

2020 한국철학자연합대회

2020

21세기 세계 속 한국철학

- 김재권을 추모하고 그 철학을 기리며 -

Corean Philosophy in the 21stC. World

UCCCP

Unified
Congress of
Corean
Philosophers

자율세션_1권

10.23.(금)~24.(토)

서울시립대학교(자연과학관 2층 국제회의장)

주관 한국철학회

공동주최 대동철학회, 대한철학회, 동양철학연구회, 범한철학회, 새한철학회, 서양근대철학회, 인도철학불교학연구소, 인도철학회, 철학연구회, 한국기독교철학회, 한국니체학회, 한국동서철학회, 한국분석철학회, 한국양명학회, 한국여성철학회, 한국유교학회, 한국윤리학회, 한국철학교육학회, 한국철학사연구회, 한국해석학회, 한국환경철학회, 한국철학회

후원 한국연구재단, 서울시립대학교

2020 한국철학자연합대회

2020

21세기 세계 속 한국철학

- 김재권을 추모하고 그 철학을 기리며 -

Corean Philosophy in the 21stC. World

UCCCP

Unified
Congress of
Corean
Philosophers

자율세션_1권

10.23.(금)~24.(토)

서울시립대학교(자연과학관 2층 국제회의장)

주관 한국철학회

공동주최 대동철학회, 대한철학회, 동양철학연구회, 범한철학회, 새한철학회, 서양근대철학회, 인도철학불교학연구소, 인도철학회, 철학연구회, 한국기독교철학회, 한국니체학회, 한국동서철학회, 한국분석철학회, 한국양명학회, 한국여성철학회, 한국유교학회, 한국윤리학회, 한국철학교육학회, 한국철학사연구회, 한국해석학회, 한국환경철학회, 한국철학회

후원 한국연구재단, 서울시립대학교

- 주관 한국철학회
- 공동주최 대동철학회, 대한철학회, 동양철학연구회, 범한철학회, 새한철학회, 서양근대철학회, 인도철학불교학연구소, 인도철학회, 철학연구회, 한국기독교철학회, 한국니체학회, 한국동서철학회, 한국분석철학회, 한국양명학회, 한국여성철학회, 한국유교학회, 한국윤리학회, 한국철학교육학회, 한국철학사연구회, 한국해석학회, 한국환경철학회, 한국철학회
- 후원 한국연구재단, 서울시립대학교

이 발표논문집은 2020년도 정부재원(교육부)으로 한국연구재단의 지원을 받아 발간되었음.

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government.

인사말

한국의 철학자 여러분! 안녕하십니까!

동서양의 철학회가 모두 한 자리에 모이는 철학의 향연인 한국철학자 연합대회에 여러분을 초대합니다.

금번 연합 학술대회는 그동안 한국의 철학인으로서 세계 철학의 발전에 크게 기여하고 그 학문적 깊이와 명성으로 한국 철학자 사회의 위상을 드높인, 작고하신 선배 철학자들의 업적을 되새기고 뜻을 기리는 추모의 자리이기도 합니다. 2019년 작고하신 김재권 선생님의 철학을 필두로 박이문 선생님, 정화열 선생님의 철학을 되새기고 그 메시지를 함께 공유하고자 합니다. 또한 이를 계기로 세계 속의 한국 철학의 위상과 현주소도 함께 되돌아보고자 합니다. 나아가 21세기 한국 사회에 기여할 한국의 실천적 철학 담론에 대해서도 함께 고민해보고자 합니다.

2008년 세계철학자대회가 한국에서 열린 이후로 한국의 모든 철학인들이 함께 하는 학술공동체 향연은 다소 주춤했습니다. 그로 인한 학문적 소통과 연대의 갈증이 깊었기에, 많은 학회가 적극 참여하였고 연합 학술대회의 성공적 개최를 위해 각고의 노력으로 준비하였습니다. 이를 좋은 계기로 삼아 한국 철학자 사회가 세계 속의 한국 철학으로 한 단계 더 도약할 수 있기를 소원합니다.

코로나 19라는 전대미문의 사태에도 불구하고 한국철학자 연합대회의 성공을 위해 적극 참여해 주신, 모든 발표자와 토론자분들 그리고 공동주최 학회의 모든 임원 및 관계자분들의 노고에 다시 한번 머리 숙여 감사드리며, 10월 23-24일 양일 간 서울시립대학교에서(온/오프라인 병행으로) 개최되는 연합 학술대회에 적극 참여해 주시길 당부드립니다. 감사합니다.

(개최와 관련한 상세한 내용은 아래의 포스터와 초대장을 참조해 주시기 바라며, 연합대회 전체 세션과 학회별 자율 세션에 참가 또는 접속하는 방법은, 현재 연합대회 홈페이지를 별도로 구축하고 있사오니 구축되는 즉시 별도 안내를 드리도록 하겠습니다.)

2020 한국철학자 연합대회 상임운영위원회

2020 한국철학자연합대회(2020 UCCP)

프로그램

10월 23일(금)

09:30-10:30	개회식	사회_김성민(건국대)
09:30-10:00	대회 참가자 등록	
	개회사	이중원(한국철학회 회장, 서울시립대)
	축사 I	서순탁(서울시립대 총장)
	축사 II	손동현(우송대 석좌교수)
10:30-11:30	기조발제(Keynote Speech)	사회_백훈승(범한철학회 회장, 전북대)
	21세기 초까지의 한국현대철학 - 서양철학 수용 120여년, 한국현대철학의 과제를 성찰하기	김정현(대한철학회 전임회장, 원광대)
	김재권과 그의 철학 - 물리세계 안의 행위주관성	정대현(이화여대 명예교수)
11:30-11:40	휴식	
11:40-12:20	[전체 발제 1부]	사회_하중호(고려대)
	김재권과 세계 형이상학 및 심리철학(Plenary Session Part1)	
	인간에 대한 이해와 설명 - 김재권의 미완의 과제	원치욱(광주과기원)
	형이상학 르네상스와 김재권 - 설명 이론을 중심으로	이재호(중앙대)
12:20-14:00	중식	
14:00-14:40	[전체 발제 2부]	사회_박병준(철학연구회/한국철학상담치료학회 회장, 서강대)
	세계 속의 한국철학(Plenary Session Part2)	
	정화열의 몸의 정치와 진리지평의 확장 - '테오리아의 독재'에 대한 저항의 논리	송석량(한국동서철학회 회장, 목원대)
	박이문의 '동지'에서 날아오른 것	강학순(안양대)
14:40-15:00	휴식	
15:00-15:40	[전체 발제 3부]	사회_이상훈(한국철학회 전임회장, 대전대)
	세계를 향한 한국으로부터의 철학적 메시지(Plenary Session Part3)	
	서양권 연구를 통해서 본 퇴계학의 생명력	장윤수(대동철학회 회장, 대구교대)
	현대와 세계에서 본 다산 정약용의 철학	
	- 유곡(幽谷)으로부터 날아오른 한 마리 새가 교목(喬木)으로 옮겨 가다.-	방 인(경북대)
15:40-16:00	휴식	

	[종합담론] 사회_김교빈(호서대) 21세기 한국현대철학, 어디로?
	세계적 차원의 철학 담론에서 한국계 철학자들의 담론생성력 홍윤기(동국대)
16:00-17:20	21세기 한국 철학의 문제 담론 1) 한국철학은 한국의 철학인가? / 권영우(한국외대) 2) 한국 철학계의 교육 및 학문 후속세대 활성화 방안 / 홍윤기(동국대) 3) 한국 사회 갈등 해소를 위한 사회적 가치 연구 활성화 방안 이중원(서울시립대. 한국철학회 회장) 4) 한반도 평화와 통일을 위한 철학의 역할 / 연효숙(연세대. 한국철학사상연구회 회장) 5) 한국사회의 미래전망과 여성주의적 가치 / 양선숙(경북대. 한국여성철학회 회장)
17:20-17:30	휴식
	제29회 열암철학상 시상식 사회_허남진(서울대. 열암철학상 운영위원장)
17:30-17:50	박종홍 철학과 열암철학상
17:50-18:00	2020 철학올림피아드 시상식 사회_이지애(이화여대. 철학올림피아드위원장)
18:00-20:00	만찬

10:20-10:55	[밀레니얼세대의 취미와 유교] “재밌는” 유교, 유교와 게임의 만남 발표_안승우(한양대), 토론_오준원((주)잼블로)
10:55-11:30	[밀레니얼세대의 직업과 유교] 유교적 관점에서 본 밀레니얼 세대의 직업윤리와 직업의식 발표_강보승(성균관대), 토론_전지열(성균관대)
11:30-12:05	[밀레니얼세대의 사랑과 유교] 학이시습(學而時習), ‘사랑하기’를 통한 치유와 성장의 날갯짓. - 자기애(自己愛)와 방탄소년단(BTS) 열풍을 중심으로 발표_윤민항(성균관대), 토론_윤부한(한국문화번역원 문학진흥본부장)
[2부] 종합토론	
12:10-13:00	좌장_김재경(조선대)
새한철학회	
10:00-13:00	‘나’를 찾는 행복 / 황종환(한남대)
	『화엄경』 「십정품」에 담긴 철학적 사유 - 10가지 삼매를 중심으로 도업 강기선(동국대)
	중국어이슬람의 사상적 특징: 유학적 이슬람 / 권상우(계명대)
서양근대철학회	
서양근대철학에서 다시 보는 인과론의 문제 황수영(홍익대)	
10:00-13:00	데카르트 철학에서 자연과 목적의 문제 - “감각들 속에는 신의 능력과 선함을 보여 주지 않는 것은 하나도 없다(AT VII, 87, 26-28)”에 대한 해석의 한 시도 / 이재훈(세종대)
	흠의 인과 추론의 정당성 문제 : 『인성론』1권과 『인성론』2권, 3권 사이의 관계를 중심으로 / 김병재(대구경북과학기술원)
	칸트의 인과성 개념에 대한 비판적 고찰 / 박제철(서울시립대)
인도철학회 / 인도철학불교학 연구소	
인도철학에서 언어의 문제	
10:10-10:40	바르뜨리하리의 언어론 - 스포따(sphoṭa)를 중심으로 / 박수영(동국대)
10:40-11:10	유식의 구조와 언어의 상관성에 관한 고찰 / 김재권(능인불교대학원대)
11:10-11:40	인도 사상에 있어 표현할 수 없는 것에 대한 언어표현 / 박기열(동국대)
11:40-12:10	시의 논리 - 『카비야 알람카라』Kāvyaśāstra 제5장을 중심으로 류현정(서울대)
12:10-13:00	종합토론 사회_안필섭(동국대)
한국기독교철학회	
10:00-13:00	평화주의는 인식론이다 - 존 요더의 평화주의 인식론 탐구 / 김기현(로고스 서원)
	앙리 베르그손이 만난 하나님 - 이성적 신과 직관의 신 / 양선진(충남대)
	한국의 기독교철학자 100년(1) - 일제강점기의 기독교철학자 / 최태연(백석대)

한국니체학회		장소_서울스퀘어 중회의실 1
14:00	개회	사회_김주휘(교원대)
14:00-14:10	개회사 정낙림(한국니체학회장)	
14:10-14:50	클로소프스키의 니체 해석 / 발표_임건태(대진대), 논평_강용수(고려대)	
14:50-15:30	니체의 실험철학과 철학적 의사 - 두 개념의 사상적 연관성의 해명을 중심으로 발표_이상범(원광대), 논평_김주휘(교원대)	
15:30-15:40	휴식	
15:40-16:20	쇼펜하우어와 니체의 예술론에 대한 비판적 고찰 - 개체화의 원리를 중심으로 발표_남재민(제주대), 논평_이주향(수원대)	
16:20-17:00	니체 철학에서의 수치심 / 발표_이선(전북대), 논평_김선희(강원대)	

한국동서철학회	
21세기 오늘, 한국의 시선으로 본 동서철학	
09:40-10:00	등록 및 개회 사회_조원일(전남대)
10:00-12:30	주제발표
	아리스토텔레스에 있어서 우애, 그리고 사랑 / 손윤락(동국대)
	플라톤과 절제(sōphrosynē)의 인간학 / 서영식(충남대)
	윤리학으로서의 메틀로-폰티 철학에 대한 국내 연구의 관점과 새로운 해석의 가능성 / 한우섭(중부대)
	의심의 '한국'철학, 한국에서도 철학을 하는가? - 세계철학을 꿈꾸는 한 철학도의 자전적 고백 / 박치원(한국외대)
	방촌 황희의 사상과 전인교육 / 황수영(목원대)
	조선후기의 도가사상 / 이진경(청운대)
12:30-13:30	최근 서양철학의 존재론적 전회(ontological turn) 경향과 한국철학의 생명존재론 - 신유물론과 사변적 실재론, 홍대용의 '인물균' 사상을 중심으로 / 이원진(연세대)
	지정종합토론 좌장_장영란(한국외대)
12:30-13:30	논평 및 토론 전종윤(전주대), 박규철(국민대), 심귀연(경상대), 이종성(충남대), 김창경(서원대), 정우진(경희대), 조성환(원광대)

한국양명학회	
21세기, 양명학 연구의 새로운 모색	
09:30-09:50	등록
[1부] 사회_김윤경(조선대)	
09:50-10:00	인사말 및 축사
10:00-10:30	왕양명의 용장오도 다시 읽기 - 용장의 생활을 중심으로 발표_이우진(공주교대), 토론_박길수(강원대)
10:30-11:00	양명학을 활용한 대학 인성교육 개발에 관한 연구 발표_조지선(충남대), 토론_고재석(성균관대)

11:00-11:10	휴식	
[2부]		사회_박성호(원광대)
11:10-11:40	한국에서의 중국양명학파 연구의 현황과 과제 발표_김세정(충남대), 토론_임흥태(성균관대)	
11:40-12:10	하린의 양명학 전승과 발전 발표_천병돈(대진대), 토론_박현정(성균관대)	
12:10-12:20	휴식	
12:20-13:10	종합토론	좌장_김용재(성신여대)
한국여성철학회		
사회적 제문제에 대한 여성주의적 응답		사회_양선숙(경북대)
10:00-10:40	불교문학 속에서 생물학적 여성 / 원혜영(동국대)	
10:50-11:30	수치심과 인간의 존엄성에 관한 소고 - 친족성폭력 피해 여성의 경험을 바탕으로 / 전해정(국민대)	
11:40-12:20	또 하나의 빈곤 포르노그래피 : 소녀와 생리대 김양희(국민건강보험공단 건강보험연구원)	
한국윤리학회		
21세기 한국 사회의 윤리적 가치 분석과 구성		사회_강준호(경희대)
10:00-10:05	개회사	
10:05-10:45	낙태의 도덕성에 대한 덕 윤리의 관점 / 장동익(공주교대), 논평_최훈(강원대)	
10:45-11:25	The Prudential Rationality of Near-Term Bias 최동용(캔사스대), 논평_김상수(UC Santa Barbara)	
11:25-11:40	휴식	
11:40-12:25	전파감염병에 대한 윤리학적 분석 / 강철(서울시립대), 논평_성창원(고려대)	
12:25-12:40	종합토론	
한국철학교육학회		
초중등 교육현장과 철학교육		
13:00-16:00	고등학교 철학 2015 개정 교육과정에 대한 비판적 논의 윤상철(경희여고), 논평_지혜인(중동고)	
	동양철학과 철학교육 / 이지(이화여대), 논평_박정원(이화여대)	
	'IAPC 어린이철학 모델에 관한 국내 연구'에 대한 진단과 비판 박현주(이화 철학연구소), 논평_강재린(이화 철학연구소)	
	철학의 방법, 철학함의 과정: 립만의 창의적 사고를 중심으로 한지희(명지대), 논평_박연숙(송실대)	
	엔택트 시대의 철학 교사 : 한국철학교사모임 / 윤이정(안양예고)	

한국해석학회	
21세기 한국의 해석학적 인간학 사회_곽영윤(강원대)	
10:00-13:00	수양(Bildung)개념과 자기이해 - 디지털 시대에서 해석학적 치유과정의 정립을 위하여 / 손세훈(부산외대)
	정조의 수양론과 인간학적 의미 / 안효성(대전대)
	가브리엘 마르셀의 실존적 인간이해 / 이희용(서울신학대)
한국환경철학회 장소_토즈(서울교대점) K-6	
인간+공간+건축의 환경철학적 모색 사회_김완구(충북대)	
10:00-10:10	개회(학회장 인사말)
10:10-10:40	생태 친화적 거주함으로서 건축 공간과 산책 / 김민수(동서울대)
10:40-11:10	'같이 따로' 사는 삶에서 '따로 같이' 사는 삶으로의 변화 - 스마트시티의 의미에 대한 철학적 연구 / 변순용(서울교대)
11:10-11:40	자연과 인간의 공존에 대한 생태철학적 이해 : 자원, 공간과 지속가능성 정민걸(공주대)
11:40-12:00	중간 휴식
12:00-13:00	종합토론 (발표자+김명식(진주교대)+노희정(광주교대))
13:00	폐회

차 례

인사말 / iii

[10월 23일(금)]

한국분석철학회

발표 1	왜 ‘나의 마누라’가 아니라 ‘우리 마누라’인가? 최성호	5
논평	‘우리 마누라’와 비진리조건적 의미’에 대한 논평문 이풍실	21
발표 2	The Identity of Proofs and the Criterion for Admissible Reductions Seungrak Choi	29
논평	“The Identity of Proofs and The Requirement of a Proper Reduction”에 대한 논평 소한중	57
발표 3	Knowledge as Objectively Justified Belief Byeong Deok Lee	61
논평	이병덕 교수의 Knowledge as Objectively Justified Belief에 대하여 오성	83
발표 4	부작위는 행위와 도덕적으로 동일한 지위에 있는가 -대안가능성과 이중결과원리의 검증도구로서의 적합성- 김효은	87
논평	“부작위는 행위와 도덕적으로 동일한 지위에 있는가”에 대한 토론문 신상규	105

[10월 24일(토)]

대동철학회

발표 1	조선 성리학의 철학적 가치 유원기	115
------	-----------------------	-----

발표 2	圖像學에서 본 이황의 『聖學十圖』와 그 해석의 생명력-매체철학의 논법에 착안하여- 김연재	133
발표 3	최석정 『구수락』의 이진법 독법 배선복	147
발표 4	퇴계의 향약 분석을 통한 조선시대 자치조직 연구 김상현	159

동양철학연구회 / 한국유교학회 / 한국철학사연구

발표 1	“재밋는” 유교, 유교와 게임의 만남 안승우	191
논평	「“재밋는” 유교, 유교와 게임의 만남」에 대한 토론문 오준원	213
발표 2	유교적 관점에서 본 밀레니얼 세대의 직업윤리와 직업의식 강보승	215
논평	유교적 관점에서 본 밀레니얼 세대의 직업윤리와 직업의식 전지열	235
발표 3	학이시습(學而時習), ‘사랑하기’를 통한 치유와 성장의 날갯짓 -자기애(自己愛)와 방탄소년단(BTS) 열풍을 중심으로- 윤민향	239
논평	「학이시습(學而時習), ‘사랑하기’를 통한 치유와 성장의 날갯짓」에 대한 논평 윤부한	257

새한철학회

발표 1	‘나’를 찾는 행복 황종환	263
발표 2	『화엄경』 「십정품」에 담긴 철학적 사유-10가지 삼매를 중심으로- 강기선	281
발표 3	중국어이슬람의 사상적 특징-유학적 이슬람- 권상우	291

서양근대철학회

- 발표 1** 데카르트 철학에서 자연과 목적의 문제-“감각들 속에는 신의 능력과 선함을 보여 주지 않는 것은 하나도 없다(AT VII, 87, 26-28)”에 대한 해석의 한 시도- 이재훈 309
- 발표 2** 흄의 인과 추론의 정당성 문제-『인성론』 1권과 『인성론』 2권, 3권 사이의 관계를 중심으로- 김병재 329
- 발표 3** 칸트의 인과성 개념에 대한 비판적 고찰 박제철 343

인도철학회 / 인도철학불교학 연구소

- 발표 1** 바르뜨리하리(Bhartrhari)의 언어론-스포따(sphoṭa)를 중심으로- 박수영 369
- 발표 2** 유식의 구조와 언어의 상관성에 관한 고찰 김재권 381
- 발표 3** 인도 사상에 있어 표현할 수 없는 것에 대한 언어표현 박기열 393
- 발표 4** 시의 논리-『카비야 알람카라』(*Kāvyaḷamkāra*) 제5장을 중심으로- 류현정 405

한국기독교철학회

- 발표 1** 평화주의는 인식론이다-존 요더의 평화주의 인식론 탐구- 김기현 425
- 발표 2** 앙리 베르그손이 만난 하나님-이성적 신과 직관의 신- 양선진 451
- 발표 3** 한국의 기독교철학자 100년(1)-일제강점기의 기독교철학자- 최태연 469

한국니체학회

발표 1	클로소프스키의 니체 해석 임건태	485
논평	임건태 박사의 「클로소프스키의 니체 해석」에 대한 논평 강용수	509
발표 2	니체의 실험철학과 철학적 의사-두 개념의 사상적 연관성의 해명을 중심으로- 이상범	515
논평	니체의 실험철학과 철학적 의사-두 개념의 사상적 연관성의 해명을 중심으로- 김주휘	529
발표 3	쇼펜하우어와 니체의 예술론에 대한 비판적 고찰-개체화의 원리를 중심으로- 남재민	531
논평	니체로부터 쇼펜하우어의 의지를 구하다?!-“쇼펜하우어와 니체의 예술론에 대한 비판적 고찰에 대한 논평: 개체화의 원리를 중심으로”를 읽고- 이주향	543
발표 4	니체 철학에서의 수치심 이선	547
논평	이선 선생님의 「니체 철학에서의 수치심」에 대한 토론문 김선희	569

2020 한국철학자연합대회

2020
UCCP

2020.10.23.(금)

2020 한국철학자연합대회

2020
UCCP

—
한국분석철학회

왜 ‘나의 마누라’가 아니라 ‘우리 마누라’인가?

최성호

경희대

초록

한국 분석철학계에서 진행된 ‘우리 마누라’에 관한 최근 논의는 한국어 ‘우리’의 용법과 의미에 대한 학술적 이해를 심화하는데 일조했다고 자부할만하다. 그 논의를 통하여 적어도 ‘우리 마누라’가 ‘우리의 마누라’와 동의어가 아니라는, ‘우리 마누라’에서 ‘우리’가 단순히 복수 일인칭 지표어의 소유격이 무표적으로 사용된 용례가 아니라는 점에 대해서는 폭넓은 공감대가 형성된 것으로 필자는 파악한다. 그럼에도 ‘우리 마누라’의 정확한 의미가 무엇인지, 그것의 용법을 기존의 의미론적 틀 안에서 어떻게 파악할지, 그 용법에서 화용론은 어떤 역할을 하는지 등은 여전히 오리무중이다. 본 논문의 목적은 그 주제들에 대한 언어철학적 탐구를 위한 하나의 초석을 마련하는 것이다. 필자는 먼저 『표준국어대사전』의 관련 항목을 참고하여 ‘우리’의 용법을 분류할 것이다. 이후의 논의는 ‘우리 마누라’의 형태를 갖는 구문에서 ‘우리’가 어떤 용법으로 사용되는지에 대하여 집중될 것인데, 그 용법을 올바르게 이론화하기 위하여 비진리조건적 의미(non-truth-conditional meaning) 개념을 고려할 필요가 있다는 것이 본 논문의 핵심 논지이다. 그렇게 비진리조건적 의미 개념을 도입하여 ‘우리 마누라’에 나타나는 ‘우리’의 의미를 면밀히 분석하고, 그를 통하여 ‘우리 마누라’의 의미론과 화용론에 대한 밑그림을 제시하는 것이 본 논문의 목적이다.

1. 우리 X 구문

한국어 화자들이 관형어 ‘우리’와 명사(구)의 결합으로 이루어진 구문을 상당히 폭넓게 사용한다는 것은 잘 알려진 사실이다 - 편의를 위하여 이를 ‘우리 X 구문’이라 명명하자. ‘우리 집’, ‘우리 나라’, ‘우리 남편’, ‘우리 찬호’, ‘우리 담임 선생님’, ‘우리 마누라’, ‘우리 서울시민

들’ 등이 그 예인데, 언어철학자나 언어학자들의 최근 연구는 한국어 화자들이 우리 X 구문을 사용하는 방식을 정교하게 이론화하는 것이 결코 만만치 않은 작업임을 여실히 증명한다. 우리 X 구문을 복수 일인칭 지표어의 소유격 구문으로 간주하는 것은 초견적으로(*prima facie*) 상당한 호소력을 지닌다. ‘우리’는 일인칭 지표어 ‘나’의 복수형이고 ‘우리 담임 선생님’은 ‘우리의 담임 선생님’의 준말이라는 가정하에서 ‘우리 담임 선생님’은 기본적으로 복수 일인칭 지표어의 소유격 구문이라는 것이 그 접근이다. 그러나 적어도 일부 우리 X 구문의 사례에서 이러한 초견적 접근은 매우 불합리한 결과를 발생시킨다.

대표적으로 한국어 화자가 자신의 부인을 가리키기 위하여 사용하는 ‘우리 마누라’라는 표현을 고려해 보자. 우리 X 구문에 대한 초견적 접근에 따르면 ‘우리 마누라’는 그 표준적인 용법에서 ‘우리의 마누라’의 준말이고, 그러한 한에 있어서 그것은 복수 일인칭 지표어의 소유격 구문이다. 그러나 필자는 (최성호 2017b)에서 ‘우리 마누라’가 ‘우리의 마누라’의 준말이라는 가정, 즉 ‘우리 마누라’와 ‘우리의 마누라’가 서로 동의어라는 가정을 받아들여지게 되면 마누라 공유 공동체의 존재를 긍정할 수밖에 없는데, 한국 사회에서 마누라 공유 공동체는 존재하지 않기에 그 가정은 거부되어야 한다는 것을 논증한 바 있다.¹⁾

좀 더 상세히 서술하자면 그 논증의 핵심 아이디어는 다음의 두 논제로부터 모순을 이끌어 낼 수 있다는 관찰에 근거한다.

동의어 논제. ‘우리 마누라’와 ‘우리의 마누라’는 동의어이다.

공동체 논제. ‘x가 우리의 마누라이다’는 오직 x를 공유하는 마누라 공유 공동체가 존재할 때에만, 즉 어떤 공동체가 있어서 x가 그 공동체의 구성원 각각의 마누라일 때에만 참이다.

이 두 논제로부터 모순이 도출된다는 사실은 이 두 논제 중 적어도 하나가 거짓이라는 것을 뜻한다. 그런데 필자는 앞서 언급한 논문에서 공동체 논제가 참이라고 볼 유력한 논거를 제시하였고, 그 논거를 받아들일 때 동의어 논제는 거부되어야 한다. 하여 필자는 ‘우리 마누라’를 ‘우리의 마누라’의 동의어가 아니라는, ‘우리 마누라’가 복수 일인칭 지표어의 소유격 구문이 아니라는 입장을 견지하고 있다.

이러한 필자의 논의로부터 따라 나오는 결론은 우리 X 구문이 복수 일인칭 지표어의 소유

1) 정대현(2009)과 강진호(2010)는 동의어 논제에 대한 자신들의 입장을 명시적으로 밝히지는 않지만 그럼에도 그들이 동의어 논제를 수용하는 것으로 그들의 입장을 해석하는 것이 가장 자연스럽다고 필자는 믿고 있다. 그런 점에서 동의어 논제와 관련하여 그들은 필자와 상반된 입장을 개진하고 있다는 것이 그들에 대한 필자의 해석이다. 정대현과 강진호에 대한 필자의 비판은 (최성호 2016)과 (최성호 2017a)를 참고하라.

격 구문이라고 보는 그에 대한 초견적 접근이 모든 우리 X 구문의 사례들에 대하여 예외 없이 성립하지는 않는다는 것이다. 위에서 살펴본 바와 같이 필자는 우리 X 구문에 대한 초견적 접근이 ‘우리 마누라’에 대해서는 성립하지 않는다는 것을 논증하였다. 그러나 이러한 필자의 논증으로부터 우리 X 구문에 대한 초견적 접근이 성립하는 사례가 아예 존재하지 않는다는 결론이 곧장 따라 나오지는 않는다. ‘우리 담임 선생님’에서 ‘우리’를 복수 일인칭 지표어의 소유격이 무표적으로 표기된 경우로 해석하고 나아가 ‘우리 담임 선생님’을 ‘우리의 담임 선생님’의 준말로 해석하는 것은 꽤 그럴듯하다.

실제로 ‘우리 마누라’의 경우에서와 달리, ‘우리 담임 선생님’의 경우 다음의 두 논제로부터 모순이 도출되지 않고, 그에 따라 그 두 논제를 동시에 참으로 수용하는 것이 얼마든지 가능하다는 것이 필자의 판단이다.

동의어 논제*. ‘우리 담임 선생님’과 ‘우리의 담임 선생님’은 동의어이다.

공동체 논제*. ‘x가 우리의 담임 선생님이다’는 오직 x를 공유하는 담임 선생님 공유 공동체가 존재할 때에만, 즉 어떤 공동체가 있어서 x가 그 공동체의 구성원 각각의 담임 선생님일 때에만 참이다.

이는 비록 우리 X 구문에 대한 초견적 견해가 ‘우리 마누라’에 대해서는 성립하지 않지만 그럼에도 그로부터 그 견해가 ‘우리 담임 선생님’에 대해서도 성립하지 않는다는 결론이 곧장 따라 나오지는 않는다는 것을 뜻한다. 우리 X 구문에 대한 이론적 접근이 ‘우리 마누라’나 ‘우리 담임 선생님’과 같은 개별 사례에 따라 상당히 다른 양상을 띠 수밖에 없다는 것이다. ‘우리 마누라’와 같은 개별 사례에 대한 탐구의 결과를 모든 우리 X 구문의 사례로 곧장 일반화할 수 없다는 교훈을 이끌어낼 수 있는 대목이다.

사실 필자는 우리 X 구문의 모든 사례들에 대하여 참되게 적용되는 언어철학적 이론을 일반적 수준에서 구축하는 것은 상당히 지난한 작업이 될 것으로 추측한다. 논문 한두 개로 완성할 수 있는 작업이 아니라 장기적인 연구프로젝트로 접근해야 하는 작업이라는 것이다. 그러한 장기적 연구프로젝트의 전망을 염두에 두며 본 논문은 ‘우리 마누라’에 대한 한국어 화자들의 표준적 용법을 분석하는 것에 집중할 것이다. 그렇게 논문의 초점을(우리 X 구문 일반이 아닌) ‘우리 마누라’라는 개별 사례에 집중하는 목적은 우리 X 구문과 관련한 복잡다단한 고려사항들 중 일부를 배제함으로써 이하의 논의를 단순화하는 것이다. 그러한 집중과 단순화의 전략은, 비록 우리 X 구문에 관한 다양한 고려사항들을 모두 포섭하지 못한다는 결함

을 갖지만, 그럼에도 복잡한 실타래와 같은 이 주제를 조심스럽게 그리고 엄밀하게 접근하는 탐구의 전략으로는 안성맞춤이다.

2. ‘우리 마누라’와 ‘나의 마누라’, ‘문재인’과 ‘문재앙’

‘우리 마누라’와 같은 일상어에 관한 탐구에서 연구자들이 특히 조심해야 할 사항 한 가지를 지적하며 논의를 시작해 보자. 그것은 어떤 문구가 우리들의 귀에 얼마나 자연스럽게 들리는지, 일상적인 담화에서 얼마나 자주 사용되는지에 대한 즉각적이고 직관적인 판단으로부터 고도의 주의가 요망되는 이론적인 주장을 성급하게 이끌어내는 것이다. 그런 즉각적이고 직관적인 판단은 어디까지나 언어적 현상에 대한 이론화 과정에서 고려되어야 하는 최초의 원재료일 뿐, 그것으로부터 이론적으로 유의미한 결론을 이끌어내기 위해서는 그 원재료 자체에 대한 엄밀한 분석과 평가가 선행되어야 한다. 그러한 분석과 평가 없이 원재료로부터 부주의하게 이론적 주장으로 나아갈 때 연구자는 수많은 혼란과 실수에 직면할 수밖에 없다. 문제는 이런 부주의가 ‘우리 마누라’에 관한 기존의 문헌에서도 종종 엿보인다는 것이다.

‘우리 마누라’의 가장 적절한 영어번역이 ‘My wife’라는 사실에 이견을 달기 힘들다. 실상 ‘우리 마누라’는 다른 누구도 아닌 바로 화자의 마누라를 가리키기 위해서 사용되는 만큼, 주어진 발화 맥락에서, ‘우리 마누라’와 ‘나의 마누라’는 동일한 지시체를 갖는 두 한정기술구(definite descriptions)이다. 이 지점에서 자연스럽게 떠오르는 질문이 그들이 지시체를 공유할 뿐만 아니라 언어적 의미까지 공유하는지 여부이다. ‘우리 마누라’와 ‘나의 마누라’가 동일한 대상을 가리킬 뿐만 아니라 서로 동의어의 관계를 갖는지 여부이다. 이에 대해서 그들은 분명 동일한 지시체를 갖는 두 한정기술구이지만 그럼에도 그들의 관계가 동의어의 관계까지는 아니라고 보는 입장이 상당히 직관적이다.²⁾ 그리고 이러한 입장은 한국어 화자들이 다음의 두 문장에 대하여 상당히 상이한 반응을 내놓는다는 사실에 의해 정당화되곤 한다.

1a. 우리 마누라가 요즘 바람을 피운다.

#1b. 나의 마누라는 요즘 바람을 피운다.

2) 이와 관련한 일부 언어학자들의 서술은 상당히 혼란스럽다. 가령 김혜령(2016, pp. 282-283)은 ‘우리 개’와 ‘내 개’의 사례에 대하여 “‘우리’와 ‘내’가 교체되는 것이 가능하지만 ‘우리 개’와 ‘내 개’는 그 의미가 같지 않다”고 서술한다. 이러한 서술은 상당히 혼란스러운데, 그것은 만약 ‘우리’와 ‘내’를 서로 교체하는 것이 가능하다면 ‘우리 개’와 ‘내 개’의 의미가 같아야 할 것이기 때문이다. ‘우리 개’와 ‘내 개’의 의미가 서로 상이함에도 불구하고 그 표현들에서 ‘우리’와 ‘내’가 교체 가능하다는 김혜령의 주장이 일말의 설득력을 지니기 위해서는 김혜령이 ‘교체 가능하다’는 용어를 통해서 무엇을 뜻하는지를 좀 더 명확히 밝혀야 할 것이다.

(1a)는 문법적으로 올바른 뿐만 아니라 한국어 화자들의 귀에 아주 자연스럽게 들리는, 일상적 담화에서 상용적인 문장이다. 이에 반해 (1b)는, 비록 문법적으로는 올바르고 나아가 그 진리조건이 무엇인지도 분명하지만, 그럼에도 한국어 화자들의 귀에 부자연스럽게 들리고 그래서 그런지 일상적 담화에서 거의 사용되지 않는 문장이다.³⁾

그런데 많은 국어학자나 언어학자들이 관찰하는 바와 같이 이것이 ‘우리 마누라’에서만 나타나는 독특한 현상이 아니다(윤재학 2003, 2009; 이한규 2007; 김세령 2016). 우리 X 구문에서 ‘우리’를 ‘내’로 교체할 때 기존과 동일한 지시체를 갖지만 그럼에도 한국어 화자의 귀에 상당히 부자연스럽게 들리는, 그래서 일상적 담화에서 거의 사용되지 않은 한정기술구를 갖는 경우는 실상 부지기수이다. 비록 새롭게 획득된 한정기술구가 문법적으로 올바른 뿐만 아니라 상당히 명료한 진리조건적 의미를 가짐에도 그러하다 - 진리조건적 의미 개념에 대해서는 이하에서 상세한 설명이 제시될 것이다. 몇 가지 사례를 더 고려해 보기로 하자.

2a. 우리 학교는 다음 주부터 방학이야.

#2b. 내 학교는 다음 주부터 방학이야.

3a. 우리 형님이 이번에 결혼합니다.

#3b. 내 형님이 이번에 결혼합니다.

‘우리 학교’나 ‘우리 형님’은 모두 우리 X 구문의 사례들이다. 이때 한국어 화자에게 (2a)는 문법적으로 올바른 뿐만 아니라 상용적인 문장인 반면 (2b)는 비록 문법적으로는 올바르지만 그럼에도 비문이다. 이와 동일한 사항이 (3a)와 (3b)에 대해서도 성립한다.

문제는 일부 언어이론가들이 이런 현상에 대한 단순 관찰로부터 우리 X 구문에 대하여 과도한 결론을 이끌어낸다는 사실에 있다. 이번 절의 서두에서 언급했던 부주의가 이 지점에서 나타난다는 것이다. 윤재학(2003, pp. 11-12, p. 15, p. 22; 2009, p. 100-101)은 (1a)가 상용적인데 반하여 (1b)가 상용적이지 않다는 사실로부터 곧장 ‘우리 마누라’와 ‘나의 마누라’가 동의어가 아니라는 결론을 추론한다. 이와 유사한 결론을 (2a)와 (3a)가 자연스러운 한국어 문장인데 반하여 (2b)와 (3b)가 비문이라는 사실로부터도 이끌어낸다. 그 사실로부터 ‘우리 학교’와 ‘내 학교’가 동의어가 아니라는, ‘우리 형님’과 ‘내 형님’이 동의어가 아니라는 결론을 추론해 내는 것이다. 아래에서 자세히 살펴보겠지만 이러 결론을 발판 삼아 윤재학은 우리 X

3) 김준걸(미발표원고)는 (1b)가 ‘적절한 상황에서’ 한국어 사용자에게 자연스러운 문장이 수 있다고 관찰한다. 그러나 그러한 관찰에도 불구하고 김준걸 역시 많은 맥락에서 (1a)는 자연스럽게 사용될 수 있지만 (1b)는 자연스럽게 사용될 수 없다는 것을 받아들인다.

구문에서 ‘우리’는 복수 일인칭 소유격의 무표적 표현이 아니라 화자가 ‘X’에 의해 지칭되는 개체와 동일한 내집단을 형성한다는 사실을 알리는 지표 역할을 수행한다고 제안한다.

그러나 이러한 추론은 상당히 성급한 것이 아닐 수 없다. (1b), (2b), (3b)가 비문이고 (1a), (2a), (3a)이 비문이 아니라는 사실로부터 곧장 ‘우리 마누라’와 ‘나의 마누라’가, ‘우리 학교’와 ‘내 학교’가, ‘우리 형님’과 ‘내 형님’이 동의어가 아니라는 결론이 따라 나오지 않기 때문이다. 어떤 문장이 문법적으로 올바르고 나아가 상당히 명료한 진리조건을 가짐에도 불구하고 일상적 담화에서 자주 사용되지 않는 현상에 대해서는 다양한 언어내적인 혹은 언어외적인 설명이 가능하다. 한국어 화자의 귀에서 ‘옆집 총각이 취직했어’는 자연스러운 문장으로 들리는데 반하여 ‘옆집 결혼하지 않은 남자가 취직했어’는 그렇지 않다. 그러나 이로부터 ‘총각’과 ‘결혼하지 않은 남자’가 동의어가 아니라는 결론이 곧장 따라 나오지는 않는다. 두 표현이 서로 동의어임에도 불구하고 일상적 담화에서 한국어 화자들이 첫 번째 문장은 빈번하게 사용하지만 두 번째 문장은 그렇게 사용하지 않는 것에 대하여 화용론적 설명이 가능하기 때문이다. 두 번째 문장이 첫 번째 문장에 비하여 훨씬 간결하고 경제적이라는 사실은 그러한 화용론적 설명을 제공한다.

그런데 그러한 설명이 반드시 언어내적인 성격을 가질 필요도 없다. ‘이번 신규채용에 젊은 여성들이 많이 지원했습니다’는 아무런 하자가 없는 문장이지만 ‘이번 신규채용에 아가씨들이 많이 지원했습니다’는 청자의 눈살을 찌푸리게 만드는 문장이다. 그런데 두 문장 사이의 이러한 차이에도 불구하고 ‘젊은 여성’과 ‘아가씨’가 서로 동의어라는 점에 대해서는 의문의 여지가 없다. 그렇다면 두 문장 사이의 차이는 어디에서 비롯되는 것일까? 그것은 ‘아가씨’라는 용어가 대개 유흥업과 관련된 맥락에서 많이 사용되었다는 언어 사용의 역사에서 비롯된다. 그러한 언어 사용의 역사가 유흥업에 대한 일반인들의 거부감과 결합하여 일상적인 담화에서 ‘아가씨’의 사용을 기피하게 만드는 결과를 만든다. 여기서 두 문장 사이의 차이에 대한 설명은 한국인들의 사회 심리와 같은 언어외적인 요소를 포함하고 있다.

요컨대, 문장 S에 문법적으로 아무런 하자가 없음에도 불구하고 화자들이 S를 비문으로 간주하고, S를 부자연스럽다거나 혹은 어색하다고 여기며, 일상적 담화에서 S의 사용을 기피한다는 현상에 대해서는 다양한 언어내적인 혹은 언어외적인 설명이 가능하다. 이런 다양한 설명이 가능한 만큼 그 현상으로부터 곧장 과감한 의미론적 결론을 이끌어내는 것은 가능하지 않다. (1a), (2a), (3a)는 한국어 공동체에서 상용적인 문장인 반면 (1b), (2b), (3b)는 그렇지 않다는 관찰로부터 ‘우리 마누라’, ‘우리 학교’, ‘우리 형님’이 각각 ‘나의 마누라’, ‘내 학교’, ‘내 형님’과 동의어가 아니라는 결론으로 나아간 윤재학의 추론이 타당하지 않은 까닭이다.

전술한 바와 같이 ‘우리 마누라’와 ‘나의 마누라’는 동일한 대상을 가리킨다. 이와 관련하여 우리가 주의 깊게 살펴봐야 할 점이 하나 있는데, 그것은 그들이 세계에 우연적으로(contingently) 성립하는 사실 덕분에 동일한 대상을 가리키는 것이 아니라는 것이다. 가령 ‘아리스토텔레스’와 ‘알렉산더 대왕의 스승’이 동일한 인물을 가리키는 것은 마침 아리스토텔레스가 알렉산더 대왕의 스승이었다는 세계의 우연적 사실 덕분이다. 아리스토텔레스가 알렉산더 대왕의 스승이 아닐 수도 있었다. 그리고 아리스토텔레스가 알렉산더 대왕의 스승이 아닌 상황에서 ‘아리스토텔레스’와 ‘알렉산더 대왕의 스승’은 동일한 대상을 가리키지 않을 것이다. 마찬가지로 콰인(W.V.O. Quine 1951, p. 21)의 유명한 사례에도 성립한다. 마침 심장을 가지고 있는 모든 생명체가 신장을 가지고 있고, 또 그 역도 성립한다고 가정해 보자. 그때 ‘심장을 가지고 있는 생명체(creatures with a heart)’와 ‘신장을 가지고 있는 생명체(creatures with a kidney)’는 동일한 대상을 가리킬 것이다. 그리고 그 두 기술구가 동일한 대상을 가리키는 것은 세계에 우연적으로 성립하는 사실 덕분이다. 혹시라도 심장을 갖지만 신장을 갖지 않는 생명체 혹은 신장을 갖지만 심장을 갖지 않는 생명체가 존재한다면 두 기술구는 동일한 대상을 가리키지 못할 것이다.

그런데 여기서 우리가 유념해야 할 점은 ‘우리 마누라’와 ‘나의 마누라’가 동일한 대상을 가리키는 것이 세계에 우연적으로 성립하는 사실 덕분이 아니라는 것이다. 좀 더 구체적으로 말하자면, 화자가 ‘우리 마누라’를 적절히 발화하는 모든 현실적인 혹은 가능한 맥락에서 그것은 화자의 마누라를 가리키고, 그에 따라서 그것은 ‘나의 마누라’와 동일한 대상을 가리킨다. 그런 점에서 ‘우리 마누라’와 ‘나의 마누라’가 세계에 성립하는 우연적인 사실들과 무관하게 동일한 대상을 가리킨다. 그들이 동일한 대상을 가리키는 것이 필연적이라는 말이다. 그 둘 사이의 관계가 ‘아리스토텔레스’와 ‘알렉산더 대왕의 스승’ 사이의 관계 혹은 ‘심장을 가진 생명체’와 ‘신장을 가진 생명체’ 사이의 관계와는 근본적으로 다르다는 것을 시사한다.

그렇다고 ‘우리 마누라’와 ‘나의 마누라’가 동일한 대상을 가리키는 것이 크립키(Saul Kripke 1980)가 그의 『명명과 필연(Naming and Necessity)』에서 제시한 것으로 유명한 후험적 필연성(necessity a posteriori)의 사례가 될 수도 없다. 왜냐하면 유창한 한국어 사용자라면 누구나 ‘우리 마누라’와 ‘나의 마누라’가 동일한 대상을 가리킨다는 것을 알 것이기 때문이다. 이는 ‘우리 마누라’와 ‘나의 마누라’가 동일한 대상을 가리키는 것을 알기 위하여 그 두 표현의 의미를 아는 것 이상의 경험적 탐구가 요구되지 않는다는 것을 의미한다. 그들이 동일한 대상을 가리킨다는 것이 후험적이지 않다는 말이다. 그러므로 ‘우리 마누라’와 ‘나의 마누라’가 동일한 대상을 가리키는 것은 후험적 필연성의 사례가 아니다.

‘우리 মানুষ라’와 ‘나의 মানুষ라’가 세계에 우연적으로 성립하는 사실과 무관하게 그들이 발화되는 모든 맥락에서 동일한 대상을 가리킨다는 관찰, 그들의 의미를 파악하는 한 그들이 동일한 대상을 가리킨다는 것을 모를 수 없다는 관찰은 많은 이들의 성급한 부인에도 불구하고 그들이 사실은 동의어라는 주장에 커다란 힘을 실어 준다. ‘총각’과 ‘결혼하지 않은 남자’는 세계에 우연적으로 성립하는 사실과 무관하게 동일한 대상을 가리킨다. 뿐만 아니라 그들의 의미를 파악하는 이가 그들이 동일한 대상을 가리킨다는 것을 모르는 것은 불가능하다. ‘총각’과 ‘결혼하지 않은 남자’가 정확히 동일한 의미를 갖는 동의어라는 점에 비추어 그들 사이에 이런 관계가 성립한다는 것은 전혀 놀랍지 않다. 마찬가지로 ‘우리 মানুষ라’와 ‘나의 মানুষ라’에 대해서도 성립하는 것 아닐까? ‘우리 মানুষ라’와 ‘나의 মানুষ라’가 진정 세계에 우연적으로 성립하는 사실과 무관하게 그것이 발화되는 모든 맥락에서 동일한 대상을 가리킨다면, 그리고 그렇게 그들이 동일한 대상을 가리킨다는 것을 알기 위해 그들의 의미를 파악하는 것 이상이 요구되지 않는다면, 그 둘이 실상 동일한 의미를 지닌다고 결론으로 나아가는 것이 가장 합당하지 않을까?

이 지점에서 우리를 어리둥절하게 만드는 것은 진정 ‘우리 মানুষ라’와 ‘나의 মানুষ라’가 동일한 의미를 지니는 동의어라면, 그 둘을 사용할 때 관찰되는 어감에서의 차이는 도대체 어디에서 비롯하는가이다. 그리고 도대체 무엇 때문에 (1a), (2a), (3a)는 한국어 공동체에서 상용적인 문장인 반면 (1b), (2b), (3b)는 그렇지 않은지에 대하여 설명이 필요해 보인다는 것이다. 아래에서 상술하겠지만 ‘나의 মানুষ라’가 갖지 않는 어떤 부가적인 의미를 ‘우리 মানুষ라’가 가지고 있다는 것은 많은 학자들이 공유하는 관찰이다. 그렇다면 그 부가적인 의미의 성격은 무엇인가? 어떻게 ‘우리 মানুষ라’는 그런 부가적인 의미를 가짐에도 불구하고 ‘나의 মানুষ라’와 동일한 대상을 가리킬 수 있단 말인가? 이 질문들에 대한 상세한 답변을 위해서는 언어적 표현의 ‘의미’가 도대체 무엇인가라는 언어철학의 가장 근본적인 주제에 대한 성찰이 요구된다.

결론부터 말하자면, 필자는 우리가 일반적으로 언어적 표현의 의미라고 여기는 것에는 진리조건적 의미 요소와 비진리조건적 의미 요소가 있고, 그 둘을 명확히 구분하는 것이 ‘우리 মানুষ라’와 ‘나의 মানুষ라’의 관계에 대한 명료한 이해를 위한 첫 출발점이 되어야 한다고 믿고 있다. 필자는 기본적으로 언어 안에서 표현 e를 어떻게 사용하는 것이 적절한지에 관한 규약화된 규칙(conventionalized rules of use)이 존재하고 나아가 e가 그 규칙에 관한 정보를 부호화(encoding)된 형태로 저장하고 있을 때 그 정보를 표현 e의 의미로 간주할 것이다. 이러한 e의 의미 중 일부는 e를 포함하는 정언문장의 진리조건에 영향을 미친다. 이때 표현 e의 진리

조건적 의미는 e를 포함하는 정언문장의 진리조건에 e가 기여하는 바에 의하여 표상된다. 여기서 중요한 논점은 e의 의미가 모두 진리조건적 의미인 것은 아니라는 것이다. 표현 e의 의미를 형성하는 구성 요소 중에는 e를 포함하는 문장의 진리조건과 무관한 것들이 존재한다. 것처럼 표현 e의 규약화된 사용 규칙에 관한 정보 중에서 e의 진리조건적 의미를 배제한 나머지를 필자는 ‘비진리조건적 의미’라 명명할 것이다.

이렇게 의미의 두 가지 요소를 구분할 때 ‘우리 마누라’와 ‘나의 마누라’ 사이의 관계는 두 고유명 ‘문재인’과 ‘문재양’ 사이의 관계와 유사하다고 필자는 제안한다. ‘문재인’과 ‘문재양’ 모두 문재인을 가리키는 고유명이다. 만약 고유명의 진리조건적 의미가 그것의 지시체라고 보는 밀주의(Millianism)가 옳다면, ‘문재인’과 ‘문재양’은 그들을 포함하는 문장의 진리조건에 기여하는 바가 정확히 동일하다. 그들은 문재인이라는 동일한 대상을 가리키고, 그 문재인이 그들의 진리조건적 기여에 해당한다. 뿐만 아니라 그들은 크립키가 정의한 바의 고정지시어(rigid designators)인 만큼 모든 가능 세계에서 동일한 대상을 가리킨다. ‘문재인’과 ‘문재양’이 동일한 대상을 가리킨다는 것이 필연적으로 성립한다는 말이다. 마지막으로 ‘문재인’과 ‘문재양’의 의미를 아는 화자라면 그 둘이 동일한 대상을 가리킨다는 것을 모를 수 없다. 이는 그 둘이 동일한 대상을 가리킨다는 것이 선택적임을 뜻한다. 이러한 고려로부터 ‘문재인’과 ‘문재양’이 동의어라는 결론이 일견 자명하게 따라 나오는 것처럼 보인다. 그러나 필자는 이러한 결론을 거부한다.

앞 문단의 고려에서 따라 나오는 것은 단지 ‘문재인’과 ‘문재양’이 정확히 동일한 진리조건적 의미를 갖는다는 것뿐이다. 그러나 앞서 지적한 바와 같이 언어적 표현의 의미는 진리조건적 의미 요소와 함께 비진리조건적 의미 요소로 구성된다. 그리고 설사 ‘문재인’과 ‘문재양’이 동일한 진리조건적 의미를 갖는다 하더라도 그들이 상이한 비진리조건적 의미를 갖는 것이 얼마든지 가능하다. 실제로 ‘문재인’과 ‘문재양’은 그들의 비진리조건적 의미에서 중요한 차이를 갖는다. ‘문재양’의 의미에는 화자가 문재인을 경멸하는 태도를 지닌다는, ‘문재인’의 의미에는 포함되지 않는 부가적인 정보가 포함되어 있고, 그것이 ‘문재양’의 비진리조건적 의미를 구성한다.

이러한 필자의 논점을 데이빗 카플란(David Kaplan 미발표원고)이 표현적 용어(expressives)를 논의하며 제시하였던 것과 유사한 논증의 형태로 형식화해 보기로 하자.

<문재양-문재인 논증>

문재양은 2020년 대한민국의 대통령이다

따라서, 문재인은 2020년 대한민국의 대통령이다

<문재인-문재양 논증>

문재인은 2020년 대한민국의 대통령이다

따라서, 문재양은 2020년 대한민국의 대통령이다

이 두 논증의 차이는 전제와 결론의 위치에 있다. 첫 번째 논증의 전제가 두 번째 논증의 결론이고 두 번째 논증의 전제가 첫 번째 논증의 결론이다. 핵심 질문은 <문재양-문재인 논증>과 <문재인-문재양 논증> 중 무엇이 타당한 논증이고 무엇이 부당한 논증이냐는 것이다. 이에 대한 카플란의 답변은 <문재양-문재인 논증>은 타당한 논증이지만 <문재인-문재양 논증>은 타당하지 않은 논증이라는 것이다.⁴⁾ 아쉽게도 카플란 자신은 이러한 답변에 대하여 아무런 정당화도 제시하지 않는다. 두 논증의 타당성에 대하여 필자 자신은 본 논문에서 어떤 특정한 입장을 취하지 않을 것이다. 단지 독자들에게 비진리조건적 의미가 무엇인지를 설명하는 작업의 편의를 위하여 <문재양-문재인 논증>은 타당한 논증이지만 <문재인-문재양 논증>은 타당하지 않은 논증이라는 카플란의 입장을 잠정적으로 받아들일 것이다. 그 두 논증이 왜 타당성에서 차이를 보이는지에 대한 카플란의 설명에서 비진리조건적 의미가 어떻게 활용되는지 살펴봄으로써 비진리조건적 의미가 무엇인지에 대한 독자들의 이해를 돕는 것이 지금 필자의 주된 관심사이다.

<문재양-문재인 논증>의 타당성과 <문재인-문재양 논증>의 부당성은 무엇에서 비롯하는 것인가? 이에 대한 카플란의 답변은 ‘문재양’이 ‘문재인’보다 더 풍부한 의미를 지닌다는 관찰에 근거한다. (고유명의 진리조건적 의미에 대한 밑주의를 가정할 때) ‘문재인’은 단순히 문재인으로 표상되는 진리조건적 의미만을 갖는데 반하여 ‘문재양’은 그러한 진리조건적 의미와 함께 화자가 문재인에 대하여 부정적인 태도를 취하고 있다는 부가적인 비진리조건적 의미를 갖는다는 것이다. 그런 만큼 ‘문재양은 2020년 대한민국의 대통령이다’은 ‘문재인은 2020년 대한민국의 대통령이다’보다 더 많은 정보를 담고 있고, 그에 따라서 전자로부터 후자를 타당하게 이끌어낼 수 있지만 후자로부터 전자를 타당하게 이끌어낼 수 없다는 것이 카

4) 카플란이 그의 논문에서 실제 고려하는 것은 두 문장 ‘Kaplan was promoted’와 ‘The damned Kaplan was promoted’ 사이의 추론 관계이다. 카플란은 후자로부터 전자가 타당하게 따라 나오지만 전자로부터 후자가 타당하게 따라 나오지는 않는다는 입장을 밝힌다. 이러한 카플란의 입장은 필자가 제시한 <문재양-문재인 논증>과 <문재인-문재양 논증>에도 자연스럽게 적용될 수 있고 그 때 전자가 타당하지만 후자가 타당하지 않다는 입장을 얻게 된다.

플란의 제안이다.

전술한 바와 같이 본 논문에서 필자는 <문재양-문재인 논증>의 타당성과 <문재인-문재양 논증>의 부당성에 대한 카플란의 판단과 관련하여 시시비비를 가리지 않을 것이다. 그럼에도 필자는 ‘문재인’과 ‘문재양’의 의미에 대한 카플란의 분석은 기본적으로 옳다고 믿는다. 카플란의 분석이 두 논증의 타당성 혹은 부당성에 대한 그의 판단까지 정당화하는지는 분명치 않지만 그럼에도 그 분석 자체는 옳다는 것이 필자의 생각이다.

필자는 이와 정확히 동일한 논점이 ‘우리 마누라’와 ‘나의 마누라’에 대해서도 제시될 수 있다고 본다. 다음의 두 논증을 검토해 보자.

<우리-내 논증>

우리 마누라가 요즘 바람을 피운다.

따라서, 나의 마누라가 요즘 바람을 피운다.

<내-우리 논증>

나의 마누라가 요즘 바람을 피운다.

따라서, 우리 마누라가 요즘 바람을 피운다.

<우리-내 논증>과 <내-우리 논증> 중 무엇이 타당하고 무엇이 부당한가? 이에 대하여 앞서 <문재양-문재인 논증>과 <문재인-문재양 논증>에 대하여 카플란이 개진한 것과 유사한 입장을 개진하는 것이 가능하다. ‘우리 마누라’가 ‘나의 마누라’보다 더 풍부한 비진리조건적 의미를 갖는다는 근거에서 <우리-내 논증>은 타당하지만 <내-우리 논증>은 타당하지 않다는 결론을 이끌어내는 것이 그것이다. 그 입장에 따르면 ‘우리 마누라’와 ‘나의 마누라’는 그들의 지시체를 결정하는 진리조건적 의미에 있어서는 정확히 동일하고, 바로 그런 이유에서 그들은 선형적으로 그리고 필연적으로 동일한 대상을 가리킨다. 이처럼 ‘우리 마누라’와 ‘나의 마누라’는 진리조건적 의미를 공유하지만 비진리조건적 의미에 있어서 그들 사이에 중요한 차이가 존재하고, 그 차이로 인하여 <우리-내 논증>은 타당하지만 <내-우리 논증>은 타당하지 않다는 것이 그 입장의 핵심 아이디어이다. 비록 ‘문재양’과 ‘문재인’이 동일한 진리조건적 의미를 갖지만 ‘문재양’이 ‘문재인’보다 더 풍부한 비진리조건적 의미를 갖는다는 사실로부터 <문재양-문재인 논증>은 타당한 논증이지만 <문재인-문재양 논증>은 타당하지 않은 논증이라는 결론을 이끌어내는 것과 기본적으로 동일한 논리이다.

‘우리 마누라’와 ‘나의 마누라’ 사이의 관계가 ‘문재인’과 ‘문재앙’ 사이의 관계와 유사하고, 그 관계를 올바르게 분석하기 위해서는 진리조건적 의미와 비진리조건적 의미의 구분이 필요하다. 필자의 논제는 그 자체로 상세한 부연설명과 엄밀한 정당화를 요구하는 대담한 논제이고, 그 대담성만큼 그 요구에 응답하는 것은 결코 만만한 과제가 아닐 수 없다. 그러나 그러한 부연설명이나 정당화보다 먼저 선행되어야 할 것이 바로 한국어 ‘우리’와 관련한 복잡다단한 논의의 지형에서 필자의 논제가 어떤 위치를 점하는지를 명료하게 밝히는 것이다. 대부분의 한국어 사전에서 ‘우리’는 복수 일인칭 지표어로 정의되고 있다. 그러한 정의와 이번 절에서 제시한 필자의 논제 사이의 관계를 명확히 정리하는 것 역시 결코 간단치 않은 작업이 될 것이다. ‘우리’의 용법에 대한 『표준국어대사전』의 분류를 검토하는 다음 절은 그러한 작업을 위한 단초를 마련해 줄 것이다.

참고문헌

- 강진호, 2010, “‘우리 마누라’의 의미”, 『철학적 분석』 21, pp. 153-64.
 _____, 2013, “백락주의와 ‘공유된 내용’의 문제”, 『철학적 분석』 27, pp. 31-57.
 김정남, 2003, “한국어 대명사 ‘우리’의 의미와 용법”, 『한국어 의미학』 13, pp. 257-74.
 김준걸, “‘내’의 공손한 표현으로서의 ‘우리’” 『철학적 분석』 43: pp. 91-114.
 김혜령, 2016, “1인칭 대명사의 의미 분석 - 의미 실현 조건 분석을 중심으로 -”, 『민족문화연구』 70, pp. 267-294.
 선저이 꾸마르, 2015, “한국어에 나타난 유교문화적 요소 일고찰”, 『韓國學研究論文集』 4, pp. 323-38.
 윤재학, 2003, “단수적 용법의 ‘우리’”, 『언어와 정보』 7, pp. 1-30.
 _____, 2009, “소유의 의미유형: 한, 영 소유구문의 의미차이”, 『언어와 정보』 13, pp. 93-125.
 이한규, 2007, “한국어 대명사 ‘우리’”, 『담화와인지』 14, pp. 155-78.
 정경옥, 2005, “한국어에 있어서의 ‘우리’의 사용에 대하여”, 『한국어교육』 16, pp. 405-22.
 정대현, 2009, “‘우리 마누라’의 문법”, 『철학적 분석』 20, pp. 69-83.
 최성호, 2016, “‘우리 마누라’와 험티덤편티 문제”, 『철학적 분석』 36, pp. 143-67.
 _____, 2017a, “강진호 교수에게 ‘우리 마누라’란 무엇인가”, 『철학 사상』 64, pp. 183-215.
 _____, 2017b, “‘우리 마누라’, ‘우리의 마누라’, 그리고 마누라 공유 공동체”, 『철학적 분석』 38, pp. 291-326.
 저자미상, 1999, “우리”, 『표준국어대사전』, 두산동아.
 Ameka, F., 1992, ‘Interjections: The Universal yet Neglected Part of Speech’, *Journal of Pragmatics* 18: 101-118.
 Bach, K., 1999, ‘The Myth of Conventional Implicature’, *Linguistics and Philosophy* 22: 327-366.
 _____, 2006, ‘The Top 10 Misconceptions about Implicature’, in B. J. Birner & G. Ward (eds.), *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*, Amsterdam: John Benjamins, pp. 21-30.

- _____, 2018, 'Loaded Words: On the Semantics and Pragmatics of Slurs', in D. Sosa (ed.), *Bad Words*, Oxford: Oxford University Press, pp.77-107.
- Blakemore, D., 1989, 'Denial and Contrast: A Relevance Theoretic Analysis of "but"', *Linguistics and Philosophy* 12: 15-37.
- _____, 2000, 'Indicators and Procedures: Nevertheless and but', *Journal of Linguistics* 36: 463-486.
- Cappelen, H. & Lepore, E., 2005, *Insensitive Semantics*, Oxford: Blackwell.
- Carston, R., 1996, 'Metalinguistic Negation and Echoic Use', *Journal of Pragmatics* 25: 309-30.
- Chapman, S., 1996, 'Some Observations on Metalinguistic Negation', *Journal of Linguistics* 32: 387-402.
- Coleman, J., 2006, 'Design Features of Language', in K. Brown, *Encyclopedia of Language & Linguistics* (2nd edition), Oxford: Elsevier, pp. 471-475.
- Cruse, D. A., 1986, *Lexical Semantics*, Cambridge: Cambridge University Press.
- Davidson, D., 1967, 'Truth and Meaning', *Synthese* 17: 304-323, reprinted in A. P. Martinich (ed.), *The Philosophy of Language*, Oxford: Oxford University Press, 1985, pp. 79-90.
- _____, 1977, 'The Method of Truth in Metaphysics', *Midwest Studies in Philosophy* 2: 244-254, reprinted in P. French, T. Uehling & H. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*, Minneapolis: University of Minnesota Press, 1979, pp. 294-304.
- _____, 1980, 'Reality without Reference', in M. Platts (ed.), *Reference, Truth and Reality: Essays on the Philosophy of Language*, London: Routledge & Kegan Paul, pp. 134-135.
- Frege, G., 1892/1997, 'On Sinn and Bedeutung', in M. Beaney (ed.), *The Frege Reader*, Malden: Blackwell, pp. 151-171.
- _____, 1897/1997, 'Logic', in M. Beaney (ed.), *The Frege Reader*, Malden: Blackwell, pp. 227-250.
- _____, 1918/1997, 'Thought', in M. Beaney (ed.), *The Frege Reader*, Malden: Blackwell, pp. 325-345.
- _____, 1953, *The Foundations of Arithmetic*, J. L. Austin (trans.), Oxford: Blackwell.
- Goldberg, A. E., 2016, 'Compositionality', in N. Reimer, *The Routledge Handbook of Semantics*, New York: Routledge, pp. 419-433.
- Grice, P., 1961, 'The Causal Theory of Perception', *Proceedings of the Aristotelian Society* 35: 121-52, reprinted in (Grice 1989). (본문의 페이지 번호는 재출간된 판본을 기준으로 작성되었다.)
- _____, 1969, 'Utterer's Meaning and Intention', *The Philosophical Review* 78: 147-177, reprinted in (Grice 1989). (본문의 페이지 번호는 재출간된 판본을 기준으로 작성되었다.)
- _____, 1975, 'Logic and Conversation', in P. Cole & J. Morgan (eds.), *Syntax and Semantics, Vol 3: Speech Acts*, New York: Academic Press, pp. 41-58.
- _____, 1989, *Studies in the Way of Words*, Cambridge, MA: Harvard University Press.
- Gutzmann, D., 2013, 'Expressives and Beyond: An Introduction to Varieties of Use-conditional Meaning', in D. Gutzmann & H. Gärtner (eds.), *Beyond Expressives: Explorations in Use-Conditional Meaning*, Leiden: Brill, pp. 1-58.

- Hockett, C., 1958, *A Course in Modern Linguistics*, New York: Macmillan.
- Hylton, P., 1993, *Russell, Idealism, and the Emergence of Analytic Philosophy*, Oxford: Clarendon Press.
- Janssen, T., 2001, 'Frege, Contextuality and Compositionality', *Journal of Logic, Language and Information* 10: 115-136.
- Jeshion, R., 2013, 'Expressivism and the Offensiveness of Slurs', *Philosophical Perspectives* 27: 232-259.
- _____, 2016, 'Slur Creation and Bigotry Formation: The Power of Expressivism', *Phenomenology and Mind* 11: 130-139.
- _____, 2018, 'Slurs, Dehumanization and the Expression of Contempt', in D. Sosa (ed.), *Bad Words*, Oxford: Oxford University Press, pp.77-107.
- Kaplan, D., 1989a, 'Demonstratives', in J. Almog, H. Wettstein & J. Perry (eds.), *Themes from Kaplan*, New York: Oxford University Press, pp. 481-563.
- _____, 1989b, 'Afterthoughts', in J. Almog, H. Wettstein & J. Perry (eds.), *Themes from Kaplan*, New York: Oxford University Press, pp. 565-614.
- _____, 1999, 'The Meaning of 'Ouch' and 'Oops'', Presented at the Cornell Conference on Context Dependency.
- Kim, J. B. & Sells, P., 2007, 'Korean Honorification: A Kind of Expressive Meaning', *Journal of East Asian Linguistics*, 16: 303-336.
- Kripke, S., 1980, *Naming and Necessity*, Oxford: Oxford University Press.
- Leech, G., 2014, *The Pragmatics of Politeness*, Oxford: Oxford University Press.
- McCawley, J., 1991, 'Contrastive Negation and Metalinguistic Negation', *Chicago Linguistic Society* 27: 189-206.
- Partee, B., 1984, 'Compositionality', in F. Landman & F. Veltman (eds.), *Varieties of Formal Semantics*, Dordrecht: Foris, pp. 281-311.
- Pelletier, F., 2001, 'Did Frege Believe Frege's Principle?', *Journal of Logic, Language and Information* 10: 87-114.
- Potts, C., 2007a, 'Conventional Implicatures, a Distinguished Class of Meanings', in G. Ramchand & C. Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*, Oxford: Oxford University Press, pp. 475-501.
- _____, 2007b, 'The Expressive Dimension', *Theoretical Linguistics* 33: 165-197.
- Predelli, S., 2013, *Meaning without Truth*. Oxford: Oxford University Press.
- Quine, W. V. O., 1951, 'Two Dogmas of Empiricism', *The Philosophical Review* 60: 20-43.
- Recanati, F., 2004, *Literal Meaning*, Cambridge: Cambridge University Press.
- _____, 2010, *Truth-Conditional Pragmatics*, New York: Oxford University Press.
- Richard, M., 2008, *When Truth Gives Out*, New York: Oxford University Press.
- Saka, P., 2007, *How to Think About Meaning*, Berlin: Springer.
- Sander, T., 2019, 'Two Misconstruals of Frege's Theory of Colouring', *The Philosophical Quarterly* 69: 374-392.

- Searle, J., 1979, *Expression and Meaning*, Cambridge: Cambridge University Press.
- _____, 1980, 'The Background of Meaning', in J. Searle, F. Kiefer & M. Bierwisch (eds.), *Speech Act Theory and Pragmatics*, Dordrecht: Reidel, pp. 221-232.
- Stanley, J., 2008, *Language in Context*, Oxford: Clarendon Press.
- Travis, C., 1975, *Saying and Understanding*, Oxford: Blackwell.
- Yule, G., 2006, *The Study of Language* (3rd edition), Cambridge: Cambridge University Press.

“우리 마누라’와 비진리조건적 의미’에 대한 논평문

이풍실

경희대

논문에서 “우리 마누라’와 비진리조건적 의미’에서 최성호 교수님(이하 존칭 생략)은 카플란(David Kaplan)과 집단 비하어 등에 대한 기존의 연구 성과를 수용하면서도 이에 대한 한층 더 심화된 이론적인 구체화를 바탕으로 ‘우리 마누라’의 의미라는 새로운 주제에 접근하고 있다. 기존의 성과들에 대한 비판적 수용과 이론적 구체화, 그리고 한국어 표현 ‘우리 마누라’가 갖는 특수한 성격에 대한 참신한 분석 모두 중요한 학문적 기여라고 나는 생각한다. 다만 최성호의 논증들과 관련하여 몇 가지 의문점들이 있는데, 논문에 대한 요약을 제시한 후에 이 의문점들 중 한 가지만 다루기로 하겠다.

이 논문에서 최성호는 ‘우리 마누라’의 의미를 분석하면서 이 표현이 두 종류의 의미를 갖는다고 주장한다. 하나는 화자의 여성 배우자인 사람을 지시체로 결정하는 의미이며, 다른 하나는 이 표현이 사용되는 문장의 진리조건과는 무관하지만 여전히 의미론적 내용에 해당하는 의미이다. 최성호는 ‘우리 마누라’가 이러한 의미구조를 갖는다는 주장을 정당화하는 논증에 주력하며, 후자의 의미가 정확하게 어떤 것인지를 만족스럽게 규명하는 것을 차후의 과제로 남겨둔다. 최성호는 ‘우리 x’의 형태의 구문의 의미에 대한 『표준국어대사전』의 세 가지 의미 구분을 검토하며, ‘우리 마누라’의 의미는 이 중 세 번째와 깊은 관련성을 갖는다고 판단한다. 세 번째 규정에 의하면, ‘우리 x’는 x가 지칭하는 것에 대해서 화자가 갖는 친밀감을 표현한다는 것이다. 또한 최성호는 ‘우리 x’가 화자와 x의 지시체가 특정한 내집단(in-group)에 속한다는 표지로 기능한다는 운재학의 입장 역시 이 세 번째 규정과 맥을 같이 하는 관점으로 간주한다. 최성호는 이러한 입장들의 취지를 발전시켜 보다 세련된 이론적 설명을 제안한다. 여기서 중요한 논점은 친밀감이나 내집단과 같은 요소가 ‘우리 마누라’의 비진리조건적 의미에 해당한다는 것이다.

최성호는 ‘우리 마누라’가 진리조건적 의미와 비진리조건적 의미를 모두 갖는다고 주장한다. 이 두 종류의 의미는 모두 한 표현의 의미론적 내용에 속한다. 어떤 표현 e의 의미론적 내

용은 e 속에 부호화되어 있는 언어적 규약이다. 의미론적 내용은 e를 사용하는 유창한 화자의 선택적 앎의 대상이다. ‘나’와 같은 지표어(indexical)를 예로 들자면, 유창한 화자는 ‘나’의 특성(character), 즉 ‘나’ 사례(token)는 그 사례를 발화하는 화자를 지칭한다는 것을 안다. ‘나’가 무엇을 가리키는지가 이렇게 결정된다는 것은 언어적 규약이기도 하다. 어떤 표현 e의 진리조건적 의미는 e가 사용되는 문장이 갖는 진리조건에 e가 기여하는 의미이고, 비진리조건적 의미는 진리조건에 기여하는 바는 없지만 여전히 언어적 규약에 의해서 부호화된 것이라는 점에서 의미론적 내용에 속한다. 앞서 예로 들었던 ‘나’의 특성은 진리조건적 의미에 해당한다.

최성호는 비진리조건적 의미를 두 종류로 구분한다. 하나는 ‘그러나’와 같은 표현이 그 예가 되는 바 그라이스(Paul Grice)를 통해서 익히 알려진 규약적 함축(conventional implicature)이고, 다른 하나는 표현적 내용/정보(expressive content/information)이다. ‘우리 마누라’의 의미와 관련해서 최성호가 주목하는 것은 후자이다. 우선 이 두 가지의 비진리조건적 의미는 전이불가능성(non-displaceability)와 독립성(independence)의 기준에 의해서 구별된다. 다수의 언어 표현들이 전달하는 정보는 해당 표현이 사용되는 시제를 변경한다거나, 표현을 조건문의 조건절 속에 넣는다거나, 간접화법의 보어절 속에 넣는 등 언어적 문맥을 전환함으로써 변경될 수 있다. 이것을 전이가능성이라고 한다. ‘철수는 밥을 먹는다’가 전달하는 정보는 ‘철수가 밥을 먹었다’ (과거 시제), ‘철수는 밥을 먹지 않았다’ (부정), ‘만일 철수가 밥을 먹는다면 영이도 밥을 먹을 것이다’ (조건), ‘철수가 밥을 먹는다고 영수가 말하고 있다’ (간접 화법) 등의 다양한 문맥의 변경을 통해 달라질 수 있다. ‘철수가 밥을 먹는다’는 화자가 참이라고 주장하는 정보, 즉 철수가 발화 시점에 밥을 먹고 있다는 정보를 전달하지만, 위에서 예시한 간접 화법에서는 영수가 현재 철수가 발화 시점에 밥을 먹고 있다는 주장을 한다는 정보를 전달하며 화자는 이러한 주장 자체를 받아들일 필요가 없다. 그런데 일군의 언어 표현들은 전이불가능성을 갖는다. 예를 들어, ‘우리 마누라가 바람을 피고 있어’가 ‘우리 마누라가 바람을 피고 있다고 영수가 말했어’라는 간접 화법 문맥 속에 나타난다고 하더라도 ‘우리 마누라’는 여전히 영수가 아닌 화자의 여성 배우자에 대한 것이다. 독립성은 의문문 발화에 대한 긍정 또는 부정의 대답을 통해서 의문문 속에서 사용되는 표현 e가 전달하는 특정한 정보에 대한 긍정이나 부정이 가능하지 않다는 사실과 관련된다. ‘우리 마누라가 바람을 피웠니?’에 대한 ‘예’ 또는 ‘아니오’의 대답은 질문자의 여성 배우자가 바람을 피웠는지에 대한 긍정이나 부정일 수 있을 뿐이지, ‘우리 마누라’가 전달하는 지시체 외적인 어떤 정보에 대한 긍정이나 부정일 수는 없다. 이러한 전이불가능성과 독립성에 대한 최성호의 논증은 기본적으로 비하

적 단칭어(slurring singular term)에 해당하는 ‘문재양’과 같은 표현이 전이불가능성과 독립성을 갖고 있음을 드러내는 사례들과 ‘우리 마누라’의 사례들이 동일한 패턴을 보인다는 관찰에 근거한다. (예: ‘문재양이 부동산 정책을 망쳤다고 영수가 말했다’ (간접화법이라고 해도 영수가 아닌 화자의 문재인에 대한 부정적, 비하적 태도가 표현됨), ‘문재양의 부동산 정책이 성공했니?’ (이 질문에 대한 긍정과 부정 모두 문재인의 부동산 정책의 성공 여부에 대한 것이지 ‘문재양’이 전달하는 비하적 의미에 대한 것일 수 없음.) 이러한 관찰은 또한 집단에 대한 비하어가 갖는 중요한 특성들과 관련된 기존 연구의 도움을 받고 있다. 이러한 종류의 표현들(예: ‘검둥이’, ‘짱깨’)은 문맥 변경에도 불구하고 집단에 대한 부정적 속성을 전달하는 정보를 그대로 유지한다. (예: ‘철수가 짱깨들이 전염병을 퍼뜨린다고 말했어. 나는 중국인들에 대한 편견이 없지만 말이야.’) 이것은 ‘문재양’의 경우에도 동일하다. (예: ‘철수가 문재양이 부동산 가격을 잡지 못한다고 말했어. 그러나 문재인 대통령은 훌륭한 분이시.’)

최성호의 주장에 의하면 ‘우리 마누라’는 ‘나의 마누라’와 동의어이며, 여기서 ‘우리’는 그 표면적 문법에도 불구하고 일인칭 복수 대명사가 아니다. ‘우리 x’ 구문이 화자와 청자를 포함한 집단을 가리키거나 (예: ‘우리 밥 먹으러 가자.’) 청자를 배제한 집단을 가리키는 경우 (예: ‘우리 경상도 사람들은 그런 경우에 ‘조심해라’라고 하지 않고 ‘단디해라’라고 말합니다.’)에 ‘우리’는 지표어로 기능하며 맥락을 입력값으로 받아서 일정한 내포 함수를 출력하는 특성을 갖는다. 최성호는 ‘우리 x’에서의 ‘우리’가 지표어로 기능하지 않으며 아예 진리조건적 의미를 결여하고 비진리조건적 의미만을 갖는 경우가 있다고 주장한다. 예를 들어, ‘우리 손흥민 선수가 골을 넣었습니다’는 ‘손흥민 선수가 골을 넣었습니다’와 정확히 동일한 진리조건을 가지며 여기서 ‘우리’는 손흥민에 대한 화자의 친밀감을 표현하는 비진리조건적 의미만을 전달한다. 그러나 ‘우리 마누라’의 경우에 ‘우리’는 표현적 내용 외에도 진리조건적 의미 역시 갖는다. ‘우리 마누라’의 진리조건적 의미는 ‘나의 마누라’와 동일한데, 여기서의 ‘우리’는 표면적인 문법적 형태에도 불구하고 일인칭 복수 소유격이 아니라 ‘나의’와 마찬가지로 일인칭 단수 소유격으로 기능한다. ‘우리 마누라=나의 마누라’는 분석적, 선형적, 필연적으로 참이며, 이러한 사실은 여기서 ‘우리’와 ‘나의’의 진리조건적 의미가 동일하다는 주장을 통해서 설명될 수 있다.

‘나의 마누라’와는 달리 ‘우리 마누라’는 비진리조건적 의미 또한 갖는다. ‘우리 마누라’의 비진리조건적 의미는 규약적으로 결정되지만 (따라서 의미론적 내용이지만) 여기서 ‘우리’는 ‘나’와 같은 지표어의 특성처럼 맥락이 주어지면 가능세계와 발화 시점에 상대적으로 외연을 결정하는 내포 함수를 산출하는 그런 의미 구조를 갖지는 않는다. 간단히 말해서 ‘우

리 마누라'에서의 '우리'는 지표어가 아니다. '문재앙'이 '문재인'과 정확히 동일한 진리조건적 의미를 갖지만 추가적으로 비하적 의미를 갖는 것과 마찬가지로 '우리 마누라'는 화자의 여성 배우자에 대한 일정한, 그리고 진리값 평가가 불가능한(not truth-apt) 표현적 정보를 전달한다. 정확히 이 정보가 어떤 것인지를 해명하는 것은 앞으로의 과제이지만, 적어도 이러한 표현적 정보가 『표준국어대사전』의 세 번째 의미 규정(친밀감 표현)이나 윤재학의 내집단 표지와 밀접하게 관련되어서 이해될 수 있을 것이라고 최성호는 생각한다.

'우리 마누라'는 비진리조건적 의미로서 표현적 내용/정보를 가지며, 이것은 화용론적인 함축이 아니라 의미론적 내용이라고 최성호는 주장한다. 최성호는 정확하게 이 표현적 내용이 무엇인지는 차후의 과제로 남기고 있지만 여전히 이 내용이 의미론적이라는 근거를 제시할 수 있다고 말한다. 최성호에 의하면 의미론적 내용의 핵심적 특성은 그것이 규약적으로 결정되어 있다는 것이다. '문재앙'의 표현적 정보인 문재인을 비하하는 내용이 의미론적인 내용인 이유는 이 표현을 제대로 이해하는 사람들은 모두 그러한 비하적 내용을 파악하기 때문이다. 적어도 최성호의 입장에 의하면 이러한 비하적 내용은 '문재앙'에 적용되는 규약이다. (물론 이러한 규약의 존재가 이 표현의 사용이 사회적, 정치적으로 문제가 없다는 것을 함축하지는 않는다.) 마찬가지로 '우리 마누라'를 제대로 이해하는 유창한 화자들은 이 구문이 전달하는 특정한 표현적 정보를 공통적으로 파악하고 있으며, 이 때문에 '우리 마누라'의 표현적 정보는 의미론적 내용이 된다. 실제로 최성호는 "'우리 마누라'의 의미론이 '문재앙'의 의미론과 동일한 구조를 지닌다"고 주장한다. (p.71)

당연히 이러한 주장의 궁극적인 정당화는 이 표현적 내용이 무엇인지를 해명함으로써 이루어져야 할 것이다. 이미 지적했듯이, 친밀성 이론이나 내집단 이론을 거론하는 것 외에 최성호 자신이 정확하게 이 표현적 정보가 무엇인지를 해명하고 있지는 않다. 이 때문에 최성호가 앞으로의 과제로 남겨둔 부분을 따지는 것이 적절하지 않을 수도 있겠으나, 개인적으로 다음과 같은 우려가 있다. 무엇보다도 '문재앙'이나 '검둥이' 같은 비하어들에 대해서는 유창한 언어 사용자들 사이에 그 표현적 정보가 무엇인지에 대한 정확한 이해가 공유되고 있는 것 같다. (개인/집단에 대한 비하) 그러나 '우리 마누라'의 경우는 그렇지 않아 보인다. 최성호는 친밀성 이론이나 내집단 이론을 들어서 일정하게 공유되는 표현적 내용의 존재를 받아들일 수 있다고 생각하지만, 과연 그러한지 의문이다. 만일 '우리 마누라'의 표현적 내용이 규약적으로 결정되지 않는다면, 더 나아가 모든 발화에서 고정적인 표현적 내용이 존재하지 않는다면 이러한 내용이 의미론적 내용이라는 최성호의 주장을 받아들이는 것은 쉬운 일이 아니다. 적어도 내가 보기에 '우리 마누라'의 모든 사용에 공통적인 규약적인 표현적 내용/정

보가 있는 것 같지 않으며, ‘우리 마누라’는 맥락에 따라서 다양한 비진리조건적 의미를 전달할 수 있는 것 같다. 이런 점에서 ‘문재양’과 ‘우리 마누라’는 의미구조가 다른 것으로 보인다.

최성호 자신이 사용하는 ‘우리 마누라가 요즘 바람을 피운다’를 생각해보자. 맥락과 무관하게 화자는 언제나 ‘우리 마누라’를 사용해서 자신의 배우자에 대한 친밀감을 표현하는가? 발화자가 자신의 마누라를 지극히 증오하면서 이 문장을 발화하는 경우를 상상하는 것은 어렵지 않다. 화자나 청자 모두 여기서의 ‘우리 마누라’를 친밀감의 표현이라고 생각하지는 않을 것이다. 이런 경우에는 친밀감이 아니라 단순히 ‘나의’가 표현하는 소유격을 강조하는 것 이상의 의미가 없어 보인다. 그리고 배우자에 대한 적대감을 표현하는 사람이 ‘우리 마누라’를 사용하는 것이 반드시 부적절한 것인지는 분명하지 않다. (예: ‘철수야 네 마누라가 아니라 우리 마누라랑 영수가 바람을 피고 있다고!’ ‘우리 마누라가 이제는 열 두 번째 애인을 만 들었군!’)

최성호의 기준에 따르면, 어떤 표현의 규약적인 내용이 있다면 그것에 부합하지 않는 해당 표현의 사용은 유창한 화자들에게 그러한 사용이 부적절하다는 직관을 낳아야 할 것으로 보인다. ‘문재양’과 같은 비하어의 경우에는 그 부적절성의 파악이 매우 분명하다. (여기서의 부적절성은 비하어 자체가 도덕적, 정치적으로 그 사용이 용인되지 않는다는 점에서의 부적절성과는 다른 개념이다.) 예를 들어, 문재인 지지자가 ‘문재양께서 반드시 부동산 가격을 잡아 주실꺼야’라고 말한다는 이는 명백히 부적절하다. 그러나 ‘우리 마누라’의 경우는 이렇게 명백한 직관이 없는 경우들이 있는 것 같다. 다음을 생각해보자. 최성호는 친밀감 이론을 지지하는 사례로서, 이혼남이 자신의 전 부인을 ‘우리 마누라’라고 부르지 않는다는 점을 든다.

어떤 의미로 위장결혼을 한 부부가 있을 때 남편이 “우리 마누라와 나는 법적으로만 부부이지 사실 남이나 마찬가지로입니다”라고 말한다면 그는 ‘우리 마누라’라는 구문을 잘못 사용하고 있다는 지적을 받기 십상이다. (p.72)

사실 나는 위의 발화가 ‘우리 마누라’의 잘못된 사용인지 잘 모르겠다. 수시로 바람을 피우는 배우자에 대한 극도의 분노를 표출하고 있는 화자가 ‘우리 마누라가 또 바람을 피고 있다’라고 말한다고 해서 ‘우리 마누라’의 사용이 부적절한 것으로 보이지 않는 것처럼, 위의 사례에서도 ‘우리 마누라’의 사용이 부적절하지 않아 보인다고 말할 유창한 한국어 화자들이 꽤 있을 것 같다. 물론 최성호 측에서는 이런 경우의 ‘우리 마누라’의 사용은 반어적인 것이라고 주장하겠지만 이 역시 문제가 있다. 어떤 문재인 지지자도 반어적으로 ‘문재양’을 사용하지

않는다. 하지만 이 반론에 따르면, ‘우리 마누라’는 반어적으로 사용될 수 있다. ‘문재앙’과 같은 비하어는 그 비하적인 의미에 해당하는 표현적 정보가 반어의 경우에도 약화되지 않으며, 문재인 지지자가 이 단어를 사용할 때는 언제나 스스로의 문재인에 대한 비하적 태도를 표현하게 된다. 그러나 ‘우리 마누라’가 반어적으로 쓰였다고 한다면, 여전히 화자는 자신의 마누라에 대한 어느 정도의 친밀감을 표현하게 되겠지만 이러한 친밀감은 반어적이지 않게 쓰이는 경우와는 달리 많이 약화된다. (예: 드디어 우리 마누라께서 한 달도 참지를 못하시고 다시 바람을 피셨구만.)

물론 이 경우에도 내집단 이론은 여전히 옹호될 수 있다는 반론이 가능할 것이다. 내집단의 결정은 친밀감의 정도차와 무관하게 성립한다고 주장할 수 있기 때문이다. 하지만 나는 윤재학의 내집단 개념이 정확하게 어떻게 정의되는지 이해하는 데 어려움을 겪었다. ‘우리 마누라’의 내집단은 무엇인가? 화자와 배우자? 화자와 배우자와 친정어머니? 아니면 화자와 배우자와 시아버지? (예: 시아버지를 극진히 모시던 자신의 배우자가 바람을 피고 있다는 사실을 알게 된 화자가 이 사정을 알고 있는 청자에게 이렇게 말한다. ‘우리 마누라가 바람을 피는 장면을 목격한 아버지께서 쓰러지셨다.’) 내집단을 단순히 화자와 그 배우자를 반드시 포함하는 불특정 집합 또는 맥락적으로 추가 원소가 결정되는 집합으로 보면 되지 않을까? (이 경우에 맥락을 입력으로 받아서 추가적 원소를 결정하는 함수가 일종의 특성으로 존재한다고 할 수도 있겠다.) 하지만 이 경우에도 위에서 내가 말한 적절성 기준 문제를 똑같이 제기할 수 있다. 앞서의 위장 결혼 사례를 다시 생각해보자. 왜 ‘우리 마누라’를 이 경우에 쓴다면 부적절한 사용이 되는가? 이 부적절성이 정확하게 내집단의 결정과 어떤 관련이 있는가? 법적으로 부부 관계인 배우자를 내집단 원소로 포함하는 ‘우리 마누라’ 사용은 왜 부적절한가? 남이나 마찬가지로 배우자를 내집단에 소속시키는 것은 부적절하다고 응수한다면, 이 대답은 친밀감 이론과 크게 다른 점이 없어 보인다.

‘문재앙’과 같은 비하어의 중요한 특성 중 하나는, 진리조건에 직접적으로 기여하지는 않지만 간접화법과 같은 문맥의 경우에도 여전히 부정적 화자의 태도와 연관되어 전달되는 의미 요소를 그대로 가진다는 것이다. 이러한 의미 요소는 맥락 독립적으로 항상 표현되며 규약적으로 결정되어 있다는 점에서 의미론적이다. 최성호는 비하어뿐 아니라 ‘우리 마누라’ 역시 정확하게 동일한 속성을 갖는다고 생각하며, 이것은 최성호의 전체 입장을 지탱하는 중요한 기초 역할을 한다. 하지만 내가 보기에 ‘문재앙’과는 달리 ‘우리 마누라’는 규약적으로 결정되어 있는 고정된 표현적 정보를 전달하는 것처럼 보이지 않는다. 오히려 사용 맥락에 따라서 특정한 규약에 구애됨 없이 화자의 의도 등에 따라서 다양한 방식으로 결정되는 화용

론적 함축만이 있는 것 같다. ‘우리 마누라’의 표현적 정보가 그라이스의 규약적 함축에 해당하지 않는다는 최성호의 논증이 성공적이더라도, 여전히 규약적 함축이 아닌 다른 종류의 화용론적 함축의 존재를 인정할 수 있다. ‘나의 마누라’와는 달리 ‘우리 마누라’에서의 ‘우리’의 사용은 규약적으로 화자의 의도와 관련된 추가적인 화용론적 함축의 존재를 표시하지만, 이것은 의무적인 것은 아니며 일종의 허용적 규칙에 해당할 수 있다. 그리고 각 발화 맥락에서 ‘우리 마누라’의 화용론적 함축은 다양한 복합적 요소에 의해서 결정될 뿐 하나의 단일한 규칙이 존재하지 않을 수도 있다.

The Identity of Proofs and the Criterion for Admissible Reductions

Seungrak Choi

Department of Philosophy, Korea University

Abstract

Dag Prawitz (1971) put forward the idea that an admissible reduction process does not affect the identity of proofs represented by derivations in natural deduction. The idea relies on his conjecture that two derivations represent the same proof if and only if they are equivalent in the sense that they are reflexive, transitive and symmetric closure of the immediate reducibility relation.

Schroeder-Heister and Tranchini (2017) accept Prawitz's conjecture and propose the triviality test as the criterion for admissible reductions. In the present paper, we will consider two main troubles of the triviality test. The first is the obscurity of a method of evaluating admissible reductions. The second is the circularity problem that the triviality test already assumes the set of admissible reduction procedures. For the solution of the problems, we will propose the spoiler test which immunizes the problems of the triviality test and has the role of the criterion for admissible reductions. At last, we shall cover the plausible problem of the spoiler test that can be caused by Crabbé's case.

Keywords: The identity of proofs, admissible reductions, Ekman's paradox, Church-Rosser's confluence property, Crabbé's case.

1. Introduction

Many mathematical theorems have more than one proof. For example, Euclid's theorem which asserts that there are infinitely many prime numbers has at least three versions, such as Euclid's, Euler's, and Erdős's proof. Similarly, there is more than one formalization of the same proof in proof theory. Proof-theorists have investigated which derivation is a *canonical representation* of the proofs represented in the system considered as the way that the numerals are canonical notations for natural numbers.

Dag Prawitz (1971, Sec. I.3.5.6) has proposed the idea that normal derivations serve the special role of canonical representations. Normal derivations can be obtained by certain reduction procedures. Any derivations which can be a normal derivation with the finite steps of reductions are called normalizable derivations. A normalizable derivation shares the special role of a canonical representation of its normal derivation. Prawitz explained in terms of the equivalence class that normal and normalizable derivations represent the same proofs. He considered the equivalence relation between derivations that is reflexive, transitive and symmetric closure of the immediate reducibility relation between derivations. Then, he conjectured that two derivations represent the same proof if and only if they are equivalent. He thought that the right-to-left direction is relatively unproblematic and so claimed that an admissible reduction does not affect the identity of a proof represented. Hence, an admissible reduction preserves the identity between proofs. In this perspective, the admissible reductions can serve as the criterion for the identity of proofs.

On the other hand, assuming that we intuitively recognize which formal derivations represent the same informal proof, Schroeder-Heister and Tranchini (2017) accept Prawitz's conjecture and suggest the criterion for admissible reduction procedures. The idea is to evaluate whether, after a new reduction is added in a set of reductions for an intended system, it can be shown that two derivations of the same conclusion previously belonging to two distinct equivalence classes still belong to the distinct equivalence classes. If it can, they say that the new reduction does not trivialize the identity of proofs; otherwise, it *trivializes the identity of proofs*. A newly added reduction procedure is *admissible* if it does not trivialize the identity of proofs; otherwise, it is inadmissible. We call their evaluation method for admissible reductions, the *triviality test*.

The present paper aims to point out the problems of the triviality test and to propose an improved approach. In Section 2, we shall introduce Ekman's paradox and the triviality test. We will see how the triviality test blocks Ekman's reduction. Section 3 put forward two main problems of the triviality test: the obscurity of a test method and the circularity problem. With the solutions to the problems, we will introduce the spoiler test as the refined criterion for admissible reduction procedures in Section 4. The spoiler test can restrict the application of the Ekman reduction process.

Schroeder-Heister and Tranchini (2018) have considered their triviality test has a defect that it cannot solve the problem of Tennant's proof-theoretic criterion for paradoxicality occurred by Crabbé's case. Since every reduction process of Crabbé's case passes the spoiler test, the spoiler test may be seen as having the same defect. Section 5 argues that not every problem with Tennant's criterion is raised by inadmissible reductions. When we distinguish between the problems of Tennant's criterion and of inadmissible reductions, the spoiler test need not solve every problem of Tennant's criterion but it only has the role of evaluating admissible reductions.

2. Ekman's paradox and the triviality test

Neil Tennant (1982) has proposed a proof-theoretic criterion for paradoxicality that a genuine paradox is a derivation of an unacceptable conclusion which employs a certain form of *id est* inferences and generates an infinite reduction sequence. Schroeder-Heister and Tranchini (2017) have suggested the triviality test to solve the problem of Tennant's criterion. They have considered Ekman's paradox taken from Jan Ekman (1998) to show that Tennant's criterion overgenerates in the sense that there exists a derivation which is intuitively non-paradoxical but satisfies the criterion. To solve the problem of overgeneration, Tennant (2016, 2017) has refined his criterion and suggested an additional condition that all elimination rules stated in generalized form. Unfortunately, Schroeder-Heister and Tranchini (2018) argue that Tennant's refined version of the criterion over-generates again by showing that there is a generalized version of Ekman's paradox satisfying the refined criterion.

Schroeder-Heister and Tranchini (2017) diagnose that Ekman's paradox applies too loose reduction procedure, called the *Ekman reduction*. Since they focus on the issue of which

reduction process is an admissible one, they suggest the triviality test as the criterion for the admissibility of reduction procedures in order to restrict inadmissible reductions.

In this section, we set aside the issue of Tennant’s criterion for paradoxicality and center on the triviality test. After we have preliminary notions and rules in Section 2.1, we shall see the triviality test and how it blocks the Ekman reduction process in Section 2.2.

2.1 Preliminaries: some terminologies and natural deduction rules

To begin with, we will introduce some terminologies and natural deduction rules. Our language has constants and quantifiers, \wedge , \rightarrow , \perp , and \neg for conjunction, implication, absurdity, and negation respectively. We use φ , Ψ , and σ for arbitrary formula. Let D be a derivation of a given natural deduction system, used in the same manner as ‘deduction’ in Prawitz (1965). Also, we use the following conventions: if a derivation D ends with a formula φ , we write as shown on the left below and φ is called an ‘end-formula.’ If it depends on a formula Ψ , we write as shown on the right below.

$$\begin{array}{ccc} & & \Psi \\ & & D \\ D & & D \\ \varphi & & \varphi. \end{array}$$

Then, we have rules for \wedge and \rightarrow in the natural deduction style proposed by Prawitz (1965).

$$\begin{array}{cccc} & & [\varphi]^1 & \\ \begin{array}{c} D_1 \quad D_2 \\ \frac{\varphi_1 \quad \varphi_2}{\varphi_1 \wedge \varphi_2} \wedge I \end{array} & \begin{array}{c} \frac{\varphi_1 \wedge \varphi_2}{\varphi_i} \wedge E_{(i=1,2)} \end{array} & \begin{array}{c} D_1 \\ \frac{\Psi}{\varphi \rightarrow \Psi} \rightarrow I_1 \end{array} & \begin{array}{c} D_2 \\ \frac{\varphi \rightarrow \Psi \quad \varphi}{\Psi} \rightarrow E \end{array} \end{array}$$

$\neg\varphi$ is defined by $\varphi \rightarrow \perp$. We call the formulas directly above the line in each rule, ‘premise,’ and the formula directly below the line, ‘conclusion.’ *Assumptions* which can be discharged are in the square brackets, e.g. $[\varphi]$. When the same formula φ is introduced as assumptions with different superindexes, they are different types and not discharged simultaneously by a single rule. For instance, the following derivations have the same formula φ as their assumptions, yet their assumptions are discharged at different places.

$$\frac{\frac{[\varphi]^1 \quad [\varphi]^1}{\varphi \wedge \varphi} \wedge I}{\varphi \rightarrow (\varphi \wedge \varphi)} \rightarrow I_1 \quad \frac{\frac{[\varphi]^1 \quad [\varphi]^1}{\varphi \wedge \varphi} \wedge I}{\varphi \rightarrow (\varphi \wedge \varphi)} \rightarrow I_2}{\varphi \rightarrow (\varphi \rightarrow (\varphi \wedge \varphi))} \rightarrow I_1$$

The *open* assumptions of a derivation are the assumptions on which the end-formula depends. A derivation is called *closed* if it contains no open assumptions, otherwise it is called *open*. A *major premise* of the elimination rule for a constant is the premise containing the constant in *the elimination rule and all other premises are minor premises*. The *maximum formula* is the conclusion of an introduction rule at the same time the major premise of an elimination rule.

We accept the standard reduction procedures of Prawitz (1965, pp. 36-38). Let us consider any two derivations D_1 and D_2 having the same end-formula. We say that a derivation D' is an *immediate subderivation* of D_1 if D' is an initial part of D_1 ending with a premise of the last inference step in D_1 . Let $D_1 \triangleright D_2$ mean that D_1 *reduces* to D_2 by applying a single reduction step to an immediate subderivation D' of D_1 . For our convenience's sake, we will introduce standard reduction procedures for \wedge and \rightarrow .

1. The standard reduction procedure for \wedge .

$$\frac{\frac{D_1 \quad D_2}{\varphi_1 \quad \varphi_2} \wedge I}{\varphi_i} \wedge E_{(i=1,2)} \quad \triangleright_{\wedge} \quad \frac{D_i}{\varphi_i}$$

2. The standard reduction procedure for \rightarrow

$$\frac{\frac{[\varphi]^1}{D_1} \quad \frac{\Psi}{\varphi \rightarrow \Psi} \rightarrow I_1 \quad \frac{D_2}{\varphi} \rightarrow E}{\Psi} \rightarrow E \quad \triangleright_{\rightarrow} \quad \frac{D_2}{\varphi} \rightarrow E \quad \frac{D_1}{\Psi}$$

The main role of these standard reduction procedures is to eliminate the maximum formula. When a derivation has no maximum formula, we say that it is in *normal form*. Let \mathbb{R} be a set of

reduction procedures. A reduction procedure $\frac{\varphi_1, \dots, \varphi_n}{D} \triangleright \frac{\varphi_1, \dots, \varphi_n}{D'}$ in \mathbb{R} is *closure under substitution* iff, for any derivation $\frac{D_1}{\varphi_1, \dots, \varphi_n} \triangleright \frac{D_n}{\varphi_1, \dots, \varphi_n}$, a reduction procedure $\frac{D_1}{D} \triangleright \frac{D_n}{D'}$ is

in \mathbb{R} as well. Every reduction procedure in \mathbb{R} is to be closed under substitution of derivations for open assumptions, and the notions of ‘normal’ and ‘normalizable’ are defined in the following ways¹⁾

Definition 1. A sequence $\langle D_1, \dots, D_i, D_{i+1}, \dots \rangle$ of derivations is a *reduction sequence* relative to \mathbb{R} iff $D_i \triangleright D_{i+1}$ relative to \mathbb{R} where $1 \leq i$ for any natural number i . A derivation D_1 is *reducible* to D_i ($D_1 \triangleright D_i$) relative to \mathbb{R} iff there is a sequence $\langle D_1, D_2, \dots, D_i \rangle$ relative to \mathbb{R} where for each $j < i$, $D_j \triangleright D_{j+1}$; D_1 is *irreducible* relative to \mathbb{R} iff there is no derivation D' to which $D_1 \triangleright D'$ relative to \mathbb{R} except D_1 itself.

Definition 2. A derivation D is *normal* (or in *normal form*) relative to \mathbb{R} iff D is irreducible relative to \mathbb{R} , i.e. D has no maximum formula. A reduction sequence *terminates* iff it has a finite number of derivations and its last derivation is in normal form. A derivation D is *normalizable* relative to \mathbb{R} iff there is a terminating reduction sequence relative to \mathbb{R} starting from D .

In the next section, we shall introduce Schroeder-Heister and Tranchini’s triviality test.

2.2. The triviality test and the Ekman reduction procedure

Schroeder-Heister and Tranchini (2017) have suggested the triviality test in order to restrict inadmissible reductions. Their target was the following form of reduction procedure

1) In Definition 1, for any term x and y , let $x \leq y$ mean that x is less than or equal to y . For Definition 2, for our convenience’s sake, we drop the ‘relative to \mathbb{R} ’ in the suggested notions if there is no misunderstanding.

introduced by Ekman (1998):

$$\frac{\frac{[\Psi \rightarrow \varphi]}{\varphi} \quad \frac{[\varphi \rightarrow \Psi] \quad \varphi}{\Psi} \rightarrow E}{\varphi} \rightarrow E \quad \begin{array}{c} D \\ \hline \end{array} \quad \begin{array}{c} D \\ \hline \end{array} \quad \begin{array}{c} \triangleright_E \\ \hline \end{array} \quad \varphi$$

We call \triangleright_E the *Ekman reduction* process. Ekman (1998) considered Ψ in the last $\rightarrow E$ -rule to be a similar kind of maximum formula. We call Ψ in the last $\rightarrow E$ -rule an *Ekman maximum formula*.

The triviality test is based on the thesis of Prawitz (1971, p. 257) that an admissible reduction should not affect the identity of proofs represented by derivations in the same equivalence class. He conjectures that two derivations represent the same proof if and only if they are equivalent. Prawitz's equivalence relation \sim is defined via the reducibility relation $>$. We borrow the notion of the equivalence relation between derivations from Prawitz (1971, p. 255). Let R be any set of reduction procedures.²⁾

Definition 3. A derivation D_1 is *equivalent* to D_i ($D_1 \sim D_i$) relative to Riff $D_1 > D_i$ or $D_i > D_1$ where $1 \leq i$ for any natural number i ; otherwise, they are not equivalent ($D_1 \not\sim D_i$). Let S be any natural deduction system. The equivalence class of D_1 under \sim in S , denoted by $\llbracket D_1 \rrbracket$, is defined as $\llbracket D_1 \rrbracket = \{ D_i \in S \mid D_1 \sim D_i \}$.

Then, the relation \sim is clearly reflexive, symmetric, and transitive. Let \hat{D} be an informal proof represented by a formal derivation D in natural deduction. Prawitz's thesis and conjecture are summarized as below.

Prawitz's thesis: An admissible reduction procedure does not affect the identity of proofs represented by derivations in the same equivalence class.

2) Although Prawitz (1971) considers a set of standard reduction procedures for rules of first-order intuitionistic and classical logic, since we focus on the test of an admissible reduction procedure in general, we consider arbitrary set of reduction procedures.

The Conjecture for the Identity of Proofs(CIP): for any derivation D_1 and D_2 , $\hat{D}_1 = \hat{D}_2$ if and only if $D_1 \sim D_2$.

While Schroeder-Heister and Tranchini (2017) take Prawitz's thesis and CIP, they suggest the triviality test for evaluating whether a newly added reduction makes derivations representing different proofs belong to the same equivalence class. Let \mathbb{R} be a set of reduction procedures. A set \mathbb{R}' is an *extension* of \mathbb{R} if \mathbb{R}' results from \mathbb{R} by adding reduction procedures which are closed under substitution in \mathbb{R} . Let $\underline{\triangleright}$ be a reduction procedure not in a set \mathbb{R} of reductions and \mathbb{R}' be an extension of \mathbb{R} by adding $\underline{\triangleright}$. We use an abbreviation ' $S(\mathbb{R})$ ' for a natural deduction system S relative to \mathbb{R} . Let us consider the case that, for any three derivations Π_1 , Π_2 , and Π_3 , Π_1 is reducible to Π_2 but not to Π_3 in $S(\mathbb{R})$ whereas Π_1 is reducible to both Π_2 and Π_3 in $S(\mathbb{R}')$. Suppose that $\hat{\Pi}_2 \neq \hat{\Pi}_3$. Then, $\Pi_2 \simeq \Pi_3$ by CIP. However, in $S(\mathbb{R}')$, $\Pi_1 \sim \Pi_2$ and $\Pi_1 \sim \Pi_3$, and so $\Pi_2 \sim \Pi_3$ in $S(\mathbb{R}')$. While $\hat{\Pi}_2 \neq \hat{\Pi}_3$, it should be possible to show that Π_2 is not equivalent to Π_3 . Since it is possible to show that $\Pi_2 \sim \Pi_3$ in $S(\mathbb{R})$ but not in $S(\mathbb{R}')$, we may say that an additional reduction $\underline{\triangleright}$ affects the identity of proofs. Following Prawitz's thesis, we may conclude that $\underline{\triangleright}$ is not an admissible reduction procedure. From this perspective, Schroeder-Heister and Tranchini (2017, p. 575) propose the triviality test for an admissible reduction process that a newly added reduction procedure should not trivialize the identity of proofs.

A natural requirement of the addition of a new reduction could be that of not trivializing identity of proof, in the sense that it should always be possible to exhibit two derivations of the same conclusion belonging to two distinct equivalence classes.

We say that, for any derivations D_1 and D_2 where $\hat{D}_1 \neq \hat{D}_2$ and so $D_1 \simeq D_2$ in $S(\mathbb{R})$, a reduction process $\underline{\triangleright}$ *trivializes the identity of proofs* in S iff it is not possible to show in $S(\mathbb{R}')$ that $D_1 \simeq D_2$; otherwise, $\underline{\triangleright}$ *does not trivialize the identity of proofs* in S .³⁾ Then, their

3) Another interpretation of the notion trivializing the identity of proofs may be that a newly added reduction trivializes the identity of proofs iff it is possible to show in $S(\mathbb{R}')$ that every provable derivation of the same conclusion is equivalent. This interpretation is stronger than the one we have proposed, making the triviality test more difficult to apply. Since Schroeder-Heister and Tranchini (2017, p. 575) explicitly state the meaning of 'not

triviality test can be summarized as below:

Triviality Test: Let S be any natural deduction system. Let \triangleright be a new reduction procedure not in \mathbb{R} and \mathbb{R}' be an extension of \mathbb{R} by adding \triangleright . \triangleright is an *admissible reduction* procedure for S iff \triangleright does not trivialize the identity of proofs in S .

Schroeder-Heister and Tranchini (2017, pp. 575-577) evaluate whether the Ekman reduction \triangleright_E is admissible in accordance with the triviality test. Let S_T be a natural deduction system containing \wedge - and \rightarrow -rules. S_T has a set \mathbb{R}_T of standard reductions for \wedge and \rightarrow . \mathbb{R}'_T is an extension of \mathbb{R}_T by adding the Ekman reduction \triangleright_E . If the Ekman reduction \triangleright_E does not trivialize the identity of proofs, \triangleright_E is an admissible reduction. They provide an example showing that it is not. First, they consider two derivations D_1 and D_2 of the same formula which represent different proofs, i.e. $\hat{D}_1 \neq \hat{D}_2$. $D_1 \simeq D_2$, and so are not equivalent, i.e. Then, they argue that whence \triangleright_E is added in \mathbb{R}'_T , it is not possible to show that $D_1 \simeq D_2$ in $S_T(\mathbb{R}'_T)$.

For any two derivations D_1 and D_2 of φ in S_T , Schroeder-Heister and Tranchini give a derivation Π_1 of $\varphi \wedge \varphi$ as below:

$$\frac{\frac{\frac{[\varphi]^1 \quad [\varphi]^1}{\varphi \wedge \varphi} \wedge I}{\varphi \rightarrow (\varphi \wedge \varphi)} \rightarrow I_1}{\frac{\varphi \wedge \varphi}{\varphi} \wedge E_2} \rightarrow I_1 \quad \frac{\frac{\frac{[\varphi \wedge \varphi]^2}{\varphi} \wedge E_1}{(\varphi \wedge \varphi) \rightarrow \varphi} \rightarrow I_2}{\varphi} \rightarrow E \quad \frac{D_1 D_2}{\varphi \wedge \varphi} \wedge I \rightarrow E$$

$S_T(\mathbb{R}'_T)$ has a reduced derivation Π_2 from Π_1 by applying $\triangleright_{\rightarrow}$ and \triangleright_{\wedge} with regard to $\wedge E_1$ -rule.

trivializing identity of proofs,' we will use the proposed stipulation in this paper.

$$\begin{array}{c}
D_1 D_1 \\
\frac{\varphi \quad \varphi}{\varphi \wedge \varphi} \wedge I \\
\frac{\varphi \wedge \varphi}{\varphi} \wedge E_2
\end{array}$$

On the other hand, $S_T(\mathbb{R}'_T)$ has not only Π_2 but also the following Π_3 through the application of $\underline{\triangleright}_E$.

$$\begin{array}{c}
D_1 D_2 \\
\frac{\varphi \quad \varphi}{\varphi \wedge \varphi} \wedge I \\
\frac{\varphi \wedge \varphi}{\varphi} \wedge E_2
\end{array}$$

A proof represented by derivations in $[[\Pi_1]]$ should represent the same proof. Then, by the application of \triangleright_{\wedge} with regard to $\wedge E_2$ -rule, we have Π'_2 from Π_2 on the left below and Π'_3 from Π_3 on the right below in $S_T(\mathbb{R}'_T)$ as follows:

$$\begin{array}{cc}
D_1 & D_2 \\
\varphi & \varphi
\end{array}$$

By Definition 3, $\Pi_1 \sim \Pi'_2$ and $\Pi_1 \sim \Pi'_3$. Hence, $\Pi'_2 \sim \Pi'_3$. This means that, for any derivations D_1 and D_2 of φ , $D_1 \sim D_2$, and so $\hat{D}_1 = \hat{D}_2$. At this point, Schroeder-Heister and Tranchini's argument depends on an assumption that not every proof of φ is the same. That is, it is not the case that, for any derivation D_1 and D_2 of φ in S_T , $D_1 \sim D_2$. If the assumption is true, there are derivations of φ in S_T that are not equivalent. However, since $D_1 \sim D_2$ in $S_T(\mathbb{R}'_T)$, it is not possible to show in $S_T(\mathbb{R}'_T)$ that $D_1 \simeq D_2$. Therefore, the Ekman reduction $\underline{\triangleright}_E$ trivializes the identity of proofs. The triviality test says that E is not admissible.

As we have noted, the triviality test presumes Prawitz's thesis and the conjecture for the identity of proofs. If these are required for a correct natural deduction system, the triviality test would be an admissible reduction checker. However, there are still obstacles for the

triviality test to be a promising criterion for the admissibility of reduction procedures.

3. Problems of the triviality test

There are three problems of the triviality test to be solved. A minor problem is that the triviality test is based on *CIP*. Prawitz (1971, p. 257) and Kreisel (1971, pp. 114-116) noticed that the right-to-left direction of *CIP* that, for any derivations D_1 and D_2 , $D_1 \sim D_2$ implies $\hat{D}_1 = \hat{D}_2$, is convincing, but the opposite direction that $\hat{D}_1 = \hat{D}_2$ implies $D_1 \sim D_2$ is not. Provided that we have intuitively admissible reduction procedures, Prawitz's thesis says that an admissible reduction does not affect the identity of proofs represented by derivations in the same equivalence class. Equivalent derivations represent the same proof. That is, $D_1 \sim D_2$ implies $\hat{D}_1 = \hat{D}_2$. On the other hand, the same proof can be represented by different derivations. It is not obvious that inequivalent derivations always represent different proofs. If the left-to-right direction of *CIP* is necessary for the triviality test, it should be explicated why $\hat{D}_1 = \hat{D}_2$ implies $D_1 \sim D_2$ to have the triviality test as a promising admissible reduction checker.

It is unclear whether the triviality test must presume both directions of *CIP*. Schroeder-Heister and Tranchini seem to accept both directions. First, when Schroeder-Heister and Tranchini (2017, p. 574) introduce the perspective of the identity of proofs, they say, '[f]or simplicity, proofs might ... be considered to be equivalence classes with respect to derivations.' Although they intend to simplify the discussion for the triviality test, it can be interpreted that they admit both directions of *CIP*.

Moreover, when they introduce a *typical* example of two derivations which represent two distinct proofs, they use two derivations of the same conclusion which their assumptions are discharged at different places.

$$\frac{\frac{\frac{[\varphi]^1}{\varphi \rightarrow \varphi} \rightarrow I,1}{\varphi \rightarrow (\varphi \rightarrow \varphi)} \rightarrow I,\phi \quad D}{\varphi \rightarrow \varphi} \quad E \qquad \frac{\frac{\frac{[\varphi]^1}{\varphi \rightarrow \varphi} \rightarrow I,\phi}{\varphi \rightarrow (\varphi \rightarrow \varphi)} \rightarrow I,1 \quad D}{\varphi \rightarrow \varphi} \quad E$$

Two derivations above belong to two different equivalent classes in the sense of Prawitz's equivalence relation. In the case of the empty discharge, the reduction process for \rightarrow has the following form.

$$\frac{\frac{\frac{D_1}{\Psi} \rightarrow I, \phi \quad D_2}{\varphi \rightarrow \Psi} \rightarrow E}{\Psi} \quad \triangleright_{\rightarrow(\phi)} \quad \frac{D_1}{\Psi}$$

$\triangleright_{\rightarrow(\phi)}$ -reduction is an instance of $\triangleright_{\rightarrow}$. Then, the reduced derivations below are obtained by $\triangleright_{\rightarrow}$ respectively.

$$\frac{[\varphi]^1}{\varphi \rightarrow \varphi} \rightarrow I, 1 \quad \frac{D}{\varphi \rightarrow \varphi} \rightarrow I, \phi$$

The reduced derivations are different and so the original derivations are not equivalent.

If Schroeder-Heister and Tranchini interpret that the typical example shows that inequivalent derivations represent different proofs, their interpretation is based on the one half of Prawitz's *CIP* that $\hat{D}_1 = \hat{D}_2$ implies $D_1 \sim D_2$, i.e. $D_1 \simeq D_2$ implies $\hat{D}_1 \neq \hat{D}_2$. Therefore, it is better to think that they accept the whole conjecture.

Prawitz (1971, pp. 256-257) introduced *CIP* as a philosophical consequence of the strong normalization theorem that every derivation is reducible to a unique normal derivation regardless of the order in which reductions are applied. If the strong normalization theorem holds in an intended system, every equivalent derivation has the same normal derivation. Inequivalent normal derivation cannot represent the same proof. Therefore, when the strong normalization theorem holds, inequivalent derivations represent distinct proofs, i.e. $D_1 \simeq D_2$ implies $\hat{D}_1 \neq \hat{D}_2$.

Unfortunately, there is no guarantee that the strong normalization theorem will be established if new rules are added in S or new reduction procedures are introduced. Also, we are in the context of finding the general criterion of an admissible reduction concerned with not only

normalizable derivations but also paradoxical non-normalizable derivations. It is better not to accept the left-to-right direction of *CIP* that $\hat{D}_1 = \hat{D}_2$ implies $D_1 \sim D_2$ unless there is a clear verification of the left-to-right direction.

We may partly accept *CIP*, i.e. $D_1 \sim D_2$ implies $\hat{D}_1 = \hat{D}_2$, and use the triviality test. However, the triviality test has two more problems that have to be solved. The first problem is the obscurity of a test method. The triviality test uses modal notions such as ‘possibility.’ It is hard to say that the triviality test has a clear characterization of its application. The triviality test only states that when a new reduction is added, it *should be possible* to show that two derivations originally inequivalent still belong to different equivalent classes. Schroeder-Heister and Tranchini’s argument finds a case that every derivation of the same formula becomes equivalent, but it does not the only way to establish the impossibility of showing that derivations representing different proofs belong to distinct equivalence classes. We may merely show that inequivalent derivations become equivalent after a new reduction process is added.

Moreover, it is not desirable to use modal notions as well as an informal notion, such as ‘proof,’ in order to give a clear characterization of the criterion for admissible reductions. The obscurity of the test method leads to an undesirable consequence that some standard reductions for primary constants are inadmissible because the triviality test does not explicate which derivations are fundamentally inequivalent.

Let S_t be a natural deduction system having only \wedge -rules. S_t has a set \mathbb{R}_φ of a reduction procedure which only has the following reduction process.

$$\frac{\frac{D}{\varphi \wedge \varphi} \wedge E_1}{\varphi} \quad \frac{\frac{D}{\varphi \wedge \varphi} \wedge E_1}{\varphi} \wedge I}{\varphi \wedge \varphi} \quad \frac{D}{\varphi \wedge \varphi} \triangleright_e$$

Suppose that, for any two inequivalent derivations D_1 and D_2 ending with φ in $S_t(\mathbb{R}_\varphi)$, there is a derivation Σ_1 as follows:

$$\begin{array}{c}
\begin{array}{c} D_1 \quad D_2 \\ \frac{\varphi \quad \varphi}{\varphi \wedge \varphi} \wedge I \\ \hline \varphi \end{array} \wedge E_1 \quad \begin{array}{c} D_1 \quad D_2 \\ \frac{\varphi \quad \varphi}{\varphi \wedge \varphi} \wedge I \\ \hline \varphi \end{array} \wedge E_1 \\
\hline
\frac{\varphi \wedge \varphi}{\varphi} \wedge E_2
\end{array}$$

In $S_t(\mathbb{R}_\forall)$, Σ_1 reduces to the following derivation Σ_2 by \triangleright_e .

$$\begin{array}{c}
D_1 \quad D_2 \\
\frac{\varphi \quad \varphi}{\varphi \wedge \varphi} \wedge I \\
\hline
\varphi \wedge E_2
\end{array}$$

Let us consider the below derivation Σ_3 .

$$\begin{array}{c}
D_1 \quad D_1 \\
\frac{\varphi \quad \varphi}{\varphi \wedge \varphi} \wedge I \\
\hline
\varphi \wedge E_2
\end{array}$$

$\Sigma_1 \sim \Sigma_2$ but both Σ_1 and Σ_2 are not equivalent to Σ_3 in $S_t(\mathbb{R}_\forall)$ because Σ_1 and Σ_2 are not reducible to Σ_3 in $S_t(\mathbb{R}_\forall)$. Now, we add the standard reduction for \wedge , \triangleright_\wedge , to \mathbb{R}_\forall and have \mathbb{R}'_\forall . Then, Σ_1 reduces to Σ_3 and Σ_2 reduces to the left derivation Σ_4 below. Also, Σ_2 reduces to the right derivation Σ_5 below.

$$\begin{array}{c}
D_1 \quad D_2 \\
\varphi \quad \varphi
\end{array}$$

Since $\Sigma_1 \sim \Sigma_4$ and $\Sigma_1 \sim \Sigma_5$ in $S_t(\mathbb{R}'_\forall)$, $\Sigma_4 \sim \Sigma_5$. However, $D_1 \not\equiv D_2$ in $S_t(\mathbb{R}_\forall)$ and so $\Sigma_4 \not\equiv \Sigma_5$ in $S_t(\mathbb{R}_\forall)$. It is not possible to show in $S_t(\mathbb{R}'_\forall)$ that $\Sigma_4 \equiv \Sigma_5$, \triangleright_\wedge trivializes the identity of proofs in S_t and so is not an admissible reduction for St . Hence, the triviality test

makes the standard reduction process \triangleright_{\wedge} inadmissible. Should we say that \triangleright_{\wedge} is an inadmissible reduction process?

The above result shows that the notion of an inequivalent derivation is relative to our choice of the set of reductions. If we use the triviality test with the set of reductions which already have inadmissible reductions, it can evaluate that even standard reduction procedures commonly considered admissible are inadmissible. In this regard, the base set of admissible reductions should be properly chosen. Unfortunately, at this point, the triviality test faces the second trouble.

The second problem is the circularity problem. The triviality test is (partially) based on *CIP*, especially that $D_1 \sim D_2$ implies $\hat{D}_1 = \hat{D}_2$. The notion of an inequivalence relation between two derivations has the significant role of the triviality test. The matter is that the equivalence relation is dependent on which set of reductions we choose. The evaluation of whether a newly added reduction process trivializes the identity of proofs or not is relative to our choice of the base set of admissible reductions. By the way, since the triviality test should be an admissible reduction checker, the admissible reductions in the base set should also be evaluated by it. The circularity problem arises. If admissible reductions are dependent on our choice of reductions, we can simply choose without worrying about which reduction process is admissible. In the next section, we shall attempt to solve these problems and revise the triviality test.

4. The spoiler test and solutions

In order to solve the circularity problem of the triviality test, we need to fix the base set of admissible reduction procedures. Although the Ekman reduction process gets rid of unnecessary steps, its role is different from the role of reductions for \wedge and \rightarrow . Schroeder-Heister and Tranchini did not distinguish such roles. Unlike the Ekman reduction, the reduction procedures for \wedge and \rightarrow has the main role to eliminate a maximum formula in accordance with the inversion principle which is the core principle of Gentzen-Prawitz's natural deduction. In Section 4.1, we will distinguish between the standard and auxiliary reductions and define a *base set* of reductions. Then, the circularity problem will be solved. For the problem of the

obscurity of a test method, we revise the triviality test and propose the spoiler test in Section 4.2.

4.1 The solution to the circularity problem: the standard and auxiliary reduction procedures.

When Gentzen (1935) introduces a natural deduction system, he explains the roles of introduction and elimination rules as below:

The introductions represent, as it were, the ‘definitions’ of the symbols concerned, and the eliminations are no more, in the final analysis, than the consequences of these definitions. This fact may be expressed as follows: In eliminating a symbol, we may use the formula with whose terminal symbol we are dealing only ‘in the sense afforded it by the introduction of that symbol.’ (Gentzen, 1935, p. 80)

Gentzen’s idea is often interpreted as the meaning of a logical constant is exhaustively determined by its introduction rule and determines its elimination rule.

Prawitz (1965) borrowed the idea and has developed it. To realize Gentzen’s idea that an elimination rule is determined by the meaning of the conclusion of an introduction rule, there must be a certain requirement that fixes the elimination rule as the inverse of the corresponding introduction rule. For such requirement, Prawitz (1965, p. 33) suggests the inversion principle which states that whatever follows from a formula must follow from the direct ground for deriving that formula. We borrow his principle in the following way.

The Inversion Principle: Let D_i be any immediate subderivation of an introduction rule for deriving the major premise of an elimination rule, D_j be any derivation of minor premises of the elimination rule, and φ be any conclusion of the elimination rule. D_i together with D_j already derives φ without the application of the elimination rule. (i.e. any consequences of the major premise is derivable by D_i together with D_j .)

The inversion principle reflects Gentzen’s idea and says that nothing is gained by the application of an elimination rule when its major premise has been derived by means of an

introduction rule. In order to show that a pair of introduction and elimination rules of each constant (or operator) satisfy the inversion principle, Prawitz (1965, pp. 35–38) proposes reduction procedures for logical constants. When a reduction procedure satisfies the inversion principle, it provides a method to eliminate a maximum formula. Thus, any reduction procedure eliminating a maximum formula in accordance with the inversion principle will be called a *standard reduction* procedure. The examples of standard reduction procedures are the reduction procedures for \wedge and \rightarrow introduced in Section 2.1.

The Ekman reduction is not a reduction which fits the inversion principle. Ekman maximum formula is not a maximum formula in the standard sense because it is neither a conclusion of an introduction rule nor a major premise of an elimination rule. Therefore, the Ekman reduction is not a standard reduction. Rather, it is a process to reduce the length of a derivation. We will call it an auxiliary reduction process.⁴⁾

Since the inversion principle reflects Gentzen’s idea, it explains not only the primary roles of introduction and elimination rules but also those of the standard reductions in natural deduction. With respect to the primary roles, Dummett (1991, p. 250) treats ‘the eliminability of [maximum formulas] as a criterion for intrinsic harmony.’ The intrinsic harmony is often regarded as a minimum requirement for acceptable rules. It can be defined via the inversion principle.

Definition 4. (Intrinsic Harmony) Let \circ be a constant. Introduction and elimination rules for \circ are *intrinsically harmonious* iff every pair consisting of an $\circ I$ - and $\circ E$ -rules satisfies the inversion principle.

From the history of the development of natural deduction, it is convincing that we have minimal criteria of acceptable rules and of admissible reduction procedures through the inversion principle. Therefore, we say that, for any system S , the *base set* \mathbb{R} of admissible

4) The degree $d(\varphi)$ of a formula φ is defined by $d(\perp)=0$, $d(\alpha)=0$ for an atomic formula α , $d(\varphi \circ \Psi)=d(\varphi)+d(\Psi)+1$ for binary operators \circ , $d(\circ \varphi)=d(\varphi)+1$ for unary operators \circ . The length of a derivation D is the number of formula occurrences. There are at least three sorts of auxiliary reductions which (i) lessens the degree of a major premise or (ii) lessens the length of a derivation, or (iii) changes the order of subderivations of a derivation. The example of (i) is a reduction process for the rule of classical reductio suggested by Prawitz (1965, p. 40). The Ekman reduction is the example of (ii). The example of (iii) is a permutation conversion introduced in Prawitz (1965, p. 51; 1971, p. 253). The target of these reductions (i), (ii), and (iii) are not to eliminate a maximum formula. Therefore, the standard reduction and auxiliary reductions are to be distinguished.

reduction procedures of S consists of all standard reduction procedures for constants of intrinsically harmonious rules in S . Then, the circularity problem suggested in Section 3 is solved. When S_t has intrinsically harmonious rules for \wedge , it has the base set \mathbb{R}_b consisting of the standard reduction for \wedge . Since $\Sigma_1 \sim \Sigma_3$ in $S_t(\mathbb{R}_b)$, we do not arrive at the conclusion that \triangleright_{\wedge} is inadmissible. Moreover, while an intended system has intrinsically harmonious rules, its base set of admissible reductions exists and the circularity problem does not occur.

4.2 The spoiler test

The circularity problem can be solved, but the triviality test still has an unclear characterization of its test method. For instance, as we have noted in footnote 3 of Section 2.2 and Section 3, it remains obscure whether we should show that every provable derivation of the same conclusion is equivalent in order to evaluate the triviality of a given reduction. Therefore, we propose the spoiler test as an improvement of the triviality test on Prawitz's perspective of the identity of proofs.

Although Prawitz (1971, pp. 256-257) suggests his thesis and *CIP* as the philosophical consequence of the strong normalization theorem, our criterion for admissible reductions should be applicable to a system that the strong normalization theorem does not hold. For instance, the derivation of a set-theoretic paradox introduced in Prawitz (1965, pp. 94-95) yields a non-terminating reduction sequence and so is not normalizable. The strong normalization theorem cannot be established in the system containing Prawitz's derivation of a set-theoretic paradox. However, it is hard to say that the derivation uses an inadmissible reduction process. The strong normalization theorem does not need to be requested for finding the criterion for admissible reduction procedures.

For persuing the admissible reduction checker based on the identity of proofs, especially the one half of *CIP*, i.e. $D_1 \sim D_2$ implies $\hat{D}_1 = \hat{D}_2$, the Church-Rosser's confluence property (*CRCP*) is more suitable than the strong normalization.

The Church-Rosser's Confluence Property(*CRCP*): Let D_1, D_2, D_3, D_4 be any derivations in an intended natural deduction system. If $D_1 > D_2$ and $D_1 > D_3$ where $D_2 \simeq D_3$, then there is D_4 such that $D_2 > D_4$ and $D_3 > D_4$.

Church-Rosser theorem states something concerning reduction sequences in the lambda

calculus, however the property was proved in an intuitionistic natural deduction system by Jean-Yves Girard (1987, pp. 135-150). Every normalizable derivation satisfying *CRCP* has at most one normal derivation. If two distinct derivations reduce to different normal derivations, they are not reduced from the same derivation. Moreover, a system having a paradoxical derivation which generates a non-terminating reduction sequence and so is not normalizable can satisfy *CRCP*. For instance, the derivation of a set-theoretic paradox introduced in Prawitz (1965, pp. 94-95) is not normalizable but it satisfies *CRCP* because *CRCP* does not require that every derivation should be normalizable.

While we only consider derivations satisfying *CRCP*, it is guaranteed that there are no distinct normal derivations reduced from the same derivation. Using *CRCP*, we introduce the spoiler test.

The Spoiler Test: Let S be any natural deduction system satisfying *CRCP*, and \mathbb{R} be a base set of standard reductions for S . Let $\underline{\triangleright}$ be a reduction procedure not in \mathbb{R} and \mathbb{R}' be an extension of \mathbb{R} by adding $\underline{\triangleright}$. $\underline{\triangleright}$ *spoils the identity of proofs in $S(\mathbb{R})$* if there is a derivation in $S(\mathbb{R}')$ which does not satisfy *CRCP*; otherwise, it does not. $\underline{\triangleright}$ is an *admissible reduction* procedure in $S(\mathbb{R})$ if it does not spoiler the identity of proofs in $S(\mathbb{R})$; otherwise, it is inadmissible.

Any reduction procedures which spoil the identity of proofs are called a ‘spoiler.’ Let us consider the same system $S_T(\mathbb{R}_T)$ and $S_T(\mathbb{R}'_T)$ introduced in Section 2.2. \mathbb{R}_T is the base system for S_T . \mathbb{R}'_T is an extension of \mathbb{R}_T by adding the Ekman reduction $\underline{\triangleright}_E$ to \mathbb{R}_T . We only need the right-to-left direction of *CIP* that $D_1 \sim D_2$ implies $\hat{D}_1 = \hat{D}_2$. Now, we have an argument that the Ekman reduction $\underline{\triangleright}_E$ is a spoiler and is inadmissible.

Let D_1 be any closed normal derivation of Ψ . For any distinct closed normal derivations D_2 and D_3 of φ , we have the following derivation Δ_1 of φ .

$$\frac{\frac{\frac{\varphi}{\Psi \rightarrow \varphi} \rightarrow I, \phi \quad \frac{\frac{D_3}{\varphi \rightarrow \Psi} \rightarrow I, \phi \quad \frac{D_1}{\Psi} \rightarrow E}{\varphi \rightarrow \Psi} \rightarrow E \quad \frac{D_2}{\varphi} \rightarrow E}{\varphi} \rightarrow E}{\varphi} \rightarrow E$$

The application of the Ekman reduction $\underline{\triangleright}_E$ to \mathcal{A}_1 yields the left derivation \mathcal{A}_2 below and the right derivation \mathcal{A}_3 below is given by the application of $\triangleright_{\rightarrow(\phi)}$ to \mathcal{A}_1 .

$$\begin{array}{cc} D_2 & D_3 \\ \varphi & \varphi \end{array}$$

Since D_2 and D_3 are distinct normal derivations, there is no single derivation reducible from both D_2 and D_3 . Therefore, there is a derivation in $S(\mathbb{R}^n)$ which does not satisfy *CRCP*. Therefore, the Ekman reduction process $\underline{\triangleright}_E$ is inadmissible because it is a spoiler.

5. An objection: Crabbé's case

Schroeder-Heister and Tranchini (2018) have considered an example observed by Marcel Crabbé which arises the problem of overgeneration of the proof-theoretic criterion for paradoxicality. So to speak, the case of Crabbé shows that the criterion for paradoxicality makes a non-paradoxical derivation paradoxical. They said that the triviality test fails to block the reduction procedures used in Crabbé's case and so it is a defect that the test cannot solve the overgeneration problem. If it is a defect of the triviality test, the spoiler test may have the same defect.

After we will introduce a brief summary of the development of the proof-theoretic criterion for paradoxicality in Section 5.1, we will see in Section 5.2 that every reduction used in Crabbé's case passes the spoiler test. Yet, we shall argue that not every problem of overgeneration is occurred by inadmissible reduction. When we focus on the method of evaluating admissible reduction, the spoiler test can be the criterion for admissible reduction procedures.

5.1 A brief summary of the development of the proof-theoretic criterion for paradoxicality

Prawitz (1965, Appendix B) first investigated that a derivation of \perp from the set-theoretic paradox falls into a non-terminating reduction sequence, i.e. such derivation is not normalizable.

Tennant (1982, p. 283) has claimed that the non-terminating reduction sequence is the distinguishing feature of the genuine paradoxes and proposed a proof-theoretic criterion for paradoxicality. Let M be any model and $\theta(M)$ be a set of sentences relative to M . He said,

A set of sentences is paradoxical relative to M iff there is some proof of $[\perp]$ from $\theta(M)$, involving those sentences in *id est* inferences, that has a looping reduction sequence.

His *id est* inferences may be any inferences having a formula interdeducible with its own negation (or its predication), such as inferences from φ to $\neg\varphi$ and $\neg\varphi$ to φ . Although his first stipulation of the criterion is not purely described in proof-theoretic fashion because of using the notion of a ‘model,’ Tennant’s early version of the proof-theoretic criterion for paradoxicality, *PCP*, can be summarized as below.

Tennant’s Proof-Theoretic Criterion for Paradoxicality(*PCP*_T): Let S be a natural deduction system relative to a set \mathbb{R} of reduction procedures. Let D be any derivation in S . D is a *P-paradox* if and only if

- (i) D is a (closed or open) derivation of \perp ,⁵⁾
- (ii) *id est* inferences (or rules) are used in D ,
- (iii) applications of the reduction procedures in \mathbb{R} to D yield a non-terminating reduction sequence.⁶⁾

*PCP*_T has a counterexample which raises the problem of overgeneration. When *PCP*_T overgenerates, it includes a non-paradoxical derivation into the realm of P-paradoxes. Ekman’s paradox suggested by Schroeder-Heister and Tranchini (2017) causes the overgeneration problem. Tennant (2017) has attempted to solve the problem and proposed an additional

5) \perp is not the only unacceptable conclusion. We can use a propositional variable p as an unacceptable conclusion while formulating Curry’s paradox. For the examination of other cases, the reader can consult Tennant (1982)

6) Tennant (1982, 1995) has introduced two types of a non-terminating reduction sequence; a looping reduction sequence and a spiraling reduction sequence. Tennant (1995, p. 207) said that a looping reduction sequence is the main feature of the self-referential paradoxes whereas a spiraling reduction is generated by non-self-referential paradoxes, such as Yablo’s paradox.

condition that every elimination rule is to be stated in generalized form. However, as Schroeder-Heister and Tranchini (2018) note, there is an Ekman-type paradox using generalized elimination rules which shows that Tennant’s revised criterion overgenerates. We follow Schroeder-Heister and Tranchini’s diagnosis that Ekman’s paradox uses too loose reduction procedure. So instead of considering Tennant’s additional criterion, we have the revised version of the proof-theoretic criterion for paradoxicality, PCP_R , by adding the following condition to PCP_T .

(iv) every reduction procedure in \mathbb{R} is admissible.

Then, our discussions of the criterion for admissible reductions in the present paper can be melted into the discussions of PCP_R .

5.2 Should the spoiler test solve the overgeneration problem occurred by Crabbé’s case?

The criterion for paradoxicality is one thing and the criterion for admissible reductions is another. Schroeder-Heister and Tranchini (2018) have regarded the triviality test as the solution for every overgeneration problem of PCP_R . The failure of solving the problem of overgeneration occurred by Crabbé’s case becomes a defect of the triviality test.

On the other hand, we claim that the spoiler test need not be a general solution to all overgeneration problems. After we shall see that every reduction in Crabbé’s case passes the spoiler test, we will argue that not every overgeneration problem is raised by inadmissible reductions.

To begin with, Crabbé’s case may raise the overgeneration problem of PCP_R . It has rules for Zermelo’s separation axiom. We propose the rules for Zermelo’s separation axiom as the rules for \in in the following way.

$$\begin{array}{ccc}
 D_1 & & \\
 \frac{t \in s \quad \varphi[t/x]}{t \in x \in s \mid \varphi(x)} \in I & \quad \frac{t \in x \in x \mid \varphi(x)}{t \in s} \in E_1 & \quad \frac{t \in x \in s \mid \varphi(x)}{\varphi[t/x]} \in E_2
 \end{array}$$

The standard reduction procedure \triangleright_{\in} for $\in I$ - and $\in E$ -rules are as below:

$$\begin{array}{ccc}
\frac{D_1 \quad D_2}{\frac{t \in s \quad \varphi[t/x]}{t \in x \in s} \mid \varphi(x)} \in I & & \\
\frac{\quad}{t \in s} \in E_1 & \triangleright_{\in 1} & \frac{D_1}{t \in s} \\
\\
\frac{D_1 \quad D_2}{\frac{t \in s \quad \varphi[t/x]}{t \in x \in s} \mid \varphi(x)} \in I & & \\
\frac{\quad}{\varphi[t/x]} \in E_2 & \triangleright_{\in 2} & \frac{D_2}{\varphi[t/x]}
\end{array}$$

We call both $\triangleright_{\in 1}$ and $\triangleright_{\in 2}$, ‘Crabbé reduction.’ For any set b , we define Z_b as a set $x \in b \mid \neg x \in x$. We take $\neg x \in x$ for φ in $\in I$ - and $\in E$ -rules and for terms t and s we take Z_b and b respectively. Then, the following rules are the instances of $\in I$ - and $\in E$ -rules.

$$\frac{D_1 \quad D_2}{\frac{Z_b \in b \quad \neg Z_b \in Z_b}{Z_b \in x \in b} \mid \neg x \in x} \in I \quad \frac{Z_b \in x \in b \mid \neg x \in x}{Z_b \in b} \in E_1 \quad \frac{Z_b \in x \in b \mid \neg x \in x}{\neg Z_b \in Z_b} \in E_2$$

Moreover, the instances of the reduction procedures \triangleright_{\in} can be suggested in a similar way. Let S_{\in} be a system having \in - and \rightarrow -rules. Since $\triangleright_{\rightarrow}$ and \triangleright_{\in} are standard reductions, S_{\in} has a base set of reductions including $\triangleright_{\rightarrow}$ and \triangleright_{\in} . Then, the following result is called, ‘Crabbé’s case.’

Lemma 1. *There is an open derivation Θ of \perp from the assumption $[Z_b \in b]$ in S_{\in} relative to \mathbb{R}_{\in} which generates a non-terminating reduction sequence, and so is not normalizable.*

Proof. See Schroeder-Heister and Tranchini (2018).

Proposition 2. *There is a closed derivation of $\neg \exists y (Z_y \in y)$ in S_{\in} . Proof.* By using Lemma 1, it is easily proved.

\mathbb{R}_{\in} does not have auxiliary reductions but have standard reductions for $\triangleright_{\rightarrow}$ and \triangleright_{\in} .

According to the spoiler test, all reductions in \mathbb{R}_{\in} are admissible. Also, it is readily seen that every derivation in S_{\in} satisfies *CRCP*.⁷⁾ Proposition 2 states that no set contains its own subset and is a commonsense result in a consistent Zermelo's paradox, but the derivation Θ of Lemma 1 satisfies *PCP_R*. As Crabbé's case does not formulate a paradox, *PCP_R* makes it a P-paradox. Hence, Crabbé's case shows that *PCP_R* over generates and the spoiler test can not solve the over generation problem.

Since Schroeder-Heister and Tranchini (2018) think that the triviality test has a defect that it is unable to solve the overgeneration problem by Crabbé's case, they may say that the spoiler test has the same defect. However, the criterion for admissible reductions does not need to solve every problem of overgeneration by ruling out inadmissible reduction procedures. Not every overgeneration problem is caused by inadmissible reductions. For arbitrary φ , we consider the following odd rules for φ and their standard reduction \triangleright_{φ} .

$$\frac{[\varphi]^1}{D_1} \quad \frac{D_1}{\frac{\varphi}{\perp} \varphi E}$$

$$\frac{\perp}{\varphi} \varphi I_1$$

The standard reduction for φ is stated as below.

$$\frac{[\varphi]^1}{D_1} \quad \frac{D_2}{\varphi} \quad \frac{\perp}{\varphi} \varphi I_1 \quad \frac{D_2}{\varphi} \quad \frac{\perp}{\varphi} \varphi E \quad \triangleright_{\varphi} \quad \frac{D_1}{\perp}$$

Let us consider a system having \rightarrow - and φ -rules with their standard reductions. In the system, the following derivation is proved.

7) For the proof, the reader can consult Girard (1987, pp. 135-150). The cases concerning \triangleright_{\in} are similar to the cases concerning \triangleright_{\wedge} .

$$\frac{\frac{\frac{[\varphi]^1 \quad [\varphi]^1}{\perp} \rightarrow I,1 \quad \varphi E}{\varphi \rightarrow \perp} \text{def} \quad \frac{\frac{[\varphi]^2 \quad [\varphi]^2}{\perp} \varphi E \quad \frac{\perp}{\varphi \rightarrow \perp} \rightarrow I,2}{\neg \varphi} \text{def}}{\perp} \rightarrow E \quad \frac{\frac{\perp}{\varphi} \varphi I,3}{\varphi} \rightarrow E}{\perp} \rightarrow E$$

The derivation does not represent any paradox but satisfies PCP_T and PCP_R . It shows that both criteria overgenerate. However, we should not say that \triangleright_φ is the culprit of generating a non-terminating reduction because \triangleright_φ relies on the form of φI -rule. φI -rule already contains a constant that it is intended to introduce in its conclusion. A reduced derivation by \triangleright_φ still has a maximum formula φ . The reduction process does not terminate.

Similarly, the form of Crabbé reduction, \triangleright_\in , relies on the form of $\in I$ -rule. Since $\in I$ -rule has the premise $\neg Z_b \in Z_b$ whose degree is greater than that of the conclusion $Z_b \in \{x \in b \mid \neg x \in x\}$, the reduced derivation has the formula $\neg Z_b \in Z_b$ by eliminating the maximum formula $Z_b \in \{x \in b \mid \neg x \in x\}$. Therefore, the real problem may not be Crabbé reduction itself but be the form of $\in I$ -rule. From these cases, we conclude that not every problem of overgeneration is raised by inadmissible reductions. When we focus on the way to assess admissible reductions, the spoiler test can be regarded as a promising criterion for admissible reductions.

6. Conclusion and other alternatives

In this article, we introduce Schroeder-Heister and Tranchini's triviality test and two main problems of the test: the obscurity of a test method and the circularity problem. In order to solve the two problems, we first have distinguished between the standard and auxiliary reduction procedures and proposed the way to have a base set of standard reduction procedures. Second, we have introduced the spoiler test as an improvement on the triviality test. The spoiler test can be a promising criterion for admissible reduction procedures.

Of course, there are other possibilities. It is often claimed that a natural deduction system is isomorphic to a sequent calculus system if the generalized eliminations are used. (Cf.

Tennant (2002), Negri and von Plato (2001) and von Plato (2011)). Especially, Negri and von Plato (2001, Ch.1 and Ch. 8) propose an inductive definition of the translation algorithm between an intuitionistic natural deduction system with generalized elimination rules and an intuitionistic sequent calculus system with independent contexts. Let Γ, Δ, Θ be a finite multiset, i.e. a list with multiplicity but no order, of assumptions in sequent calculus. We use a binary derivation symbol \Rightarrow and ' $\Gamma \Rightarrow \varphi$ ' means that the antecedent Γ derives the succedent φ . For convenience, we introduce the following left and right rules for \rightarrow , and cut-rule in sequent calculus.

$$\frac{\Gamma \Rightarrow \varphi \quad \Psi, \Delta \Rightarrow \sigma}{\varphi \rightarrow \Psi, \Gamma, \Delta \Rightarrow \sigma} L_{\rightarrow} \rightarrow \frac{\varphi, \Gamma \Rightarrow \Psi}{\Gamma \Rightarrow \varphi \rightarrow \Psi} R_{\rightarrow} \rightarrow \frac{\Gamma \Rightarrow \varphi \quad \varphi, \Delta \Rightarrow \Psi}{\Gamma, \Delta \Rightarrow \Psi} Cut$$

φ in *Cut*-rule is called a *cut-formula*. The reduction process for \rightarrow in the style of generalized elimination rule is as follows.

$$\frac{\frac{\frac{[\varphi]^1}{D_1} \rightarrow I_1 \quad \Psi}{\varphi \rightarrow \Psi} \quad \frac{D_2 \quad D_3}{\varphi \quad \sigma} \rightarrow E}{\sigma} \quad \triangleright_{\rightarrow} \quad \frac{D_2 \quad \varphi \quad D_1 \quad \Psi \quad D_3}{\sigma}}$$

$\triangleright_{\rightarrow}$ -reduction in natural deduction is translated as below by Negri and von Plato's algorithm.

$$\frac{\frac{\frac{D'_1}{\varphi, \Gamma \Rightarrow \Psi} R_{\rightarrow} \quad \frac{D'_2}{\Delta_1 \Rightarrow \varphi} \quad \frac{D'_3}{\Psi, \Delta_2 \Rightarrow \sigma}}{\varphi \rightarrow \Psi, \Gamma, \Delta_1, \Delta_2 \Rightarrow \sigma} L_{\rightarrow} \quad \frac{\frac{D'_2}{\Delta_1 \Rightarrow \varphi} \quad \frac{D'_3}{\varphi, \Gamma \Rightarrow \Psi}}{\Gamma, \Delta_1 \Rightarrow \Psi} Cut \quad \frac{D'_3}{\Psi, \Delta_2 \Rightarrow \sigma}}{\Gamma, \Delta_1, \Delta_2 \Rightarrow \sigma} Cut \quad \triangleright_{\rightarrow} \quad \frac{\frac{D'_2}{\Delta_1 \Rightarrow \varphi} \quad \frac{D'_3}{\varphi, \Gamma \Rightarrow \Psi}}{\Gamma, \Delta_1 \Rightarrow \Psi} Cut \quad \frac{D'_3}{\Psi, \Delta_2 \Rightarrow \sigma}}{\Gamma, \Delta_1, \Delta_2 \Rightarrow \sigma} Cut$$

As the maximum formula $\varphi \rightarrow \Psi$ is eliminated by $\triangleright_{\rightarrow}$ -reduction in the original derivation, the cut-formula in the translated derivation is removed by the above reduction. If a natural

deduction with generalized elimination rules has an isomorphic translation to a sequent calculus, the isomorphic translation of normal derivations preserves the order of rules such that an introduction rule turns into corresponding right rule and an elimination rule into a left rule. The isomorphic translation guarantees the correspondence between normal and cut-free derivations. In this sense, if one of the main roles of an admissible reduction is to eliminate the maximum formula or to lessen the degree of it, an elimination of the cut-formula or a lessening of the degree of it plays the same role.

One of the interesting points is that Ekman's reduction stated in generalized form seems to be a cut-making process. Let us consider the following form of the Ekman reduction stated in generalized form with open assumptions.

$$\frac{\frac{\frac{D_1}{[\varphi \rightarrow \Psi] \quad \varphi \quad [\Psi]^1} \rightarrow E_1 \quad D_2}{\sigma} \rightarrow E_2}{\sigma} \rightarrow E_2 \quad \triangleright_{E_g} \begin{array}{l} D_1 \\ \varphi \\ D_2 \\ \sigma \end{array}$$

The derivation is translated as below.

$$\frac{\frac{\frac{D'_1}{\Gamma \Rightarrow \varphi \quad \Psi \Rightarrow \Psi} L \rightarrow \quad D'_2}{\Gamma, \varphi \rightarrow \Psi \Rightarrow \Psi} \quad \varphi, \Delta \Rightarrow \sigma}{\varphi \rightarrow \Psi, \Psi \rightarrow \varphi, \Gamma, \Delta \Rightarrow \sigma} L \rightarrow \quad \text{reduces to} \quad \frac{\frac{D'_1}{\Gamma \Rightarrow \varphi} \quad \varphi, \Delta \Rightarrow \sigma}{\Gamma, \Delta \Rightarrow \sigma} \quad D'_2 \quad \text{Cut}$$

Interestingly, the translated derivation on the left side has no cut-formula but after applying the reduction process, the cut-formula φ appears on the right side derivation. One who accepts that there is an isomorphic translation between natural deduction and sequent calculus can claim that the Ekman reduction is wrong because it is a maximum (or cut) formula-making process. On the other hand, one may say that the case shows the failure of the isomorphic translation between natural deduction and sequent calculus. Further investigation is needed to clarify the relation between two systems and to pursue the criterion for admissible reductions with regard to the isomorphic translation.

References

- Dummett, M. (1991). *Logical Basis of Metaphysics*. Harvard University Press.
- Ekman, J. (1998). Propositions in propositional logic provable only by indirect proof. *Mathematical Logic Quarterly*. vol. 44(1), pp. 69-91.
- Gentzen, G. (1935). Investigations into logical deduction. In M. E. Szabo (Eds.), *The Collected Papers of Gerhard Gentzen*. Amsterdam and London:North-Holland, pp. 68-131.
- Girard, J. (1987). *Proof Theory and Logical Complexity*. Bibliopolis.
- Kreisel, G. (1971). A survey of proof theory II. In J. E. Fenstad (Ed.), *Proceedings of the Second Scandinavian Logic Symposium*. Amsterdam: North-Holland, pp. 109-170.
- Negri, S. and von Plato, J. (2001). *Structural Proof Theory*. Cambridge: Cambridge University Press.
- Prawitz, D. (1965). *Natural Deduction: A Proof-Theoretical Study*. Dover Publications.
- Prawitz, D. (1971). Ideas and result in proof theory. In J. E. Fenstad (Ed.), *Proceedings of the Second Scandinavian Logic Symposium*. Amsterdam: North-Holland, pp. 235-308.
- Schroeder-Heister, P. and Tranchini, L. (2017). Ekman's paradox. *Notre Dame Journal of Formal Logic*. vol. 58(4), pp. 567-581.
- Schroeder-Heister, P. and Tranchini, L. (2018). How to Ekman a Crabb'e-Tennant. *Synthese*. <https://doi.org/10.1007/s11229-018-02018-3>.
- Tennant, N. (1982). Proof and paradox. *Dialectica*. vol. 36, pp. 265-296.
- Tennant, N. (1995). On paradox without self-reference. *Analysis*. vol. 55, pp. 199-207.
- Tennant, N. (2002). Ultimate normal forms for parallelized natural deduction. *Logic Journal of IGPL*. vol. 10(3), pp. 299-337.
- Tennant, N. (2016). Normalizability, cut eliminability and paradox. *Synthese*. <https://doi.org/10.1007/s11229-016-1119-8>.
- Tennant, N. (2017). *Core Logic*. Oxford University Press.
- von Plato, J. (2011). A sequent calculus isomorphic to Gentzen's natural deduction. *The Review of Symbolic Logic*. vol. 4(1), pp. 43-53.

“The Identity of Proofs and The Requirement of a Proper Reduction”에 대한 논평

소한중
고려대

증명론적 관점에서 형식 체계의 도출들은 비형식적인 증명들을 포착하기 위한 형식적 상대물들로 간주된다. 어떤 도출에 적용되어서 같은 결론을 가지는 다른 도출을 낳는 절차들을 환원(reduction)이라고 하자. 환원들의 어떤 집합 R 이 주 어지면 R 에 대한 반사 추이적 폐포(reflexive, transitive closure)로서 환원가능성 관계 $> R$ 를 정의하고 다시 $> R$ 에 대한 대칭 폐포(symmetric closure)로서 도출들 사이에 동치관계 $\sim R$ 를 정의하자. 프라우츠는 다음과 같이 추측한다.

증명의 동일성에 대한 추측 (i) 동일한 동치류들에 속하는 도출들은 모두 같은 증명을 표상하며, (ii) 동일한 증명을 표상하는 도출들은 모두 같은 동치류에 속한다.

테넨트는 역설들이 종료되지 않는 환원 열(reduction sequence)을 가지는 도출로 표상된다고 말함으로써 역설을 특징지으려고 한다. 에크만은 종료되지 않는 환원 열을 만들어 내지만 그 자체로는 문제없어 보이는 어떤 환원을 제시하였으며 이는 테넨트의 분석이 부적절함을 보이는 것으로 이해된다. 슈뢰더하이스터와 트란치니는 에크만의 환원이 적절한 환원인 아님을 보이고자 하며 프라우츠의 추측에 기반하여 허용 가능한 환원에 대한 다음과 같은 검사법을 제시한다.

사소성 검사법 환원 집합에 새로운 환원이 추가되었을 때 증명들 사이의 동일성 관계를 사소하게 만들지 않아야 한다. 즉 환원들 사이에 동치관계를 사소한 관계로 만드는 환원은 받아들일 수 없다.

최승락 선생님(이하 필자)은 사소성 검사법이 동일성 추측 중 더 불확실한 부분인 (ii)에 의존하고 있다는 의혹을 제기하며, 이에 더해서 사소성 검사법에 두 가지 주요한 문제가 있다고 말한다. 첫째는 환원의 허용가능성에 대한 평가가 환원 집합에 상대적으로 달라지기 때문에 그 검사법이 불분명하다는 것이며, 둘째는 어떤 환원이 허용가능하지 않음을 보이기 위해서 허용 가능한 환원들의 집합을 미리 가지고 있어야 한다는 점에서 그 검사법이 순환적이라는 것이다. 이후 필자는 그가 제시한 문제들에 대한 해결책으로써 허용가능성을 평가하는데 기본이 되어야 환원 집합을 사소성 검사법과 독립적으로 제시하고 이를 정당화하며, 이에 더하여 사소성 검사법에 대한 대안으로 스포일러 검사법을 제안한다.

1. 필자는 슈뢰더하이스터와 트란치니가 추측 (i)과 (ii)를 모두 받아들이고 있다고 지적하며 이것을 사소성 검사법의 문제 중 하나로 거론한다. 하지만 논평자가 보기에 사소성 검사법이 (ii)를 필요로 한다는 게 그렇게 분명치 않다. 만약 사소성 검사법이 (ii)를 사용하지 않고 정당화될 수 있다면 오히려 그렇게 이해해줘야 하며, 그럴 수 있을 경우 슈뢰더하이스터와 트란치니가 실제로 (ii)에 동의한다고 하더라도 그 사실이 사소성 검사법에 대한 문제라고 하긴 어렵다. 논평자가 이해하기로 사소성 검사법은 다음과 같은 생각에 기반한다.

- (a) 기존 환원 집합에 환원 \triangleright 을 추가해 환원 집합 R 을 얻었을 때 R 로부터 정의되는 동치 관계가 사소한 관계가 된다고 하자 즉 임의의 문장 ϕ 에 대해서 ϕ 를 결론으로 가지는 모든 닫힌 도출들이 하나의 동치류에 속하게 된다고 하자.
- (b) 동치인 도출은 같은 증명을 표상하기 때문에 ϕ 에 대한 증명은 모두 같은 증명이 된다.
- (c) 하지만 증명들 간의 동일성 관계는 사소한 관계가 아니다 다시 말해서 어떤 문장 ϕ 에 대한 증명들이 모두 같은 것은 아니다.
- (d) 따라서 \triangleright 는 받아들일 수 있는 환원이 아니다.

동일성 추측이 역할을 하는 것은 (b)이며 (b)를 위해서는 (i)만이 필요하다. (c)를 보이기 위해 슈뢰더하이스터와 트란치니는 서로 다른 증명을 표상하는 것으로 여겨지는 두 도출을 전형적인 예시(typical example)로 제시했다. 필자가 지적하듯이 슈뢰더하이스터와 트란치니는 합류성 속성이 성립하는 체계에서 그것들이 동치가 아님을 보임으로써 (ii)를 방증하려고 했다. 그러나 슈뢰더하이스터와 트란치니는 그 도출들이 서로 다른 증명을 표상한다는 것을 보이기 위해 (ii)에 의존하지는 않았다.

2. 논평자가 보기에 필자가 지적한 사소성 검사법의 두 주요 문제는 슈뢰더하이스터와 트란치니의 주장을 그들의 의도보다 훨씬 강하게 읽은 것에서 기인한다.

필자는 슈뢰더하이스터와 트란치니가 허용가능한 환원을 거부하기 위한 필요충분조건으로서 사소성 검사법을 제시한 것으로 해석하고 있다. 그러나 슈뢰더하이스터와 트란치니는 어떤 환원을 거부하기 위한 충분조건으로서 사소성 검사법을 언급했을 뿐이다: 다시 말해서 슈뢰더하이스터와 트란치니는 어떤 환원을 허용 가능하지 않다고 말할 수 있는 하나의 기준으로서 사소성 검사법을 제시한 것이며, 사소성 검사법을 통과하는 환원을 모두 허용해야 한다고 주장하는 것도 아니고 사소성 검사법이 환원들의 허용가능성을 평가하기 위한 유일한 기준이라고 주장 하지도 않는다.

그런 점에서 슈뢰더하이스터와 트란치니에 대한 필자의 공격은 약간 부당한데 명시적으로 얘기하진 않지만 그들도 소위 표준환원들을 사소성 검사법을 위해서 기본이 되는 환원으로 염두에 두고 있는 듯이 보이기 때문이다. 그들이 하는 설명이나 그들이 드는 예시들은 모두 표준환원들을 사용하고 있다. 그 표준환원들의 집합이 필자가 문제들에 대한 해결책으로서 제시한 바로 그 집합이며, 논평자의 생각으로는 슈뢰더하이스터와 트란치니 역시 필자와 같은 고려들을 통해서 표준 환원들을 받아들였을 것이다.

사소성 검사법이 환원을 거부하기 위한 충분조건만을 제시하는 것으로 이해한다면 사소성 검사법을 사용하기 위해 기본이 되는 표준환원들의 집합을 사소성 검사법에 따라서 평가해야 할 필요는 없으며 필자가 지적한 순환성 문제는 발생하지 않는다. 그리고 표준환원들의 허용가능성이 다른 식으로 정당화할 수 있다면 표준환원 집합과 다른 환원 집합에 기초했을 때 환원들의 허용가능성에 대한 평가가 달라진다는 것 역시 슈뢰더하이스터와 트란치니에게 문제가 될 것 같지는 않다.

3. 필자가 제안하는 스포일러 검사법은, 간단히 말해서, 기존 환원 집합에 \triangleright 을 추가해 환원 집합 R' 을 얻었을 때 R' 로부터 정의되는 환원가능성 관계 $>R'$ 가 합류성 속성(confluence property)을 만족하지 않는다면 \triangleright 을 받아들일 수 없다는 것이다. 필자는 그의 검사법을 사소성 검사법보다 더 나은 검사법이라고 주장하지만 논평자가 보기에는 정말로 그런지 꽤 의심스럽다. 필자가 스포일러 검사법을 제안하는 동기는 다음과 같아 보인다.

(7) 사소성 검사법은 증명의 동일성에 대한 프라우츠의 추측(ii)에 의존하며, 프라우츠는 강한 정규화 정리(strong normalization theorem)를 통해 (ii)를 방증한다.

(ㄴ) 환원가능성 관계가 합류성 속성을 만족하지만 강한 정규화정리가 성립하지 않는 체계들이 있으며 그런 체계들에 대해서 사소성 검사법은 적용될 수 없다.

(ㄷ) 따라서 합류성 속성에 기반해서 환원들의 허용가능성을 평가하는 게 더 적절하다.

(ㄱ)에 대해: 앞서 얘기한 논평자의 의견이 맞다면, 사소성 검사법은 (i)에만 의존한다.

(ㄴ)에 대해: 프라우츠는, 강한 정규화 정리에 따르면 서로 다른 정규형태(normal form)를 가지는 두 도출이 동치가 아니라는 사실과 서로 다른 정규형태를 가지는 두 도출이 서로 다른 증명을 나타내는 것처럼 보인다는 예시를 통해 (ii)를 방증하려고 했다. 서로 다른 정규형태를 가지는 두 도출이 동치가 아니라는 것은 환원가능성 관계가 합류성 속성을 만족한다는 것만으로도 따라 나온다. 따라서 사소성 검사법이 (i)에 의존한다고 하더라도 합류성 속성만을 만족하는 체계에서 그 검사법을 적용하는 것이 강한 정규화 정리가 성립하는 체계들에서 그 검사법을 사용하는 것보다 더 문제가 될 것 같지는 않다.

필자는 합류성 속성을 만족하지 않게 만드는 환원을 증명의 동일성을 망치(spoil)는 환원이라고 부르며 그의 검사법을 증명의 동일성 개념과 관련해서 환원을 거부하거나 허용하는 그런 검사법으로 소개하고 있다. 그러나 합류성 속성은 체계에서의 도출들 사이의 환원가능성 관계에 대해 그저 얘기할 뿐이다. 환원가능성이 증명의 동일성과 어떻게 연관되는지 그리고 환원이 증명의 동일성을 망친다는 게 무슨 의미인지를 파악하기에는 본문의 설명만으로는 너무 부족하다.

Knowledge as Objectively Justified Belief

Byeong Deok Lee
Sungkyunkwan Univ.

Abstract

According to Lehrer's defeasibility account of knowledge, we can understand knowledge as undefeated justified true belief. But this account faces many serious problems. One important problem is that one can hardly determine whether one's justification of a belief is undefeated. Another problem is that it can hardly accommodate the externalist intuition that the epistemic status of a belief is not entirely determined by factors that are internal to the subject's perspective. The purpose of this paper is to offer an alternative account of knowledge which can successfully deal with these problems. This account is based on a Sellarsian social practice theory of justification, according to which a proposition is justified just in case we can answer all objections raised against it in our social practice of demanding justification and responding to such demands. On the basis of this social practice theory of justification, I argue that we can understand knowledge as objectively justified belief. I also argue that this new account of knowledge can successfully deal with the problems with the defeasibility account.

1. Preliminary Remarks

The purpose of this paper is to offer a new account of knowledge which explains knowledge in terms of objectively justified belief. There are important similarities between my account of knowledge and Keith Lehrer's defeasibility account [1969; 2000; 2003; 2005]. On the latter, we can understand knowledge as undefeated justified belief.¹⁾ One important

problem with this defeasibility account is that one can hardly determine whether one's justification of a belief is undefeated. Another problem is that it can hardly accommodate the externalist intuition that the epistemic status of a belief is not entirely determined by factors that are internal to the subject's perspective. I accept a Lehrerian conception of coherentist justification, according to which our concept of coherence is to be understood in terms of 'answering all objections' or 'beating all competitors'. In addition, I agree with the basic idea of the defeasibility approach: knowledge that p requires that there be no hidden truths that defeat one's evidence for believing that p . As we will see, however, my account of knowledge can successfully deal with the problems of the defeasibility approach.

In addition to the defeasibility approach, there are other competing approaches to knowledge, such as the causal approach (e.g., Goldman [1967; 1976]), the reliabilist approach (e.g., Goldman [1986]), and the virtue-theoretic approach (e.g., Sosa [2007]). But I will not discuss those competing approaches in this paper for the following reasons. As noted, the main goal of this paper is to offer a new account of knowledge which is similar to Lehrer's account of knowledge. And a proper discussion of other competing approaches would take me too far afield from the main goal of this paper.

This paper proceeds as follows. In section 2, I briefly discuss Lehrer's defeasibility account of knowledge. As we will see, although there are important similarities between his account and mine, there are important differences as well. In section 3, on the basis of a Sellarsian social practice theory of justification, I argue that we can understand knowledge as objectively justified belief. Finally, in section 4, I argue that my account of knowledge can successfully deal with the problems of Lehrer's defeasibility account. And I bring out the distinctive characteristics of my account further by comparing it with Michael Williams's contextualist view.

1) The concept of epistemic defeasibility was originally introduced by Roderick Chisholm [1964], based on an analogy with the ethical concept of a defeasible obligation. And the defeasibility account of knowledge is also defended by many philosophers such as the following: Lehrer and Paxson [1969], Klein [1980; 2008], Swain [1981; 1998], Pollock [1986], Hilpinen [1988], and De Almeida & Fett [2016].

2. Lehrer's Defeasibility Account of Knowledge

According to Lehrer's defeasibility account, knowledge is undefeated justified true belief. And he understands justification in accordance with his coherence theory of justification. Thus, let me briefly explain his theory of justification.

To begin with, what is coherence? Coherence is usually understood as a matter of how the components in a system of beliefs fit together or dovetail with each other. Except for this sort of metaphorical characterization, however, there is no generally accepted definition of coherence. In this frustrating situation, Lehrer offers us a very illuminating account of this elusive concept. On his proposal, we can understand coherence in terms of 'answering all objections' or 'beating all competitors'. On this view, one's belief that *p* coheres with one's evaluation system just in case one can answer all objections raised against it on the basis of one's evaluation system (see Lehrer [2000: 170]). In addition, on Lehrer's view, justification consists of two levels. The first level is *personal justification*, and the second level is *undefeated justification*. A subject, *S*, is personally justified in believing that *p* just in case she can answer all objections raised against it on the basis of her evaluation system. And "undefeated justification is personal justification that is not based on error" (Lehrer [2000: 153]).

At this point, it is important to note that even undefeated justification is based on personal justification. This means that Lehrer adopts a subjective model of justification. And based on this subjective model of justification, he defines knowledge as follows. *S* knows that *p* just in case the following four conditions hold:

- (i) It is true that *p*.
- (ii) *S* believes that *p*.
- (iii) *S* is personally justified in believing that *p*.
- (iv) *S* is justified in believing that *p* in a way that is undefeated.²⁾

2) Lehrer distinguishes between belief and acceptance (see [Lehrer 2000: 13]). On his view, belief may result from the pursuit of some purpose, but it is not *defined* in terms of any purpose, whereas acceptance requisite to knowledge is a special kind of acceptance for an intellectual purpose concerned with truth. For this reason, on his view, knowledge requires an acceptance condition rather than a belief condition. Fortunately, this distinction would not affect the main arguments of this paper. Thus, I will ignore this distinction in this paper for the sake of brevity.

One important thing to note about the above account is that the fourth condition of knowledge (that is, the no-defeater condition) makes the first condition (that is, the truth condition) superfluous. The reason is straightforward. Suppose that S is personally justified in believing that p on the basis of evidence e . Then this justification is defeated if there is a true proposition d such that the conjunction of e and d fails to justify S in believing that p . And if ' p ' happens to be false, there is a true proposition d (namely 'not- p ') that defeats the justified status of S's belief that p . This is because the conjunction of e and not- p fails to justify S's belief that that p (see [Lehrer 2000: 171-172]). But we can hardly rule out the possibility that any given (contingent) proposition might turn out to be false, no matter how strongly it is currently justified. If so, how can one's personal justification of a belief be connected with truth in a way that yields knowledge?

Lehrer addresses the above question by appealing to what he calls 'the Principle of Trustworthiness'.

T. I am trustworthy (worthy of my own trust) in what I accept with the objective of accepting something just in case it is true. [Lehrer 2000: 138]

On Lehrer's view, S is personally justified in believing that p just in case she can answer all objections raised against it on the basis of her evaluation system. S can answer an objection raised against her belief that p if it is more reasonable for her to accept that p than the objection. But her belief that p will not be truth-conducive if she is not trustworthy in judging this kind of comparative reasonableness. Therefore, to meet objections, S has to accept the Principle of Trustworthiness (see Lehrer 2005: 421)). How then does Lehrer defend this principle?

On his view, one's being justified in holding a belief depends on Principle (T). As a consequence, one cannot justify T on the basis of one's evaluation system. For this reason, Lehrer admits that (T) applies to itself. If, however, (T) applied to itself, then the justification of (T) would be circular. Nonetheless, Lehrer argues that the circularity involved in the justification of (T) is *virtuous*, rather than being vicious. Besides, he argues that the circularity involved in the justification of (T) is *explanatory*, rather than being argumentative

(see Lehrer 2000: 143]).

However, Richard Manning [2003] argues that Lehrer's retreat from justification to explanation fails in the context of defending (T). One important reason is this. Every explanation consists of two distinct components: the explanandum and the explanans. And when we are trying to explain something, we begin with the assumption that that thing needs to be explained. Thus, as Manning [2003: 208] points out, in the context of explaining something, we start with the assumption that we can take the explanandum for granted. As a consequence, we can explain why it is reasonable to accept (T) on the basis of the truth of (T), insofar as we can take the explanandum for granted. If so, the explanandum in question, namely that it is reasonable to accept (T), must already have a positive justificatory status. For this reason, when Lehrer retreats from circular justification to circular explanation, he must start by assuming what needs to be justified. In other words, what needs to be justified is assumed as the explanandum, that is, something which we can take for granted. According to Manning, this is not just changing a strategy for the game we have been playing for the purpose of justifying (T), but rather changing the game itself. The reason is straightforward. If the reasonableness of (T) were taken for granted, Lehrer would not have faced the demand for justifying (T) in the first place.

Despite the above objection, Lehrer [2003: 343-344] still wants to defend (T) in the following way. If one does not accept (T), one can explain nothing. This would be an epistemic disaster. By contrast, if one accepts (T), one can explain why one is reasonable to accept many things about which one deeply cares. And it is certainly better to explain as much as we can than to explain nothing. Therefore, this kind of explanatory power makes the circularity involved in (T) virtuous. But this line of argument amounts to the 'this or nothing' argument. And one important problem with this argument is that it does not give us any illuminating insight as to why the Principle of Trustworthiness is correct. For these reasons, I agree with Manning that Lehrer is not successful in justifying (T).

In addition, Lehrer's defeasibility account is not successful in avoiding Gettier-style counterexamples (see [Gettier 1963]). Let me explain. On Lehrer's view, undefeated justification is personal justification that is not based on any error. How then can one evaluate whether one's belief is justified in a way that is undefeated? Lehrer addresses this question by

appealing to what he calls ‘ultrasystem’. The ultrasystem of a person is what remains when everything false is eliminated from the person’s evaluation system. And a person’s justification of a belief is undefeated just in case the belief is justified on the basis of his ultrasystem. But then, how can a person be in a position to judge that his justification of a belief is based on the ultrasystem? In this regard, Lehrer writes:

One might worry that the ultracritic in the ultra justification game imagined above is merely a useless fiction because no one will actually be in a position to play the role of the ultracritic. The objection is, however, unwarranted. I may know enough about what another person accepts to play the winning role of an ultracritic against him. [Lehrer 2000: 162]

The above reply is implausible, however. Suppose that S’s personal justification for believing that p is undefeated throughout her entire life. Note that this is compatible with the possibility that some defeating evidence becomes available only after S’s death. In such a case, someone will be in a position to play the role of the ultracritic only after S dies. This means that during her lifetime, S is not in a position to judge that her personal justification for believing that p is based on an error. For this reason, the fact that S has done everything to reach the truth from her point of view is compatible with the possibility that the belief might be true merely by luck, that is, due to an accidental feature of her situation that has nothing to do with her evidence for believing that p . Along these lines, it can be argued that Lehrer’s defeasibility account is still susceptible to Gettier-style counterexamples.

The above point could be put somewhat differently. Lehrer’s defeasibility account faces a serious dilemma with regard to how to understand the no-defeater condition. If, on the one hand, this condition is understood as something that a subject can evaluate from the subject’s perspective, as argued above, Lehrer’s account is susceptible to Gettier-style counterexamples. If, on the other hand, this condition is understood in such a way that it implies the truth to the effect that it can avoid Gettier problems, we can argue that it is not something which one can epistemically evaluate from one’s subjective point of view. In other words, one can hardly evaluate whether the no-defeater condition holds or not. But one is entitled to say that S knows that p just in case one is entitled to say that the aforementioned four conditions of

knowledge hold. Accordingly, if one is not entitled to say that the no-defeater condition holds, then one is not entitled to say that S knows that p , either. This means that Lehrer's account fails to explain why we are often entitled to make knowledge attributions.³⁾

One more important problem with Lehrer's account is concerned with the externalist intuition that the epistemic status of a belief is not entirely determined by factors internal to the subject's perspective. For example, suppose that a child has an ordinary perceptual belief that an apple is in front of him. Suppose also that this child is unable to offer any suitable justification for this belief because he is not yet epistemologically sophisticated. Even in such a case, it seems, the child *knows* that an apple is in front of him. Our epistemic goal is usually understood as having true beliefs and avoiding false beliefs. Thus, a justified belief must be truth-conducive. The child's belief in question has been formed by his reliable perceptual mechanism to the effect that it is truth-conducive. Thus, according to reliabilists, true beliefs can amount to genuine knowledge even when the candidate knower is unable to offer any suitable justification. Following Robert Brandom [2000: 97], let us call this *the Founding Insight of reliabilism*. But Lehrer's defeasibility account of knowledge can hardly accommodate this insight. He writes:

The evaluation system of the person provides justification for what a person accepts by providing answers to some objections and neutralization of others. ① But what is essential is that the person understands how to meet and neutralize the objections to what he accepts. [Lehrer 2000: 200-201]

But the aforementioned child would not be able to meet the demand for justifying his belief in question because he is not yet epistemologically sophisticated. Hence, Lehrer's account can hardly explain why the child can be taken to know that an apple is in front of him.

3) De Almeida and Fett [2016] argue that the most important objections raised against the defeasibility of account so far are the ones put forward by Feldman [2003], Foley [2012], and Turri [2012], and also that the defeasibility account can successfully tackle these objections. In this paper, I do not want to dispute that the defeasibility account could handle those objections. On my view, however, the defeasibility account is still vulnerable to the problem of how to evaluate whether one's personal justification of a belief is undefeated.

3. Objectively Justified Belief

In this section, on the basis of a Sellarsian coherence theory, I will offer an alternative account of knowledge, according to which we can understand knowledge as objectively justified belief.⁴⁾

Above all, I accept a Lehrerian conception of coherence. As pointed out in the previous section, Lehrer explains coherentist justification in terms of answering all objections. On this view, one is justified in believing that p just in case one can answer all objections raised against it on the basis of one's evaluation system. But I accept a Sellarsian social practice theory of justification. On this view, our concept of *being justified in holding a belief* has been developed on the basis of our social practices of demanding justification and responding to such demands. Thus, we should understand our concept of justification in accordance with this social practice model of justification. As a consequence, we have to modify Lehrer's conception of coherence in accordance with this intersubjective model of justification. On this social practice theory of justification, we are justified in believing that p just in case we can answer all objections raised against it in our social practice of justification.

In addition, as Brandom argues, our social practice of justification requires the default-and-challenge structure of justification (see [Brandom 1994: esp. 204-206]). The reasons are roughly as follows. In the first place, the infinite regress of justification is not possible in our social practice of justification. In the second place, genuine doubt can take place only against a background of beliefs that are not doubted at the same time; to put the point another way, there must be a conceptual framework within which doubts and settlement of doubt take place. Therefore, the possibility of one's defending something requires that some claims be treated as having default justification unless we are given positive reasons to doubt them; that is, there must be some claims for which the burden of proof (or justification) shifts to any challenger.

With the above points in mind, let us now consider a certain belief that p which currently has a positive justificatory status in our social practice of justification. In this case, we cannot rule out the possibility that our belief that p might be based on some overriding error. For

4) For a detailed discussion and defense of the Sellarsian theory of justification, see Lee [2017; 2019a; 2020b]. See also Sellars [1963, esp. chapters 4 and 5; 1979].

example, it is possible that we are currently omitting or neglecting some defeating evidence against this belief. Or some evidence that is not currently taken to defeat our belief that p could be recognized as a defeater at a later time. Thus, we can lose the justification for believing that p by acquiring new evidence that together with the old evidence no longer justifies our belief that p . I agree with the defeasibility account on this point. Another thing worth pointing out is this: it is a regulative ideal of our epistemological pursuit that any rational being is not excluded from our justification practices. If any rational being shows us why we are mistaken in believing that p , we ought to give up the belief for the sake of achieving our epistemic goal of having true beliefs and avoiding false beliefs. In this sense, there are no membership restrictions to our social practice of justification to the effect that any rational being could, in principle, participate in our social practice of justification. Even the future generations are not excluded from this membership. Thus, if a certain belief of ours is based on some overriding error, then some rational being could, in principle, point out the error to us someday, and so the belief could thereby lose its positive justificatory status in our social practice of justification. If, on the other hand, a certain belief of ours is not based on any overriding error, then its positive justificatory status would not be lost in the future. In the latter case, we may say that it is not merely taken to be justified, but really justified. Along these lines, we can distinguish between *being merely taken to be justified* and *being really (or objectively) justified* by recourse to our forever ongoing, dynamic practice of justification. And on the basis of this distinction, we can argue that if a belief has a positive justificatory status in our social practice of justification and if its positive justificatory status does not depend on any overriding error, then it is not merely taken to be justified, but rather it is really (or objectively) justified.

Along the above lines, we may argue that objective justification can be understood as a special kind of intersubjective justification. For example, currently, our belief that the earth goes around the sun has a positive justificatory status in our social practice of justification. Suppose that there is no defeating evidence which we are omitting or neglecting for this belief. Then it follows that our belief that the earth goes around the sun is objectively justified. Admittedly, we cannot rule out the possibility that its positive justificatory status depends on some overriding error. Nonetheless, considering an overwhelming amount of

evidence for the belief, this is an extremely unlikely possibility. Therefore, we can make a *fallible* claim that our belief that the earth goes around the sun is objectively justified. In this regard, it is important to notice that *fallibility* is compatible with *being objectively justified*. In other words, the possibility of error does not imply the impossibility of objective justification. And insofar as this fallible claim, as a matter of fact, is not based on any overriding error, we are really (or objectively) justified in believing that the earth goes around the sun.

If the above considerations are on the right track, we can claim that S knows that *p* just in case we can claim the following three:

- (i) It is true that *p*.
- (ii) S believes that *p*.
- (iii) S is justified in believing that *p*.

The above three conditions are similar to the traditional tripartite account of knowledge. But we have to understand the above three conditions on the basis of the Sellarsian coherence theory and the deflationary conception of truth in the following way.

Let us begin with the second condition of knowledge, that is, the belief condition. Clearly, if S does not believe that *p*, we cannot say that she knows that *p*. Thus, this condition is virtually uncontroversial.

The third condition of knowledge is the justification condition. As mentioned before, the Sellarsian coherence theory understands justification in terms of a Sellarsian social practice model of justification. On this account, we are justified in believing that *p* just in case we can answer all objections raised against it in our social practice of justification. And we can evaluate whether S is (intersubjectively) justified in believing that *p* in accordance with this intersubjective model of justification. As an illustration, let us consider a Gettier case. Suppose that, as far as Smith can remember, Jones has always owned a Ford, and recently Smith got a ride in a Ford Jones claimed was his. Thus, on the basis of this evidence, Smith believes the following proposition:

- (1) Jones owns a Ford.

In addition, on the basis of (1), Smith also believes (2):

(2) Either Jones owns a Ford or Brown is in Barcelona.

But the fact of the matter is that Jones sold his old Ford and the Ford Jones currently drives is a rental car. Nonetheless, (2) is still true, because Brown is in Barcelona by mere coincidence. In this case, Smith is not blameworthy for believing (2). Nevertheless, we cannot say that he knows (2), because his belief in (2) is true merely by luck. And my account provides a proper explanation as to why Smith fails to know (2). Let me explain.

In the above case, Smith believes (1) on the grounds that as far as he can remember, Jones has always owned a Ford, and recently he got a ride in a Ford Jones claimed was his. However, there is defeating evidence, namely that the Ford Jones currently drives is in fact a rental car. Here the reason why Smith is not blameworthy for believing (2) is that the defeating evidence is not available to him. Nevertheless, the defeating evidence can be available to us, and so we can put forth the defeating evidence to Smith. In that case Smith's justification for (2) is undermined by the defeating evidence. For his belief in (2) is based on (1) which he is no longer justified in holding. And the fact that (2) happens to be true has nothing to do with Smith's evidence for believing (2). Therefore, although Smith's belief in (2) could be taken to be justified before the defeating evidence is available to us, he is not really justified in believing (2). In other words, my third condition of knowledge is not satisfied. This is why Smith does not know (2).

Here it might be worth considering a possible objection. On my account, we are justified in believing that p just in case we can answer all objections raised against it in our social practice of justification. But it might be objected that to require of any subject that he or she be able to answer all objections raised against a belief is to impose an unrealistic burden on the subject in order to be justified. At this point, it should be noted that being non-susceptible to this objection is an important merit of my account. On my social practice model of justification, it is not required of every subject that the subject alone should be able to answer all objections in order to be justified. For we can engage in the social division of epistemic labor in the matter of meeting objections. Let me illustrate this point. Suppose that our belief that p has a positive justificatory status in our social practice of justification because this belief is supported by adequate evidence e , and there is no defeating evidence as far as we know. Suppose also that S believes that p on the basis of evidence e . Suppose further that she

is not yet epistemologically sophisticated, and so she fails to meet some objections, such as a demand for justifying the reliability of her senses and a skeptical challenge about the external world. Under these conditions, we cannot say that she is not justified in believing that p . Of course, if our belief that p is successfully challenged to the effect that it loses its positive justificatory status, then S cannot be intersubjectively justified in believing that p . But insofar as this belief maintains its positive justificatory status in our social practice of justification, what is required of S is just that she believes that p in accordance with our epistemic norms. Note that insofar as she believes that p in a way that does not violate any epistemic norm, we cannot demand that she give up the belief, because it is adequately supported by her evidence e , and we do not have any defeating evidence. Therefore, S can be taken to be (intersubjectively) justified in holding the belief, even if she herself fails to meet some objections raised against the belief. I will say more on this point in the next section.

Let us now turn to the truth condition on knowledge. I endorse the deflationary conception of truth, which denies that truth is a substantial concept. Let me briefly explain the main motivation for this conception of truth. Consider the following equivalence schema:

(3) ' p ' is true if and only if p .

On the deflationary conception of truth, there is complete cognitive equivalence between the left-hand side of the biconditional above and its right-hand side. In this regard, observe that if one person asserts (4), and another person asserts (5), what they assert is exactly the same.

(4) Caesar was murdered.

(5) It is true that Caesar was murdered.

Along this line of thought, on the deflationary conception of truth, there is nothing else to say about truth other than what the truth predicate does; and the truth predicate serves only as a vehicle of generalization, semantic ascent, and certain other logical or expressive functions, and so truth is not a substantial concept.⁵⁾ For this reason, there is no substantial norm of truth

independent of the norms of justification. And this has an important consequence: the justification condition on knowledge makes the truth condition redundant for the following reason. Suppose that we are entitled to deny that ‘*p*’ is true. Then it is implied that we have some evidence which defeats S’s justification for believing that *p*. Thus, if we can deny that ‘*p*’ is true, we can also deny that S is intersubjectively justified in believing that *p*. This means that if we can hold that S is intersubjectively justified in believing that *p*, we can also hold that ‘*p*’ is true. When then is the role of the truth condition? It can play an *expressive* role: by saying that ‘*p*’ is true, we can express *explicitly* that we agree with S’s belief that *p*.

Finally, let me briefly explain how objective justification is related to objective truth by comparing the following three statements:

- (1) We are objectively justified in believing that *p*.
- (2) We are objectively justified in believing that ‘*p*’ is true.
- (3) It is objectively true that *p*.

If (6) holds, then (7) also holds due to the aforementioned equivalence schema. And there is no epistemically significant difference between (7) and (8). Let me explain. Suppose that our belief that *p* is objectively justified, and so its positive justificatory status will not be lost in our forever, ongoing, dynamic practice of justification. Now, recall that if we are entitled to assert that *p* in our social practice of justification, we are also entitled to assert that ‘*p*’ is true. Therefore, if our belief that *p* does not lose its positive justificatory status in our forever, ongoing, dynamic practice of justification, then there are no circumstances in which we can claim that ‘*p*’ is not true. To put it another way, if our belief that *p* is objectively justified, the possibility that there might be some defeating evidence against the belief is not epistemologically significant. Let me elaborate on this point a bit further.

Suppose that our belief that *p* is really (or objectively) justified in our social practice of justification due to the fact that its positive justificatory status does not depend on any overriding error. As mentioned before, our belief that the earth goes around the sun is very

5) It is beyond the scope of this paper to defend the deflationary conception of truth properly. For a detailed discussion and defense of this conception of truth, see Lee [2017]. And it is also noteworthy that this conception of truth is very influential among truth theorists in the philosophy of logic.

likely to be such a case. Under this condition, let us consider the possibility that its contradictory belief ‘ $\sim p$ ’ is equally really justified in some alien conceptual framework. If ‘ $\sim p$ ’ is equally justified in this way, there must be some reason R for believing that $\sim p$. Now, suppose that we can construe R as a reason for believing that $\sim p$. But this condition conflicts with our assumption that we are really justified in believing that p in our social practice of justification. This is because R would refute our previous justification for believing that p . This time, suppose that we cannot construe R as a reason for believing that $\sim p$. Then we have no reason to think that R is a reason for believing that $\sim p$. In this regard, it is noteworthy that anything which we cannot, even in principle, construe as a reason for believing a proposition is simply unintelligible in our epistemic discourse.⁶⁾

Along the above lines, we can argue that there is no epistemically significant difference between (7) and (8). And we can understand objective justification as a special kind of intersubjective justification. Recall that our belief that p is objectively justified just in case its positive justificatory status will not be lost in our forever, ongoing, dynamic practice of justification. If this is correct, we can understand knowledge as objectively justified belief.⁷⁾

4. Merits and Distinctive Characteristics of My Account

In this section, I argue that my account of knowledge can successfully handle the problems with Lehrer’s defeasibility account. And I bring out the distinctive characteristics of my account a bit further by comparing it with Williams’s view.

As pointed out in section 2, one important problem with Lehrer’s account is that it can hardly accommodate the Founding Insight of reliabilism. By contrast, my account of knowledge can accommodate this insight. Consider the following epistemic principle:

6) For a detailed discussion and defense of this claim, see Lee [2017].

7) If what I have argued so far are on the right track, then we can understand (8) in accordance with (7). But this should not be misunderstood as claiming that ‘ x is objectively true’ can be *defined* as ‘ x is objectively justified’. As noted before, if we are entitled to claim that p , then we are also entitled to claim that ‘ p ’ is true. This is why (7) holds just in case (6) holds. Nonetheless, on the deflationary conception of truth, truth is not a substantial concept. By contrast, justification is a substantial concept. Therefore, we need to distinguish an analysis of ‘is true’ from specifying the circumstances in which the predicate ‘is true’ can be rightly applied. And the fact that a belief has a positive justificatory status in our social practice of justification provides only a criterion for the application of the predicate ‘is true’, instead of an analysis of ‘is true’.

EP₁: Our perceptual judgments are generally reliable.

My Sellarsian coherence theory can defend EP₁. On this theory, it is inevitable to address any justification question on the basis of our social practice of justification. EP₁ is no exception, for it is a rational question whether or not EP₁ is correct (or justified). And we can defend EP₁, roughly, in the following way:

First, EP₁ has so far been successfully employed as an epistemic norm in our social practice of justification. Second, we have no positive reason to think that EP₁ is false or defective.

Under these conditions, it can be argued that we have no alternative but to accept EP₁ in order to rationally pursue our epistemic goal. Then we can take EP₁ as having default positive justificatory status in our social practice of justification. Let me elaborate on this point a bit further.

To begin with, as noted before, our social practice of justification requires the default-and-challenge structure of justification. In addition, our epistemic goal includes having true beliefs about the world around us. And it is a minimum presumption for our epistemic discourse that it is reasonable for us to pursue our epistemic goal. Now, it is important to recognize that we ultimately obtain information about the world *through our senses*. In other words, our senses are the ultimate source of information about the world. Thus, if we could not rely on our perceptual abilities in order to obtain information about the world, then we could not realize our epistemic goal. In this regard, it is noteworthy that denying the general reliability of our perceptual abilities is tantamount to closing our ultimate epistemic door to the world. Therefore, if we reject the general reliability of our perceptual judgments, we cannot get our epistemic pursuit off the ground. Therefore, unless we are given positive reasons to think that EP₁ is false or defective, we can argue that we have no alternative but to accept EP₁ in order to rationally pursue our epistemic goal.

On the basis of the above considerations, we can take EP₁ as enjoying default positive justificatory status in our social practice of justification, insofar as the aforementioned two conditions hold. Hence, it is reasonable for us to accept EP₁, unless we are given some positive reasons for not so doing.⁸⁾ If we can defend EP₁ in this way, we can argue that a

child's perceptual belief that there is an apple in front of him has default positive justificatory status in our social practice of justification. In this regard, it is important to note that the child, as a member of our society, can engage in the social division of epistemic labor in our social practice of justification. Accordingly, he himself does not have to justify EP_1 . This is because he can defer to some relevant experts in his society on the matter of justifying EP_1 .⁹⁾ For these reasons, we can say that unless we have some positive reasons to doubt this belief, the child's perceptual belief maintains its positive justificatory status in our social practice of justification. Hence, unless some challenger provides us with positive reasons to doubt this belief, we can say that the child knows that p . It is in this way that my account of knowledge can accommodate the Founding Insight of reliabilism.

In addition, my coherentist account can successfully handle the most important problem of the defeasibility account, namely that from one's subjective point of view, one can hardly evaluate whether the no-defeater condition of knowledge holds or not. In this regard, it is worth recalling that even if S's belief that p is undefeated throughout her entire life, the possibility that the belief might be defeated by some contrary evidence available in the future is still epistemologically significant. By contrast, my coherentist account can cope with the above problem because it understands the justification condition in terms of the Sellarsian social practice model of justification. Let me explain. On my account, we can claim that S knows that p just in case we can claim the following three:

- (i) It is true that p .
- (ii) S believes that p .
- (iii) S is justified in believing that p .

And we should understand (iii) in accordance with the aforementioned social practice model of justification. Accordingly, it is clear what we should do for evaluating condition (iii). We should evaluate whether S is justified in holding a belief in terms of our epistemic norms which have positive justificatory status in our social practice of justification. As a

8) For a detailed discussion and defense of this view, see Lee [2014; 2019b].

9) For a detailed discussion of this point, see Lee [2020a].

consequence, if we can judge that S believes that p in a way that does not violate any of our epistemic norms, we can say that condition (iii) holds. Therefore, when we evaluate the justification condition in the context of our social practice of justification, we have, at least in principle, no real difficulty in evaluating this condition.

Of course, one's knowledge attribution is fallible so that we cannot rule out the possibility that it might turn out to be false in the future. But it is important to note that at any point in time, we have no real difficulty in evaluating this condition. As mentioned before, if we are justified in believing that p in our social practice of justification, and if S believes that p in a way that does not violate any of our epistemic norms, we can say that condition (iii) holds. And insofar we are entitled to say so, we are also entitled to claim that S knows that p . In addition, as also pointed out before, the fallibility of such an epistemic evaluation is compatible with the case that S is really justified in believing that p . In other words, we can claim that our belief that p is objectively justified, while allowing the possibility that this claim might turn out to be false. For example, we can make a fallible claim that we are really justified in believing that the earth goes around the sun. For this reason, to endorse the justification condition, we don't have to rule out the possibility that this endorsement might be defeated someday. Hence, we can make a fallible knowledge claim, such as that we know that the earth goes around the sun. And we are entitled to make such a knowledge claim unless we are given positive reasons to doubt it.

At this point, it might be worth emphasizing the key difference between Lehrer's defeasibility account of knowledge and my account. As previously pointed out, even if S's belief that p is undefeated throughout her entire life, the possibility that the belief might be defeated is still epistemologically significant. By contrast, my Sellarsian social practice theory of justification can cope with this epistemic limitation by engaging in the social division of epistemic labor and also by appealing to our forever, ongoing, dynamic practice of justification. The first thing to note is that by engaging in the social division of epistemic labor, we can rely on our epistemic norms which have so far been successfully employed in our social practice of justification. In addition, as noted before, if our belief that p is really justified, and so its positive justificatory status will not be lost in our forever, ongoing, dynamic practice of justification, the possibility that there might be some defeating evidence

against the belief is not epistemologically significant. Therefore, to say that our belief that p will not lose its positive justificatory status in our forever, ongoing, dynamic practice of justification is tantamount to saying that ' p ' is objectively true. Along these lines, we can explain why there is an appropriate connection between justification as conceived by the Sellarsian coherence theory and objective truth.

Finally, let me bring out the distinctive characteristics of my account a bit further by comparing it with Williams's view [1991; 2001]. In a similar way as I have combined a Sellarsian view with the deflationary conception of truth, Williams holds a Sellarsian view while embracing the deflationary conception of truth. In addition, he also accepts a default and challenge conception of justification. But there are important differences as well. First, Williams [2001: 177] defends a partially externalist view, which clearly separates his Sellarsian view from coherentism. Second, Williams [2001: 171] endorses what he calls 'a pragmatic conception of epistemic norms', according to which epistemic norms are fixed by us in light of our practical interests, projects, and assessment of our situation. Third, Williams takes a contextualist approach to knowledge, according to which standards for claiming or attributing knowledge can vary depending on context. It is beyond the scope of this paper to fully discuss Williams's account of knowledge. Thus, let me confine myself to briefly mentioning three reasons why I do not accept his account.

First, I do not agree with Williams's partially externalist view that the justificatory status of a belief may depend on some non-doxastic contextual factor of which we are unaware. As argued before, our concept of *being justified in holding a belief* has been developed on the basis of our social practice of demanding justification and responding to such demands to the effect that our epistemic assessments are not external to such social practice. Thus, we should be capable of assessing whether or not our beliefs are justified within our social practice of justification, and we can only appeal to what is ascertainable intersubjectively in such assessments. For these reasons, the alleged non-doxastic contextual factor of which we are unaware has no place in our social practice of justification. In this regard, it is worth noting that my Sellarsian coherence theory is distinguished from standard internalist views for the reason that this intersubjective model of justification allows us to take into consideration evidence that might not be available to the subject. However, it should be noted that my

account is not an externalist view. For it denies that the epistemic status of a belief is determined in part by factors that are not accessible to any person. On my view, the epistemic status of a belief can be determined only by what is, at least in principle, ascertainable intersubjectively.

Second, I do not accept Williams's pragmatic conception of epistemic norms. Instead, I defend a coherentist justification for epistemic norms. Here it is noteworthy that there is an important reason to justify epistemic principles in terms of epistemic rationality, rather than in terms of practical rationality. Let me explain. The first thing to note is that epistemic justification (or rationality) is different from practical justification. What is distinctive about epistemic justification is that it is essentially connected with our epistemic goal of having true beliefs and avoiding false beliefs. On Kant's view (1788), we are rational beings who can engage in theoretical and/or practical reasoning in order to determine what to believe and/or what to do. And when we try to determine what to believe by engaging in theoretical reasoning, we take an *epistemic (or theoretical) point of view*, which is concerned (primarily) with *determining how things actually are in the world*. By contrast, when we try to determine what to do by engaging in practical reasoning, we take a *practical point of view*, which is concerned with *realizing what is desired or desirable*. Therefore, it is one thing to make an epistemic evaluation as to whether a certain proposition is true of the actual world; and it is another thing to make a practical evaluation as to whether doing a certain action is the means for realizing what is desired or desirable. To put the point another way, epistemic justification is essentially connected with our epistemic goal of having true beliefs and avoiding false beliefs, whereas practical justification has no direct bearing on this epistemic goal of truth. The second thing to note is that epistemic principles such as EP₁ are also essentially connected with the epistemic goal of truth. In this regard, it is important to recognize that we evaluate whether or not our beliefs are justified (at least in part) in accordance with such epistemic principles, and these epistemic principles are our means for achieving the epistemic goal of truth. Along these lines, we may argue that practical rationality has no direct bearing on our evaluation of whether epistemic principles such as EP₁ are conducive to the epistemic goal of truth, which is concerned with how things actually are in the world. Therefore, it is desirable to defend epistemic principles in terms of epistemic rationality,

rather than in terms of practical rationality. And if what I have argued so far are correct, my account explains why epistemic principles such as EP₁ are epistemically justified on the basis of a theory of *epistemic justification*, that is, a Sellarsian coherence theory. In other words, my account defends epistemic principles in terms of *epistemic rationality* rather than practical rationality.

Third, and finally, I do not agree with Williams's contextualist view. On this view, standards for claiming or attributing knowledge can vary depending on context. For example, it is possible that a person knows that he has hands in an ordinary context, but he may lose this knowledge in a context where he faces a certain skeptical challenge (see e.g., Williams [2001: 195]). By contrast, on my view, standards for claiming or attributing knowledge do not vary from context to context. Suppose that our epistemic norms like EP₁ are successfully challenged by the skeptic to the effect that they lose their positive justificatory status in our social practice of justification. Then, even in an ordinary context, we are not entitled to claim that we have hands. As I have argued elsewhere, however, we can meet such a skeptical challenge partly on the grounds that EP₁ has a default positive justificatory status in our social practice of justification and so we can shift the burden of proof to the skeptic.¹⁰⁾ In short, on my view, we can handle the challenge of skepticism without taking a contextualist approach.

To conclude, my account of knowledge, which is based on an intersubjective model of justification, is much more plausible than Lehrer's defeasibility account, which is based on a subjectivist model of justification. And we can understand knowledge as objectively justified belief.

References

- Brandom, Robert 1994. *Making It Explicit*, Cambridge: Harvard University Press.
- Brandom, Robert 2000. *Articulating Reasons: An Introduction to Inferentialism*, Cambridge: Harvard University Press.
- Chisholm, Roderick 1964. The Ethics of Requirement. *American Philosophical Quarterly*, 1: 147-153.
- De Almeida, Claudio and J.R. Fetf. 2016. Defeasibility and Gettierization: A Reminder. *Australasian Journal of Philosophy*, 94/1: 152-169.

10) It is beyond the scope of this paper to defend this view properly. For a detailed discussion and defense of this view, see Lee [2014]

- Feldman, Richard 2003. *Epistemology*, Upper Saddle River: Prentice Hall.
- Foley, Richard 2012. *When Is True Belief Knowledge?* Princeton: Princeton University Press.
- Gettier, Edmund L. 1963. Is Justified True Belief Knowledge? *Analysis*, 23: 121-23.
- Goldman, Alvin I. 1967. A Causal Theory of Knowing. *The Journal of Philosophy*, 64/12: 357-372.
- Goldman, Alvin I. 1976. Discrimination and Perceptual Knowledge. *The Journal of Philosophy*, 73/20: 771-791.
- Goldman, Alvin I. 1986. *Epistemology and Cognition*, Cambridge: Harvard University Press.
- Hilpinen, Risto 1988. Knowledge and Conditionals, in *Philosophical Perspectives 2*, ed. James Tomberlin, Atascadero: Ridgeview: 157-182.
- Kant, Immanuel 1788 (1996). Critique of Practical Reason, in *Practical Philosophy*, trans. and ed. Mary J. Gregor, Cambridge: Cambridge University Press: 135-271.
- Klein, Peter 1980. Misleading Evidence and the Restoration of Justification, *Philosophical Studies*, 37/1: 81-90.
- Klein, Peter 2008. Useful False Beliefs, in *Epistemology: New Essays*, ed. Quentin Smith, Oxford: Oxford University Press: 97-130.
- Lee, Byeong D. 2014. Epistemic Principles and Epistemic Circularity. *Philosophia*, 42: 413-432.
- Lee, Byeong D. 2017. The Truth-conduciveness Problem of Coherentism and a Sellarsian Explanatory Coherence Theory. *International Journal of Philosophical Studies*, 25: 63-79.
- Lee, Byeong D. 2019a. A Kantian-Brandorian View of Concepts and the Problem of a Regress of Norms. *International Journal of Philosophical Studies*, 27/4: 528-543.
- Lee, Byeong D. 2019b. "Gupta on Sellars's theory of perception." *Dialogue: Canadian Philosophical Review*. <https://doi.org/10.1017/s0012217319000180>
- Lee, Byeong D. 2020a. Ginsborg on a Kantian-Brandorian View of Concepts. *International Journal of Philosophical Studies*, 28/1: 56-74.
- Lee, Byeong D. 2020b. A Coherentist Justification of Induction. *Erkenntnis*. <https://doi.org/10.1017/s10670-019-00183-6>
- Lehrer, Keith 2000. *Theory of Knowledge*, Second Edition. Boulder: Westview Press.
- Lehrer, Keith 2003. Coherence, Circularity and Consistency: Lehrer Replies, in *The Epistemology of Keith Lehrer*, ed. Erik. J. Olsson, Dordrecht: Kluwer: 309-356.
- Lehrer, Keith 2005. Coherence and the Truth Connection, *Erkenntnis*, 63: 413-423.
- Lehrer, Keith and Thomas D. Paxson, Jr. 1969. Knowledge: Undefeated Justified True Belief, *Journal of Philosophy*, 66: 225-37.
- Manning, Richard 2003. The Dialectical Illusion of a Vicious Bootstrap, in *The Epistemology of Keith Lehrer*, ed. Erik. J. Olsson, Dordrecht: Kluwer: 195-216.
- Pollock, John 1986. *Contemporary Theories of Knowledge*, Totowa: Rowman & Littlefield.
- Sellars, Wilfrid 1963. *Science, Perception and Reality*, Atascadero: Ridgeview.
- Sellars, Wilfrid 1979. More on Givenness and Explanatory Coherence, in *Justification and Knowledge*, ed. George S. Pappas, Dordrecht: Reidel: 169-182.
- Sosa, Ernest. 2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge*, New York: Oxford

- University Press.
- Swain, Marshall 1981. *Reasons and Knowledge*, Ithaca: Cornell University Press.
- Swain, Marshall 1998. The Defeasibility Theory of Knowledge, in *Routledge Encyclopedia of Philosophy*, Vol. 5, ed. Edward Craig: 276-280.
- Turri, John 2012. In Gettier's Wake, in *Epistemology: The Key Thinkers*, ed. Stephen Hetherington, London: Continuum: 214-229.
- Williams, Michael 1991. *Unnatural doubts: Epistemological realism and the Basis of Scepticism*, Cambridge: Blackwell.
- Williams, Michael 2001. *Problems of Knowledge: A Critical Introduction to Epistemology*, Oxford: Oxford University Press.

이병덕 교수의 Knowledge as Objectively Justified Belief에 대하여

오성
서울대

해당 논문에서 이병덕은 레러(Lehrer)의 내부주의적 정합론을 소개하며 그가 ‘정합coherence’이라는 인식론적 개념에 대해 의미 있는 구체적 규정을 제시하고 있지만 순수한 내부주의적 정당화 모델에서 비롯하는 여러가지 한계를 가진다는 점을 지적한다. 그에 따라 저자는 셀라스적 사회 수행 모델에 의거해 일종의 간주관적(intersubjective) 정당화 개념을 제시하여 해결책을 모색하고 있다. 나는 본 논평에서 이병덕이 훌륭한 정합론적 모델을 제시하고 있기는 하지만 그가 제시하는 사회 수행 모델에 의거한 정당화 규정이 몇가지 우려를 불러일으킨다는 점을 지적하고 싶다.

먼저 순수하게 내부주의적인 정합론적 정당화 개념이 어려움에 처한다는 이병덕의 주장은 정당하다. 레러에 따르면 어떤 믿음이 정합적으로 정당화 된다는 것은 믿음의 주체가 모든 반박에 대답할 수 있다는 것 혹은 모든 경쟁 믿음을 압도한다는 것을 의미한다. 하지만 이러한 정합론적 정당화가 순수하게 1인칭적 내부주의 관점에서 규정될 경우 이러한 정당화는 믿음 소유자의 사후에 반박되어 정당화의 위상을 잃게 될 가능성을 내재하게 된다. 셀라스적 사회 수행 모델은 영원하고 지속적이고 동적인 정당화 수행 모델을 제시함으로써 어떻게 하면 정당화의 주체가 지속적으로 반박에 답변할 수 있는지를 보여주어 이러한 사후 반박 가능성이 정당화를 훼손할 가능성을 차단한다. 이러한 모델은 정합론적 정당화를 규정하는 데에 있어서 레러의 것보다 설득력 있는 것임에는 분명하다.

이병덕이 스스로를 정당화와 관련해서 외부주의자가 아니라 간주관주의자로 규정하는 부분은 상당히 흥미롭다. 그의 정당화 개념은 사회적인Default-and-Challenge 모델에 기반하고 있으나 어디까지나 사회 구성원들의 내적 정당화(구체적으로 주어지는 반박에 대한 응답)에 의거한 것이지 그 외의 외적 기준에 의거한 것이 아니라는 것이다. 이는 우리의 내부주의적 직관을 충분히 옹호하면서 1인칭적 관점의 한계를 보완하는 좋은 시도로 보인다. 그러

나 그 와중에 어째서 그가 신빙론적인 직관을 만족시키는 설명을 추구해야 한다고 주장했는지는 다소 의아한 부분이기는 하다. 그는 논문의 전반부에서는 어린 아이의 지각적 믿음이 정당화 된다고 생각하는 우리의 관점이 신빙론적 직관에 의거해 있다고 하면서 후반부에는 그러한 사실상 간주관적 정당화 모델을 통해 어린아이의 지각적 믿음이 정당화 되는 이유를 설명하는데 이러한 논의의 흐름에 사실상(사소하지만) 다소간의 비일관성이 보인다.

한편으로 어린 아이의 지각적 믿음이 정당화 되기 위해서는 간주관적인 정당화 모델이 도입되어야 한다는 그의 관점이 다소간에 강하게 느껴지는 것이 사실이다. 그는 ‘사과가 앞에 있다’는 어린 아이의 믿음이 인식론적으로 정교하지 못하기 때문에 주관적 모델 하에서는 정당화의 기준을 만족시키지 못한다고 이야기하는데 이는 개별 지각적 믿음이 지각 경험에 의해서 *즉각적으로* 정당화된다는 우리의 일상적 직관과는 사뭇 간극이 있는 견해인 듯 보인다. 물론 우리는 지각적 판단이 믿을만하지 않을 회의론적 가능성을 고려해야 하고, 그런 측면에서 “우리의 지각적 판단이 일반적으로 믿을만하다”는 원칙(EP1)을 고려하지 않은 어린 아이의 지각적 믿음이 정당화되지 않는다고 볼 수도 있다. 하지만 원칙적으로 회의론적인 가능성은 어느 경우에도 열려 있고, 극단적인 회의론적 가능성은 간주관적인 정당화 관점을 도입하더라도 극복할 수 없는 것이며, 극단적인 회의론적 가정을 극복하기 위해 이병덕이 셀라스적 사회 수행 모델을 제시하며 그렇게 했듯이 ‘기본적 양의 정당화 위상 a default positive justificatory status’에 대해 고려한다면, 어린 아이가 *인식주관적인 관점에서* 그런 상태에 있지 않다고 단정지을 이유 역시 없다. 왜냐하면 적어도 어린아이에게도 자신의 지각적 판단을 의심할 이유가 없다는 점에서 어느 정도 양의 정당화 위상의 조건을 만족하기 때문이다. 달리 말하면 우리는 사회적 수행 모델을 가정하지 않고도 일종의 유사 Default-and-Challenge 정당화 구조를 구성할 수 있다. 이는 지각하는 주체 개인이 고립된 환경에서 스스로의 지각적 판단을 형성하는 능력에 대한 메타적인 믿음을 형성할 수 있다는 점을 고려했을 때에 더 분명해진다. 예를 들어서 어떤 사람이 시력이 나쁘다면 그 사람은 지속적으로 잘못된 시각적 믿음을 형성할 것이고 그에 따라서 원하는 인식론적 목적을 달성하는 데에 지속적으로 실패해 왔을 것이다. 그런 경험이 계속해서 반복되면 그는 “나의 시각이 별로 믿을만 하지 않다”는 믿음을 보다 명시적으로 가지게 될 것이고 그에 따라 그의 지각적 판단은 논박될 것이다. 이 경우 그가 “나의 그 외 감각은 믿을만하다”는 믿음도, *상대적으로 덜 명시적이지만*, 가지고 있다고 보는 것이 그렇게 무리한 결론은 아닌 것 같다. 덧붙여서 어린 아이가 특별히 이러한 인식론적 과업의 수행에 있어서 발달적으로 미진하다고 볼 이유 역시 없다. 왜냐하면 어린 아이가 지속적인 인식론적 목적 달성 실패에 따라 자신의 감각이 신뢰할만하지 않다는 메

타적 믿음을 가지게 되는 것이 불가능하지는 않기 때문이다.

고유한 1인칭적 특권성을 가진 믿음에 대해서는 셀라스적 모델이 어떤 방식으로 대답할 수 있을지도 상당히 궁금한 부분이다. 예를 들어 ‘나는 눈앞의 사과를 본다’는 의식 경험 자체에 대한 믿음의 경우 상식적으로는 반박의 여지가 없는 것으로 여겨지지만 이 또한 극단적인 회의론의 가능성에서는 자유로울 수 없다. ‘나는 눈앞의 사과를 본다’는 믿음의 소유에 대해 기본적 양의 정당화 위상을 인정한다 하더라도 데카르트의 악신 사례와 같이 현상 경험에 대한 믿음이 도전 받을 가능성은 늘 존재하기 때문이다. 이 경우 추가적 정당화가 사회적 수행 모델에 의해 부여될 수 있는 것인지는 다소 의문의 여지가 있다. 만일 가능하다고 한다면 이러한 종류의 믿음이 가진 1인칭적 특권성이 훼손되지는 않을지 우려된다.

내가 가지고 있는 마지막 우려는 지식의 조건을 규정할 때에 이병덕이 진리에 대한 축소주의적인 관점을 가정하고 있다는 사실과 관련한다. 이러한 축소주의적 관점에 따르면 어떠한 사실이 참이라는 것과 그것이 사회적 수행 모델에 의해 객관적으로 정당화 되었다는 것은 동일한 의미를 지니게 되는데 이는 객관적으로 정당화 되었다는 것이 ‘틀릴 가능성(fallibility)’과 양립 가능하다는 이병덕의 관점을 고려했을 때에 다소 어색하게 느껴지는 부분이다. 이에 대해 ‘축소주의적인 참 개념은 틀릴 가능성과 양립 가능하다’고 말한다면 과연 이러한 참 개념이 ‘참’에 대한 우리의 직관을 만족시키는가에 대해 의문의 여지가 있다. 또한 이런 객관적 정당화 개념에 의해 규정된 ‘참’ 개념을 견지할 경우 비록 이것이 회의주의의 가능성을 차단하지만, 달리 생각하면 궁극적인 회의론적 가능성에 대한 우리의 근본적 철학적 직관을 무력화하기 보다는 그저 무시하는 시도가 아닌가 하는 우려가 있다.

부작위는 행위와 도덕적으로 동일한 지위에 있는가

-대안가능성과 이중결과원리의 검증도구로서의 적합성-

김효은
한밭대학교

I. 들어가는 말

이 논문은 행위 및 부작위¹⁾에 도덕적 책임을 비대칭적으로 평가하는데 사용하는 ‘대안가능성 여부’와 ‘이중결과원리’가 검증도구로서 적합한지를 문제제기한다. 이를 통해 행위와 부작위의 비대칭성은 다소 임의적인 부분이 있다고 할 것이다. 대안가능성을 통해 인과적 경로를 검토하는 방식은 도덕적 책임 판단의 두 요소인 인식적 요소와 제어가능성 요소를 독립적으로 다루면서 전자를 배제 한다. 한 결과로 행위-부작위는 인과적 비대칭성에서 도덕적 비대칭성으로 자동 이행하게 만들며 도덕적 책임에 대한 설명을 일관적으로 하기 어렵게 만든다. 또, 이중결과원리에서 ‘의도’와 ‘단순한 예견’의 구분은 모호하다. 이 논문은 첫 번째 도구에 대해서는 수정된 시나리오를 고안하여 인식적 요소의 고려없이 설명이 불가능함을 제시하고, 두 번째 도구에 대해서는 동기와 예견을 나타내는 두 개념은 사실상 한 인지상태의 질적 정도의 차이일 수 있다라고 주장할 것이다. 결론적으로 행위-부작위의 도덕적 비대칭성은 인과구조의 서술을 동등하게 만들거나 동기의 내용을 구체화하는 경우 더 이상 유효하지 않을 가능성이 있다.

1) ‘부작위’(不作爲, omission)란 어떤 상황에서 요구되거나 기대되는 행위를 하지 않거나 또는 금지되는 행위를 하지 않는 것을 의미한다. 태만, 방조, 방관, 방치, 유기 등이 여기 모두 포함된다. 부작위라고 해서 모든 경우에 책임이 면제되는 것은 아니나 행위와 비교하여 상대적으로 책임을 덜 부여하곤 한다. 그러나 부작위도 의도적 부작위나 비의도적 부작위라 할지라도 중요한 사회적 기대나 의무와 관련한 부작위에는 행위만큼이나 중한 도덕적, 법적 책임을 부여한다. 서양 고대에서는 부작위를 소극적(negative) 의미의 행위 혹은 소극적 원인(negative cause)으로 해석하였다.

II. 행위-부작위의 인과경로와 도덕적 책임

행위-부작위의 도덕적 책임에 있어서의 비대칭성 주장은 죽임은 도덕적으로 비난받을만 하지만 죽게 내버려둠은 도덕적으로 허용되거나 상대적으로 덜 비난받는다라는 직관을 더 잘 받아들이게 만든다. 죽임은 적극적 행위이며, 죽게 내버려둠은 적극적 행위를 하지 않음으로 도덕적 책임 또한 전자에 부여되고 후자에는 부여되기 어렵다.²⁾ 도덕적 책임이 부여되는 ‘행위’의 경우 행위에 따른 결과라는 인과관계가 물리적으로 잘 보여지므로 달리할 가능성을 파악하고 그에 따른 도덕적 책임을 부여하기가 비교적 용이하다. 대조적으로 ‘부작위’³⁾는 행해지는 바가 보여지지 않으므로 인과관계를 파악하기가 상대적으로 어렵다.⁴⁾ 그런데 필자는 이 비대칭성의 대조가 적절하게 이루어졌는지에 대해 문제를 제기한다. 행위자가 자신의 행위/부작위의 결과 등에 대해 어떤 방식으로 인지하는가에 따라 달리할가능성 및 인과적 귀적이 달라진다고 제안하게 될 것이다. 즉, 도덕적 책임의 인식적 요소는 인과의 틀로 형식화가능한 제어조건에 영향을 준다.

달리할 가능성이 있는가 여부로 도덕적 책임을 판단하는 기준이 행위와 부작위에 다르게 적용된다는 점이 행위-부작위 도덕적 비대칭성 논의의 핵심이다. 달리할가능성 원칙(PAP: Principle of Alternative Possibility)은 미처 하지 않은 행위를 할 능력이 있어야만, 즉 실제와 달리 행위할 수 있었는데 하지 않았을 경우 부작위에 도덕적 책임을 지을 수 있으며, 그렇지 않은 경우 도덕적 책임은 부여되지 않는다.(Fischer and Ravizza 1993, 125)는 원칙이다. 행위-부작위 간 도덕적 책임의 비대칭성을 주장하기 위해 사토리오(Sartorio 2017) 아래와 같이 프랑크푸르트 식의 두 사례를 대조하여 행위와 부작위의 도덕적 책임이 비대칭적임을 제시한다. 달리할가능성 원칙(principle of alternative possibilities)이 부작위에는 적용되나 행위에는 적용되지 않는다는 것이다. 프랑크푸르트의 설명에 따르면, 사례 1은 달리할 가능성이 없어도 도덕적 책임을 부여하는 경우⁵⁾이며, 사례 2는 달리 할 가능성이 없다면 도

2) ‘도덕적 책임있음’, ‘비난가능성’, ‘허용불가능성’이 동의어인가에 대해서는 이론의 여지가 있으나 이에 대해서는 별도의 논문에서 다루겠다.

3) ‘omission’은 부작위 외에도 ‘부재’로 번역(김성수, 「인과와 도덕적 책임」, 철학사상 54, 2014, 201~226.)되기도 한다. 그런데 ‘부재’는 인과적 역할을 반영하지 않는 ‘absence’와 혼동할 여지가 있다. ‘아이가 물에 빠지는데 구하지 않았다’에서 ‘구하지 않음’은 단지 어떤 행위를 ‘하지 않음’의 묘사일 뿐만 아니라 더 나아가 ‘해야 할 어떤 행위’라는 사회적 기대 및 의무 등의 맥락이 포함되어 있다. 이 논문은 ‘omission’에 대해 법에서 사용하는 공식어인 ‘부작위’를 사용하여 안락사 등과 같은 실제 예들처럼 인과적 역할이 있는 부작위 사례를 포괄하고 법의 윤리적 기반과 용어의 일관성을 유지하고자 하였다.

4) 이러한 입장을 취하는 더위(P. Dowe, *Physical Causation*, New York, NY: Cambridge UP, 2000.)는 에너지 보존량을 가지는 대상의 귀적이거나 그 보존량의 교환과 같은 과정이 부작위에는 없으므로 부작위의 인과적 역할이 없다고 본다.

5) 따라서 전통적인 달리할가능성 기준이 작동하지 않는 반례로 간주된다.

덕적 책임을 부여하지 않는 경우이다.

<사례 1> 행위(프랭크+신경과학자) 사례 (Frankfurt 1969)

한 아이가 연못에서 수영을 하고 있다. 프랭크는 그가 죽길 원해서 물에 뛰어들어 아이가 익사할 때까지 아이의 머리를 물 속으로 밀어넣었다. 또한 아이가 죽기를 원했던 사악한 신경과학자는 프랭크가 자는 동안 뇌에 삽입한 칩을 통해 프랭크의 뇌 활동을 감시하고 있었고 프랭크의 선택을 조작 하는 데도 사용할 수 있다고 말했다. 프랭크가 아이를 죽이기로 결심했을 때 마음이 흔들렸다면 신경과학자는 프랭크의 뇌 활동을 조작하는 방식으로 개입하게 될 것이다.

<사례 2> 부작용(존+악어) 사례 (Fischer 1986)

한 아이가 연못에서 수영을 하고 있다. 존은 아이가 물에 잠기기 시작했다는 것을 알아차렸고 프랭크와 마찬가지로 그는 아이가 죽기를 원한다. 그래서 아이를 쉽게 구출할 수 있다고 생각하더라도 물에 뛰어들지 않기로 결정했다. 아이는 결국 죽는다. 존은 몰랐지만 연못에 배고픈 악어가 있어 그를 공격하여 아이를 구하지 못하게 했었을 것이다.

사례 2에서 존이 만약 물에 뛰어들어 쉽게 아이를 구한다 할지라도 악어 때문에 결과적으로 아이를 구하지 못했을 것이다. 사토리오는 이에 대해 결과적으로 달리 되었으리라는 가능성이 없으므로 존은 아이를 구하지 않은 ‘부작용’에 대해 도덕적 책임이 없다고 한다. 반면, 사례 1에서는 달리 할 가능성 여부가 도덕적 책임 여부에 영향을 미치지 않는다고 주장한다.

1. 인과구조의 비동등성 I

필자는 사례 1-사례2 대조항이 다음의 두 측면에서 동등한 인과구조를 가지고 있지 않다고 생각한다. 첫 번째 측면은 달리할가능성이 개입되는 단계의 차이이다. 사례 1-행위(프랭크+신경과학자) --에서 신경과학자의 개입은 프랭크의 ‘의사결정’ 단계에 관련된 반면, 부작용(프랭크+악어)에서 악어들의 개입은 ‘결과’ 단계에 관련된다. 부작용(프랭크+악어) 상황에서 한 상황에 대해 ‘달리할 가능성이 없다’는 것은 행위의 ‘결과’ 뿐만 아니라 행위자의 ‘행위’나 ‘의도’ 등이 모두 달리할 수 없음의 상황이어야 한다. 그런데 사례 1, 2 는 의지에서 결과에 이르기까지의 일련의 상태들 중 일부 상태, 그것도 서로 다른 단계에 대한 달리할 수 없음을 서술한다.

사례 2(부작위:존+악어)에서 존은 아이를 구할 의도를 가지지 않은 것에 대한 도덕적 책임이 있을 ‘수도’ 있고, 또 아이를 구하지 않았다고 의사결정하거나 행위하지 않은 점에 대한 도덕적 책임이 있을 ‘수도’ 있다고 할 것이다. 그러나 아이를 구하는 것, 즉 결과에 대해서는 달리 할 수 없었기 때문에 그에 대해서는 책임이 없다는 것이다.

이는 의사결정이나 행위에 대한 책임은 논외로 하거나, ‘결과에 대한 도덕적 책임’ 여부는 거꾸로 ‘행위에 대한 도덕적 책임’ 여부와 ‘의사결정에 대한 도덕적 책임’ 여부로 전이된다고 가정하는 것이다. 결과중심의 달리할가능성 해석은 도덕적 행위/부작위의 상황에서 인과적 경로의 각 단계인 의도-의사결정-행위-결과 단계들 간의 ‘너무 쉬운’ 전이(transfer)를 반성없이 전제하고 있다.⁶⁾

만약 도덕적 책임을 부여할 수 없는 ‘달리할 가능성 없는’ 상황이 되려면, 그 상황의 인과적 경로의 모든 단계에 반사실적 상황들이 개입하는 상황이어야 완전한 달리할가능성이 없는 사례라고 할 수 있다. 사례 1과 사례 3의 경우 신경과학자는 행위자의 행위/부작위와 인과적으로 연결된 모든 상황에 개입해서 달리할가능성을 완전히 차단한다고 보기 어렵다. 신경과학자는 필요한 경우에만 의사결정 단계와 관련하여 개입한다. 그러면, 왜 우리는 사례 1에 있어서 행위자에게 도덕적 책임을 부여하는 직관을 가지는가? 이에 대한 한 설명방식은 실제로 신경과학자가 개입하지 않은 사례이기 때문에 실제 일어난 상황을 고정하고 반사실적 관계를 검토함으로써 인과관계를 판단하기도 한다.(Halpern and Pearl 2001; 김성수 2014, 215~219쪽) 이 방식은 프랭크의 행위/부작위 혹은 신경과학자의 개입이라는 선접적(disjunctive) 원인이 배타적 선접의 구조이든 포괄적 선접의 구조이든 선접원인의 문제를 깔끔하게 해석한다는 점에서도 의미있는 해결방식으로 보인다. 그런데, 실제 일어나지 않는 상황을 고정하는 방법은 사후적 정당화 및 설명의 특징을 지니므로 실제 여러 달리할가능성을 미리 고려하거나 일어나지 않은 상황을 고려한 도덕적 책임을 판단해야 하는 상황에서는 불충분한 기준이다.

비대칭성 주장을 옹호하는 측면에서 인과적 경로의 각 단계인 의도-의사결정-행위-결과 단계들 간의 전이가 인과적 책임의 측면에서는 적절하다고 일단 가정해보자.⁷⁾ 그렇게 가정하더라도 그러한 인과적 책임이 도덕적 책임의 경우에도 그대로 전이되는가? 의지와 의사

6) 필자는 이런 편향, 즉 결과 중심의 도덕적 책임 논의는 인과적 책임과 도덕적 책임을 구분하지 못하게 만든다. 우리가 한 행위자에게 도덕적 책임을 물을 수 있는 이유는 행위자의 행위(action)나 행위하지 않음(omission)이 원인이고, 도덕적 책임에 인과관계가 필수적인 고려사항이기 때문이다. 그러나 인과관계가 있음이 곧 도덕적 책임을 함의하는 것은 아니다.

7) 이 논문에서 언급하는 전이의 문제는 인과경로의 각 단계 간의 전이, 그리고 ‘인과적 책임’에서 ‘도덕적 책임’으로의 전이가 필연적이지 않다는 점이다. 여기서 더 나아가 ‘인과적 책임’의 전이성은 심지어 물리적 전이와도 독립적이라는 주장은 김성수(김성수, 「인과와 도덕적 책임」 철학사상 54, 2014, 222)를 참조.

결정조차 제어되는 경우와 의지만은 제어되지 않는 경우에 대해 우리가 도덕적 책임을 부여하는 직관이 다른 이유는 도덕적 책임이 일련의 인과절차에 의해서만 결정되지 않기 때문이다.

2. 인과구조의 비동등성 II

비동등성의 두 번째 측면은 개입 양상이 인과 경로에 미치는 내용 구조의 차이이다. 사례 1에서 결정 단계에서의 신경과학자의 개입은 반사실적 가정으로 위 시나리오에서는 결과적으로 실제 개입이 발생하지 않았으며 개입이 필연적이지 않다. 반면, 사례 2에서 결과 단계에서 악어의 개입 시나리오의 경우 애초에 악어가 ‘이미’ 개입이 되어 있는 ‘필연적 인과’ 경로가 만들어져 있는 상태이다. 이 두 측면의 결합은 두 시나리오 간의 인과 구조를 크게 다르게 만든다.

시나리오를 동등한 인과적 경로로 재구성했을 경우 비대칭성은 사라질 가능성이 있다. 필자는 여기서 사례를 빌어오거나 재구성해서 동등한 인과경로로 재구성해보겠다. 추가로 행위와 부작용에서 인과적 역할을 표현하는 방식이나 강도를 동등하게 표현했을 경우 비대칭성이 완화 혹은 사라질 수 있다. 아래의 두 표에서의 대조는 인과경로의 단계를 동등하게 재구성하여 대조시킨 것이다. 필자는 사례1-사례2의 행위-부작용 대조는 처음부터 잘못된 대조이며 아래의 사례들의 대조를 통해 논의를 시작해야 한다.

사례 3을 보자. 사례 1과는 행위-부작용에서만 차이가 있을 뿐 인과경로나 구조에서 동일하다. 사례 3의 경우 신경과학자의 만약의 경우에서의 개입 때문에 결과를 달리할 가능성은 여전히 없을 것이다.⁸⁾ 이 두 사례에서 인과경로의 의도-결정-행위-결과의 각 단계별로 도덕적 책임을 견주어보는 것이 정당하다. 비대칭성을 주장하는 이들은 결과중심의 달리할 가능성을 가정하는 경우 사례 3의 경우에 있어서도 사례 2와 마찬가지로 프랭크는 도덕적 책임이 없다고 할 것이다. 사토리오(2005)는 수정된 사례 3이 사실상 사례 2와 유사하다고 해석한다. 그러나 필자는 사토리오처럼 2-3을 유사하게 해석하고, 사례 1과 2를 적절한 행위-부작용 대조항으로 비교하여 그 도덕적 비대칭성을 주장하는 입장은 인과적 효력을 결과중심으로 보는 가정을 전제할 때에만 가능하다고 본다. 또, 사토리오는 사례 3에서 아이가 물에 빠져 익사한 것은 프랭크가 어떤 결심을 ‘해서’가 아니라 어떤 것을 하겠다고 ‘결심하지 않아서’인

8) 만약 행위-부작용의 도덕적 책임에 대해 대칭성을 강하게 주장하는 철학자가 있다면 위 사례에서 프랭크가 결과를 달리할 가능성이 없다 할지라도 프랭크는 결과 뿐만 아니라 그의 결심이나 행위에 도덕적 책임이 있다고 주장할 수 있다. 그러나 이런 주장은 비대칭성을 주장하는 이들과처럼 결과에 대한 책임이 의사결정이나 행위에 대한 책임으로 전이된다는 생각을 전제한다. 본고에서는 비대칭성 주장에 대한 비판적 검토가 주된 작업이므로 이에 대한 논의는 참고로만 남겨둔다.

부분에 대하여 “부정적인 심적 상태는 책임을 질 수 있는 상태가 아니라”(Sartorio 2005) 고 단정한다. 필자는 사토리오의 설명은 하나의 상황에 대한 고정된 기술 (description) 을 가정할 때에만 가능한 생각이며 ‘아이를 구하겠다고 결심하지 않은’은 ‘아이를 구하지 않겠다’는 결심을 달리 기술한 것에 불과하다고 생각한다. 이는 구체적으로 다음 절에서 인과관계에 대한 지각이 인식이나 기술에 상대적임을 검토함으로써 유효하지 않을 수 있음이 드러날 것이다.

<사례 1> 행위(프랭크+신경과학자) 사례 (Frankfurt 1969)	<사례 3> 부작용(존+신경과학자) (Sartorio 2017)
<p>한 아이가 연못에서 수영을 하고 있다. 프랭크는 그가 죽길 원해서 물에 뛰어들어 아이가 익사할 때까지 아이의 머리를 물 속으로 밀어넣었다.</p> <p>또한 아이가 죽기를 원했던 사악한 신경과학자는 프랭크가 자는 동안 뇌에 삽입한 칩을 통해 프랭크의 뇌 활동을 감시하고 있었고 프랭크의 선택을 조작하는데도 사용할 수 있다고 말했다. 프랭크가 아이를 죽이기로 결심했을 때 그 결정을 반복하려고 마음이 흔들리는 경우 신경과학자는 프랭크의 뇌 활동을 조작하여 죽이려는 결정을 유지하는 방식으로 개입하게 될 것이다.</p>	<p>한 아이가 연못에서 수영을 하고 있다. 존은 아이가 물에 빠지기 시작한다는 것을 알아차린다. 그는 아이가 물에 빠져 죽기를 원하기 때문에, 뛰어들지 않기로 결심한다.</p> <p>물속에 악어들은 없었지만 신경과학자가 아이가 죽기를 원했던 사악한 신경과학자는 존이 자는 동안 뇌에 삽입한 칩을 통해 존의 뇌 활동을 감시하고 있었고 존의 선택을 조작할 수 있다. 존이 뛰어들지 않기로 했지만 그 결정을 반복하려고 마음이 흔들리는 경우 신경과학자는 존의 뇌 활동을 조작하여 죽이려는 결정을 유지하는 방식으로 개입하게 될 것이다.</p>

<사례 4> 행위(프랭크+악어) 사례 (새로 구성)	<사례 2> 부작용(존+악어) 사례 (Fischer 1986)
<p>한 아이가 연못에서 수영을 하고 있다. 프랭크는 그가 죽길 원해서 물에 뛰어들어 아이가 익사할 때까지 아이의 머리를 물 속으로 밀어넣었다. 아이는 결국 죽는다. 존은 몰랐지만 연못에 배고픈 악어가 있어 그가 물속으로 밀어넣지 않아도 아이는 결국 공격당해 죽었을 것이다.</p>	<p>한 아이가 연못에서 수영을 하고 있다. 존은 아이가 물에 잠기기 시작했다는 것을 알아차렸고 프랭크와 마찬가지로 그는 아이가 죽기를 원한다. 그래서 아이를 쉽게 구출할 수 있다고 생각하더라도 물에 뛰어들지 않기로 결정했다. 아이는 결국 죽는다. 존은 몰랐지만 연못에 배고픈 악어가 있어 그를 공격하여 아이를 구하지 못하게 했었을 것이다.</p>

III. 인지적 구성물로서의 행위 -부작위 개념

인과적 역할이 있는가의 여부를 기준으로 행위와 부작위에 도덕적 책임 여부가 다르다고 비대칭성을 주장하는 견해에서 더 명료하게 해야 할 부분이 있다. 행위에서는 인과적 역할이 분명한 반면 부작위에서는 인과적 역할이 미미하거나 부재하다는 견해가 성립하려면 다음과 같은 요소들이 먼저 전제되어야 할 것이다.

- (1) 행위와 부작위에 도덕적 책임을 부여하려면 도덕적으로 침해되는 규칙이 명료하게 표현되어야 한다.
- (2) 행위와 부작위에 관련된 도덕적 책임을 ‘비교하여’ 이야기하려면 관련한 도덕적 침해 유형과 강도가 행위와 부작위 간에 유사해야 비교가 정당하다.

(1)의 요소가 중요한 이유는 행위는 그 특성상 물리적 촉발이 분명하게 표현되는 경향이 있는 반면 부작위는 그렇지 않기 때문이다. 행위와 부작위의 존재론적 양상이 다르든 같은 원인으로서의 역할에 대해 판단할 때 표현이나 기술(description)에 의해 왜곡되는 것은 방지할 필요가 있다. 관련 행위와 부작위와 관련한 상황만 주어질 경우 이를 읽거나 바라보는 사람들이 동일한 ‘도덕적 책임’ 혹은 ‘도덕적 비난가능성’을 이야기할지라도 그러한 도덕적 평가와 관련한 윤리적 규칙의 내용은 다르게 해석될 수 있다. 행위 혹은 부작위가 관련된 결과의 원인이 되는지 여부를 검토하려면 해당 행위 및 부작위가 침해한 도덕 규칙이 명시될 필요가 있다. (2)의 요소 또한 마찬가지로 이유로 행위와 부작위에 대한 인간의 인식적 편향을 견어내고 강도를 동일하게 하여 비교할 필요가 있다.

만약 (1)과 (2)의 조건에 따라 행위와 부작위에 동등한 조건들이 갖추어지면 행위와 부작위가 인과적 결과를 초래했는가의 정도가 행위 쪽에 더 많이 평가되는 것이 아니라 동등해질 수 있다. 또, 규칙이 동등한 강도로 기술되면 부작위 효과, 즉 작위와 부작위에 대한 도덕적 평가의 비대칭성은 더 이상 존재하지 않을 수 있다. 어떤 행위가 도덕적으로 잘못되었다는 증거는 제시하기 쉬운 반면, 부작위가 도덕적으로 잘못되었다는 증거는 상대적으로 제시하기가 어렵다는 태생적 비대칭성을 고려하여 가중치를 더하거나 빼거나 해야 기술이나 묘사에 의해 왜곡되기 쉬운 인간의 지각을 보완할 수 있다. 만약 그러한 근본적 비대칭성이 있다면 인간이 행위와 부작위에 대해 가지고 있는 비대칭성의 직관은 행위-부작위 자체에 기인하는 것이 아니라 행위-부작위가 기술된 방식에 기인한 것일 가능성이 있다.

행위-부작위의 도덕적 책임에서의 비대칭성을 주장하는 이들은 주로 행위나 부작위의 상

황을 묘사하는 시나리오에 기반하여 그 달리할 가능성과 인과적 경로를 검토함으로써 도덕적 책임 여부를 따지는데, 이는 행위가 도덕규칙을 위반하는 것처럼 그리고 부작용의 경우 도덕규칙을 위반한 것이 아닌 것처럼 기술되어 있기 때문에 비대칭성을 주장하게 될 가능성이 있다. 위의 두 조건을 동등하게 한 실험에서 부작용을 도덕적으로 관대하게 평가하는 현상은 어떤 물리적 증거가 부작용을 선택 하게끔 한다는 것을 보여주었을 때 사라졌다. (DeScioli, P. et al., 2011)

더 나아가 위 조건 2에서처럼 규칙의 강도와 종류를 유사하게 기술하는 경우 참가자들은 규범이 동등한 힘을 갖는 것으로 간주할 수 있다. 이렇게 두 요소를 보완하여 만든 행위-부작용 시나리오를 제시하여 피험자들에게 규칙의 중요성, 인과관계 여부, 칭찬받을만한지 등을 여러 강도로 표시하여 비교하였을 때, 행위와 부작용은 의미있는 차이를 보이지 않았다는 사실이 최근(Willemsen, et al. 2016) 제시되었다. 이처럼 원인-결과 관계 그리고 행위-부작용의 상황에 대해 명백하게 기술할 경우 인간의 그에 대한 지각은 이전과 달라졌고 비난가능성과 같은 도덕적 판단의 강도도 변화하였다. 행위-부작용의 구분은 한 사건을 표현하는 인간 능력에서의 차이일 수 있다. 이는 실재론이나 반실재론이라는 견해의 차이로 해석될 수도 있지만 우리는 일반적으로 한 사태나 사건 그 자체 보다는 그것이 표현되는 방식을 통해 그 사태를 인식하며, 서술방식에 따라 다르게 인식한다.⁹⁾ 이렇게 인과성이나 행위-부작용의 비대칭성이 인식 상대적이라는 점은 더 나아가 행위-부작용의 구분 자체도 상대적인 것은 아닌지 문제제기하게 만든다. ‘행위’와 ‘부작용’라는 구분은 물리적 촉발이 있는지의 여부에 따른 범주라는 데에는 의심의 여지가 없는 것으로 간주하는 반면 위의 반례들은 ‘물리적 촉발이 있는지’ 혹은 ‘기존의 상태에 변화를 주는 것인지’의 기준 자체가 명확한지의 근본적 질문을 던지게 한다.

일반적으로 행위와 부작용 개념은 구분되지만 그 경계가 뚜렷한지에 대해서는 의심의 여지가 있다. 행위와 부작용의 구분에 대한 근본적인 문제제기는 제임스 레이첼스의 ‘근소한 차이 논변 (the bare difference arguments)’(Rachels 1992) 에서 볼 수 있다. 레이첼스는 행위로 간주되는 적극적 안락사와 부작용으로 간주되는 소극적 안락사의 구분 근거가 사실상 설득력이 없다고 주장한다. 첫 번째 사례가 행위로 간주되는 사례이고, 두 번째 사례가 부작용으로

9) 이를 더 현저하게 보여주는 것은 행위-부작용의 도덕적 책임에서의 비대칭성이 ‘사라진’ 경우를 넘어, 역비대칭성이 나타나기도 한다는 사실이다. 즉, 부작용보다 행위의 경우에 도덕적으로 관대하거나 도덕적 책임을 덜 부여하고 행위에 도덕적 책임을 덜 부여할 가능성이 있다. 최근 코넬리의 연구(T. Connolly/J. Reb, “Omission bias in vaccination decisions: Where’s the ‘omission’? Where’s the ‘bias’?”, *Organizational Behavior and Human Decision Processes*, 91, 2003, 186-202.)는 부작용이 아니라 행위에 오히려 도덕적 책임을 덜 부여하는 맥락을 제시한다.

간주된다.

<유산상속 사례 1>

스미스는 그의 여섯 살 난 조카가 사망한다면 거액의 유산을 상속받을 수 있다. 어느 날 저녁 조카가 목욕을 하는 동안 스미스는 욕실로 잠입하여 조카를 익사시키고 마치 사고로 죽은 것처럼 꾸민다.

<유산상속 사례 2>

존스 역시 그의 여섯 살 난 조카가 사망한다면 거액의 유산을 상속받을 수 있다. 스미스와 마찬가지로 목욕 중인 조카를 익사시킬 계획으로 욕실에 잠입한다. 그가 막 욕실에 들어서는 순간 조카가 미끄러져 머리를 찰고 넘어지면서 얼굴이 물속에 잠기는 것을 목격한다. 존스는 기뻐하며 만약 필요 하다면 조카의 머리를 다시 물속으로 밀어 넣을 준비를 한 채 옆에서 대기하고 있다. 그러나 존스가 아무런 행동도 취하지 않고 그냥 지켜보고 있는 가운데 조카는 단지 약간의 몸부림과 함께 전적으로 혼자서 익사한다.

위의 두 경우를 비교해 보면 스미스는 조카를 죽였고 존스는 조카를 죽게 방치했다는 것 이외에는 그들 사이에 어떤 차이점도 존재하지 않는다. 사례 1과 2 모두 동일한 동기와 목적에 의해서 행 해졌으므로 도덕적인 차이를 둘 수 없다. 물론 안락사의 경우 유산상속 사례와 차이가 있을 수 있다. 안락사 시술 동기가 시술자 자신의 이익에 있지 않다. 하지만 레이첼즈에 따르면 의사가 인간적인 이유에서 환자를 죽게 방치했더라도 인간적인 이유에서 환자에게 독극물을 주입시켰을 때와 동일한 도덕적 입장이라는 것이다. 안락사의 경우에 있어서도 죽임과 죽게 방치함의 구분은 특별한 의미가 없다고 주장한다.¹⁰⁾

행위-부작위 구분에 대한 보다 일상적이고 쉬운 예로 풍장에 대한 인식을 들 수 있다. 전통 티벳이나 아메리카 원주민은 부모가 사망하면 거리에 두고 짐승이나 새들의 먹이가 되어 자연스럽게 시신을 깨끗하게 하도록 하였다. 즉 “아메리카 원주민은 사망한 부모를 길에 내버려둔다”는 ‘풍장(風葬)’이라는 적극적 행위이지 부작위가 아니다. 게다가 아메리카 원주민은 이에 대해 도덕적 책임이 없다. 이러한 판단은 도덕적 책임이라는 평가 뿐만 아니라 행위와 부작위라는 개념 또한 사회적 기대나 의무와의 관련성, 서술방식 속에서의 ‘규정’이며 그

10) 필자도 레이첼즈의 의견에 동의하지만 행위와 부작위의 구분이 모호한 점이 있다는 측면으로 받아들일 뿐 두 양상이 실제로는 같다는 해석은 채택하지 않는다. 레이첼즈의 논변에 대한 비판적인 논의로는 구영모(구영모, 「죽임/죽게 내버려둠에 관한 레이첼즈의 견해에 대하여」, 철학사상 6, 1996, 165~179.)을 참조하라.

자체로 본래적인 것은 아니라는 점을 보여 준다. 이를 소극적 안락사와 같은 의료사례에서도 볼 수 있다.¹¹⁾

이 절에서는 행위-부작위의 도덕적 비대칭성이라는 주장에 대한 문제제기로, 그러한 비대칭성이 서술에 따라 변경될 수 있고 이에 따라 행위-부작위 구분이나 개념 또한 기술방식에 따라 달라짐을 보였다.

IV. 의도/예견 구분의 모호성

앞 절에서 서술된 것처럼, 행위-부작위의 차이가 서술되는 방식에서 기인한다면 부작위로 아이를 죽게 내버려둔 존의 ‘결정’단계에 대한 도덕적 책임 가능성은 남는다. 그렇다면 아이의 죽음에 대한 온전한 도덕적 책임과 아이를 내버려두겠다는 의도, 결심 등에 대한 도덕적 책임을 분리해서 고려하고 이를 전체 과정 안에서 평가할 수 있다. 이 인과적 경로 중 특히 동기는 인과적 책임/법적 책임 뿐만 아니라 도덕적 책임 부여에 대한 중요한 평가 근거가 된다.

토마스 아퀴나스에서 연원한 이중결과원리(the doctrine of the double effect)는 한 행위가 도덕적으로 문제있는지를 그 동기가 어떤지를 통해 평가하는 도구이다. 이중결과원리는 동기를 맥락에 따라 두 종류로 구분함으로써 동기에 따른 결과를 세분화한다. 하나는 행위자가 의도한 결과이고, 다른 하나는 예견되긴(foreseen)하지만 의도한 것은 아닌 결과이다.¹²⁾ 예컨대 죽임이라는 행위의 경우 행위자는 죽임을 당하는 사람이 사망하기를 ‘의도’하는 반면, 죽게 내버려둠이라는 부작위는 그러한 ‘의도’가 없고 단지 예견할 뿐이라는 점이다. 그러나 이러한 구분이 상황서술에 따라 변동될 수 있음은 부분적으로 앞절, 특히 풍장의 예에서 명백히 보았다.

11) 일반적 행위-부작위에 있어서 그 속성의 본래성을 가정하고 그에 의존하는 것처럼, 안락사 논의 역시 적극적, 소극적 안락사가 분명히 구분된다는 견해는 임의로 설정된 ‘본래 상태’에 호소한다. 인과와 인과적 책임을 근거로 도덕적 책임을 논하는 관점에서는 ‘물리적으로 영향을 주는’ 원인을 주된 기준으로 삼는데, 소극적 안락사는 투여되던 약을 중지하는 경우이다. 이는 기존 상태에 분명한 물리적 변화를 주는 것이다. 이는 물리적 원인의 견지에서는 부작위가 아닌 ‘행위’로 규정될 수 있다. 물론 환자의 본래 상태는 약이 투여되는 상태가 아니라 약이 투여되지 않고 죽어가던 상태였으므로 본래 상태에 변화를 주는 것이 아니라고 반박할 수도 있다. 그러나 ‘본래 상태’란 과연 무엇인가? 죽어가던 상태 이전에는 건강하게 살아있는 상태였고, 그 이전은 더 젊고 체력이 좋은 상태였을 것이며, 그 이전의 더 본래적 상태는 갓 태어난 연약한 아기의 상태이다. 어느 단계가 본래상태인가?

12) 이중 결과론은 Foot (P. Foot, “Morality, Action, and Outcome”, *Morality and objectivity*, ed. T. Honderich, New York, NY: Routledge, 2011, 23~38., P. Foot, “The problem of abortion and the doctrine of the double effect”, *Applied ethics: critical concepts in philosophy*, 2, 2002(1967), 187.)에서 언급되었으며 뜻에 따르면 카톨릭 교도들이 임신중절과 관련한 딜레마를 다룰 때 의사결정의 기준으로 제시된 것이다. 이중결과원리에 대해서는 국내에서 강철(강철, 트롤리문제와 도덕판단의 세 가지 근거들, 철학연구 90, 2013, 137~171.)에서 자세하게 다루어졌으며, 주로 생명윤리와 같은 실제 사례에 대한 적용과 분석이 이루어졌다.

이중결과원리는 주로 전쟁상황에서 의도하는 주된 결과 외에 다른 부수적 결과도 발생하는데 (이중결과) 의도와 단순한 예견을 구분함으로써 민간에 대한 부수적 피해가 있음에도 불구하고 적진을 타격하는 것을 정당화하는 것이 가능하다. 즉 두 종류의 결과를 나눔으로써 예견되는 것에 의한 결과에 대해서는 “간접적 의도에 따라 야기되는 것이 때로는 허용”(Foot 2002) 될 수밖에 없는 상황을 반영한다. 필자는 ‘의도’와 ‘단순한 예견’은 인지상태의 질적 정도의 차이이며 이중결과원리는 적절한 검증도구가 되지 못한다고 주장한다. 그 이유는 ‘의도’와 ‘단순한 예견’의 구분이 도덕적 판단에 차이를 항상 만들어내는 것은 아니기 때문이다. 이는 두 개념의 구분이 모호할 뿐만 아니라 그 개념 자체도 애매하기 때문이다. 구체적으로, 의도없는 단순한 예견이어도 도덕적 책임이 부여되는 맥락이 있으며, 그렇지 않은 맥락이 있다.

형법에서는 분명한 의도가 아니라 할지라도 단순히 예견된 (merely foreseeing) 바에 의해 일어난 결과에 대해서는 무죄가 아니라 일정 정도의 책임을 지우고 있다. 이는 인과적 책임 판단에서 달리할 가능성을 고려할 때 단순히 결과 중심의 원인 뿐만 아니라 행위자/부작위자의 인지적 요소를 고려하여 인과 경로에 포함하고 있다는 것이다. 달리할 가능성을 고려할 때 결과 뿐만 아니라 의도를 넘어서, 단순한 예견 등 인지적 요소가 포함된다는 의미이다. 앞서 논의한 오른쪽으로 운전대를 돌린 사례에서 만약 ‘단순한 예견’이 포함된 것으로 평가할 경우, 이는 단순히 결과 중심의 달리할 가능성 분석으로는 도달할 수 없는 분석이며 사회적 기대나 행위자의 인식적 요소를 고려하지 않으면 불가능한 판단이다. 도로 위를 걷던 사람이 있는 방향으로 운전대를 돌린 운전자의 사례로 돌아가보자. 이 상황에서 오른쪽에 인도가 있고, 그 위에 사람이 있을 것이며 따라서 인도 위의 사람에게 상해를 입힐 수 있다는 예견이 가능할 수도 있을 것이다. 여기서는 논의를 위해 도로 위의 사람을 칠 수 있는 가능성에 대한 인지가 가능한 상황을 고려해본다.¹³⁾ 이 경우 상해와 관련한 분명한 의도를 가지지도 않았지만, 그러한 상황에 완전히 무지하지는 않으므로 이러한 상황을 어떻게 해석해야 하는 지에 대한 논란이 가능하다. 법적으로는 단순한 예견이 가능하므로 명백한 의도를 가질 경우의 처벌보다는 약한 정도의 법적 책임이 부여될 것임을 분명하다.

반면, 도덕적 책임 부여의 경우 ‘명백한 의도’와 ‘단순한 예견(merely foreseeing)’을 어떻게 구분할 수 있는지가 판단의 주요 요소 중 하나가 될 것이다. 이 경우는 트롤리 딜레마의 여러 버전 중 선 로가 합쳐지는 버전의 딜레마와 유사한 구조를 가진다. 널리 알려진 트롤리 딜

13) 이 상황은 편의상 행위 사례를 중심으로 논의하지만 부작위 사례에도 동일하게 적용가능하다. 이는 현실세계에서는 인간보다는 무인자동차처럼 빠른 의사결정을 내릴 수 있는 인공지능의 판단에 적용가능한 상황이 될 것이다.

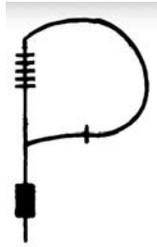
레마의 여러 시나리오 중 기차 선로를 변경해도 다시 본래의 선로와 합쳐지는 상황을 보자. 이 시나리오는 행위와 부작위의 경우 둘 다 그리고 여러 맥락의 달리할가능성을 포함하고 있어서 선로변경기를 돌린 사람에게 도덕적 책임을 부여할 수 있는지를 검토하기에 좋은 플랫폼 품이라고 할 수 있다. 도덕적 책임 여부는 인식적 요소를 고려하는 양상에 따라 맥락마다 달라질 것이다.

다음은 이중결과원리를 구체적 조건으로 상세화한 것이다.

- (1) 행위 자체가 도덕적으로 선한 행위이거나 적어도 중립인 행위이어야 한다.
- (2) 행위자의 의도가 나쁜 결과에 있지 않고 좋은 결과에 있어야 한다
- (3) 좋은 결과는 나쁜 결과가 수단이 되어 얻어져서는 안 된다.
- (4) 나쁜 결과를 허용하는 것이 불가피할 만큼 중대한 이유가 있어야 한다. (Mangan, 1949)

널리 알려진 트롤리 시나리오 중 하나를 이중결과원리의 네 요소에 비추어 평가해보자. 필자는 단순화를 위해 딜레마를 먼저 제거하겠다. 고장난 트롤리가 달려오는 선로 옆에서 나는 서있다. 사람이 없는 다른 선로로 갈 수 있도록 변환기를 돌리지 않고 내버려두었을 때 선로 위의 다섯명이 사망하게 되는데 이 사망에 대해 ‘변환기를 돌리지 않음’이라는 부작위는 도덕적으로 비난받을 만한가? 단순히 의도-예견 구분으로 판단하면, 변환기를 돌리지 않는 경우 선로 위의 다섯 사람을 고의로 죽게 할 의도는 없었을 것이다. 그러나 부작위의 경우 발생할 결과(다섯사람의 사망)는 충분히 예견가능하고 선로를 돌려 다섯명을 살릴 수 있었지만 그렇게 하지 않았으므로 도덕적 비난을 받을 수 있다. 또한 이중결과원리의 네 조건 모두를 충족시키지 못하므로 그 행위는 허용가능한 행위로 간주하기 어렵다.

그러나 모든 상황이 이렇게 쉽게 판단되지 않는다. 이중결과원리는 전쟁이나 의료와 같은 현실상황에서처럼 우선사항이 분명한 사례에서는 의도와 단순한 예견의 구분이 비교적 명확하여 적용하는 데 큰 무리가 없다. 그러나 추상적이고 구체화 정도가 낮은 사고실험의 여러 맥락에 적용하기에는 부족하다. 또다른 버전의 트롤리 사례를 보자. 아래에서 루프선 사례는 ‘의도’와 ‘단순한 예견’의 구분에 대해 명확한 개념을 보유하고 있지 않다는 것을 드러내어준다.



〈그림 1〉 루프선 사례 (Thompson 1986, p.1402)

루프선 사례는 다음과 같다. 고장난 트롤리가 달려오는데 그대로 직진하면 첫 번째 선로에 있는 다섯 사람이 죽게 된다. 다섯 명을 살리려고 선로를 바꾸면 다른 선로에는 몸집이 아주 큰 사람이 한명 있다. 이 사람은 아주 크고 무겁기 때문에 트롤리를 멈추게 한다. 만약 멈추지 않는다면 트롤리는 루프선을 돌아 첫번째 선로에 있던 다섯사람에게로 돌진하게 된다. 이 상황에서의 선로변경행위의 목표는 다섯사람을 살리기 위해서이다. 그리고 두 번째 선로가 옆편에 보이므로 최소한 그에 대해 ‘단순한 예견’이 가능한 상태이다. 선로변경을 통해 두번째 선로에 있던 몸집 큰 사람이 죽게 되는 경우, 몸집 큰 사람의 사망은 ‘의도’에 의해 ‘필연적’으로 발생한 것이 아니라 ‘단순한 예견’된 ‘부수적인’ 결과라고 하자. 여기까지는 이중결과원리의 구체적 조건을 충족시킨다고 할 수 있다. 그런데, 이중결과원리에 따르면 몸집 큰 사람의 죽음은 발생하는 결과의 수단이 되어서는 안된다. 그런데 몸집 큰 사람이 트롤리를 막게 된 결과로 다섯사람이 죽음을 피할 수 있게 되므로 몸집 큰 사람은 수단의 역할을 하게 된다. 즉 선로변경에 따른 큰 사람의 죽음은 단순한 부수적 결과가 아니라 의도된 수단이 된다. 이 상황에서는 의도-수단이 불분명해지며 이중결과원리로는 행위에 대한 도덕적 판단을 내릴 수 없다.

이에 대해 선로를 변경한 사람의 애초 의도는 다섯사람을 구하려는 의도였을 뿐이며 제 2의 선로에 묶여있던 사람이 몸집이 커서 ‘우연히’ 다섯사람을 살리는 수단의 역할을 하게 된 것이라고 반박 할 수 있다. 하지만 이 설명이 설득력이 있으려면 선로변경 결정을 내리는 사람은 어떤 이유로든 두 번째 선로에 묶인 사람이 트롤리를 멈출 정도로 크다는 점을 모르는 경우여야 할 것이다. 하지만 위 시나리오의 선로에 묶인 사람이 몸집이 매우 크고 이를 안다는 점이 가정되어 있으므로 충분히 그 사망을 인지/예견(의도없는 ‘단순예견’과 다른 의미에서)할 수 있다. 그러면 이에 따라 (선로변경에 따른) 몸집 큰 사람의 사망을 ‘의도’된 수단이라고 하자. 그러나 애초에 첫 번째 목표는 다섯사람을 살리는 것이며, 몸집 큰 사람을 죽이려는 것은 다섯사람을 살리려는 첫 번째 목표에 따른 ‘부수적 목표’라고 하면서 의도는 아니었

다고 반박할 수 있다. 이러한 상황에서 큰 사람의 사망에 대해 의도가 아닌 ‘단순 예견’이라는 설명은 미흡하다.

루프선에서의 선로변경과 관련한 이러한 논의는 과연 ‘의도’란 무엇이며 의도없는 ‘단순한 예견’이란 무엇인가를 다시 제기하도록 만든다. 첫 번째 선로 위의 다섯사람을 살리기 위해 두 번째 선로에 있는 몸집 큰 사람을 죽게 만들었다고 해도 이 때 선로변경시 몸집 큰 사람을 ‘죽이려는 의도’가 전혀 없다고 할 수 있을까? 온전하고 명백한 의도가 아니었다면 반(半; semi) 의도라고 해야 하는가?¹⁴⁾ 이중결과이론의 측정도구인 ‘의도’와 ‘단순히 예견함’의 구분으로는 위의 사례를 설명하기 어려울 뿐만 아니라 그 구분 간의 경계가 다소 모호함을 알 수 있다.¹⁵⁾ 또, 사촌의 재산을 상속하려는 의도로 욕실의 물에 잠겨 허우적대는 사촌을 ‘적극적으로 죽이는’ 행위나 물에 잠겨 결국은 죽도록 ‘내버려두는’ 부작위는 두 경우 모두 적극적인 의도를 가지고 행해지는 것이므로 의도-단순예견의 구분으로 행위-부작위를 특징짓거나 도덕적 허용 여부를 평가하기 어렵다.

이러한 문제는 ‘의도’를 어떻게 파악하는가와 관련이 된다. ‘의도’와 ‘단순예견’의 상관관계가 다소 모호하다는 점 외에도, ‘의도’ 개념 자체가 모든 이에게 명백하게 단일하게 지각된다는 가정이 전제 되어야 하는데 그렇지 않다는 점은 상황을 더 어렵게 만든다. 우리 인간은 과연 원인으로서의 의도를 객관적으로 지각할 능력을 가지고 있는가? 결과가 좋은지 나쁜지에 따라 동일한 의도를 달리 해석할 가능성(Knobe 2003)이 있다는 사실은 ‘의도’가 자의적으로 해석될 가능성이 있다는 점을 보여 준다.

또, 이중결과이론에서 ‘단순히 예견함(merely foreseeing)’은 의도가 전혀 없는 예견함을 가정하는데 이것은 과연 가능한가? 욕조에 빠져 죽어가는 사촌을 보험금 때문에 구하지 않는 친척에(인과적, 법적 책임은 물론) 도덕적 책임을 부여하는 사례에서 친척장본인은 죽이는 행위를 하지 않았기 때문에 적극적 ‘의도’가 아니라 ‘단순히 예견’했다고 반박할 수 있다. 반면 앤서니(1968)은 예견은 의도와 달리 원하는 바에 의한 것이 아니므로 예견이 의도를 포함하지 않는다고 하면서 ‘의도없는 단순한 예견’의 가능성을 주장한다. 그러나 필자는 보험금 때문에 사촌의 죽음을 ‘의도’하지는 않았지만 ‘예견’했을 뿐이라고 설명할 때 사촌이 물

14) 이러한 필자의 해석은 ‘단순한 예견(merely foreseeing)’에 대한 또다른 해석과 유사하다. 또다른 해석은 의도와 단순예견을 각각 ‘직접적 의도’와 ‘간접적 의도’로 지칭(Quinn 1989)한다. 뜻 또한 전자와 후자를 한 덩어리로 파악해야 하며 후자를 ‘의도되지 않은 것’으로 다루게 될 경우의 문제점을 지적하고 있다.

15) 의도와 예견 개념 구분의 어려움은 형사사건에 대한 엇갈리는 판결에서도 보인다. 계모의 무자비한 폭행으로 숨을 거둔 8살 아이의 계모에 대해서는 폭행과정에서 살인을 인식하지 못했다고 보아 상해치사 혐의를 적용했지만 경찰관을 차에 매단채 질주하다가 경찰관이 떨어져 숨진 미국의 사건에서는 의도적 살인죄를 적용했다.

에서 허우적거리다가 그대로 두면 사망한다는 인과 경로를 추측하고 이해하는 ‘예견’이 포함되며 이를 그대로 두었다는 부작용은 ‘원하는 의도’가 내재된 것으로 해석한다. 이 경우 예견이 의도를 포함하는 경우이다. 이렇게 ‘의도’ 만큼이나 ‘단순한 예견’ 개념 또한 애매한 해석에 열려있다.

그렇다면 ‘의도’와 ‘단순한 예견’으로 도덕적 허용가능성 여부를 평가하거나 행위-부작용을 구분 및 그 도덕적 책임의 다름을 판단하기보다는 ‘의도’와 ‘예견’의 구분이 의미있을 때와 의미없을 때를 구분하는 것이 더 생산적인 논의방식 일 수 있다.

V. 나가기

이 논문에서 행위-부작용 비대칭성의 인과 경로나 동기-예견을 논하는 것은 결과적으로 도덕적 책임을 논하기 위함이다. 그렇다면 부작용에 있어 도덕적 책임이란 과연 무엇으로 구성되는가? 비대칭성 주장은 도덕적 책임을 논하는 데 있어서 주로 인과관계, 인과적 책임을 논하는 데 집중했으나, 도덕적 책임의 논의는 존재론 뿐만 아니라 인식론, 가치론적 문제까지 포함된 복합적 맥락을 고려해야 하는 주제이다. 도덕적 책임은 단순히 원인 여부를 분석하는 차원을 넘어서 그것을 원인으로 보게 하는 인식적, 인지적 맥락을 도외시할 수 없다. 부작용은 그저 ‘내버려두는 것’의 의미 뿐만 아니라 ‘무언가 하도록 기대되는데’ 안하는 것이라는 사회적 맥락을 포함한다. 이는 풍자에 대한 인식에서 잘 드러난다. 따라서 어린이가 물에 빠졌는데 그냥 두는 상황이나 트롤리 딜레마의 부작용 상황도 서술되는 내용의 구체성 정도 등에 따라 다르게 해석될 수 있다. 이는 거꾸로 우리가 행위와 부작용의 인과경로나 달리 할가능성 여부를 해석하고 판단하는데 영향을 미친다. ‘무언가 하도록 기대되는’ 인식적 요소를 고려하지 않을 경우, ‘부작용(omission)’와 ‘아무것도 안함(inaction)’ 그리고 ‘부재(non-existence)’는 존재론적으로 거의 차이가 없다.

참고문헌

- 강철, 「트롤리문제와 도덕판단의 세 가지 근거들」, 『철학연구』 90, 2013, 137~171.
구영모, 「죽임/죽게 내버려둠에 관한 레이첼스의 견해에 대하여」, 『철학사상』 6, 1996, 165~179.
김성수, 「인과와 도덕적 책임」, 『철학사상』 54, 2014, 201~226.
류지한, 「안락사의 윤리적 쟁점」, 『철학논총』 31, 2003, 43~73.
Anscombe, G. E. M.(1957), *Intention*, Ithaca, N.Y. Cornell University Press.
Connolly, T./Reb, J., “Omission bias in vaccination decisions: Where’s the ‘omission’? Where’s the ‘bias’?”, *Organizational Behavior and Human Decision Processes*, 91, 2003, 186~202.

- DeScioli, P./Bruening, R./Kurzban, R., “The omission effect in moral cognition: Toward a functional explanation”, *Evolution and Human Behavior*, 32(3), 2011, 204~215.
- Dowe, P., *Physical Causation*, New York, NY: Cambridge UP, 2000.
- Feinberg/Joel, *Doing and Deserving*, Princeton: Princeton University Press, 1970.
- Fischer, J. M.(ed.), *Moral Responsibility*, Cornell University Press, 1986.
- Fischer, J. M./Ravizza, M., *Responsibility and Control: A Theory of Moral Responsibility*, Cambridge University Press, 1998.
- Fischer, J. M./Ravizza, M.(eds.), *Perspectives on Moral Responsibility*, Cornell University Press: Ithaca, 1993.
- Foot, P., “Morality, Action, and Outcome”, *Morality and Objectivity*, ed. T. Honderich, New York, NY: Routledge, 2011, 23~38.
- Foot, P., “The problem of abortion and the doctrine of the double effect”, *Applied Ethics: Critical Concepts in Philosophy*, 2, 2002, 187.
- Frankfurt, Harry., “Alternate Possibilities and Moral Responsibility”, *Journal of Philosophy* 66, 1969, 829~839.
- Graham, Peter., “The Epistemic Condition on Moral Blameworthiness: A Theoretical Epiphenomenon”, Robichaud and Wieland, *Responsibility: the Epistemic Condition*, Oxford: Oxford University Press, 2017, 163~179.
- Halpern, J./Pearl, J., “Causes and Explanations: A Structural- Model Approach Part 1: Causes”, *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*., 2001, 194~202.
- Kenny, Anthony(1968), “ Intention and Purpose in Law,” In Robert Summers(ed.), *Essays in Legal Philosophy*, Oxford: Blackwell.
- Knobe, J., “Intentional action and side effects in ordinary language”, *Analysis*, 63, 2003, 190~194.
- Knobe, J./Fraser, B., “Causal judgment and moral judgment: Two experiments.”, *Moral Psychology*, ed. W. Sinnott-Armstrong, Vol. 2 Cambridge, MA: MIT Press, 2008, 441~8.
- Levy, Neil., “Culpable Ignorance and Moral Responsibility: A Reply to FitzPatrick”, *Ethics*, 119(4);, 2009, 729~741.
- Malle, B. F./Guglielmo, S./Monroe, A. E., “A theory of blame.”, *Psychological Inquiry*, 25, 2014, 147~186.
- Mandel, D./Vartanian, O., “Taboo or tragic: Effect of tradeoff type on moral choice, conflict, and confidence”, *Mind and Society*, 7, 2007, 215~226.
- Mele, Alfred., “Moral Responsibility for Actions: Epistemic and Freedom Conditions”, *Philosophical Explorations*, 13(2);, 2010, 101~111.
- Moore, M., *Causation and responsibility: An Essay in Law, Morals, and Metaphysics*, Oxford: Oxford University Press, 2010
- Quinn, Warren., “Actions, intentions, and consequences: the doctrine of double effect,” *Philosophy and Public Affairs*, 18(4);, 1989, 334~351.
- Rachels, J., “Active and passive euthanasia”, *The New England journal of Medicine*, 292(9), 1975.

- Sartorio, C., “A New Asymmetry Between Actions and Omissions”, *Nous* 39:, 1975, 460~482.
- Sartorio, C., “The Puzzle (s) of Frankfurt-Style Omission Cases”, *The Ethics and Law of Omissions*, 2017, 133~147.
- Willemsen, P./Reuter, K., “Is there really an omission effect?”, *Philosophical Psychology*, 29(8), 2016, 1142~1159.

“부작위는 행위와 도덕적으로 동일한 지위에 있는가”에 대한 토론문

신상규

이화여대

김효은의 논문 “부작위는 행위와 도덕적으로 동일한 지위에 있는가”는 부작위와 행위의 대칭성 여부를 검토하는 논문으로, “대안가능성과 이중결과원리의 검증도구로서의 적합성”이란 부제를 달고 있다. 이 논문은 크게 3개의 절로 구성되어 있으며, 행위-부작위의 인과경로와 도덕적 책임, 인지적 구성물로서의 행위-부작위 개념, 의도/예견 구분의 모호성이란 제목을 가지고 있다. 토론자는 이 논문을 여러 번 읽었는데, 사실 아직도 내용을 잘 파악했는지 확신이 들지 않는다. 그나마 논의의 구도를 일부라도 약간 이해하게 된 것은 참고문헌으로 인용되고 있는 김성수의 「인과와 도덕적 책임」을 읽고 난 이후였다. 이하의 문제 제기는 토론자가 이해한 한도 내에서 제기하는 의문점임을 미리 밝혀 둔다.

이 논문에서 가장 문제라고 생각되는 것은 필자가 비판하고자 하는 입장이 정확히 무엇인지가 분명하지 않다는 것이다. 이는 유사한 결과를 산출하는 모든 형태의 행위와 부작위에 대해서 책임의 비대칭성을 주장하는 입장인가? 아니면 비록 유사한 결과를 산출하지만, 행위와 부작위에 대해서 책임의 비대칭성이 성립하는 경우가 일부 존재한다는 입장인가? 만약 전자라면 그런 주장을 하는 사람은 도대체 누구인가? 혹은 행위와 부작위의 구분이 사태에 내재하는 본래적 구분이라는 주장을 비판하고 싶은 것인가? 또 필자가 옹호하려는 입장이나 논지도 고정되어 있지 않고 각 절의 논의에 따라 변하는 것처럼 보인다. 필자는 유사한 결과를 산출하는 모든 형태의 행위와 부작위에 대해서 책임의 비대칭성을 주장하는가? 혹은 행위와 부작위의 구분은 사회적 문화적 관행에 의존하는 인식적/언어적 구분이어서, 책임의 문제와 관련된 한 유의미한 차이를 낳지 않는다는 주장인가? 이러한 불분명성은 필자가 비판하고자 하는 타겟의 주장이나 논거가 무엇인지를 명시적으로 잘 정리하고 있지 않기 때문으로 보인다.

일단 논문이 처음에 문제 삼는 바는 유사한 결과를 낳는 행위와 부작위에 대해서 동일한 도덕적 책임을 물을 수 있는가(책임의 대칭성이 성립하는가)의 문제이다. 필자가 비판의 대상으로 명시적으로 언급하고 있는 견해는, 행위에는 책임이 귀속되지만 부작위의 경우에는 책임이 귀속이 어렵다는, 책임 귀속의 비대칭성을 주장하는 사토리오의 견해이다. 사토리오는 행위와 부작위 각각과 관련된 프랑크푸르트 사례의 경우에 대해서, 전자에는 책임의 귀속되지만 부작위의 경우에는 책임을 귀속할 수 없다는 주장을 한다.

‘프랑크푸르트 사례’는 ‘달리할 가능성’이 책임 귀속의 필요조건이 아님을 보이기 위하여 프랑크푸르트가 고안한 사례들이다. 즉, 달리할 가능성이 없어도 책임의 귀속이 가능하다는 것이다. 논문의 (사례1)과 (사례2)는 각각 행위와 부작위와 관련하여 ‘달리할 가능성’이 없는 사례들이다. 프랑크푸르트는 두 경우 모두에 비록 달리할 가능성은 없지만 책임이 귀속된다고 주장한다.(김성수 논문 참조) 그런데 김효은은 프랑크푸르트 자신이 2에 대해서 책임의 귀속이 안되는 사례로 설명하고 있다고 하는데, 확인이 필요해 보인다.

그런데 사토리오는 1의 경우에는 책임의 귀속이 가능하지만, 2의 경우에는 불가능하다고 하여 두 경우 책임의 비대칭성이 성립한다고 주장한다. 그런데 이 논문에서 사토리오가 왜 그런 주장을 하는지의 논거에 대한 분석은 자세히 제공하지 않고 있으며, 단지 ‘결과가 달라질 가능성’ 여부가 주요한 이유임을 시사하고 있을 뿐이다. 그런데 여기서 한가지 의문스러운 점은 사례1의 경우에도 결과가 달라질 가능성은 없다는 것이다. 프랑크푸르트가 두 경우 다른 대안 가능성이 없음에도 책임의 귀속이 가능하다고 주장했다는 사실을 기억하자. 그런데 그럼에도 불구하고 왜 사토리오는 사례1과 사례2의 비대칭성을 주장하는 것일까? 그런데 필자는 이에 대한 별다른 분석 없이 두 사례 사이의 인과구조의 비동등성에 대한 논의로 바로 넘어간다.

김성수의 논문에 따르면, 사토리오는 행위와 부작위의 경우, 각각의 인과적 구조 혹은 행위와 부작위가 수행하는 인과적 역할에 차이가 있다고 주장한다. 사토리오에 따르면, 행위의 경우 행위자는 기존의 위협을 증진시키거나 새로운 위협을 창출하지만, 부작위의 경우 행위자는 기존의 위협이 그대로 진행되도록 내버려 두거나 그것의 경로를 단지 살짝 바꾸기만 하는 차이를 가지고 있다.

필자는 사례2의 경우에도 책임의 귀속이 가능함을 주장하고자 한다. 그런데 필자는 이런 비대칭성을 주장하는 사토리오 같은 철학자들의 주장에서 ‘대안가능성’이나 ‘이중결과원리’가 중요한 작용을 하고 있다고 판단하는 듯 하다. <이 논문은 행위 및 부작위에 도덕적 책임을 비대칭적으로 평가하는데 사용하는 ‘대안가능성 여부’와 ‘이중결과원리’가 검증도구

로서 적합한지를 문제제기한다.> p.1. 그런데 이 논문에서는 사토리오 등이 어떤 방식으로 ‘대안가능성’이나 ‘이중결과원리’를 활용하고 있는지를 구체적으로 논의하고 있지 않아서, 과연 이런 방식으로 문제를 설정하는 것이 온당한 전략인지에 대해 의구심이 든다. 만약 모종의 비대칭성을 주장하는 사람들의 견해에서 이들 기준이 큰 역할을 하지 않는다면, 논문의 전체 구도가 허수아비 공격일 가능성이 있기 때문이다.

김성수의 논문을 참조하면, 행위와 부작위에 차이에 대한 사토리오 견해의 핵심은 행위와 부작위의 인과적 역할 혹은 인과적 영향을 끼치는 방식이 서로 다르다는 데에 있는 것처럼 보인다. 그런데 이 논문에서는 p5에 이르러서야 “인과적 역할이 있는가의 여부를 기준으로 행위와 부작위에 도덕적 책임 여부가 다르다고 비대칭성을 주장하는 견해에서 더 명료하게 해야 할 부분이 있다.”라는 언급이 등장한다. 사실 이 논문을 처음 읽었을 때, 왜 갑자기 ‘인과적 역할이 있는가의 여부’에 대한 언급이 나오는지 잘 이해하지 못했다. II절의 논의로부터 행위와 부작위의 차이를 이들이 수행하는 인과적 역할의 차이로 설명한다는 논의를 읽어낼 수 없었기 때문이다.

II절에서 필자의 논의는 제목들이 암시하듯이 행위와 부작위 사이의 인과적 경로나 구조가 서로 동등하지 않음을 보이는 것에 집중하고 있다. (2절에서 논의하는 경로나 구조에 대한 이야기는 사토리오가 말하는 인과적 역할과는 다른 논의처럼 보인다.) 여기서 잘 이해할 수 없는 것은, 필자에게 필요한 것은 행위와 부작위의 인과적 경로나 구조의 동등성(대칭성)으로 보이는데, 왜 비동등성을 주장하는가 하는 부분이다. 가령 사토리오의 경우, 이들 사이에 성립하는 인과적 역할의 비대칭성을 토대로, 책임의 비대칭성을 주장한다. 그렇다면 필자가 따라야 할 전략은, 그러한 (표면적) 비대칭성을 부정하고 이들 사이에 성립하는 (또 다른 수준의) 인과적 대칭성을 토대로 책임의 대칭성을 주장하는 것이 자연스러운 전략으로 보이기 때문이다. 그런데 필자는 “인과구조의 비동등성 I”에서는 이들 사이의 비대칭성만을 논의하고 있다. 그 자체로 나름 의미있는 논의일 수는 있겠으나, 사토리오가 취하는 전략에 대한 효과적인 대응인지는 최소한 이 논문 안에서는 판단이 불가능하다.

필자는 “인과구조의 비동등성 II”에 이르러서야 사례1과2의 시나리오에서 인과적 경로를 재구성하면 행위와 부작위 사이의 비대칭성이 사라진다고 주장한다. 문제는 사례2와 재구성한 사례3의 경우인데, 필자도 말하고 있듯이 사토리오의 도덕적 책임 문제와 관련하여 이들 사이에 아무런 차이가 없다고 해석한다. 그렇다면, 원래 사토리오가 주장한 행위와 부작위 사이의 인과적 역할에 대한 비대칭성의 문제는 전혀 해소되지 않고 남아 있다. 즉, 사례2가 되었건 사례 3이 되었건, 이들은 모두 부작위의 경우로서 사례 1과는 그 인과적 역할이 서로

다를 것이라 주장할 것이기 때문이다.

III절의 핵심 주장은 “행위와 부작위(의 책임)에 대해 가지고 있는 비대칭성의 직관은 행위-부작위 자체에 기인하는 것이 아니라 행위-부작위가 기술된 방식에 기인한 것일 가능성이 있다.”(p.5)는 진술로 보인다. 이와 관련하여, 필자는 “어떤 행위가 도덕적으로 잘못되었다는 증거는 제시하기 쉬운 반면, 부작위가 도덕적으로 잘못되었다는 증거는 상대적으로 제시하기가 어렵다는 태생적 비대칭성을 고려하여 가중치를 더하거나 빼거나 해야 기술이나 묘사에 의해 왜곡되기 쉬운 인간의 지각을 보완할 수 있다.”고 주장한다. 그런데 만약 행위와 부작위의 사이의 차이가 존재론적/형이상학적 차이인데도 불구하고, 위와 같은 인식론적/언어적 표준화(normalization)를 함으로써 오히려 사태를 왜곡할 가능성은 없는가?

필자는 이러한 논의를 뒷받침하기 위하여 레이첼스의 유산상속 사례를 언급하고 있다. 여기서 드는 의문은, 필자는 이와 같은 논의를 통하여 겨냥하고 있는 타겟이 도대체 어떤 입장인가 하는 점이다. 사토리오의 대안적 가능성이 없는 프랑크푸르트 사례와 관련하여, 행위와 부작위 사이의 책임의 비대칭성을 주장하였다. 말하자면, 사토리오의 유사한 결과를 산출하는 모든 형태의 행위와 부작위 사이의 비대칭성을 주장하지는 않았다는 것이다. 즉 사토리오는 행위와 부작위 사이에 비대칭성이 성립하는 사례들이 일부 존재한다는 주장을 하고 있을 뿐이다. 유산상속 사례의 경우는 프랑크푸르트 사례에 해당하지 않으며, 다른 대안 가능성들이 열려 있다. 그렇다면 사토리오의 입장에서 두 경우 모두에 대해서 책임의 귀속이 가능하다고 주장할 가능성은 얼마든지 열려 있다. 그렇다면, 여기서 필자가 비판하고 있는 정확한 타겟 입장은 무엇인가?

뿐만 아니라, III절에서 필자의 논의는 유사한 결과를 산출하는 행위/부작위 사이의 대칭성을 확보하는 것이 아니라, 행위/부작위 구분 일반의 정당성을 문제 삼는 것으로 그 논점이 완전히 바뀌고 있다. 가령 풍장의 사례에 대한 필자의 다음 주장을 보자.

<행위-부작위 구분에 대한 보다 일상적이고 쉬운 예로 풍장에 대한 인식을 들 수 있다. 전통 티벳이나 아메리카 원주민은 부모가 사망하면 거리에 두고 짐승이나 새들의 먹이가 되어 자연스럽게 시신을 깨끗하게 하도록 하였다. 즉 “아메리카 원주민은 사망한 부모를 길에 내버려둔다”는 ‘풍장이라는 적극적 행위이지 부작위가 아니다. 게다가 아메리카 원주민은 이에 대해 도덕적 책임이 없다. 이러한 판단은 도덕적 책임이라는 평가 뿐만 아니라 행위와 부작위라는 개념 또한 사회적 기대나 의무와의 관련성, 서술방식 속에서의 ‘규정’이며 그 자체로 본래적인 것은 아니라는 점을 보여준다.> p.6~7.

필자는 이 짧은 언급을 통하여 굉장히 많은 내용을 읽어내기를 기대하고 있는 듯 하다. 상

당히 논란이 많을 수 있는 주제인 것 같은데, 필자는 왜 풍장이 행위와 부작위의 구분이 본래적인 것이 아니며, 사회적 기대나 의무와의 관련성, 서술방식 속에서의 ‘규정’임을 명백히 보여주고 있다고 생각하는가? 충분한 설명이나 정당화가 필요한 듯 보인다.

IV절의 이중결과원리에 대한 논의도 혼란스럽기는 마찬가지이다. 이중결과원리는 어떤 행위를 통해 의도한 (선한) 결과와, 비록 의도하지는 않았지만 행위에 수반하는(예견되는) 나쁜 결과를 구분하고, 몇 가지 조건을 만족시켰다는 전제하에, 행위에 수반하는 나쁜 결과를 정당화하는 논리이다. III절과 마찬가지로, 여기서도 이중결과원리를 이용하여 행위와 부작위의 비대칭성을 주장하는 이는 누구이며, 어떤 방식으로 그 논의를 전개하는지에 대한 분석이 전혀 제시되어 있지 않다. 평가를 위한 맥락이 제시되어 있지 않다 보니, 필자의 논의가 어떤 성과를 거두고 있는지도 덩달아 불분명하다.

이에 더불어 이중결과원리에 대한 필자의 논의 또한 이중결과원리가 문제가 되는 맥락과 분리되어 단순히 의도와 예견의 구분만을 문제삼고 있는 듯이 보인다. 의도와 예견 구분의 모호성을 통하여 필자가 입증하는 바는 무엇인가? 원래 필자가 문제삼고 있는 상황들, 예컨대 프랭크나 존의 사례, 그리고 스미스나 존스의 사례조차도 애당초 이중결과원리가 요구하는 조건들을 전혀 충족하지 못하기에, 이중결과원리가 적용 가능한 사례들이 아니다. 그러므로 단순히 의도와 예견의 구분만을 가져와서, 이들 사례에 적용하고 당사자들이 의도와 예견에 대해 각기 다른 주장을 할 수 있다는 것이 보여주는 바가 무엇인지 잘 모르겠다.

p3의 중간 부분에 사례3을 언급하지만, 실제로 사례3이 소개되는 것은 p.4이다. p.8에서도 “앞서 논의한 오른쪽으로 운전대를 돌린 사례”가 언급되지만, 그에 대한 논의는 어디에서도 찾아볼 수 없었다. 완성본에서 수정이 필요해 보인다.

2020 한국철학자연합대회

2020
UCCP

2020.10.24.(토)

2020 한국철학자연합대회

2020
UCCP

—
대동철학회

조선 성리학의 철학적 가치

유원기
계명대학교

1. 들어가는 말

매년 수백 건 이상 산출되는 국내의 조선 성리학¹⁾ 관련 연구는 정확히 어떤 의미 또는 가치를 갖는가? 그것은 진리의 탐구에 도움이 되는가, 삶을 살아가는 데 도움이 되는가, 아니면 단순한 역사적 잔재로서 과거에 대한 우리의 관심과 흥미를 충족시키는 데 도움이 되는가? 조선 성리학에 대한 논의가 오늘날 어떤 가치를 갖느냐는 이런 물음은 간헐적이거나 계속 제기되어왔으나 그에 대한 답변은 아직도 명확하지 않다. 학문의 가치를 어디에서 찾을 수 있느냐에 대한 다양한 의견이 있었지만, 우주와 인간 등에 대한 진리를 탐구한다는 진리성의 측면과 삶에 유용한 지혜를 제공한다는 유용성의 측면이라는 두 가지 기준을 중심으로 할 때, 우리는 다음의 네 가지 견해를 생각해볼 수 있다.²⁾ 즉, 조선 성리학에는 a) 진리성과 유용성이 모두 있다는 견해, b) 진리성은 없으나 유용성은 있다는 견해, c) 유용성은 없으나 진리성은 있다는 견해, 그리고 d) 진리성도 없고 유용성도 없다는 견해이다.

1) 주지하듯이, 중국 유학은 시기에 따라 선진유학, 한당유학, 송명유학 등으로 구분되며, 특히 기존의 유학보다 더 체계적이고 이론적인 성격을 지니는 새로운 유학이라는 점에서 신유학(新儒學)이라고도 불리는 송대의 유학은 송학(宋學, 송나라의 유학), 성리학(性理學, 性과 理를 핵심 개념으로 삼는 유학), 정주학(程朱學, 정호·정이 형제와 주희가 체계화한 유학), 주자학(朱子學, 주희가 체계화한 유학) 등으로도 불린다. 한편, 송명유학은 ‘리학(理學)이나 도학(道學)이라고도 불리는 주자학’과 ‘심학(心學)이라고 불리는 육왕학(陸王學, 육구연과 왕수인 계통의 유학)과 양명학(陽明學, 왕수인이 체계화한 유학)’을 모두 포함한다. 조선 성리학은 주로 주자학 계열의 유학을 가리킨다. 참조: 안유경(2015), pp.18-20; <https://terms.naver.com/entry.nhn?docId=388668&cid=41978&categoryId=41985>.

2) 진리는 반드시 삶에 유용한 것도 아니고 또한 삶에 유용한 것이 반드시 진리는 아니지만, 그렇다고 해서 진리성과 유용성이 완전히 배타적이지는 않으며, 그것들은 서로 일치할 수도 있고 일치하지 않을 수도 있다. 한편, 여기에서 학문의 진리성은 어떤 학문이 반드시 진리를 담고 있어야 한다기보다 단지 그것이 진리를 규명하기 위해 탐구한다는 의미로서, 탐구를 통해 진리가 아닌 결과에 도달하거나 또는 진리에 도달하지 못하는 학문이라 할지라도 그것은 진리성을 지닌다고 말할 수 있다. 그리고 유용성은 도구 생산 등에 필요한 실질적 지혜 뿐만 아니라 삶에 대한 반추나 반성 등과 같이 더 나은 삶을 사는 데 도움 되는 정신적 지혜도 포함한다고 말할 수 있다.

성리학은 주로 철학이라는 학문 영역에서 논의되지만, 철학 연구자들은 그들의 연구 분야가 동양철학인가 또는 서양철학인가에 따라 서로 다른 견해를 갖기도 한다. 최소한 대부분의 (또는 모든) 동양철학 연구자들은 d)를 거부하고, a)나 b) 또는 c) 가운데 어떤 한 가지를 근거로 성리학의 가치를 주장할 것이다. 반면에 최소한 일부 (또는 대부분의) 서양철학 연구자들은 a)를 부정하는 d)를 수용함으로써 또는 진리성을 부정하는 b)를 수용함으로써 성리학의 가치를 거부할 것이다. 위 견해들 가운데 특히 b)는 해결되지 않은 철학의 정의와 관련된 문제를 담고 있다. 그것은 바로 철학이란 진리만을 추구하는 학문인가, 아니면 유용성만을 갖는 학문도 철학이라 볼 수 있는가에 대한 문제이다. 서양철학 연구자는 진리성을 갖지 않는다는 점에서 성리학이 철학임을 부정하며, 동양철학 연구자는 유용성을 갖는다는 점에서 그것이 철학임을 주장한다. 이것은 일정한 철학의 기준을 인정하고 어떤 것이 그 기준에 맞는 지 또는 맞지 않는지에 대한 논의가 아니라 철학의 기준 자체에 대한 논의이므로 어떤 한쪽이 양보하지 않는 한 해결될 수 없는 문제이며, 이 논의에 끼어들어 어떤 결과를 얻을 수 있으리라 기대하기는 어렵다.

여기에서 우리는 철학이 진리성을 갖춘 학문이거나 또는 최소한 진리가 무엇인가를 밝히려고 추구하는 학문이라고 전제할 것이다. 이것은 단순히 논의의 편의를 위한 것이지 철학의 성격만을 고집하면서 다른 어떤 학문 영역을 배제하기 위한 것이 아니다. 따라서 만약 성리학이 진리에 관한 학문이고 이 기준에 따르면 성리학이 철학에 포함되지 않기 때문에 이 전제에 동의하지 않는 사람이 있다면, 철학이라는 명칭을 성리학에 내어주고 진리에 관한 학문을 다른 어떤 이름으로 불러도 무방하다. 또한 위와 동일한 이유에서 이제 성리학이 철학에 포함되지 않는다는 것을 인정하고 또한 철학이라는 명칭에 연연하지 않는다면, 성리학을 철학이 아니라 그저 성리학이라고 불러도 무방하다. ‘철학’이라는 명칭이 중요한 것은 아니며, 따라서 그 명칭에 연연할 필요는 없다.

중국 학계에서는 2000년도를 전후해서 이른바 “중국철학의 합법성(中國哲學的合法性)” 문제가 집중적으로 논의되었다.³⁾ 이것은 유교·도교·불교 등을 포괄하는 중국철학을 서양철학과 동일한 의미에서 ‘철학’이라 볼 수 있는가에 대한 문제 제기이다. 이 문제는 한국철학과 관련해서도 제기되고 고찰할만한 중요한 문제이지만 국내 학계에서는 비중 있게 다루어지지 않았다.⁴⁾ 이 글은 중국철학의 적절성에 대한 문제가 제기된 이유와 경과를 먼저 살펴보

3) 양일모(2011, p.118)와 박영미(2012, p.175)는 ‘합법성’을 ‘정당성’으로 옮기지만, 우리말로 ‘중국철학의 적절성 또는 적합성(the legitimacy of Chinese Philosophy)’으로 옮기는 것이 자연스러워 보인다.

4) Yoo(2013), p.183. 이것은 중국 외에도 한국이나 베트남이나 태국 또는 다양한 중동 국가들과 같은 아시아 국가들은 물론이고, 사실상 유럽의 몇몇 국가들과 아마도 미국 정도를 제외한 미주 대륙과 아프리카 대륙의 거

고, 그런 뒤에 철학이 진리에 관한 학문이라는 전제 위에서 조선 성리학⁵⁾이 과연 철학으로 분류될 수 있는가를 논의하며, 또한 이 논의를 토대로 성리학에 대한 현대적 논의가 과연 어떤 가치를 갖는가, 그리고 그러한 가치를 고양할 방법이 있다면 어떤 것이 있는가를 고찰한다.

2. ‘중국철학의 적절성’과 ‘철학의 정의’

중국 학계에서 ‘철학’이란 단어가 처음 사용된 것은 1902년이었으며, 중국철학을 서양철학의 반열에 올려놓기 위한 실질적인 노력은 이미 1930년대에 풍우란(馮友蘭) 등에 의해 시작되었다.⁶⁾ 그 후 우시아오밍(Wu Xiao-ming)의 논문(1998)이나 <신중국철학오십년(新中國哲學五十年)>이라는 제목의 학술대회에서 정가동(鄭家棟)과 진래(陳來) 등이 발표문(1999)에서도 “중국철학을 ‘철학’이라고 부르는 것이 과연 적절한가?”라는 중국철학의 적절성 문제가 여전히 논의되고 있었다.⁷⁾ 그러다가 2년 뒤인 2001년도에 중국 상해를 방문했던 데리다(Jacques Derrida, 1930-2004)가 “중국에는 철학이 없고 사상이 있을 뿐이다.”라고 발언하면서 그 문제에 대한 논란이 가속화되었다.⁸⁾

의 모든 국가들과 관련해서도 제기될 수 있는 문제이다(아래 특히 2.2장 참조). 한편, 중국철학이 적절한 의미에서의 철학임이 입증되면, 그와 유사한 한국철학도 입증된 것으로 주장할 여지도 있겠지만, 그것을 남의 손에만 맡겨 놓는 것은 상당히 소극적인 태도이다. 중국 학계는 그 문제를 국가적 자존심의 문제로 간주하는 반면에, 한국 학계는 개인의 생계유지 문제로 간주하는 경향이 더 강해 보인다. 그러나 어떤 경우든 한국철학의 생존을 위한 노력이 필요한 것은 분명한 것이다. 국내 학계는 한국철학이 철학인가에 대한 논의보다는 그것이 철학임을 먼저 전제하고 그것의 장점이 무엇인가를 밝히는 데 초점을 맞추었고, 또한 한국철학이 진리를 담고 있거나 진리 탐구에 유용하다는 점을 강조하기보다 삶의 지침이나 윤리적 규범으로서 인간의 삶에 보탬이 된다는 점을 강조하는 경우가 많았다. 이런 접근방식의 이유와 문제점에 대해서는 아래 참조.

- 5) 이 글은 ‘조선 성리학’에 대해 고찰하며, 다른 논문이나 저서의 ‘동양철학’이나 ‘한국철학’이란 명칭을 불가피하게 그대로 인용한 경우에도 그것은 불교, 도교 등을 배제한 ‘조선 성리학’을 가리킨다.
- 6) 박영미(2012, p.177)에 따르면, 1873년에 일본인 서주(西周)가 ‘필로소피(philosophy)’를 번역한 ‘철학(哲學)’이라는 번역어가 중국에서 1902년 『신민만보(新民晚報)』에 처음으로 사용되었고, 북경대학에서 1914년에 설립되었던 ‘중국철학문(中國哲學門)’은 1919년에 ‘중국철학계(中國哲學系)’로 명칭이 변경되었다. 1915-1918년에 북경대학의 중국철학문에 재학했던 풍우란에 의하면(평유란 지음, 2011, pp.335-337), 중국철학문의 교육과정은 중국철학사·제자학(諸子學)·송학(宋學, 송명철학사)이었고, 서양철학과 관련된 교육은 전혀 제공되지 않았던 것으로 보인다.
- 7) Wu(1998), pp.406-452; 정가동과 진래의 발표문 서지 사항은 박영미(2012), pp.178-179와 각주 3 참조.
- 8) 필자가 영국의 분석철학자인 A. Woodfield에게 보내는 메일에서 이 사실을 언급하자 상당히 흥미로워했다. 주지하듯이, 현대서양철학의 두 분파인 영미분석철학과 대륙철학은 서로 다른 논증 방법, 판단 기준 등에 의존한다는 몇 가지 이유에서 “피차간의 논문도 읽지 않고 피차간의 학술대회에도 참석하지 않”을 뿐만 아니라(Charlton, 1991, 특히 pp.2-4), 심지어 영미분석철학의 입장에서는 대륙철학을 정통 철학이 아니라고 보는데, 그런 대륙철학에 속한 데리다가 다른 나라의 철학을 폄하했다는 것이 상당히 역설적으로 들린다는 것이다.

2.1 철학의 적절성 문제

“어떤 것이 X인가?”에 답변하려면 먼저 ‘X’라는 어떤 것이 있고, 또한 그것을 ‘X’라 판단하는 ‘기준’이 있어야 한다. 중국철학이 철학이냐는 질문에서 ‘철학’은 ‘서양철학’을 가리키며, ‘기준’은 ‘서양철학의 기준’을 의미한다. 이것은 서양철학을 철학의 기준이 되는 ‘보편철학’ 또는 ‘세계철학’으로 간주하고, 이를 기준으로 하여 검증되는 다른 모든 것들을 ‘특수철학’ 또는 ‘지역철학’으로 간주한다고 이해될 수도 있고, 이런 맥락에서 본다면 중국철학을 마치 변방의 철학처럼 간주하여 그것의 적절성 여부를 검증하는 일 자체가 중국 학자들에게 전혀 유쾌한 일이 아니었음을 추측하기란 어렵지 않다. 중국 학계 외부에서 중국철학의 존재를 부정하는 철학자들이 이전에 없었던 것은 아니며, 칸트·헤겔·후설·하이데거 등을 비롯한 몇몇 서양철학자들이 이미 중국철학의 존재를 부정했던 바 있었지만,⁹⁾ 특히 새천년 벽두에 중국 내에서 중국철학의 존재 자체를 부정하는 듯한 데리다의 발언은 중국 학자들의 ‘민족적 자존심(national pride)’에 깊은 상처를 남긴 사건으로 평가된다.¹⁰⁾

중국철학의 적절성 문제에 대한 중국 학자들의 대응은 다음과 같이 몇 가지 유형으로 구분할 수 있다.¹¹⁾ 그것들은 (a) 1916년에 출간된 사무량(謝無量)의 『중국철학사』에서처럼 “서양철학과 중국철학이 분류와 내용에 실제로 차이가 없다”는 입장, (b) 1931년에 출간된 풍우란(馮友蘭)의 『중국철학사』에서처럼 “서양의 철학을 표준으로 하여 엄격하게 중국의 의리지학 가운데 그것에 대응하거나 부합하는 것을 취하”는 입장, (c) 진래(陳來)와 장대년(張岱年)처럼 “서양철학은 특수한 것일 뿐 표준은 아니”라고 보면서 철학의 외연을 확장하는 입장, (d) 정가동(鄭家棟)이나 모종삼(牟宗三)처럼 “중국전통철학의 특수성을 주장하는” 한편 “‘철학’이라는 진리의 보편성을 서양전통으로 귀속시켜 분석하려는 시도를 반대하는” 한편 철학의 외연을 확장하기보다 “‘철학’ 개념의 내적 함의를 풍부하게 하고 심화”하는 입장 등이다. 결국 (a)와 (b)는 서양철학의 우선성 또는 기준성을 인정하고 그에 부합하는 중국철학의 요소를 찾으려는 노력이며, 그와 달리 (c)와 (d)는 모두 서양철학과 중국철학이 서로 다르다는 것을 전제하지만 (c)는 철학의 의미를 재정립하여 서양철학과 중국철학을 포괄하도록

9) Defoort(2006), p. 626 참조. 이남영(1986, pp.161-163)은 헤겔에 대해, Wu(1998, 특히 pp.411-432)는 헤겔, 후설, 하이데거에 대해, Defoort(2006, p. 626)는 칸트에 대해, 그리고 Rosker(2017, pp.1-9)는 헤겔의 견해를 언급한다.

10) Defoort(2001, p.393; 2006, p. 626)는 중국철학의 존재가 거부되면 중국철학 연구자들의 존재 가치가 없어질 뿐만 아니라 그들의 민족적 자존심에도 충격을 받게 된다고 말한다. 대체로 학자들은 데리다가 서양철학의 우월함이나 중국철학의 열등함을 지적하려 했던 것은 아니지만, 그의 발언이 중국학자들에게 ‘경악’과 ‘고통’을 동시에 안겨주었다는 점에 동의한다(양일모, 2011, p.118; Rosker, 2017, p.4). 데리다의 발언에 대해서는 Defoort(2006), p. 626; 양일모(2011), p.116; 박영미(2012), p. 176 각주 1; Rosker(2017), p.3 등 참조.

11) 박영미(2012), pp.179-182; 아래 명명된 학자들의 서지 사항은 ‘참고문헌(pp.200-201)’ 참조.

외연을 확장하는 반면에 (d)는 서양철학과 무관한 중국철학의 개별적 또는 독자적 노선을 수행하려는 노력이다. 사실상 (a)는 중국철학이 서양철학과 동일한 위상을 지닌다고 주장하는 반면에, (b)의 입장을 취했던 풍우란은 “중국의 전통 철학은 항상 중국학의 일부로 간주되었으며, 철학과는 아무런 관련이 없는 것으로 생각되었다.”고 말함으로써, 서양에서 중국철학을 철학으로 인정하지 않는 것에 대한 실망감 또는 좌절감을 드러내기도 했다.¹²⁾ 중국철학과 서양철학이 사실상 서로 다르다는 결론에 도달했다는 점에서, 그는 (c)와 (d)의 학자들과 다르지 않다. 결국 (b), (c), 그리고 (d)는 모두 중국철학이 서양철학적인 의미의 철학이 아니라는 입장이다.

하지만 중국철학의 적절성 문제에 대한 중국 학계의 대응 방식과 데리다의 발언에 대한 중국 학자들의 감정적 반응, 즉 서양철학을 기준으로 중국철학의 적절성을 평가하려는 시도와 철학의 소유 여부를 자존심의 문제로 받아들인 것이 필자에게는 모두 부적절한 대응으로 보인다. 그 이유를 간단히나마 살펴보자.

2.2 철학의 정의 문제

‘철학(철학)’이란 한자어는 서양철학을 지칭하는 ‘필로소피(philosophy)’의 번역어이므로,¹³⁾ 서양철학을 하나의 기준으로 제시하고, 그와 유사해 보이는 다양한 학문들이 그것과 실제로 어떤 유사성을 갖는가를 확인해보는 것은 자연스러운 호기심의 발로라 할 수 있다. 다시 말해서, ‘철학’은 원래 서양의 학문 분야를 지칭하는 용어였고, 뒤늦게 발견된 동양의 어떤 학문과 그것의 유사성을 비교해보는 것은 자연스럽다는 것이다. 그런데 “이러저러한 것이 서양철학이다.”라는 식으로 서양철학의 조건이나 기준을 제시한다는 것은 그것의 정의를 제시하는 일이지만, 여기에는 두 가지 근본적인 문제점이 있다.

서양철학의 정의와 관련된 첫 번째 문제는 그것이 정확히 어느 시기 어떤 철학을 가리키는가에 대한 것이다. 아마도 우리는 서양철학사에 나타나는 내용을 서양철학으로 간주할 것이다. 대체로 서양철학사는 시대순으로 철학적 사조를 열거하는데, 고대 시기에는 그리스 철학을 제시하고, 중세로부터 현대 시기까지는 영국, 독일, 프랑스를 비롯한 몇몇 유럽 국가의 철학을 제시하며, 그리고 현대에는 유럽대륙에 속하지 않는 미국의 철학을 예외적으로 포함한다. 그러나 이런 구성에는 몇 가지 궁금증이 제기된다. (i) 그리스 철학은 그리스 식민지였던 (특히, 밀레토스 학파의 출신지인 밀레토스 섬을 비롯한) 터키 주변의 섬 출신 철학자들의 사

12) Defoort(2006), p. 625; 유원기(2018), p.24 각주 3 참조.

13) 위 각주 6 참조.

상과 그리스 본토 출신 철학자들의 사상을 가리키는데, 이것이 왜 서양철학 전체, 즉 서양에 속하는 국가들 전체의 출발점으로 간주되어야 하는지 분명하지 않다. 무엇보다 고대 영국 철학, 고대 독일 철학, 고대 프랑스 철학 등을 비롯한 많은 유럽 국가들의 고대 철학은 전혀 고려되지 않는 이유도 궁금하다. (ii) 더 나아가 서양철학사에서 영국, 독일, 프랑스를 제외한 대부분의 유럽 국가를 배제하는 이유가 불분명하고, 또한 (iii) 남미와 북미에 속한 대부분의 미주 국가들을 배제하면서 굳이 미국을 포함하는 이유도 분명하지 않다. 서양철학은 지도상으로 또는 지역별로 서양이라는 지역에 속한 철학을 가리키는 듯이 보이지만, 엄밀하게 말하자면 그것은 많은 서양 국가들 가운데 일부 시기의 일부 국가에 속한 일부 철학자들의 철학만을 가리키는 것에 불과하다. 그러나 동일한 시기에 속하거나 동일한 국가에 속하는 철학자들이라 하더라도 개인에 따라 철학의 주제와 내용 등에 차이가 있을 수 있기 때문에, 그 가운데 일부만을 선택하여 그것을 전체로 간주하는 것은 분명히 적절하지 않아 보인다. 다시 말하자면, 우리가 현재 서양철학이라 부르는 것은 결코 서양에 속한 모든 국가들에 공통된 철학을 가리키는 것이 아니며, 더구나 어떤 철학을 보편철학처럼 간주하고자 한다면 그렇게 간주하는 충분한 기준과 근거가 명확하지 않다. 그런 것이 제대로 제시되지 않는다면, 서양철학은 임의적이거나 독단적인 지칭으로 볼 수밖에 없다.

서양철학의 정의와 관련된 두 번째 문제는 바로 정의 자체에 대한 것, 즉 과연 정의가 가능한가에 대한 것이다. 흔히 모두가 동의할만한 정의는 없으며, 사실상 철학의 정의는 철학자의 수만큼이나 많다고 말해진다. 무엇보다도 철학 주제의 범위, 철학의 역할, 철학의 방법 등에 대한 이견이 많기 어렵기 때문이다. 솔로몬(Solomon)에 따르면, 철학은 ㉠ 진리를 목표로 하는 것, ㉡ 정의나 행복 등과 같은 중요한 개념들을 명료하게 규명하고 정의하는 것, ㉢ 삶에 의미를 부여하고 올바른 삶을 살도록 이끄는 도덕과 종교에 가까운 것, ㉣ 증명과 논증을 강조하는 것, ㉤ 비판 기술이나 개념 체계 구축 기술 등과 같은 것, ㉥ 경험을 연구하는 것, ㉦ 경험을 믿지 않는 원리를 연구하는 것, ㉧ 실용적인 것, ㉨ 순수하게 관념적인 것, ㉩ 직관과 통찰력을 신뢰하는 것 등으로 정의된다고 말한다.¹⁴⁾ 이처럼 아주 다양한 규정들을 모두 수용하여 정의할 수도 없고, 또한 그 가운데 어떤 하나를 선택하는 경우에는 다른 것들을 거부하는 이유를 모두 밝혀야 하는 어려움이 있다.

그럼에도 하나의 정의를 선택하고 그것을 기준으로 ‘철학’이라는 학문 분야에 속하는 것과 그렇지 않은 것을 구분해보는 작업은 의미가 있다. 그러나 이미 말했듯이, 서양철학 자체를 규정하고 한정하는 것이 거의 불가능하기 때문에, 따라서 우리가 정의할 것은 지역적이고

14) Solomon(1997), pp.12-13; (2001), pp.100-103 참조.

개별적이고 한정적인 ‘서양철학’이 아니라 세계적이고 보편적인 ‘철학’이라는 점이다.

** 미완 상태이며, 아래 1과 2에 대한 논의가 첨가될 예정임.

1. (특수철학 // 보편철학 설명, --- 이동준, 최영성, 이철승, 이찬, 김영건...)
2. (서양철학을 정의하는 것이 아니라 ‘철학’을 정의해야 하며, 그 정의는 ‘진리 탐구’이다.)

이처럼 명확하게 정의되지 않은 어떤 것을 기준으로 하여 다른 것의 적절성 여부를 논의하는 것은 부적절하다. 따라서 중국철학이 서양철학과 많은 유사점을 갖는다는 결론이 나오든 또는 차이점을 갖는다는 결론이 나오든, 그 결론에 대해 중국 학계나 학자들이 민감하게 반응할 필요는 없을 것이다.

간혹 서양철학은 객관적이고 보편적인 진리를 탐구하는 반면에, 중국철학은 주관적이고 특수한 지혜를 탐구한다고 말해진다. 여기에서 이런 규정의 옳고 그름을 논의할 필요는 없으며, 다만 학문의 탐구에서 우리에게 필요한 것이 무엇인가를 생각해보자.

이미 언급했듯이, 서양철학사에서 발견되는 서양철학의 전부가 진리만을 탐구했던 것은 아니지만, 여기에서 ‘진리의 탐구’를 철학의 핵심 기능으로 규정하는 것은 상당히 중요한 일이다. 더 나아가 이것을 ‘서양철학’의 특징이 아니라 ‘철학’ 일반의 특징으로 보는 것이 중요하다. 즉, “철학은 진리를 탐구하는 학문”이라고 규정하자는 것이다. 여기에서 ‘진리’는 “무엇이 옳은가?”라는 인식론적 진리가 아니라 “그것은 무엇인가?”라는 아주 일반적인 질문에 대한 올바른 답변이라는 의미에서의 진리이다.

3. 조선 성리학의 적절성 문제: 발전 과정과 연구방법

한국철학의 많은 부분이 중국철학에서 파생되었거나 영향을 받았음은 사실이지만, 발전 과정에서 중국철학과 다른 독창적인 성격을 갖게 되거나 또는 중국철학과 발생적으로 무관한 독자적 형식과 내용을 갖는 경우도 있기 때문에, 중국철학의 모든 문제들을 한국철학이 공유하지는 않는다.¹⁵⁾ 우리의 논의를 조선 성리학으로 한정할 때, 그것은 중국철학의 영향과 한국철학의 독창성이라는 특징을 모두 갖는다. 국내에서 한국철학이 전반적으로 어떤 발전 과정을 거쳐왔고, 또한 그 연구방법은 어떻게 변화되었는가를 살펴보자.

15) 유원기(2018), pp.37-40.

3.1 한국철학의 발전 과정

현재 한국의 철학 교육은 동양철학과 서양철학을 포함하지만, ‘한국철학’이라는 명칭은 국내에서 연구되는 서양철학 전체와 한국 지역의 철학을 제외한 동양철학의 대부분을 제외하고, 주로 유교·불교·도교 등의 전통사상만을 가리키는 듯하다. 그렇지만 ‘한국철학’이란 용어 자체가 사실상 무엇을 지칭하는지 정확히 알기 어려운 모호한(vague) 용어이다. 국내 또는 국외 학자가 국내 또는 국외에서 동양 또는 서양의 사상을 연구하는 경우, 즉 연구자의 국적·연구자의 거주 지역(특히, 국가)·연구자의 연구 대상을 고려하는 경우, 우리는 다음의 8가지 경우를 갖게 된다. 이 가운데 어떤 것(들)을 한국철학으로 승인할 것인지 살펴보자.¹⁶⁾

- ㉠ 한국인 학자가 한국에서 연구한 한국전통사상
- ㉡ 한국인 학자가 외국에서 연구한 한국전통사상
- ㉢ 외국인 학자가 한국에서 연구한 한국전통사상
- ㉣ 외국인 학자가 외국에서 연구한 한국전통사상
- ㉤ 한국인 학자가 한국에서 연구한 서양사상
- ㉥ 한국인 학자가 외국에서 연구한 서양사상
- ㉦ 외국인 학자가 한국에서 연구한 서양사상
- ㉧ 외국인 학자가 외국에서 연구한 서양사상

우리는 먼저 한국인의 사고방식이나 정서와 무관한 ㉠와 ㉢를 한국철학에서 배제할 수 있을 것이다. 그런데 ㉡와 ㉥에서 한국인의 사고방식을 갖는 ‘한국인 학자’들이 국내외에서 연구한 것에 대한 배제 여부가 그리 간단히 결정되지 않는다. 이것은 연구 대상이 아니라 연구자의 연구 성향에 대한 문제로서, 연구자의 사고방식이나 정서가 한국적이라면 그가 연구하는 서양사상에 한국적 사고가 어느 정도 반영될 텐데 그런 것을 한국철학이 아니라고 단정 짓기가 어렵다는 것이다. 그런데 ㉢와 ㉣는 어떤가? 외국인 학자가 국내외에서 연구한 한국사상은 한국철학에서 배제해야 하는가? 이것은 연구 성향이 아니라 연구 대상과 관련된 문제이다. 물론 외국인 학자가 연구한 한국사상은 외국인의 사고방식과 관점이 반영되었겠지만, 그가 한국사상을 연구하고 새로운 발전적인 관점을 제시하는 등의 공헌을 한다면 그것을 한

16) 여기에서 한국인 학자는 한국인의 사고방식을 갖고, 외국인 학자는 외국인의 사고방식을 갖는다고 가정한다. 왜냐하면 한국인의 사고방식을 갖거나(예를 들어, 외국에서 태어나) 한국인의 사고방식을 갖지 않는 한국인 학자, 또는 외국인의 사고방식을 갖거나(예를 들어, 한국에서 태어나) 외국인의 사고방식을 갖지 않는 외국인 학자가 있을 수 있기 때문이다.

국철학이 아니라고 단정 짓기가 어렵다. 반면에, 한국인 학자가 한국에서 연구한 한국사상을 일컫는 ㉑가 한국철학이라는 이름이 가장 적합하다는 데는 아무런 이견이 없겠지만, ㉒의 경우와 같이 한국인 학자가 연구한 한국사상이라 하더라도 그가 국내가 아니라 국외에서 연구하는 경우에 그것을 한국철학으로 승인할 것인지의 문제가 즉각적으로 분명하지는 않다. 다시 정리하자면, ㉑ 한국인 학자가 한국에서 연구한 한국전통사상을 ‘한국철학’이라 부르는 것이 정서적으로 가장 거부감이 없을 뿐만 아니라 실제로 그 이름이 적절해 보이는 한편, ㉒ 한국인 학자가 국외에서 연구한 한국사상은 물론이고 ㉓-㉔ 외국인 학자가 국내외에서 연구한 한국사상과 ㉕-㉖ 한국인 학자가 국내외에서 연구한 서양사상이 ‘한국철학’에 포함되는 가를 단정 짓기가 어렵다. 무엇보다 연구자의 국적·연구자의 거주 지역(특히, 국가)·연구자의 연구 대상 등과 관련하여, 전반적이고 복합적인 검토가 필요해 보인다. 위의 문제에 대한 별다른 고찰 없이, 국내에서는 전통적으로 ㉑를 ‘한국철학’으로 간주해온 것으로 보인다.

간략하게나마 한국철학의 발전 과정을 살펴보자. 국내에서 서양철학의 교육은 1926년 경성제국대학(1924년 설립)에 국내 최초로 철학과가 설립되면서 시작되었다.¹⁷⁾ 이철승은 일제 강점기에 철학 교육을 받은 이른바 현대 한국의 철학 1세대들은 (a) 서양에 대한 “근대 일본 지식인들의 사대주의적 태도”(p.43)와 구체적인 현실과 동떨어진 “관념적이고 추상적인 이론을 장려하는 일본 제국주의의 정책”(p.53)으로 인해 경성제국대학에서 서양철학, 특히 서양의 관념론적 철학을 배운 사람들이었거나, 또는 (b) “독일, 프랑스, 미국 등 서양과 일본에서” “서양식 근대 교육을 받은 사람들”(p.42)이었으며, 더 나아가 “당시의 많은 사람들에게 ‘서양철학’은 보편적인 철학이고, 전통철학은 특수한 ‘동양철학’으로 여겨졌다.”(p.42)고 말한다.¹⁸⁾ 이 시기에는 철학이라 불릴만한 수준의 전통사상에 대한 교육은 이루어지지 않았고, 다만 혜화전문학교(1906년 설립, 동국대학교 전신)와 명륜전문학원(1910년 설립, 성균관대학교 전신) 등에서 경학이나 한학 등을 가르쳤다.¹⁹⁾ 해방 직후인 1946년 현대적 대학으로 출발한 성균관대학교에 동양철학과가 개설되긴 했으나, 1960년대까지도 소속 교수들은 여전히 경학이나 한문학을 서당식으로 강의했고, 논문이나 저술 등의 학술 활동은 거의 하지 않았다고 평가된다.²⁰⁾

17) 최영성(2000), p.179. 비교: 이철승(2016, pp.53-54)도 이 내용에 동의하지만, 그는 천주교의 영향으로 국내에 ‘서양철학’의 내용이 처음 소개되었으며(p.49), 국내에서 “‘철학’이란 용어가 일반인들에게 사용되기 시작한 것은 19세기 후반”이었다고 말한다(p.50).

18) 이철승(2016), pp.42-53.

19) 최영성(2000), p.179.

20) 최영성(2000), pp.182-183. 그는 사실상 1980년대까지도 동양철학 개론서의 “내용 역시 서양철학 개론서에 비교할 수 없을 정도로 수준 이하 것이 대부분”이었다고 말하는데, 뒤에서는 1949년에 출간된 현상운의 『조선유학사』를 비롯하여 1980년대까지 출간된 몇몇 저서들의 특징과 우수성을 인정하는 것으로 보아 ‘개

그러나 1970년대까지도 서양철학 일변도의 교육이 지배적이었던 한편,²¹⁾ 이때부터 한국 유학 관련 박사학위 논문이 산출되기 시작하여 1970년대에 9건(퇴계-6, 율곡-3)에 불과했던 퇴계와 율곡에 관한 박사학위 논문이 1980년대에는 12건(퇴계-4, 율곡-8)이고 1990년대에는 43건(퇴계-25, 율곡-18)이 될 정도의 양적 발전이 이루어졌다는 것은 상당히 주목할만한 일이다.²²⁾ 무엇보다도 1970년대부터 유학 연구자들의 수가 증가하였고, “유학 연구의 경향은 대체로 조선 성리학의 쟁점과 전개 과정을 해명하는 데 관심이 집중되어, ‘성리학=유학’이라는 등식이 성립될 정도”였으며,²³⁾ 이러한 연구 경향이 오늘날 조선 성리학을 한국철학의 대표 주자로 인식되게 만드는 데 중요한 역할을 했다고 볼 수 있다. 그러나 이처럼 한국철학이 양적으로 크게 성장했다는 점은 분명하지만, 질적으로도 그만큼 성장이 이루어졌느냐는 질문에 대한 답변에는 선뜻 답하기가 어렵다.

3.2 한국철학 연구방법의 동향

사실상 한국사상 연구가 과연 ‘철학’이라는 이름에 걸맞은 방법으로 연구되었느냐는 질문에 대한 답변은 대체로 부정적이다. 이미 보았듯이, 20세기 초부터 현대에 이르기까지 그토록 오랜 기간을 서양철학이 교육되었다는 점을 고려하면 국내 서양철학 연구의 수준이나 방법이 상당히 발전되었으리라 기대하는 것은 당연하며, 또한 시대적 상황으로 인해 전통철학 연구가 억압되었던 것은 어쩔 수 없었더라도 최소한 해방 이후 오늘날까지 그러한 서양철학 토대 위에서 동양철학이 상당한 수준에 도달했으리라 기대하는 것도 당연할 것이다. 그러나 그 기대는 충족되지 않는다.

허남진과 최영성은 전통철학에 대한 연구방법이 크게 두 가지 시기와 단계를 거쳤다고 설명한다. 즉, 그들은 (1) 1960년대에서 1970년대 초반까지는 “서양철학의 방법론과 개념에 입각하여 동양의 전통철학을 연구”했던 시기로서, “동양철학에서 서양철학과 일치되는 면을 찾던 단계”이고, (2) 1970년대 중반 이후로 2000년까지는 “동양철학의 독자적인 연구방법론과 분석틀을 모색”했던 시기로서, 특히 1980년대의 동양철학계는 서양철학의 방법론을 통한 동양철학 연구의 한계를 자각하면서 동양철학 고유의 연구방법을 찾기 위해 노력했던 단계로 설명한다.²⁴⁾ 이 내용을 한국철학에 한정하면, 서양철학을 통해 한국철학을 이해하려

문서’의 수준만을 비판했던 것으로 보인다(p.189 이하 참조).

21) 박상리(2004), p.88.

22) 유원기(2018), pp.67-74. 필자가 파악한 바로는 해방 이후부터 2016년까지 발표된 퇴계와 율곡에 관한 박사학위 논문은 총 134건(퇴계-84, 율곡-50)이 발표되었다. 1970년대의 유학연구에 대해서는 최영성(2000), pp.194-197과 박상리(2004), pp.86-87 참조.

23) 최영성(2000), pp.193-194.

시도했으나 적절하지 않다고 판단되었고, 그 후 한국철학의 독자적인 연구방법을 모색하기 시작했다는 것이다.

하지만 그들의 설명처럼 한국철학에 대한 서양철학적 검토가 실제로 이루어졌는지, 그리고 실제로 이루어졌다면 어떤 방식으로 어느 정도나 이루어졌는지는 분명하지 않다. 1960년대부터 한국철학과 서양철학의 비교철학적인 논문들이 발표되긴 했으나 그 수는 많지 않으며, 더구나 동양철학계 내부의 많은 학자들이 한국철학(특히, 성리학)에 대한 현대철학 또는 서양철학적 검토의 필요성을 역설하고 있는 점을 고려하면,²⁵⁾ 서양철학적 검토가 충분히 이루어지지 않았다고 보아야 할 것이다. 사실 국내의 동양철학계와 서양철학계 사이에는 서로 겉으로 분명하게 드러내 보이지는 않으나 모두가 알고 있는 갈등이 숨어 있다. 그것은 단적으로 서로에 대한 경계심이라고 말할 수 있을 것이다.

“동양철학을 철학이 아닌 무당 푸닥거리로 폄하해서 본 분석철학 일각의 시각”²⁶⁾이라는 허남진의 표현과 “동양철학을 ‘철학’으로 인정하지 않는 서양철학계의 편향된 시각”²⁷⁾이라는 최영성의 표현이 바로 동양철학에 대한 서양철학계의 경계심을 보여준다면, 동양철학에는 논리로서 해석할 수 없는 요소들이 있음에도 불구하고 서양철학은 논리적 잣대만을 이용하여 모든 것을 분석하고 해체하려 한다는 것이 서양철학에 대한 동양철학계의 경계심을 보여준다.²⁸⁾ 이 경계심들은 사실상 모두 동양철학계에서 느끼는 위기감이라 할 수 있다. 동양철학이 철학이 아니라고 판정되거나 또는 동양철학의 논의가 비논리적이라고 판정된다면, 그것은 당연히 철학계에서 퇴출될 것이고 또한 동양철학 전공자들도 철학계에서 퇴출될 수밖에 없기 때문이다. 이처럼 동양철학에 대한 서양철학적 검토의 문제는 사실상 많은 동양철학 전공자들의 학문적 좌절감은 물론이고 생계의 문제와도 직결되는 문제이다. 하지만 동양

24) 최영성(2000, pp.199-200)은 따옴표 안의 글에서처럼 종종 자신의 표현을 사용하지만, 전반적인 내용은 허남진(1995, p.176)의 글에 대한 요약이다. 이에 덧붙여, 허남진(1995, p.177)은 (3) 1970년대에는 총론적인 연구가 이루어졌던 반면에, 1980년대에는 “연구분야가 세분화되고 확대되는 각론”적인 연구가 이루어졌다고 말한다.

25) ‘비교철학 논문’과 ‘한국철학에 대한 서양철학적 검토의 필요성을 주장하는 학자와 논문’은 유원기(2018), p.83 각주 130과 pp.85-86 참조.

26) 허남진(1995), p.176.

27) 최영성(2000), p.201.

28) ‘경계심’ 또는 ‘우려’라 할 수 있다. 실제로 필자의 한국철학 박사학위 논문심사과정에서 심사위원들은 서양철학을 먼저 전공했던 필자가 한국철학 주제의 논의에서 지나칠 정도로 논리 분석, 개념 분석 등에 치중하자 이런 우려 섞인 조언을 했던 바 있다. 이에 대해, 필자는 “한국철학의 모든 것을 논리적으로 분석하는 것은 가능하지도 않고 또한 그렇게 해서도 안 되겠지만, 그래도 할 수 있는 데까지 논리적으로 분석해보는 시도도 필요하지 않겠습니까?”라 반문하여 심사위원들의 승인을 받았던 바 있다. 비교: “서양철학으로 동양철학을 해석하고 동양철학에서 서양철학의 요소를 찾아내는 작업”의 결과 “서양철학으로 해석할 수 없는 면이 동양철학에 내재한다는 것도 알게 되었다.”는 허남진(2000, p.176)의 말이 이런 의미로 이해된다.

철학이 철학이 아니라는 판정이 즐겁거나 기대했던 결과가 아니라 할지라도, 철학이 아닌 것을 철학이라고 호도하면서 끝까지 진실을 외면하고 자신을 속이거나 또는 모르는 체 방관하는 것은 진정한(철)학자의 태도가 아닐 것이다.

3.3 한국철학의 연구방법 모색

한국철학을 비롯한 동양철학 일반에 대한 연구물, 연구방법 또는 연구 자세 등과 관련하여 다양하고도 많은 비판들이 제기될 수 있는데, 한국철학의 고찰할만한 가치가 있는 몇 가지를 살펴보자. 첫째는 타전공자들이 읽고 이해할만한 한국철학 논문이 없다는 비판이다.²⁹⁾ 논문은 전공자 또는 전문가를 위한 글이므로 이 사실 자체는 문제가 되지 않는다는 것이 필자의 견해이다. 비전공자가 수학이나 물리학 전공서를 읽으면서 어렵다고 불평하는 경우는 많지 않은데 반해, 철학서가 소설처럼 쉽게 읽히지 않는다고 불평하는 경우는 많다. 하지만 철학도 수학이나 물리학처럼 하나의 학문 분야로서 적절한 교육을 받지 않은 비전공자가 쉽게 모두 이해할 수 있는 학문은 아니며, 따라서 철학서나 논문이 어렵다는 비전공자의 평가는 어느 정도 당연한 일이라 할 수 있다. 그러나 정작 문제는 철학을 오랫동안 연구해온 많은 서양 철학 전공자가 한국철학서에 대해 어렵다고 평가한다는 것이다.³⁰⁾ 물론 철학 전공자라 할지라도 자신의 주요 전공 분야가 아닌 영역을 이해하는 데 어려움을 겪을 수도 있을 것이다. 하지만 최소한의 기본적인 독해도 가능하지 않고 매 문장과 매 단락에서 이해에 어려움을 겪는다면, 그것은 독자의 이해력 문제가 아니라 쉽게 이해하도록 집필하지 못한 저자의 문제라고 할 수 있다. 특히, 한문을 사용하는 전통철학의 경우에 종종 전공자에게는 익숙하지만 일반인에게는 그렇지 않거나 또는 일상에서 흔히 사용되지 않는 한자어들을 풀어 설명하지 않고 그대로 사용하는 경우가 있는데, 이것은 그리스철학 전공자가 그리스어의 뜻을 설명하지 않고 그대로 사용하는 것과 크게 다르지 않다. 이런 경우에 전공자가 아닌 독자가 독해에 어려움을 겪는 것은 당연할 것이며, 또한 한자어가 여러 의미로 해석되는 다의적인 용어나 개념의 의미들을 명확하게 구분하지 않고 사용하는 경우에도 한국철학서의 독해가 어려워질 수밖에 없다.³¹⁾ 물론 이 문제의 해결책은 비전공자라도 이해할 수 있는 쉽고 친절한 글을 쓰기 위해 노력을 기울이는 것이다.

둘째는 성리학적 논의를 지나치게 신봉하여 객관적 태도를 잃는 경우가 있다는 비판이다.

29) 허남진(1995), p.189.

30) 비교: 김영건(2005), p.113.

31) 이상호(2001), p.212; 김영건(2005), pp.122-125; 유원기(2018), pp.47-55와 p.87 등 참조.

실제로 동양철학계에서는 선현들 또는 성인들에 대한 존경심으로 말미암아 그들의 말을 무조건 수용하면서 비판 자체를 거부하거나 부담스러워하는 듯하다. 동양철학 관련 논문이나 서적에서, 연구자들은 대체로 공자, 맹자, 주자와 같은 중국의 학자들은 물론이고 퇴계와 율곡과 같은 국내 학자들의 경전들을 인용하기도 하고 전거로 사용하기도 하지만, 그들의 이론이 진정 타당한가를 평가하려는 시도조차 하지 않으며, 더 나아가 그들의 견해가 틀릴 수도 있다고 주장하는 경우를 찾아보기도 어렵다. 그렇다고 동양철학 전공자들이 비판이나 평가를 전혀 하지 않는 것은 아니다. 하지만 주로 그들의 비판과 평가는 성인의 이론이 타당한가에 대한 것이 아니라 성인의 이론에 대한 연구자의 해석이 적절한가에 대한 것이다. 이것은 적절한 학자의 태도가 아니다. 서양철학의 경우에 아무리 저명하고 탁월하다고 말해지는 철학자라 하더라도 비판 대상이 되지 않는 경우는 없다. 어떤 철학자가 훌륭한 인격을 갖추었다고 해서 그의 사상이나 견해가 반드시 옳은 것은 아니며, 또한 그것을 무조건 지지하고 옹호해야 하는 것도 아니다. 한편으로, 한국철학의 이론들 가운데 유효하지 않은 것임에도 불구하고, 단지 우리 것이라는 이유에서 보존하려고 하는 태도도 적절하지 않다.³²⁾ 선현들에 대한 존경심을 버릴 필요도 없고 옛것이라고 해서 반드시 버려야 하는 것도 아니지만, 유효한 것은 취하고 그렇지 않은 것은 버리는 취사선택의 태도가 필요할 것이다.

지금까지 살펴본 두 비판은 동양철학계 내부나 외부에서 모두 제기될만한 비판인 반면에, 다음 비판은 주로 동양철학계 내부에서 제기되는 비판이다. 세 번째 비판은 한국철학에 대한 논리적 분석이나 서양철학적 검토는 불필요할 뿐만 아니라 사실상 절대 해서는 안 된다는 비판이다.³³⁾ 이런 비판을 하는 사람들은 서양철학의 도구를 통한 한국철학의 검토가 한국철학 본연의 것을 왜곡하거나 망칠 수 있다는 주장을 한다. 하지만 한국철학에 대한 (그것이 정확히 무엇을 의미하든) 서양철학적 검토는 절대 안 된다는 이런 주장은 한국철학의 모든 것을 서양철학으로 분석해야 한다는 주장만큼이나 부적절한 주장이다. 좀 더 현실적이고 합리적인 자세는 한국철학의 내부에 서양철학의 도구로 분석할 수 없는 요소들도 있으니 그런 점을 감안하면서 분석해야 한다는 식의 경고를 하는 것이다. 하지만 동양철학계 내부에서 이런 경고를 하기 위해서는 서양철학의 도구로 분석할 수 없는 요소들이 무엇인가에 대한 논의가 선행되어야 한다. 서양철학으로 분석되지 않는 것이 무엇인가를 제시하지 못하면서, 말로 표현할 수 없는 그런 어떤 것이 있다는 식의 주장은 받아들여지기 어려울 것이다.

1980년대 박종홍은 사단칠정론에 대해, 1990년대 윤사순은 한국유학 전반에 대해, 2000

32) 비교: 김영진(2005), p.112, pp.122-127, pp.139-140 등.

33) 이상호(2001, p.210)는 동양철학에 대해 논리적 정합성을 지나치게 강조하는 것을 거부하는 반면에, 윤용남(2005, pp.285-286)과 유권종(2011, pp.29-30)은 동양철학에 대한 서양철학적 분석 자체를 거부한다.

년대 장숙필은 율곡의 이론에 대해, 서양철학을 통한 비판적 검토의 필요성을 역설한 바 있다.³⁴⁾ 서양철학적 검토라는 표현이 정확히 무엇을 어떻게 하겠다는 것인가에 대해서는 논란의 여지가 있겠지만,³⁵⁾ 서양철학적 검토가 필요하다는 주장의 의도는 한국철학의 보편성을 검토하겠다는 것이다.³⁶⁾ 여기에서 우리가 기억할 점은, 그것이 한국철학의 보편성을 반드시 확보해야 한다는 의지의 표명이 아니라 단지 한국철학이 진정한 철학적 문제를 해결하는 데 보탬이 될 수 있는 그러한 보편성이나 유용성을 갖느냐는 것을 검토하겠다는 열린 태도 또는 중립적인 태도에서 출발하겠다는 의미이다. 또한 “이른바 ‘현대철학적 평가’라는 표현은 사실상 하나의 논변을 상식적인 선에서 이해 가능한 논변으로 구성해야 한다는 것으로 이해해도 좋을 것이다.”³⁷⁾

4. 결론을 대신하여

한국철학에 대한 서양철학적 검토가 어려운 이유는 부분적으로는 두려움 때문이고, 부분적으로는 역량의 부족 때문으로 볼 수 있다. 먼저 두려움 때문이라는 것은 한국철학계 내부에서는 그런 검토를 통해 한국철학이 더 이상 철학의 영역에 속하지 않는다는 결과가 나올지도 모른다는 두려움을 갖는다는 것이다. 귀신 잡는 해병대의 역할은 당연히 귀신을 잡는 것이며, 따라서 귀신이 존재하지 않는다는 결론이 나오면 해병대의 존립 이유가 없게 될 것이다. 이와 마찬가지로 맥락에서, 한국철학이 더 이상 철학이 아니라는 결론이 나오면 한국철학으로 생계를 유지하는 많은 연구자들이 생계의 터전을 잃게 될 수도 있다는 두려움이 있을 수도 있다. 하지만 이런 두려움은 두 가지 이유에서 불필요하다. ‘한국철학’이라 불렀던 것이 이제 더 이상 철학이라는 특수한 학문 분야에 속하지 않게 되더라도 그것은 ‘한국사상’이라는 다소 포괄적인 학문 분야에 속하고 그와 관련된 연구가 지속되어 갑작스러운 생계의 문제가 발생하지는 않으리라 생각된다.³⁸⁾ 그러나 사실은 이런 생각 자체가 불필요하다. 왜냐하면 한국철학, 특히 조선 성리학은 상당히 이론적인 학문 분야이고 또한 성리학자들이 논의들

34) 유원기(2018), pp.85-86.

35) Yoo(2013, pp.191-194)는 조선성리학이 보편성을 확보하기 위한 최소한의 조건들로 일관성(consistency), 타당성(validity), 명료성(clarity)을 제안하고, 이것들이 서양철학의 일부로 인식되는 ‘논리학의 조건들’로 인식되지만 사실은 일반적인 학문에서 요구되는 ‘상식적인 조건들’이라 주장한다. 비교: 유원기(2018), pp.87-90.

36) 비교: 유원기(2018), p.85.

37) 유원기(2018), p.89.

38) 윤사순(1992, pp.133-134)은 사상을 “인간이 생활환경 속에서 직접·간접으로 부딪히는 수많은 난관으로서의 문제를 해결하기 위하여 짜내어진 생활관으로서의 지성의 산물”로 규정하는 한편, 한국사상을 “과거부터 지금까지의 모든 한국인들, 즉 한민족의 사유능력이 전개하여 온 생활관으로서의 사유세계”로 규정한다.

이 오늘날 기대되는 만큼 엄밀하고 체계적이지는 않을지라도 충분히 현대적 논의로 수용될 만한 논변들로 구성하는 것이 가능하기 때문이다. 즉, 한국철학에 대한 서양철학적 검토라는 것에 대한 두려움은 아무 근거 없는 두려움이므로, 그런 검토를 적극적으로 수용하고 검토할 수 있는 데까지 검토하고 불충분한 것은 충분하게 만들며 그렇게 만드는 것이 불가능한 부분이 발견된다면 그것을 기꺼이 포기하는 대범함이나 자신감을 갖는 것도 중요하다. 또한 이미 언급했듯이, 서양철학적 검토가 불필요하다고 여겨지는 부분들이 있다면 그런 부분들이 정확히 어떤 부분들이고 그런 검토가 왜 불필요한가를 분명히 밝히려는 노력도 있어야 할 것이다.

한편, 역량의 부족이라는 것은 동양철학계 내부에서 서양철학적 검토라는 것을 수행할만한 역량을 갖춘 연구자들이 많지 않다는 것이다. 우리가 잘 알듯이, 대체로 국내의 철학자들은 학부과정에서 서양철학을 중점적으로 교육하며, 동양철학 교과목들보다 서양철학 교과목들을 훨씬 더 제공한다. 대부분의 학부 학생들은 당연히 서양철학을 더 빨리 그리고 더 많이 접할 것이다. 그런데 그런 학생들이 대학원에 입학하여 동양철학을 전공하게 되면 기이한 일이 발생하게 된다. 그 학생들은 학부과정에서 배웠던 서양철학에 대한 지식을 활용하지도 않고, 또한 대학원에서는 서양철학을 더 이상 배우려 하지도 않고 관심을 갖지도 않는다는 것이다. 서양철학을 전공하는 대학원생들의 경우에는 정확히 반대의 현상이 발생하며, 그들은 더 이상 동양철학에 관심을 갖지 않는다. 이때부터 동양철학과 서양철학의 조화나 결합은 가능하지 않고, 또한 동서양 철학을 함께 배우으로써 서로의 장점을 강화하고 단점을 보완할 학문적 시너지 효과라는 것을 전혀 기대할 수 없게 된다. 동서양 철학을 모두 배울 수 있는 가장 좋은 환경을 가진 우리나라에서 그런 시너지 효과를 얻지 못한다는 것은 그만큼의 크나큰 손실이 아닐 수 없다. 서양에서 서양철학을 전공하고 귀국한 홍콩이나 대만 학자들이 아주 자연스럽게 중국철학에 대한 연구물을 발표하는 경우를 보게 된다. 하지만 국내 학자들의 경우는 아주 다르다. 국내외에서 서양철학을 전공한 학자들은 서양철학만을 연구하고, 동양철학을 전공한 학자들은 동양철학만을 연구하는 식으로 자신들의 구역을 굳건히 지키고 있을 뿐만 아니라 상대방들이 자신들의 영역을 침범하는 것을 용납하지 않는다. 동양철학에 관심을 갖는 서양철학 연구자나 서양철학에 관심을 갖는 동양철학 연구자가 흔히 듣게 되는 말은 “한 가지나 똑바로 하라!”는 말이다. 물론 자신의 전공을 열심히 하는 것은 중요한 일이다. 그러나 다른 어떤 학문 분야보다도 특히 철학 분야에서는 다양한 이론과 견해를 고찰함으로써 자신의 철학적 문제를 해결할 기회를 가질 수도 있으며, 따라서 자신의 전공만을 열심히 하는 것으로는 충분하지 않다. “신학문을 한 사람은 고전을 잘 모르고, 반대로 고전을 아는 사람은 새로운 학문에 어두웠다.”³⁹⁾

모두가 동의할만한 철학의 정의가 불가능하다고 말해지긴 하지만, 그래도 진리의 탐구라는 측면에서 본다면 타인의 것을 무조건 거부하고 자신의 것만을 고집하기보다 좀 더 겸손하고 포용력 있는 학문적 태도를 갖는 것이 무엇보다 중요해 보인다. 이런 태도를 통해 동양철학과 서양철학의 조화와 협력이 가능할 수도 있겠지만, 여기에서 궁극적으로 제안하고자 하는 것은 조화와 협력이 아니라 진리를 탐구하고 획득하려는 노력이다. 중립적인 태도를 갖자는 것은 박쥐처럼 회색분자가 되자는 주장이 아니라 옳은 것을 옳다고 말하고 그른 것을 그르다고 말할 수 있는 열린 태도를 갖자는 것이다. 구체적이고 바람직한 학문적 태도가 정확히 무엇인가에 대해서는 단정하기 어렵지만, 최소한 학문적으로 엄밀하고 체계적이고 명료하고 일관적인 태도를 갖는다면 우리는 최소한 진리 탐구를 위한 출발점에서 있다고 할 수 있을 것이다. 한국철학, 특히 조선 성리학에 어떤 철학적 가치가 있느냐는 물음에 대한 답변은 그것이 오늘날 우리가 관심을 갖는 철학적 문제의 탐구에 관여하는가에 달려있다. 여기에서 철학적 문제는 서양철학만의 문제나 동양철학만의 문제가 아니라 보편적인 철학의 문제를 일컫는다. 하지만 그것은 어떤 한 나라나 개인만이 관심을 갖는 문제도 아니고, 또한 세계 전체가 반드시 관심을 가져야 하는 문제도 아니다. 그것은 다만 세계 전체가 관심을 가질만한 문제를 말한다. 그리고 철학적 가치는 그런 문제의 해결이나 진리의 획득에 반드시 공헌하는 데서 찾아지는 것이 아니라 다만 진리의 탐구를 위해 얼마나 노력을 하는가에서 찾아진다. 다시 말해서, 진리를 획득하도록 돕는 학문도 철학적 가치를 갖지만, 진리를 획득하지 못하더라도 진리의 탐구를 위해 지속적으로 노력하는 학문도 철학적 가치를 갖는다는 것이다. 따라서 조선 성리학이 전자의 의미에서 철학적 가치를 지니느냐 또는 후자의 입장에서 철학적 가치를 지니느냐의 문제는 크게 중요하지 않다고 할 수 있다.

참고문헌

- Borchert, D.M. (ed.) (2006) *Encyclopedia of Philosophy*. 2nd ed. Vol. 7, Thomson Gale.
- Charlton, W. (1991) *The Analytic Ambition*. Blackwell.
- Defoort, C. (2001), 'Is There Such a Thing as Chinese Philosophy? Arguments of an Implicit Debate,' *Philosophy East and West* Vol. 51:3, pp. 393-413.
- Defoort, C. (2006), 'Is "Chinese Philosophy" a Proper Name?,' *Philosophy East and West* Vol. 56:4, pp. 625-660.
- Fung Yu-Lan (1952) *A History of Chinese Philosophy*. Vol. I, D. Bodde (tr.), Princeton: Princeton University Press.

39) 이동준(1997), p.266.

- Rosker, J.S. (2017) ‘Chinese Philosophy - Fact or Fiction?,’ *Sobre Ontens*, pp. 1-9.
- Solomon, R. C. (1997) *Introducing Philosophy*. 6th ed. Harcourt Brace & Company.
- Solomon, R. C. (2001) “‘What Is Philosophy?’” *The Status of World Philosophy in the Profession,* *Philosophy East and West* Vol. 51:1, pp. 100-104.
- Wu, Xia-ming (1998), “Philosophy, Philosophia, Zhe-xue,” *Philosophy East & West* vol. 48:3, pp. 406-452.
- Yoo, Weon-Ki (2013) “A Research Methodology for Korean Neo-Confucianism,” *Acta Koreana* vol. 16:1, pp. 177-197.
- 김영건(2009), 『동양철학에 관한 분석적 비판』, 서울: 라디오.
- 박상리(2004), 「해방이후 30여년의 한국 유학연구(I)」, 『동양철학연구』 제37집, pp.79-109.
- 박영미(2012), 「‘중국철학’은 정당한가?」, 『시대와철학』 제23권 제3호, pp.175-202.
- 양일모(2011), 「중국철학의 명운」, 『철학과현실』 제90집, pp.115-125.
- 유원기(2018), 『조선성리학 논쟁의 분석적 탐구: 사단칠정론과 인심도심론』, 서울: 역락.
- 윤사순(1992), 『동양사상과 한국사상』, 서울: 을유문화사.
- 윤사순(1993), 「한국유학에 대한 철학적 이해의 문제-그 회고와 전망」, 『철학』 제39집, pp.6-35.
- 최영성(2000), 「동양철학 연구 오십년사」, 『한국사상과 문화』 제10집, pp.175-295.
- 이상호(2001), 「조선조성리학 연구방법론 시고」, 『동양철학연구』 제25집, pp.189-215.
- 황의동(2002), 『유교와 현대의 대화』, 서울: 예문서원.
- 이동준(1997), 『유교의 인도주의와 한국사상』, 서울: 한올아카데미.
- 심재룡 외(1986), 『한국에서 철학하는 자세들』, 서울: 집문당.
- 이남영(1979), 「동양철학—그 연구방법 두 가지」, 심재룡 외(1986), pp.157-173.
- 김재권(1985), 「韓國哲學이란 가능한가?」, 심재룡 외(1986), pp.83-96.
- 이규호(1979), 「韓國哲學의 定立을 위한 摸索」, 심재룡 외(1986), pp.37-58.
- 이철승(2016), 「우리철학의 현황과 과제(1)」, 『인문학연구』 제52집, pp.39-63.
- 허남진(1995), 「서양철학 수용기의 동양철학연구 현황과 과제」, 『철학사상』 제5집, pp.175-190.
- 유권중(2011), 「21세기 유교 담론의 새로운 구도와 전망」, 『유교사상연구』 제44집, pp.5-36.
- 윤용남(2005), 「주자 심설의 체용이론적 분석」, 『동양철학연구』 제41집, pp.283-320.

圖像學에서 본 이황의 『聖學十圖』와 그 해석의 생명력 - 매체철학의 논법에 착안하여 -

김연재
공주국립대학

I. 들어가는 말: 매체철학의 문제

불확실성의 시대에 인간은 사회와 격리된 파편화된 자아가 아닌 주체적 자아로서의 모습을 지녀야 한다. 인간은 자기조절의 능력을 갖고서 타인과의 연속적인 만남을 통해 공감적으로 소통하고 조정하면서 공동체적 의식을 갖게 된다. 사회가 모순과 갈등을 통해 상대적인 협화가 극단으로 치달을 때에 소통과 통합의 사회적 본능이 구심적 역할을 하기 마련이다. 사회의 발전적 역량이라는 것도 결국에 공동체적 의식이 얼마나 발휘될 수 있는가 하는 문제로 귀착된다.

소통과 통합의 사회적 본능을 발휘하는 데에 매체는 중요한 수단 중의 하나이다. 언어, 그림 등의 매체는 주체와 객체를 연결하는 사회적 통용의 방편이다. 매체는 인식의 객체로서 주체인 인간과 일정한 관계를 맺는다. 이 관계에서 도상은 언어나 문자보다 재현이나 표현의 측면에서 더 포용성을 지닌다. 그것은 관념성이나 추상성의 한계를 넘어서며 인간의 사고방식을 보다 강화시키며 그 배후에 있는 사회의 통합적 인식을 정착하거나 확산시키는 역할을 한다.¹⁾ 여기에 매체철학의 생명력이 있다.

이러한 문제의식에서 본고에서는 동아시아의 사상과 문화와 그 원류에서 소통, 조정, 통합 등의 공동체적 방식에 관한 방법론적 발상에 착안한다. 논의의 핵심은 도덕의 이론적 원리나 실천적 강령을 어떻게 그림이나 형상으로써 전달하고 소통할 것인가 하는 것이다. 그 논점은 성리학적 세계관의 圖像化 혹은 圖式化에 관한 일종의 매체철학의 문제이다. 여기에는 도상이나 도식의 매체와 그 함축된 정보에 관한 이해와 해석의 문제가 있다. 그 대상은 매체철학의 중요한 논제들 중의 하나로서 圖像學(Iconology)²⁾이며 그 속에 담긴 계도나 교화의 기제

1) 임흥빈, 「비판적 매체철학의 관점에서 본 공적 합리성」, 『철학연구』, 20(2005), 168-170쪽.

이다.

본고에서는 이항의 『성학십도』에서 도상의 매체와 그 해설을 도상학의 현대적 시각에서 조망하고 그 속에서 성리학적 세계관과 그 시대정신을 밝히고자 한다. 그것은 聖人 혹은 帝王의 길을 위한 학문과 정치의 원리를 설파하면서 이른바 內聖外王의 도덕수양의 경지를 집약적으로 계도한다. 여기에서 성리학적 통치적 질서와 그 계도의 시대적 요청이 무엇인가를 성찰하고, 또한 그 고유한 정체성(identity)을 어떻게 확보하고 현대의 유교사회에서 그 지속가능한 현재성(presentness)을 어떻게 담보할 것인가 하는 문제의식을 지닌 해결의 실마리를 도출해내고자 한다. 여기에는 도상의 세계에서 역대로 混種과 創新의 시대적 변용을 통해 어떻게 해석의 생명력을 끊임없이 발휘할 수 있는가 점을 밝히고자 한다.

II. 동아시아의 圖書學과 도상학의 세계

인간은 삶에서 사고하고 행위하며 신념이나 믿음과 같은 의식을 갖고서 살아간다. 인간의 시선은 자연계에서 삼라만상의 스펙트럼을 다각도로 인지하고 이해하고 해석하여 이루어지는 의식의 산물이다. 시선이 미치는 경계가 바로 세계이고 세계는 주체와 객체의 일정한 관계에서 형성된다. 인간은 자신의 시선으로써 삼라만상의 각양각색의 스펙트럼을 보는 것이다. 이러한 시선의 영역이 인간의 삶과 관련된 세계이다. 여기에는 인식론적, 존재론적, 가치론적 내용이 담겨있다. 이른바 ‘天下’의 공간과 같은 동아시아의 세계관도 우리가 시선의 프리즘에 의해 집중적으로 초점을 맞춘 신념의 투사면이다. 우리는 신념의 방식을 통해 인간의 총체적 삶과 그 방식, 즉 사회나 문화의 현상 전반을 조명하는 것이다. 이러한 세계의 실재(reality)와 관련하여 국내외적으로 주도권이나 패러다임의 문제가 대두되고 민족주의, 보편주의, 제국주의 등과 같은 사조나 시대정신이 형성되는 것이다.

그렇다면 세계는 인간이 자신의 시선으로써 세상을 경계지워진 단계가 세계이다. 자연계의 변화 속에서 살아가는 인간의 삶과 그 주위의 모습을 투사한 프리즘의 세상이 바로 세계의 실재(reality)이다. 세계의 실재는 생명의 연결망 속에 삼라만상의 형형색색의 스펙트럼을 특정의 통일성의 질서에 초점을 맞추어진 결과이다. 세계에 대한 시선을 통한 관점은 세계와 인간의 관계 속에서 거시적이고도 미시적으로 동시에 조망하는 인식의 출발점이다. 그렇다면 세계관은 인간이 시선을 통해 삶을 투사한 일종의 프리즘이라고 말할 수 있다.

2) 이는 서양예술의 방법론 중의 하나이다. 그것은 예술론에서 작품의 구조적 문제를 발생론이나 위상론과 연계하여 논의하는 것이다. 여기에는 파노프스키(E. Panofsky)의 상징해석의 문제나 제들마이어(H. Sedlmayr)의 이론 등이 있다.

이러한 프리즘의 세계는 도상학의 분야에서 접근될 수 있다. 도상은 형상화의 이미지를 통해 주체와 객체의 내재적 관계를 연결하는 매체이다. 매체는 당시에 그것은 우주관, 인생관, 세계관 등을 서로 소통시키고 그 지평을 넓히는 데에 활용되는 형식이다. 여기에서는 세계에 대한 인간의 공통적 경험이 매체의 형식을 통해서 서로 주장되고 소통되고 이해되는 것이다. 우리가 도상에 표현되거나 재현된 관념이나 의미를 파악하는 데에 도상 혹은 도식의 특성이 일종의 담지체의 역할을 한다.

동아시아의 사상과 문화에서 圖像學은 유구한 역사적 전통을 지닌다. 당시에 그것은 우주관, 인생관, 세계관 등을 서로 소통시키고 그 지평을 넓히는 데에 일종의 매체의 형식이 유행하였다. 이와 관련한 현대적 학문이 바로 이른바 도상학이다. 도상학은 ‘그림과 그에 대한 해설(圖說)’을 통한 담론의 체계를 지닌다. 그것은 상징적으로 象, 數 등의 기호나 부호의 체계를 지니며 구조적으로 동심형, Tree형, 方圓形 등의 도식의 형태로 나타난다. 여기에는 세계에 대한 인간의 공통적 경험이 매체의 형식을 통해서 서로 소통되고 이해되는 과정이 있다. 우리가 도상에 표현되거나 재현된 관념이나 의미를 파악하는 데에 象이나 數의 물리적 특성이 이들을 간직한 일종의 담지체로서 역할을 한다.

동아시아의 사상과 문화를 볼 때, 도상학의 영역에 속한 독특한 분야가 송대에 이른바 圖書學이다. 도서학 중에서 특히 『주역』의 도식에 관련된 易圖學이 유행하였다. 그것은 象과 數의 매체를 통해 각종의 도상들 이를테면 太極圖, 先天圖, 河洛圖 등을 구상하고 그 속에 자신의 시각을 투영하면서 자신의 철학적 세계를 제시하였다. 이들은 우주의 본원, 생성 및 변화의 과정을 설명하고, 궁극적으로는 인간의 존재와 삶의 가치를 고양시키는 해석적 작업의 일환이었다. 이는 이러한 도상의 특수한 형상을 통해서 天人合一의 보편적 이념을 그려낸 것이다. 이는 매체철학의 넓은 의미에서 도상학의 범주에 속한다. 도상학은 대상의 본질과 내용을 효과적으로 그리고 계도적으로 전달하기 위한 방법으로서, 공통적으로 해석할 수 있는 매체의 방식을 지닌다. 그것은 태극도, 선천도, 하락도 등의 특수한 표상을 매개로 하여 철학의 사변성을 재현(representation)하거나 표현(expression)하는 방식이다. 이들의 도상은 당시에 세계관이나 인생관을 이해하고 접근할 수 있는 방법론적 통로를 제공하였다.

역도학은 『주역』과 관련한 역학계통의 도상 혹은 도식에 관한 학문체계를 가리킨다. 전통적으로 그것은 陰陽이나 五行의 개념과 그 범주에 기반한 태극도, 선천도, 하락도 등을 통해 우주의 본원, 생성 및 변화의 과정을 설명하고, 궁극적으로는 인간의 존재와 삶의 가치를 고양시키는 해석적 작업의 일환이었다. 태극도는 태극의 본원과 그로부터 만사만물의 생성문제를, 선천도는 卦象과 방위의 관계를 통해 인간의 선험적 준칙을, 그리고 하락도는 河圖와

洛書의 그림을 통해 인간을 포함한 세계의 도식을 상징적으로 그려낸 것이다. 이 도식들은 현대적 표현기법의 넓은 의미에서 일종의 ‘알레고리’의 범주에서 이해될 수 있다. ‘알레고리’는 대상의 본질과 내용을 효과적으로 그리고 계도적으로 전달하기 위한 방법으로서, 그 속에 반영된 공통적으로 해석할 수 있는 방식을 가리킨다. 이러한 방식이 바로 取象歸類 혹은 立象盡意로서 역도학의 기제에 속한다.

매체철학의 입장³⁾에서 보면, 역도학은 세계의 실재에 대한 인간의 경험을 바탕으로 수성되고 인간의 상상력을 통해 특정의 의도나 취지를 계도하거나 교화한다. 도상의 매체에서는 형식과 내용의 내재적 관계를 어떻게 이해하고 해석하는가 하는 문제가 중요하다. 여기에는 상징이나 기호의 체계, 알레고리적 의미, 주제의 내용, 취지의 강령 등이 중요하다. 도상의 구성에는 실재(reality)의 선형성과 경험성 및 그 양자의 관계를 어떻게 합리적으로 설정할 것인가 하는 효과적인 기법이 중요하며, 또한 도상의 이해에는 그 속에 담긴 역사적 자료를 이해하고 그 요체를 직관적으로 파악하여 전체내용을 심리적으로 수용할 수 있는 것이다.

도상학은 천인합일의 보편적 이념 하에서 시대적 혹은 사회적 욕구를 표출하고 충족하려는 시도의 단면을 보여준다. 도상학에서 진정한 세계관이나 참다운 인간관을 계도하고 계몽하는 효과를 기대하고 이를 충분히 이론화할 수 있었던 것도 이러한 욕구를 충족시킬 수 있었던 도상 안에서 구체적이고도 세부적인 재현의 영역을 분석하고 도상 전체의 표현의 의미를 이해함으로써 그 속에서 그 전체적인 유형구조가 투영되어 있다. 구체적으로 말해, 도상들 각각에서 象과 數의 연계성이나 중층적인 의미의 구조가 이해되고 그 구조 속에 전체적인 의미의 연관을 도출한 것이다. 이는 도상의 세부적 분석은 물론이고 태극, 음양, 오행, 팔괘 및 이들의 상관성에 따른 방위, 절기 등으로 구성되는 도상의 구성요소들의 구조, 그 함의 및 역사적인 발생의 관계를 바탕으로 하여 역학의 역사에서 共時的이고도 通時的인 종합적 해석이 담겨있는 것이다. 태극도, 선천도 및 하락도와 같은 매체의 역할과 이를 해석하는 도상학적 내용을 파악했기 때문일 것이다.

도상학은 인간이 인식한 것을 묘사하는, 단순히 수동적 의미의 재현적 도구만이 아니라 인간의 인식의 규준을 제공하는 능동적 의미의 표현적 담지체이다. 특히 象과 數의 물리적 특성은 의미의 담지체로서 인간의 주체가 세계의 객체를 인식하는 단서가 된다. 그러므로 도상학이 인간 주체의 단순한 발명품이라든가, 아니면 인간의 의지에 의해 사용되는 중립적 도구가 아니다. 도상은 인간과 세계를 매개하는 양식으로서 인간과 세계의 관계망에서 쌍방향성의

3) 새넬과 위버(Shannon & Weaver)은 의사소통의 문제를 기술적 문제, 의미론적 문제 및 효과론적 문제로 나누고, 현대의 정보이론이나 맥루한의 매체이론에서는 기술적 문제를 가장 우선시한다. 김상호, 「맥루한 매체이론에서 인간의 위치」, 『언론과학연구』, 8.2(2008), 94-95쪽.

연결통로의 역할을 한 것이다. 그것은 당시의 세계관에 관한 정보와 지식을 제공한 것이다. 여기에는 ‘정보의 도상화’와 ‘도상의 정보화’라는 양측면이 존재한다. 전자가 매체의 재현적 방식과 관련된다면 후자는 매체의 표현적 방식과 관련된다. 그 양자의 結晶體로서의 역도학은 천인합일의 성리학적 사유의 담지체이며 그 체계에 대한 해석의 방법론을 제공하였다.

그러므로 도상학의 등장은 당시 사회의 공적인 합리적 인식 즉 세계관을 형성하는 주요 계기가 되었으며, 더 나아가 그 시대에 관념적 사유능력을 필요로 하는 경험적 사태와 맞물려 있다. 물론 글자나 언어로도 충분히 추상적 사유가 가능하지만, 도상이 언어적 맥락을 넘어서 관념론적 사유와 그 전개과정을 표달하거나 이해하는 데에 보다 효과적인 기제로 작용할 수 있다. 이와 같이 도상학은 당시에 현실인식의 일정한 경향성, 즉 사회적 이념을 형성하는 데에 가치적 혹은 실체적 무게를 두고 세계관의 확립과 이에 대한 의사소통에 중요한 축을 담당하게 되었다.

III. 性理學의 세계와 도덕형이상학적 지평

송대에 동아시아의 전통적 문화와 사상을 대변하는 宋學⁴⁾은 송대의 학문과 사상뿐만 아니라 그 이후의 유교사회의 학문적 세계도 대표한다. 당시에 북방의 거란족의 침입, 사대부의 권력의 위축 등과 같은 국내외의 정치적 혹은 사회적 상황이 바뀔에 따라 역사의식과 자존심이 침식당하고 지리적 공간에서도 커다란 변화가 일어났다. 이러한 상황에서 사상의 측면에서는 도교와 불교가 흥성하고 있었고 다른 한편으로는 유가의 전통적 권위가 무너졌다는 위기의식이 고조되고 있었다.⁵⁾ 그러므로 어느 시대보다도 인간사회의 문제의식이 고양되고 그에 따라 도덕규범의 강령에 대한 사회적 요구가 한층 강화되었다.

이러한 분위기에서 학자들은 유가의 보편적 이념을 구체적으로 표상할 특성의 대상과 방식을 모색하였다. 초기에 권력층이자 지식층이었던 사대부는 道學의 원리를 현실에 구현하고자 그 실천의 강령으로서 內聖外王 혹은 皇極經世의 구호를 내걸었다. 인간의 삶에서 內聖 혹은 皇極이 道의 내면적 규준이라면 外王 혹은 經世는 도의 외면적 실천이다. 그들은 내성과 외왕의 관계 혹은 황극과 경세의 관계를 어떻게 서로 원만하게 조화시키는가 하는 점을 자신들의 과제이자 사명감으로 여겼다. 그러나 그들은 정치권력에서 물러나면서 외왕 혹은 경세의 외면세계를 추구하기 보다는 내성 혹은 황극의 내면세계에 치중하게 되었다. 여기에는 인

4) 송학, 도학, 宋明理學 등의 용어들은 송대 이후 유교사회의 학문적 체계를 대변한다. 이 용어들의 개념 및 정의를 이해하려면 董玉整 주편, 『中國理學大辭典』(暨南大學, 1996), 294, 298, 595쪽 참조.

5) 朱漢民/肖永明, 『宋代《四書》學與理學』, 53-73쪽.

간이 외부의 세계의 실재를 어떻게 주체적으로 수용하여 실천하는가 하는 도덕형이상학적 문제가 중요하였다.

당시에 공통의 세계관을 찾는다면 그것은 ‘천인합일’의 보편적 이념이다. 학문의 체계에서 ‘天人合一’의 이념은 인간의 자아실현을 위한 보편적 강령이었다. 유학자들은 天道와 人道 사이에는 어떤 동일성 즉 ‘천인합덕(天人合德)’의 논거가 있다고 보고, 이를 인간이 구현해야 할 보편적 이념으로 삼아 인간의 정체성(identity)과 대의명분을 찾으려 했다. 이러한 정체성은 인간의 존재론적 문제와 형이상학적 가치론으로 귀결된다. 그것은 천체 혹은 자연계의 운행과 그 질서의 원리 속에서 인간의 존재론적 의미와 의의를 찾고 이에 관한 윤리적 필연성과 타당성을 확보하려는 것이다. 여기에는 인간이 생명의 존재가치를 인식하고 이를 어떻게 삶의 원칙으로 삼는가 하는 보편적 문제의식이 함축되어있다. 우주의 궁극적 차원에서 인간의 현실적 존재를 설명하면서도 인간의 도덕적 가치를 우주의 궁극적 차원으로 끌어올림으로써 인간의 본질, 의미 및 의의를 찾고자 하였다.

유가의 도덕철학에서는 宇宙論의 조망 속에서 인간사회의 공동체적 삶과 그 합당한 방식을 모색하였다. 특히 세계의 실재를 설명하는 데에 우주의 본원이나 이치를 상정함으로써 인간의 주체적식을 생명 혹은 생명력의 본체라는 관념론적 범주에서 접근하였다. 여기에는 인간의식의 통합적 지평, 즉 궁극성(ultimacy), 본원성(primacy) 및 잠재성(potency)의 삼차원적 구조를 지닌다. 그들은 이러한 삼차원적 구조를 통해 인간의식의 문제에서 심성론의 도덕형이상학적 체계를 열어놓았다. 이는 우주의 인식론적인 조망 속에서 인간 삶의 존재론적 내용을 찾아내고 이를 바탕으로 하여 궁극적으로 개인과 사회의 가치론적인 의의를 추구하는 것이다.

유학자들은 모든 생명체의 존재론적 규준을 천체의 운행과 그 질서의 방식에서 모색하고, 특히 인간을 포함한 만물의 모든 존재를 우주의 본원적 실체로부터 파생되어 변화하는 일련의 시공간적 흐름의 과정 속에서 이해하였으며 그 속에서 인간의 존재와 그 본질적 가치의 문제를 도출하고자 하였다. 그들은 道와 器, 형이상학과 형이하학, 理와 氣 등의 범주를 설정하고 이들을 통해 심성론의 본체론적 논변을 진행하였다. 이는 도덕철학의 사유방식에서 ‘理一分殊’, ‘體用一源’ 등과 같은 방법론을 낳았으며 ‘本體論’이라는 사유방식의 지평을 열었다. 본체론적 사유방식에는 이른바 ‘본체와 그 활용(體用一源)’의 논법이 활용된다. 여기에는 인간이 생명의 존재와 가치를 올바르게 인식하고 이를 어떻게 삶의 원칙으로 수용할 수 있는가 하는 삶의 원칙이 담겨있다. 그것은 우주의 본체에서 인간의 존재론적 의미와 가치론적 의의를 찾고, 더 나아가 인간사회에서 공동체의 윤리적 필연성과 정당성을 확보하려는 방법론인 것

이다.

유학자들은 도상의 매체라는 특수성에 시공간적으로 함축된 모종의 의식의 차원에 주목하였다. 그 내용으로는 우주의 원리, 세계의 변화, 사회현상 등이 있다. 그 속에서 그들은 인간의 경험 너머에 있는 천인합일의 보편적 이념을 모색하고 을 조망하고 유가의 세계관과 그 실현의 가능성을 천착하였다. 이러한 도상학적 세계는 사회의 공동체적 의식의 차원에서 주체가 타자와의 의사소통과 그 실천적 합리성을 확보하려는 계기가 되었다. 이처럼 보편성과 특수성의 관계에 관한 도상학적 세계에서 송대 상수역학적 사유의 수준을 단적으로 엿볼 수 있다.⁶⁾

易圖學은 역사적으로 取象歸類 혹은 立象盡意와 같은 도상 혹은 도식의 고유한 해석체계와 그 접근법을 지니고 있었고, 그렇기 때문에 후대에 이들을 전달 혹은 소통되는 과정에서 그 나름대로의 실체성을 그대로 간직할 수 있었다. 이는 역도학이 시공간을 초월하여 경전과 그 풀이과정 사이에서 동일한 이해의 기제로 작동했음을 의미한다. 그러므로 역도학에서 알레고리와 같은 표상적 방식 이론적 사유의 현재성(presentness)을 지속적으로 유지하는 데에 중요한 역할을 한다. 그러므로 송대 이후에 역도학과 관련한 역대의 논변은 세계관에 함유된 이론적 사유의 본령과 내용을 효과적으로 전달하고 소통시킬 수 있었다. 특히 표상의 방식이 특정의 암유적 계도나 교훈을 효과적으로 전달하거나 소통시킨다는 관점에서 보자면, 역도학은 이른바 도학의 취지와 내용을 간명하고도 포괄적으로 함축하는 담지체이며 오늘날에서도 성리학적 세계관의 성격과 특색을 이해하는 도상학적 해석의 측면을 제공하고 있다.

IV. 『聖學十圖』의 강령과 도상학적 해석

『성학십도』는 전통적 성리학을 핵심 내용, 즉 우주론, 심성론, 수양론 등에 관한 10가지 도상과 해설⁷⁾을 총괄적이면서도 일목요연하게 정리해 놓고 있다. 『성학십도』는 성리학의 학문체계를 일목요연하게 집약해놓은 10장의 그림과 그 해설이다. 그 속에는 유학의 근본적 원리와 수양의 방법을 총망라하였으며, 특히 그가 말년에 성리학의 전통적 그림과 해설에 자신의 의견을 덧붙였다. “『성학십도』 모두가 敬으로써 주를 삼고 있다.”⁸⁾라고 주장하는 것처럼 이황은 경을 성인으로 나아가는 학문의 요지로 생각하였다. 그는 17세인 宣祖에게 강론하

6) 김연재, 『『洪範皇極內篇』에 나타난 蔡沈의 數本論과 그 세계관 - 宋代 象數易學의 사유지평을 중심으로』, 『유교 사상연구』, 42(2010), 77-78쪽.

7) 이 도식은 태극도(太極圖), 서명도(西銘圖), 소학도(小學圖), 대학도(大學圖), 백록동구도(白鹿洞規圖), 심통성정도(心統性靜圖), 인설도(仁說圖), 심학도(心學圖), 경재잠도(敬齋箴圖), 숙흥야매잠도(夙興夜寐箴圖)이다.

8) 李滉, 『退溪先生文集』 卷7, 「進聖學十圖節」: 今茲十圖, 皆以敬爲主焉.

기 위해 저 유명한 『聖學十圖』와 그것을 임금께 올리는 글인 「進聖學十圖劄」를 같이 올렸다. 선조도 이황의 뜻을 고맙게 받아들여 그것을 병풍으로 제작해 놓았다.

전반의 다섯 그림은 우주론적 도식에 입각하여 天道에 입각하여 인륜 및 덕행을 역설한 것이라면, 후반의 다섯 그림은 일종의 심성론적 도식에 입각하여 내면적 세계의 함양을 통해 도덕적 실천을 강조한 것이다. 특히 주돈이의 태극도의 도상을 필두로 하여 나머지 9가지 그림은 도식화를 통해서 聖學 혹은 道學의 강령을 나타내고 있다. 이러한 도상화 혹은 도식화는 당시의 조선시대에 맞는 공통의 이념적 사조, 즉 ‘시대정신’을 투영하고 있다. 시대정신은 특정의 시대를 살아가는 인간의 진정한 모습의 척도로서, 인생관이나 세계관과 같은 삶의 지침이자 목표가 된다. 유가의 도덕철학에서 그것은 天人合一의 이념을 목표로 하여 修己治人の 원리를 추구하는 것이다. 이는 개인과 사회, 개별성과 공공성 등과 같은 인간관계의 연속선상에서 經世主義的世界觀이 크나큰 기초를 이룬다.

이러한 도식화에서는 우주의 본체와 그 파생의 과정을 도상으로 가시화하고 이 도상을 기점으로 하여 理一分殊의 방식에 입각하여 유가의 도덕철학을 해석하는 도덕형이상학적 지평을 펼쳐내고 있다. 理一分殊의 방식은 ‘본체와 그 활용(體用一源)’의 논법을 활용한다. 이러한 해석의 지평에는 인간이 생명의 존재와 가치를 올바르게 인식하고 이를 어떻게 삶의 원칙으로 수용할 수 있는가 하는 삶의 원칙이 담겨있다. 그것은 인간의 존재론적 의미와 가치론적 의의를 찾고, 더 나아가 사회공동체의 윤리적 필연성과 정당성을 확보하려는 취지를 지닌다.

여기에서 주목할만한 점은 이황이 학문과 정치의 원리를 피력하고 이를 통해 聖學이란 무엇인지를 보여주고 있다. 특히 그는 聖人 혹은 帝王의 길을 제시하는 데에 心法에 따른 이른바 內聖外王의 도덕수양의 경지를 집약적으로 천명하고 있다. 그 전체적인 근본적 구도는 천인의 도(天人之道)에 있으며 그 핵심은 바로 敬의 관념을 중심으로 하는 수양론에 있다.

이 열 가지 도식을 매체철학의 도상학적 맥락에서 설명하면 다음과 같다.⁹⁾

「太極圖」에서 이황은 10가지 그림 중에서 제일 먼저 주돈이의 「太極圖說」과 그에 대한 주희의 해설을 첫머리에 도입한다. 여기에서 우주의 생성론의 도식을 통해 세계의 실재에 대한 인간의 시선을 드러낸다. 특히 우주의 본원과 그로부터 만물이 생성되는 과정, 즉 태극에서 음양과 오행을 거쳐 인간과 만물이 생성되는 과정을 보여주고 있다. 그는 태극과 음양 및 그 관계로부터 만물의 원리를 밝히고, 이러한 도의 원리가 바로 만물의 존재론적 근거라고 생각한다. 무극, 태극 및 음양과 오행은 모두 동일한 ‘본체’의 다른 표현이 된다. 매개체가 없는 경

9) 이 10가지 도식과 그에 대한 해설은 금장태, 『성학십도와 퇴계철학의 구조』(서울: 서울대학교출판부, 2003), 3-280쪽을 참조할 것.

힘은 존재하지 않는다는 매체철학의 차원에서 보자면, 이 태극도에서 우주의 생성의 관념이나 의미를 재현하는 데에 태극이나 음양 혹은 오행의 물리적 특성이 인간의 경험의 일종의 담지체가 된다. 이는 존재의 보편성의 문제를 특수성의 문제로 환원한 것이라고 말할 수 있다.

이화는 태극도를 제일 앞에 놓고 天人合一의 보편적 이념을 크나큰 기초로 삼고 있다. 이른바 천인합일에서 ‘천인’은 주체와 객체, 주관과 객관, 자아와 타자, 사유와 존재, 존재와 가치 등의 관계가 반영된 세계를 가리킨다. 하늘(天)이 세계에 있는 만물 또는 자연에 있는 삼라만상의 세계를 광범위하게 가리키고, 인간(人)은 그 속에 있는 생명체들 중의 하나이면서도 이들과 다른 독특한 존재를 가리킨다. 천인합일의 세계에는 기본적으로 인간과 자연을 기본적으로 구분하고 생명의 이치를 바탕으로 하여 인간과 자연과의 합치를 자각하고 자연의 통일적 질서를 인간의 가치로 수용하는 자연과 인간의 궁극적 조화의 차원을 추구하는 것이다. 그러므로 태극도에서는 천인합일의 보편적 이념을 우주에서 만물이 생성되는 일련의 과정으로 도상화한 것이다. 여기에는 태극, 음양, 오행, 乾坤의 범주를 통해 만물의 존재론적 내용을 도상화한 것이다.

그 다음은 張載의 「西銘圖」이다. 이는 원나라 때에 程復心이 도식화한 것에 기초하고 있다. 전체적으로 하나의 리가 다양한 현상으로 드러나는 과정을 理一分殊의 원리를 밝힌 것이다. ‘理一’은 理는 태극의 본원의 리로서 삼라만상에 보편적으로 존재한다. ‘分殊’는 리와 기의 관계를 통해 인간을 포함한 삼라만상의 개별적 본성을 형성한다. 이러한 全一論의 구도에 따르면, 자연계는 태극의 본원적 理와 이로부터 음양의 현상적 氣가 분화하는 일련의 유기체적 과정으로 이해된다. 여기에서는 하늘과 땅은 부모로 삼아 인간과 만물이 태어난다는 萬物一體 혹은 物我一體의 통합적 지평을 피력하고 있다.

이 도식은 위와 아래의 두 단계로 이루어진 전형적인 Tree형의 모습이다. 위의 도식은 이 일분수의 원칙을 적용한다. 그 구체적 강령으로서 仁의 一理에 따른 五倫의 分殊를 가치론적 내용을 나타내고 있다. 아래의 도식은 오륜의 分殊의 차원에서 구체적으로 事親의 誠과 事天의 道를 설명한다. 이 도식을 전체적인 구조에서 보자면 부모를 봉양함으로써 하늘을 섬기는 도리를 설명하고 있다.

「小學圖」는 이황이 소학의 체계를 도식화한 것으로서, 이 그림의 글은 주희의 「小學題辭」이다. 소학은 본래 대학을 배우기 위한 입문서이다. 立教, 明倫, 敬身の 세 강령을 기본 틀로 하여 稽古를 귀감으로 삼아 그 다음에 Tree형으로써 嘉言과 善行으로 확충하고 있다. 이는 오륜의 유교규범과 그 실천궁행의 방법을 도식으로 나타낸 것이다.

「大學圖」에서는 權近이 『대학』의 1장을 도식화한 것을 기초하고 있다. 『대학』은 본래 3

강령과 8조목을 주요 내용으로 하고 있다. 3강령은 明明德, 新民, 止至善이고, 8조목은 格物, 致知, 誠意, 正心, 修身, 齊家, 治國, 平天下이다. 이황은 이것을 내용으로 하여 體用과 知行의 구조, 공부와 그 실현과정을 제시하는데, 이는 窮理의 학문방법론과 居敬의 수양론의 기반이 된다.

「白鹿洞規圖」는 주희가 백록동 서원의 학도들에게 학규로 제시했던 白鹿洞規를 이황이 도식화한 것이다. 그 해설은 주희의 洞規後敍의 내용이다. 여기에서는 주희가 중점을 두는 학문의 의 방법을 학문교육의 기본원리로 삼고 있으며, 인륜의 실천을 전제로 하는 단계적 학문방법의 체계를 제시하고 있다. 특히 그는 學, 思, 辨, 行을 致知와 力行的 인식적 단계로 간주하고 그것을 중시하고 있으며, 窮理, 修身, 處事, 接物의 구체적인 실천영역을 강조한다.

이상의 전반부 5가지 도식에서는 대체로 우주론적 도식 속에서 인간의 삶의 강령을 설명하고자 한 것이다. 주자학의 차원에서 보자면, 천인합일의 보편적 이념은 性理의 가치론적 지평으로 귀결된다. 인간의 본성에서 군신과 부자의 본질적 이치는 현실의 상황에서 군신과 부자의 현상적 관계에 담겨 있다. 인간이 君臣이나 父子의 예의범절을 실천하는 것이 자신의 도덕적 본성을 발휘하는 것이며 天命의 정당성을 확보하는 방법이라고 생각하는 것이다. 이러한 道學은 결국에 內聖外王(혹은 修己治人)의 실천적 강령으로 귀결된다.

다른 한편, 이러한 우주론적 도식과 달리 심성론적 도식을 활용한다. 여기에는 인간의 존재론적 의미와 가치론적 의의 속에서 공동체의 윤리적 필연성과 정당성을 확보하려는 방법론이다.

우선, 「心統性情圖」는 상, 중, 하의 세 가지 도식으로 되어 있다. 상의 도식은 정복심의 「心統性情圖說」을 이황이 약간 수정한 것이고, 중과 하의 도식은 이황 자신이 그린 것이다. 전자에서는 性과 情을 마음의 體와 用으로 통합시킨 본체론적 사유를 담아내고 있다. 후자에서는 성이 발동하여 정으로 나타나는 마음의 존재 양상을 두 가지 인식의 입장에서 각각 분석한 것이다. 즉 중간의 도식에서는 본성과 마음의 理를 위주로 하는 본질적 측면을 중시한 반면에, 가장 아래의 도식에서는 理와 氣의 합일적 측면을 중시하고 있다. 이것은 마음이 리와 기가 결합된 것이라는 이황의 관점(心合理氣說)을 드러낸 것이다. 이황은 마음, 성, 정의 개념적 구성을 통해 심성론의 인식체계를 확립하고자 했으며, 그 속에 사실상 자신의 四端七情論을 총괄하고 있고 따라서 조선 후기에 심성론 논쟁의 핵심 중의 하나가 되었다.

「仁說圖」는 주희의 「仁說圖」와 「仁說」을 토대로 한 것으로서 四德의 근본이 되는 仁 개념과 그 실현적 양상을 밝히고 있다. 구체적으로 말해, 천지의 마음과 인간의 마음은 사덕을 지니고 있고 그 중에 仁의 덕은 四德 전체를 포괄한다는 것이다. 그러한 포괄적 양상은 바로 마

음의 未發과 已發에 따른 體와 用의 방식에 입각하고 있다. 특히 이황은 仁의 본체가 인간이 천지만물과 일체가 되는 근간이 되고 仁의 실현이 궁극적으로는 物我一體까지 확장될 수 있다고 생각한다. 따라서 그는 仁이라는 본연의 마음이 인간 삶의 관건이 되는 人道의 근본원리가 됨을 피력하고 있다.

「心學圖」는 정복심의 「心學圖」와 「心學圖說」을 기본으로 한 것이다. 이것은 마음이 육체의 주재이고 敬은 마음의 주재가 되는 구조를 밝혔다. 이는 마음과 경을 그림의 위와 아래의 양 축으로 삼으며 心學이 경을 근본으로 하고 있음을 밝힌 것이다. 여기에서 이황은 자신의 수양론을 펼치고 있다. 즉 마음이 체와 용의 구성, 즉 理와 氣, 性과 情의 결합으로 되어 있으며, 리와 기의 상관관계에 따라 마음이 발동하는 것을 善과 惡의 개념으로 파악하고, 따라서 선 즉 마음의 본체의 理를 실현하는 것이 心性論의 핵심이라고 생각한다.

「敬齋箴圖」는 주희의 「敬齋箴」을 王栢이 도식화를 기초로 한 것이다. 그것은 마음을 그림의 중심에 놓고 인간의 행위에서 敬을 실천하는 방법과 원리를 보여주고 있다. 「경재잡도」에서는 경의 실천적 방법과 원리가 집중적으로 거론되고 있어 이황의 수양론과 학문방법론의 핵심적 자료라 할 수 있다. 그는 경이 動과 靜의 마음의 방식 그리고 窮理와 居敬의 인식과정을 관통하고 있다고도 생각하여 그 양자 사이에 상호 보완적 실천을 도덕수양의 방법과 기준으로 삼고 있다. 따라서 그가 강조하는 敬이란 마음이 대상 세계의 무한한 변화에 대응하면서도 의식의 흠어짐(즉 욕망)을 억제하고 天理와 합일하게 되는(主一無適) 일종의 정신의 집중 상태를 의미한다.

「夙興夜寐箴圖」는 陳栢의 「숙흥야매잠」을 이황이 도식화한 그림이다. 여기에서는 敬을 그림의 중심에 놓고 새벽부터 밤까지 매 순간마다 敬으로써 일관할 것을 요구한다. 「경재잡도」가 학습태도에 따른 공부론을 언급한 것이라면, 「숙흥야매잠도」는 학습시기에 따른 공부론을 언급한 것이라 할 수 있다. 이황이 자신의 수양론에서 시종일관 중시한 경의 수양방법이란 動과 靜의 흐름에 따른 體用的 방식에 따른 마음과 사물 간의 통일적 관계를 중시하는 것이다. 이것을 전제로 하여 이황은 「숙흥야매잠도」에서 배움(學)과 사고(思)를 겸비하는 공부가 도덕적 수양을 실현할 수 있는 방법이라고 생각한다. 이것이 바로 居敬의 수양론의 본질이다. 따라서 「숙흥야매잠도」는 학문하기 위한 일련의 도식화(즉 『성학십도』)의 궁극적 취지이자 최종적 귀결인 셈이다.

종합적으로 말하자면, 『성학십도』에 담긴 10가지 도식은 전통적 성리학을 우주론, 심성론, 수양론 등에 관한 10가지 도해와 해설로써 총괄적이면서도 일목요연하게 정리해 놓고 있다. 그는 道學, 즉 학문과 정치의 원리를 피력하고 이를 통해 聖學이란 무엇인지를 보여주고

있다. 특히 그는 聖人 혹은 帝王의 길을 제시하는 데에 心法에 따른 이른바 內聖外王의 도덕수양의 경지를 집약적으로 천명하고 있다. 이처럼 『성학십도』는 그가 말년에 성리학의 전통적 내용을 기존의 그림과 말씀으로 정리하여 거기에서 자신의 의견을 덧붙였다는 점에서 이항의 학문적 세계의 원숙한 결정체라고 해도 과언이 아니다.

그는 天, 人, 心, 性의 개념을 통해 기본적으로 天道와 人道의 관계 설정을 전제로 하여 기존의 성리학의 문제들, 즉 理와 氣, 未發과 已發, 善과 惡 등의 범주에 따라 유가의 심성론에 접근한다. 내용상에서는 體用의 본체론적 사유와 밀접하게 관련되며, 우주생성론의 방식, 理一分殊의 원리, 心統性情의 원칙 등을 토대로 하고 있다. 따라서 그는 孔孟學에 관한 道統과 義理의 變奏의 방법을 통해 한편으로 인간본성에서 도덕적 함양의 본체가 무엇인지를 밝히며, 다른 한편으로 우주의 궁극적인 존재론적 차원에서 先秦시대 유가의 본원적 가치를 모색하였다.

도상은 인간과 세계를 매개하는 양식으로서 세계관에 관한 정보와 지식을 제공한 것이다. 여기에는 매체의 재현적 혹은 표현적 방식을 통해 성리학의 본체론적 사유를 상징적으로 밝히는 해석의 방법론이 담겨있다. 이 방법론에 따르면, 가상의 본체를 상징하여 구체적 현상을 설명하는 데에 특정의 도상을 활용하고 그 속에 전달하고자 하는 취지나 의도를 반영하고 특정의 교화나 계도를 피력하는 것이다. 이는 聖學의 정보와 도상을 동일선상에 놓는, ‘정보의 도상화’와 ‘도상의 정보화’라는 양측면을 지닌다.

도상 혹은 도식의 구성에는 실재(reality)의 선형성과 경험성 및 그 양자의 관계를 어떻게 합리적으로 설정할 것인가 하는 효과적인 기법이 중요하며, 또한 도상의 이해에는 그 속에 담긴 역사적 자료를 이해하고 그 요체를 직관적으로 파악하여 전체내용을 심리적으로 수용할 수 있다. 유가의 도덕철학에서 보자면, 도상학은 세계의 실재에 대한 인간의 시선 속에서 도덕형이상학적 지평을 피력하는 방편이 된다. 인간은 세계관 속에서 사회의 전체적 구조, 그 하부의 계층들 및 이들의 관계를 통해 대체, 해체, 유지, 변화, 재건축 등의 과정을 거치면서 통일적인 가치의 지향성을 모색해간다. 이로한 통일성의 과정에서 혼돈과 질서가 경계지워진다. 즉 주체와 객체의 일정한 관계를 인지하고 파악하는 단계에서 이러한 관계를 구현하고 실현하는 단계로 나아간다. 유가의 도덕철학에서 보자면, 이러한 과정에는 존재와 가치가 통합되는 경계, 즉 도덕형이상학적 경계가 있다. 이는 인간의 내면적 세계에서 도덕성의 본체가 무엇이며 그것을 어떻게 함양할 것인가 하는 문제의식과 관련한다. 여기에는 孔孟學의 계승과 발전이라는 道統意識의 시대적 정신 하에서 도덕적 실천과 그 수양의 강령을 담아내고 있다.

V. 나오는 말: 聖學의 도상학적 경계

매체철학의 논법에서 圖像學(Iconology)은 문헌 혹은 경전의 내용을 이해하기 위해 그것과 표현된 이미지와의 연관성을 규명하는 것이다. 즉 도상에 담겨진 내용의 정보전달 전반과 관련한 의미의 연관성, 그 해석의 가치 등을 총체적으로 파악하는 것이다. 그것은 형상화되거나 도식화된 이미지와 그 내용을 어떻게 이해할 것인가 하는 문제와 관련된다. 여기에는 도상이나 도식의 구성, 해석, 가치 등을 전반적으로 분석하며 또한 그 속에 반영되거나 투영된 역사적 배경, 문화적 성격 및 사회적 구조까지도 共時的이고도 通時的으로 해석하는 과정이 있다. 그러므로 도상학에서 도상을 해석하는 과정에서 ‘생명의 창조력’을 끊임없이 구성하면서 다른 한편 ‘창조의 생명력’을 끊임없이 만들어낸다.

도상학의 기제는 시대적 혹은 사회적 욕구를 표출하고 충족하려는 시도의 단면을 보여준다. 도상학에서 참다운 인간학적 대강령, 즉 聖學을 계도하고 교화하는 문명의 효과를 기대하고 이를 충분히 이론화할 수 있었던 것도 이러한 욕구를 충족시킬 수 있었던 태극도, 선천도 및 하락도와 같은 매체의 역할과 이를 해석한 도상학적 내용을 파악할 수 있었기 때문이다.

도상학은 인간이 인식한 것을 묘사하는, 단순히 수동적 의미의 재현적 도구만이 아니라 인간의 인식의 기준을 제공하는 능동적 의미의 표현적 담지체이다. 특히 도상이나 도식에 반영된 의미의 結晶體로서 인간의 주체가 세계의 객체를 인식하는 단서가 된다. 그러므로 도상학이 인간 주체의 단순한 발명품이라든가, 아니면 인간의 의지에 의해 사용되는 중립적 도구가 아니다. 도상은 인간과 세계를 매개하는 양식으로서 인간과 세계의 관계망에서 쌍방향성의 연결통로의 역할을 한다. 그것은 당시의 세계관에 관한 정보와 지식을 제공하는 것이다. 여기에는 ‘정보의 도상화’와 ‘도상의 정보화’라는 양측면이 존재한다. 전자가 매체의 재현적 방식과 관련된다면 후자는 매체의 표현적 방식과 관련된다. 그 양자를 유기적으로 통합한 圖書學 혹은 易圖學은 성리학의 본체론적 사유의 담지체이며 그 체계에 대한 해석의 방법론을 제공하였다. 이러한 해석의 방법론은 고전적 가치를 현대적 시각으로 접근하고 이해하는 소통하는 매체의 방식인 것이다.

도상학은 세계의 실재에 대한 인간의 인식에 상과 수의 형식적 요소를 매체로 한다. 상과 수의 매체를 통한 연역의 방식은 만사만물의 존재를 설명하기 위한 최초의 전제이자 그 내용을 드러내기 위한 기본적 기준이 된다. 인간의 경험과 지식은 도상이나 도식의 형식을 통해 나타낼 수 있으며, 따라서 그 방식은 천하의 經綸을 표명하는 중요한 척도가 된다. 유가의 시대정신에서 보면, 이는 天人合一의 보편적 이념, 즉 天道의 필연성에 따라 人道의 정당성을 확보하고자 하는 것이다. 이를 개인과 사회의 관계로 확장하면 天道의 가치론적 지향성을 人道

의 공동체의식에서 실현하는 것이다. 그 양자의 관계성은 공동체의 가치화와 가치의 공동체화의 지평에서 이해될 수 있다. 따라서 이러한 매체철학의 도상학적 해석은 항상 끊임없이 변화하는, 결코 완결되거나 완성되지 않는 연속적이고도 지속가능한 과정이라고 말할 수 있다.

결론적으로 말해, 聖學의 도상학적 경계에는 도상이나 도식의 구성에 투영된 시선, 즉 세계의 선형성과 경험성, 通時性和 共時性 등의 관계를 어떻게 합리적으로 설정할 것인가 하는 논법이 담겨있다. 이는 해석의 기제가 어떻게 작동하는가 하는 매체철학의 문제로 환원된다. 여기에는 매체의 재현으로서의 ‘정보의 도상화’와 매체의 표현으로서의 ‘도상의 정보화’라는 양측면이 작동하고 있다. 이러한 점에서 『성학십도』의 도상학적 해석은 시대마다 끊임없는 생명력을 발휘할 수 있다. 우리는 이를 통해 유가정신의 지평을 비교철학의 논점에서 접근하고 현대의 관점에서 전통문화의 지평을 모색할 수 있는 것이다.

최석정 『구수략』의 이진법 독법

배선복
한국체육대학교

개요

(易)의 사상(四象)체계를 가감승제(加減乘除)로 보는 것은 소강절 이론이다. 사상 팔괘를 음양으로 환원시켜 보는 것이 이진법 독법이다. 최석정은 먹과 붓 뚜껑으로 음양 -- --에서 ●○을 찍고, = :: == ≡ 사상을 일월성신(日月星辰) 현상에 대응시켜, 가감승제로 + - × ÷ 사칙연산을 하였다. 음양 -- --을 오로지 0과 1로만 읽는 이진법독법은 복희씨의 팔괘와 음양 차서도와 팔괘중첩의 64괘 구성을 망라한다. 라이프니츠는 이진법을 창안하고 잠수함을 설계하고, 컴퓨터까지 제작해보므로, 불(J. Boole)에서 추체(K. Zuche), 에니악(Eniac)을 거쳐 인공지능머신까지 디지털지식의 토양을 제공하였다. 최석정이 동양수학의 구조를 바꾸려고 『구수략』을 썼다는 점만 강조되는데, 그는 음양 ●○을 영락(零落)없이 이진법으로 읽고 가감승제로 사상을 들여다보았다. 라이프니츠는 역을 숙지한바 없고 한자는 몰랐고, 최석정은 이진법은 몰랐지만 역시 서양수학을 도입하므로, 두 인물은 같은 년도에 다른 대륙에서 태어나 비슷한 선구적인 생각으로 디지털 세계를 열었다. 이진법에서 소강절 육십사괘 방원도 독법 경위, 부베가 강희제에게 제시한 「천존지비도(天尊地卑圖)」, 최석정 『구수략』의 「태양지수」의 천지지교(天地之交)의 음양분해까지는 일목요연한 연관관계가 성립한다.

1. 서론

최석정의 음양 독법과 이진법의 동일성 논의에는 잘 알려지지 않는 역사적 배경이 있다.¹⁾ 먼저 부베는 역의 구성이 이진법으로 읽힌다는 점을 알고, 라이프니츠에게 소강절의 육십사괘 방원도를 보내면서 원자자의 승인을 구한다.²⁾ 라이프니츠는 부베의 요구와 복희씨의 팔

1) 국내 수학계에서 김용운, 홍성사 등은 그들의 대표적 논문에서 최석정을 사대부 수학에 머물렀다고 비난한다.

괘의 귀속을 인정하면서 루돌프 대공에게 보낸 「신년하례연하장」에서 역의 이진법 배경을 해명한다. 부베는 이진부호 0과 1 대신 음양 ●○ 상징으로 강희제에게 「천존지비도」에서 1은 ○, 2는 ●●, 3은 ○○○으로, 10은 ●의 열 개로, 이런 방식으로 1에서 10까지 하도의 수열을 분해한다.³⁾ 최석정은 부베에 준하는 「태양지수」에서 아라비아 숫자 대신 붓 뚜껑과 먹물로 음양 ●○을 찍고, 천지지교의 가감승제를 + · - × ÷ 사칙연산 하므로 라이프니츠 「신년하례연하장」의 생각에 도달한다.⁴⁾ 그것은 비트겐슈타인이 『논고』에서 말한 16가지 명제 논리학적 진리근거와 상통하는 것이다. 경전에 따라, 태극에 이미 음양이 있으므로, 순수 수는 0과 1 외에 다른 수가 없다. 순수 수가 태양과 태음이면, 소음과 소양은 상교 하고, 어떤 상(象)도 순수 수를 갖는 한, 사상이 팔괘를 낳고, 팔괘가 중첩(重疊)되면 64괘는 우주의 총체상이다. 64괘 해체는 음은 감이고 양은 가이며, 가는 나가고 감은 물러서는 진퇴로 완벽하게 음양 소식으로 대응한다.

최석정은 가현의 삼각형 패턴에 따라, 꼭짓점 초열은 천 ○을 찍고 둘째 열은 지 ●●를 찍고, 셋째 열은 천 ○○○을 찍고, 마지막 열여섯째 열은 열여섯 개의 천 ○을 찍어 16괘의 상징체계를 구성하여 태양지수에서 순수 음양의 전개과정을 보여준다. 이는 ●○에 기초하여 1에서 16까지 천지지교를 가감승제로 풀이하므로 16 괘의 천지 값 ●○을 가감승제로 분할하고 천지지교의 태양지수를 음양으로 환원시킨 것이다. 태양지수도상에는 숫자가 적혀있지 않아서 어떻게 음양이진과정으로 유추되는지 알 수 없지만, 천지지수는 온통 음양으로 변하며 천지시간이 바뀐다. 최석정이 『구장산술』 각 장을 임의 형식으로 가감승제 일월성신 가감승제에 할당하였다지만 그렇다고 수학의 전체 구조를 음양에서 분해하려고 한 의도가 분명하게 드러나지 않은 것은 아니다. 그는 일월오성 음양소식을 이진논리학 부호 값 11 10 01 00 과 팔괘만으로도 원활하게 셀 수 있는 기반을 마련하였다.

2. 공통출발점

비트겐슈타인은 『논고』에서 진리배경에 대하여 16가지 명제 논리학적 진리결합 가능성을 밝혔다.⁵⁾ 그 근거는 『프린키피아 마테마티카』이지만, 가감승제의 음양사상 팔괘의 이진법독법과 연관된다. 임의 명제 P는 이가윌리에 따라 T, F로, P와 Q는 (T, T), (T, F), (F, T), (T,

2) G. W. Leibniz(2006).

3) G. W. Leibniz(2006).

4) 『구수략』.

5) L. Wittgenstein(1969), 44-45.

T)로, T, F를 이진법부호 0과 1로 변환하면, 명제논리진리표를 계산할 수 있다. 한 명제의 언표의 T, F의 진리치에 대하여 동일한 상징적 추론형식으로 음양 -- — 과 사상 ::, ==, ≡, = 은 0, 1와 00, 01, 10, 11로 변형된다. 음양사상의 대수적 변용은 전산수학의 논리회로 학습, 유전자코드의 구조의 설명, 명제논리학의 진리표, 컴퓨터 공학, 정보통신 아스키코드 등에서 광범위하게 활용된다. 라이프니츠는 오로지 0과 1로만 -- — 음양을 읽으므로, 복희씨 팔괘와 음양차서도와 팔괘증첩의 64괘의 이진법독법에 도달하였다. 그는 1697년 루돌프 대공에게 보내는 「신년하례연하장」에서 음양을 0과 1로 변환하는 이진법 독법을 널리 소개하고 0과 1의 이진법으로 도해된 음양 사상 십육괘를 세계창조의 상(imago)이라 불렀다. 대표적으로 소강절은 「음양차서도」에서 음양소식(陰陽消息)의⁶⁾ 16괘를 나타냈고, 최석정도 ●○음양으로 16열 삼각형 형태로 태양지수를 쌓고 이를 천지지교라 불렀다. 최석정의 ●○표시는 소강절의 『황극경세서』에 문자는 없으나 성(聲)이 있으면 ○, 음(音)은 있으나 문자가 없으면 □, 성도 문자도 없으면 ●, 음과 문자도 없으면 ■로 성음창화를 표시하였을 경우와는⁷⁾ 무관하나, 부베가 강희황제에게 1711년에 「천존지비도」에서 하도의 수 체계를 지시하기 위하여 사용하였던 ●○표시와 일치한다.⁸⁾ 부베는 라이프니츠와의 서신교환에서 역의 상징체계를 이진법으로 해독할 수 있는 길을 공유했지만 정작 강희황제 앞에서는 0과 1의 이진법의 수적 표현 없이 ●○로 이진수의 존재를 간접적으로만 암시하였다. 최석정은 부베가 표시한 ●○과 같이 명백한 자의식으로 16괘의 태양지수의 천지지교를 이진법적으로 분해하였다.

3. 라이프니츠

라이프니츠는 1646년 7월 1일 독일 라이프찌히에서 태어나 1716년 11월 14일 죽기까지 『모나드론』이라는 주저를 포함하여 오늘날까지도 전권이 나오지 않는 가운데 여전히 편집되고 있는 헤아릴 수 없이 많은 작품을 후세에 남겼다. 쿠플레(1623-1693)가 1687년 『중국 철학자 공자』의 부록에서 음양 사상 팔괘를 포함하는 역을 최초로 유럽에 소개하자, 라이프니츠는 1688년 페르비스트 후임인 그리말디를 로마에서 만나면서 그에게 보낸 1693년 편지에서 역의 이진법독법을 언급하였다.⁹⁾ 그 후 그는 루돌프 대공에게 보낸 「신년하례연하장」

6) <https://www.eee-learning.com/article/5061>

7) 「관물내편」35에서 50까지 성음창화도(聲音唱和圖) 참조.

8) Sophie Ling-chia Wei(2015), 180.

9) 라이프니츠와 그리말디 사이의 로마회동비망록은 모두 30가지 관심사항에 대한 정보교환인데 반해 1695년 편지에 이진법해명이 있는 것으로 보아 선교사업에서 정보의 암호화 측면에서 이진표기를 언급했을 가능성이 높다.

에서 음양, 사상, 팔괘와 십육괘가 이진법 수열로 등장하는 세계 창조의 상을 제시한다. 나중으로 동메달로 제작된 이 연하장에서 좌측은 2진법으로 0, 1, 2, 3, 4, 5, 6, 7까지 8괘의 수열을 (0, 0), (1, 1), (10, 02), (11, 3), (100, 04), (101, 5), (110, 6), (111, 7)로 표기한다. 십진법이 1, 0, 4, 8, 13이면, 이진법 자릿수 표기는 1, 00, 100, 1000, 1101이다. 따라서 (1000, 8), (1001, 9), (1010, 10), (1011, 11), (1100, 12), (1101, 13), (1110, 14), (1111, 15)의¹⁰⁾ 사상은 이진자릿수에 따라 16괘 변화주기를 담는다.

부베는 1697 4월 파리에서 라이프니츠의 『최신중국소식』편지를 받고 서신교환을 하였다.¹¹⁾ 그는 1701년 2월 15일 라이프니츠의 편지를 받고, 1701년 11월 4일 북경에서 라이프니츠의 이진법 발견에 답신을 하였다. 1703년 4월 1일 라이프니츠에게 도착한 이 편지에는 복희씨 육십사괘방위도 목판 사본이 들어있었다. 라이프니츠는 부베의 편지를 받자 1주일 이내에 파리학술원 비서관 비농(A. Bignon)에게 파리학술원 비망록(*Mémoires*)에 실을 「이진법 산수의 해명」이라는 논문을 보냈다. 그는 건 태 리 진 손 간 감 곤의 팔괘에서 이진법 수학의 기능을 수학적 철학적 신학적 진리의 발견수단이라 하였다.¹²⁾ 곤 ≡ 태 ≡ 리 ≡ 진 ≡ 손 ≡ 감 ≡ 간 ≡ 건 ≡ 은 000 001 101 011 100 101 110 111로 표기된다. 이는 0, 1, 2, 3, 4, 5, 6, 7까지 주기로 도 레 미 파 솔 라 시 음정이나 빨 주 노 초 파 남 보 무지개 색 등, 8로 순환하는 모든 주기를 표현한다. 이진법 110010101101을 십진법으로 표기하려면 밑수 2에 대한 자릿수 표기를 순서대로 더해 나타낸다. $1 \times 2^{11} + 1 \times 2^{10} + 0 \times 2^9 + 0 \times 2^8 + 1 \times 2^7 + 0 \times 2^6 + 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 = 3245$ 이다. 이 표기법은 거꾸로도 타당하다. 베르누이는 $1701 = 2^{10} + 2^9 + 2^7 + 2^5 + 2^2 + 1 = 11010100101$ 로 읽으며 라이프니츠가 파리학술원으로부터 공인받은 이진법 발견년도를 기념하였다. 라이프니츠는 “-은 통일로서 1, --는 깨진 선으로 0을 보이기에 충분하다.”고 하면서, “나는 2년 전에 북경의 부베에게 보냈더니 그는 1701년 11월 4일 회신하였고 의사소통가운데 나의 제안으로 역의 상징체계를 해독하였다.”고 하였고, “역을 지구상에 존재하는 인류의 가장 오랜 기념비적 작품”으로 인정하였다.

10) 다음 칸에는 (10000, 16), (10001, 17), etc.로 16괘와 이후의 진행을 나타냈다. 위에는 ‘무로부터 만물을 창조하기 위하여서는 1이면 충분하다 *omnibus ex nihilo ducendum Sufficit unum*’, 아래는 ‘1697년 세계창조의 상 *imago creationis Anno Chr. MDGXXVII*’이라 적었다.
 11) 부베는 강희제의 위임으로 파리에서 머물면서 1698년에 8명의 천문학지리학자를 데리고 중국으로 들어온다.
 12) 라이프니츠는 1698년 3월 29일과 5월 17일 요한 크리스티안 슐렌베르크에게 편지에서 반복으로 설명하였다.

4. 최석정

최석정은 본관이 전주이고 1646년 7월 2일에 충북 진천군 초평면 금곡리에서 태어나 1715년 12월 6일까지 9번이나 영의정에 오른 조선의 명신이다. 그는 12세에 『주역』을 통달하였고, 17세에 초시장원을 하고 1671년에 급제하여 『에기유편』, 『명곡집』 36권, 『구수략』 등의 저작을 남겼다. 그는 1687년에 이민철과 송이영의 『혼천시계』에 대한 선기옥형 개보수에 참여하여 「자명종명」을 남겼다.¹³⁾ 1685년에는 호조 참판 겸 연행부사로, 1697년에는 우의정 왕세자 주청사로 북경을 다녀왔는데¹⁴⁾ 당시 중국에는 페르비스트, 그리말디, 토마스, 부베, 슈토프, 자르투 등이 있었다. 그들은 과학기술전문가 집단으로 중국전국지도 제작을 위한 『황여전람도』 계획에 참여하였고, 최석정은 아담샬의 『적도남북총성도』를 모사한 『곤여만국전도』를 제작하여 1708년 숙종에게 진상(進上)하며 국제적인 학술동향에 어둡지 않았다.¹⁵⁾ 혼천의 천문시계 개보수 작업과 서양천문도에 기초한 예술적 세계지도제작을 주관한 것은 국가과학 입국이라는 비전과 전문지식이 없이는 나설 수 없는 일이다.

부베의 저술이 최석정의 음양이진법 독법에 어떤 연관관계가 있고 영향을 미쳤는지를 추측하기는 어렵다. 부베는 역과 한자어에 담긴 비전의 기독교진리의 해석에 전념하고 1719년 수아레즈(J. Suarez, 蘇霖, 1609-1677)와 같이 동지부사 조영복(趙榮福, 1627-1728)을 방문하면서 조선학자들과 교류를 한 것으로 보인다.¹⁶⁾ 그는 1720년에는 두세 차례나 조선 진사 이기지(李器之, 1690-1722)의 답방을 받고 필담을 나누며 조선의 지리와 목극등과 함께 가본 장백산, 조선의 바다에서 나는 해삼에 관하여 담론을 나누는 적이 있다.¹⁷⁾

학자들의 추정대로, 최석정은 1710년에 관직을 접으면서 1710년에서 1715년 사이의 기간에 『구수략』은 집필하였다. 최석정은 소강절이 태양은 해, 태음은 달, 소양은 별, 소음은 신이 되는 것을 하늘의 사상¹⁸⁾이라 한데에서 진일보하여 사상을 일월성신에 대응시킨 가감승제

13) 『明谷集』숙종실록 18권, 숙종 13년 7월 15일에 上命崔錫鼎、李敏哲、改修審幾玉衡이라는 구절이 등장한다. 최석정은 이듬해 「齊政圖記」작성하고, 남구만과 같이 관상감 천문학자의 추산법 학습을 제안하고, 1695년 역법모임 결성에 도움을 준다. 영의정에 오른 최석정은 1701년 역법연구모임을 부활하였고, 1705년 동지사행에 천문학자 파견을 주관한다. 최석정과 우의정 김창집은 1706년 허원이 일월오성초면추보지법(日月五星初面推步之法)을 학습하였으니 상을 주자고 하므로, 허원은 중6품에서 파격적으로 정3품 통정대부에 가차된다. 최석정은 관상감 영사로서 1708년 「서양건상곤여도이병총서」를 작성하여 진상한다.

14) 『조선왕조실록』숙종 23년, 9월 6일 주청사 최석정이 북명하였다. 최석정이 1685년 연행부사로 갔을 때 그리말디, 페르비스트가 있었다. 1697년 중국에 갔을 때 부베는 파리에 있었고, 그리말디는 활동 중이었으나, 부베는 1708년에는 만주의 흑룡강 근처에 머물렀을 것이다.

15) 조선왕조실록』병조판서 이이명도 숙종 32년 1706년에 「요계관방도遼蕪關防圖」를 작첩(作帖)하여 진달한다.

16) 신익철(2013), 17-18, 조영복(1998), 114-115.

17) 신익철(2013), 28-117. 1708년에는 괴글러가 있었다.

18) 『皇極經世書』, 「纂圖指要」, 太陽爲日 太陰爲月 少陽爲星 少陰爲辰 日月星辰 交而 天文體盡之矣. 『구수략』에는 陽

산법을 만들었다.¹⁹⁾ 그것이 1이 태극, 1이 양의를 낳고, 양의가 사상을 낳을 때, 사상이 일월 성신을 썬하는 음양사상 독법이다. 이 썬법의 목표는 천지지수를 음양으로 환원시키고²⁰⁾ 사상을 가감승제로 썬하므로, 지구의 태양공전인 기영(氣盈)과²¹⁾ 달의 지구공전인 삭허(朔虛)²²⁾, 즉, 기영삭허 문제를 남의 손을 빌리지 않고 자립적으로 계산하자는 것이다. 최석정은 이 썬법을 새로운 발견이라고 하였다.²³⁾

5. 하도낙서의 수원(數源)

수원은 도에서 유래하며²⁴⁾ 하도의 수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10이다. 1, 3, 5, 7, 9는 홀수 천수(天數), 2, 4, 6, 8, 10는 짝수 지수(地數)이다. 전자는 양수, 후자는 음수이다. 가감승제에서 더하는 수는 곱하므로 늘어나는 승수이나, 덜 하는 수는 나누는 수로 사라지는 제수이다. 가하면 천수는 $1+3+5+7+9=25$, 지수는 $2+4+6+8+10=30$ 이다.²⁵⁾ 각각의 합은 25, 30이다. 1, 2, 3, 4, 5 가운데 1, 3, 5는 양의 천수이고, 2, 4는 음의 지수이다.²⁶⁾ 중앙 5에서 1은 북쪽, 겨울, 수, 2는 여름, 남쪽, 화이다. 수와 화는 반대방향에 서있다. 3은 동쪽, 봄, 목, 4는 가을, 서쪽, 금이다. 목과 금은 반대방향에 있다. 5는 방위기준으로, 각 방향으로 수를 더하면 $5+1=6$, $5+2=7$, $5+3=8$, $5+4=9$, $5+5=10$ 이다. 이때 1, 2, 3, 4, 5는 생수, 6, 7, 8, 9, 10는 성수이다. 생수는 양수, 성수는 음수인데, (1,2), (3,4), (5,6), (7,8), (9,10)은 상득이다.²⁷⁾ 드러난 방위는 (1,6)=수, (2,7)=화, (3,8)=목, (4,9)=금, (5,10)=토로 나타난다. 해석기하학적으로 5를 중심으로 보면, 좌로 음, 우로 양, 위로 양, 아래로 음이 움직인다.²⁸⁾ 양수와 음수, 생수와 성수는 자

之陽爲太陽 陰之陰爲太陰 陰之陽爲少陽 陽之陰爲少陰으로 적었다.

19) 『구수략』 加減乘除爲算法之四象 分以言之 四象各具四數 陽之陽爲太陽 陰之陰爲太陰 陰之陽爲少陽 陽之陰爲少陰 太陽爲日 太陰爲月 少陽爲星 少陰爲辰 天地之間只有四象而已 數之理雖至深至蹟亦豈外於此哉.

20) 『구수략』 嗚呼. 天有萬化而不外乎陰陽 易有萬變而不外乎剛柔 筮有萬術而不外乎乘除 陰陽正位而不亂 剛柔迭用而成章 正位法乎天 迭用則乎易者 其乘除之術乎 若由是而引而伸之 觀小道而悟大德者 存乎其人.

21) 『구수략』 太陽之數는 參加이다. 一歲有十二月 每月三十日 一歲日數幾何. 1년은 12달, 매달은 30일이다. 일 년은 며칠이나?

22) 『구수략』 太陰之數는 參減이다. 問一歲月數幾何. 일 년은 몇 달이나?

23) 『구수략』 今發四象新義以解九章諸法 觀者毋謂創說而忽諸.

24) 『구수략』 【갑】 「수원」.

25) <https://ctext.org/book-of-changes/xi-ci-shang/zh>.

26) 생수 1, 3, 5를 가하여 쌓으면 9 태양 ==, 2, 4를 가하여 쌓으면 6 태음 == 이다.

27) <https://ctext.org/book-of-changes/xi-ci-shang/zh>. 數五, 地數五, 五位相得, 而各有合.

28) 낙서의 수는 10이 없는 가운데, 1, 2, 3, 4, 5, 6, 7, 8, 9로, 전자는 홀수이고 후자는 짝수이다. 낙서의 방위는 중앙에 5를 두고 상하좌우로 양의 홀수와 음의 짝수가 상합을 이루고, 북 1 간, 서남 2 곤, 동쪽 3 진, 동남 4 손, 북서 6 건, 서 7 태, 동북 8 간, 남 9 리로 방위를 잡는다. 5를 중심으로 1, 3, 7, 9를 천수, 2, 4, 6, 8을 지수 가로 세로 대각선으로 합이 15가 되는 마방진이다.

연수의 연속이 아니다.

최석정은 「하도변수」에서 천지지교를 음양으로 표기하고 삼각형 퇴타 형식에서 적(積)으로 태양지수를 구했다. 삼각형 형식에 수를 쌓은 태양지수는 피타고라스 테트라틱스($\tau\epsilon\tau\rho\alpha\kappa\tau\acute{\upsilon}\varsigma$)나 파스칼 삼각형과 같지만 고법칠승방도, 곧 가현의 삼각형 패턴을 따른다는 점은 의심할 여지가 없다. 그는 처음 초열에 천으로 ○을 찍고 다음 둘째 열에 지로 ●●를 찍고, 셋째 열에 천으로 ○○○을 찍은 순서로 마지막 열여섯째 열에 열여섯 개의 천 ○을 찍어서 16괘 상의 상징체계를 도해하였다.²⁹⁾ 이는 ●○에 기초하여 1에서 16까지 수열의 천지지교를 가감승제로 더하면 홀, 덜하면 짝, 혹은 더하면 짝, 덜하면 홀로서, 사상이 합을 분해하여 표현하는 16괘의 이진수열을 해석한 것이다.³⁰⁾ 16괘의 천지 값 ●○을 가감승제로 분할하므로 생겨나는 붓 뚜껑의 미학이 보여 주는 대로, 천지지수의 구조가 온통 음양으로 변하고, 천지기간이 음양으로 바뀐다.

부베는 「천존지비도」에서 음양을 ●○로 표기하면서 하도의 1에서 10까지를 구체적으로 도해하였다. 1은 ○, 2는 ●●, 3은 ○○○으로, 내려가면 10은 ●의 열 개로, 1에서 10까지 천지천지천지천지천지의 삼각형 형태로 내려간다. 부베가 나타낸 이진수가 1에서 10까지의 하도의 수를 목표로 분석하였다면, 최석정의 태양지수는 16괘와 관련 된 라이프니츠의 「신년하례연하장」세계창조의 상과 일치한다. 최석정은 ‘하늘은 3으로 하고 땅은 2로 하여 수를 의지함’³¹⁾은 ‘태극이 1이고, 1은 2를 낳고, 2는 3을 낳으나, 3과 2로 수가 서다’³²⁾라며 천지기간의 방원곡직의 수가 서 있음을 지칭하므로, 천원지방에서 행성 주기를 측정하여 밤낮의 길이를 재고 달이 차고 기우는 주기와 4계절의 길이를 재는 계산을 정당화한다.³³⁾

6. 강설(降雪)과 자유낙하의 수원

하늘을 마름질한 천문학의 이론적 토대이며, 수원의 해명은 케플러(J. Kepler, 1571-1630)의 강설분석, 갈릴레이(G. Galilei, 1564-1642)의 자유낙하실험에서 비롯한다. 케플러는 자연현상에 감추어 ‘숨어있는’ 근접원인에서 원 안에 n번째 내접한 다각형면적을 구했고, 갈릴

29) 조선 산학에서 법(法)과 실(實)은 $\frac{\text{실}}{\text{법}}$ 으로 표현한다. 실은 나누어지고 법은 나누는 수로, 분자와 분모이다.

최석정은 갑 16-1=15, 가 2+15=18, 승 15×18=270, 제 270÷2=135을 통하여 ∴ 135+1=136을 구했다.

30) 『구수략』 數原於一一爲太極一生二兩儀也 二生四四象.

31) 『구수략』 參天兩地而倚數.

32) 『구수략』 太極者一也, 一生二, 二生三, 參兩而數立矣.

33) 고희민(2017). 「觀物外篇」에서 팔괘 방위도는 천지개벽(天地開闢)과 일월출입(日月出入)을 도출하고 춘하추동(春夏秋冬) 회현삭망(晦朔弦望) 주야장단(晝夜長短) 행도영축(行度盈縮)에 관련된다.

레이는 원둘레를 10,000각 정도로 쪼개는 것으로 다각형면적을 구했다. 케플러는 1, 1, 2, 3, 5, 8, 13, 21, 등으로 비율로 떨어지는 강설에서 진행되는 5의 8은 8의 13에 대한 것, 8의 13은 13의 21, 등에 대한 피보나치(1180-1250) 수열에 숨은 규칙성을 발견하여 강설 비율 ϕ 를 1.618로 보았다. 케플러는 여섯 봉의 깃털모양을 갖춘 육각 형태를 갖춘 눈송이 형태에서 천 지지간에 육각형에 신적 비밀이 깃들여 있을 것이라 추측하면서 내리는 눈을 감상하였다. 최석정이 태양지수에서 제시한 수열에 빗금을 쳐서 합산하며 읽어 가면, 케플러가 발견한 동일한 피보나치수열을 확인할 수 있다. 라이프니츠는 하위헌스가 1673년 『진자시계』를 완성한 다음 분수 수열의 합을 구하라는 문제제기에서 이항정리 $(a + b)^n$ 의 계수를 나타내는 ‘파스칼의 삼각형’의 역순을 취하면서 파이 π 를 4등분하는 조화 삼각형을 발견하였다. 최석정은 태양지수에서 피보나치수열과의 연관성을 명백하게 인식하지 못했지만,【정】편 이후 부록에서 「의장일면첨타(倚牆一面尖塚)」라는 퇴타 형식 등을 적설이나 성벽 쌓기 같은 수학문제로 분명하게 인식하고 있었고 「낙서육고도(洛書六觚圖)」는 최대밀집농도의 수학적 가능성을 탐색한 것이다. 나아가 최석정은 주세걸도 증명 없이 옮겨 적어놓은 가현삼각형에 대하여, 가현삼각형을 비스듬하게 순서대로 읽으면 나타나는 초교타, 감각타, 삼각살성타, 삼각살성낙일형타, 등을 증명 없이 옮겨 적었는데, 이는 서양수학의 발전과는 다른 조선수학의 독립적 위상이 존재하였음을 보여준다. 초교타는 자연수 수열, 삼각타는 삼각수 수열, 삼각성형타는 파스칼 삼각형같이 피라미드 수열로 알려지는데, 아라비아 수 체계의 도움이 없이 오직 전통 산법의 산대만을 이용하여 기하학 도형을 대수적으로 해석하였기 때문에 잘 알려지지 않은 채 전래되었다.

7. 복희씨육십사괘방원도와 음양이진사상독법

최석정은 『구수략』에서 각 장을 임의 형식으로 분류하고 수학의 전체 구조를 음양에서 분해하였다. 『구장산술』구수(九數)에서 가는 일(日), 태양 ☰, 방전(方田), 감은 월(月), 소음 ☷, 속미(粟米), 소광(少廣), 승은 성(星), 소양 ☱, 상공(商工), 쇠분(衰分), 영육(盈朒), 제는 신(辰), 태음 ☵, 균수(均輸), 구고(勾股), 방정(方程)으로 분해된다. 순수 수는 0과 1이다. 태양과 태음은 순수 수이다. 태극에 이미 음양이 있으므로, 순수 수는 0과 1 외에 다른 수가 없다. 소음과 소양은 상교 한다. 어떤 상도 순수 수를 갖는다면, 사상이 팔괘를 낳을 때, 팔괘의 중괘로 생성된 64괘는 우주의 총체상이다. 64괘의 해체는 음은 감이고 양은 가이며, 가감이 진퇴에 완벽하게 음양소식으로 대응한다.³⁴⁾

라이프니츠는 1703년 4월 1일 말 그대로 원 안에 사각형에 육십사괘가 있는 복희씨육십사

괘방원 도를 대면하였다. 라이프니츠는 방원도의 방향을 읽기 위하여 아래로 $\kappa\acute{\alpha}\tau\omega$ 위로 $\acute{\alpha}\nu\omega$ 라는 희랍어 문구를 적었다. 원도는 아래의 곤(坤)에서 반시계방향으로 밑에서 위로 구(姤)까지 올라가고, 다시 밑의 복(復)에서 건(乾)까지 올라가므로 육십사괘를 완성한다. 방도는 가로 우에서 좌로 상괘는 건태리진손감간곤으로 배열되고, 세로 아래에서 위로 하괘는 건태리진손감간곤으로 배열되므로, 상하괘의 조합이 곤에서 건까지 사각형으로 조합을 이룬다. 가감승제에 맞추어 이 방원도를 해석하면, 역은 5를 중앙으로 홀수와 짝수의 상합으로 (1,6)=수, (2,7)=화, (3,8)=목, (4,9)=금, (5,10)=토로 동서남북 방위를 도출한다. 일 월 화 수 목 금 토 일주가 일월오성의 방위를 가리키는 것이라면 복희씨 육십사괘방원도 독법은 라이프니츠의 모나드 철학의 핵심물음이다.

8. 결론

이진논리학의 11 10 01 00, 곤 사상의 이진부호는 가감승제와 사칙연산으로 명제진리계 산표, 전산수학 회로계산, 아스키코드, 유전자 염기서열코드 등의 분야에 논리적 구조를 해명한다. 지금까지는 라이프니츠에서 불에 이르기까지 0과 1의 부호의 상징체계의 논리의 대수화, 이성적 계산의 기계화로 컴퓨터 직렬회로를 규명하는 전산수학의 원리가 생겨났다고 설명하였는데, 음양 -- — 상징체계를 0과 1로 대수화될 수 있다는 것을 최석정이 보인 것은 새로운 사실이다. 그는 방원곡직의 문제에 가감승제의 기본은 천지시간을 재미있게 마름질 하는 관점이자 셈의 문제로 보면서, 독자(讀者)로서 관자(觀者)가 신 학설을 창시하였다고 홀시하지 말라고 한다. 방원 도 아래 곤에서 위의 구까지 음이 자라서 양이 사라지기까지 올라가고, 밑바닥의 복에서 위의 건까지 위로 양이 차오르며 음이 사라지기까지 올라간다. 이는 건이 서북쪽에 있고 곤이 동남쪽에 있다고 말하는 것이며, 이미 생긴 수에 오지 않은 수에 이르기까지 상을 셈하기 때문이다.

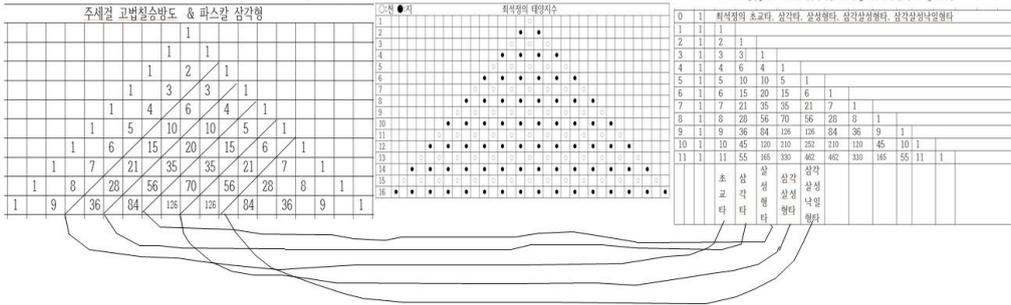
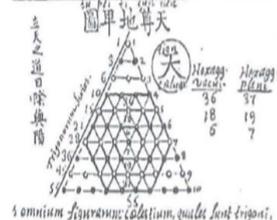
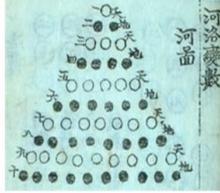
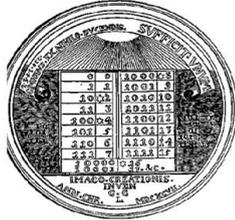
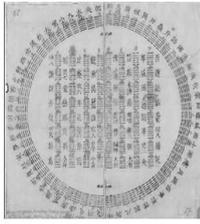
34) 『구수략』 筭有四法 加減乘除 是已 加減二法俗云行筭 加者添而息之 減者損而消之 加爲陽減爲陰 加法以減環原 減法而加還原 乘法而除還原 除法而乘還原.

복희씨육십사괘방
원도

1697년 신년하례연하장

하락변수 천지지도

천존지비도



<그림 1> 구글캡처

참고문헌

Hong-Yeop Song(2008), “Choi’s orthogonal Latin Squares is at least 67 years earlier than Euler’s”, in: *2008 Global KMS International Conference*, 2008.10.23.-10.25, Jesu ICC, Korea.

Leibniz, G. W.(1999), *Sämtliche Schriften und Briefe. Reihe 6 Philosophische Schriften, Bd. 4, Teil A*, N.56, N.57, N.58, N.59, N.60, N.61, N.63, N.64. Akademie Verlag, Berlin.

Leibniz, G. W.(2006), *Der Briefwechsel mit den Jesuiten in China(1689-1714)*, hrsg. v. R. Widmaier, uebersetzt v. M.-L. Babin, Hamburg.

Loly, P. D.(2002), “A Logical Way of Ordering the Hexagrams of the Yijing and the Trigrams of the bagua”. *The Journal of Yijing Studies*, vol. 2, No. 12, 2-13.

Sophie Ling-chia Wei(2015), *Trans-Textual Dialogue in the Jesuit Missionary Intra-Lingual Translation of the Yijing*, Dissertation of University of Pennsylvania Scholarly Commons.

Wittgenstein, L.(1969), *Schriften 1*, Frankfurt am Main.

韓國科學技術史大系 數學編

고희민(2011), 『소강절의 선천역학』, 광신환 옮김, 예문서원.

박우석외(2015), 『논리학의 역사 1, 2』, 한길사.

신익철(2013), 『연행사와 북경천주당』, 보고사.

정해남(2012), 「최석정의 『구수략』에 대한 연구」, in: *History and Pedagogy of Mathematics*.

조영복(1998), 『燕行日錄』, 경기도박물관 학술총서.

조희영(2017), 「『九數略』, 도서상수학으로 짜인 조선산학서」, 『정신문화연구』40권.

최석정(2006), 『구수략』, 조선시대 산학총서, 정해남, 허민 역, 교우사.

<https://zh.wikisource.org/wiki/%E6%B8%85%E5%8F%B2%E7%A8%BF/%E5%8D%B7506>

<https://www.eee-learning.com/article/5061>

<https://ctext.org/book-of-changes/xi-ci-shang/zh>

퇴계의 향약 분석을 통한 조선시대 자치조직 연구

김상현
경북대학교

1. 서론

문재인 정부의 100대 국정과제 중 여섯 번째 국정과제를 살펴보면 ‘국민 인권을 우선하는 민주주의의 회복과 강화’로 명시되어 있다. 이를 위해 정부는 시민사회발전기본법을 제정하고, 시민사회발전위원회를 설치하며, 민주시민 교육체계를 수립함으로써 시민사회가 성장할 기반을 마련하고, 자원봉사 및 기부문화를 활성화 시켜 시민사회의 인적, 물적인 토대 기반을 안정시키려는 노력을 하고 있다. 이와 동시에 지방정부의 자치권을 강화함으로써 수도와 지방이 상생할 수 있는 지방분권형 공화국을 건설하기 위해 ‘지방분권형개헌’도 추진하고 있다.

이러한 시대적 분위기 속에서 현재 우리사회의 바람직한 지역 자치조직의 모습을 그려볼 필요성이 제기된다. 그동안 한국에서도 시민의 자치조직과 관련 연구는 상당히 진척되었다. 하지만 대부분은 헤겔을 비롯한 토크빌, 하버마스 등의 서양사회의 자치조직 이론을 연구하는데 치중되었을 뿐, 우리 전통내부에서 존재했던 자치조직의 모습과 그 이론에 대해서는 그다지 주목하지 못했다. 서양사회에 존재했던 자치조직과 그 조직에 담긴 이론을 분석하는 것 역시 매우 중요하다고 생각한다. 하지만 그것보다는 과거 우리사회에 자치조직이 존재했다면 그것에 대한 연구가 우선되어야 할 것이다. 이러한 문제의식에서 본 논문은 우리 전통 내부 존재해왔던 자치조직의 흔적을 찾아보고 그것에 대한 정당한 평가를 통해 우리사회에 적합한 자치조직의 모습을 그려보고자 한다. 이와 같은 목표를 이루기 위해 우리는 조선향약의 기원을 열었다고 평가받는 퇴계의 향약에 주목할 것이다. 좀 더 구체적으로 말하면, 퇴계의 「향입약조」의 분석을 통해 조선시대 향촌자치조직의 규약인 향약의 모습이 어떠한지, 그리고 그 속에서 얻을 수 있는 가치가 무엇인지를 살펴봄으로써 현재 한국사회에 적합한 자치조직의 방향을 제시해보고자 한다.

그런데 이와 같은 목표를 이루기 위해서는 다음과 같은 문제가 해결되어야 한다. 현재까지 퇴계의 「향입약조」가 향약인지에 대한 논란이 있다. 그렇기 때문에 먼저 퇴계의 「향입약조」를 향약으로 볼 수 있는지를 규명해야 한다. 다음으로 그것을 향약으로 볼 수 있다고 한다면 조선향약으로서 대표성을 지닐 수 있는지를 살펴보아야 한다. 마지막으로 조선향약을 대표할 만한 지위를 가진 것으로 볼 수 있다면 과연 그 속에서 긍정적인 요소를 발견할 수 있는지, 그리고 그것이 바람직한 자치조직을 건설하는데 일조할 수 있는지 등을 확인할 수 있어야 한다.

이와 같은 문제 해소하기 위해, 본 연구는 먼저 향약 개념의 모호성이 어디에서 비롯된 것인지에 대해 분석할 것이다. 이러한 분석을 통해 퇴계의 「향입약조」를 향약으로 보아도 무방하다는 점을 밝힐 것이다. 다음으로 향약의 전형으로 평가받는, 그래서 조선사회에도 많은 영향을 미쳤던 주자의 『증손여씨향약』과 퇴계의 향약을 비교하여 분석할 것이다. 이 과정에서 퇴계의 향약은 주자의 향약과 다른 특징을 지니고 있으며, 그러한 특징은 조선사회가 가진 독특한 역사에서 나온 것이기 때문에 퇴계 향약이 조선향약의 기원을 열었다고 보아도 무방하다는 점을 밝힐 것이다. 마지막으로 조선향약의 기원을 열었다고 평가되는 퇴계 향약에 담긴 가치를 규명함으로써 조선향약의 의의를 밝히고, 오늘날 바람직한 자치조직을 수립하는데 일조할 수 있는 퇴계 향약의 긍정적 가치를 소개할 것이다.

2. 퇴계 「향입약조」의 정체성

1517년(중종 12년)에 의정부가 중종에게 여씨향약과 관련하여 다음과 같은 말을 한다. “여씨향약이 비록 『소학』에 수록되어 있지만 왕께서 밝은 가르침을 내리시어 특별히 실시하지 않는다면 백성들이 대수롭지 않게 보아 걸치레가 될 것입니다. 각도의 감사(監司)에 명령하여 널리 반포하게 하시는 것이 어떠하겠습니까?”¹⁾ 이와 같은 의정부에 요청을 중종은 허락한다. 이 기록에 따르면 여씨향약은 『소학』을 통해 우리나라에 비교적 일찍부터 소개되었던 것으로 보인다.²⁾ 그리고 그 후 『주자대전』이 들어오면서 주자에 의해 수정·보완된 『증손여씨향약』이 소개되었다.³⁾ 그런데 ‘향약’(鄉約)이라는 용어가 『조선왕조실록』에 처음 등장

1) 『중종실록』 28권, 중종 12년 7월 26일 기사. 이하 『조선왕조실록』의 글은 국사편찬위원회의 번역을 인용한다. 다만 가독성을 위해 필요한 경우 문장 표현은 수정하기도 하였다.

2) 그런데 『소학』에 수록된 내용은 다음과 같다. “藍田呂氏鄉約曰, 凡訂約者, 德業相勸, 過失相規, 禮俗相交, 患難相恤. 有善則書于籍, 有過若違約者, 亦書之, 三犯而行罰, 不悛者絕之.” 따라서 『소학』을 통해 소개되었던 향약은 4대 강목과 향약조직의 간략한 운영에 대한 것뿐이다.

3) 여씨 형제들의 『여씨향약』 자체가 우리나라에 소개된 것이 아니라 주자의 『증손여씨향약』이 소개된 것이다. 윤인숙, 「조선전기 향약의 구현을 통한 ‘사문화’의 확산-김안국의 인적 네트워크를 중심으로」, 『대동문화연구』81, 2013, 19-20쪽 참고.

한 때는 조선왕조가 개국된지 100년도 더 지난 1517년(중종 12년)이 되어서다. 김안국이 경상도 감사로 재직할 때 『증손여씨향약』을 한글로 번역한 『여씨향약언해본』을 간행하여 경상도 전역에 배포했다는 기사가 처음으로 등장한다. 이후 여씨향약은 2년도 되지 않아 전국적으로 배포되고⁴⁾, 비로소 조선에 ‘~향약’이라는 용어가 사용되기 시작한다. 그럼에도 여전히 조선의 향촌사회에서는 자치조직의 규약을 ‘~향약’으로 부르거나 표기하지 않은 것으로 보인다.⁵⁾

그렇다면 우리나라에 향약이 언제부터 만들어져 실시된 것일까? 중종 이전에는 향약이 없었던 것일까? 중종 12년 이후에 작성된 퇴계의 「향입약조」는 향약으로 볼 수 있을까? 이러한 물음에 답하기 위해서는 먼저 향약이란 개념의 의미부터 파악해야 할 것이다. 개념규정 확정되어야 우리나라에 그러한 개념에 부합하는 규약이 언제부터 존재했는지, 퇴계의 「향입약조」가 그러한 개념에 부합한지를 확인할 수 있기 때문이다.

『한민족대백과사전』⁶⁾에 따르면 향약은 다음과 같이 규정되어 있다. “넓은 의미로 향촌규약, 향규(鄕規), 일향약속(一鄕約束), 향약계(鄕約契), 향안(鄕案), 동약(洞約), 동계(洞契), 동안(洞安), 족계(族契), 약속조목(約束條目) 등의 다양한 의미를 가진다.” 그러면서 향약의 특징에 대해 “원칙적으로 향약은 조선시대 양반들의 향촌자치와 이를 통해 하층민을 통제하기 위한 것이었지만, 다른 한편으로는 송유배불정책에 의하여 유교적 예절과 풍속을 향촌사회에 보급하여 도덕적 질서를 확립하고 미풍양속을 진작시키며 각종 재난(災難)을 당했을 때 상부상조하기 위한 규약이라고 할 수 있다.”라고 설명한다. 요약하면 향약의 명칭과 의미는 다양하지만, 조선시대 향촌지역 세력가, 즉 재지품관(在地品官)⁷⁾들이 향촌사회를 자치적으로 통제하기 위해 작성한 규약으로 이것을 통해 ‘화민성속’(化民成俗)을 추구했다는 것이다.

그런데 향촌지역 세력가인 재지품관은 조선초기부터 이미 존재했었다. 그렇다면 이들은 향촌사회에서 자신들의 세력을 유지하기 위해 자치조직을 결성하여 향촌사회를 통제하고자

4) 중종대 『증손여씨향약』이 이렇게 짧은 시간에 전국적으로 확산될 수 있었던 이유는 바로 성리학적 이념을 향촌사회에 실현하고자 한 사람 출신의 감사가 주도적으로 이 일을 담당했기 때문이다. 김안국은 1517년에는 경상도 감사를, 1519년에는 전라도 감사를 지냈고, 그의 동생 김정국은 1518년 황해도 감사를 지냈고, 유운은 1518년 충청도 감사를 지냈다.

5) 『영남향약자료집성』(오세창 외 편저, 영남대학교 출판부, 1986.)을 참고하면, 김안국이 『증손여씨향약』을 선도적으로 배포했던 영남지역에서도 ‘~향약’ 혹은 ‘향약~’로 표기된 명칭은 거의 보이지 않는다. ‘~契約’, ‘~洞約’, ‘~約條’, ‘~鄕約契’, ‘~洞憲’, ‘~約法’, ‘~鄕(洞)規’, ‘~契約’ 등 매우 다양하다.

6) 『한민족대백과사전』, 정신문화연구원, <https://100.daum.net/encyclopedia/view/14XXE0062948>

7) 고려말 첨설직(添設職), 동정직(同正職) 등 산직(散職)을 제수받은 자들과 중앙에서 관직 생활을 하다가 왕조 교체 및 기타 사정으로 인해 실직(實職)에서 낙향한 자들을 통칭하여 ‘지방에 있는 품관’이라는 뜻에서 ‘재지품관’ 혹은 ‘유향품관’이라고 한다. 이들은 16세기 중반기를 넘어서면서 서서히 성리학적 지식을 무기로 ‘사족’이라는 특수한 신분층을 만들면서 ‘재지사족’(在地士族)으로 거듭난다. 더 자세한 내용은 다음 논문을 참고하기 바란다. 정재훈, 「조선중기 사족의 위상」, 『조선시대사학보』 73, 조선시대사학회, 2015.

했을 것이고, 이들은 대체로 유교 지식인이라는 점에서 향촌사회의 ‘화민성속’을 추구했을 것이다. 그렇다면 주자의 향약이 조선에 유입되기 이전부터 조선사회에는 ‘화민성속’을 추구했던 자치조직의 규약이 있었다고 추정할 수 있다. 그런데 문제는 그러한 것을 향약으로 부를 수 있느냐는 것이다.

이러한 이유로 향약에 대한 선행 연구자들의 개념 규정 역시 다양하다.⁸⁾ 일제강점기에 선도적으로 향약을 연구했던 다카하시 토오루는 조선을 건국한 태조가 1398년(태조 7년)에 자신의 고향인 풍패향⁹⁾에 향헌(鄉憲) 41조를 제정한 후, 효령대군에게 그것을 증보(增補)하여 선목(善目) 21개조와 악목(惡目) 35개조로 구성된 향헌 51조를 반포하도록 하였는데, 이것을 조선향약의 시초라고 주장했다.¹⁰⁾ 하지만 유홍렬은 다카하시 토오루의 이러한 주장에 동의하지 않는다. 그는 다카하시 토오루가 향약의 시초로 보았던 것을 ‘향헌’(鄉憲)으로 규정한 후, 그것을 향약이 아닌 유향소의 규정 정도로 파악했다. 그리고 향약은 16세기 초(중종대)에 송대 향약이 도입된 이후 명종대를 거쳐서 선조대로 진입하면서 한국적 특성을 분명히 하는 조선향약이 성립되었다고 한다.¹¹⁾ 오늘날 상당수의 연구자들은 유홍렬의 주장에 동조한다. 이후 김용덕은 유홍렬의 주장을 밀고 나가 향규(鄉規)와 향약을 서로 다른 것으로 구분하여 정의한다. 향규는 향안에 수록된 향원(鄉員)들 간의 규칙으로 유향소의 임원 구성과 선출 방식 및 절차, 향촌 계(契)의 영수인 향선생(鄉先生; 鄉大夫, 鄉憲), 그를 보좌하는 향유사(鄉有司)의 선임절차와 그 업무, 향안 가입절차 등과 관련된 규약으로 보고, 향약은 주자의 향약처럼 ‘화민성속’의 교화를 목적으로 한 것으로 보았다.¹²⁾ 박익환은 김용덕의 연구를 더 밀고 나가

8) “향약은 기본적으로 유자들이 경제사상을 실천적으로 펼치기 위한 전통사회의 민간자치조직으로 향촌사회 교화를 목적으로 만들어진 것이다.”(장동표, 「조선시기와 명청대의 향약 시행과 그 성격 비교 연구」, 『한국민족과 문화』 58, 부산대학교 한국민족문화연구소, 2016, 238쪽.) “향약의 종류는 다양하여 주도세력 혹은 참여층에 따라 사족향약, 수령향약, 상하향약 등으로 나누기도 한다.”(김무진, 「조선시대 향약 및 향안조직에서의 의사결정구조」, 『한국학논집』 32, 계명대학교 한국학연구원, 2005, 84쪽.) “북송의 여씨형제가 남천현에서 시행하기 위해 처음 제정한 향약은 남송의 주자에 의해 증손되었다. 조선시대 이래 자치규범으로 널리 시행된 향약은 바로 주자의 『주자증손여씨향약』이다.”(이광우, 「조선시대 정부의 향약 시행 논의와 그 성격」, 『동아인문학』 39, 2017, 140쪽.) “향약은 명종대 후반을 거쳐 선조대에 진입하면서 확산 정도도 심화될뿐더러 한국적 특성을 분명히 하기에 이른다. 이른바 ‘조선향약(朝鮮鄉約)’이 성립되기에 이른 것이다.”(유홍렬, 「조선향약의 성립」, 『진단학보』 9, 1933.)

9) 태조는 즉위 2년 9월에 자신의 출생지인 화령부(和寧府)를 영흥부(永興府)로 고친 후, 다시 한나라 고조의 출생지 풍패현을 본떠 그것을 풍패향(豐沛鄉)이라고 명명하여 자신의 출생지를 미화했다. 그리고 그곳의 경제소(京在所)는 왕족들이 책임을 맡아서 담당했다고 한다.

10) 지교현, 최문성, 박균섭, 『조선조향약연구』, 민속원, 1991, 199쪽.

11) 유홍렬은 ‘향헌’은 고려말 조선초 향촌자치 조직으로 구성된 유향소와 결합함으로써 빠르게 전파되었다고 보고 있다. 유홍렬, 「조선향약의 성립」, 『진단학보』 9, 진단학회, 1938, 96쪽 및 138쪽 참고.

12) 김용덕, 「향규연구」, 『한국사연구』 54, 한국사연구회, 1986, 29쪽 참고. 김용덕은 앞서 유홍렬이 향헌으로 보았던 태조와 효령대군의 향헌을 향규로 보았고, 퇴계의 ‘향입약조’와 율곡의 ‘해주일향약속’도 모두 향규로 보았다. 그런데 태조와 효령대군의 향헌, 퇴계의 향입약조 등을 비롯해 김용덕이 향규로 보고 있는 많은

서 향규와 향약을 구분한 후, 중간 명칭으로 향규약(鄕規約)을 제시한다. 그리고는 앞서 김용덕이 향규로 규정한 것의 더 적절한 명칭은 향규가 아니라 향규약이라고 주장한다.¹³⁾

향약에 대한 개념규정이 이처럼 다양하다는 사실을 통해 우리는 퇴계의 「향입약조」가 무엇이나에 대해서도 다양한 이견이 있을 것이라고 추론할 수 있다. 유흥렬은 퇴계의 「향입약조」와 율곡의 「서원향약」 등이 중국향약과 다른 조선향약이며, 이전 태조 및 효령대군의 향현은 조선향약을 출현시키는 토대 역할을 했다고 주장한다. 이러한 주장에 대해 김용덕은 유흥렬이 향현으로 보았던 것을 큰 틀에서 향규에 포함시킨 후, 퇴계의 「향입약조」와 율곡의 「해주일향학속」 등을 향약이 아니라 향규에 포함시킨다. 이러한 주장에 대해 박익환은 퇴계의 「향입약조」와 율곡의 「해주일향학속」 등을 향규로 부르기보다는 향규약으로 부르는 것이 더 적절하다고 말한다. 이에 반해 다카하시 토오루는 향약이라는 개념을 넓게 사용하여 유흥렬, 김용덕, 박익환 등이 향약과 구분했던 향현, 향규, 향규약 등을 모두 향약의 개념 속에 포함시키고, 퇴계의 「향입약조」를 향약으로 본다. 이렇듯이 오늘날 많은 학자들은 퇴계의 「향입약조」를 다양하게 규정하고 있다. 왜 이러한 결과가 생긴 것일까? 이 물음에 대한 답을 찾기 전에 먼저, 당시 학자들은 퇴계의 「향입약조」를 어떻게 보았는지에 살펴보자.

김성일의 제자인 이민환(1573~1649)은 「제향규후(題鄕規後)」에서 “오로지 퇴계 선생의 향중약조(鄕中約條)에 의거했다”¹⁴⁾라고 하여 ‘향중약조’로 보았고, 정칙(1601-1663)은 「동약서(洞約序)」에서 “퇴도 선생의 향현(鄕憲)에 따라 한 동네에서 행할만한 것 16조목을 선택했다.”¹⁵⁾라고 하여 ‘향현’이라 불렀다. 김용(1557-1620)은 「동중입약서(洞中立約序)」에서 “우리 퇴계 선생께서 시대적 상황을 참고하여 향규(鄕規) 약간 조목을 제정하셨다.”¹⁶⁾라고 하여 ‘향규’라 보았다. 그리고 신원록(1516~1576)은 「서향약후(書鄕約後)」에서 “[퇴계] 선생께서 손수 향약(鄕約)을 제정하셨다.”¹⁷⁾라고 하여 ‘향약’이라 불렀다. 이렇듯이 16세기에 서 17세기에 활동했던 학자들 역시 퇴계의 「향입약조」를 다양한 명칭으로 불렀다는 것을 확인할 수 있다.

그러나 18세기로 넘어가면서 정조는 「일득록」에서 “이 문순공의 예안향약(禮安鄕約)은 간략하면서도 질박하고, 이 문성공의 석담 향약(石潭鄕約)은 번다하면서도 섬세하다.”¹⁸⁾라

약조에 그가 향규의 특징으로 보았던 ‘유향소의 조직, 향원가입 방법, 향임의 선출 절차 등과 관련된 내용이 전혀 보이지 않는다.

13) 박익환, 「조선전기 향규와 향규약고」, 『사학연구』 38, 한국사학회, 1984, 각주 74번 참고.

14) 『紫巖集』 卷3 「題鄕規後」: “專依退溪先生鄕中約條.”

15) 『愚川集』, 「洞約序」: “就退陶先生鄕憲, 擇其可行於一洞者十六條.”

16) 『雲川集』, 「洞中立約序」: “我退溪先生酌其時宜, 定爲鄕規若干條.”

17) 『梅堂集』 卷2, 「書鄕約後」: “先生手定鄕約.”

18) 『弘齋全書』 卷169, 「日得錄」.

고 하여 퇴계의 향약을 아예 ‘예안향약’으로 부르고, 안정복도 이리동약 서문에서 “이것은 남전여씨(藍田呂氏)의 향약(鄉約)과 같은 것이다. 우리나라 선배 학자들 가운데 수령직을 맡아 향촌에 거주하신 분들도 모두 의심 없이 향약을 실행하였다. 예를 들어 일두(一蠹)는 안음(安陰)에서, 퇴계는 예안(禮安)에서, 울곡(栗谷)은 석담(石潭)에서 이것을 실행했던 것이다.”¹⁹⁾라고 하여 주저 없이 향약이라 부른다. 이러한 두 사례를 통해 곧바로 일반화할 수는 없겠지만, 이전에 다양한 명칭으로 불리던 것이 18세기로 넘어오면서 퇴계의 「향입약조」를 향약이라 불렀다는 사실을 확인할 수 있다.

그런데 퇴계의 「향입약조」는 울곡의 「해주향약」과 형식적인 측면에 매우 상이하다. 뒤에서 자세히 살펴보겠지만, 퇴계의 「향입약조」는 별목과 처벌규정을 중심으로 구성된 반면 울곡의 「해주향약」은 ‘입약범례’, ‘증손여씨향약문’, ‘회집독약법’으로 구성되어 있다. 그럼에도 불구하고 정조와 안정복은 두 사람이 제정한 향촌 자치규범을 동일한 부류로 보았던 것이다. 그것은 바로 ‘화민성속’을 목적으로 하는 향촌사회의 자치규약이라는 공통점 때문으로 보인다. 조선 후기로 가면서 조선초기와 중기에 보였던 많은 명칭들이 사라지고 점차 주자가 사용했던 명칭인 ‘향약’으로 수렴된 것은 그것들 속에 모두 ‘화민성속’을 목적으로 하는 향촌사회의 자치규약이라는 의미가 담겨 있었기 때문으로 볼 수 있다.

그렇다면 앞서 보았던 혼란은 어떻게 해서 생긴 것일까? 그 원인은 크게 두 가지 측면에서 살펴볼 수 있을 것이다. 하나는 향약이라는 용어의 애매성으로 인한 차이 때문으로 볼 수 있고, 다른 하나는 향약이라는 개념을 어떻게 규정하느냐의 차이 때문으로 볼 수 있다. ‘향약’이라는 표현에는 여씨향약 혹은 주자의 『증손여씨향약』을 가리키는 고유명사적 용례와 ‘향촌사회의 자치 규약’ 일반을 가리키는 보통명사적 용례가 혼재되어 있었던 것이다. 실록을 보면 실제 조선시대에 향약이라는 표현에 두 용례가 혼재되고 있음을 확인할 수 있다. 여기에 더해 당시 실재했던 ‘향촌사회의 규약’을 앞서 살펴보았듯이 당시 사람들이 향약 이외의 다양한 명칭으로 불렀던 원인도 있다. 이렇게 다양하게 불리던 것이 18세기 이후 점차 개념화 과정을 거치면서 대표적인 용어인 ‘향약’으로 부르게 되었던 것이다.

후대 학자들이 향약에 대한 각자 상이하게 개념규정한 것도 이러한 혼란을 낳는데 일조했다. 향약이라는 개념을 넓은 의미로 규정할 것이냐, 좁은 의미로 규정할 것이냐에 따라 어떤 것이 향약이 되기도 하고 되지 않기도 하기 때문이다. 다카하시 토오루의 경우는 향약이라는 개념을 넓게 보아 ‘화민성속’을 목적으로 하는 향촌사회의 자치규약이라면 향약이라 부를 수 있다고 보았고, 다른 학자들은 의도와 목적, 운영주체와 적용대상, 규약의 내용, 수령과의

19) 『順菴集』 권18, 「慶安二里洞約序」.

관계 등의 차이를 근거로 미세하게 구분했던 것이다. 향약이라는 개념을 미시적으로 규정하는 것 역시 의미 있는 작업이겠지만, 이러한 작업은 자칫 연구자의 자의성이 개입될 가능성이 높고, 개념과 그 개념이 가리키는 실제 대상과의 부조화를 낳을 가능성이 높다고 판단된다. 다시 말해, 개념 규정은 머릿속에서 이루어지는 것이지만, 그 개념이 지시하는 대상은 현실 속에 존재하는 것이기 때문에 개념의 분류에 맞게 대상을 분류하는 것이 매우 어려울 수 있다는 것이다. 개념이 실재와 짝을 이루지 못하면 그 개념은 실재와 유리되어 신화가 되어 버린다. 논자는 조선시대 향촌자치조직의 규약을 미시적으로 규정하는 것 역시 이와 같은 위험에 노출되어 있다고 생각한다. 이러한 이유에서 향헌, 향규, 향규약 등을 모두 향약의 일종으로 보는 것이 더 적절하다고 판단하며, 아울러 퇴계의 「향입약조」 역시 향약으로 보아야 한다고 판단한다.

이하에서는 향약의 전형으로 평가받는, 그래서 조선사회에도 많은 영향을 미쳤던 주자의 『증손여씨향약』과 퇴계의 향약을 비교 분석할 것이다. 이러한 분석을 통해 퇴계의 향약이 중국의 주자 향약과 다른 특징을 지니고 있고, 그런 점에서 충분히 조선향약의 기원을 열었다고 평가할 수 있다는 점을 밝힐 것이다.

3. 주자 향약과 퇴계 향약의 비교 분석

3.1 등장 배경과 목적

잘 알려져 있듯이, 여씨향약은 북송시대 남전현에 살았던 여씨 4형제들 가운데 여대균이 중심이 되어 만든 향촌자치규약이다. 하지만 사서(四書)와 마찬가지로 주자의 눈을 통해 세상이 드러나 주목받게 된다.²⁰⁾ 이러한 사정은 “향약이라는 규범은 송나라 유학자 여대균에서 시작되었지만, 주자가 그 글을 토대로 가감하여 새롭게 만들어내니 참으로 만세에 통행하는 훌륭한 법이 되었습니다. 여씨가 작성한 전문은 오늘날에 전해지지 않지만 주자가 가감해서 작성한 『증손여씨향약』은 『주자대전』에 실려 있습니다.”²¹⁾라는 실록의 기록을 통해서도 확인할 수 있다.

그렇다면 주자는 무엇 때문에 여씨의 향약을 수정, 보완하여 새롭게 작성했던 것일까? 이

20) 이 말은 여씨향약이 북송시대에는 그다지 큰 호응을 불러일으키지 못하였다는 것이다. 이렇게 보는 이유는 여씨 형제들에 의해 편찬된 「여씨향약」이 세상에 주목을 받기 시작한 것은 바로 주자가 『증손여씨향약』을 세상에 내어놓으면서부터였다는 점과, 여씨향약이 세상에 나온 이후 북송시대로부터 남송 초년에 이르기까지 향약이 실시된 사례도 거의 발견되지 않았다는 점 때문이다. 이근명, 「주희의 『증보여씨향약』과 조선사회」, 『중국학보』 제45집, 2002, 277쪽 참고.

21) 『정조실록』 정조 21년 6월 2일 기사.

물음에 대한 답은 첫째 주자가 『대학』의 삼강령 중 하나인 ‘친민’(親民)을 ‘신민’(新民)으로 해석했다는 점, 둘째 주자향약이 ‘덕업상권’, ‘과실상규’, ‘예속상교’, ‘환난상휼’이라는 4대 강목을 중심으로 구성되었다는 점을 통해 찾을 수 있다.

‘친민’은 민과의 친밀감을 강조한 표현인 반면, ‘신민’은 민의 교화를 강조한 표현이다. 그러므로 주자가 『대학』의 두 번째 강령을 ‘신민’으로 해석했다는 사실을 통해 그가 성리학적 지식인을 교화를 담당하는 주체로 보고, 민을 교화의 대상으로 간주했다는 사실을 읽을 수 있다. 이것은 주자가 유교의 핵심 이념인 치인(治人)을 ‘화민성속’으로 보았고, 그것을 향촌 사회 자치 규범인 향약에 담고자 했다는 점을 시사한다. 이러한 사실은 주자가 향약을 4대 강목이라는 덕목을 중심으로 구성하였다는 점, 그리고 향약 말미에 부록 형식으로 향회시의 의례를 첨부하고 있다는 점을 보면 더 분명해진다. 주자는 여씨향약에는 없었던 향회시의 의례를 첨가함으로써 향약의 실천을 독려하고자 했다.²²⁾ 결국 주자가 향약을 제정한 핵심 목적은 바로 ‘화민성속’에 있었다고 하겠다.

그렇다면 조선에서 향약이라는 규범이 작성되고 실시되었던 배경과 목적은 무엇일까? 이 물음에 대한 답을 찾기 위해서는 조선 초기 향촌에서 자체적으로 조직되었던 유향소라는 조직에 대한 검토가 필요하다. 왜냐하면 유향소²³⁾라는 기관을 운영했던 주체들이 만들었던 향촌자치의 규약인 향규(鄕規) 역시 향약으로 볼 수 있고, 이러한 자치조직의 규약은 주자의 향약이 본격적으로 도입되기 이전부터 조선사회에 존재하고 있었기 때문이다.

유향소라는 기관만큼 조선사회에서 혁파와 부활이 반복되었던 기관도 없었을 것이다. 유향소는 조선이 건국된지 불과 14년 뒤인 1406년(태종 6년)에 혁파되었다가 1428년(세종 10년)에 부활된다. 그리고 1467년(세조 13년) 세조에 의해 다시 혁파되었다가, 1488년(성종 19년)에 다시 부활된다.²⁴⁾ 유향소는 왜 이렇게 자주 혁파와 부활을 반복했던 것일까? 그리고 그 이유는 무엇일까? 이 물음에 대한 답은 『조선왕조실록』에서 유향소 혁파와 재설치를 둘러싸고 진행되었던 논의를 검토함으로써 확인할 수 있다. 대개 유향소 혁파를 주장한 사람들

22) 이성무, 「여씨향약과 주자중손여씨향약」, 『진단학보』 71, 72, 진단학회, 1991, 69쪽.

23) 유향소가 언제 처음으로 조직되었는지 보여주는 명확한 자료는 없는 것으로 보인다. 하지만 유향령을 비롯한 많은 학자들은 고려말에서 조선으로 넘어오는 시기에 조직되었던 것으로 본다. 유향소는 재지품관들이 자발적으로 조직한 것으로 알려져 있다. 재지품관을 비롯한 향촌사회의 현족들은 그들만의 모임인 계(契)를 조직하였는데 계의 구성원을 ‘향원’(鄕員), 구성원 명부를 ‘향안’(鄕案)이라 한다. 향원 가운데 나이가 많고 덕망이 높은 사람을 선출하여 향대부 또는 향선생이라 불렀고, 이들을 도울 사람으로 향유사를 선발하였는데, 이들을 ‘향집강’(鄕集綱)이라 불렀다. 이들은 향리 규찰과 향풍 교화 등의 업무를 수행할 기관으로 유향소를 설치했던 것이다.(『한민족대백과사전』 참고)

24) 유향소의 혁파와 복설 이유에 대해서는 배기현의 다음 논문을 참조하기 바란다. 배기현, 「16세기 향촌지배 질서와 유향소의 성격」, 『대구사학』 35, 대구사학회, 1988.

이 내세웠던 주요 근거는 유향소가 수령권을 침탈하고 그 결과 향리와 백성들의 고충을 심화시켰다는 것이다.²⁵⁾ 조선초기 향촌사회는 향촌사회에서 대대로 세력을 유지해온 재지사족들의 향촌사회자치의 움직임과 수령을 파견함으로써 중앙정부의 직접적 통치 사이에 긴장감이 흐르고 있었다. 그리고 양자 사이에 주, 부, 군, 현의 행정기관에서 행정실무를 담당했던 향리계층들이 있었다. 이러한 상황에서 재지품관 세력과 수령은 향촌사회의 주도권을 놓고 때로는 협력하기도 했고 때로는 경쟁하기도 했다. 조선초기 중앙정부는 재지품관들이 자치기관으로 설치한 유향소를 묵인했는데, 그것은 수령만으로 향리를 비롯한 향민들의 통제가 힘들었고, 더욱이 향촌사회의 ‘화민성속’을 이룰 수 없었기 때문이다.²⁶⁾ 하지만 국가의 묵인 아래 향촌사회의 지배권을 어느 정도 확보한 유향소가 공권력을 침해하고 그 결과로 발생한 폐단이 향리와 백성들에 전가되었기 때문에 얼마 가지 않아 혁파되었던 것이다. 이러한 사실은 유향소 운영 주체들이 건전하다면 사실 유향소는 당시 조선에서 필요한 기관이란 말이다.²⁷⁾ 결국 중앙정부가 유향소를 부활시킨 주된 이유는 주자가 향약 실시를 통해 목적으로 삼았던 향촌사회의 ‘화민성속’을 이루는 것에 더해 향촌사회 통치에 도움을 받고자 하는데 있었던 것이다. 이러한 사실을 고려한다면, 주자 향약이 본격적으로 도입되기 이전에 이미 존재했던 향촌자치조직의 규약인 향규에는 화민성속의 목적과 함께 향촌사회 자치의 목적도 포함되었다고 할 수 있다.

그런데 퇴계의 「향입약조」(이하 퇴계향약으로 표기) 역시 유향소와 밀접하게 관련된 규약이다. 이것은 퇴계의 다음과 같은 말을 통해서 확인할 수 있다. “오늘날의 유향소[留鄕]는 바로 옛날 향대부의 뜻이 담긴 제도이다. (...) 효, 제, 충, 신의 도덕이 막혀서 실행되지 못하게 된다면 예의(禮義)를 버리고 염치가 없어지는 것이 날로 심해지게 될 것이다. 그렇게 되면 결국 점점 오랑개(夷狄)나 금수(禽獸)의 세계로 돌아갈 것이다. 이것은 정말 국가정치(王政)의 큰 걱정인데, 이러한 것을 꾸짖고 바로잡는 책임이 유향소[鄕所]로 돌아오니, 아아! 그 책임 또한 막중하다.”²⁸⁾ 앞서 살펴보았듯이, 재지품관들이 유향소를 설치한 목적에는 ‘화민성속’

25) 『태종실록』 6년 6월 9일 기사; “주, 부, 군, 현에는 각각 수령이 있는데, 향원(鄕憲; 약질 토호)들 가운데 일벌 이기를 좋아하는 무리들이 유향소를 조직하여 수시로 모여서 수령을 헐뜯고 사람과 물건을 함부로 올리거나 내치며, 백성들을 괴롭히는 것이 교활한 향리보다도 더 심합니다. 모두 혁파하시어 오래된 폐단을 없애주시기 바랍니다.”

26) 『성종실록』 19년 3월 2일 기사; “예전에는 1향 가운데에 정직한 품관 한두 명을 택하여 향유사(鄕有司)를 삼아서 풍속을 바로잡게 하고 이름하기를 유향소라고 하였는데, 혁파한 이래로 향풍이 날로 투박하여졌다고 합니다.”; 같은 책 5월 을해: “중앙의 경재소와 지방의 유향소가 서로 들은 대로 규찰(糾察)하여 간활한 향리를 억제시키고 시골의 풍속을 유지시킨다면 풍속을 좋은 방향으로 개선하는 계기에 있어서 도움이 될 것으로 여깁니다.”

27) 이러한 이유로 시간이 지나면서 유향소는 점차 관의 하부기관으로 들어가게 된다.

28) 『퇴계집』 권42, 「향입약조서」.

의 향풍교정과 함께 향촌사회의 자치도 들어있었다. 그러므로 퇴계의 향약 역시 일차적 목적은 주자와 동일하게 ‘화민성속’에 있다고 볼 수 있지만, 퇴계의 향약에는 여기에 더해 ‘향촌사회의 자치’의 목적도 어느 정도 들어있다고 하겠다.

3.2 체제와 구성

이제 두 향약의 체제 및 구성에 대해 검토하고자 한다. 먼저 주자 향약의 구성에 대해 간단히 살펴보자. 주자 향약은 내용상 ‘머리말’에 해당하는 글이 먼저 나온다. 여기서는 향약의 4대 강목으로 알려진 ‘덕업상권’, ‘과실상규’, ‘예속상교’, ‘환난상휼’을 소개한 후, 향약조직의 임원[鄉旻]인 [도]약정 1인, 부[약]정 2인, 직월 1인 선출 방식을 소개한다. 그리고 마지막으로 직월이 관리하는 향약조직의 장부인 ‘선적’ 1책과 ‘악적’ 1책, 그리고 ‘회원명부’ 1책에 대해 기술한다.²⁹⁾ 그리고 이어서 곧바로 본문에 들어간다. 본문 부분에서는 4대 강목을 ‘덕업상권’에서부터 ‘환난상휼’에 이르기까지 차례로 소개한다.

먼저, ‘덕업상권’부분을 보면 다음과 같다.

〈표 1〉 덕업상권

덕(德)	업(業)
<p>선한 일을 보면 반드시 실천하고, 잘못을 지적받으면 반드시 고친다. 몸을 잘 다스리고, 가정을 잘 다스린다. 부모와 형을 잘 섬기고 자제를 잘 가르친다. 종을 잘 다스리고 정사와 교육을 엄숙하게 한다. 어른을 잘 섬기고, 천척 및 오랜 친구들과 화목하게 지낸다. 좋은 사람을 잘 가려서 사귀고, 청렴함과 절개를 잘 지킨다. 널리 은혜를 베풀고, 다른 사람의 부탁을 잘 들어줄 수 있다. 어려운 처지에 있는 사람들을 잘 구제해주고, 사람들을 선할 일로 이끌어 준다. 다른 사람의 잘못을 잘 타이르고, 다른 사람을 위해 일을 모의한다. 공동체의 일을 위해 최선을 다한다. 싸움을 잘 해결하고, 시비를 잘 판단한다. 이로운 일은 일으키고 해로운 일은 잘 제거한다. 관직에 있을 때에는 그 직무를 잘 수행한다.</p>	<p>집에서는 부모와 형을 섬기며, 아우와 자식을 가르치며, 아내와 첩을 예로서 대한다. 밖에서는 어른과 높은 사람을 섬기며, 친구와 사귀고, 젊은 사람들을 가르치며, 노비를 통제한다. 글을 읽고 밭을 관리하며, 집안을 경영하고 물을 구제한다. 법령을 두려워하며 조세와 공물을 바치고, 예약 사어서수의 류를 좋아한다.</p>

29) 선적, 악적, 회원명부라는 명칭은 논자가 편의상 그렇게 표기한 것이다.

주자는 ‘덕업상권’ 부분을 ‘덕’에 해당하는 것과 ‘업’에 해당하는 것을 구분하고 있다. 덕의 항목이 업의 항목보다 훨씬 많은데, 업의 항목 가운데 상당수는 덕과 중복되기도 한다. 다만 차이는 덕의 항목은 ‘반드시’ 혹은 ‘할 수 있는 능력’을 뜻하는 ‘必’과 ‘能’으로 표현되어 있다는 것이다.

과실상규는 크게 ‘일반적 도덕규범에 어긋난 과실’, ‘향약의 4대 강목을 위반한 과실’, ‘자신을 수양하지 못해 생긴 과실’ 이렇게 3가지로 분류한 후, 다시 아래와 같이 하부 별목을 제시하고 있다.

〈표 2〉 과실상규

‘일반적 도덕규범[義]에 어긋난 과실	‘향약의 4대 강목’을 어긴 과실	‘자신을 수양하지 못한’ 과실
술주정, 노름, 싸움, 송사 등을 자주 벌이는 것 행동이 예의를 무시하며 법을 어기는 것 행동이 공손하거나 겸손하지 않은 것 말이 진실되지 않거나 미덥지 못한 것 말을 만들어 죄 없는 사람을 무고하는 것 너무 심하게 자신의 이익만을 탐하는 것	덕과 업을 서로 권하지 않은 것 잘못을 서로 규제하지 않은 것 예의와 법도의 풍속을 지켜 서로 사귀지 않는 것 어려운 일에 서로 돕지 아니하는 것	어질지 못한 사람을 사귀는 것 쓸데없이 놀면서 덕업 닦기를 게을리 하는 것 예의에 맞지 않게 행동하는 것 일을 조심해서 처리하지 않는 것 재물을 절약에서 사용하지 않는 것

주자 향약은 과실의 별목이 매우 많다고 할 수 있다. 왜냐하면 덕업상권의 ‘덕’목과 ‘업’목도 그 세부 조항이 매우 많은데, 과실의 조항은 덕업상권을 포함해서 향약 4대 강목을 잘 수행하지 못한 것을 모두 망라하고 있기 때문이다. 그리고 여기에 더해 ‘일반적 도덕규범에 어긋난 것’과 ‘자신의 수양부족으로 생긴 것’까지 포함하고 있기 때문이다. 그리고 과실상규 부분에는 과실에 대한 처벌 규정도 소개되어 있는데, 처벌 규정은 크게 세 단계로 이루어진다.³⁰⁾ 첫 번째 단계는 주의 조치이다. 작은 과실은 조용히 처리하고 큰 과실은 여럿이 처리한다는 차이만 있을 뿐이다. 두 번째 단계는 첫 번째 단계에 조치에 대해 승복하지 않을 때에만 진행한다. 이 단계는 집회 때 약정이 주관하는데, 사과하고 고치려는 의향이 있으면 악적에 기록한 후 고치기를 기다린다. 하지만 두 번째 조치에 대해서도 승복하지 않는다면 마지막

30) 『朱熹集』 卷74, 「增損呂氏鄉約」; “동약인이 과실을 스스로 성찰하고 서로 규제하며 과실이 작을 때는 조용히 주의 주고(密規) 클 때는 여러 사람이 나서서 질책한다. 주의나 질책을 받아들이지 않을 때에는 집회일에 직월이 약정에게 알린 후, 약정이 의리로써 가르치고 깨우쳐준다. 사과하면서 고치려고 한다면 악적에 기록한 후 기다린다. 변명만 늘어놓은 채 승복하지 않아 끝내 개선의 여지가 없으면 향약에서 제명한다.”

세 번째 단계로 들어간다. 최종 단계인 세 번째 단계의 처벌은 향약조직에서 퇴출시키는 출약(出約)이다.

예속상고부분에서는 크게 네 가지를 제시하고 있다. 첫째는 나이 기준으로 향원들의 서열을 5등급의 구분하는 예이고, 둘째는 향원들 간에 절과 읍하는 예이고, 셋째는 손님을 청하고 맞이하는데 필요한 예이고, 넷째는 축하와 조문하는데 필요한 예를 밝히고 있다. 환난상훈부분에서는 먼저 수재와 화재, 도둑맞음, 질병, 초상, 고아, 무고, 가난과 같은 7가지 어려운 상황에 대해 밝힌 후, 각각 경우 어떻게 도와주어야 하는지를 밝히고 있다. 그리고 이어서 환난상훈과 관련된 행동지침과 그러한 행동지침을 어겼을 때 가하는 처벌 등에 대해 밝히고 있다. 그리고 마지막으로 향약의 집회와 관련된 복잡한 의식을 ‘월단집회독약지례’라는 표제 아래 소개하고 있다.

주자의 향약에 비하면 퇴계의 향약은 그 구조가 매우 간결하다. 머리말에 해당하는 「서문(鄉立約條序)」과 본문에 해당하는 「約條」로만 구성되어 있다. 앞서 살펴보았듯이 「향입약조서」에서는 향약의 중요성, 향약을 제정하게 된 동기, 향약 구성체제에 담긴 의도 등에 대해 밝히고 있다. 본문부에 해당되는 「약조(約條)」³¹⁾에는 벌목(罰目)과 처벌 규정을 담고 있다.

31) 오늘날 전하는 퇴계 향약에는 두 가지 본이 있다. 하나는 『퇴계전서』본에 수록된 것이고, 다른 하나는 『社會史研究史料叢書(1)』(김인걸, 한상권 편집, 1986)에 수록된 것이다.(내용이 다름에도 여전히 출처는 “『퇴계전서』 권42”로 표기되어 있다. 오기로 보인다.) 두 본의 차이는 전자에 없는 것이 후자에는 추가되어 있다는 것이다. 추가된 내용은 5가지 정도이다. ①극벌의 처벌 규정에 “上罰, 告官司科罪, 不通水火, 中罰, 削籬, 不齒鄉里, 下罰, 損徒, 不與公會.”가 추가되어 있다.(김기 향약은 극벌 벌목 조항에는 “侵暴小民, 私門用杖者”도 추가되어 있다) ②중벌 벌목 조항에 “多接人戶, 不服官役者.”와 “不謹租賦, 圖免徭役者” 두 조항이 더 추가되어 있으며, ③중벌 처벌 규정에 “鄉中從輕重施罰”이 추가되어 있다(김기 향약의 중벌 처벌 조항에는 “上罰告官司科罪, 中下從輕重施罰.”로 표기되어 있다.). ④하벌 처벌 규정에 “座中或面責施罰”이 추가되어 있다(김기 향약의 하벌 처벌 규정은 “或面責施罰”로 표기되어 있다.), ⑤특별 조항의 처벌 규정에 “已上, 隨聞見摘發, 告官依律科罪.”이 추가되어 있다. 그렇다면 이와 같은 차이를 어떻게 이해할 수 있을까? 그리고 어느 것이 퇴계의 의중을 온전히 반영한 것으로 볼 수 있을까? 「향입약조」를 역주했던 이장우는 『퇴계전서』본에 수록된 것을 토대로 역주했으며, 후자에 대해서는 전혀 언급하지 않았다. 김명진, 황병곤, 박인환 등도 전자를 근거로 연구논문을 발표했다. 그리고 『영남향약자료집성』(1986)에서도 김기 등의 향약과 비교한 퇴계의 향약을 『퇴계전서』본의 것을 사용했다. 반면 정영진(1986)은 임진왜란 이전에 간행된 것으로 보이는 『여씨향약』(출처 불분명?)이라는 책에 수록된 「향입약조서」, 「약조」인 후자의 자료를 더 원형에 가까운 것으로 평가하고, 변영호(2011) 역시 퇴계의 사상을 고려할 때, 전자는 다소 어색하지만 후자는 완성도가 높기 때문에 후자를 원문으로 판단한다.(김명진, 「예안향약에 관한 소고」, 『정법논총』13, 건국대학교, 1978; 황병곤, 「예안향약을 통해서 본 퇴계의 향민자치관」, 『퇴계학보』52, 퇴계학연구원, 1986; 박인환, 「조선전기 향규와 향규약고」, 『사학연구』38, 한국사학회, 1986; 정진영, 「16세기 향촌문제와 재지사족의 대응」, 『민족문화논총』7, 영남대학교 민족문화연구소 1986 ; 변영호, 「이퇴계에게 있어서의 정치-재지사족의 시점으로부터」, 『퇴계학논집』9, 영남퇴계학연구원, 2011.) 논자는 정진영 및 변영호와 다른 근거에서 후자를 퇴계의 의중을 더 온전하게 반영한 자료라고 판단한다. 논자가 이렇게 판단한 근거는 다음과 같다. 첫째, 퇴계의 제자 김기가 “별조는 퇴계선생의 향약을 오로지 사용하였다.”(『북애집』 권3 「잡저, 향약」)라고 하면서 기술한 처벌규정이 후자와 거의 동일하다는 것, 둘째 영남퇴계학의 학맥을 이은 것으로 평가되는 이현일의 동생 이승일의 「鄉約定規」(『恒齋集』 권4)의 과실상규 부분에도 후자와 거의 동일한 자료가 수록되어 있다는 것, 셋째 새롭게 추가된 것 가운데 중벌의 벌목 조항 “多接人戶, 不服官役者.”와 “不謹租賦, 圖免徭役者”는 퇴계의 평상시 지론을 잘 보여주는

별목은 처벌 수위에 따라 ‘극벌’ 조항과 그 처벌 규정, ‘중벌’ 조항과 그 처벌 규정, ‘하벌’ 조항과 그 처벌 규정, 이렇게 세 단계로 나누어지고 있으며, 마지막에 특별 조항으로 향약 외부인에 대한 별목과 처벌 규정을 밝히고 있다. 그리고 특별 조항의 벌칙 규정을 제외하고는 각 단계별로 처벌 규정은 다시 다시 ‘상’, ‘중’, ‘하’로 세분되고 있다.

먼저, 극벌 7조항부터 차례로 살펴보자.

- ① 부모의 뜻을 따르지 않는 자.[불효 죄는 나라에서 정한 형벌이 있기 때문에 그 다음 죄목만 들었음.]
- ② 형제끼리 서로 싸우는 자.[아우에게는 비록 잘못이 없다고 하더라도 형의 잘못으로 인해 형제가 서로 싸우게 되면 둘 모두 동일하게 처벌함. 동생이 잘못하여 싸움이 발생한 경우라면 동생만 처벌함. 두 사람 모두의 잘못으로 인해 싸움이 발생한 경우라면 형은 가볍고 아우는 중하게 처벌함.]
- ③ 집안의 법도를 어지럽히는 자.[부부가 치고받고 싸우는 경우, 정처(正妻)를 쫓아내는 경우(아내가 사납게 거역한 경우는 죄를 감등함), 남녀가 분별이 없이 행동하는 경우, 적처(嫡妾)를 뒤바꾸는 경우, 첩으로 처를 삼은 경우, 서자[孽]로 적자[適]를 삼는 경우, 적자가 서얼을 보살피지 않는 경우, 서얼이 도리어 적자를 능멸하는 경우]
- ④ 관청의 일에 간섭하여 영향력을 행사하고, 고을의 풍속[鄉風]을 어지럽힌 자.
- ⑤ 경거망동 위세를 부려서 관청을 제멋대로 주무르고 자기 마음대로 행하는 자.
- ⑥ 향장(鄉長; 고을의 어른)을 능욕하는 자.
- ⑦ 수절(守節)하는 과부를 유혹하여 더럽힌 자.

극벌 별목 조항을 보면, 퇴계는 규범이 적용되는 영역을 기준으로 윤리규범을 먼저 가족공동체와 관계된 윤리, 지역공동체와 관계된 윤리, 국가공동체와 관계된 윤리로 구분한다. 이렇게 구분한 후, 각 윤리 규범을 위반한 별목을 나열한 것으로 보인다. 예를 들어 ①, ②, ③은 가족공동체와 관계된 윤리 위반 조항이고, ④와 ⑤는 관(官), 즉 국가공동체와 관련된 윤리

것이라는 점, 넷째 퇴계의 「향입약조」는 결과적으로 시행되지 못했는데 그 이유가 “그것을 꺼리는 사람들이 있어서”(『鶴峯先生文集續集』卷之五, 『退溪先生言行錄』)라는 것이다. 퇴계의 향약 가운데 당시 예안의 재지사족들이 꺼릴만 내용으로는 추가된 중벌 조항과 다소 엄격한 처벌조항일 개연성이 높다. 그런데 이 내용은 퇴계의 평소 지론을 잘 보여주는 것으로 없는 것이 다소 이상하다고 여겨질 정도이다. 이러한 점을 고려한다면, 퇴계의 문인들 사이에는 퇴계의 진의가 전달되었지만, 문집을 간행하는 단계에서는 논란이 될 것을 우려하여 고의로 누락해서 이와 같은 결과가 생기게 된 것으로 볼 수 있을 것이다. 이 문제에 대한 정밀한 논의는 다음 기회로 미루겠다.

위반 조항이며, ⑥과 ⑦은 고을, 즉 사회공동체와 관계된 윤리 위반 조항으로 볼 수 있다. 물론 이러한 조항은 퇴계가 서문에서 강조했던 효, 제, 충, 신의 덕목 위반 사례로 수렴될 수 있다. 이런 측면에서 볼 때 퇴계의 향약은 간략하면서도 체계성을 갖춘 것으로 평가할 수 있을 것이다. 다만 특이한 점은 부모, 형, 남편, 정치, 적자를 자식, 동생, 아내, 첩, 서자보다 우선시키고 있다는 것이다.³²⁾ 물론 시대적 한계가 있겠지만, 이것은 오늘날의 관점에서 본다면 다소 아쉬움이 남는 대목이다.

이어서 극벌의 처벌 규정을 다음과 같이 명시한다. “상벌은 관청에 알리고 일체 교류를 하지 않는다[不通水火]. 중벌은 회원 명부에서 이름을 빼고[削籍] 고을 일에 끼워주지 않는다[不齒鄉里]하며, 하벌은 회원자격의 발탁하고[擯徒] 공공의 모임에 함께하지 않는다.” 처벌 규정을 볼 때, 퇴계는 심각한 과실을 범한 경우, 향중 자치적 처벌과 관의 공권력에 의한 처벌을 함께 진행하고 있다는 것을 알 수 있다. 이러한 점은 공권력을 최대한 존중하면서도 동시에 가능한 범위 내에서 향촌조직의 자발적인 처벌도 강조하고 있다는 것이다. 이러한 관점은 주자의 향약과 구별되는 퇴계 향약의 특징이라고 할 수 있다.

중벌 별목 조항은 18개인데, 이 숫자는 극벌과 하벌, 그리고 특별 조항을 모두 합한 것보다도 더 많다. 그러면 먼저, 중벌 별목 조항에서 대해 살펴보자.

- ① 친척들과 화목하게 지내지 않는 자
- ② 본처를 박대하는 자[처에게 죄가 있는 경우는 죄를 감한다.]
- ③ 이웃과 화합하지 않는 자
- ④ 또래끼리 서로 치고받고 싸우는 자
- ⑤ 옆치를 돌보지 않고 선비의 풍속[土風]을 허물고 더럽히는 자
- ⑥ 강(強)함을 믿고 약한 이를 능멸하고 침탈(侵奪)하여 다투는 자
- ⑦ 무뢰배와 모임(당)을 만들어 횡포한 일을 많이 행하는 자
- ⑧ 공적, 사적 모임을 불문하고 관청의 정사[官政]에 대해 시비하는 자
- ⑨ 말을 만들고 거짓으로 사람을 죄에 빠뜨리게 하는 자
- ⑩ 환란(患亂)을 보고 힘이 미치는 데도 가만히 보기만 하고 구하지 않는 자
- ⑪ 관가의 임명을 받고 공무를 빙자하여 폐해를 만드는 자
- ⑫ 혼인(婚姻)과 상제(喪祭)에 아무 이유 없이 시기를 넘기는 자

32) 반면 율곡이 청주목사시절 제정했던 『서원향약』(1571년 36세)의 악목 부분에 보면 “不孝不慈”, “不友不悌”, “夫婦無別”라고 하여 ‘효’와 ‘자’, ‘우’와 ‘제’, ‘유별’의 덕목을 강조하고 있어 부자, 형제, 부부 관계를 퇴계보다 수평적으로 바라보고 있다.

- ⑬ 집강(執綱, 座首)을 업신여기며 유향소의 명령을 따르지 않는 자
- ⑭ 향론에 복종하지 않고 도리어 원망을 품는 자
- ⑮ 집강(執綱)으로서 [향론도 없이] 사사롭게 외부인을 향안(鄉案)에 들인 자
- ⑯ 구관(舊官)을 전송하는데 특별한 일 없이 참석하지 않는 자
- ⑰ 많은 인호(人戶)를 자신의 호(戶)에 예속시켜 놓고 관역을 따르지 않는 자³³⁾
- ⑱ 세금납부(租賦)에 힘쓰지 않고 요역(徭役)을 면하길 도모하는 자

중벌 벌목 조항 역시 극벌 벌목 조항처럼 윤리 규범을 크게 가족공동체와 관계된 윤리, 지역공동체와 관계된 윤리, 국가공동체와 관계된 윤리로 구분한 후, 이들 각 윤리 규범을 위반한 경우를 구체적 사례를 들어 밝히고 있다. ①, ②는 가족공동체와 관계된 윤리 위반 조항이고, ③~⑩, ⑬, ⑭, ⑮는 사회공동체와 관계된 윤리 위반 조항이며, ⑪, ⑫, ⑯, ⑰, ⑱은 국가공동체와 관계된 윤리 위반 조항이다. 극벌이 가족공동체 내의 질서와 국가공동체에 대한 사적 도전, 지역사회에서 인륜을 저버린 일에 대한 벌목 조항이었다면, 중벌은 상대적으로 가족공동체를 벗어나 일반적인 사회 윤리, 유향소 및 국민의식 결여 관계된 벌목이 대부분이다. 중벌 벌목 조항에서 특히 눈에 띄는 것 중 하나는 유향소와 관련된 벌목 조항이 2개나 보인다는 것이고, 다른 하나는 백성들의 의무 가운데 핵심인 세금(地租, 貢納) 납부, 요역(徭役)과 군역(軍役) 등 국가공동체에 대한 의무를 불성실하게 이행함으로써 국가의 건전성을 약화시키는 자에 대한 벌목 조항이 들어 있다는 것이다. 특히 이 부분은 당시 지역 사회에서 온갖 특권을 누리면서도 공적 의무는 방기함으로써 그 피해를 고스란히 힘없는 양민에게 전가시켰던 재지사족(즉 토호)들에 대한 경고이기도 하다. 이것은 퇴계가 서문에서 밝힌 “옛날 향대부(鄉大夫)의 직무는 덕행과 도예(道藝)로써 백성을 인도하고 그것을 따르지 않는 자는 형벌(刑)으로써 규탄하는 것이었다. 선비들은 반드시 집안에서 수신하고 고을에서 자신을 드러낸 후에 국정에 참여하여 나라를 발전시켰던 것이다.”라는 말이 벌목 조항으로 구체화된 것으로 볼 수 있다. 퇴계는 사회 지도층일수록 모든 면에서 솔선수범할 것을 주장했던 것이다.³⁴⁾

하벌 벌목 조항은 모두 5개인데 모두 향약 집회와 관련된 것이다. 구체적으로 살펴보면,

33) 지방 유력자, 즉 토호(土豪)가 자신의 집안 내에 여러 집을 몰래 들여놓으면 그 집은 호적대장에서 누락되어(漏戶) 그 집이 부담해야 할 지조, 공납, 군역, 요역 등으로부터 자유롭게 된다. 그렇게 되면 토호는 국역으로부터 자유로운 누호의 노동력을 자신의 부를 불리는데(개간, 농사 등) 사적으로 이용했던 것이다.

34) 다음은 퇴계가 아들 준에게 철저한 세금납부를 당부하면서 보낸 가서(家書)이다. “은계의 여러 집에서는 세금으로 납부할 쌀을 내리 가는 사람들이 언제 간다고 하더냐? 마땅히 그 날짜와 그 되와 말의 숫자를 알아야 할 것이다. [창고가 있는] 용산으로 사람을 보내, 같은 날 출발하도록 하여라. 일시에 나라의 창고에 법대로 납부한다면 더욱 좋을 것이니 소홀히 거둬 소홀히 하지 말도록 하여라.”(『退溪先生續集』 권7, 「答子稿」)

“①공회에 늦게 참석한 자. ②문란하게 앉아 예의를 잃은 자. ③좌중에서 떠들썩하게 다투는 자. ④자리를 비워 놓고 물러가 자기의 편리를 도모한 자. ⑤특별한 일 없이 먼저 집회를 떠난 자.”이다. 하벌이기 때문에 처벌 역시 향회 때 좌중(座中)에서 면책(面責)하는 것으로 비교적 가볍다.

이와 같은 극벌, 상벌, 중벌 조항 외에 퇴계는 특별 조항 4개를 더 두고 있다. 그것은 “①악행을 저지른 향리, ②향리로서 민가(民家)에 폐를 끼치는 자, ③공물(貢物)을 취급하는 자로서 물건의 값을 지나치게 징수하는 자, ④서인(庶人)으로서 사족을 능멸하는 자”이다. 특별 조항은 모두 향회 구성원이 아닌 향리와 상민(常民)을 대상으로 한 것이다. 이러한 이유로 몇몇 학자들은 퇴계의 향약을 유향소의 향규로 분류하기도 하고, 사족향약으로 분류하기도 했던 것이다.³⁵⁾ 아무튼 예안지역의 사족들이 유향소를 중심으로 향약을 운영하고자 했던 것은 분명해 보인다. 그렇기 때문에 향리 규찰의 내용이 포함되었던 것이다. 이러한 점을 통해 볼 때, 퇴계는 향리 규찰과 상민 통제를 위한 특별 조항을 향약에 포함시킴으로써 향촌자치조직의 자치권을 강화하고자 했다고 평가할 수 있다.

3.3 두 향약의 특징

앞서 우리는 주자 향약과 퇴계 향약의 출현 배경, 작성 목적, 그리고 그 내용 및 구성체제까지 살펴보았다. 지금부터는 두 향약의 특징에 대해서 살펴보려고 한다. 향약의 특징은 다음과 같은 키워드를 중심으로 분석할 것이다. 첫째는 향약 조문의 체제, 둘째는 구성원과 규약의 적용대상이고, 셋째는 관, 즉 국가와의 관련성이며, 넷째는 과실에 대한 처벌 방식과 수위에 대해 살펴볼 것이다.

1) 향약 조문의 구성체제

주자의 향약은 ‘4대 강목을 소개한 후 곧바로 향임(鄕任)의 직책과 선출 방식을 소개하고 장부 세 권에 대해 안내’하는 서문 격의 글과 ‘4대 강목(綱目)의 세부적 내용을 체계적으로 소개’하는 본론 격의 글, 그리고 자신이 여씨향약의 글을 다른 글을 참고하여 수정·보완하였다는 결론 격의 글로 이루어져 있다. 그리고 여기에 더해 말미에 부록 형식의 향약 집회 의례를 추가하고 있다. 그렇기 때문에 주자의 향약은 형식적인 측면에서 보면 매우 체계적인 구조를 지니고 있다고 평가할 수 있다. 그런데 문제는 형식적인 측면에서 보면 체계적인 구조를 갖

35) 박현순, 「조선시기 향벌의 내용과 추이」, 『국사관논총』 105, 국사편찬위원회, 2004, 7쪽, 참고.

고 있는 것처럼 보이지만 세부 내용으로 들어가면 다소 중복적이고 산만한 느낌이 든다. 이렇게 평가하는 이유는 크게 두 가지 사항 때문이다. 먼저, 주자는 ‘덕업상권’ 부분을 ‘덕’과 ‘업’으로 나누어 서술하는데 무엇 때문에 이렇게 ‘덕’과 ‘업’으로 나누어 서술하는지가 분명하게 설명되고 있지 않다. 상대적으로 더 적은 조항으로 이루어진 ‘업’의 몇몇 조항은 ‘덕’의 조항과 중복되기도 하지만 또 어떤 조항은 전혀 중복되지도 않는다.³⁶⁾ 둘째, 일반적으로 생각할 때 ‘덕업상권’의 ‘덕’목 조항과 ‘과실상규’의 ‘과실’ 조항은 서로 상반되어야 한다. 왜냐하면 덕목을 실천하지 못하게 되면 대체로 과실로 이어지고, 행동에 과실이 없게 되면 덕목을 갖춘 인물로 평가되기 때문이다. 그러므로 데칼코마니처럼 정확하게 일치하지는 않더라도 내용상 덕목과 과실은 서로 짝이 되어야 한다.³⁷⁾ 그런데 주자 향약에서 ‘덕업상권’의 덕목과 ‘과실상규’의 과실 목록은 앞서 확인했듯이 그렇지 않다. 여기에 더해 주자 향약이 갖는 체제 면에서 독특한 특징으로는 집회시의 의례(월단집회독약회)가 추가되어 있다는 것이다. 여씨향약에는 없던 집회 시 의례 부분을 주자가 추가한 것은 향약 실천의 효과를 거두고 조직 구성원들 간의 결속을 다지기 위한 목적 때문으로 보인다.³⁸⁾

반면 서문과 별목 및 처벌 규정만으로 구성된 퇴계의 향약은 형식적인 측면에서 많은 것이 생략되어 있어 마치 하나의 완비된 글이 아닌 것처럼 보인다. 주자 향약을 비롯하여 여타의 향약에서 흔히 보이는 향임 선출 절차와 방법에 대한 것, 덕목에 대한 것, 환란(患亂)시 상호 부조에 대한 것, 그리고 집회시 의례에 대한 것이 전혀 보이지 않기 때문이다. 퇴계의 경우 이미 주자의 향약을 보았을 텐데³⁹⁾ 무엇 때문에 이와 같이 향약을 구성했던 것일까? 퇴계는 주

36) 업을 기준으로 보면 집 안에서 하는 일(事父兄, 教子弟)과 집 밖에서 하는 일 가운데 몇 조항(事長上, 接朋友, 教後生, 御僮僕)은 덕의 조항과 중복되기도 하지만, 그 나머지 “글을 읽고 받을 관리하며 (...)” 이하의 조항은 덕목에 보이지 않는 완전히 새로운 것이다.

37) 대개 여러 향약을 검토한 결과 과실의 조항이 덕목 조항 수보다 많다. 그 이유는 과실은 처벌로 이어지기 때문에 덕목 보다 좀 더 구체적으로 표현되어야 하기 때문으로 보인다. 가령 덕목 조항에 ‘부부사이에 예가 있어야 한다’라는 것이 있으면, 과실 조항에서는 “부부가 치고 받고 싸우는 행위”, “본처를 쫓아내는 행위”, “부부가 분별없이 행동하는 경우”, “적처를 뒤바꾸는 행위”, “첩을 처로 삼는 행위” 등과 같이 좀 더 구체적으로 표현할 수 있다. 하지만 이러한 경우 숫자에 차이가 있지만 내용상 서로 짝을 이루고 있다고 볼 수 있다. 물론 덕목과 과실의 관계는 과실의 성질에 따라 다르게 평가할 수 있다. 다시 말해 과실의 종류가 도덕적 차원의 것이면 양자가 상반될 가능성이 높고, 반면 과실이 법률적 차원의 것이면 상반되지 않을 수도 있다. 예를 들어 과실상규 ‘犯義之過’조에 보면 ‘술주정과 노름, 싸움과 송사를 자주 벌이는 과실’이 나온다. 우리는 이러한 과실을 범하지 않는다고 해서 곧바로 그 사람에게 덕이 있다고는 말하지 않는다. 더 나아가 행위가 있어야 하기 때문이다. 하지만 주자 향약의 과실상규 부분에 나오는 과실의 경우는 그렇게 하지 않은 경우에는 대부분 덕이 있다고 말할만한 성질의 과실이다. 이에 반해 효령대군의 「증제향현목」(선목 21조와 악목 35조로 구성)과 울곡의 「서원향약」에서는 비교적 분명하게 덕목과 과실의 상반성을 확인할 수 있다. 전자는 선목 21조와 악목 35조를 수록하고 있어 악목이 조금 더 많은데, 선목 21조와 상반되는 항목을 악목 35조에서 거의 다 찾을 수 있다. 그리고 후자 역시 「條目」 1조에서 선목과 악목에 대해 소개하고 있는데, 대체로 선목과 상반되는 내용을 악목에 수록하고 있다.

38) 이성무, 앞의 논문, 같은 곳.

자를 매우 존경했지만, 향약의 경우에는 국가의 정치적 상황, 향촌이라는 지역적 특성 등에 맞게 구성하는 것이 더 적절하다고 판단했던 것으로 보인다. 다카하시 토오루가 조선향약의 시초라고 평가했던 태조의 향약,⁴⁰⁾ 퇴계의 향약보다 대략 100년이나 앞선 광주향약⁴¹⁾ 역시 퇴계의 향약과 형식적 측면에서 매우 유사하다.⁴²⁾ 그리고 영남지역의 많은 향약 역시 퇴계 향약을 기본 모델로 하고 있으며, 적어도 과실상규 부분에서는 퇴계의 향약을 상당 부분 인용하고 있다.⁴³⁾

퇴계의 향약이 별목이 나오고 해당 별목에 대한 처벌 규정으로 극벌, 중벌, 하벌을 제시하고 있다면, 태조의 향약은 거꾸로 처벌 규정이 먼저 나오고 그러한 처벌에 해당되는 별목을 제시하고 있다.⁴⁴⁾ 그렇기 때문에 태조의 향약 역시 별목과 처벌규정만으로 구성된 향약으로 볼 수 있다. 그렇다면 태조와 퇴계의 향약은 다른 향약에서 흔히 보이는 덕목, 그리고 향임 선

39) 연보에 따르면, 퇴계는 43세 무렵 왕명으로 인해 『주자대전』 교정작업을 수행하였다. 그렇기 때문에 이때 『주자대전』에 수록된 『증손여씨향약』을 보았을 것이다.

40) 광무 7년인 1903년에 간행된 『鄉憲』의 기록에 따라, 강점기 이후 학자들은 영종의 「太祖高皇帝御製惠目」(일명 풍패헌목)은 1398년에 태조가 작성하였고, 함흥의 「효령대군향헌목」(일명 함흥향헌)은 효령대군 李補(1396~1486)가 작성한 것으로 알려져 왔다. 이에 대해 田川孝三은 작성주체를 태조가 아니라 정도전으로 보고, 작성 시기 또한 태조 말년 내지 태종원년경으로 보고 있다. 하지만 배기현은 이와 같은 기존 주장에 대해 다음과 같은 반론을 제기하면서 작성주체를 모두 효령대군으로 보고, 작성시기는 전자는 1458년으로, 후자는 1469년으로 보고 있다. 배기현, 「16세기 향촌지배질서와留향소의 성격」, 『대구사학』 35, 1988, 대구사학회, p.10 각주 48번 참고. 본고에서는 명칭에 근거해 편상 전자를 ‘태조의 향약’으로 부르기로 한다.

41) 『수암지(秀巖誌)』의 「광주향약조목」에 보면 이선제가 1451년에 향약조문을 작성하여 실시하였다는 기록이 나온다. 매우 특이한 점은 1451년 이선제가 작성한 것으로 보이는 광주 향약이 1556년 퇴계가 작성한 향약과 거의 동일하다는 것이다. 이 부분에 대한 연구는 향후 필요할 것으로 보인다.

42) 박익환도 “퇴계는 예안향림약조의 조목을 세움에 있어 중국의 여씨향약보다는 위의 두 향헌과 안동향규구조(鄉規舊條)를 더 참고했음을 곧 알아볼 수 있다.”라고 하였다. 박익환, 「조선전기 향규와 향규약고」, 『사학연구』 38, 한국사학회, 1984, 339쪽.

43) 17세기에서 19세기 후반 영남지역이 대표적인 향약으로 金坵鄉約(17세기 초, 김기), 苞山鄉約(17세기 초, 金世濂), 鄉約定規(17세기 후반, 李嵩逸), 望亭鄉約(18세기 중반), 夫仁洞約(18세기 중반, 崔興遠), 退老洞約(19세기 후반, 李翊九) 등을 들 수 있다. 이 가운데 형식적으로 주자향약의 체제를 따른 것은 김기의 것밖에 없는데, 김기 역시 과실상규 부분은 퇴계의 것을 그대로 수용하였다. 다음으로 부인동약 역시 비교적 주자향약의 체제를 모방했지만, 퇴계 향약처럼 부록이 없고 과실상규는 퇴계의 것을 상당히 수용하였다. 그 밖에 향약 역시 과실상규 부분은 퇴계향약을 대체로 수용하고 있는 것으로 보인다. 정진영, 「영남지역 향약의 형성과 변천」, 『향토사연구』 4, 한국향토사연구전국협의회, 1992, 10쪽-15쪽 참고)

44) 논자는 태조의 향약의 경우, 5개의 처벌규정과 40 조항의 별목과 특별규정으로 구성된 것으로 본다. ①영원히 향록에서 삭제하여 사족으로서의 신분을 박탈하는 처벌(永爲削籍)과 별목 8조항. ②자신에 한하여 향원의 자격을 정지시키는 처벌(限己身損徒)과 별목 8조항. ③영구히 수신이나 입신을 포기한 상태에 해당하기 때문에 선비의 신분은 그대로 두지만 향임에 나아갈 자격을 박탈시키는 처벌(永爲棄身事)과 별목 8조항. ④자신은 마땅히 향임에 나아갈 수 없지만 그것이 자손의까지 적용하지 않는 처벌(當身不用而勿棄子孫事)과 별목 8조항. ⑤1년 손도시켰다가 원일에 자신이 지은 죄와 그 처벌에 대해 말한 후 죄를 용서(赦)하는 잔치를 거행해야 하는 처벌, 연소자는 미허좌하고 이에 벌로 태 40을 맞은 뒤 본디 자리를 허락하는 처벌(兼年損徒, 而元日召僉懲罪後, 受宴赦身事. 年少者, 未許座, 而罰四十度後, 許本座事)과 별목 8조항. 그리고 부록으로 「향안록」(향중의 사족만을 대상으로 한 것으로 보인다)에 가입하는 절차 및 손도시벌을 받은 후 복귀하는 절차에 관해 적시하고 있다.

출 절차와 방법, 집회시 의례에 관한 내용이 누락된 이유는 어디에 있는 것일까? 우선 별목을 갖추지 않은 것에 대해서는 퇴계가 다음과 같이 밝히고 있다. 우선 덕목 교육은 국가 교육기관에서 이미 상세히 교육하고 있고⁴⁵⁾, 둘째 덕목은 인간 본성과 관계된 것이기 때문에 굳이 세세히 알려주지 않더라도 이미 알고 있다는 것이다. 퇴계는 사람들이 덕목을 몰라서 악한 존재가 되는 것이 아니라 덕목을 실천하지 않기 때문에 문란해진다고 보았던 것이다.⁴⁶⁾ 더욱이 퇴계가 생각할 때 덕목은 별목과 상반되는 것이기 때문에 별목이 있다면 덕목은 굳이 필요 없을 수도 있고, 다른 한편 별목을 부각시키는 것이 구성원들의 실천력을 높이는 데 도움이 된다고 생각했을 것이다. 그 밖에 내용이 누락된 것은 조선초기부터 향촌자치조직으로 활동했던 유향소와 관련이 있는 것으로 보인다.⁴⁷⁾ 이미 예안지역의 향회에서 향임 선출 절차와 방법, 집회시 의례에 관한 내용과 운영기관으로서 유향소가 이미 갖추어져 있었기 때문에 퇴계가 향약을 작성하면서 굳이 새롭게 제정할 필요가 없었기 때문으로 볼 수 있다. 이렇게 본다면 퇴계의 향약을 간략하면서도 향촌사회의 자치와 ‘화민성속’의 목적을 드러내는 데 손색이 없다고 평가할 수 있다.

2) 향약의 구성원과 규약의 적용대상

주자 향약의 구성원과 규약의 적용대상은 모두 향촌 재지사족으로 볼 수 있다. 구성원의 범위는 ‘덕업상권’에 보이는 덕목의 대상이 누구인지, ‘예속상교’의 예가 어떤 사람에게 적용되는 것인지, 향약의 집회시 의례가 어떤 계층 사람들에게 해당되는 것인지 등을 통해 파악할 수 있다. 그리고 규약의 적용 대상은 과실에 따른 처벌 조항 및 규정을 보면 알 수 있다. 주자의 향약은 덕업상권의 덕(德) 가운데 “중을 다스리고 정사와 교육을 엄숙하게 한다.”, “관직에 있을 때 그 직무를 잘 수행한다.”, 그리고 예속상교의 내용⁴⁸⁾과 향약 활동의 구체적

45) 『退溪集』卷42, 「鄉立約條序」; “효, 제, 충, 신의 덕목은 사람이 타고난 아름다운 성품에서 근원한 것이며, 더구나 나라에서 상(庠)과 서(序)를 설치하여 교육을 하는데 모두 이러한 것을 권하고 인도하는 방법이 아닌 것이 없다.”

46) 『退溪集』卷42, 「鄉立約條序」; “맹자께서 ‘도는 가까운 데 있는데도 그것을 먼 곳에서 구하려 하고, 일을 쉽게 할 수 있는데 어렵게 하는구나! 사람마다 부모를 부모로 대하고 어른을 어른으로 대하면 천하가 편안해진다.’라고 하였는데, 이것이 바로 공자께서 말한 지극한 덕목[至德]이며 핵심적인 법도[要道]요, 선왕께서 사람들의 마음을 착하게 만들었던 이유이다.”

47) 1588년에 작성된 안동부 유향소의 「향규」를 보면 이 역시 퇴계의 향약을 거의 그대로 수용하고 있다. 유향소의 규약이기 때문에 향임의 선출절차와 방법, 덕목, 집회시 의례 등이 있어야 하는데도 불구하고 생략되어 있다. 향원가입절차 및 명부인 「향안인록」이 「향규」 이전에 앞서 작성되었듯이, 향임의 선출절차와 방법, 덕목, 집회시 의례 등도 이미 별도로 작성되었던 것으로 볼 수 있을 것이다.(오세창, 정진영, 권대웅, 조강희 편저, 『영남향약자료집성』, 영남대학교출판부, 1986, 8쪽 참고) 이렇듯이 유향소는 조선중기 이후 본격적으로 전개되었던 향약과도 불가분의 관계를 형성하고 있는 것으로 보아야 한다. 이석희, 「조선시대 지방자치론」, 『중앙행정논집』 창간호, 중앙대학교 국가정책연구소 중앙행정학연구회, 1987, 299쪽.

인 내용을 담고 있는 ‘월단집회독약지례’ 등을 살펴볼 때 일반 평민이 아닌 사족을 구성원으로 하고 있다는 것을 확인할 수 있다. 그리고 최고의 처벌이 ‘출약’이라는 것을 볼 때 규약의 적용대상 역시 향약 구성원을 벗어나지 않고 있다. 이러한 사실을 통해 주자의 향약은 철저하게 재지사족 중심의 화민성속에 초점이 맞추어져 있다고 할 수 있다.

그렇다면 퇴계의 향약은 어떨까? 퇴계 향약은 주자 향약과 달리 향약의 체제가 벌목과 과실에 대한 처벌규정 중심으로 구성되어 있기 때문에 다른 방식의 접근을 통해 향약의 구성원과 규약의 적용대상을 확인해야 한다. 먼저, 규약의 적용대상은 벌목의 적용 범위와 처벌 대상의 범위 등을 살펴봄으로써 확인할 수 있다. “집강(執綱)으로서 [향론도 없이] 사사롭게 위부인을 향안(鄕案)에 들인 자.”, “많은 인호(人戶)를 자신의 호(戶)에 예속시켜 놓고 관역을 따르지 않는 자.” 등은 규약의 적용대상이 재지사족에 해당된다는 것을 보여주고, “원악향리”, “아전으로 민가에 폐를 끼친자.” 등은 지방 행정실무를 담당하는 향리에 해당된다는 것을 보여주고, “서인으로 사족을 능멸한 자”는 바로 상민(常民)에 해당된다는 것을 보여준다. 그러므로 퇴계 향약의 규약 적용 대상은 예안 지역에 거주하는 평민에서부터, 향리, 그리고 사족 전체에 해당된다고 할 수 있다. 그렇다면 실질적으로 향약을 운영했던 구성원은 누구였을까? 퇴계 향약에는 향원(鄕員) 관련 내용이 별도로 없기 때문에 관련 내용을 통해 간접적으로 추론해야 한다. 서문과 특별조항의 내용을 고려할 때, 퇴계가 향약 운영의 주체로 유향소를 거론했다는 점은 확인할 수 있다. 그리고 특별 조항에 따르면, 구성원은 향리를 규찰할 수 있고, 상민을 통제할 수 있는 계층이어야 한다. 이러한 점을 고려할 때 퇴계 향약의 구성원은 재지사족, 즉 예안지역 향안(鄕案)에 가입된 사람으로 한정된다고 할 수 있다. 결국 퇴계 향약은 향약을 집행하고 운영하는 주체와 향약 규범이 적용되는 대상이 상이한 특징을 갖고 있다고 할 수 있다. 이것은 향촌사회 자치와 향촌사회의 화민성속을 동시에 추구했던 조선 향약의 특징으로 볼 수 있다.

3) 관, 국가와의 관련성

주자 향약은 관, 즉 국가와의 관련성이 상대적으로 약하다고 할 수 있다. 그것은 첫째, 구성원이 향약의 규약을 위반했다고 하더라도 관에 보고하여 처벌하는 규정이 전혀 없고, 모든 처벌을 자치조직 내에서 진행하였다는 점을 통해 확인할 수 있고, 둘째는 향약조직이 향리와

48) 주자는 여씨향약을 가감하여 『증손여씨향약』을 작성하면서 예속상고 부분을 강화한다. 이때 그가 참고한 자료가 바로 여대균의 「향의」이다. 그런데 「향의」의 의관조는 인이나 진사 등이 관혼상제의 의식에서 예복의 하나로 갖추어야 할 두건인 ‘복두를 쓰는 사람’들을 대상으로 한 향촌내 의례이다. 이근명, 앞의 논문, 278쪽 참고.

같은 하급관리를 규찰한다거나 상민을 규제한다는 내용이 전혀 보이지 않는다는 점을 통해 확인할 수 있다. 주자향약은 철저히 재지사족 중심의 ‘화민성속’을 주요 목적으로 했던 것이다. 예를 들어 덕과 업에는 “관직에 있을 때 그 직무를 잘 수행한다.”, “법령을 두려워하며 조세와 공물을 잘 납부한다.” 등과 같은 국가공동체와 관련된 윤리 조항이 있음에도 이러한 윤리 조항을 어긴 것에 대한 처벌에 관에 보고하여 치죄한다는 내용이 보이지 않는다. 이러한 점을 통해 주자의 향약은 국가와의 관련성이 상대적으로 약하다는 것을 확인할 수 있다.

반면 퇴계의 향약은 국가와의 관련성이 상대적으로 강하다고 할 수 있다. 그것은 첫째, 퇴계의 향약이 유향소와 긴밀히 연계되어 있다는 점을 통해 확인할 수 있고, 둘째는 과실 정도가 상대적으로 엄중한 경우 그 처벌로 “관에 알려 법에 따라 처벌한다”라는 내용이 있다는 것을 통해 확인할 수 있다. 그 밖에도 퇴계 향약의 별목에는 수령 밑에서 지방행정을 담당하고 있는 향리를 규찰하는 내용도 포함되어 있는데, 이것 역시 퇴계의 향약이 주자의 것보다 국가와의 관련성이 높다는 것을 보여준다. 이러한 특성 역시 유향소와 밀접하게 결부된 퇴계 향약의 특징으로 볼 수 있다.⁴⁹⁾

4) 과실에 대한 처벌 방식과 수위

주자 향약은 과실에 대한 처벌 수위가 상대적으로 약하다. 앞서 살펴보았듯이 주자 향약의 처벌 규정은 크게 세 단계로 이루어진다. 첫 번째 단계는 ‘주의 조치’이다. 작은 과실은 ‘조용히’ 주의를 주고 큰 과실은 ‘여럿이’ 주의를 준다는 차이만 있을 뿐이다. 두 번째 단계는 첫 번째 단계의 조치에 대해 승복하지 않을 때에 한해서 진행된다. 이 단계는 집회 때 약정이 타이트한 후, 과실을 저지른 사람이 사과하고 개선하려고 하는 의향이 있으면 ‘악적’에 기록한 후 고치기를 기다린다. 그러나 끝내 향약조직의 처벌에 승복하지 않는다면 마지막 세 번째 단계로

49) 1488년(성종 19년) 성종은 ‘화민성속’을 이루기 위해 결국 유향소를 다시 설치하게 된다. 이때 부활된 유향소는 시간이 흐르면서 점차 향촌사회의 지배권을 확보하려는 노력과 동시에 조직의 존속을 위해 국가의 수령권 안정화 정책도 수용하게 된다. 이 과정에서 유향소는 향촌사회의 지배권 확보라는 자신들의 목표를 이루기 위해 수령권 안정화라는 중앙 정부의 움직임과 대결하는 대신 적절한 타협책을 모색한다. 이러한 움직임은 임진왜란 후에 가속화된다. 전란으로 인해 향안과 향청이 불타버린 데다가 1603년(선조 36년)에는 유향소의 감독 및 지원기관이었던 경계소까지 혁파된다. 경계소의 혁파로 인해 유향소의 향임을 이제 수령이 임명하게 된다.(이석희, 앞의 논문, 301쪽) 그 결과 유향소는 관의 하부조직으로 점차 편입된다. 이렇게 하여 유향소는 향촌사회의 자치조직으로 기능이 점차 약화된다. 유향소 향임들은 합법적으로 수령과 결탁함으로써 읍사(邑事)에 깊이 개입하게 되고, 수령의 하수인으로 전락하게 되자, 사람들은 유향소를 아관(亞官) 또는 아청(亞廳)으로 대우했던 것이다. 배기현, 앞의 논문, 118쪽 참고. 유향소의 관변화화 동시에 향약 역시 17세기로 접어들면서 관변화의 성격을 띠기 시작한다. 이러한 점은 안동향약의 “각 면의 상유사는 옛날에 약정의 직임인데 지금은 관가의 크고 작은 정령이 맡겨지고 있다. 상유사를 세운 본 뜻이 어찌 말단의 심부름꾼으로 만드려는 데 있었겠는가? 참으로 한심한 노릇이다.”라는 구절을 통해서도 확인할 수 있다. 우홍준, 「조선 후기 관변자치제도인 ‘수령향약’의 특성」, 『한국행정사학회지』 29, 2011, 213쪽 참고.

들어간다. 최종 단계인 세 번째 단계의 처벌은 향약조직에서 퇴출시키는 ‘출약’(出約)이다. 과실에 대한 주자 향약의 처벌 수위는 문자 그대로 ‘서로 규찰하고 그 정도가 너무 심하면 조직에서 퇴출하는 것’에 그칠 뿐이다. 그 외에 관에 처벌을 호소하거나, 자체적으로 형벌을 가하거나 하지는 않는다. 이것 역시 주자 향약이 재지사족을 중심으로 한 향풍교정에 초점을 맞추었기 때문으로 보인다.

반면 퇴계의 향약은 처벌 수위가 상대적으로 강하다. 퇴계 향약의 처벌 규정은 앞서 살펴 보았듯이 극벌 상, 중, 하, 중벌 상, 중, 하, 하벌 상, 중, 하 이렇게 9단계로 세분화되어 있고, 여기에 더해 특별 조항의 처벌 규정까지 있다. 이 가운데 극벌 ‘상’에 해당되는 처벌은 과실 정도가 매우 엄중하기 때문에 국법에 따른 처벌과 자체 처벌인 ‘교류금지’(不通水火) 조치가 결합되어 있다. 이러한 처벌은 분명 주자 향약의 최고 처벌인 ‘출약’보다 훨씬 더 강력한 처벌이라고 할 수 있다. 극벌 ‘중’에 해당하는 과실을 처벌하는 규정인 ‘삭적’(회원자격 박탈)은 주자 향약의 최고 처벌인 ‘출약’과 비슷하다고 할 수 있지만, 퇴계의 향약은 여기에 더해 ‘고을의 일 처리에 함께 하지 않는다’(不齒鄉里)는 자체 처벌까지 더 하고 있기 때문에 극벌 ‘중’의 처벌도 주자 향약의 최고 형벌보다 강하다고 할 수 있다. 퇴계 향약에서 보이는 높은 수위의 처벌 역시 조선향약의 특징으로도 볼 수 있다.

4. 조선향약으로서 퇴계향약의 가치

퇴계가 예안 지역의 향촌자치 규약으로 작성했던 향약은 물론 시행되지 못했다. 하지만 우리는 퇴계의 향약이 예안 지역에서 시행되었더라면 당시 예안 지역의 향촌사회가 어떠한 가치를 추구하고자 했는지에 대해서는 답을 찾을 수 있을 것이다. 왜냐하면 퇴계의 향약 속에 당시 향촌자치조직이 추구했을 법한 가치가 충분히 담겨져 있을 것이기 때문이다. 아래에서는 퇴계 향약이 당시 향촌자치조직의 규약으로 추구하고자 했던 가치를 살펴보고자 한다.

4.1 공권력과 자치권의 공존 추구

주자의 향약과 달리 퇴계의 향약은 국가와 긴밀한 관련성을 갖는다. 주자 향약의 경우에는 재지사족들의 향촌사회 향풍교정에 초점을 맞추어져 있기 때문에 향촌사회 구성원의 과실에 대한 처벌 역시 국가의 공권력을 빌리지 않았던 것으로 보인다. 하지만 우리의 경우는 중국과 사정이 조금 다르다. 조선을 건국한 세력은 중앙집권적 정치체제를 무사히 안착시키고 강화하기 위해 지방을 주, 부, 군, 현 체제로 정비한 후, 그곳에 수령을 파견하여 국가가 직접

통치하고자 했다. 반면 오랫동안 향촌사회를 실질적으로 이끌었던 재지품관들도 자신들 나름대로 향촌사회를 올바른 방향으로 이끌기 위해 자치조직을 결성하는 움직임을 보였다. 이렇게 하여 중앙정부가 향촌사회를 직접적으로 통치하려는 움직임과 향촌사회조직이 자치적으로 향촌사회를 운영하려는 움직임 사이에 갈등이 있었던 것이다. 그 결과로 조선초기 향촌사회 자치조직의 운영기관이었던 유향소의 혁파와 재설치가 반복되었던 것이다. 그런데 이러한 갈등 국면에서 중앙정부는 향촌사회 자치조직의 자치를 어느 정도 인정해주고, 향촌사회 자치조직은 중앙정부의 요구에 협조하면서 해결책을 모색하기 시작한다.

중앙정부의 경우 향촌사회를 직접 통치하기 위해서는 넓은 향촌지역의 행정업무를 도와 줄 사람(즉 향리, 면리임, 권농관, 감고 등)의 도움이 반드시 필요한데, 외부에서 파견된 수령이 그 지역에서 오랫동안 활동하며 기반을 닦아온 이들을 효과적으로 통제하기 어려웠던 것이다. 그렇기 때문에 중앙정부에서는 특히 토성 출신의 재지품관들의 도움이 절실했다. 그들의 협조 없이는 수령이 효과적으로 향촌사회를 통치하기란 매우 어렵기 때문이었다.

한편, 향촌사회의 실질적 세력가인 재지품관 역시 자신들 나름대로 향촌사회에서 주도권을 잡고, ‘화민성속’을 이루기 위해서는 자치조직의 결성과 조직의 운영기관으로 유향소의 설치가 필요했다. 그런데 중앙집권제를 표방하는 조선에서 중앙정부의 인정 없이 향촌자치조직이 자치활동을 성공적으로 수행하기란 매우 어렵다. 이러한 상황에서 향촌자치조직은 유향소 복설(復設)이라는 조치를 통해 향촌사회에서 자치를 인정받는 동시에 중앙정부에서 파견한 수령의 향촌사회 통치에 협조적인 자세를 취한다. 이들은 이 과정에서 자신들 조직의 규약을 자체적으로 마련하는데, 이것이 바로 조선향약의 탄생으로 볼 수 있다. 이렇듯이 조선향약에는 향촌사회 구성원들이 주체가 되어 정치적 측면에서의 자치와 윤리적 측면에서의 교화를 동시에 추구하고자 했던 것이다.

이와 같은 조선향약의 특색이 퇴계의 향약에도 그대로 녹아 있다. 향약이 향촌사회구성원들의 윤리적 교화에만 초점이 맞추어져 있다면 높은 수위의 처벌이나 관에 고발하는 처벌 규정이 필요하지 않을 수도 있었다. 하지만 정치적 측면에서 자치(自治)를 효과적으로 실현하기 위해서는 공권력의 힘을 빌리지 않을 수 없었다. 향약은 향촌사회구성원들의 자치규약이긴 하지만, 어디까지나 국가의 테두리 내에 있는 조직의 규약이기 때문에 중앙정부와의 연대성 없이는 원활한 활동을 유지하기 어렵다. 정치적 측면에서 향촌사회구성원들을 효과적으로 통치하기 위해서는 국법 준수의 엄중함을 일깨워주어야 하고, 윤리적 측면에서 향촌사회구성원들을 교화하기 위해서 윤리 규범 준수의 엄중함을 일깨워주어야 한다. 그리고 이와 같은 법과 윤리적 규범 준수는 반드시 향촌사회 지도층이 앞장서야 한다. 향촌사회에서 세력이

있다고 해서 관청의 공무에 간섭한다거나 사적인 위세를 부려 관청을 소란스럽게 한 것을 공권력에 호소해서 처벌하지 않고 다만 향촌사회 내에서 자체적으로 처벌한다면 이것은 향촌 자치라는 미명하에 국가권력을 무시하는 처사가 되어버린다.

국가체제 속에서 공권력에 대한 무시는 향촌자치조직의 존립을 위태롭게 만들며, 그렇게 되면 향촌자치조직의 자치는 결국 무산될 수밖에 없다. 그렇기 때문에 중앙정부의 통치와 향촌자치조직의 자치가 공존하기 위해서는 중앙정부 차원에서는 향촌자치조직을 존중하고 인정해주는 것이 필요하고, 향촌자치조직의 차원에도 중앙정부의 공권력을 존중해주는 것이 반드시 필요하다. 공권력과 자치권의 공존을 추구한 퇴계 향약의 특성은 공권력의 상징인 중앙정부 통치와 자치조직의 자치가 원만하게 공존할 수 있는 길을 제시한 것으로 평가할 수 있다.

4.2 가족윤리에서 사회윤리로, 그리고 국가윤리로의 확장

앞서 살펴보았듯이, 퇴계 향약의 별목 조항을 규범이 적용되는 영역을 기준으로 분류하면 가족공동체 윤리, 사회공동체 윤리, 국가공동체 윤리에 관한 것으로 구분할 수 있다. 이 가운데 퇴계는 윤리의 근간이 되는 가족공동체의 윤리와 외연 범위가 가장 넓은 국가공동체 윤리를 가장 엄중한 것으로 간주한다. 그렇기 때문에 그러한 윤리를 어겼을 때 극별로 처벌한다. 그리고 가족공동체와 국가공동체 중간 영역인 사회공동체와 관련된 윤리는 중별로 처벌하고, 향회 집회시에 국한되는 행위와 관련된 윤리는 하별로 처벌한다. 퇴계가 별목에서 가족공동체의 윤리나 사회공동체의 윤리 외에 국가공동체의 윤리를 수용하고, 국가공동체 윤리를 위반했을 때 엄중하게 그 책임 묻고 있는데, 이것은 퇴계 향약이 비록 향촌사회구성원들의 자치 규약이긴 하지만 향촌사회의 근간이 되는 가족공동체의 윤리가 국가공동체와 관계된 윤리까지 확장되고 있다는 것을 보여주는 것이다. 이와 같은 관점은 유교 윤리의 본질을 담고 있는 것이기도 하다.

유교에서 말하는 인(仁)은 자신과 가까운 사람을 그렇지 않은 사람보다 친하게 대우하는 ‘친친’(親親)의 사랑이며, 그렇기 때문에 공자는 부모에 대한 사랑을 실천한 효(孝)와 형제에 대한 사랑을 실천한 제(悌)를 인(仁)을 실천하는 근본이라고 했던 것이다.⁵⁰⁾ 즉 나와 가장 가까운 사람에 대한 사랑을 먼저 실천하고, 이러한 사랑을 점점 외부로 넓혀서 사회공동체와 국가공동체로까지 확장하라는 것이다. 퇴계는 이와 같은 가치를 자신이 입안한 향약에 담고자 했던 것이다. 나를 낳아주고 길러준 부모가 소중하다면 마땅히 누군가를 낳고 길러준 이웃의 부모도 소중할 것이고, 우리들의 소중한 부모, 형제가 안심하고 생업에 종사할 수 있도

50) 『論語』, 「學而」: “孝弟也者, 其爲仁之本與.”

록 기반을 마련해주고 보호해주는 공동체의 리더도 소중한 것이다. 친친의 윤리가 가족공동체에만 머물고 사회공동체와 국가공동체로 확장되지 못한다면 사회공동체와 국가공동체는 건전하게 운영될 수 없을 것이다. 인간은 사회생활을 할 수밖에 없는 존재이다. 그러므로 사회공동체와 국가공동체의 문란(紊亂)은 결국 부메랑이 되어 돌아와 공동체의 토대하고 할 수 있는 가족공동체의 존립마저 위협하게 될 것이다.

극벌 중 “관청의 일에 간섭하여 영향력을 행사하고 고을의 풍속을 어지럽힌 자”와 “위세를 부려 관청을 소란스럽게 만들고 자기 마음대로 행동하는 자”, 중벌의 “공적, 사적 모임을 불문하고 관청의 일에 시비를 거는 자”, “공무를 빙자하여 폐해를 만드는 자” 등을 벌목으로 규정하여 강하게 처벌하는 내용은 주자 향약에서는 볼 수 없는 퇴계 향약의 특징이라고 할 수 있다. 특히 “많은 인호를 자신의 집에 숨겨놓고 국역(官役)을 회피하는 자”와 “세금납부(租賦)에 힘쓰지 않고 국역(徭役)을 회피하기 위해 꾀를 부리는 자”에 대한 처벌 규정은 국가공동체 유지에 필수적인 것으로 향촌사회구성원들의 자율적 자기통제를 한 단계 높은 차원으로 끌어올리고 있다. 조선초기부터 재지품관의 주도로 설치되었던 유향소가 성종대 공인되기 이전까지 여러 차례 혁파되었던 가장 큰 이유는 향촌의 유력자인 재지품관이 향원(鄉愿) 노릇을 했기 때문이다. 공권력의 사각지대에 놓여있는 향촌 유력자들이 자기절제 없이 폭주하게 되면 수령권은 약화될 것이고, 수령권이 약화되면 당연히 공권력의 보호를 받아야 할 많은 백성들이 피해를 보게 될 것이다. 조선초기의 사회로부터 이와 같은 교훈을 얻은 퇴계는 가족공동체의 윤리로부터 국가공동체의 윤리까지 확장되는 내용을 향약에 담아 각 공동체의 지속가능한 공존을 모색했던 것이다.

4.3 법과 제도보다 도덕성 중시

조선시대 향촌사회구성원들이 스스로 향촌사회를 자치(自治)하기 위해 설치한 기관이 바로 유향소이다. 유향소가 강조했던 향촌사회 통치방식은 중앙정부의 공권력을 앞세운 법과 제도에 기반한 통치가 아니라 도덕성을 앞세운 자치였다. 물론 그렇다고 법과 제도에 기반한 통치를 거부했다는 말은 아니다. 그렇게 할 수도 없고, 그렇게 해서도 안 된다. 이러한 이유로 향촌사회를 건전한 방향으로 발전시키기 위해 향촌자치조직은 도덕성에 기반한 자치의 한계를 공권력에 기반한 법과 제도라는 외적 강제를 통해 보완하고자 했던 것이다. 그럼에도 여전히 강조점은 도덕성에 입각한 자치에 있었다.

조선의 향약, 특히 퇴계의 향약도 향촌자치조직의 기관이었던 유향소와 밀접한 관계를 맺고 있다. 그러므로 퇴계 역시 법과 제도라는 외적 강제장치를 동원해 향촌사회를 운영하기

보다는 윤리와 도덕에 기반한 내적 강제 장치를 사용하여 향촌사회를 운영하고자 했던 것이다. 이것은 “세월이 흐르면서 선왕이 제정한 법과 제도는 사라졌지만, 사람으로서 지켜야 할 도덕까지 사라진 것은 아니다. 그러므로 어떻게 시대(古今)적 상황에 따른 마땅함을 헤아려서 [백성들을] 권면하고 꾸짖지 않을 수 있겠는가!”⁵¹⁾라는 퇴계의 발언을 통해서도 확인할 수 있다. 법과 제도는 가변적이고 부차적인 것이지만 인간으로써 마땅히 지켜야 할 도덕은 영원하고 본질적이라는 말이다.

퇴계의 향약에는 두 가지 목표가 들어있다. 하나는 ‘화민성속’, 즉 향촌사회의 풍속을 성리학적 이념에 따라 아름답게 교정하는 것이고, 다른 하나는 중앙정부의 일방적이고 완전한 통치에서 벗어나 향촌사회를 자치적으로 운영하는 것이다. 그런데 이러한 두 가지 목적은 별개의 항목으로 분리되어 있는 것이 아니라 서로 연결되어 있다. 왜냐하면 성리학적 이념에 따르면 정치의 최종 목적이 결국 ‘화민성속’에 있기 때문이다.⁵²⁾ 이것은 『대학』의 조목이 ‘수신’에서 ‘제가’로, ‘제가’에서 ‘치국’으로, ‘치국’에서 ‘평천하’로 확장되는 구도를 통해서도 확인할 수 있다. 개인의 인격적 성숙을 통해 가족공동체의 윤리를 확립하고, 가족공동체의 윤리를 확장하여 국가공동체의 윤리를 확립한다는 것이다. 그렇게 하여 전체 구성원들이 유교적 가르침에 따라 교화하고 그로 인해 올바른 사회 풍속을 확립할 수 있게 되면 세상은 자연스럽게 평화롭게 된다는 것이다. 여기서 우리는 퇴계 향약의 두 가지 목표가 결국에는 ‘화민성속’을 지향하고 있다 것을 확인할 수 있다.

그런데 향촌사회의 자치를 통해 ‘화민성속’이라는 목표를 이루기 위한 방법에 두 가지가 있을 수 있다. 하나는 법과 제도를 통해 강제하는 것이고 다른 하나는 윤리와 도덕을 통해 강제하는 것이다. 윤리와 도덕은 내적 강제력을 가진 장치이지만⁵³⁾, 법과 제도는 외적 강제력을 가진 장치이기 때문에 목표달성이라는 현실적인 측면에서 보면 법과 제도가 훨씬 더 매력적으로 보인다. 하지만 퇴계는 법과 제도를 통한 자치보다는 윤리와 도덕을 통한 자치를 더 중요하고 근원적인 것으로 간주한다. 효과적인 면에서 다소 효율성이 떨어질 것으로 보임에도 도덕을 통한 자치를 강조했던 이유는 무엇일까? 이 물음에 대한 대답은 “백성들을 법으로 이끌어 따르지 않는 자를 형벌로 다스린다면 백성들은 벌을 면하려고만 할 뿐 자신의 행동에

51) 『退溪集』卷42, 「鄉立約條序」.

52) 이것은 주자가 『대학』의 삼강령 중 하나인 ‘親民’을 ‘新民’으로 해석한 데서 잘 드러난다. 여기서 주희는 백성들을 친밀하게 대해야 할 공동체의 구성원으로 보기보다는 유교적 교양을 갖춘 지식인, 즉 사대부가 교화해야 할 대상으로 보았던 것이다. 그러므로 성리학자들은 치인을 화민성속과 동일한 것으로 간주한다고 할 수 있다. 한재훈, 「향에 대한 퇴계의 이해와 실천」, 『한국서원학보』 8, 2019, 85쪽 참고.

53) 물론 법과 도덕이 선명하게 구분되지 않았던 전통사회에는 이와 같은 표현이 적절하지 않을 수도 있다. 하지만 오늘날만큼은 아니더라도 유교사회였던 조선 역시 법제(法制)의 영역과 도(道), 인륜(人倫)의 영역을 어느 정도 구분하고 있었던 것으로 볼 수 있다.