

컴퓨터 문헌 분석 기법을 활용한 <적벽가> 이본의 계통 분류 연구

A Study on the Classification of *Jeokbyeok-ga's* Version by the Computer Analysis Technique of Bibliographies

이진오*, 김동건**

금오공과대학교 교양교직과정부*, 경희대학교 후마니타스칼리지**

Jin-O Lee(dltkd00@hanmail.net)*, Dong-Keon Kim(dehi@khu.ac.kr)**

요약

본 연구는 컴퓨터 문헌 분석 기법을 활용하여 <적벽가> 전체 이본의 계통을 파악하는 한편, 기존 이본론의 연구 성과를 검토하는 데에 목적을 둔다. 우선 분석의 기초자료를 마련하기 위해 <적벽가>의 이본 중 46종을 선정하고 이를 대상으로 원시 코퍼스를 구축하였다. 이를 통해서 <적벽가>의 공통 서사단위는 5개의 계층으로 파악할 수 있었으며, 146개의 개별 내용 단락을 추출할 수 있었다. 다음으로 인코딩 된 코퍼스를 바탕으로 이본간의 유사도와 거리 측정을 시도하였는데, <적벽가>의 계통과 이본간의 거리를 시각적으로 제시할 수 있었다. 이후 다차원 척도법, 계층적 군집 분석, 계통의 분기 분석 기법을 활용하여 이본군의 분포를 확인할 수 있었다. 이처럼 컴퓨터 문헌 분석 기법을 종합하여 적용해보면 <적벽가>의 이본군은 완판 계열과 창본 계열로 양분되어 있음을 확인할 수 있는데, 기존의 논의보다 더 세밀한 계통 구분을 할 수 있었다. 또한 판소리의 유파와 전승에 따른 영향 관계도 검토할 수 있었다.

■ 중심어 : | 적벽가(赤壁歌) | 이본(異本) | 컴퓨터 문헌 분석 | 다차원 척도법 | 계층적 군집 분석 | 분기 분석 |

Abstract

The purpose of this study is to examine the system of the Jeokbyeok-ga's version using the Computer analysis technique of bibliographies and to examine the achievements of the Jeokbyeok-ga's version studies. First, in order to provide basic data for analysis, a raw corpus was constructed for 46 species of Jeokbyeok-ga. Through this, the common narrative units of the Jeokbyeok-ga were identified as 5 layers, and thus 146 individual paragraphs could be extracted. Based on the encoded corpus, we tried to measure the similarity and the distance between the two. Next, we applied the Multidimensional scaling method, Hierarchical cluster analysis and Cladistic analysis method of the system to confirm the distribution of versions group and it was possible to visually grasp the distance between versions and the system of the work. As a result of analyzing Computer analysis technique of bibliographies, it was found that version's group of the Jeokbyeok-ga was divided into a Wanpan(完板) series and Changbon(唱本) series. Also, it was possible to examine the influence relationship between the Pansori's traditions and transmission

■ keyword : | Jeokbyeok-ga | Versions | Computer Analysis Technique of Bibliographies | Multidimensional Scaling | Hierarchical Clustering | Cladistic Analysis |

I. 서론

〈적벽가〉는 전래의 『삼국지연의』를 기반으로 조선 후기에 등장한 작품군을 통칭한다. 이 작품은 소설과 판소리로 향유되면서 조선 후기부터 근대에 이르기까지 다양한 형태의 이본을 남겼다.

〈적벽가〉의 초기 논의에서는 이 작품이 여타의 판소리 작품과 달리 소설본이 먼저 형성했다는 주장이 주를 이루었으나[1][2], 이후 간략한 내용의 판소리가 성립되고 나서 『삼국지연의』의 내용이 수용되었다라는 견해가 일반화 되었다. 이러한 논의의 흐름에 따라 〈적벽가〉 작품군의 성립과 이본 계통을 고찰한 연구도 꾸준히 진행되어 왔다. 〈적벽가〉와 관련하여 진행된 이본 연구는 대체로 특정 단락의 유무, 공통 단락의 유형, 작품의 각 이본군의 성격을 규명하는 것으로 요약할 수 있다[3-5]. 이러한 연구들이 〈적벽가〉 연구에 상당한 성과를 가져왔음은 부인할 수 없는 사실이다. 하지만 작품의 계열을 파악하는 데에 연구자의 직관에 의존한 면이 없지 않기 때문에 기준을 어디에 두느냐에 따라 다소 다른 의견이 제출되기도 했다. 따라서 보다 객관적인 근거를 통해 작품의 이본 계통을 살필 필요성이 제기된다고 하겠다.

한편 컴퓨터의 텍스트 분석 기법이 국내에 소개되면서 이공계뿐 아니라 인문학 분야에도 이를 적용한 연구가 활발히 진행되고 있다. 근래 인문학 분야에서 관련 연구로는 셰익스피어의 〈로미오와 줄리엣〉을 코퍼스 언어학으로 분석한 논의[6], 근 20년간의 영어교육 관련 석박사 학위논문과 연구논문의 경향성을 코퍼스 프로그램으로 파악한 시도[7], 승례문 방화 사건을 전후하여 언론사별 기사를 분석한 연구[8] 등이 주목되었다. 한편 판소리 작품군을 대상으로 컴퓨터 문헌 분석 기법으로 분석한 연구성과도 제출된 바 있다[9][10].

본 연구에서는 이러한 컴퓨터 문헌 분석 기법을 활용하여 〈적벽가〉 이본군의 분포를 분석하고 각 작품군의 성격을 규명하고자 한다. 또한 기존 연구를 검증하는 작업과 더불어 〈적벽가〉의 이본 현황에 대한 보다 객관적인 근거를 제공하고자 한다.

II. 자료의 유형 분석과 유사도 측정

1. 자료의 분류와 유형 인코딩

본 연구에서는 『적벽가 전집』 1~7권(박이정 출판사 발행)에 수록된 작품의 텍스트를 기초 자료로 삼는다 [11]. 『적벽가 전집』에는 총 54종의 이본이 실려 있는데, 이중에 작품의 계통 분류에서 통계상의 의미가 뚜렷하게 드러나지 않는 작품과 동종본은 분석 대상에서 제외하였다.

최종적으로 〈적벽가〉의 계통과 특징을 뚜렷이 보여줄 수 있는 이본 46종을 연구대상으로 선정하였다. 연구 대상이 되는 작품을 식별기호를 부여하여 정리하면 아래와 같다.

표 1. 연구 대상 이본과 식별기호(ID) 목록

	이본명	ID
1	신재호본 적벽가	CSJH
2	이선유 창본 적벽가	CLSU
3	임방울 창본 적벽가	CLBU
4	정광수 창본 적벽가	CJKS
5	김연수 창본 적벽가	CKYS
6	한승호 창본 적벽가	CHSH
7	정권진 창본 적벽가	CJKJ
8	박동진 창본 적벽가	CPDJ
9	송순섭 창본 적벽가	CSSS
10	박봉술 창본 적벽가	CPBS
11	원판 화흥도 83장본 정미구동산간본	EJUN01
12	원판 화흥도 84장본 무신서계개각본	EJUN02
13	원판 화흥도 84장본 무신양채방본	EJUN03
14	박순호소장 89장본 화흥도	SPSH08
15	박순호소장 42장본 화흥도	SPSH03
16	박순호소장 57장본 화흥도가	SPSH04
17	박순호소장 81장본 화흥도	SPSH05
18	박순호소장 59장본 화흥도	SPSH06
19	박순호소장 46장본 화흥도	SPSH07
20	김종철소장 40장본 화흥도	SKJC01
21	김종철소장 44장본 적벽가화흥도	SKJC02
22	김종철소장 75장본 화흥도	SKJC03
23	국립도서관소장 40장본 화흥도	SNAL01
24	국립도서관소장 19장본 적벽가	SNAL02
25	사재동소장 35장본 적벽대전	SSJD02
26	한중연소장 낙장88장본 화흥도	SHJY01
27	한중연소장 94장본 화흥도	SHJY02
28	김광순소장 23장본 화흥도	SKKS02
29	배연형소장 29장본 적벽가	SBYH01
30	배연형소장 45장본 적벽가	SBYH02
31	조동일소장 63장본 화흥도	SJDI

32	홍윤표소장 65장본 적벽가	SHYP
33	김동욱소장 24장본 화룡도전	SKDU
34	단국대소장 78장본 화용도	SDGL01
35	단국대소장 낙장82장본 화용도	SDGL02
36	단국대소장 81장본 화룡도	SDGL03
37	단국대소장 68장본 화룡도	SDGL04
38	경북대도서관소장 13장본 화용도가	SKBL
39	송실대도서관소장 낙장57장본 화룡도	SSSL
40	영남대도서관소장 37장본 화룡도연산별곡	SYNL
41	연세대도서관소장 낙장27장본 적벽가	SYSL01
42	연세대도서관소장 29장본 화룡도전	SYSL02
43	단국대소장 39장본 화룡도	SDGL05
44	단국대소장 21장본 화룡도	SDGL06
45	단국대소장 27장본 적벽전	SDGL07
46	유일서관본 적벽가	PYIB

분석 대상인 <적벽가>의 텍스트는 [표 1]에서 확인할 수 있는 것처럼 작품 형태에 따라 C(판소리 창본), P(활자본), S(필사본), E(판각본)의 식별 ID를 부여하였다. 그리고 소장자 또는 소장처에 따라 알파벳 약자를 결합하였다. 활자본의 경우는 출판사의 영어 약자를 식별기호로 부여하는 방식을 취했다.

컴퓨터 문헌 분석 방법론의 기반을 마련하기 위해서는 <적벽가>의 내용 단락과 단락별 유형을 반영한 전체 이본의 데이터화 작업을 수행할 필요가 있다. 이러한 기초 자료를 마련하기 위하여 작품의 서사단락 구분과 단락을 유형별로 인코딩 하는 기반 작업을 진행하였으며, 작품 내에서 변별되는 단락은 추출하여 원시 코퍼스(raw corpus)로 작성하였다.

ID	CSJH	TITLE	신재호 40장본 적벽가
<1-앞>赤壁歌 天下大勢 分久必合 合久必分이 聖嘆先生の 万古 確論이라 漢靈帝 建寧 二年 四月 望日 溫德殿에 전좌하여 百官朝會 바드실 欣난 뒤업는푸른 비암 標上으로 괴어나려 御榻을 두르더니 因忽不見 간 닥업시 雷聲大雨 雨霰호고 其後四年 二月日에 洛 陽에 地振하여 海水가 넘찌르고 其後 光和元년에 암담이 슈닥되며 六月에 거문 기운 七月에 무지 게오 五原山이 문어지니 잇되 天下 紛紛 하여 四方 兵이 일어날 제 黃巾賊도 어렵거든 十七陣 웬일 인고...			

그림 1. <적벽가> '정세 제시와 인물소개' 단락의 원시 코퍼스

[그림 1]은 CSJH(신재호 40장본 적벽가)의 '정세 제시와 인물소개' 단락을 원시 코퍼스로 작성한 것이다. 이

원시 코퍼스를 통해서 신재호본의 '정세 제시와 인물소개'의 단락이 어떻게 구성되는지 파악할 수 있다. 이와 같은 방식으로 <적벽가> 46종의 모든 이본을 내용 단락으로 분류하였고, 각각의 원시 코퍼스를 작성하였다.

2. 내용 단락 분석과 유형 분류

앞에서 예로 들었던 <적벽가>의 제1 단락 '정세 제시와 인물소개'는 좀 더 작은 단위의 내용 단락으로 세분화할 수 있다. 예컨대 CSJH(신재호 40장본 적벽가)의 '정세 제시와 인물소개' 단락은 '정세 제시', '인물소개', '도원결의', '공명천거'로 나눌 수 있다. 또한 '공명천거'는 더 세분화하여 '서서의 공명천거', '공명천거' 단락처럼 하위 단락으로 나누어진다.

SID	L1	L2	L3	L4	L5	L1	L2	L3	L4	
S0001	1	010	000	000	000	정세 제시와 인물소개	정세 제시			
S0002	1	015	000	000	000	인물 소개				
S0003	1	020	000	000	000	도원결의				
S0004	1	030	010	010	000	공명 천거	서서의 공명 천거	서서 이별		
S0005	1	030	020	020	000		공명 천거			
S0006	1	030	020	000	000			수강신성의 공명 천거		
S0007	2	010	010	000	000	상고초려	인물지역	유비		
S0008	2	010	020	000	000			관우		
S0009	2	010	030	000	000			장비		
S0010	2	020	000	000	000		와룡강 경계 물이			
S0011	2	030	000	000	000		일크			
S0012	2	040	010	000	000		이크	제갈공 명날		
S0013	2	040	020	000	000			유비 만지		
S0014	2	040	030	000	000			유비 시작		
S0015	2	040	040	000	000			활승인 등 만날		
S0016	2	050	000	000	000		상고			
S0017	2	060	010	000	000		공명 명날	대흥시		
S0018	2	060	020	000	000			첫 만남		
S0019	2	060	030	000	000			공명 명날		
S0020	2	060	040	000	000			공명 명날		
S0021	2	060	050	000	000			공명 명날		

그림 2. <적벽가> 내용 단락의 구조 분류

[그림 2]는 <적벽가>의 내용 단락 전체를 층위별로 분류하여 도식화한 것이다. <적벽가>의 각 이본별로 개별 단락의 미세한 차이까지 측정하기 위해서는 층위에 따른 단락 구분이 필요하다. 내용 단락의 분류 기준은 상호 대비와 관계를 살피기 위하여 서사 단락을 등장 인물의 역할과 사건의 분화에 따라 세분화하였다.

<적벽가>의 전체 내용 단락을 층위별로 분류하면 '대 단락(L1)', '중단락(L2)', '소단락(L3)', '소소단락(L4)', '소소소단락(L5)'의 다섯 층위의 구조로 나눌 수 있다. 대 단락(L1)은 6개로 구분되며, 중단락(L2)은 22개, 소단락(L3)은 67개, 소소단락(L4)은 140개, '소소소단락(L5)은 146개로 집계된다. 위의 표에서 SID는 <적벽가>의 내용 단락의 일련번호를 의미한다. 이와 같이 <적벽가>의 내용 단락을 분류한 다음 전체 이본을 유형별로 인코딩 하

였다. 그리고 각 작품의 유형 표지를 추출하였고, 변별되는 내용 단락을 기호로 나타낸 작업 테이블을 구축하였다.

SID	ID															
	1	2	3	4	5	42	43	44	45	46	47	48	49	50	51	52
EUN02	SDGL02	SDGL03	SDGL04	SKK502	CS55	CHSH	CJKJ	CKYJ	CKYS	PYVB
50001	x	x	x	x	x	b1	d	b2	b2	x
50002	x	x	x	x	x	b1	d	b2	b2	x
50003	a1	a1	a1	a1	a1	b1	b1	b2	b2	a1
50004	x	x	x	x	x	x	x	b1	b1	x
50005	x	x	x	x	x	b2	b2	b2	b2	x
50006	x	x	x	x	x	b2	b2	b2	b2	x
50007	x	x	x	x	x	b2	b2	b1	b1	x
50008	x	x	x	x	x	b1	b1	b1	b1	x
50009	x	x	x	x	x	b2	b2	b1	b1	x
50010	a1	a2	a1	a1	a1	b1	b6	b7	b2	b2
50011	a1	a1	a1	a1	a1	b2	a2	b	a2	x
50012	a1	a1	a1	a1	a1	d	a3	c	c	x
50013	x	x	x	x	x	x	x	c	c	x
50014	x	x	x	x	x	x	x	x	x	x
50015	x	x	x	x	x	b1	x	x	x	x
50135	a1	a2	a1	a1	ㄴ	b1	x	x	b3	x
50136	a1	b	a1	a1	ㄴ	c1	c3	d2	e	a3
50137	x	x	a	a	ㄴ	b2	x	x	a2	a2
50138	c	x	a	a	ㄴ	d1	b1	b1	b2	a
50139	a	a	a	a	ㄴ	c1	c1	c1	c4	a2
50140	a	b	b	b	ㄴ	x	x	x	x	x
50141	o	o	o	o	ㄴ	x	x	x	x	x
50142	o	o	o	o	ㄴ	x	x	x	x	x
50143	o	x	o	o	ㄴ	x	x	x	x	x
50144	o	o	o	o	ㄴ	x	x	x	x	x
50145	o	x	o	o	ㄴ	x	x	x	x	x
50146	x	x	x	x	ㄴ	a2	a2	a2	c	x
50147	o	x	o	o	ㄴ	x	x	x	x	o
50148	x	x	x	x	ㄴ	a2	a2	a2	c	x
50149	o	x	o	o	ㄴ	x	x	x	x	o

그림 3. 이본별 내용 단락 유형 테이블

[그림 3]은 <적벽가>의 이본별 내용 단락을 유형화 하여 정리한 것이다. <적벽가> 전체 이본의 내용 단락마다 고유의 일련번호를 부여하였다.

일련기호(SID)는 알파벳과 숫자의 조합으로 단락의 변별과 차이를 나타내며, 알파벳이 다른 경우는 전혀 다른 내용 단락을 의미한다. 같은 알파벳에 결합한 숫자는 내용 단락은 같으나 미세한 차이가 발견되었을 경우 부여한 기호이다. 이와 같은 방법을 통해 전체 이본의 내용 단락 유형은 가장 작은 내용 단락인 소소소단락과 마찬가지로 총 146개로 구성된다.

3. 이본간의 거리 행렬과 유사도 측정

다음으로 <적벽가> 이본 간의 거리와 상호 관련성을 파악하기 위하여 내용 단락 유형 테이블[그림 3]을 바탕으로 이본간 거리 행렬과 유사도를 측정하였다. 이를 통해 특정 이본을 중심으로 다른 이본과의 거리를 수치로써 환산할 수 있으며, 유사도를 시각적으로 표현할 수 있었다.

<적벽가> 이본의 거리 행렬 산출은 내용 단락의 상이에 따라 0에서부터 4까지의 척도를 부여하는 방식을 취하였다. 이본별 내용 단락을 유형별로 정리한 뒤 동일한 단락은 0, 내용 단락의 변이는 정도에 따라 1, 2, 3으로

표시하였고, 내용의 유사성이 없는 단락은 4와 같이 수치를 부여하였다. 이 과정을 통해 각 이본의 거리를 수치로 표시할 수 있다.

	CHSH	CJKJ	CKJS	CKYS	CLBU	CLSU	CPBS	CPDJ	CSJH	CSSS	EJUN01	EJUN02	...
CHSH	0	0.9	1.17	1.03	1.09	1.18	0.38	1	1.62	0.64	2.38	2.4	...
CJKJ	0.9	0	1.06	0.5	1.11	1.06	0.9	0.66	1.6	0.83	2.4	2.38	...
CKJS	1.17	1.06	0	1.12	0.3	1.35	0.99	1.17	1.72	0.87	2.49	2.51	...
CKYS	1.03	0.5	1.12	0	1.17	1.05	0.87	0.78	1.53	0.79	2.32	2.29	...
CLBU	1.09	1.11	0.3	1.17	0	1.26	0.9	1.27	1.66	0.79	2.4	2.43	...
CLSU	1.18	1.06	1.35	1.05	1.26	0	1.24	1.23	1.47	1.21	2.22	2.25	...
CPBS	0.38	0.9	0.99	0.87	0.9	1.24	0	0.85	1.4	0.16	2.34	2.32	...
CPDJ	1	0.66	1.17	0.78	1.27	1.23	0.85	0	1.64	0.79	2.36	2.33	...
CSJH	1.62	1.6	1.72	1.53	1.66	1.47	1.4	1.64	0	1.32	2.18	2.16	...
CSSS	0.64	0.83	0.87	0.79	0.79	1.21	0.16	0.79	1.32	0	2.26	2.23	...
EJUN01	2.38	2.4	2.49	2.32	2.4	2.22	2.34	2.36	2.18	2.26	0	0.06	...
EJUN02	2.4	2.38	2.51	2.29	2.43	2.25	2.32	2.33	2.16	2.23	0.06	0	...
...

그림 4. 단락 유형 차이에 따른 거리 행렬

[그림 4]는 CHSH(한승호 창본)을 중심으로 각 이본간의 거리를 측정하여 거리 행렬의 수치로 표시한 것이다. 가로와 세로의 행렬에 적힌 척도를 통하여 이본간의 거리를 가늠할 수 있다.

CHSH는 CJKJ와 행렬값이 0.9로 나타나며, CSSS와는 행렬값이 0.64로 산출된다. 따라서 CHSH는 CJKJ, CSSS와 가까우며, 특히 CSSS와는 매우 밀접한 거리에 있다고 해석할 수 있다. 결국 CHSH는 두 이본과 가까운 거리에 있기 때문에 친연성이 높다고 할 수 있다. 이에 반해 CHSH는 EJUN01과 행렬값이 2.38, EJUN02와 행렬값이 2.4로 산출된다. 따라서 CHSH(한승호 창본)는 두 이본과 상당한 거리로 떨어져 있는 작품임을 알 수 있다.

CHSH(한승호 창본)을 기준으로 보자면, CJKJ(정권진 창본)와 CSSS(송순섭 창본)와는 내용 단락상 가까운 위치에 있으며, EJUN01(완판 화릉도 83장본 정미구동신간본), EJUN02(완판 화릉도 84장본 무신서계개각본)와는 거리가 먼 작품임을 알 수 있다. 특히 EJUN02(완판 화릉도 84장본 무신서계개각본)의 경우는 가장 관련성이 떨어진다고 해석할 수 있다.

다음으로는 완판본을 기준으로 하여 이본들의 유사도를 측정하는 작업을 진행하였다. 완판본을 기준으로 삼

은 까닭은 텍스트가 고정되어 있으므로, 계열 파악에 용이하기 때문이다. 우선 [그림 4]와 같은 방식으로 완판본 중심의 거리 행렬값을 구한 뒤, 이를 아래의 그래프처럼 시각화하여 완판본과 여타 이본의 관련성을 확인하였다.

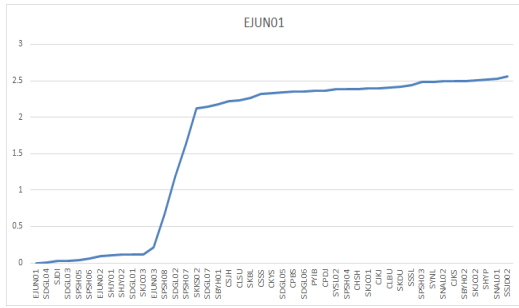


그림 5. 'EJUN01'과 여타 이본의 유사도

[그림 5]의 그래프는 EJUN01(완판 화릉도 83장본 정미구동신간본)을 기준으로 이본들의 거리를 측정한 것이다. 세로축의 '0'에 가까울수록 유사도가 높은 이본이고, 세로축에서 높이 위치할수록 이본간의 유사도가 멀어지는 이본이다.

위의 그래프에서 EJUN01부터 EJUN03까지는 EJUN01과 유사도가 높은 이본으로 완판 계열의 이본이며, SPSH07부터 SSJD02까지는 유사도가 먼 이본으로 창본 또는 창본 계열의 이본이다. 그리고 중간에 기울기의 구간은 완판 계열과 창본 계열의 중간 성격을 띤 이본으로 추정할 수 있다. 뒤에서 언급할 사항이지만, 중간 구간에 위치한 SKKS02, SPSH07, SDGL02는 책의 분권, 개작, 낙장으로 인하여 상대적으로 완판본과 유사도가 낮게 나타난 것이다.

이처럼 거리 유사도 측정을 통해서 <적벽가>의 이본은 크게 완판 계열과 창본 계열로 나누어져 있음을 확인할 수 있다.

III. 컴퓨터 문헌 분석 기법의 적용

1. 다차원 척도법에 의한 분석

다차원 척도법은 다차원 공간에서 정의되는 유사성(similarity, dissimilarity)을 저차원의 공간에 기하학적

으로 나타낸 것이다[12]. 즉 이본간의 상호 관계를 탐색하는 기법이라고 할 수 있다.

이러한 방식을 <적벽가>에 적용하여 이본간의 거리를 저차원 공간의 척도에 위치시키면 아래와 같은 그래프를 얻을 수 있다.

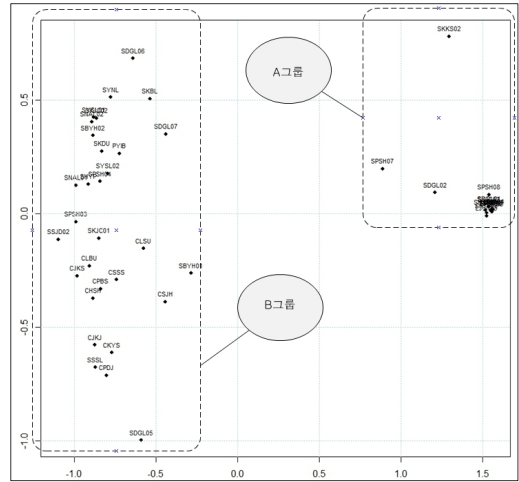


그림 6. 다차원 척도법에 의한 거리 행렬 분석

[그림 6]의 그래프는 [그림 4]의 거리 행렬을 2차원 공간에 구현하여 이본간의 거리를 시각화한 것이라 할 수 있다. 이를 통해 이본간의 유사성을 공간적 거리로 확인할 수 있다. 위의 좌표는 <적벽가>의 각 이본이 친소 관계에 따라 어떠한 거리 행렬로 작품군을 형성하고 있는지를 명확하게 보여주는데, A그룹과 B그룹으로 양분되어 있음을 확인할 수 있다.

A그룹에 밀집된 이본군은 완판 계열의 작품들로 파악된다. A그룹 중에서도 산발적으로 흩어져 있는 SKKS02, SPSH07, SDGL02는 완판 계열의 작품 중에서도 독자성을 띤 작품들이라고 할 수 있다. 이 세 작품은 146개의 전체 내용 단락 중 각각 100개, 59개, 4개의 단락이 누락되어 있다. 텍스트를 직접 확인해본 결과, SKKS02는 완판본의 후반부에 해당하는 단락이 없고, SPSH07은 전반부에 해당하는 단락이 없으며, SDGL02는 일부 낙장을 확인할 수 있다. 따라서 이 앞의 두 이본은 분권된 텍스트 또는 낙장본임을 알 수 있으며, SDGL02 개작본이면서 일부 텍스트가 낙장된 이본으로 확인할 수 있게 된다.

B그룹에 위치한 이본들은 창본이나 창본 계열의 작품으로, 세로축의 0.0 위쪽에 찍힌 좌표들은 창본 계열의 필사본과 활자본이다. 0.0 아래쪽에 찍힌 좌표들은 창본과 창본 계열의 필사본임을 알 수 있다. 한편 B그룹의 작품 중에서 CSJH, SBYH01은 가로축의 0.0에 가까운 비교적 오른쪽 좌표에 위치하는데, 상대적으로 완판 계열과 가까운 작품이라는 것을 의미한다. 즉 이들 이본은 완판 계열의 내용 단락을 수용한 작품들로 평가할 수 있다.

2. 계층적 군집 분석

계층적 군집 분석(hierarchical cluster analysis)은 군집 분석의 일종으로, 가까운 거리를 차례로 묶어 군집을 파악하는 기법이다[1]. 또한 형성된 군집에 위계를 주어 자료 간 상호 관계를 표시하기 때문에 유사한 속성을 지닌 이본들의 성향을 가늠할 수 있다.

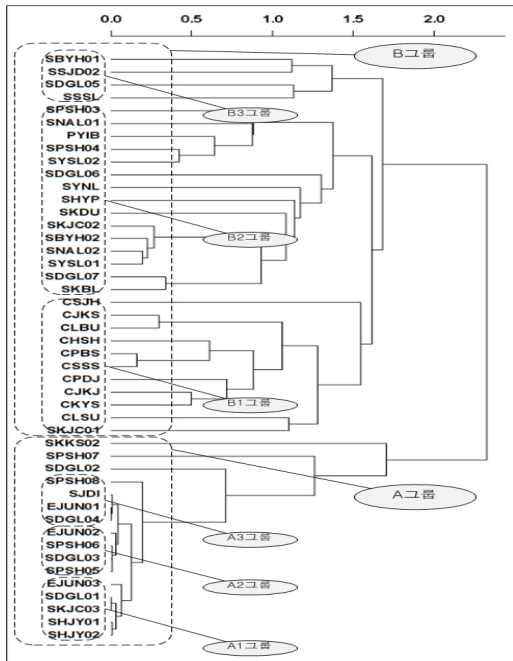


그림 7. <적벽가> 이본의 계층적 군집 분석

[그림 7]은 계층적 군집 분석을 통해 <적벽가> 이본의 군집을 시각화 한 것으로, [그림 6]을 수형도(樹形圖)와 같은 형태로 표현한 것이다. 이 계층적 군집 분석을 통해서 <적벽가> 이본의 군집은 크게 A그룹과 B그룹으로 나

누어짐을 확인할 수 있다. 두 이본이 같은 군집에 속하더라도 거리가 길면 상대적으로 공유하는 단락이 적고, 그만큼 이본간의 거리와 격차도 크다는 것을 의미한다.

우선 아래의 A그룹에 속하는 군집은 완판 계열의 작품으로 파악되는데, 하위 세 그룹이 확인된다. A3그룹에 속하는 이본은 완판 화룡도 83장본 정미구동신간본(EJUN01)과 주변 이본이며, A2그룹에 속하는 이본은 완판 화룡도 84장본 무신서계개각본(EJUN02)과 주변 이본으로 판단된다. 그리고 A1그룹은 완판 화룡도 84장본 무신양책방본(EJUN03)과 그 파생작으로 볼 수 있다. SKKS02, SPSH07, SDGL02는 앞서 다차원 척도법으로 해석했던 것처럼 완판 계열의 작품이지만, 낙장이거나 내용 단락의 변이로 독자적인 성격을 갖는 작품으로 확인된다.

위의 B그룹에 분포한 작품들은 창본이거나 그와 유사한 내용 단락을 공유하는 작품들이 분포하는데, 세 그룹으로 나누어진다. B3그룹은 창본 계열의 작품이면서 비교적 장편에 속하는 작품들로, 문어체의 문장이 주로 사용되는 것으로 분석된다. B2그룹도 창본 계열의 작품인데, 완판본의 서사 단락을 수용한 작품이다. 이 작품에는 단가, 가사 등 가요사설이 적잖이 포함되어 있으며 창본에 가까운 필사본과 활자본이 위치한다. B1그룹에는 창본과 그와 같은 내용 단락을 공유하는 이본이 위치하는데, 이 그룹은 현전 판소리 창본과 가장 유사하다고 볼 수 있다.

3. <적벽가> 이본의 계통 분기 분석과 검토

다음으로는 계통 분기 분석을 통하여 <적벽가> 이본의 계통별 분기를 시각하기로 한다. <적벽가> 이본의 분기 분석을 통해서도 앞의 분석 결과들과 마찬가지로 <적벽가>의 이본이 크게 A그룹과 B그룹으로 나뉘어져 있음을 확인할 수 있다.

[그림 8]은 <적벽가> 이본의 분기 분석의 결과를 시각화한 것으로 각 이본의 분류와 영향관계를 가늠할 수 있게 한다. A그룹은 이본들이 밀집되어 있는 부분과 결과지로 분기된 형태를 띠고 있으며, B그룹은 3개의 군집을 형성하고 있다.

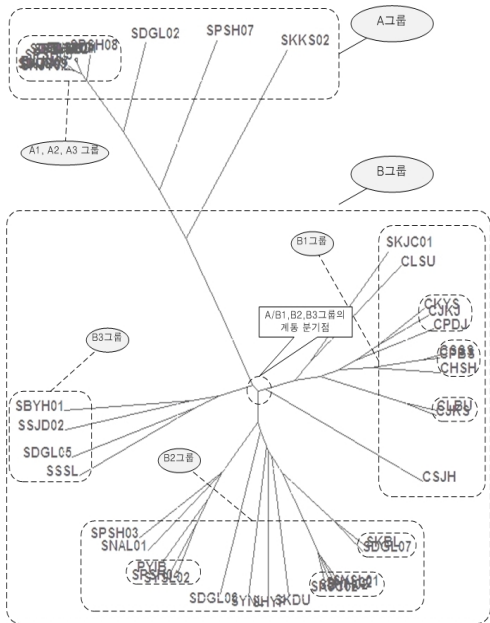


그림 8. 계통의 분기 분석에 따른 <적벽가>의 이본 분포

A그룹은 완판본을 중심으로 형성되어 있다. 완판 화로도 83장본 정미구동신간본(EJUN01), 완판 화로도 84장본 무신서계개각본(EJUN02), 완판 화로도 84장본 무신양책방본(EJUN03)은 A1, A2, A3의 소그룹으로 매우 가까운 분기점에서 모여 있다. 이들 완판본이 보여주는 밀집 형태는 이 작품들이 같은 내용 단락을 공유하고 있음을 보여준다. 즉 개별적 이본의 특징을 지니고 있지만 작품간의 편차가 크지 않음을 의미한다. SKKS02, SPSH07, SDGL02는 내용 단락을 공유하지 않는 독자적인 이본이거나 낙장본이라고 볼 수 있다. 앞서 확인했던 것처럼 이들 작품은 애초 분권된 텍스트일 가능성이 있다.

B그룹은 창본 계열의 이본이 세 갈래로 분기되어 있다. B1그룹은 임방울 창본(CLBU)과 정광수 창본(CJKS), 박봉술 창본(CPBS)과 송순섭 창본(CSSS), 정권진 창본(CJKJ)과 김연수 창본(CKYS)으로 서로 가깝게 위치하여 밀접도가 높은 편이다. 이들은 스승과 제자 사이이거나 사설을 공유하는 관계였기 때문에 텍스트로도 큰 차이가 없는 것으로 생각된다. 박동진 창본(CPDJ), 이선유 창본(CLSU)은 좀 떨어진 분기에 위치하는데, 이들 창본은 중고제 혹은 고제 판소리와 관련이 있을 것

로 보인다. B2그룹도 창본 계열의 작품이지만 활자본인 유일서관본(PYIB)을 중심으로 활자본의 저본이 되었던 작품이거나 파생작으로 판단된다. B3그룹은 『삼국지연의』와 관련성이 높은 작품들로, 비교적 장편이며, 이본간의 관련성은 떨어진다고 할 수 있다.

IV. 기존 연구와의 비교

<적벽가>의 대표적인 이본 연구는 김상훈, 김기형, 이기형에 의하여 진행되었다[3-5].

김상훈은 <적벽가>의 이본을 창본 계열과 소설본 계열로 나누어 살폈는데, ‘박막과 전투’, ‘동작대부’, ‘황개사설’, ‘메초리사설’ 등의 유무를 두 계열의 변별점으로 삼았다[1]. 김기형은 <적벽가>를 형태상 판소리 창본, 필사본, 완판본으로 나누고 필사본을 다시 창본 계열과 완판 계열로 분류하였으며, 창본 계열을 다시 ‘민적벽가’ 계열과 ‘적벽가’ 계열로 나누었다. 창본 계열의 분류는 ‘삼고초려’ 대목의 유무를 기준으로 삼은 것이다[2]. 이기형은 <적벽가>의 이본을 형태에 따라 창본과 필사본, 완판본, 구활자본으로 구분하였고, 필사본을 창본 계열, 완판 계열, 창본과 완판의 교합본으로 분류하였다[3].

이중에 김상훈의 연구 <적벽가>의 이본 형태에 주목했던 것으로 김기형, 이기형의 이본 분류와 크게 다르지 않다. 따라서 김상훈을 제외한 두 연구자의 이본 분류를 제시해보고 본 연구의 결과와 비교 검토해보기로 한다.

표 2. <적벽가> 이본론에서의 계통 분류

	김기형	이기형
민적벽가 계열	CLBU	
적벽가 계열	CSJH CJKS CHSH CPBS CPDJ CPDJ CJKJ CKYS CLSU SKJC01 SKJC02 SNAL01 SNAL02 SPSH03 SPSH04 SSJD02 SKBL	CLSU CLBU CJKS CPDJ CSSS CPBS CKYS CHSH CJKJ CSJH SDGL07 SNAL01 SKJC01 SYSL02 SNAL02 SKJC02

	SKDU SBYH02 SBYH01 PYIB	SYSL01 SSJD02 SBYH01 SKBL SBYH02 SHYP SPSH03 SDGL05 SSSL SDGL06 SPSH04 SYNL
완판 정미구동 계열	EJUN01 SJDI SDGL04	EJUN01 SJDI
완판 무신서계 계열	EJUN02 SPSH05	EJUN02 SPSH05 SPSH06 SPSH07 SDGL03 SKKS02
완판 무신양책 계열	EJUN03	EJUN03 SPSH08 SDGL01 SHJY02 SKJC03

김기형과 이기형은 <적벽가>의 이본 계통을 크게 창본 계열과 완판 계열로 나누었다는 점에서 대동소이하다 [4][5]. 완판 계열을 세 가지로 나눈 것도 비슷한 양상을 보인다. 김기형의 연구는 분석 대상의 이본수가 이기형의 연구보다 상대적으로 적은 편이지만, 임방울 창본을 별도로 '민적벽가'로 설정했다는 점에서 변별적 의미를 갖는다고 할 수 있다. 두 연구자의 계통 분류는 큰 틀에서 보자면 유사도 측정[그림 5]에서 그 타당성을 확인할 수 있으나, 세부적인 부분에서 다른 결과를 보인다.

우선 완판 계열부터 검토하자면, 계층적 군집 분석[그림 7]에서 완판본은 A3(완판 화룡도 83장본 정미구동신간본), A2(완판 화룡도 84장본 무신서계개각본), A1(완판 화룡도 84장본 무신양책방본)으로 크게 분류되고 있다는 점은 두 연구자의 결과와 같다. 분기 분석에 있어서 세 개의 계열은 친밀도가 분명히 확인된다. 하지만 SKKS02(김광순 소장 23장본 화용도), SPSH07(박순호 소장 46장본 화용도), SDGL02(단국대소장 낙장 82장본 화용도)는 완판 계열에 속하면서도 상대적으로 거리가 떨어져 있는 작품으로 분석된다. 이러한 결과는 이들 작품이 완판 계열의 파생작이거나 분권본, 낙장본임을 알려주는 표지가 된다.

창본 계열의 경우에도 김기형과 이기형의 분류와 유사

점을 찾을 수 있다. 하지만 본 연구 결과를 통해서는 보다 다양한 스펙트럼을 확인할 수 있었다. 계층적 군집 분석[그림 7]을 통해 볼 때 창본 계열은 B3, B2, B1 군집으로 나누어짐을 확인할 수 있다. 분기 분석[그림 8]에 의해 파악해보면 B3 그룹은 창본이면서 완판본의 단락을 수용한 이본으로 확인된다. B2 계열은 창본에 가까운 완판본의 영향을 받은 이본군으로 파악되며, B1 군집에는 다수의 창본이 분포함을 알 수 있다. 특히 B1의 군집에서는 CLSU(이선유 창본)는 독립적인 위치에 놓여 있어 중고제 또는 고제의 텍스트임을 알 수 있으며, CLBU(임방울 창본)과 CJKS(정광수 창본)처럼 친밀도가 매우 높은 이본들을 확인할 수 있다. 이처럼 친밀도가 높은 창본의 군집이 나타난 까닭은 판소리 명창의 사승관계를 고려하여 추론할 수 있었는데, 이는 스승에게 배운 텍스트가 비교적 온전히 제자에게 전해졌던 현상을 보여주는 것이라고 해석할 수 있다.

V. 결론

<적벽가>는 선행 연구의 축적으로 작품의 계통과 계열을 이루고 있는 이본의 성격에 대해서 어느 정도 논의가 이루어졌다고 할 수 있다. 하지만 기존의 연구에는 상당 부분 연구자들의 직관이 반영되었다는 점은 부정하기 어렵다. 그렇기 때문에 <적벽가> 계통 분류는 좀 더 과학적인 방법으로 해석해야 할 필요성이 제기된다. 우선 연구의 진행을 위해 <적벽가> 전체 이본간의 거리를 파악하여 각 이본군의 계통을 확인하는 작업을 진행하였다. 그 다음으로 본격적인 컴퓨터 문헌 분석법을 적용하였는데, 다차원 척도법에 의한 분석, 계층적 군집 분석, 개별 이본의 계통 분기 분석을 활용하였다.

본고에서는 컴퓨터 문헌 분석 기법을 활용하여 <적벽가> 이본의 계통 분류를 시도하였는데, 종합적인 분석 결과 기존의 <적벽가>의 창본 계열과 완판 계열의 분류는 대체로 일치한다는 것을 확인할 수 있었다. 그렇지만 좀 더 세밀한 부분으로 따져 보자면, 기존 연구와의 차이점도 발견할 수 있다.

기존의 연구에서 <적벽가>의 이본은 대체로 창본 계열과 완판 계열로 분류하였다. 필사본의 경우 창본 계열과

완판 계열을 나누는 시각이 다르게 적용되기도 하였다. 앞서 유사도 측정[그림 5]과 다차원 척도법[그림 6]에 의한 분석을 통해 그동안 <적벽가>의 계통 분류가 실제 어느 정도 부합하는 것을 확인할 수 있었다. <적벽가> 이본의 계층적 군집 분석[그림 7]과 <적벽가> 이본 계통의 분기 분석표[그림 8]는 기왕의 연구보다 <적벽가>의 이본군에 대한 좀 더 상세하고 분명한 정보를 보여준다는 점에서 의미가 있다고 할 수 있다.

본 연구를 통하여 <적벽가>의 이본간 거리를 확인함으로써 각 이본의 개별적인 의미를 도출할 수 있는 계기가 마련되었다고 할 수 있다. 특히 기존 <적벽가>의 이본론에서 제외된 작품들을 주요 이본과의 거리와 관계망 속에서 새롭게 의미를 부여할 수 있는 계기를 마련하였으며, 그간 모호했던 판소리 <적벽가>의 전승과 사사관계를 파악할 수 있는 단서를 제공하였다는 데에서 의미가 있다.

본 연구의 성과는 기존 이본론을 통계화하여 구체적으로 점검할 수 있다는 데에 있다. 이후 추가적인 분석을 통하여 <적벽가> 이본의 통시적인 의미도 함께 고찰하기를 기대한다.

참 고 문 헌

[1] 강한영, *신재효 판소리 사설집(全)*, 민중서관, p.25, 1971.
 [2] 김흥규, *판소리의 이해*, 창작과비평사, p.17, 1979.
 [3] 김상훈, *적벽가의 이본과 형성 연구*, 인하대학교, 박사학위논문, pp.12-58, 1992.
 [4] 김기형, *적벽가 연구*, 민속원, pp.29-50, 2000.
 [5] 이기형, *필사본 화용도 연구*, 민속원, pp.93-128, 2003.
 [6] 하명정, “코퍼스에 기반한 문학텍스트 분석,” 한국콘텐츠학회논문지, 제13권, 제9호, pp.440-447, 2013.
 [7] 최원경, “초등영어교육 연구 논문의 변천: 코퍼스 기반 분석,” 한국콘텐츠학회논문지, 제19권, 제2호, pp.11-21, 2019.
 [8] 김민정, 김철주, “텍스트마이닝을 활용한 승례문 관련 기사의 트렌드 분석,” 한국콘텐츠학회논문지, 제17권, 제3호, pp.474-485, 2017.

[9] 최운호, 김동진, “컴퓨터를 이용한 고전문학 디지털콘텐츠의 유사도에 따른 이본 계통 분류 연구,” 한국정보기술학회지, 제12권, 제7호, pp.27-38, 2003.
 [10] 최운호, 김동진, “컴퓨터 문헌분석 기법을 활용한 <토끼전> 이본 연구,” *우리문학연구*, 제58집, pp.123-154, 2018.
 [11] 김진영, 김현주, 이원걸, 안영훈, 김지연, 김태희, 이기형, 백미나, 차충환, 김동건, 서유석, *적벽가 전집 1-7*, 박이정, 1998-2003.
 [12] 신상민, 김은성, 최용석, “분할법에 의한 가상점을 활용한 다차원 척도법,” *응용통계연구*, 제28집, 제6호, p.1171, 2015.

저 자 소개

이진오(Jin-O Lee)

정회원



- 2003년 2월 : 고려대학교 국어국문학과(문학사)
- 2014년 8월 : 고려대학교 국어국문학과(문학박사)
- 2014년 9월 ~ 현재 : 금오공과대학교 교양교직과정부 강사

<관심분야> : 고전문학, 판소리, 문화콘텐츠

김동건(Dong-Keon Kim)

정회원



- 1995년 2월 : 경희대학교 국어국문학과(문학사)
- 2001년 2월 : 경희대학교 국어국문학과(문학박사)
- 2016년 3월 ~ 현재 : 경희대학교 후마니타스칼리지 교수

<관심분야> : 고전문학, 판소리, 정보학