

학술논문의 저자키워드 출현순서에 따른 저자키워드 중요도 측정을 위한 네트워크 분석방법의 적용에 관한 연구

A Study on the Application to Network Analysis on the Importance of Author Keyword based on the Position of Keyword

권선영 (Sun-Young Kwon)*

초 록

본 연구는 학술논문의 저자키워드 출현순서에 따른 저자키워드의 중요도를 측정해보고자 하는 연구이다. 먼저 출현순서에 따른 저자키워드의 특징을 분석한 후 네트워크 분석 방법의 연결정도중심성, 근접중심성, 매개중심성, 위세중심성, 그리고 네트워크의 구조적공백성의 효과크기와 같은 지수를 사용하여 학술논문의 저자키워드 출현순서에 따른 저자키워드의 중요도를 측정해보았으며 각각의 네트워크 지수와 저자키워드의 출현순서와의 상관관계분석을 수행하였다. 네트워크 분석 지수 중 연결정도중심성 지수, 매개중심성 지수의 경우 각 학문분야별 저자키워드의 출현순서와의 상관관계의 결과에서의 유의한 분야의 수가 비교적 다른 지수에 비해 많았다. 이와 같은 결과를 통해 저자키워드의 중요도를 단지 출현빈도만으로 판단했던 것에서 벗어나 저자키워드의 중요도 측정을 위한 방법으로 연결정도중심성 지수, 매개중심성 지수도 고려해 볼 수 있음을 알 수 있었다.

ABSTRACT

This study aims to investigate the importance of author keyword with analysis the position of author keyword in journal . In the first stage, an analysis was carried out on the position of author keyword. We examined the importance of author keyword by using degree centrality, closeness centrality, betweenness centrality, eigenvector centrality and effective size of structural hole. In the next stage, We performed analysis on correlation between network centrality measures and the position of author keyword. The result of correlation analysis on network centrality measures and the position of author keyword shows that there are the more significant areas of the result of the correlation analysis on degree centrality, betweenness centrality and the position of keyword. In addition, These results show that we need to consider that the possible way as measuring the importance of author keyword in journal is not only a term frequency but also degree centrality and betweenness centrality.

키워드: 저자키워드, 저자키워드 출현순서, 네트워크분석, 중심성지수, 매개중심성
author keyword, author keyword order, network analysis, centrality measures,
betweenness centrality

* 성균관대학교 정보관리연구소 선임연구원(freesson@gmail.com)

■ 논문접수일자: 2014년 5월 24일 ■ 최초심사일자: 2014년 5월 28일 ■ 게재확정일자: 2014년 6월 20일
■ 정보관리학회지, 31(2), 121-142, 2014. [http://dx.doi.org/10.3743/KOSIM.2014.31.2.121]

1. 서론

1.1 연구의 목적

학술논문의 저자키워드는 학술논문의 내용을 요약하여 핵심적인 단어로 표현되는 학술논문의 매우 중요한 요소이다. 그렇기 때문에 연구자는 항상 본인이 집필한 논문에서 가장 핵심적이고 중요하다고 판단한 용어를 고심하여 선택하기 마련이다. 이렇게 선택된 용어는 다른 여러 필수적인 항목들과 함께 학술논문을 구성하게 된다. 연구자에게 있어 저자키워드의 선택과정은 실로 중요하다고 할 수 있는데, 저자키워드는 연구자뿐만 아니라 이용자, 그리고 데이터베이스 측면에서도 다양하게 활용될 수 있기 때문이다. 이용자는 학술논문의 본문을 살펴보기 전, 제목과 초록뿐만 아니라 저자키워드와 같은 항목을 반드시 살펴본다. 특히 이용자는 저자키워드가 논문을 대표하는 가장 핵심적인 용어로 구성되어 있다는 것을 인지하고 있다. 그러므로 이용자 입장에서는 저자키워드를 신속하게 확인하여 논문을 선택할 수 있으며, 미처 알지 못하였던 관련 키워드를 확인하여 추가적인 확장 검색에 활용할 수 있다. 연구자 입장에서는 본인의 학술논문이 다른 동료 연구자에게 선택되어 피인용될 경우 학술논문의 피인용 빈도가 증가할 것이다. 피인용 빈도는 결국 학술논문의 질을 나타내는 객관적 근거가 되므로 이를 높이기 위해서는 결국 학술논문이 이용자에게 선택이 될 수 있는 요소를 부각시키는 것이 중요할 것이다. 다른 한편으로 데이터베이스 측면을 살펴보면, 저자키워드는 색인어로 활용될 수도 있다(문혜원, 2000;

Taghva, Borsack, Nartker, & Condit, 2004; 박수현, 박경영, 2011). 물론, 저자키워드는 비통제어휘어의 특징을 지니기 때문에 기계적인 방식만으로는 한계는 있지만, 형태소 분석이나 일련의 정제 과정을 통해 확장 검색 및 전문용어 사전의 구축의 데이터베이스로도 활용될 수 있다(고영만 외, 2013).

저자키워드를 직접 부여하는 논문의 저자뿐만 아니라, 학술논문의 이용자 또한 저자키워드가 중요하다는 것에 대해서는 누구라도 동의할 것이다. 이는 사실상 저자키워드가 어떤 규정에 의해 정의된 것이 아닐지라도, 당연하게도 저자키워드 부여시 ‘중요한’ 용어를 부여하는 것이 필수적이고 당연시 여기는 행위가 인식의 바탕에 존재한다고 볼 수 있으며 학술커뮤니케이션의 모든 사람들은 이러한 행위 및 현상에 대해 이미 암묵적 동의를 하고 체화되었기 때문이다.

그동안 학술논문의 저자키워드와 관련한 연구들을 살펴보면 다음과 같다. 먼저 정보검색을 위한 색인어로서의 활용과 관련한 연구들을 찾아볼 수 있다(Taghva, Borsack, Nartker, & Condit, 2004; Gross & Taylor, 2005; 정영미, 2012; 김관준, 이재윤, 2012). 다음으로 저자키워드의 비통제어휘적 특징에 대한 연구 또한 찾아볼 수 있다. 문혜원(2000), 박수현과 박경영(2011)은 저자키워드와 MeSH용어와의 비교를 통해 저자 키워드 비통제어휘로서의 특징을 살펴본 바 있다. 최근에는 학술논문의 저자키워드에 대한 네트워크 분석방법을 적용한 연구들을 살펴볼 수 있다. 이러한 연구들은 학술논문에서 저자키워드를 추출하고 클러스터링 하여 학술논문이 속한 주제 분야의 지적구조 및

연구 동향을 밝히고자 한 연구들이다(이용덕, 2006; 박옥남, 2011; 이해영, 박승진, 2011; 조재인, 2011; 강범일, 박지홍, 2013; 서선경, 정은경, 2013; 장령령, 홍현진, 2014). 그 외에도 저자키워드 표기의 표준화를 통해 관련 논문들을 상호 연계 활용할 수 있는 방안을 제시한 양창진(2010)의 연구와 저자키워드의 출현빈도를 통해 학술용어사전의 구축대상용어로 선정하는 방법을 연구한 고영만 외(2013)의 연구가 있다.

이처럼 많은 연구들은 공통적으로 학술논문의 저자키워드의 중요성을 강조하면서 그 중요성을 측정하고자 논문 또는 분야 내 용어의 출현빈도를 측정하는 방법을 사용하였다. 이는 많은 출현빈도를 가진 용어가 적은 출현빈도를 가진 용어에 비해 더 중요한 용어라는 것을 전제로 한다. 출현빈도는 단순한 통계적 검증 뿐 아니라, 네트워크 분석 기반으로 활용된다는 것을 살펴볼 때 분명 의미가 있는 방법이라고 할 수 있다. 하지만 출현빈도 이외에도 용어의 중요성을 측정하기 위한 다른 방법을 고려해 볼 수 있다. 권선영(2013)은 특정 용어 네트워크 내에서 네트워크 분석을 수행하여 용어의 위치를 파악하고 이를 통해 용어의 중요성을 판단한 바 있다. 또한 이를 기반으로 용어 사전 구축 시 용어의 선정 방법으로 제안한 바 있다. 이외에도 장령령과 홍현진(2014)은 네트워크 분석 방법을 사용할 때 저자키워드의 순서를 살펴봐야 함을 강조하였다. 이 연구에서는 기존의 연구들이 저자키워드의 순서를 무시한 채 분석함을 지적하고 '수직 가중치'란 개념을 제안하여 키워드 순서에 따른 키워드의 중요성을 측정한 바 있다.

학술논문의 저자키워드를 살펴보면 단일 키워드가 아니라 대체로 복수의 키워드가 부여되어 있으며 나열되는 형태를 띠고 있다. 즉, 이와 같은 형태는 키워드가 출현한 순서에 따라 어떤 의미적 전개를 되고 있는지 살펴볼 필요성이 있다. 왜냐하면 연구자들은 복수의 저자키워드를 부여하면서 연구자만의 어떤 기준을 적용하기 때문이다. 이러한 기준은 이를테면 가장 핵심적인 키워드 즉, 가장 중요하다고 여기는 키워드를 가장 먼저 부여한 후 중요도 순으로 나열할 수도 있다. 반대로 핵심적이지 않은 키워드부터 부여하여 가장 핵심적인 저자 키워드를 마지막에 부여하기도 할 수도 있으며 이와는 다르게 어떤 규칙이 없이 키워드를 나열할 수도 있다.

본 연구에서는 연구자는 키워드의 중요도순에 따라 저자키워드를 부여한다는 것을 가정하여 '저자키워드는 단순한 나열이 아니라 중요도 값에 따라 순서대로 나열되어 있다.'라는 문제를 검증해 보고자 한다.

저자키워드의 중요도를 측정하기 위해서는 객관적으로 측정할 수 있는 지표가 필요하며 이를 위한 방법으로 네트워크 분석 방법의 연결정도중심성, 근접중심성, 매개중심성, 위세중심성, 그리고 네트워크의 구조적공백성의 효과 크기와 같은 지수를 사용하고자한다. 이러한 지수간의 비교를 통해 출현순서에 따른 용어의 중요도를 살펴보고 지수별 특징, 분야별 특징 등을 살펴보며 결과적으로 어떠한 지수가 출현순서에 따른 저자키워드의 중요도를 살펴보는 데 효과적인지 제시하고자 한다.

1.2 연구 문제

본 연구에서 살펴보고자 하는 연구문제는 다음과 같다.

첫째, 출현순서에 따른 저자키워드의 특징은 어떻게 나타나는가?

둘째, 출현순서에 따른 저자키워드의 중요도를 측정하기 위한 네트워크 지수 값은 분야별로 어떠한 특징을 지니고 있으며 출현순서와 각각의 네트워크 지수는 상관관계가 나타나는가?

셋째, 네트워크 분석 방법 중 어떠한 지수가 저자키워드의 출현순서를 구분하는데 효과적인지 수 있으며 어떠한 의미를 가지는가?

2. 연구 설계 및 방법

2.1 저자키워드 선정

본 연구는 한국연구재단 등재학술지에 등재된 2007년부터 2011년도 사이의 인문학 분야 학술논문 88,874건의 저자키워드 826,905개와 사회과학 학술논문 125,046건 중 985,580개의 저자키워드를 대상으로 이루어졌다. 이 중 인문학과 사회과학 분야별로 전체대비 논문 수 비율이 높은 상위 4개의 분야를 선정하였다. 인문학 분야의 경우 논문 수 비율이 높은 상위 4개 분야는 '한국어와 문학'(18.10%), '역

사학'(14.40%), '기타인문학'(9.30%), '영어와 문학'(8.40%)이었는데 이중 '기타인문학'은 특정분야로 분류되지 않는 다양한 분야 및 복합적인 분야의 학술지가 속한 경우이므로 분야의 특징을 살펴보기 어려운 문제가 있어 제외하였으며, '영어와 문학'의 경우 저자키워드의 경우 한글 키워드 없이 영문 키워드가 부여되어 있거나, 영문키워드가 더 많이 부여되어 있는 등의 특징을 보였기 때문에 제외하였다. 본 연구에서는 한글 키워드 외에는 분석 대상에서 제외하였으므로, '일본어와 문학'(7.60%)의 분야 또한 동일한 기준에 의해 제외하고 다음으로 '철학'(7.30%) 분야와 '언어학'(4.60%) 분야를 선정하였다. 사회과학 분야의 경우 인문학과 같은 문제가 발견되지 않았으므로 상위 4개의 분야인 '교육학'(20.0%), '법학'(16.5%), '경영학'(9.7%), '경제학'(5.3%) 분야를 선정하였다.

선정된 분야의 저자키워드는 정제되어있지 않은 형태이므로 형태적 통일과정과 같은 정제 작업을 거쳤다. 본 연구에서는 분석대상을 한글 키워드로 제한하였는데, 외국어 키워드의 대부분이 한글 키워드와 대응하여 나타나 의미적으로 중복되기 때문이다(학술논문에는 한글 키워드와 외국어 키워드가 같이 표기되는 데 일반적으로 한글 키워드를 나열한 후 외국어 키워드를 나열하고 있다. 용어를 살펴보면 대체로 한글 키워드에 대응되는 의미로 외국어 키워드를 나열한 경우가 많은 것을 살펴볼 수 있었다). 이와 같은 과정을 거친 저자키워드의 현황은 <표 1>, <표 2>와 같다.

〈표 1〉 인문학 분야 저자키워드 현황

분야	하위 분야	논문수	학술지 종수	학술지당 평균 논문수	키워드수	한글고유키 워드수	논문당 평균 키워드수
인문학	역사학	12,832	96	134	136,951	38,391	10.7
	철학	6,473	44	147	37,918	16,744	5.9
	언어학	4,106	17	242	35,886	6,867	8.7
	한국어와 문학	16,124	86	187	186,341	44,436	11.6
	합	39,535	243	177.50	397,096	106,438	9.22

〈표 2〉 사회과학 분야 저자키워드 현황

분야	하위 분야	논문수	학술지 종수	학술지당 평균 논문수	키워드수	한글고유키 워드수	논문당 평균 키워드수
사회과학	법학	20,640	101	204.4	208,918	53,523	10.1
	경제학	6,685	41	163.0	49,621	14,276	7.4
	경영학	12,183	62	196.5	86,405	19,151	7.1
	교육학	25,056	124	202.1	164,410	41,061	7.1
	합	64,564	328	191.50	509,354	128,011	7.92

2.2 저자키워드의 출현순서 산출 방법

저자키워드의 출현순서를 살펴보기 위하여 인문학 4개 분야 '역사학', '철학', '언어학', '한국어와 문학' 그리고 사회과학 4개 분야 '법학', '경제학', '경영학', '교육학' 분야로 나누어 저자키워드를 수집한 후 학술논문에서 키워드 필드에 출현한 순서를 살펴보았다. 출현 순서를 살펴보는 데 있어서 본 연구에서는 출현 빈도를 제한하지 않았다. 그 이유로 각 출현 순서에 따른 용어의 전체 개수를 살펴보기 위함이다. 다만 출현 순서를 10번째까지로 제한하였는데 이는 대체로 8개의 분야 모두 11번째부터는 용어가 출현하는 빈도가 전체 출현빈도의 1% 이내인 것을 살펴 볼 수 있었으며 이는 출현순서를 산정하는 데 있어 큰 의미를 둘 수 없기 때문이다.

다음으로 저자키워드에 대한 출현순서를 부여하였다. 예를 들어 〈표 3〉의 용어 중 '고구려'

는 1번째 147회, 2번째 70회 순으로 9번째의 경우 1회 나타나는 것을 볼 수 있다. 이와 같은 경우 가장 출현빈도가 높은 1번째 순서를 '고구려'의 출현순서로 보며 '첫 번째 순서에 출현하는 저자키워드'로 정의하였다. '백제' 또한 1번째 순서에서 다른 순서에 비해 더 많이 출현하였으므로 이 또한 1번째 순서에 출현하는 저자키워드로 정의하였다. 이와 같은 규칙에 따라 '민족주의', '역사교육'은 2번째, '역사교과서', '근대화'는 3번째 순서에 출현하는 저자키워드로 정의하였다. 사회과학 분야의 경우에도 이와 같은 규칙을 적용하여 '다문화 교육'과 '창의성'은 1번째, '정서지능', '다문화주의'는 2번째, '교사교육', '우울'은 3번째 순서로 적용하였다.

이처럼 다른 모든 분야 또한 동일한 규칙을 적용하여 키워드에 대한 출현순서를 부여하였으며 각 출현순서에 속하는 저자 키워드의 개수와 누적분포를 살펴보았다(〈표 5〉 참조).

〈표 3〉 용어의 출현 순서 산정을 위한 매트릭스의 예: 인문학 중 역사학 분야

분야	키워드	순서										출현순서
		1	2	3	4	5	6	7	8	9	10	
역사학	고구려	147	70	23	16	10	8	6	6	1	2	첫 번째
	백제	129	56	21	13	12	11	7	1	1	1	첫 번째
	민족주의	3	19	12	15	11	8	5	4	0	0	두 번째
	역사교육	12	19	8	12	1	1	0	1	1	1	두 번째
	역사교과서	16	13	17	4	3	1	0	0	0	0	세 번째
	근대화	5	8	14	6	9	1	1	0	1	0	세 번째

〈표 4〉 용어의 출현 순서 산정을 위한 매트릭스의 예: 사회과학 중 교육학 분야

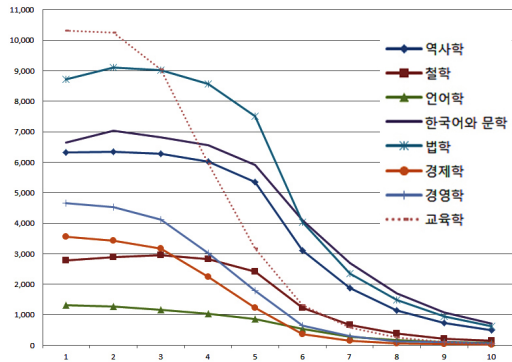
분야	키워드	순서										출현순서
		1	2	3	4	5	6	7	8	9	10	
교육학	다문화교육	138	80	34	25	7	4	2	3	0	0	첫 번째
	창의성	74	53	57	23	11	4	1	0	0	0	첫 번째
	정서지능	17	37	34	11	3	1	0	0	0	0	두 번째
	다문화주의	12	19	9	8	5	0	1	0	0	1	두 번째
	교사교육	24	28	44	29	19	3	1	0	0	0	세 번째
	우울	7	17	36	23	15	2	1	1	0	0	세 번째

〈표 5〉 출현순서가 적용된 저자키워드의 개수와 누적분포

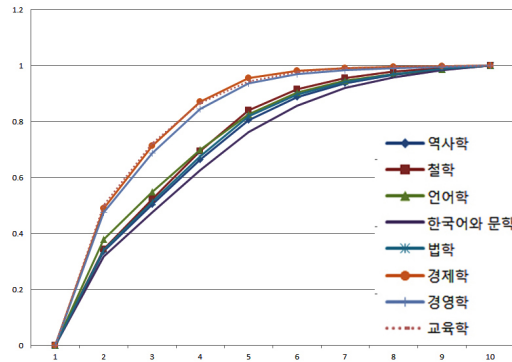
	1	2	3	4	5	6	7	8	9	10	합계
역사학	6,322	6,351	6,279	6,018	5,353	3,100	1,882	1,136	725	485	37,651
	-	33.7%	50.3%	66.3%	80.5%	88.8%	93.8%	96.8%	98.7%	100%	
철학	2,780	2,887	2,944	2,834	2,423	1,218	667	379	218	141	16,491
	-	34.4%	52.2%	69.4%	84.1%	91.5%	95.5%	97.8%	99.1%	100%	
언어학	1,314	1,262	1,153	1,025	854	539	282	161	112	90	6,792
	-	37.9%	54.9%	70.0%	82.6%	90.5%	94.7%	97.0%	98.7%	100%	
한국어와 문학	6,643	7,025	6,822	6,555	5,914	4,089	2,685	1,693	1,073	714	43,213
	-	31.6%	47.4%	62.6%	76.3%	85.7%	91.9%	95.9%	98.3%	100%	
법학	8,726	9,103	9,020	8,561	7,504	4,033	2,354	1,487	944	629	52,361
	-	34.1%	51.3%	67.6%	82.0%	89.7%	94.2%	97.0%	98.8%	100%	
경제학	3,558	3,421	3,164	2,249	1,236	354	145	58	36	23	14,244
	-	49.0%	71.2%	87.0%	95.7%	98.2%	99.2%	99.6%	99.8%	100%	
경영학	4,649	4,519	4,117	3,022	1,788	647	289	119	109	60	19,319
	-	47.5%	68.8%	84.4%	93.7%	97.0%	98.5%	99.1%	99.7%	100%	
교육학	10,319	10,246	9,043	5,962	3,170	1,321	569	247	97	44	41,018
	-	50.1%	72.2%	86.7%	94.4%	97.7%	99.1%	99.7%	99.9%	100%	

1번째부터 10번째까지의 출현순서가 부여된 저자키워드의 특징을 분야별로 살펴보면 다음과 같다. 인문학 4개 분야 사회과학 4개 분야의 출

현순서에 해당하는 저자키워드의 수는 출현순서의 앞에서 뒤로 갈수록 저자키워드의 수는 감소하는 것을 볼 수 있다. 하지만 '역사학', '철학',



〈그림 1〉 저자키워드개수의 변화



〈그림 2〉 저자키워드개수 누적분포변화

‘한국어와 문학’, ‘법학’과 같은 분야는 2번째 출현순서에 해당하는 저자키워드 수가 더 많다. 또한 분야마다 약간의 차이는 있으나 대체로 4번째 출현 순서까지는 용어의 감소폭이 그렇게 크지 않은 것을 볼 수 있다(〈그림 1〉 참조).

누적 출현빈도를 살펴보면 인문학과 사회과학 분야는 다른 특징을 보이는 데 인문학의 경우 3번째 순서까지 누적 출현빈도는 51.22%임에 비해 사회과학의 경우 65.86%이다. 이는 인문학 분야에 비해 사회과학 분야에서는 출현순서의 앞부분에 저자키워드가 더 많이 부여된다는 것을 의미한다(〈그림 2〉 참조).

2.3 저자키워드의 출현순서에 따른 평균 저자 키워드 출현 수 측정

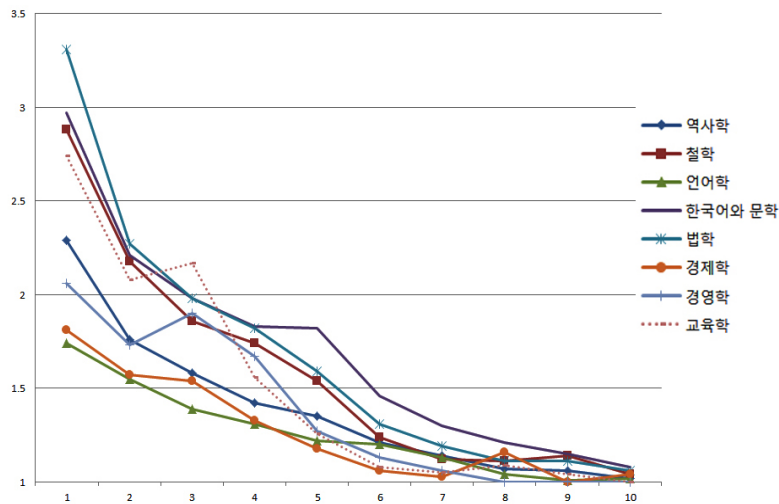
앞장에서 살펴본 각 출현순서에 따른 저자키워드의 개수는 고유한 저자키워드의 개수를 의미한다. 고유한 저자키워드는 1회 이상 출현하기 때문에 평균적으로 몇 번을 출현했는지를 살펴보기 위해서는 저자키워드의 총 출현빈도를 산정한 후 고유한 저자키워드의 수로 나누어주

어야 한다. 이를 통해 각 출현순서에 따른 저자키워드의 평균 출현 빈도의 차이를 살펴볼 수 있다.

평균 키워드 수가 1이라는 의미는 저자키워드의 수와 총 출현빈도가 동일한 것을 의미한다. 즉, 경영학 분야의 9번째와 10번째의 값은 1인데, 이는 모든 저자키워드가 단 한번만 출현한 것을 뜻한다. 인문학 분야의 경우 1번째 출현순서의 평균 저자 키워드 출현 수는 2.47회이며 이는 지속적으로 감소하다가 10번째 출현순서에서는 평균 1.04에 해당된다. 이를 살펴보면 인문학 분야에 속하는 4개의 분야 모두 유사한 패턴을 나타낼 수 있다(〈그림 3〉 참조). 하지만 ‘한국어와 문학’ 분야의 경우 평균 저자 키워드 출현수가 인문학 분야 평균 보다 높은 차이점이 있다. 사회과학 분야의 경우 1번째 출현순서의 평균 저자 키워드 출현 수는 2.48, 10번째 출현순서의 평균 저자키워드 출현 수는 1.02로 인문학 분야의 평균과 거의 유사하다. 사회과학 분야에서 속하는 4개 분야 또한 모두 유사한 패턴을 볼 수 있지만 ‘법학’ 분야의 경우 1번째 출현순서의 평균 저자 키워드 수가 3.31로 사회과학 분야 뿐 아니라, 전체 분야에서 가장 높은 것을 볼 수 있다.

〈표 6〉 저자키워드의 출현순서에 따른 평균 저자키워드 출현 수

	1	2	3	4	5	6	7	8	9	10
역사학	2.29	1.76	1.58	1.42	1.35	1.21	1.14	1.07	1.06	1.02
철학	2.88	2.18	1.86	1.74	1.54	1.24	1.12	1.11	1.14	1.04
언어학	1.74	1.55	1.39	1.31	1.22	1.20	1.13	1.04	1.01	1.02
한국어와 문학	2.97	2.21	1.98	1.83	1.82	1.46	1.30	1.21	1.15	1.08
인문학 평균	2.47	1.92	1.70	1.57	1.48	1.28	1.17	1.11	1.09	1.04
법학	3.31	2.27	1.98	1.82	1.59	1.31	1.19	1.11	1.11	1.06
경제학	1.81	1.57	1.54	1.33	1.18	1.06	1.03	1.16	1.00	1.04
경영학	2.06	1.73	1.90	1.67	1.27	1.13	1.06	1.00	1.00	1.00
교육학	2.74	2.08	2.17	1.56	1.26	1.08	1.05	1.09	1.04	1.00
사회과학 평균	2.48	1.91	1.89	1.59	1.32	1.14	1.08	1.09	1.04	1.03



〈그림 3〉 저자키워드의 출현순서에 따른 평균 저자 키워드 출현 수의 비교(분야별)

2.4 네트워크 분석 방법

출현순서에 따른 저자키워드가 네트워크에서 어떠한 위치를 차지하고 있는지를 살펴보기 위해 네트워크 분석을 수행하였다. 이 네트워크는 2.1절에서 언급한 인문학 4개 분야, 사회과학 4개 분야로 각각의 학술논문이 속한 분야를 뜻한다. 즉, 네트워크분석을 통해 학문분야

의 맥락에서 중요도가 높은 키워드를 살펴보고 이를 저자키워드의 중요도로 판단하였다. 이는 연구자가 저자키워드를 부여할 때 논문에서의 유일한 키워드뿐만 아니라, 학문 분야 내에서 중요한 용어 또한 부여한다는 것을 감안해볼 때 분야의 네트워크 분석을 통해 저자키워드의 중요도를 판단하는 것은 의미가 있다.

앞서, 출현 순서에 따른 저자키워드의 개수

〈표 7〉 네트워크 분석을 위한 저자키워드의 분야별 제한 기준

인문학	출현빈도	저자키워드수	사회과학	출현빈도	저자키워드수
역사학	5회 이상	1,423	법학	11회 이상	1,164
철학	5회 이상	945	경제학	3회 이상	1,176
언어학	2회 이상	1,104	경영학	4회 이상	1,319
한국어와 문학	10회 이상	1,200	교육학	10회 이상	1,088

를 산정할 때는 출현순서를 10번째까지로 제한하였지만 출현빈도는 제한하지 않았다. 반면, 네트워크 분석에서는 저자키워드의 수를 제한하였다. 그 이유로 네트워크 분석을 위해서는 동시출현 행렬 값이 반드시 필요한데, 저자키워드 수가 많아질수록 동시출현 행렬 값의 산출 시간 및 네트워크 분석도구에서 이를 처리하는 시간도 오래걸리기도 하고 예기치 못한 에러 또한 발생할 가능성을 배제할 수 없기 때문이다. 본 연구에서는 각 분야마다 출현빈도를 제한하여 〈표 7〉과 같이 분석 대상 키워드수를 정하였으며 이에 대해 일반적으로 많이 사용되는 네트워크 중심성 지수인 연결정도중심성, 근접중심성, 매개중심성, 위세중심성과 구조적 공백성 지수중 효과크기를 UCINET으로 측정하였다.

이와 같은 제한 기준을 정한 후 용어의 출현순서와 매핑을 한 결과, 철학분야는 10번째 출현순서에 해당하는 용어가 없고, 법학 분야의 경우, 8번째, 10번째, 경제학의 경우 9번째와 10번째, 경영학의 경우 8~10번째, 교육학 분야의 경우 6~10번째에 해당하는 용어를 발견할 수 없었다. 그러나 각 순서에 해당하는 용어를 모두 매핑 하기 위해 출현빈도를 하향하는 기준을 적용할 경우 네트워크 분석을 해야 하는 대상(분석 용어의 수)이 급격히 증가하므로 분석

의 어려움에 직면하게 된다. 본 연구에서는 연구 진행시 이러한 한계점이 있었다는 점을 네트워크 분석의 적용에 앞서 언급하고자 한다.

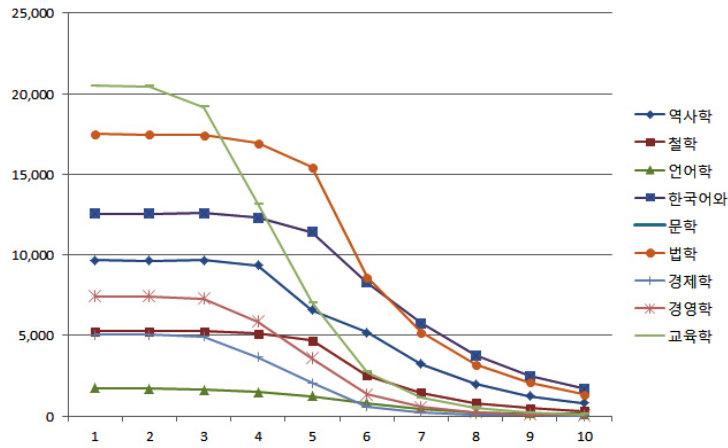
3. 연구 결과 분석

3.1 저자키워드의 출현순서와 출현빈도와의 상관관계

저자키워드의 출현순서와 저자키워드의 총 출현 빈도와 어떠한 관계가 있는지를 살펴보기 위하여 각 출현순서에 해당하는 키워드의 총 출현 빈도를 산출하였으며(〈표 8〉 참조) 저자키워드의 총 출현빈도와 출현순서에 대한 Spearman 상관분석을 수행하였다. 분석결과 인문학, 사회과학 분야 모두 강한 음(-)의 상관관계가 나타나며 0.01 수준에서 모두 유의한 상관관계가 나타나는 것을 볼 수 있었다(〈표 9〉 참조). 이는 출현순서의 앞에 위치하는 저자키워드일수록 출현빈도가 높은 것을 의미하며 출현순서의 뒤에 위치할수록 저자키워드의 출현빈도가 낮다는 것을 의미한다. 총 출현빈도의 변화를 살펴보면 대체로 4번째 출현순서까지는 큰 변화가 없다가 5번째 출현순서부터는 급격한 감소가 있는 것을 살펴볼 수 있다.

〈표 8〉 출현순서에 따른 분야별 빈도 결과

	1	2	3	4	5	6	7	8	9	10	합계
역사학	9,684	9,640	9,653	9,364	6,578	5,193	3,228	1,973	1,224	812	57,349
철학	5,269	5,268	5,263	5,128	4,671	2,532	1,450	818	492	303	31,194
언어학	1,722	1,706	1,620	1,471	1,228	824	435	249	159	127	9,541
한국어와 문학	12,536	12,534	12,568	12,287	11,421	8,278	5,740	3,770	2,482	1,692	83,308
법학	17,513	17,474	17,407	16,913	15,417	8,586	5,178	3,208	2,079	1,352	105,127
경제학	5,044	5,038	4,924	3,655	2,070	608	238	100	53	30	21,760
경영학	7,412	7,439	7,276	5,878	3,570	1,346	557	239	105	58	33,880
교육학	20,473	20,432	19,175	13,154	7,018	2,742	1,122	507	206	87	84,916



〈그림 4〉 출현 빈도 변화

〈표 9〉 저자키워드의 출현순서와 출현순서에 해당하는 키워드의 총 출현 빈도와의 상관관계

			출현빈도	출현순서
	역사학	출현빈도	Spearman 상관계수	1
	출현순서	Spearman 상관계수	-.988**	1
철학	출현빈도	Spearman 상관계수	1	-.100**
	출현순서	Spearman 상관계수	-.100**	1
언어학	출현빈도	Spearman 상관계수	1	-.100**
	출현순서	Spearman 상관계수	-.100**	1
한국어와 문학	출현빈도	Spearman 상관계수	1	-.964**
	출현순서	Spearman 상관계수	-.964**	1
법학	출현빈도	Spearman 상관계수	1	-.100**
	출현순서	Spearman 상관계수	-.100**	1
경제학	출현빈도	Spearman 상관계수	1	-.100**
	출현순서	Spearman 상관계수	-.100**	1
경영학	출현빈도	Spearman 상관계수	1	-.988**
	출현순서	Spearman 상관계수	-.988**	1
교육학	출현빈도	Spearman 상관계수	1	-.100**
	출현순서	Spearman 상관계수	-.100**	1

** p<.01

3.2 저자키워드의 출현순서와 연결정도 중심성과의 상관관계

저자키워드의 출현순서와 연결정도 중심성과의 상관관계를 살펴보기 위하여 Kendall의 타우 분석을 수행하였으며 그 결과는 <표 11>과 같다. 분석결과 인문학의 경우 '역사학', '철

학', '언어학', '한국어와 문학' 분야 모두 통계적으로 유의하였으며 양(+)의 상관관계로 나타났다. 사회과학의 경우 '법학', '경제학', '경영학' 분야에서 통계적으로 유의하였으며 양(+)의 상관관계로 나타났다. 하지만 '교육학' 분야의 경우 통계적으로 유의하지 않았다.

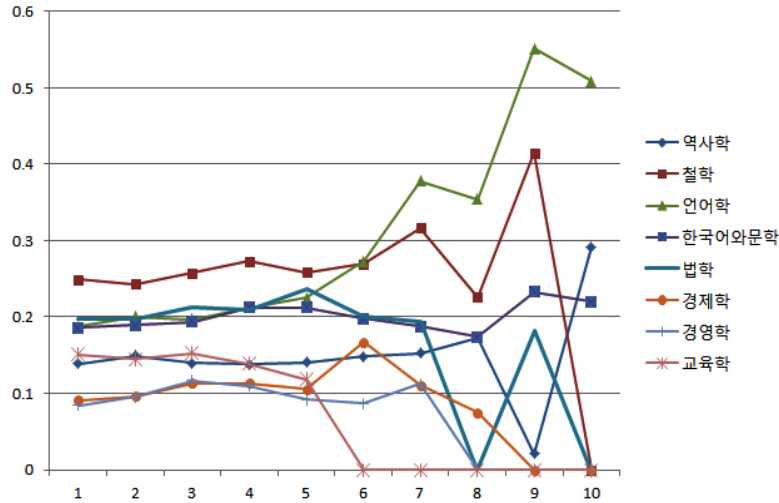
<표 10> 출현순서에 따른 분야별 연결정도 중심성 결과 평균

	1	2	3	4	5	6	7	8	9	10
역사학	0.13857	0.14909	0.13988	0.13818	0.14098	0.14824	0.15220	0.17283	0.02100	0.29200
철학	0.24935	0.24270	0.25744	0.27275	0.25867	0.26953	0.31620	0.22600	0.41450	-
언어학	0.18766	0.20019	0.19673	0.21272	0.22498	0.27309	0.37760	0.35400	0.55100	0.50833
한국어와 문학	0.18546	0.18920	0.19384	0.21225	0.21196	0.19811	0.18783	0.17355	0.23267	0.22000
법학	0.19795	0.19790	0.21250	0.20940	0.23748	0.20000	0.19460	-	0.18200	-
경제학	0.09094	0.09652	0.11285	0.11321	0.10560	0.16700	0.11100	0.07500	-	-
경영학	0.08453	0.09513	0.11642	0.10989	0.09188	0.0867	0.1125	-	-	-
교육학	0.15121	0.14504	0.1526	0.13922	0.11826	-	-	-	-	-

<표 11> 저자키워드의 출현순서와 출현순서에 해당하는 키워드의 연결정도 중심성과의 상관관계

			출현순서	연결정도중심성
역사학	출현순서	Kendall's tau_b	1	.043**
	연결정도중심성	Kendall's tau_b	.043**	1
철학	출현순서	Kendall's tau_b	1	.141**
	연결정도중심성	Kendall's tau_b	.141**	1
언어학	출현순서	Kendall's tau_b	1	.102**
	연결정도중심성	Kendall's tau_b	.102**	1
한국어와 문학	출현순서	Kendall's tau_b	1	.051**
	연결정도중심성	Kendall's tau_b	.051**	1
법학	출현순서	Kendall's tau_b	1	.064**
	연결정도중심성	Kendall's tau_b	.064**	1
경제학	출현순서	Kendall's tau_b	1	.110**
	연결정도중심성	Kendall's tau_b	.110**	1
경영학	출현순서	Kendall's tau_b	1	.103**
	연결정도중심성	Kendall's tau_b	.103**	1
교육학	출현순서	Kendall's tau_b	1	-.027
	연결정도중심성	Kendall's tau_b	-.027	1

* p<.05, ** p<.01



〈그림 5〉 각 출현순서에 따른 연결정도 중심성 평균의 변화

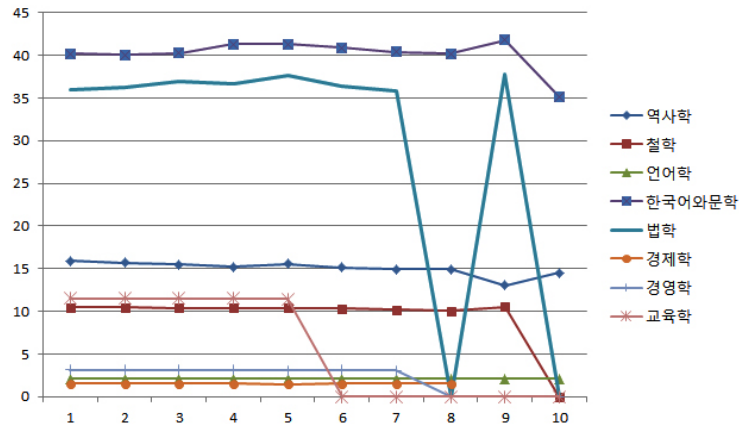
3.3 저자키워드의 출현순서와 근접 중심성과의 상관관계

저자키워드의 출현순서와 분야별 근접 중심성과의 상관관계 분석결과는 <표 13>과 같다. Kendall의 타우 분석을 수행하였으며 분석결과 인문학의 경우 ‘철학’, ‘언어학’, ‘한국어와 문

학’ 분야의 경우 통계적으로 유의한 것으로 나타났다. 양(+)의 상관관계로 나타났다. 하지만 역사학의 경우 통계적으로 유의하지 않았다. 사회과학의 경우 ‘법학’, ‘경제학’ 분야는 통계적으로 유의하였으며 양(+)의 상관관계로 나타났다. 반면 ‘경영학’, ‘교육학’ 분야는 통계적으로 유의하지 않았다.

〈표 12〉 출현순서에 따른 분야별 근접 중심성 결과

	1	2	3	4	5	6	7	8	9	10
역사학	15.90100	15.70800	15.45700	15.22300	15.53800	15.13400	14.93900	14.94200	13.02900	14.50700
철학	10.44610	10.42000	10.40900	10.40100	10.40010	10.31740	10.18900	10.04570	10.50950	-
언어학	2.03940	2.065800	2.05050	2.07030	2.08590	2.088400	2.10580	2.10525	2.09600	2.09967
한국어와 문학	40.21000	40.03500	40.27800	41.27600	41.31800	40.90800	40.42100	40.16660	41.76130	35.14100
법학	35.99690	36.30450	36.92780	36.61650	37.60180	36.32870	35.79660	-	37.81900	-
경제학	1.50223	1.52660	1.53817	1.52326	1.46283	1.52000	1.55000	1.50900		
경영학	3.08689	3.09188	3.09641	3.09231	3.07841	3.09240	3.08250	-	-	-
교육학	11.52880	11.52280	11.51560	11.51280	11.43530	-	-	-	-	-



〈그림 6〉 근접 중심성 변화

〈표 13〉 저자키워드의 출현순서와 근접 중심성과의 상관관계

			출현순서	근접중심성
역사학	출현순서	Kendall's tau_b	1	.008
	근접중심성	Kendall's tau_b	.008	1
철학	출현순서	Kendall's tau_b	1	.114**
	근접중심성	Kendall's tau_b	.114**	1
언어학	출현순서	Kendall's tau_b	1	.072**
	근접중심성	Kendall's tau_b	.072**	1
한국어와 문학	출현순서	Kendall's tau_b	1	.052*
	근접중심성	Kendall's tau_b	.052*	1
법학	출현순서	Kendall's tau_b	1	.100**
	근접중심성	Kendall's tau_b	.100**	1
경제학	출현순서	Kendall's tau_b	1	.050**
	근접중심성	Kendall's tau_b	.050**	1
경영학	출현순서	Kendall's tau_b	1	.036
	근접중심성	Kendall's tau_b	.036	1
교육학	출현순서	Kendall's tau_b	1	-.020
	근접중심성	Kendall's tau_b	-.020	1

* p<.05, ** p<.01

3.4 저자키워드의 출현순서와 매개 중심성과의 상관관계

저자키워드의 출현순서와 매개 중심성과의 상관관계 분석결과는 〈표 15〉와 같다. 저자키

워드의 출현순서와 매개 중심성과의 상관관계를 살펴보기 위하여 Kendall의 타우 분석을 수행하였으며 분석결과 인문학의 경우 '철학', '언어학', '한국어와 문학' 분야의 경우 통계적으로 유의한 것으로 나타났으며 음(-)의 상관관계

로 나타났다. 하지만 역사학의 경우 통계적으로 유의하지 않았다. 사회과학의 경우 '법학', '경제학', '경영학' 분야는 통계적으로 유의하였으며 양(+)의 상관관계로 나타났다. 반면 '교육학' 분야는 통계적으로 유의하지 않았다.

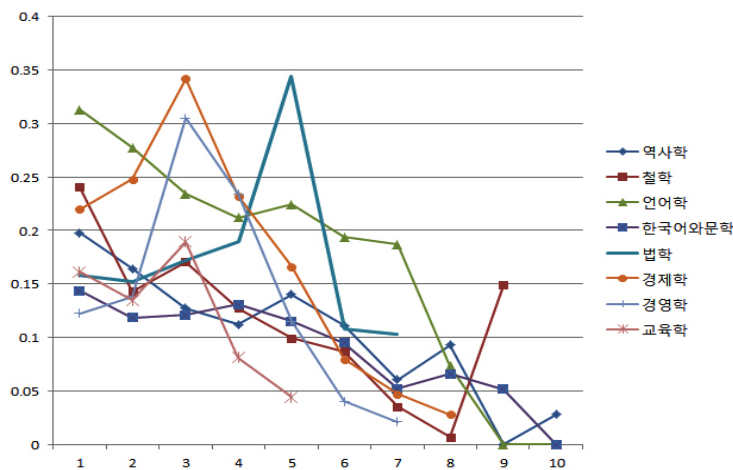
눈여겨보아야 할 점은 앞서 연결정도 중심성과 근접중심성 결과와는 달리 매개 중심성 결과에서는 인문학의 경우는 음(-)의 상관관계, 사회과학의 경우 양(+)의 상관관계가 나타났다는 점이다.

사회과학 분야의 매개 중심성 결과를 살펴보면 모든 분야에서 3번째 순서에서 값이 가장 높

은 것으로 나타난 점은 매우 특이할 만한 점이다. 이는 다른 출현순서에 비해 3번째 순서에 매개 중심성 값이 높은 용어가 많다는 것을 의미한다. 이러한 용어를 살펴보면 경제학 분야의 경우 '경제성장'(10,194), '고용(4,261)'처럼 경제학 분야의 매개중심성 상위에 위치하는 용어가 세 번째 순서에 있는 것을 볼 수 있다. '경영학' 분야의 경우 '신뢰'(11,298), '고객만족'(6,734)과 같은 용어를 볼 수 있으며 교육학 분야의 경우 또한 '자기효능감'(4,745), '질적연구'(3,478)와 같은 용어가 이에 해당된다.

〈표 14〉 출현순서에 따른 분야별 매개 중심성 결과

	1	2	3	4	5	6	7	8	9	10
역사학	0.19788	0.16402	0.12744	0.11193	0.14005	0.11122	0.06033	0.09267	0	0.02800
철학	0.24066	0.14291	0.17062	0.12729	0.09948	0.08705	0.03520	0.00675	0.14900	-
언어학	0.31287	0.27768	0.23416	0.21216	0.22445	0.19361	0.18700	0.07325	0	0
한국어와 문학	0.14344	0.11854	0.12104	0.13057	0.11503	0.09507	0.05192	0.06564	0.05133	0
법학	0.15815	0.15232	0.17169	0.18943	0.34459	0.10758	0.10260		0.06700	
경제학	0.21987	0.24807	0.34212	0.23201	0.16598	0.07950	0.04700	0.02800	-	-
경영학	0.12234	0.13815	0.30512	0.23432	0.11494	3.09240	3.08250	-	-	-
교육학	0.16149	0.13449	0.18929	0.08112	0.04411	-	-	-	-	-



〈그림 7〉 매개중심성 변화

〈표 15〉 저자키워드의 출현순서와 매개 중심성과의 상관관계

		출현순서		매개중심성
역사학	출현순서	Kendall's tau_b	1	.002
	매개중심성	Kendall's tau_b	.002	1
철학	출현순서	Kendall's tau_b	1	-.097**
	매개중심성	Kendall's tau_b	-.097**	1
언어학	출현순서	Kendall's tau_b	1	-.075**
	매개중심성	Kendall's tau_b	-.075**	1
한국어와 문학	출현순서	Kendall's tau_b	1	-.032**
	매개중심성	Kendall's tau_b	-.032**	1
법학	출현순서	Kendall's tau_b	1	.077**
	매개중심성	Kendall's tau_b	.077**	1
경제학	출현순서	Kendall's tau_b	1	.075**
	매개중심성	Kendall's tau_b	.075**	1
경영학	출현순서	Kendall's tau_b	1	.075**
	매개중심성	Kendall's tau_b	.075**	1
교육학	출현순서	Kendall's tau_b	1	-.029
	매개중심성	Kendall's tau_b	-.029	1

* p<.05, ** p<.01

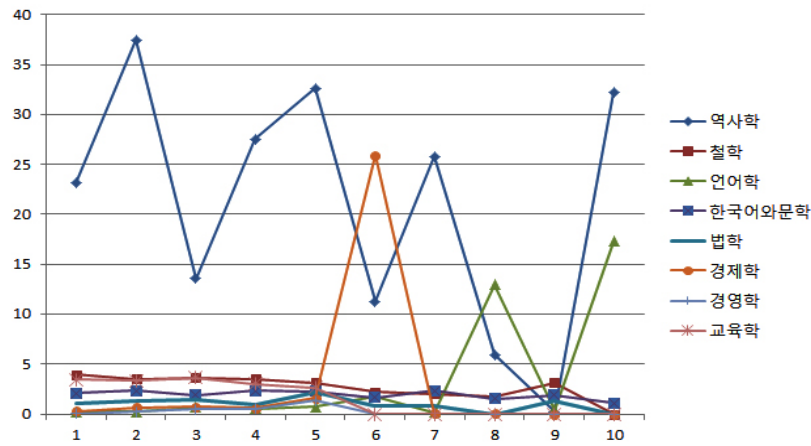
3.5 저자키워드의 출현순서와 위세 중심성과의 상관관계

저자키워드의 출현순서와 위세 중심성과의 상관관계 분석결과는 〈표 17〉과 같다. 저자키워드의 출현순서와 위세 중심성과의 상관관계를 살펴보기 위하여 Kendall의 타우 분석을 수행하였으며 분석결과 인문학의 경우 '철학', '언

어학' 분야의 경우 통계적으로 유의한 것으로 나타났으며 양(+)의 상관관계로 나타났다. 하지만 '역사학', '한국어와 문학'의 경우 통계적으로 유의하지 않았다. 사회과학의 경우 '경제학', '경영학' 분야는 통계적으로 유의하였으며 양(+)의 상관관계로 나타났다. 반면 '법학', '교육학' 분야는 통계적으로 유의하지 않았다.

〈표 16〉 출현순서에 따른 분야별 위세 중심성 결과

	1	2	3	4	5	6	7	8	9	10
역사학	23.1430	37.4850	13.5720	27.5040	32.6410	11.2800	25.7690	5.91200	0.22800	32.28000
철학	3.98132	3.49120	3.61916	3.53694	3.12265	2.20711	2.01520	1.69075	3.13750	-
언어학	0.25459	0.26004	0.65269	0.53586	0.72293	1.72871	0.17888	12.9203	0.39800	17.28470
한국어와 문학	2.07229	2.32429	1.91465	2.32224	2.27791	1.62779	2.31671	1.52691	1.87867	1.11700
법학	1.04879	1.29874	1.41911	0.92692	2.14511	0.82600	0.79140	-	1.27600	-
경제학	0.29262	0.64315	0.73596	0.62113	1.62340	25.92500	0.06800	0.03100	-	-
경영학	0.03168	0.25392	0.46804	0.54693	1.32283	0.02780	0.01400	-	-	-
교육학	3.46482	3.32470	3.64683	3.04968	2.62937	-	-	-	-	-



〈그림 8〉 위세중심성 변화

〈표 17〉 저자키워드의 출현순서와 위세 중심성과의 상관관계

			출현순서	위세중심성
역사학	출현순서	Kendall's tau_b	1	.039
	위세중심성	Kendall's tau_b	.039	1
철학	출현순서	Kendall's tau_b	1	.157**
	위세중심성	Kendall's tau_b	.157**	1
언어학	출현순서	Kendall's tau_b	1	.140**
	위세중심성	Kendall's tau_b	.140**	1
한국어와 문학	출현순서	Kendall's tau_b	1	-.003
	위세중심성	Kendall's tau_b	-.003	1
법학	출현순서	Kendall's tau_b	1	.029
	위세중심성	Kendall's tau_b	.029	1
경제학	출현순서	Kendall's tau_b	1	.089**
	위세중심성	Kendall's tau_b	.089**	1
경영학	출현순서	Kendall's tau_b	1	.064*
	위세중심성	Kendall's tau_b	.064*	1
교육학	출현순서	Kendall's tau_b	1	-.021
	위세중심성	Kendall's tau_b	-.021	1

* p<.05, ** p<.01

3.6 저자키워드의 출현순서와 구조적공백성의 효과크기와의 상관관계

저자키워드의 출현순서와 구조적 공백성의 효과크기와의 상관관계 분석결과는 〈표 19〉와 같다. 저자키워드의 출현순서와 구조적 공백성의 효과크기와의 상관관계를 살펴보기 위하여

Kendall의 타우 분석을 수행하였으며 분석결과 인문학의 경우 모든 분야에서 통계적으로 유의하지 않은 것으로 나타났다. 사회과학의 경우 '법학', '경영학' 분야는 통계적으로 유의하였으며 양(+)의 상관관계로 나타났다. 반면 '경제학', '교육학' 분야는 통계적으로 유의하지 않았다.

〈표 18〉 출현순서에 따른 분야별 구조적 공백성의 효과크기 정도 결과

	1	2	3	4	5	6	7	8	9	10
역사학	11.89850	11.79520	10.60060	10.08240	10.74770	10.19180	9.238600	11.67900	1.62200	13.13800
철학	24.63590	19.52850	20.74230	19.66860	17.46250	13.66950	10.82000	5.72450	19.75600	-
언어학	5.45656	5.27795	4.97545	5.18715	5.14164	4.64900	6.01460	4.91970	2.29200	3.15300
한국어와문학	30.65430	29.52650	30.64650	33.57280	32.06050	31.01500	24.67700	25.71600	32.06400	2.61600
법학	17.77240	17.86060	19.7428	19.67990	26.19370	17.48000	16.16700	-	16.93400	-
경제학	4.78066	5.03002	6.26932	5.498160	4.09886	2.42950	3.22150	2.12100	-	-
경영학	7.38178	8.45026	12.14200	11.54350	7.35779	5.90660	6.48500	-	-	-
교육학	19.14120	17.16040	19.62340	15.41450	11.87410	-	-	-	-	-

〈표 19〉 저자키워드의 출현순서와 효과크기와의 상관관계

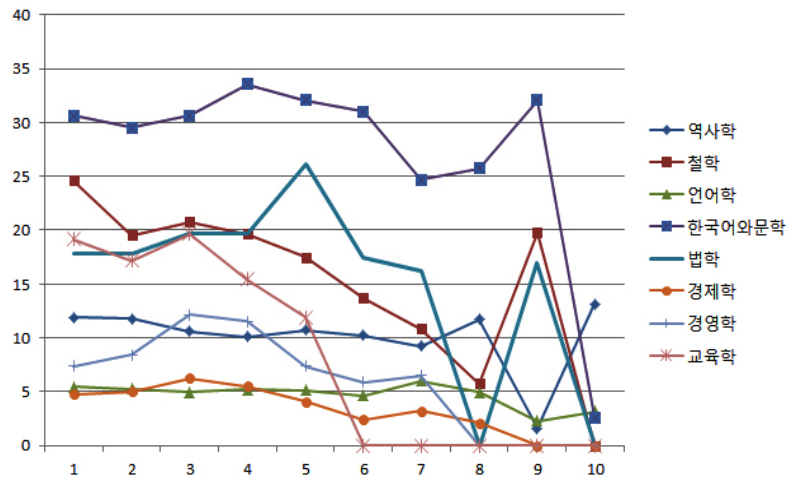
			출현순서	효과크기
역사학	출현순서	Kendall's tau_b	1	-.051
	효과크기	Kendall's tau_b	-.051	1
철학	출현순서	Kendall's tau_b	1	.043
	효과크기	Kendall's tau_b	.043	1
언어학	출현순서	Kendall's tau_b	1	-.026
	효과크기	Kendall's tau_b	-.026	1
한국어와 문학	출현순서	Kendall's tau_b	1	-.003
	효과크기	Kendall's tau_b	-.003	1
법학	출현순서	Kendall's tau_b	1	.104**
	효과크기	Kendall's tau_b	.104**	1
경제학	출현순서	Kendall's tau_b	1	.040
	효과크기	Kendall's tau_b	.040	1
경영학	출현순서	Kendall's tau_b	1	.096**
	효과크기	Kendall's tau_b	.096**	1
교육학	출현순서	Kendall's tau_b	1	-.040
	효과크기	Kendall's tau_b	-.040	1

* p<.05, ** p<.01

〈표 20〉 저자키워드의 출현순서와 네트워크 중심성과의 상관관계 결과

		분야	연결정도중심성	근접중심성	매개중심성	위세중심성	효과크기	출현빈도
출현 순서	역사학	.043**	.008**	.002	.039	-.051	-.100**	
	철학	.141**	.114**	-.097**	.157**	.043	-.964**	
	언어학	.102**	.072**	-.075**	.140**	-.026	-.988**	
	한국어와문학	.051**	.052*	-.032**	-.003	-.003	-.100**	
	법학	.064**	.100**	.077**	.029	.104**	-.989**	
	경제학	.110**	.050**	.075**	.089**	.040	-.976**	
	경영학	.103**	.036	.075**	.064*	.096**	-.920**	
교육학	-.027	-0.02	-.029	-.021	-.040	-.952**		

* p<.05, ** p<.01



〈그림 9〉 효과크기 변화

3.7 저자키워드의 출현순서와 네트워크 중심성과의 상관관계 결과 종합

저자키워드의 출현순서와 각각의 네트워크 중심성과의 상관관계 결과를 종합해 보면 <표 20>과 같다. 각 결과에서 유의한 분야를 살펴보면 연결정도중심성은 7개 분야, 근접중심성은 5개 분야, 매개중심성은 6개 분야, 위세중심성은 4개 분야, 효과 크기는 2개 분야인 것을 알 수 있었다. 또한 연결정도중심성, 근접중심성, 매개중심성, 효과크기의 경우 유의한 상관관계 결과는 모두 양(+)의 관계였지만, 매개중심성의 경우 인문학분야의 경우는 음(-)의 상관관계, 사회과학 분야는 양(+)의 상관관계를 나타냈다는 점은 특이할 만한 점이다.

상관계수를 살펴보면 모든 분야의 결과가 ±0.0에서 0.2 사이로 그 관계가 매우 미미하다는 것을 뜻한다.

다음으로 저자키워드의 출현순서와 각각의 네트워크 중심성과의 상관관계 결과와 앞서 2.3절

에서 분석한 저자키워드의 출현순서에 따른 출현빈도의 상관관계 결과와 비교해보았는데, 매개중심성 결과를 제외하고 상반된 결과를 나타낸다. 출현빈도만을 통해 용어의 중요도를 판단하는 것도 물론 의미가 있다. 하지만 이는 사실상 매우 단순한 산술적인 수치에 불과할 수 있으며 특정 분야 네트워크 내에서의 위치를 통한 중요도를 파악할 수는 없다. 이를 위해서는 네트워크 중심성의 결과를 통해 살펴볼 필요성이 있음을 뜻한다. 다시 말해 용어의 출현순서가 뒤로 갈수록 출현빈도가 현저히 줄어든다고 하여 그 용어가 네트워크 내에서의 구조적 위치가 중심에 위치하거나 변방에 위치한다고 볼 수 있는 등의 판단은 어렵다는 것이다.

4. 결론 및 제언

본 연구에서는 학술논문에 부여된 복수의 저자키워드는 단순한 나열이 아니라 중요도 값에

따라 나열되어 있을 것임을 가정하고 실제로 그러한 특성이 나타나는지를 실증적으로 검증하고자 하였다. 저자키워드의 중요도를 측정하기 위해서는 측정할 수 있는 객관적 지표가 필요하며 이를 위해 네트워크 분석 방법의 연결정도중심성, 근접중심성, 매개중심성, 위세중심성, 그리고 네트워크의 구조적공백성의 효과크기와 같은 지수를 사용하여 출현순서에 따른 저자키워드의 중요도를 측정해보았다. 연구의 결과를 요약해보면 다음과 같다.

첫째, 출현순서에 따른 저자키워드의 특징을 살펴본바 인문학 4개 분야 사회과학 4개 분야의 출현순서에 해당하는 저자키워드의 수는 출현순서의 앞에서 뒤로 갈수록 저자키워드의 수는 감소하는 것을 볼 수 있었다. 하지만 역사, 철학, 한국어와 문학, 법학과 같은 분야는 2번째 출현순서에 해당하는 저자키워드 수가 더 많았으며 분야마다 약간의 차이는 있으나 대체로 4번째 출현 순서까지는 용어의 감소폭이 그렇게 크지 않다가 이후 출현순서에서는 급속히 감소하는 것을 볼 수 있었다.

두 번째로, 출현순서에 따른 저자키워드의 중요도를 측정하기 위한 네트워크 지수 값은 분야별로 어떠한 특징을 지니고 있으며 출현순서와 각각의 네트워크 지수는 상관관계가 있는지 살펴보았다. 이를 분석하기 위해 연결정도중심성, 근접중심성, 매개중심성, 위세중심성과 구조적 공백성 지수의 효과크기와 같은 네트워크 중심성 지수를 측정 후 저자키워드의 출현순서와 상관관계분석을 수행하였다.

셋째, 네트워크 분석 지수 중 근접중심성, 위세중심성과 구조적 공백성 지수의 효과크기의 경우 일부 학문 분야에서 상관관계를 나타나기

는 하였지만 이들에 비해 비교적 저자키워드의 중요도를 잘 설명하고 있는 지수는 연결정도중심성과 매개중심성이었다. 연결정도중심성은 '교육학' 분야를 제외한 7개 분야 모두 양(+)의 상관관계를 나타냈으며 매개중심성의 경우 '역사학', '교육학' 분야를 제외한 6개 분야에서 인문학 분야는 음(-), 사회과학 분야는 양(+)의 상관관계를 나타냈다.

본 연구는 출현순서에 따른 저자키워드의 중요성을 네트워크 분석방법을 통해 실증적으로 분석하였다. 저자키워드의 출현순서에 따른 연결정도중심성, 매개중심성 값은 네트워크상에서의 용어 간의 구조적인 위치를 의미하고 이는 용어의 '중요도'와 직결되는 문제라고 할 수 있다. 다만 본 연구는 다음과 같은 한계점을 지닌다. 저자키워드의 중요도를 학문 분야 네트워크의 위치로 판단 해볼 수 있지만 그렇다고 해서 이 중요도가 반드시 논문 내 저자키워드의 중요도를 의미하지 않는다는 것이다. 만약 주제 분야의 범위를 좁혀 학문 분야 내에서도 특정 주제 분야에 속한 논문을 선정하고 그 주제의 네트워크 분석 수행을 고려해본다면 논문 내 저자키워드의 중요도와도 연관성이 높을 가능성이 있다. 다음으로, 저자키워드의 출현순서를 산출하는 데 있어 최빈값을 통해 정의하였는데, 이외에도 다른 방법을 통해 순서를 산출해 보는 것을 고려해 볼 필요성이 있다.

본 연구는 기존의 연구가 저자키워드의 중요도를 단지 출현빈도만으로 판단했던 것에서 벗어나 네트워크상에서의 위상관계로 측정하였다는 점, 그리고 출현순서라는 측면을 고려하였던 점에 의의가 있다고 할 수 있다. 다만 양적으로 분석을 시도하는 것만으로 한계가 있으며

로 실제 연구자들의 저자키워드를 부여할 때의 상황적 요인, 행동, 심리 등을 파악하여 보완하는 연구 및 최근 가중 네트워크 분석과 같은 연구가 이루어지는 만큼, 가중 네트워크 분석방법을 적용해 비교 해보는 후속 연구 또한 이루어질 필요성이 있다.

알림(Acknowledgments)

연구결과의 질을 향상시킬 수 있도록 세심한 의견과 제안을 주신 익명의 논문심사자에게 깊은 감사를 드립니다.

참 고 문 헌

- 권선영 (2013). 구조적 학술용어사전 데이터베이스 구축에 있어서 용어의 의미관계 형성에 영향을 미치는 요인에 관한 연구. 박사학위논문, 성균관대학교 대학원, 문헌정보학과.
- 강범일, 박지홍 (2013). 프로파일링 분석과 동시출현단어 분석을 이용한 한국어교육학의 정체성 분석. 정보관리학회지, 30(4), 195-213. <http://dx.doi.org/10.3743/KOSIM.2013.30.4.195>
- 김관준, 이재윤 (2012). 디스크립터 자동 할당을 위한 저자키워드의 재분류에 관한 실험적 연구. 정보관리학회지, 29(2), 225-246. <http://dx.doi.org/10.3743/KOSIM.2012.29.2.225>
- 문혜원 (2000). 한국 의학학술논문의 저자선정 주제어와 MeSH 용어의 비교 분석 연구. 석사학위논문, 숙명여자대학교 대학원, 문헌정보학과.
- 박수현, 박경영 (2011). 대한작업치료학회지 논문의 영문 주제어와 MeSH 용어의 비교분석. 대한작업치료학회지, 19(4), 131-146.
- 박옥남 (2011). 정보조직 지식구조에 대한 연구: 2000년~2011년 학술논문을 중심으로. 한국비블리아학회지, 22(3), 247-267.
- 서선경, 정은경 (2013). 단어동시출현 분석 기반 오픈 액세스 분야 지적구조에 관한 연구. 한국비블리아학회지, 24(1), 207-228.
- 양창진 (2010). 학술 논문의 주제어 표기 및 활용 방안 연구: DB 구축 및 정보연계의 관점에서. 인문콘텐츠, 19, 395-416.
- 이용덕 (2006). 일본학관련 전문학술지에 나타난 주제어 분석: 日本學報를 중심으로. 일본어문학, 31, 107-121.
- 이혜영, 박승진 (2011). 국내 학술지 논문의 주제어를 통한 학술연구분야 관계분석. 한국비블리아학회지, 22(3), 353-371.
- 장령령, 홍현진 (2014). 학술지 중요도와 키워드 순서를 고려한 단어동시출현 분석을 이용한 독서분야의 지적구조 분석. 한국비블리아학회지, 25(1), 295-318. <http://dx.doi.org/10.14699/kbiblia.2014.25.1.295>

조재인 (2011). 네트워크 텍스트 분석을 통한 문헌정보학 최근 연구 경향 분석. 정보관리학회지, 28(4), 65-83. <http://dx.doi.org/10.3743/KOSIM.2011.28.4.065>

Taghva, K., Borsack, J., Nartker, T., & Condit, A. (2004). The role of manually assigned keywords in query expansion. *Information Processing & Management*, 40, 441-458.

Gross, T., & Taylor, A. G. (2005). What have we got to lose? The effect of controlled vocabulary on keyword searching results. *College & Research Libraries*, 66(3), 212-230.

<p>• 국문 참고문헌에 대한 영문 표기 (English translation of references written in Korean)</p>

Cho, Jane (2011). A study for research area of library and information science by network text analysis. *Journal of the Korean Society for Information Management*, 28(4), 65-83. <http://dx.doi.org/10.3743/KOSIM.2011.28.4.065>

Kwon, Sunyoung (2013). A study on the factors influencing semantic relation in building a structured glossary department of library and information science. Unpublished doctoral dissertation, Sungkyunkwan University, Seoul, Korea.

Kang, Beo Mil, & Park, Ji Hong (2013). Profiling and co-word analysis of teaching korean as a foreign language domain. *Journal of the Korean Society for Information Management*, 30(4), 195-213. <http://dx.doi.org/10.3743/KOSIM.2013.30.4.195>

Kim, Pan Jun, & Lee, Jae Yun (2012). A study on the reclassification of author keywords for automatic assignment of descriptors. *Journal of the Korean Society for Information Management*, 29(2), 225-246. <http://dx.doi.org/10.3743/KOSIM.2012.29.2.225>

Lee, young dok (2006). An analysis of the key words in korean journal of japanology in Korea. *JALALIDA*, 31, 107-121.

Lee, Hye-Young, & Kwak, Seung-Jin (2011). Relation analysis among academic research areas using subject terms of domestic journal papers. *Journal of the Korea Biblia Society for Library and Information Science*, 22(3), 353-371.

Mun, Hey won (2000). Subject headings of korean medical journal papers: Author keywords vs. MeSH. Unpublished master's thesis, sookmyung Women's University, Seoul, Korea.

Park, Su-Hyun, & Park, Kyoung-Young (2011). Coincidence analysis of key words and MeSH terms in the Journal of the Korean Society of Occupational Therapy. *The Journal of Korean Society of Occupational Therapy*, 19(4), 131-146.

Park, Ok Nam (2011). Knowledge structures in Knowledge Organization Research: 2000-2011.

- Journal of the Korean Biblia Society for Library and Information Science, 22(3), 247-267.
- Seo, Sun Kyung, & Chung, Eun Kyung (2013). Domain analysis on the field of Open Access by co-word analysis. Journal of the Korean Biblia Society for Library and Information Science, 24(1), 207-228.
- Yang, Chang-Jin (2010). Study on keywords and their use of academic theses: Focused on database development and information link. Human Contents, 19, 395-416.
- Zhang, Ling Ling, & Hong, Hyun Jin (2014). Examining the intellectual structure of reading studies with co-word analysis based on the importance of journals and sequence of keywords. Journal of the Korean Biblia Society for Library and Information Science, 25(1), 295-318. <http://dx.doi.org/10.14699/kbiblia.2014.25.1.295>