu, 1998).

Mitchell(1969)은 상호 작용하는 단위들 간의 ‘관계의 유형 및 질서’로 현상을 봄으로써 행위와 구조가 지닌 효과를 설명하는데 네트워크 이론이 적절하다고 주장하였다. 이 분석은 전통적인 통계적 방법론에서 제시하는 개별 행위 주체가 독립적이라는 가정 대신에, 개개의 행위자들이 상호 의존적이라는 가정 하에 서로 간의 관계에서 도출되는 특성을 분석하는데

5) 상대빈도는 절대빈도를 100만 번 중 나타나는 빈도로 환산한 숫자이다.

적합한 방법론이다(고길곤, 2007). 우리가 사용하는 언어의 구성요소로서 단어인 대상어와 공기어가 상호 의존적이라는 사실을 염두에 둔다면, 공기어를 기반으로 한 언어 사용 양상을 분석하기 위해서 네트워크 분석 이론을 적용하는 것이 가능하다.

대상어와 공기어가 공유하는 의미에 기초한 관계를 분석하기 위해 네트워크를 이용하기도 하였다(Doerfel 1998, Doerfel & Connaughton 2009, Danowski, 1983). 예를 들어, Ferrer & Solé(2001)은 한 문장 안에서 인접하여 같이 나타나는 단어들을 연결하여 네트워크를 구축해서 인간 언어의 특성을 정략적인 방법을 이용하여 분석하는 기초를 제시하였다. Jonassen et al.(1993)은 인식의 스키마를 일련의 개념들로 구성된 네트워크라고 정의하고 있다. 이 연구에서는 신문 기사라는 텍스트에서 직업명사가 함께 사용되는 단어에 의해 관계적 의미를 형성하는지 파악하기 위해서 신문에 나타나는 직업명사와 관련어로 네트워크를 구성하여 사람들의 인식의 스키마를 살펴보기로 한다.

공기 네트워크(co-occurrence network)는 대상어와 일정 문맥 내에서 인접하여 나타나는 공기어(co-occurring word)를 그래프 알고리즘 기반으로 시각화하는 것이다. 이 네트워크는 단어가 지닌 관계적 의미를 각 단어가 갖는 내재적 의미 외에 단어와 단어 간 관계적 특성을 통해서 파악한다. 이는 단어들이 특정한 문맥에서 같이 나오므로써 관계적 의미를 발생한다는 사실에 입각한 것으로 단어들이 공기하는 정도는 네트워크 분석에서의 ‘관계’가 된다(박한우 & Leydesdorff, 2004; Rice & Danowski, 1993)6).

공기 분석은 대상어와 공기어 사이의 관계성/관련성을 측정하는 방법의 하나로 통계연관 지수를 이용하여 대상어와 한 문장에서 자주 함께 출현하는 단어를 추출한다. 이 연구에서는 문장 내의 공기 관계를 기초로 연관지수로서 t-점수(t-score)를 사용한다(강범모, 2010)7). 단

-
- 6) 이 연구에서는 신문 기사를 기본 자료로 사용하였기 때문에 연구 대상이 되는 단어의 관계적 의미는 신문 텍스트에서의 의미가 부각이 된다. 그럼에도 신문 기사를 사용한 이유는 양적, 질적 면에서 다른 텍스트보다 많은 장점이 있기 때문이다. 우선 양적인 면에서 신문 자료는 그 어떤 종류의 텍스트 자료보다 많은 양을 모을 수 있으므로 분석의 안정성을 보장할 수 있다. 질적인 면에서 신문은 현 시점의 사회와 문화를 반영하는 동시에 사회와 문화의 거의 모든 영역에 걸쳐 있다.
- 7) t-점수는 기본적으로 공기되는 어휘들의 빈도와 각각의 어휘가 출현하는 빈도를 기초로 하여 관측빈도와 기대빈도를 기초로 두 어휘의 연관성을 측정한다. 관측빈도는 문단에서 대상어와 함께 나오는 단어(w)의 빈도를 세는 것으로 구해진다. 기대빈도는 전체 코퍼스크기(C), 전체 코퍼스에서 단어(fc(w))의 빈도, 그리고 대상어와 함께 나온 문단의 어절수(p)의 비율로 계산을 한다.

$$E_w = f_c(w) \times \frac{p}{C}$$

$$t\text{-score} = \frac{O - E}{\sqrt{O}}$$

E: Expected Frequency (기대빈도)

O: Observed Frequency (관측빈도)

어쌍은 이 연관지수가 높은 순서대로 배열되어 상위에 위치하는 단어들을 행렬 형태로 변환한 다음 Pajek(Nooy, Mrvar, & Vladimir, 2005) 프로그램을 이용하여 네트워크로 단어들의 관계를 시각화하여 구성한다.

네트워크의 시각화는 대상어와 공기어 사이의 관련성을 그래프 형식(bipartite graph)에 기반을 두어 그림으로 제시하는데 이는 언어 사용자에게 단어들이 공유하는 관계의 성격과 특징을 직관적이고 용이하게 파악할 수 있게 해준다. 즉, 일반적인 수치를 보는 것보다 그래프를 보면 한 눈에 관계적 특성을 쉽게 파악할 수 있다. 예를 들어, ‘정치인’이란 대상어와 신문 기사의 한 문장 내에서 공기하는 다른 인물, 정당, 또는 지역에 대한 관계는 표로 제시되었을 때보다 네트워크를 통해 쉽게 파악할 수 있다⁸⁾.

표 5. ‘정치인’과 공기하는 고유명사

대상어	공기어	공기정도 (t-점수)	대상어	공기어	공기정도 (t-점수)
정치인	민주당	31.730	정치인	오바마	10.842
	한나라당	29.771		워싱턴	10.692
	노무현	22.815		아베	10.664
	김대중	15.818		노사모	10.564
	자민당	15.582		김근태	10.560
	청와대	15.158		정동영	10.551
	박근혜	14.571		클린턴	10.490
	열린우리당	14.393		힐러리	10.278
	자민련	13.894		이시하라	9.848
	일본	13.645		안대회	9.694
	호남	12.381		야스쿠니	9.314
	이회창	12.365		케네디	9.029
	김영삼	12.124		이명박	8.817
	공화당	11.459		사회당	8.706
	고이즈미	11.328		민주	8.575

8) ‘정치인’이라는 대상어의 관련어가 정당명, 인명, 또는 지명과 같이 특정 고유명사로 제시되는 경우 단어의 사용에 따른 사회정치 연구에 일조할 수 있는 가능성이 있다.

서 사용되느냐에 따라 그 의미가 분명해지는 경향이 있다. 어떤 단어와 의미적으로 관련되어 있는 단어가 관련되어 있지 않은 단어들보다 더 자주 동일 문맥 내에서 출현할 가능성이 높다. 동일 문맥 안에 출현할 수 있는 단어들은 상하관계, 동의/유의관계, 전체-부분관계와 같은 치환적이면서 결합적인 어휘관계도 형성할 수 있다(Lyons, 1977; Cruse, 2000).

관련어는 치환적 또는 결합적 어휘관계를 맺고 있지는 않지만 잠재적으로 또는 언어 사용자의 연상 작용으로 '상대적으로 많이 관련/연관되어 있다고 느끼는 단어'이다. 예를 들어 '병원'의 관련어는 '의사'뿐만 아니라 '환자, 간호사, 병, 수술, 약, ...' 등 범위가 한정되지 않는다. 또한 관련어들의 목록이나 관련 정도는 개인의 경험에 따라 정도성이 달라질 수 있다(강범모 2010: 3). 이처럼 관련어는 대상어인 '병원'과 상하, 동의/유의, 반의 또는 전체-부분과 같은 어휘관계는 아니지만 '병원'이라는 단어와 충분히 연상되어 함께 생각해 볼 수 있는 단어들이다. 한편 공기어는 대상어와 관련이 되는 단어이자 사물을 표상하고 더 나아가 연관되는 개념을 표상한다. Leech(1974)가 정의한 언어적 의미란 "한 표현이 문맥에 나타나 있는 다른 표현의 의미에 영향을 받아 가지게 되는 의미"로 한 단어가 그것과 공기하는 단어들의 의미에 획득하는 연상(association)들로 구성된다. 따라서 대상어는 공기어로 나타나는 관련어에 따라 다양한 의미를 가질 수 있다.

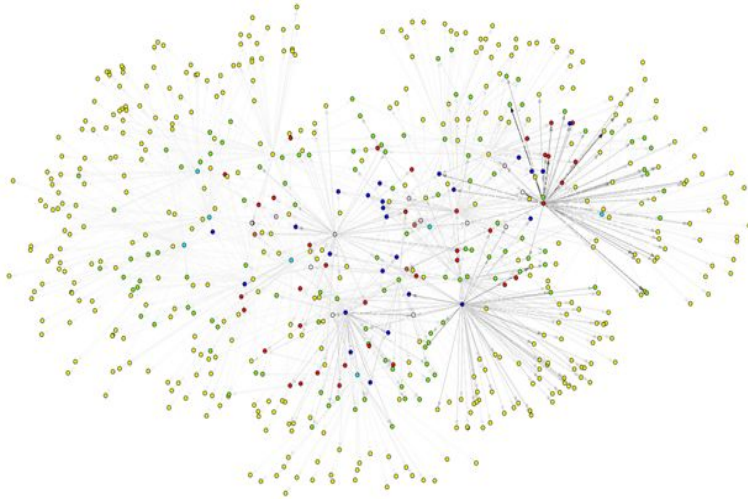
공기하는 단어들의 연관성은 통계적 유의성 지수 중 t-점수를 통하여 공기어 쌍을 추출한다. t-점수는 공기되는 어휘가 코퍼스 상에서 어휘 자체가 고빈도 어휘이기 때문에 대상어와의 공기 빈도가 높게 나오는 것과 저빈도 어휘이지만 대상 어휘와 유의미하게 함께 사용되는 경우를 구분하여 공기어를 추출시킬 수 있는 장점이 있다(강범모, 2011; 박병선, 2005; Church et al., 1991). 어휘들의 공기관계는 단어들 간의 관계를 연구하기 위해 전산 언어학이나 정보추출과 같은 분야에서 많이 사용되고 있으며 특히 단어들 사이의 관련성(word relatedness)을 측정하기 위해 사용된다. Spence & Owens(1990)는 공기 빈도가 높은 어휘들은 그 단어 간의 관련도가 강하다고 실험을 통해 밝혔다. 이렇게 추출된 공기어들을 이 연구에서는 관련어로 본다.

관련어 네트워크를 작성하는 첫 단계로, 연구대상이 되는 직업명사와 한 문장 내에서 공기하는 단어들을 추출하여 단어쌍이 공기하는 정도를 t-점수로 배열한다. 이후 연구 목적에 따라 상위 50개 또는 그 이상의 관련어를 선택한다¹⁰⁾. 이를 통해 어떤 단어들이 신문 기사에서 유의미하게 관계를 갖는지 확인할 수 있다. 다음 단계로, Pajek에서 요구되는 형식으로 대상어, 관련어 그리고 단어쌍의 관계를 나타내는 t-점수를 배열하여 그래프 기반의 네트워크로 시각화한다. 19개의 직업명사와 50개씩의 관련어 전체 네트워크는 그림 2와 같다¹¹⁾.

10) t-점수를 이용할 경우 공기 빈도에 의존하여 공기어를 산출하는 것보다 의미있게 관련성이 높은 단어들이 추출된다(박병선 2005, 강범모 2010; 2011).

11) 700여개의 단어가 한 번에 제시될 경우 복잡하게 표현이 되므로 편의상 노드의 이름, 즉 대상어와 관련어의 표시는 생략하였다.

그림 2. 19가지 직업명사와 관련어의 네트워크



이 경우 전체 노드수는 619개이며 밀도는 0.002이며 평균 연결정도는 3.07이다. 밀도(density)는 한 네트워크에서 단어들 사이의 연결된 정도를 의미한다. 다시 말하면, 네트워크 속의 관련어와 대상어 상호간 얼마나 많은 관계를 맺고 있는가를 표현하기 위한 개념이다. 평균 연결정도는 하나의 대상어가 평균 3개 정도의 다른 단어와 관련을 갖고 있다는 것을 의미한다. 연구대상이 되는 직업명사 19개와 모두 공유하고 있는 개념이나 단어는 나타나지 않는다. 직업명사와 함께 가장 많이 나타나는 관련어는 ‘말’이며 ‘정치인, 대통령, 변호사, 교수, 교사, 선생님, 회사원, 은행원’과 같은 대상어와 연결되고 있다. 관련어 중 ‘출신, 근무’도 다른 7개의 대상어와 연결이 되며 ‘학생, 전, 교육, 인사, 위원, 서울’의 경우 6개의 대상어와 연관이 되어있다.

3.2. 직업부류와 관련어

먼저 상위어에 속하는 직업부류를 나타내는 단어인 ‘정치인, 법조인, 교원, 사무직, 공무원’의 관련어 네트워크를 보도록 하자.

대상어가 되는 5개 직업부류 명사와 모두 공유되는 관련어는 나타나지 않고 ‘사무직’을 제외한 다른 4개의 대상어와 상호 관련되는 단어는 ‘인사’이다. 신문 기사에서는 4개 직업부류에 관한 관리나 직원의 임용, 해임, 평가 따위와 관계되는 행정적인 일을 기사화하고 있음을 다음의 용례를 통해 알 수 있다.

(1) 대상어와 ‘인사’가 함께 나타나는 용례

- ㄱ. 세종대 주경식 행정대학원장은 “공무원들이 신바람이 안나는 것은 지역편중 시비를 불러 일으킨 불공정한 인사 때문”이라고 지적한다.
- ㄴ. 이번 인사에서 치과 의사 출신으로 사법연수원 30기를 수료한 장연화(張宴華)씨가 서울 서부지청 검사로 임용되는 등 여성 법조인 21명이 검사로 신규 임용됐다.
- ㄷ. 그가 배속될 학교는 내년 4월 교원 정기 인사 때 정해진다.
- ㄹ. 많은 ‘문제의 경영진’ 가운데 낙하산 인사로 공기업에 간 여권(여권) 정치인 출신 임원들은 경질대상에서 빠지고 정치적 배경이 상대적으로 약한 전문경영인 출신들만이 이번 인사에 포함되어 있는 것도 많은 의혹을 자아내고 있다.

신문 기사에서 ‘사무직’이 ‘인사’와 관련되어 나타나는 경우도 있으나 다른 대상어에 비해 적은 빈도를 보이고 있어 고빈도 관련어에서 제외된 것이다. 실제 ‘공무원’과 ‘인사’가 같은 문맥에 나오는 빈도는 5004회, ‘정치인’의 경우는 1578회, ‘교원’은 1273회, ‘법조인’은 193회임에 반해 ‘사무직’은 30회에 그쳤다.

‘노조’는 ‘사무직, 교원, 공무원’의 상호관련어로 사용되었는데 지난 10년 동안의 신문 기사에서 이 직종에서 노조 문제가 사회적인 관심의 대상이 되었음을 알 수 있다.

(2) 대상어와 ‘노조’가 함께 나타나는 용례

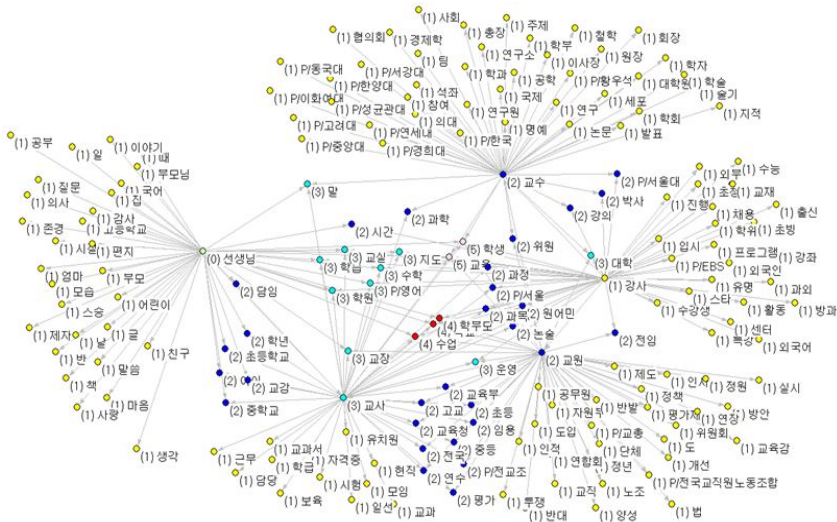
- ㄱ. 사무직들 사이에선 “노조 동의서 문제를 전체 직원투표에 부치든지 근로자 개인의 판단에 맡기자”는 주장이 높아졌다.
- ㄴ. 이 같은 갈등에도 불구하고 교원 단체 및 노조에 대한 교사들의 기대는 여전히 높다.
- ㄷ. 나라가 어려운 상황에서 국가의 최후 보루인 공무원들이 노조를 결성해 분쟁이라고도 일어난다면 나라 장래가 어떻게 될까 우려된다.

위의 두 예에서 보듯이 네트워크 분석을 이용할 경우 사회 속에서 직업 사이에 서로 공통된 관심사가 무엇인지를 쉽게 파악할 수 있다는 장점이 있다. 또한 ‘교원, 공무원’의 경우 이 직업군의 특징을 드러내는 ‘임용, 전국, 단체’ 등의 단어가 상호 관련어로 사용되어 두 대상어의 공통적인 특성을 보여준다.

‘교원’의 하위어에 해당하는 직업 중 ‘교수, 강사, 교사, 선생님’의 관련어 네트워크를 보

도록 하자. 공유하는 관련어와 특정 대상어와만 사용되는 배타적인 관련어를 통해 각 교직의 관계와 개별적 특성을 알아볼 수 있다.

그림 4. '교원, 교수, 강사, 선생님, 교사'의 관련어 네트워크



이 네트워크의 경우 5개 대상어에 대해 전체 노드수는 182개이며 밀도는 0.008이며 평균 연결정도는 2.75이다. 이전에 제시된 그림 3의 '정치인, 법조인, 교원, 사무직, 공무원'네트워크(밀도 0.05, 평균 연결평균정도 2.31)와 비교해 볼 때 밀도가 높다는 것은 상호 관련된 개념들이 조금 더 많다는 것, 즉 직업명사들 사이의 관계성이 높다는 것을 의미한다. 한편 같은 직업군 내에서 직업의 서로 다른 특색이 공유되는 개념이나 관계의 적음으로 나타날 수 있다는 것을 네트워크의 구조로 확인해 볼 수 있다.

(3) 교원의 하위 각 직업명사의 배타적 관련어 의미부류

- ㄱ. {교수}: 연세대, 성균관대, 서강대, 중앙대, 경희대, 고려대, 동국대, 이화여대 (대학교)
- ㄴ. {강사}: 과외, 전문, 스타, 수능 EBS, 입시 (입시)
- ㄷ. {선생님}: 부모님, 엄마, 부모, 어린이, 친구, 스승, 제자, 국어, 고등학교 (교육)
- ㄹ. {교사}: 유치원, 초등, 중등, 교육청, 현직, 전교조, 학급 (초중등교육체계)

(3)에 보이는 직업명사의 배타적 관련어는 개별 직업의 특성을 반영한다. 예를 들면, '교수'와 '교사'의 경우 직업의 터전이 대학교와 초중등교육이라는 차이가 그 관련어를 통해서 잘

드러난다. ‘강사’의 경우 사전적인 정의로만 해석한다면 학교와 학원에서 일하는 사람이나 대중이나 언론의 관심은 학원에 종사하는 강사, 즉 입시와 같은 시험에 보다 관심을 두고 있다. 이들 관련어는 어느 정도 우리의 상식선에서 예견될 수 있는 부분이 있다. 단어의 의미는 사전에서 제시되는 개념적인 의미 외에 사회 문화적 상황도 그 의미의 일면이다. 그 밖에 언어와 문화, 언어와 사회의 관점에서 단어가 사용되는 상황은 그 단어의 의미의 중요한 특성이다.

배타적 관련어와 달리 2개 이상의 직업명사에 공통적으로 나타나는 관계어는 그 직업들과 상관되는 공통 관심사를 반영한다. 대상어 간에 공유하는 관련어는 다음과 같다.

표 7. 교직 관련 직업명사 사이에 공유하는 관련어

대상어	관련어
교사, 교원	교육부, 교육청, 초등, 중등, 고교, 임용, 전국, 전교조, 연수, 평가
강사, 교원	전임, 교수
교수, 교원	위원
선생님, 교사	아이, 담임, 교감, 학년, 초등학교, 중학교
교수, 강사	강의, 박사, 서울대
교사, 강사	과정, 서울, 과목, 원어민, 논술, 교원
교사, 교수	과학
선생님, 강사	시간
선생님, 강사, 교원	교사
선생님, 교사, 교원	교장
교사, 강사, 교원	운영
교수, 강사, 교원	대학
교사, 선생님, 강사	수학, 영어, 학원, 학습, 지도, 교실, 학교, 학부모
선생님, 교사, 교수	말
교수, 강사, 선생님, 교사	학교, 학부모, 수업
교수, 강사, 선생님, 교사, 교원	교육, 학생

상위어인 ‘교원’과 모든 하위어들 상호 관련어는 ‘교육, 학생’으로서 교직에 종사하는 모든 직업에서 가장 핵심적인 요소와 ‘교육’과 ‘학생’임이 드러난다. ‘교원’을 제외한 다른 하위어 간의 상호 관련어는 ‘학교, 학부모, 수업’으로 이 개념들로 신문에 나타난 교직 명사의 기본 의미를 보여준다. ‘강사’의 경우 학교나 학원과 같이 장소에 따라 의미가 구체화 되는데 대상어들 사이의 상호 관련어를 통해 그 뜻이 구체화된다. 예를 들어, ‘강사’가 ‘교사, 선생님’

3개 대상어와 함께 나타나는 관련어는 법률과 의미적으로 밀접한 관계를 가진 단어이다. ‘판사’의 상위 개념어인 ‘법조인’과 상호관련어는 ‘판사’의 업무 관계적 의미를 잘 보이는 업무 시행 장소나 판사의 직급과 연관된 단어가 공유된다. ‘변호사’의 경우 최근 대중의 관심이 높아진 ‘로스쿨, 교수, 시험’ 등이 관련어로 제시된다.

전체적으로 우리가 고려한 직업명사와 그 관련어들의 네트워크는 다음과 같이 요약된다.

표 8. 관련어 네트워크 특성

부류	대상어	노드수	밀도	연결정도
사무직	회사원, 직장인, 은행원	166	0.0073019	2.4096385
정치인	대통령, 국무총리, 국회의원	145	0.0095785	2.7586206
공무원	경찰관, 소방관	134	0.0084165	2.2388059
법조인	변호사, 판사	118	0.0108648	2.5423728
교원	교수, 강사, 선생님, 교사	182	0.0075891	2.7472527

사무직 관련 직종과 정치인 관련 직종은 대상어가 모두 3개씩이나 노드수, 밀도, 연결정도가 다르게 나타나 네트워크의 구조적 차이를 비교해 볼 수 있다. 우선 노드수가 작다는 것은 상호 관련된 즉 공유하고 있는 개념이 많다는 것이며 이는 상대적으로 높은 밀도로 확인할 수 있다. 정치인 관련 직종의 경우 공유관련어가 사무직에 비해 상대적으로 적으며 각 대상어가 여러 다양한 관련어와 연결되고 있는데 이는 평균 연결정도로 알 수 있다. 공무원 관련 직종은 대상어의 숫자가 달라서 사무직이나 정치인과 직접 비교하기 힘들다. 법조인과 비교가 가능하다. 공무원에 비하여 법조인은 한정된 영역에서 역할을 하고 있음이 공유 연결어가 상대적으로 많음을 통해 드러난다.

연결망 밀도가 높고 사이즈가 작은 경우 구성원 사이에 강한 유대 관계를 형성하고 구성원들이 지식의 공유로 인해 필요한 자원들에 쉽게 접근할 수 있다는 장점이 있다 (Granovetter, 1973). 하지만 이 경우 구성원들이 비슷한 위치에 놓여있어 다양한 정보를 받아들이지 못함으로써 정보의 질과 새로운 정보 접근 기회가 적다는 점을 단점으로 꼽을 수 있다. 구성원 간의 중첩적인 정도가 낮고 성기지만 사이즈는 큰 형태를 갖는 연결망은 구성원들 사이 이질적인 정보를 갖고 있기 때문에 다양한 정보 획득이 용이하다. Granovetter(1983)는 약한 네트워크가 더 넓은 정보를 제공하며, 정보의 중첩성이 낮기 때문에 약한 네트워크를 가진 직업명사군이 그 해당 직업군에 대한 더 많은 정보나 관심의 다양성을 갖고 있다고 볼 수 있다고 하였다. 특정 명사들이 공기어로 많이 나타난다는 것은 그 명사들이 가리키는 대상들이 실제 현실에서 밀접한 위치에 있기 때문이다. 단어들의 밀도는

(5) '공무원'의 정치면과 사회면의 상호 관련어 용례

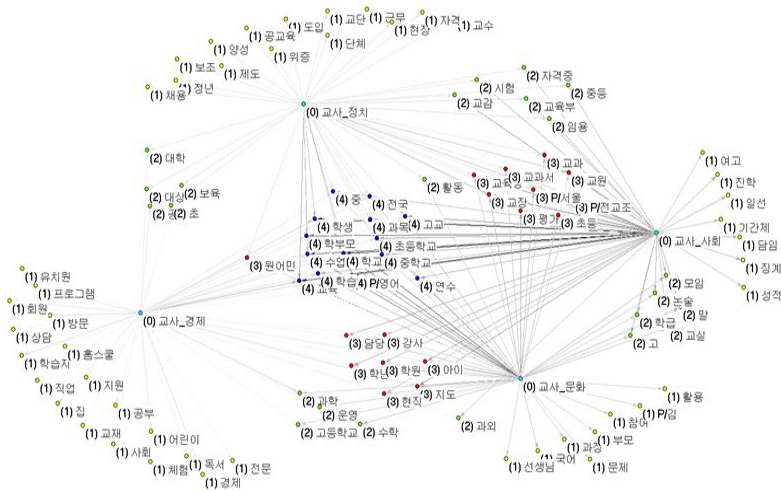
ㄱ. 경찰은 공무원들이 대학교 안에서 연 '공무원 노동자 대회 전야제' 모임에 2천여명의 병력을 들여 강제 해산하고, 공무원 772명을 강제로 데려가 '밤샘조사'까지 벌였다.

[정치면]

ㄴ. 프랑스에는 지방자치단체 소속 경찰이 없거나, 치안·교통·위생 등의 민생 업무를 담당하는 지방 공무원이 있어도 권한은 매우 제한적이다. [사회면]

다음 그림 7에서 '교사'의 각 주제별 관련어를 50개씩 추출하여 면종별 관련어 네트워크를 그려 보았다.

그림 7. '교사'의 정치면, 경제면, 사회면, 문화면 관련어



4개 주제 속에서 '교사'와 모두 함께 나타나는 관련어 양상은 다음과 같다.

(6) '교사'의 정치, 경제, 사회면, 문화면의 상호 관련어 의미부류

- ㄱ. 교육, 학교, 초등학교, 중, 중학교, 고교, 학생 (학교)
- ㄴ. 수업, 과목, 학습, 영어 (학습)
- ㄷ. 연수 (교원)
- ㄹ. 학부모, 전국 (기타)

정치면과 사회면에 공통으로 나타나는 관련어는 '자격증, 중등, 임용, 시험, 교감, 교육부'

이며, 정치면과 경제면은 ‘교육, 대학, 보육, 대상, 공무원’ 등 교육 정책의 실행에 연관된 단어들이다. 특정 대상어와 독자적으로 사용되는 배타적 관련어를 통해 주제에 따른 대상어의 특성을 고찰할 수 있다. 예를 들어, 경제면의 경우 ‘회원, 홈스쿨, 학습지, 교재, 프로그램, 방문’ 등 교육과 관련된 사업과 연관이 있는 단어가 배타적 관련어로 나타난다. 사회면에서는 ‘기간제, 징계’처럼 교사들의 사회 신분이나 처벌에 대한 개념들이 ‘교사’의 관련어로 추출된다¹⁴⁾.

4. 결론

이 연구에서 우리는 신문에서 나타나는 직업명사와 그 관련어를 공기 네트워크로 제시하여 신문 기사에서 빈번히 언급되는 직업에 대해 어떻게 대중이나 언론 매체의 관심이 반영되고 기술되는지 살펴보았다¹⁵⁾. 동아일보, 조선일보, 중앙일보, 그리고 한겨레신문에서 나타나는 직업명사와 같은 문장에서 함께 사용되는 단어를 t-점수를 이용하여 관련어를 추출하였다. 이를 통하여 직업명사들이 실제 문맥 내에서 그리고 신문 기사라는 텍스트의 특성상 사회 문화적 상황에서 어떻게 사용되고 다른 단어와 관련되어 관계적 의미를 구성하는지 살펴 보았다. 19개의 직업명사(‘정치인, 법조인, 공무원, 교원, 사무직’, ‘대통령, 국무총리, 국회의원’, ‘변호사, 판사’, ‘경찰관, 소방관’, ‘강사, 교수, 교사, 선생님’, ‘회사원, 직장인, 은행원’)를 t-점수값을 기준으로 상위 50개씩 관련어를 추출하여 네트워크 분석을 하였다. 네트워크의 밀도와 연결정도로 네트워크의 구조를 알아보고 상위 개념어와 각각의 하위어들의 배타적, 공유적 관련어를 통해 연구 대상이 되는 명사들의 의미를 직업 분야별로 알 수 있었다. 구체적으로 ‘교원, 공무원, 법조인, 정치인’은 ‘인사’라는 관련어를 공유하고 있으며 이는 신문기사에서 4개 직업부류에 관한 직원의 임용, 해임, 평가 등 인사 전반에 관계된 행정적인 일들이 기

14) 신문의 사회면에는 주로 정치, 경제, 문화의 영역을 제외한 일상사 중 가장 역동적인 사건·사고 기사가 실린다(박성희, 2004).

15) 한 심사자처럼 이러한 접근법에 대하여 근본적으로 반대하는 사람도 있을 것이다. 예를 들어 대상어의 의미가 맥락을 기반으로 해서 추출된 관련어를 통해 구체화되는데 있어 맥락의 범위를 설정하는 부분의 경우 이견이 있을 수 있는데 맥락은 병렬적 산출 방식으로 좌우 몇 단어에 한정할 수도 있고 이 연구에서처럼 문장 또는 문단까지도 하나의 맥락으로 볼 수도 있다. 그리고 신문 자료를 대상으로 단어의 의미를 알아보는 경우 자료 선정 역시 여러 의견이 있을 수 있다. 그러나 신문 자료는 다른 어떤 자료보다도 일상적이고 다양한 형식의 언어로 구성되어 있으며 코퍼스언어학의 관점에서 보면, 상당히 많은 양의 코퍼스는 신문 외에 소설이나 잡지와 같은 다양한 장르로 잘 구성된 적은 양의 코퍼스보다 가치가 있는 경우가 있다. 직관이라는 것도 결국은 언어 자료와의 경험을 통하여 형성된다는 것이 인지언어학에서도 인정되는 바이다. 이 연구의 대상 자료가 되는 코퍼스는 10년 동안 발간된 신문 텍스트로 구성되어 있어 다양한 언어현상을 살펴 볼 수 있으므로 신문에서 사용된 언어를 대상 자료로 사용된 것에 특정 텍스트로써의 한계는 없다고 본다.

사로 작성됨을 용례를 통해 알 수 있었다.

또한 직업부류와 구체적 직업을 나타내는 단어에 따른 관련어 양상을 분석하였다. 상위어에 해당하는 직업부류 명사는 하위어에 속하는 직업의 전반적인 특성을 상호 관련어를 통해 나타내는 특징을 보였다. ‘교원’의 하위어인 ‘교수, 강사, 선생님, 교사’는 배타적 관련어를 통해 각 직업의 지도 대상이 구체적으로 드러나는 것을 확인하였다. 일부 관련어는 우리의 상식 선에서 예측 가능한 단어들도 있었으나 공유하는 관련어를 통해 각 직업별 공통 관심사를 살펴 볼 수 있었다. 신문의 주제에 따라 대상명사의 관련어가 달라지는 것도 확인하였다. 예를 들어, ‘교사’의 경우 사회면에서는 ‘기간제, 징계, 성적, 일선’ 등의 관련어가, 정치면에서 ‘공교육, 자격, 정년, 채용’, 그리고 문화면에서 ‘활동, 부모, 참여, ...’ 등이 관련어로 나타나고 세 주제에 모두 공유되는 관련어는 학교에 관계된 ‘교육, 학교, 학생, ...’, 학습에 관계된 ‘수업, 과목, 학습, 영어, ...’ 그 외에 ‘학부모, 전국’ 등의 관련어가 함께 나타난다.

신문에 사용된 명사로 우리 사회에 많이 통용되는 직업에 대한 대중과 언론의 관심을 해석하려고 하였다. 코퍼스 언어학적 방법론을 사용해 특정 어휘의 앞뒤에 자주 붙는 어휘가 어떤 것인지를 알아보면 그 시대 사람들이 중요하게 생각하고 있는 것들, 즉 그 사회의 가치관이나 관심사 등을 알 수 있다는 장점이 있다. 특히 이러한 공기 네트워크 분석은 연구 대상이 되는 명사의 앞뒤에 자주 따라 붙는 관련어를 주제별로 구분해 분석해 보면서 관련된 사회적 변화를 보다 정교하게 추적해낼 수 있다는 장점이 있다.

이밖에 네트워크 분석의 장점 또는 유용성은 첫째, 어휘의 개별적 속성보다 관계성에 초점을 둔다는 점이다. 이는 기존의 어휘 네트워크 분석이 초점을 두는 개별 어휘의 속성만으로는 단어들의 의미 관계를 설명하는 데에 한계가 있다고 보기 때문이다. 즉 어휘들의 관계에 대한 분석을 통하여 개체의 속성을 밝힐 수 있다는 것이 네트워크 분석의 의의이다. 둘째, 네트워크 안에서 특정 어휘와 의미 부류에 따른 공기 관계를 통하여 의미군집에 대해 설명할 수 있다는 점과 문맥 내에서 함께 나타나는 복잡한 공기어 현상을 분석한다. 이밖에 단어 간의 관계분석을 통하여 숨어있는 구조적인 연결패턴을 발견하고 도식화할 수 있다. 즉 전체적인 수준에서 어휘간 연결 관계를 분석함으로써 공기어와 대상어의 구조적인 특성을 파악할 수 있다. 네트워크 구조는 어휘의 의미를 문맥 내에서 상호작용하는 어휘간의 관계와 질서를 파악할 수 있다는 장점이 있다.

우리가 언어생활에서 사용하는 단어도 네트워크 이론에서는 하나의 구성 성분이 되며 이 성분들은 서로 복잡하게 얽혀 있으면서 의사소통이라는 거대한 질서를 만들어 내고 있다. 네트워크 구조는 어휘의 의미를 문맥 내에서 상호작용하는 어휘간의 관계와 의미 영역을 파악하는데 도움이 될 것이라 기대한다. 향후 연구에서는 신문에 한정된 특정 텍스트 내에서의 의미 외에도 소설이나 다른 종류의 텍스트를 대용량으로 모아 유사한 연구를 시행한 후 신문 자료에 기초한 연구와 비교해 보고자 한다.

참고문헌

- 강범모. (2010). 공기 명사에 기초한 의미/개념 연관성의 네트워크 구성. *한국어 의미학*, 32, 1-28.
- 강범모. (2011). *언어 컴퓨터 코퍼스언어학* (개정판). 서울: 고려대학교 출판부.
- 강범모·김홍규. (2011). 명사 빈도의 변화, 사회적 관심의 트렌드: 물결 21 코퍼스 [2000-2009]. *언어학*, 61, 1-37.
- 고길근. (2007). 정책 네트워크 연구의 유용성과 사회연결망 이론 활용 방법의 고찰. *행정논총*, 45(1), 137-164.
- 국립국어연구원. (2003). 한국어 학습용 어휘 목록.
http://www.korean.go.kr/09_new/data/etc_view.jsp?idx=71
- 국립국어연구원. (2011). *표준국어대사전*. <http://www.korean.go.kr>
- 김용학. (2007). *사회 연결망 분석*. 서울: 박영사.
- 김동규. (2004). *직업사회학*. 서울: 건기원.
- 김병숙·강일규·진미석·장원섭·김해동·권진희 외 3인. (1998). *한국인의 직업의식 조사*. 서울: 한국직업능력개발원.
- 김홍규·강범모·이도길·김일환·홍정하·정유진 외 7인. (2010). *신문 텍스트 기반의 장기간 언어·사회·문화 연구 [물결 21] 사업 제1차 보고서*. 서울: 고려대학교 출판부.
- 문승태·박행모·장선철. (2005). *사회변화와 직업세계*. 서울: 태영출판사.
- 바라바시. (2002). *링크* 강병남·김기훈 옮김, 서울: 동아시아.
- 박성희. (2004). *신문 사회면 비교 분석 - 뉴욕타임스 요미우리신문, 조선일보를 중심으로*. 서울: 미디어연구소.
- 박한우·Leydesdorff, L. (2004). 한국어 내용분석을 위한 KrKwic 프로그램의 이해와 적용: Daum.net에서 제공된 지역혁신에 관한 뉴스를 대상으로. *Journal of the Korean Data Analysis*, 6(5), 1377-1388.
- 손태규. (2004). *서술적 보도와 해석적 보도 신문 정치면 비교 분석 - 뉴욕타임스, 요미우리신문, 조선일보를 중심으로*. 서울: 미디어연구소.
- 유홍준. (2005). *현대사회와 직업*. 서울: 그린.
- 이근희·김현석. (1995). *산업인력의 근로의욕 실태조사*. 서울: 한국생산성본부.
- 이진우. (2004). *신문 경제 섹션 비교 분석 - 뉴욕타임스, 요미우리신문, 조선일보를 중심으로*. 서울: 미디어연구소.
- 정유진·강범모. (2011). 친족명사의 공기어 양상과 네트워크 분석. *언어학*, 19(2), 209-235.
- 장홍근·한상근·이지연·정윤경·홍두승·서우석 외 1인. (2006). *한국인의 직업의식과 직업윤리 2006*. 서울: 한국직업능력개발원.

- 통계청. (2010). *한국표준직업분류*. 서울: 행법사.
- 한국생산성본부. (1995). *산업인력의 근로의욕 실태조사*.
- Bhattacharya, S., & Basu, P. K. (1998). Mapping a research area at the micro level using co-word analysis. *Scientometrics*, 43(3), 359-372.
- Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92, 1170-1182.
- Callon, M., Courtid, J-P., & Ladle, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics*, 22(1), 155-205.
- Charles, G., & Miller, A. (1989). Contexts of antonymous adjectives. *Applied Psycholinguistics*, 10, 357-375.
- Cruse, D. A. (2000). *Meaning in language*. Oxford: Oxford University Press.
- Copetake, A., & Briscoe, E. J. (1995). Semi-productive polysemous and sense extension. *Journal of Semantics*, 12, 15-67.
- Doerfel, M. L., & Connaughton, S. L. (2009). Semantic networks and competition: Election year winners and losers in U.S. televised presidential debates, 1960-2004. *Journal of the American Society for Information Science and Technology*, 60(1), 201-218.
- Danowski, J. A. (1993). Network analysis of message content. In G. Barnett & W. Richards (Eds.), *Progress in communication sciences XII* (pp. 197 - 222). Norwood, NJ: Ablex.
- Ferrer, R., & Solé, R. (2001). The small-world of human language. *Proceedings of the Royal Society of London (Series B 272)*, 561-565.
- Ferrer, R., Solé, R., & Köhler, R. (2004). Patterns in syntactic dependency network. *Physical Review*, E69, 051915.
- Firth, J. R. (1957). *Papers in linguistics*. London: Oxford University Press.
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(4), 1360-1380.
- Granovetter, M. S. (1983). The strength of weak ties: A network theory revisited. In P. Marsden (Ed.), *Social structure and network analysis* (pp. 201 - 233). CA: Lin Berverly Hills.
- Handelman, S. A. (1985). Fragments of the rock: Contemporary literary theory and the study of Rabbinic texts - a response to David Stern. *Prooftexts*, 5, 75-95.
- Jonassen, D. H., Beissner, K., & Yacci, M. A. (1993). *Structural knowledge:*

- Techniques for conveying, assessing, and acquiring structural knowledge*. Hillsdale NJ: Lawrence Erlbaum.
- Justeson, J., & Katz, S. (1990). Co-occurrences of antonymous adjectives and their contexts. *Computational Linguistics*, 17(1), 1-19.
- Kilgarriff, A. (1997). I don't believe in word senses. *Computers and the Humanities*, 31(2), 91-113.
- Leech, G. N. (1974). *Semantics: The study of meaning*. Penguin.
- Lyons, J. F. (1977). *Semantics 1*. Cambridge University Press.
- Lyons, J. F. (1995). *Linguistic semantics: An introduction*. Cambridge University Press.
- Michel, J., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., The Google Books Team, et al. (2010). Quantitative analysis of culture using millions of digitized books. *Science*, 331, 176-182.
- Mitchell, J. C. (1969). *Social networks in urban situation*. Manchester University Press.
- Nooy, W. D., Mrvar, A., & Batagelj, V. (2010). *Exploratory social network analysis with Pajek*. Cambridge University Press.
- Ostler, N., & Atkins, T. S. (1991). Predictable meaning shift: Some linguistic properties of lexical implication rules. In J. Pustejovsky & S. Bergler (Eds.), *Lexical semantics and knowledge representation* (pp 87-100). London: Springer.
- Pajek program - <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>
- Rothman, R. A. (1998). *Working: Sociological perspectives* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Rice, R. E., & Danowski, J. A. (1993). Is it really just like a fancy answering machine? Comparing semantic networks of different types of voice mail users. *Journal of Business Communication*, 30(4), 369-397.
- Scott, J. P. (2000). *Social network analysis: A handbook*. Thousand Oaks, CA: Sage Publications Ltd.
- Spence, D. P., & Owens, K. C. (1990). Lexical co-occurrence and association strength. *Journal of Psycholinguistic Research*, 19(5), 317-330.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.

정유진

136-701 서울 성북구 안암동 5가
고려대학교 민족문화연구원
전화: (02)3290-5288
e-mail: echung2@korea.ac.kr

강범모

136-701 서울 성북구 안암동 5-1
고려대학교 언어학과
전화: (02)3290-2173
e-mail: bmkang@korea.ac.kr

Received on April 11, 2012

Revised version received on June 12, 2012

Accepted on June 12, 2012