

## 페미니즘과 테러리즘<sup>\*†</sup>

이 준 효

이 논문의 목표는 오늘날 대한민국 언론 및 페미니즘 분야에서 널리 유행하고 있는 집단혐오용어 사용(혹은 혐오표현 일반)에 관한 한 가지 도덕적 논제 즉, 비대칭성 논제(= 여성 일반(혹은 소수자/약자)에 대한 혐오 표현은 도덕적으로 허용가능하지 않지만, 남성 일반(혹은 다수자/강자)에 대한 혐오 표현은 도덕적으로 허용가능하다)가 정당화될 수 없음을 보이는 것이다. 비대칭성 논제를 위한 한 가지 대표적인 논증은 적어도 일부의 남성들이 먼저 여성혐오표현을 사용해 여성들을 공격하였기에 여성들의 남성혐오표현 사용은 도덕적으로 정당하다는 것이다. 그러나 비대칭성 논제를 위한 이 논증은 오사마 빈 라덴의 테러리즘을 위한 논증과 동일한 논리적 구조를 가지고 있기에 둘은 패키지딜(package deal)이다. 비대칭성 논제를 위한 주요 논증들을 비판한 후, 나는 혐오표현의 도덕적 정당성은 무고한 사람에 대한 공격을 포함하느냐 여부에 달려 있다고 논증한다. 이러한 관점에 따르면, 여성에 대한 혐오표현뿐만 아니라 남성에 대한 혐오표현 역시 무고한 사람에 대한 공격을 포함하기에 도덕적으로 정당화될 수 없다.

**주 제** 응용언어철학, 규범 윤리학, 페미니즘

**주요어** 페미니즘, 테러리즘, 비대칭성 논제, 혐오표현, 집단혐오용어

\* 투고일: 2021. 05. 05, 심사 및 수정 완료일: 2021. 06. 08, 게재 확정일: 2021. 06. 09

† 본고의 이전 버전들을 읽고 유익한 도움을 준 이정규, 이승수 두 분 선생님께 큰 고마움을 표한다. 또한 본고에 대해 유익한 심사평을 해준 익명의 세 심사위원들에게도 진심으로 감사드린다.

## 1. 혐오표현(hate speech)과 비대칭성 논제

이 논문은 집단혐오용어(slur)와 혐오표현(hate speech)에 관한 한 가지 도덕적 논제에 대해 다룰 것이다.<sup>1)</sup> 먼저 두 용어의 정의에 대해 살펴보자. 집단혐오용어(slur)는 인종, 성, 성적 지향, 종교, 국적 등을 근거로 해당 그룹에 대한 혐오를 표현하는 단어이다 (Anderson and Lepore 2013; Popa-Wyatt and Wyatt 2018). 대표적인 예로 영어에서 “nigger”는 흑인, “chink”는 중국인, “cunt”는 여성, “faggot”은 동성애자 남자에 대한 집단혐오용어이다. 한국어에서는 “검둥이”, “짱꿀라”, “김치녀”, “된장녀”, “맘충”, “한남(충)”, “한남유충”, “똥꼬충” 등이 집단혐오용어로 분류될 수 있다.<sup>2)</sup> 다른 한편, 혐오표현(hate speech)은 일반적으로 인종, 성, 국적, 성적 지향, 종교 등 개인의 정체성에 핵심적인 요소들을 근거로 개인 혹은 집단에 대한 혐오를 표현하거나 폭력을 선동하는 말, 글 혹은 행동으로 정의된다 (United Nations Strategy and Plan of Action on Hate Speech; Cambridge English Dictionary). 이러한 정의에 따르면, 집단혐오용어를 사용하는 것(slurring)은 혐오표현(hate speech)의 일종이다.

여기서 우리는 사용(use)과 언급(mention) 구분에 대해 주의할 필요가 있다. 만약 수업시간에 한 교수가 인종차별적인 편견을 전달하기 위해 “니거(nigger)는 범죄자다”라는 말을 했다면 이는 언어표현을 사용한 것이다. 이때 “니거”라는 단어는 세계 내의 대상(즉, 흑인 집단)을 가리킬 것이다. 이에 반해 수업시간에 교수가 “‘니거’라는 표현은 인종차별적 용어이다”라는 말을 했다면 이는 언어표현을 언급한 것이다. 교수의 발화에서 “니거”는 언어적 표현 자체를 가리키는 것이지 세계 내 대상을 가리키는 것이 아니다. 첫 번째 교수의 발화는 집단혐오용어를 사용하는 것이기에 혐오표현의 일

1) 나는 이 논문에서 “slur”를 “집단혐오용어”로, “hate speech”를 “혐오표현”으로 번역하겠다. 한 가지 주의할 점은 집단혐오용어는 언어적 표현인 단어를 의미하지만 혐오표현은 말, 글, 행동, 상징 등 모든 종류의 의사소통수단을 포함하는 보다 넓은 의미를 가진다는 것이다.

2) 한국어에서의 다양한 집단혐오용어에 관해서는 박대아 (2021)를 참고하라.

중이지만, 두 번째 교수의 발화는 집단혐오용어를 언급한 것이기에 혐오표현의 일종이 아니다. 이 논문에서 나는 집단혐오용어의 사용에 초점을 맞출 것이다.

이 논문의 목표는 오늘날 대한민국 언론 및 페미니즘 분야에서 유행<sup>3)</sup>하고 있는 집단혐오용어 사용(혹은 혐오표현 일반)에 관한 한 가지 도덕적 논제에 대해 검토하고 이 논제가 지지될 수 없다는 것을 보이는 것이다. 그 논제는 다음과 같다.<sup>4)</sup>

비대칭성 논제: 여성 일반 (혹은 소수자/약자 집단)에 대한 혐오 표현은 도덕적으로 정당하지 않지만, 남성 일반 (혹은 다수자/강자 집단)에 대한 혐오 표현은 도덕적으로 정당하다 (morally justified) (= 여성 일반에 대한 혐오 표현은 도덕적으로 허용가능하지 않지만, 남성 일반에 대한 혐오 표현은 도덕적으로 허용가능하다 (morally permissible)).<sup>5)</sup>

여기서 한 가지 주의할 점은 비대칭성 논제는 도덕 혹은 윤리에 관한 논제라는 것이다. 도덕철학에서는 흔히 비도덕성 (immorality), 불법성 (illegality), 범죄성 (criminality)을 구분한다. 먼저 어떤 행위는 비도덕적이지만 법에 의해 규제 받지 않기에 불법이 아니며 형사상 범죄도 아니다. 가령, 일상적 거짓말, 친구 사이 약속 안 지키기, 운전 중 손가락 욕하기 등이 여기에 해당한다. 이러한 행위들은 도덕적 비난의 대상이 되긴 하지만 법적 처벌의 대상이 되지는 않는다. 다음으로 어떤 행위는 불법이지만 형사상 범

3) 광재훈 (2017), 박경신 (2016), 서정호 외 (2019), 윤지영 (2015), 윤지선 (2019), 정희진 (2015), 홍성수 (2018, 2020), 여성가족부 (2019).

4) 내가 아는 한 이 논제는 대한민국 페미니즘 분야에서만 독특하게 나타나며 대한민국 이외 지역에서 이러한 도덕적 논제를 옹호하려는 두드러진 움직임은 발견할 수 없었다. 2019년에는 여성가족부에서 발간하는 초중고 성평등 교수 학습 지도안 사례집에서 이와 유사한 논제 (i.e., 여성혐오는 있지만 남성혐오는 없다)를 주장해 이슈가 된 적이 있다 (남정미 2019; 손국희 2019).

5) 이 논문에서 나는 “도덕적으로 정당하다” (morally justified)는 용어와 “도덕적으로 허용가능하다” (morally permissible)는 용어를 같은 의미를 가진 것으로 사용할 것이다.

죄가 아니다. 쓰레기 무단 투기, 과속, 신호위반 등이 여기에 해당한다.<sup>6)</sup> 마지막으로 어떤 행위는 비도덕적이며, 불법이며, 형사상 범죄이다. 도둑질, 사기, 강도, 강간, 살인 등이 여기에 해당한다. 비대칭성 논제는 도덕에 관한 논제이지 불법성 혹은 범죄성에 대한 논제가 아니다. 이는 다음의 두 가지 점을 함축한다. 첫째, 만약 비대칭성 논제가 옳다면, 여성(혹은 소수자/약자 집단)에 대한 혐오표현은 도덕적 비난의 대상이 되지만 남성(혹은 다수자/강자 집단)에 대한 혐오표현은 도덕적 비난의 대상이 되지 않을 것이다. 둘째, 만약 비대칭성 논제가 틀리고 보다 상식적인 입장(즉, 어떠한 혐오표현도 도덕적으로 정당하지 않다)<sup>7)</sup>이 옳다고 하더라도, 이것이 곧바로 혐오표현이 법적으로 금지되어야 한다는 것을 의미하는 것은 아니다.

## 2. “혐오표현”에 대한 두 가지 정의

비대칭성 논제를 위한 논증들을 살펴보기 이전에 먼저 “혐오표현”의 정의와 관련된 한 가지 논란에 대해 살펴보도록 하자. 어떤 이들은 일반적으로 받아들여지는 넓은 정의에 반대해 “혐오표현”에 대한 좁은 정의를 옹호한다.

“남자가 지배하는 사회에서 남성혐오 표현이라는 말은 존재할 수 없다. 혐오 표현은 혐오스러운 표현이나 누군가에 대한 혐오를 드러내는 표현이 아니다. 우리가 도의적으로 계도하고 또는 법적으로 제재하고자 하는 혐오 표현은 사회적 약자에 대한 차별과 폭력을 일으킬 위험이 명백하고 현존하는 표현을 말한다 ... 혐오 표현은 1948년 세계인권선언(UDHR)에 ‘차별의 선동’으로 정확히 정의되어 있다.” (박경신, 2016)

6) 이러한 행위들이 비도덕적인지는 맥락에 따라 변할 것이다. 가령, 쓰레기 무단 투기가 법으로 금지된 나라에서는 이러한 행위가 비도덕적이겠지만 혼자 무인도에 사는 사람이 쓰레기를 섬 아무 곳이나 버린다고 해서 이러한 행위가 비도덕적 행위로 간주되지는 않을 것이다.

7) 이에 대해서는 몇 가지 예외적인 사례들이 있을 수 있다. 보다 자세한 사항은 4절의 Q6를 참조하라.

“여성혐오와 달리 ‘남성혐오’가 실재하지 않는 상상된 개념이라는 것은, 남성 개인 혹은 남성 전체를 증오하는 사람이 실제로는 없다는 뜻이 아니다. 일부 위마드 회원 등 ‘남성을 증오하는 사람’은 실존한다. 그러나 중요한 것은 구체적 말이나 행동이 아니라, 그에 호응하는 제도, 문화 등 사회구조의 존재다 ... ‘혐오 표현’에 대한 문명화된 사회의 합의는 이렇다. 유엔의 시민적·정치적 권리 규약(ICCPR) 20조는 ‘차별, 적의 또는 폭력의 선동이 될 민족·인종·종교적 증오의 고취는 법률에 의해 금지돼야 한다’고 규정한다. 유엔 세계인권선언(HDR) 7조는 ‘모든 사람은 이 선언에 위반되는 어떤 차별로부터도, 그러한 차별의 선동으로부터도 동등한 보호를 받을 권리를 가진다’고 정하고 있다. 즉 남성이 사회에서 차별받는 약자일 때, 그 차별을 더욱 고취하기 위해 ‘강자인 여성’이 가하는 혐오 표현이 ‘남성 혐오’일 것이다. 남성에 대한 차별이 없는데 ‘차별의 선동’은 대체 어떻게 가능한가?” (곽재훈, 2017)

“홍성수 교수는 ‘혐오 표현이란 개인적인 감정이나 태도에서 오는 말이나 표현 전부를 뜻하지 않는다’며 ‘말이나 표현이 여성, 장애인, 이주 노동자 등 약자나 소수자를 향할 때, 특히 소수자들에 대한 부정적인 이미지나 편견이 담긴 동시에 이들이 겪는 차별을 고착화하는 경우라야 혐오 표현에 해당한다’고 설명했다.” (서정호, 강재연, & 강승민, 2019)

“소수자를 비하하는 발언, 혐오표현이 문제가 되는 것은 그 표현 자체 또는 발화자의 의도 때문이 아니라, 결과적으로 그 말이 사회에 주는 영향 때문이다 ... 국가인권위원회에서 발간한 ‘혐오표현 리포트’(2019)에는 혐오표현의 개념을 정의할 때 ‘차별을 정당화·조장·강화하는 효과’라는 구절이 굳이 추가되어 있다.” (홍성수, 2020)

혐오표현에 대한 넓은 정의와 좁은 정의는 각각 다음과 같이 정식화될 수 있다:

넓은 정의: 인종, 성, 국적, 성적 지향, 종교 등 개인의 정체성에 핵심적인 요소들을 근거로 개인 혹은 집단에 대한 혐오를 표현하거나 폭력을 선동하는 말, 글 또는 행동.

좁은 정의: 인종, 성, 국적, 성적 지향, 종교 등 개인의 정체성에 핵심적인 요소들을 근거로 소수자 집단 혹은 그 집단에 속한 개인에 대한 혐오를 표현함으로써 어떤 특정한 효과(e.g., 소수자에 대한 폭력 유발, 소수자들에 대한 차별 고착화 등)를 가지는 말, 글 또는 행동.

두 정의의 주된 차이는 소수자에 대한 혐오를 드러내는 표현인가, 그리고 그것이 어떤 특정한 효과를 가지는가에 있다. 좁은 정의에 따르면, 다수자에 대한 혐오를 드러내는 표현(e.g., “한남충”, “애비충”, “개독” 등)은 진정한 혐오표현이 아니다 (홍성수 2018; 여성가족부 2019). 또한 소수자에 대한 혐오를 드러내는 표현일지라도 어떤 특정한 효과를 가지지 않는다면 진정한 혐오표현이 아닐 것이다. 물론 위 저자들은 소수자에 대한 혐오를 드러내는 것과 그것이 소수자에 대해 어떤 특정한 효과를 가지는 것이 항상 함께 간다는 것을 암묵적으로 전제하고 있다. 그러나 우리는 소수자에 대한 혐오를 드러내지만 아무런 효과도 가지지 않는 경우를 쉽게 생각해 볼 수 있다. 가령, 한 사회에서 두 소수자 그룹(e.g., 미국 사회에서 한국인 이민자와 일본인 이민자, 게이와 레즈비언 등)이 서로에 대해 혐오표현을 사용한다면 이는 서로의 사회적 지위(social standing)에 별 영향을 주지 않을 확률이 높다. 만약 그렇다면 좁은 정의는 이들 사이의 혐오표현을 진정한 혐오표현이라고 규정하지 않을 것이다.<sup>8)</sup> 다른 한편, 넓은 정의에 따르면, 다수자에 대한 혐오를 드러내는 표현들, 그리고 소수자들 사이의 혐오 표현 등도 진정한 혐오표현으로 분류된다. 그렇다면 “혐오표현”이라는 용어에 대한 넓은 정의와 좁은 정의 가운데 어떤 정의를 채택해야 할까? 나는 이 논문에서 혐오표현에 대한 넓은 정의를 채택할 것이다. 그 이유는 (좁은 정의가 수많은 반직관적인 사례들에 직면하는 것 외에도) 좁은 정의를 옹호하는 입장들은 적어도 다음의 두 가지 오류를 범하고 있기 때문이다.

첫째, 이들 입장들은 혐오표현과 혐오표현의 금지를 혼동하고 있다. 가령,

8) 이 외에도 좁은 정의는 수많은 반직관적인 사례에 직면한다. 가령, 자식이 아버지에게 “애비충”, 어머니에게 “맘충”이라고 부르는 경우, 학생이 교수에게 모욕적인 욕설을 하는 경우 등도 좁은 정의에 따르면 혐오표현이 아닐 것이다.

이들은 UN 세계인권선언 (Universal Declaration of Human Rights) 및 시민적, 정치적 권리 규약 (International Covenant on Civil and Political Rights)을 인용하며 “혐오표현”이라는 용어가 UN에 의해 “차별 및 폭력의 선동”으로 정의되어 있다고 주장한다. 그러나 이들 문서에는 “혐오표현 (hate speech)”이라는 단어조차 등장하지 않는다.<sup>9)</sup> “혐오표현”에 대한 UN의 정의는 혐오표현에 대한 전략 및 행동계획(United Nations Strategy and Plan of Action on Hate Speech)이라는 문서에 명시적으로 등장하며 이 문서에서 UN은 넓은 정의를 채택하고 있다.<sup>10)</sup> 다시 말해, UN의 정의에는 소수자에 대한 혐오만이 진정한 혐오라든지 이것이 어떤 특정한 효과를 가져야 한다든지 하는 조건이 없다. 이들이 인용하는 UN 세계인권선언과 시민적, 정치적 권리 규약의 조항들은 언제 혐오표현이 표현의 자유에 의해 보호받지 못하는가에 대한 UN의 대답으로 이해되어야 하며, 이는 혐오표현에 대한 전략과 행동계획에서 혐오표현이 차별 및 폭력을 선동할 때를 혐오표현이 법적으로 제한 혹은 금지되어야 할 경우의 한 사례로 제시되고 있는 것과 일치한다.<sup>11)</sup>

---

9) “All are equal before the law and are entitled without any discrimination to equal protection of the law. All are entitled to equal protection against any discrimination in violation of this Declaration and against any incitement to such discrimination.” (UDHR, Article 7)

“Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law.” (ICCPR, Article 20(2))

10) “The United Nations Strategy defines ‘hate speech’ as ‘any kind of communication in speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor’.” (United Nations Strategy and Plan of Action on Hate Speech, p.8 and 10)

11) “The severest forms of hate speech *are prohibited* under international law. Such expressions include, most notably: (a) “direct and public incitement to commit genocide”, as defined by international criminal law; and (b) “any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence”, as defined in article 20(2) of the International Covenant on

둘째, 이들 입장들은 오스틴(Austin, 1975)의 발화수반행위 (illocutionary act)와 발화효과행위 (perlocutionary act)를 혼동하고 있다. 오스틴의 언화행위이론(speech act theory)에 따르면, 우리는 어떤 문장 혹은 언어표현을 발화(utter)할 때 발화행위 외 수많은 다른 행위들을 같이 수행할 수 있다. 예를 들어, A가 B에게 “오늘 저녁 9시까지 돌아와!”라고 말한다면 A는 발화행위 외에 명령이라는 행위도 함께 수행하는 것이다. 이 때 문장을 말하는 행위를 발화행위, 명령을 발화수반행위라고 부른다. 발화수반행위에는 명령(ordering) 외에도 약속(promising), 요청(requesting), 경고(warning) 등이 있으며 이런 행위들의 성공적 수행 여부는 이런 행위들이 청자에게 어떤 효과를 가지는지와 독립적이다. 다시 말해, 설사 B가 A의 명령을 무시하고 9시까지 돌아오지 않더라도 A는 여전히 명령이라는 행위를 수행한 것이다. 이에 반해 발화효과행위는 한 발화가 청자에게 어떤 의도한 효과를 가질 때에만 수행한 것으로 간주된다. 예를 들어, A가 B를 놀라게 하기 위해 B앞에 갑자기 나타나며 “악!”이라고 외쳤다고 가정해보자. 이 때 A가 놀라게 하기 (surprising)라는 행위를 수행하였느냐 여부는 이 발화가 청자인 B에게 어떤 효과를 야기하였는지에 달려있다. 만약 B가 실제로 놀라지 않았다면 A는 놀라게 하기라는 행위를 수행한 것이 아니다. 발화효과행위에는 놀라게 하기 외에도 설득하기 (persuading), 당황하게 하기 (embarrassing), 영감주기 (inspiring), 겁주기 (scaring) 등이 있다.

“혐오표현”에 대한 좁은 정의에 따르면, 어떤 언어행위가 혐오표현이기 위해서는 특정한 효과(e.g., 폭력 유발 혹은 차별의 고착화)를 가져야 한다. 이 입장은 혐오가 발화효과행위라는 것을 암묵적으로 전제한다. 그러나 위에서 제시한 발화수반행위와 발화효과행위 사이의 구분 기준에 따르면, 혐오는 발화수반행위이다. 예를 들어, 오늘날 미국사회에서 어떤 네오나치대원이 자기들끼리의 모임에서 “나는 유대인들을 혐오해”라고 말하거나 유대인에 대한 집단혐오용어(e.g., “Kike”)를 사용한다면, 이는 (이 발화의 효과



와 상관없이) 성공적으로 혐오를 표현한 것으로 받아들여질 것이다. 즉, 대부분의 사람들은 네오나치대원의 발언들이 어떠한 폭력이나 차별 효과를 가지지 못할 때에도 이들의 발화를 여전히 혐오표현으로 간주할 것이다. 만약 그렇다면, 이는 혐오가 발화수반행위라는 것을 의미한다.

좁은 정의의 한층 심각한 문제점은 이 정의가 네오나치대원의 발화에 도덕적 면죄부를 준다는 것이다. 이는 좁은 정의에 따르면 이들의 발화가 혐오표현이 아닌 것으로 간주될 확률이 높기 때문이다. 오늘날 미국사회에서 네오나치는 특별한 정치적, 사회적 영향력이 없는 소수그룹 (minority group)에 가깝다. 그들은 일반적으로 사회에서 배척당하며, 어떤 언론도 그들의 입장을 대변해주지 않으며, 대중에게 공공연하게 비웃음을 당하기도 한다. 그들은 자신의 입장을 공적으로 표현할 자유를 누리지만 그럴 때마다 언론의 감시 대상이 되며 사회의 비판과 우려가 뒤따른다. 따라서 네오나치대원이 유대인들에 대한 혐오표현을 쏟아낸다고 해서 이것이 실제로 유대인들에게 위협이 될 가능성은 매우 낮을 것이다. 만약 그렇다면, 좁은 정의는 네오나치의 발화를 혐오가 아닌 것으로 판단할 것이고 따라서 우리는 이들의 발화를 도덕적으로 문제 삼기 어렵게 될 것이다.

지금까지의 논의를 요약하면, 좁은 정의를 받아들일 좋은 법적 혹은 철학적 근거는 제시된 적이 없다. 따라서 나는 이 논문에서 상식적이고 보다 널리 받아들여지는 넓은 정의를 채택하도록 하겠다. 대신 3.1절에서 나는 좁은 정의를 위한 논증들이 비대칭성 논제를 위한 논증으로 재해석될 수 있음을 보인 후 이들을 반박하도록 하겠다.

### 3. 비대칭성 논제를 위한 논증들

비대칭성 논제는 여성 일반(혹은 소수자/약자 집단)에 대한 혐오 표현은 도덕적으로 정당하지 않지만, 남성 일반(혹은 다수자/강자 집단)에 대한 혐오 표현은 도덕적으로 정당하다는 입장이다. 이 입장을 위한 논증은 크게 사회적 지위에 근거한 논증과 부정의에 대한 저항에 근거한 논증이 있다.

### 3.1 사회적 지위에 근거한 논증들

먼저 사회적 지위에 근거한 논증은 사회적 지위와 혐오표현의 도덕적 정당성 사이에 밀접한 연관이 있다고 주장한다. 이와 관련한 두 가지 논증들에 대해 살펴보자.

논증 1<sup>12)</sup>

- P1. 집단으로서의 남성은 사회적 강자/다수자이고 집단으로서의 여성은 사회적 약자/소수자이다.
- P2. 만약 집단으로서의 남성이 사회적 강자/다수자이고 집단으로서의 여성이 사회적 약자/소수자라면, 남성의 여성에 대한 혐오 표현은 여성에 대한 직접적이고 구체적인 차별, 폭력으로 이어질 위험이 있지만 여성의 남성에 대한 혐오 표현은 그러한 가능성이 없다. (보다 일반적으로, 모든 집단 X,Y에 대해, 만약 X가 사회적 강자/다수자이고 Y가 사회적 약자/소수자라면, X의 Y에 대한 혐오표현은 Y에 대한 직접적이고 구체적인 차별, 폭력으로 이어질 위험이 있지만 Y의 X에 대한 혐오표현은 그러한 가능성이 없다)
- C1. 남성의 여성에 대한 혐오 표현은 여성에 대한 직접적이고 구체적인 차별, 폭력으로 이어질 위험이 있지만 여성의 남성에 대한 혐오 표현은 그러한 가능성이 없다
- P3. 혐오표현은 직접적이고 구체적인 차별, 폭력으로 이어질 위험이 있을 때 그리고 오직 그 때에만 도덕적으로 정당화되지 않고 나아가 법적으로 금지되어야 한다.
- C. 그러므로 남성의 여성에 대한 혐오표현은 도덕적으로 정당화되지 않고 법적으로 금지되어야 하지만, 여성의 남성에 대한 혐오표현은 허용가능하다.

P1은 경험적 진술에 가깝다. 즉, P1이 참 혹은 거짓임을 보이기 위해서는 많은 경험적인 데이터들이 필요할 것이다. 하지만 이는 철학자가 잘 할 수 있는 영역을 넘어서는 일이다. 그렇기에 나는 이 논문에서 P1이 참인지 아닌지에 대해서는 다루지 않겠다.

P3는 UN의 혐오표현에 대한 전략과 행동 계획, 미대법원의 수정헌법 제1

12) 논증1은 박경신 (2016)과 콕재훈 (2017)에 나타난 논증을 재구성한 것이다.

조 해석 (US Supreme Court case *Brandenburg v. Ohio*, 395 U.S. 444 (1969)), 밀(Mill)의 해악 원칙(harm principle) 등에 의해 광범위하게 지지된다. 그렇지만 한 가지 유의할 점은 P3에서 말하는 “직접적이고 구체적인 차별 및 폭력”은 굉장히 엄격한 개념이라는 것이다. 예를 들어, UN의 혐오 표현에 대한 전략과 행동 계획에 따르면, 혐오표현은 차별, 적대, 혹은 폭력에 대한 선동 (incitement to discrimination, hostility, or violence)일 때 제한되어야 하며, 이 때 선동(incitement)은 임박한 위험(imminent risk)을 야기할 수 있는 표현으로 정의된다. 또한 미대법원의 수정헌법 제1조 해석에 따르면, 표현의 자유는 “임박한 불법행위를 선동하거나 야기할 때” (directed to inciting or producing imminent lawless action) 제한될 수 있다. 보다 구체적으로 이를 밀의 옥수수 상인(corn dealer) 사례를 통해 살펴보자 (Mill 2011, ch.3). 밀에 따르면, “옥수수 상인이 가난한 이들을 굶주리게 만든다”는 주장은 임박한 위험을 야기하지 않을 맥락(가령, 이 주장을 책 등 출판물을 통해 인쇄하는 경우)에서는 허용되지만, 옥수수 상인의 집 앞에 성난 군중들이 모여있고 이들이 언제 폭발할지 모르는 맥락(즉, 임박한 위험을 야기할 수 있는 맥락)에서는 허용되지 않는다. 이는 후자의 맥락에서 “옥수수 상인이 가난한 이들을 굶주리게 만든다”는 발화가 성난 군중들을 선동해 옥수수 상인에게 임박한(imminent) 위험을 초래할 수 있기 때문이다. 여기서 중요한 점은 이 발화의 허용 여부는 옥수수 상인과 가난한 이들 사이의 사회적 관계에 의해 결정되는 것이 아니라 이 발화가 임박한(imminent) 위험을 초래하는가에 달려 있다는 것이다. 옥수수 상인이 부자이고 사회적인 강자의 위치에 있고 가난한 이들이 사회적 약자의 위치에 있다고 할지라도 위의 발화가 옥수수 상인에게 직접적이고 임박한 위험을 초래할 경우 위의 발화는 허용되지 않는다.

“직접적이고 구체적인 폭력, 차별”이라는 용어가 이렇게 엄격한 방식으로 해석될 때 P3는 광범위한 지지를 받을 수 있다. 그러나 이 용어가 이런 식으로 해석될 때 P2 (= “만약 집단으로서의 남성이 사회적 강자/다수자이고 집단으로서의 여성이 사회적 약자/소수자라면, (1) 남성의 여성에 대한 혐오

표현은 여성에 대한 직접적이고 구체적인 차별, 폭력으로 이어질 위험이 있지만 (2) 여성의 남성에 대한 혐오 표현은 그러한 가능성이 없다”)는 거짓이 된다. 이는 P3를 지지하기 위해 사용되었던 밀의 옥수수 상인 사례가 P2를 논박하는데 사용될 수 있기 때문이다. (1) 밀에 따르면, “옥수수 상인이 가난한 이들을 굶주리게 만든다”는 주장은 임박한 위험을 야기하지 않을 보통의 맥락(가령, 출판물을 통해 이 주장이 인쇄되는 경우)에서는 허용가능하다. 이 주장을 온라인으로 출판하는 것과 오프라인으로 출판하는 것 사이에 본질적인 차이가 있지는 않을 것이다. 따라서 밀의 이론을 오늘날의 사례에 적용해보면 온라인에서 남성이 여성에 대한 혐오 표현을 사용하는 경우 이는 밀이 말하는 의미에서 직접적이고 임박한 위험을 초래하지 않을 것이다. 그렇기에 집단으로서의 남성이 사회적 강자/다수자이고 집단으로서의 여성이 사회적 약자/소수자이기 때문에, 남성의 여성에 대한 혐오 표현은 여성에 대한 직접적 차별, 폭력으로 이어질 위험이 있다는 주장은 이 맥락에서 거짓이 된다. (2) 반대로 여성의 남성에 대한 혐오표현은 남성에 대한 직접적이고 임박한 폭력으로 이어질 위험이 없다는 주장도 거짓이다. 가령, 화난 시위대 앞에서 남성 일반에 대한 혐오표현으로 폭력을 선동하는 맥락에서는 남성에 대한 혐오표현 역시 직접적이고 임박한 위험을 초래할 수 있다. 따라서 집단으로서의 남성이 사회적 강자/다수자이고 집단으로서의 여성이 사회적 약자/소수자이기에, 여성의 남성에 대한 혐오 표현은 직접적이고 임박한 폭력으로 이어질 가능성이 없다는 주장 역시 거짓이 된다.

이로부터 우리가 알 수 있는 사실은 사회적 지위와 혐오표현의 정당성 사이에 직접적인 연관이 없다는 것이다. 임박한 위험의 존재 여부는 혐오표현의 도덕적/법적 정당성과 연관이 있을 수 있다. 그렇지만 이러한 기준은 P2를 거짓으로 만든다. 왜냐하면 사회적 강자/다수자에 대한 혐오표현 역시 맥락에 따라 직접적이고 임박한 위험을 야기할 수 있기 때문이다. 다음으로 사회적 지위에 근거한 두 번째 논증을 살펴보자.

논증 2<sup>13)</sup>

- P1. 집단으로서의 남성은 사회적 강자/다수자이고 집단으로서의 여성은 사회적 약자/소수자이다.
- P2. 만약 집단으로서의 남성이 사회적 강자/다수자이고 집단으로서의 여성이 사회적 약자/소수자라면, 남성의 여성에 대한 혐오 표현은 여성의 사회적 지위 (social standing)를 약화시킬 힘이 있지만 여성의 남성에 대한 혐오 표현은 그러한 힘이 없다. (보다 일반적으로, 모든 집단 X,Y에 대해, 만약 X가 사회적 강자/다수자이고 Y가 사회적 약자/소수자라면, X의 Y에 대한 혐오표현은 Y의 사회적 지위를 약화시킬 힘이 있지만 Y의 X에 대한 혐오표현은 그러한 힘이 없다)
- C1. 남성의 여성에 대한 혐오 표현은 여성의 사회적 지위를 약화시킬 힘이 있지만 여성의 남성에 대한 혐오 표현은 그러한 힘이 없다
- P3. 혐오표현은 대상의 사회적 지위를 약화시킬 힘이 있을 때 그리고 오직 그 때에만 도덕적으로 정당화되지 않는다
- C. 그러므로 남성의 여성에 대한 혐오표현은 도덕적으로 정당화되지 않지만, 여성의 남성에 대한 혐오표현은 허용가능하다.

논증 2의 P2는 논증 1의 P2보다 얼핏 보기에 조금 더 설득력 있어 보인다. 예를 들어, 부자들의 가난한 자들에 대한 혐오표현은 가난한 자들에 대한 부정적인 편견을 강화함으로써 그들의 처지를 더욱 안 좋게 만들 가능성이 있지만 (가령, 기부금 액수를 줄인다든지, 고용을 최소화한다든지 등) 가난한 자들의 부자들에 대한 혐오표현은 부자들의 처지를 안 좋게 만들 수 있는 방법이 별로 없어 보인다. 그러나 한 가지 흥미로운 점은 논증 2의 P2는 적어도 어떤 페미니즘과 양립불가능하다는 것이다. 예를 들어, 윤지선 (2019)은 여성들의 남성들에 대한 혐오표현은 기존 질서에 대한 전복의 효과를 가지기에 정당하다고 주장한다. 만약 그렇다면, P2는 거짓일 것이다. 우리는 이 논증을 3.2 절에서 자세히 살펴볼 것이다. 여기서 내가 강조하고

13) 논증 2는 서정호, 강재연, 강승민 (2019)과 홍성수 (2018, 2020)에서 나타난 논증을 재구성한 것이다. 위 문헌들에서 언급된 “차별을 고착화하는 효과”라는 개념을 사회적 지위를 흔들 수 있는 힘이 있는가 여부로 포착하고자 하였다.

자 하는 바는 적어도 이 논증을 받아들이는 페미니스트들은 논증 2를 받아들일 수 없다는 것이다.

논증 2의 더욱 심각한 문제점은 P3에 있다. P3는 쌍조건문(biconditional)으로서 두 가지 부분으로 나뉘어질 수 있다. 첫째, (1) 만약 어떤 혐오표현이 대상의 사회적 지위를 약화시킬 힘이 있다면, 이러한 혐오표현은 도덕적으로 정당화되지 않는다 (이 주장은 “만약 어떤 혐오표현이 도덕적으로 정당화된다면, 이러한 혐오표현은 대상의 사회적 지위를 약화시킬 수 있는 힘이 없다”는 주장의 대우로서 양자는 동치이다). (1)에 대한 반례로는 미국 백인 부유층/권력가의 나치나 KKK에 대한 혐오를 들 수 있다. 어떤 사람들은 네오나치나 KKK에 대한 혐오표현은 도덕적으로 정당하다고 주장한다 (DiFranco 2017). 만약 이들의 입장이 옳다면, 미국 백인 부유층/권력가의 나치나 KKK에 대한 혐오는 이들의 사회적 지위를 약화시킬 힘이 있지만 여전히 이들에 대한 혐오표현은 정당화될 것이다.

둘째, (2) 만약 어떤 혐오표현이 대상의 사회적 지위를 약화시킬 힘이 없다면, 이러한 혐오표현은 도덕적으로 정당화될 수 있다. (2)에 대한 반례로는 네오나치나 KKK의 유대인 혐오를 들 수 있다. 오늘날의 미국사회에서 네오나치나 KKK는 특별한 정치적, 사회적 영향력이 없는 소수 그룹(minority group)에 가깝다. 그들은 사회에서 배척당하며 그들의 목소리에 귀기울여주는 언론은 찾아보기 힘들고 사람들은 그들을 공공연하게 비난하고 조롱하기도 한다. 따라서 네오나치나 KKK가 부유하고 잘 교육받은 유대인들에 대한 혐오표현을 쏟아낸다고 해서 네오나치의 사회적 지위가 올라가고 유대인들의 사회적 지위가 약화될 가능성은 거의 없을 것이다. 그럼에도 불구하고 많은 사람들은 네오나치나 KKK의 유대인들에 대한 혐오표현은 도덕적으로 정당하지 않다고 판단할 것이다. 따라서 다른 집단의 사회적 지위를 약화시킬 힘이 있느냐 여부는 혐오표현의 도덕적 정당성과 직접적인 연관이 없다.

### 3.2 부정의에 대한 저항에 근거한 논증

3.1절에서 우리는 사회적 지위에 근거해 비대칭성 논제를 옹호하고자 하는 두 가지 논증을 살펴보았고, 두 논증 모두 심각한 반례에 직면한다는 것을 살펴보았다. 이 절에서 우리는 부정의에 대한 저항에 근거해 비대칭성 논제를 옹호하고자 하는 시도를 살펴보겠다.

“‘한남유충과 한남충, 관음충’은 대한민국 남성에 대한 혐오용어로서 금기시되어야 하는가? 그렇다면 이때껏 가부장제 남성중심주의 사회에서 여성을 표적으로 했던 수많은 혐오용어들-김치녀, 된장녀, 꽃뱀, 맘충, 립나무-의 고안자와 사용자는 누구였는가? 필자가 보기에 ‘한남유충과 한남충, 관음충’의 용어들이 방사하고 있는 효과는 다음과 같다. 첫째, 해당 용어들이 표적하고 있는 계층성이 더 이상 사회적 소수자가 아닌, 동일연령 대비 (잠재적, 현실적) 사회적 다수자인 남성 집단이라는 점에서 전복적 효과를 띤다. ... 둘째, 우리는 해당 용어들의 고안자와 사용자가 누구인가에 주목할 필요가 있다. 언어와 법률, 학문과 의학, 예술과 정치, 담론과 같은 분야가 남성중심주의라는 이데올로기에 의해 영향을 받기에 혐오용어를 포함한 새로운 용어를 개발하고 고안하는 자들은 주로 남성 계층이었던 것이 사실이다. 그러나 2015년 이후 새로운 페미니즘 물결의 도래로 인해 여성혐오용어의 자장 아래 자기검열과 자기혐오의 굴레에 있던 한국여성들이 남성계층의 파괴성과 착취성, 오만성을 저격하는 반격의 용어를 스스로 고안해내고 적극적으로 사용함으로써 여성혐오문화의 지형에 균열과 저항의 가능성을 기입할 수 있었다.”(윤지선 2019)

이 논증의 핵심은 (적어도 일부의) 남성들이 먼저 여성에 대한 집단혐오용어를 사용하였고, 여성들의 남성에 대한 집단혐오용어 사용은 이에 대한 반작용이기에 정당하다는 것이다. 이 논증은 다음과 같이 정식화될 수 있다.

#### 논증 3

- P1. 적어도 일부의 남성들이 먼저 여성혐오표현을 사용해 여성들을 공격하였다.
- P2. 만약 적어도 일부의 남성들이 먼저 여성혐오 표현을 사용해 여성들을 공격하였다면, 여성들은 이에 대한 반작용으로 (부정의에 대한 전복 및 저항의 의미로) 남성혐오표현을 사용하는 것이 도덕적으로 정당하다 (보다 일반적

으로, 모든 집단 X,Y에 대해, 만약 X에 속한 적어도 일부의 사람들이 Y를 공격하였다면, Y는 이에 대한 반작용으로 (부정의에 대한 전복 및 저항의 의미로) X에 속한 사람들을 무차별적으로 공격하는 것이 도덕적으로 정당하다)

C. 그러므로 여성들의 남성혐오표현 사용은 도덕적으로 정당하다.<sup>14)</sup>

만약 여성에 대한 집단혐오용어가 시기적으로 먼저 등장하였고 남성에 대한 집단혐오용어가 이에 대한 반작용으로 등장했다면, P1은 참일 수 있다. 저자에 따르면, “한남충”, “한남유충”과 같은 표현은 “김치녀”, “된장녀”와 같은 여성에 대한 집단혐오용어에 대한 반작용으로 등장했다고 한다. 그러나 이 주장에 대해 한가지 의심스러운 점은 한국어에 “한남유충”에 대응하는 집단혐오용어는 존재하지 않는 것 같다는 점이다. 만약 그렇다면, 이는 적어도 “한남유충”이라는 표현은 여자 어린이에 대한 집단혐오용어에 대한 반작용으로 등장한 것이 아니라는 것을 시사한다.<sup>15)</sup> 그러나 P1은 경험적 진술이기에 나는 이 논문에서 이 전제의 참 거짓 여부에 개입하지 않았다.

논증 3의 문제점은 P2에 있다. 이를 다음 논증과의 비교를 통해 살펴보도록 하자. 논증 3\*는 오사마 빈 라덴(Osama bin Laden)이 테러리즘<sup>16)</sup>을 용

14) 위 인용문에서 명시적으로 드러나지는 않지만 저자는 남성들의 여성혐오표현 사용은 도덕적으로 정당하지 않다는 것을 암묵적으로 전제하고 있는 것 같다. 논의의 편의상 나는 이 절에서 여성들의 남성혐오표현 사용에만 초점을 맞출 것이다.

15) 이에 대해 어떤 이는 “룸나무”가 “한남유충”에 대응하는 표현이 아니냐고 의문을 제기할 지도 모른다. 하지만 “룸나무”는 여자 청소년을 가리키는 표현인 반면 “한남유충”은 남자 어린 아이 일반(즉, 남자 어린이와 남자 청소년)을 가리키는 표현이다. 따라서 둘 다 성인을 가리키는 “김치녀”- “한남(충)”의 쌍과는 차이가 있다. 또한 다음의 인용문은 “한남유충”이 어떤 혐오표현에 대한 반응으로 등장했다기 보다는 “한남충”의 이전 단계로 가정되었다는 것을 시사한다.

“한남유충”이란 용어의 고안은 20대 여성들이 던진 다음과 같은 질문으로부터 추동되었다. 과연 한남충은 어떻게 탄생하는가? 현재의 한남(성)충들은 성인기가 도래하자 갑자기 출현한 군집체인가, 아니면 한국여성들이 미성년자일 때부터 학교 교실에서, 학원에서 포진된 채 여성혐오놀이에 심취해있던 그 남아 군집체들의 연장성인가?” (윤지선 2019, p.263)

16) “테러리즘”의 정의에는 넓은 정의와 좁은 정의가 있다 (Primoratz 2018; Walzer 2004 Ch4&10). 먼저 좁은 정의에 따르면, 테러리즘은 (i) 어떤 정치적 목적(e.g., 공



호하기 위해 사용했던 논증이다. 오사마 빈 라덴은 9.11 공격 직전의 한 연설에서 80년에 이르는 억압과 굴종의 역사에 대해 이야기하며 테러리즘을 옹호한다 (Walzer 2004, p.133). 즉, 미국이 먼저 중동문제에 개입했고 자신의 테러 행위는 이러한 부정의에 대한 반작용으로 저항의 의미를 지닌다는 것이다.

### 논증 3\*

- P1. 적어도 일부의 미국인들이 먼저 이슬람인 및 이슬람 세계를 공격하였다.
- P2. 만약 적어도 일부의 미국인들이 먼저 이슬람인 및 이슬람 세계를 공격하였다면, 이슬람인들은 이에 대한 반작용으로 (부정의에 대한 전복 및 저항의 의미로) 미국인들을 무차별적으로 공격하는 것이 도덕적으로 정당하다 (보다 일반적으로, 모든 집단 X,Y에 대해, 만약 X에 속한 적어도 일부의 사람들이 Y를 공격하였다면, Y는 이에 대한 반작용으로 (부정의에 대한 전복 및 저항의 의미로) X에 속한 사람들을 무차별적으로 공격하는 것이 도덕적으로 정당하다).
- C. 그러므로 이슬람인들이 미국인들을 무차별적으로 공격하는 것은 도덕적으로 정당하다.

대부분의 사람들은 논증3\*에서 P2를 받아들이지 않을 것이다. 이는 P2가 무고한 (innocent) 이들에 대한 공격까지도 정당화하기 때문이다. 예를 들어, A마을과 B마을 사람들이 있다고 가정해보자. A 마을에 속한 몇몇 사람들이 B마을 사람들에게 무차별적으로 콜라를 뿌리는 잘못을 저질렀다고 가정해보자. 이런 상황에서 B 마을 사람들은 이에 대한 반작용으로 A마을 사람 아무에게나 무차별적으로 콜라를 뿌릴 자격을 가지는가? 그렇지 않을 것이다. B 마을 사람들은 A 마을 사람들 중 콜라를 뿌린 범인을 색출해 이들

---

포의 확산)을 달성하기 위해 (ii) 무고한 사람을 공격하는 (iii) 폭력행위이다. 넓은 정의에 따르면, 테러리즘은 (i) 어떤 정치적 목적 달성을 위한 (ii) 폭력 행위이다. 둘의 주된 차이는 무고한 사람에 대한 공격을 정의의 일부로 포함하느냐 여부에 있다. 윤리학에서는 주로 좁은 정의를 채택하며 이는 테러리즘과 정치적 암살을 구분하기 위함이다. 예를 들어, 만약 넓은 정의가 옳다면 안중근, 윤봉길 등 많은 독립운동가들은 테러리스트로 분류될 것이다. 그러나 이는 반직관적이다.

을 처벌할 권리를 가지지만 단순히 어떤 사람이 A 마을 사람이라는 이유로 그를 처벌할 권리를 가지지는 않을 것이다. 이는 대부분의 A 마을 사람은 콜라를 뿌리지 않은 무고한 시민들이기 때문이다. (참고로 여기서 A마을과 B마을 사이의 사회적 차이(소득 등)는 도덕적 차이를 만들지 못한다. 설사 A마을은 부유한 마을이고 B마을은 가난한 마을이라는 가정을 추가하더라도 여전히 B마을 사람들이 A마을 사람들에게 무차별적으로 콜라를 뿌릴 권리를 가지지는 않는다.)

논증 3과 논증 3\*는 정확히 동일한 논리적 구조를 가지고 있다. 따라서 만약 위의 사례가 논증 3\*의 전제 2가 거짓이라는 것을 보여준다면 이는 논증 3의 전제 2 역시 거짓이라는 것을 보여줄 것이다. 논증 3을 받아들이는 페미니즘과 테러리즘의 공통점은 무고한 사람에 대한 공격까지도 정당화한다는 데 있다. 대부분의 사람들은 테러리즘의 이러한 측면 때문에 테러리즘을 정당화하기 어렵다고 한다. 대부분의 미국 시민들은 중동 문제에 직접적으로 개입하지 않았다. 또한 미국 시민들은 중동 문제에 책임 있는 사람들(e.g., 정치인, 고위 공직자, 군인 등)을 직접적으로 통제할 수단을 가지고 있지 않다. 테러리즘에 반대하는 이 논증이 설득력이 있다고 생각하는 사람들은 같은 이유로 논증 3을 받아들이는 페미니즘에 반대할 것이다.

이러한 반론에 대해 테러리스트들은 모든 미국인은 부정의에 대한 방관자 혹은 암묵적 동조자이기에 어떠한 미국인도 무고하지 않고 따라서 미국인들을 무차별적으로 공격하는 것이 정당하다고 주장할 수 있다. 마찬가지로 어떤 페미니스트들은 모든 남자는 가부장제로 인해 발생한 부정의에 대한 방관자 혹은 암묵적 동조자이기에 어떠한 남자도 무고하지 않고 따라서 모든 남자에 대한 무차별적인 혐오표현이 정당하다고 주장할 수 있다.

하지만 대부분의 사람은 이러한 주장이 설득력이 있다고 생각하지 않을 것이다. 이는 일반 시민에 대해 너무 과도한 도덕적 요구를 하는 것이기 때문이다. 가령, 한국이 60-70년대 베트남 전쟁에 참여한 것을 이유로 어떤 테러리스트가 한국인 관광객들을 단지 한국인이라는 이유로 무차별적으로 공격한다면 대부분의 사람들은 이러한 행위가 정당하다고 생각하지 않을 것

이다. 마찬가지로 만약 어떤 페미니스트가 길거리에 지나다니는 아무 한국 남성들에 대해 단지 남자라는 이유로 언어적으로 공격한다면 대부분의 사람들은 이러한 행위가 정당하다고 생각하지 않을 것이다. 더욱 분명한 사례로는, 만약 어떤 페미니스트가 한국 남자 어린아이에 대해 단지 남자 어린아이라는 이유로 집단혐오용어인 “한남유충”을 사용해 공격한다면 이를 도덕적으로 정당화하기는 매우 어려울 것이다.

지금까지의 논의를 요약하면 논증 3을 수용하는 페미니즘과 테러리즘은 동일한 논리적 구조를 가지고 있기에 둘은 패키지딜(package deal)이다. 즉, 논증3에 의해 남성 일반 혹은 다수자 일반에 대한 혐오가 도덕적으로 정당화된다면, 테러리즘도 정당화된다. 그러나 대부분의 사람들은 테러리즘이 정당하지 않다고 생각할 것이고 그런 사람들에게 논증 3은 설득력이 없을 것이다.

마지막으로 나는 두 가지 예상 가능한 반론들에 대해 살펴보고 이 절을 마무리짓도록 하겠다. 첫째, 어떤 페미니스트들은 논증 3의 정식화에 반대할 수 있다. 그들은 오직 일부의 미국인만이 중동 문제에 개입하였지만, 대부분의 성인 남성들이 여성 일반에 대한 집단혐오용어를 사용하였기에 페미니즘과 테러리즘과의 비유는 적절하지 않다고 주장할 수 있다. 이러한 주장을 받아들인다면 논증 3은 다음과 같이 수정될 수 있다.

#### 논증 4

- P1. 대부분의 남성들이 먼저 여성혐오 표현을 사용해 여성들을 공격하였다.
- P2. 만약 대부분의 남성들이 먼저 여성혐오 표현을 사용해 여성들을 공격하였다면, 여성들은 이에 대한 반작용으로 (부정의에 대한 전복 및 저항의 의미로) 남성혐오 표현을 사용하는 것은 도덕적으로 정당하다 (보다 일반적으로, 모든 집단 X,Y에 대해, 만약 X에 속한 대부분의 사람들이 Y를 공격하였다면, Y는 이에 대한 반작용으로 (부정의에 대한 전복 및 저항의 의미로) X에 속한 사람들을 무차별적으로 공격하는 것이 도덕적으로 정당하다)
- C. 그러므로 여성들의 남성혐오 표현은 도덕적으로 정당하다.

논증 4에서 P1은 경험적 명제이다. 그런데 이는 단지 일부가 아니라 대다

수의 남성이 집단혐오용어를 사용했다는 주장이므로 논변 3의 P1보다 증명하기가 더 어려울 것이다. 그렇지만 여기서는 P1이 참이라고 가정해보자. 설사 P1이 참이라고 할지라도 P2는 여전히 참이 아니다. 예를 들어, A 마을에 속한 대부분의 혹은 모든 성인들이 B마을 사람들에게 무차별적으로 콜라를 뿌리는 잘못을 저질렀다고 가정해보자. 이런 상황에서 B 마을 사람들 중 일부가 이에 대한 반작용으로 A마을에 사는 어린이에게 콜라를 뿌린다면 이들의 행동은 정당한가? 그렇지 않을 것이다. 즉, B 마을 사람들은 A 마을 사람들 중 콜라를 뿌린 성인들을 처벌할 권리를 가질지 모르지만 잘못에 직접 가담하지 않은 어린이를 단지 A마을에 산다는 이유만으로 처벌할 권리를 가지지는 않는다. 이를 집단혐오용어 사용에 적용해보면, 만약 대부분의 혹은 모든 성인 남자들이 여성에 대한 집단혐오용어를 사용했다는 주장이 옳다면, 이는 “한남충”과 같이 성인 남성 일반에 대한 집단혐오용어 사용은 도덕적으로 정당화할 수 있을지 모르나 여전히 “한남유충”이라는 남자 어린이에 대한 집단혐오용어 사용은 정당화하지 않을 것이다. 이 사례는 단순히 숫자의 차이가 도덕적인 차이를 만들지는 않는다는 것을 보여준다. 이 사례에서 도덕적으로 관련한 요소는 숫자보다는 무고한 사람에 대한 공격을 포함하느냐 여부이다.

둘째, 어떤 페미니스트들은 논증 3과 논증 3\*에서 각각 문제가 되는 공격의 종류가 다르기 때문에 (즉, 논증 3에서 문제가 되는 공격은 언어적 공격인 반면 논증 3\*에서 문제가 되는 공격은 주로 물리적 공격이기 때문에) 논증 3\*의 P2는 거짓인 반면 논증 3의 P2는 참이라고 주장할 수 있다.<sup>17)</sup> 이 반론에 전제된 기본적인 아이디어는 물리적 공격이 초래하는 해악은 심각하기에 비도덕적인 것으로 간주되어야 하지만 언어적 공격이 초래하는 해악은 크지 않기 때문에 비도덕적인 것으로 간주될 필요가 없다는 것이다.

그러나 이런 주장은 어떤 행위가 일으키는 해악의 사소함과 행위의 비도덕성을 혼동하는 것이다. 어떤 행위가 일으키는 해악이 사소하다고 해서 그것이 그 행위를 도덕적으로 관찮은 것으로 만드는 것은 아니다. 예를 들어,

17) 이러한 반론을 제기해준 익명의 심사위원께 감사드린다.

A가 부자인 B의 지갑에서 5만원을 훔쳤다고 가정해보자. 5만원은 부자인 B에게는 사소한 금액일 수 있다. 심지어 B는 영원히 A가 훔쳤다는 것을 알아차리지 못할 수도 있다. 그렇다고 할지라도 A의 행위는 여전히 도둑질이며 따라서 비도덕적 행위이다. 왜냐하면 도둑맞은 사람이 인지하든 그렇지 못하든 그 사람의 권리(= 사유재산권)가 침해되었기 때문이다. 무고한 사람에 대한 언어적 공격이 일으키는 사회적 해악이 사소하기 때문에 도덕적으로 괜찮다는 주장은 사소한 금액의 도둑질은 도덕적으로 괜찮다고 주장하는 것과 같다. 도둑질이 타인의 권리를 침해하기에 도덕적 잘못된 것처럼, 무고한 사람에 대한 언어적 공격 역시 타인의 권리를 침해한다면 비도덕적인 것으로 간주되어야 할 것이다. 이 권리가 구체적으로 무엇인지는 다음 절에서 자세히 살펴보도록 하겠다.

#### 4. 질문과 답변

우리는 지금까지 비대칭성 논제를 옹호하는 여러 논증들을 살펴보았고 이들이 모두 실패한다는 것을 알 수 있었다. 이 중 논증 3은 논증 3\* 즉 테러리즘을 옹호하는 논증과 동일한 논리적 구조를 가지고 있기에 도덕철학적 관점에서 볼 때 흥미로운 점이 있다. 이 절에서는 “논증 3을 수용하는 페미니즘”, “테러리즘”, “혐오표현”, “도덕적 정당화”와 같은 여러 개념들 사이의 관계를 명확하게 하기 위해 합리적으로 제기될 수 있는 몇 가지 질문들을 고려해보고 이에 대한 답을 하도록 하겠다.

Q1) 논증 3을 받아들이는 페미니즘은 테러리즘을 도덕적으로 정당화하는가? 그렇다. 이는 둘 다 같은 도덕적 원칙 (즉, 모든 집단 X,Y에 대해, 만약 X에 속한 적어도 일부의 사람들이 Y를 공격하였다면, Y는 이에 대한 반작용으로 (부정의에 대한 전복 및 저항의 의미로) X에 속한 사람들을 무차별적으로 공격하는 것이 도덕적으로 정당하다)에 의존하고 있기 때문이다. 이 전제에 의존해 논증 3을 받아들이는 사람은 논증 3\*도 받아들여야 할 것이고 이는 테러리즘을 정당화 하는 게 된다.

Q2) 논증 3을 받아들이는 페미니즘은 그 자체로 곧장 테러리즘을 함축하는가? 아니다. 테러리즘은 실제 폭력<sup>18)</sup>을 행사했느냐가 기준이다. 가령, 어떤 철학자가 테러리즘이 도덕적으로 정당하다고 주장한다고 할지라도 그가 실제로 무고한 사람들에게 대해 어떤 폭력도 행사하지 않았다면 그는 테러리스트가 아닐 것이다. 마찬가지로 어떤 페미니스트가 논증 3을 받아들인다고 할지라도 실제로 혐오표현을 사용하지 않았다면 테러리스트와 같은 종류의 도덕적 잘못<sup>19)</sup>을 저지른 것은 아니다.

Q3) 논증 3을 받아들이는 페미니즘은 테러리즘과 양립가능한가? 그렇다. 테러리즘에 대한 일반적인 정의에 따르면, 테러리즘은 어떤 정치적 목적(e.g., 공포의 확산, 이데올로기 전파 등)을 위해 무고한 사람에 대해 무차별적 폭력을 가하는 행위이다 (Walzer 2004, Ch.4&10). 어떤 페미니스트가 의도적으로 특정한 정치적 목적(e.g., 특정 이데올로기의 확산, 남성에 대한 혐오의 확산, 여성 집단 내의 공포의 확산 등)을 위해 무고한 사람들 (가령, 혐오 표현에 반대하는 성인 남성, 남자 어린이 등)에게도 무차별적으로 혐오표현을 사용한다면 이는 테러리즘의 한 종류라고 볼 수 있다.

Q4) 모든 테러리즘은 비도덕적인가? 테러리즘은 일반적으로 비도덕적인 것으로 간주된다. 왜냐하면 테러리즘은 무고한 사람에 대한 공격을 포함하기 때문이다. 그렇기에 이 질문에 대한 대답은 무고한 사람에 대한 공격이 항상 비도덕적인가에 달려있다. 무고한 사람에 대한 공격이 항상 비도덕적인가 아니면 예외가 있을 수 있는가는 논란의 여지가 있다. 그렇지만 대부분의 철학자들은 무고한 사람에 대한 공격을 정당화하는 것이 매우 어려운

18) 나는 여기서 “폭력”을 물리적 폭력만이 아니라 정서적, 언어적 폭력도 포함하는 넓은 의미로 사용하고 있다.

19) 여기서 우리는 같은 종류의 도덕적 잘못과 같은 정도의 도덕적 잘못을 구별해야 한다. A가 B에게 1만원을 훔치는 경우와 1000만원을 훔치는 경우를 생각해보자. 두 경우 A는 같은 종류의 도덕적 잘못을 저질렀지만 같은 정도의 도덕적 잘못을 저지른 것은 아닐 것이다. 이와 마찬가지로, A가 B에게 인종차별적 편견에 근거해 심각한 물리적 해를 끼치는 경우와 인종차별적 농담을 하는 경우를 비교해보면, A는 두 경우 같은 종류의 도덕적 잘못을 저질렀지만 같은 정도의 도덕적 잘못을 저지른 것은 아닐 것이다.

일이라는 데 동의할 것이다. 가령, 왈저는 나치의 승리가 임박하거나 대규모 인종청소가 예상되는 등 (i) 최고비상사태 (supreme emergency)에서 적국 민간인에 대한 공격이 이러한 임박한 위협을 피할 수 있는 (ii) 유일한 수단이라면 무고한 민간인에 대한 공격이 예외적으로 정당화될 수 있다고 한다 (Walzer 2006, Ch.16). 하지만 그는 테러리즘은 역사상 임박한 위협을 피하기 위한 수단으로 사용된 적이 없으며 항상 어떤 정치적 목적(e.g., 이데올로기의 확산, 공포의 확산 등)을 달성하기 위한 수단으로 사용되었다고 한다. 또한 대부분의 경우 테러리즘은 간편한 수단으로 이용되지 최후의 수단이라거나 유일한 수단으로 이용되지 않기에 위의 두 조건을 만족시키지 못한다고 한다. 왈저에 따르면, 테러리즘이 유일한 수단이라고 주장하는 이들은 흔히 두 가지 종류의 약함을 혼동하고 있다: 적국에 비해 약한 자신의 힘, 그리고 자신의 약한 대중적 영향력. 테러리즘을 유일한 수단으로 만드는 것은 후자의 약함이다. 만약 대중을 자신의 편으로 만들 수 있다면, 비폭력 저항 운동, 총파업, 대규모 시위, 특정 인물만을 노린 게릴라전 등 테러리즘 외에도 수많은 대안들이 있을 수 있기 때문이다 (Walzer 2004, Ch.4).

왈저의 논의를 여성운동에 적용해 보자. 즉, 여성운동가들은 자신의 목적을 위해 무고한 사람을 공격(e.g., “한남유충”이라는 용어를 사용해 남자 어린이를 공격)할 권리를 가지는가? 이에 대한 대답은 “아니오”일 것이다. 이는 왈저가 제시한 두 가지 조건을 모두 만족시키지 못하기 때문이다. 먼저 “한남유충”이라는 용어를 사용하는 것은 특정한 정치적 목적을 달성하기 위함일 수는 있으나 임박한 위협을 피하기 위한 것은 아니다. 또한 여성운동가들이 대중들을 설득할 능력이 있다면 (특히나 민주주의 사회에서) 혐오 표현을 통해 무고한 사람을 공격하지 않고도 자신의 목적을 달성할 수 있는 수많은 수단들(e.g., 투표 독려, 온라인 여론전, 언론을 통한 호소, 청와대 신문고, 국민청원, 불매운동, 대규모 시위, 파업 등)이 있을 것이다.

Q5) 무고한 사람에 대한 공격은 왜 일반적으로 비도덕적인가? 왈저 (Walzer 2004, 2006)에 따르면, 모든 사람은 공격받지 않을 권리(the right not to be attacked)를 가지고 있다. 다른 사람을 공격하는 폭력 행위는 무고

한 사람의 이러한 권리를 침해하기에 비도덕적이다. 다른 한편, 다른 사람을 공격하는 혹은 곧 공격할 사람은 타인의 공격받지 않을 권리를 침해할 것이기에 본인도 공격받지 않을 권리를 상실한다.

Q6) 모든 혐오 표현은 비도덕적인가? 혐오표현은 일반적으로 비도덕적인 것으로 간주된다. 왜냐하면 모든 사람은 정당한 이유 없이 공격받지 않을 권리를 가지는데, 혐오표현은 정당한 이유 없이 다른 사람을 공격하는 일이기 때문이다. 하지만 어떤 이들은 네오나치, KKK단과 같은 인종차별주의적 그룹에 대한 혐오표현은 정당화될 수 있다고 본다 (DiFranco 2017). 이는 네오 나치나 KKK는 무고한 사람들이 아니기 때문이다. 네오나치, KKK에 대한 혐오표현과 인종, 성, 성적 지향, 국적 등에 근거한 혐오표현 사이에는 크게 다음의 두 가지 차이가 있다. (1) 네오 나치나 KKK 그룹 멤버들은 인종차별, 성차별 등과 같은 비도덕적인 목적을 공유한다. 그렇지만 남성, 여성, 백인, 동양인 등의 그룹은 어떠한 비도덕적인 목적도 공유하지 않는다. (2) 나치, KKK 멤버들은 이러한 비도덕적인 단체들에 자발적으로 가입하지만, 어떤 인종, 성, 국적에 속하는지는 개인에게 주어지는 것이지 개인이 선택할 수 있는 문제가 아니다.

우리가 만약 무고한 사람에 대한 공격이 왜 비도덕적인가에 대한 왈저의 설명을 받아들인다면, 이는 왜 네오 나치나 KKK에 대한 혐오가 도덕적으로 정당화될 수 있는지에 대한 한 가지 설명을 제공해준다. 네오 나치나 KKK는 인종, 성 등에 근거해 다른 집단에 속한 사람을 혐오하고 차별할 준비가 되어 있는 사람들이다. 즉, 그들은 다른 사람이 가진 공격받지 않을 권리를 침해하거나 혹은 침해할 것을 목표로 삼는 집단들이다. 그렇기에 그들은 공격받지 않을 권리를 보장받지 못한다는 것이다.

이 논의를 우리 사회에 적용해본다면, 우리는 흔히 성차별적 그룹으로 분류되는 일간 베스트, 메갈리아, 위마드 유저들에 대한 혐오표현(가령, “일베충”(일베+충), “메돼지”(메갈리아+돼지), “웜돼지”(위마드+돼지) 등)이 도덕적으로 정당화되는지 물어볼 수 있다. 이에 대한 답은 이들이 위에서 제시된 두 조건을 만족시키는지에 달려 있을 것이다. 만약 이들 그룹이 여성



혐오 혹은 남성혐오와 같은 비도덕적인 목적을 공유하고 이들 그룹에 대한 가입이 자발적이라면, 인종차별주의적 그룹에 대한 혐오표현이 정당화될 수 있다고 믿는 사람들은 같은 이유에서 성차별적 그룹에 대한 혐오표현도 정당화될 수 있다고 볼 것이다.

## 5. 혐오논문, 표현의 자유, 그리고 학문의 자유

이 절에서 나는 지금까지의 논의를 간략히 요약하고 그 함축에 대해 살펴 보겠다. 2절에서 나는 “혐오표현”에 대한 좁은 정의와 넓은 정의가 있는데 좁은 정의는 올바른 정의가 아니라고 주장하였다. “혐오표현”에 대한 넓은 정의(즉, 인종, 성, 국적, 성적 지향, 종교 등 개인의 정체성에 핵심적인 요소들을 근거로 개인 혹은 집단에 대한 혐오를 표현하거나 폭력을 선동하는 말, 글 혹은 행동)만이 올바른 정의일 때, 이는 윤지선 (2019)<sup>20</sup>에 대해 중요한 한 가지 함축을 갖는다. 그것은 바로 이 논문이 집단혐오용어를 단순히 언급하는 것이 아니라 사용하고 있기 때문이다. 1절에서 언급했듯이, 집단혐오용어를 사용하는 것(slurring)은 혐오표현(hate speech)의 일종이다. 이 때 “혐오표현”은 말뿐만 아니라 글, 행동까지도 포함하는 넓은 개념이다. 따라서, 우리는 집단혐오용어를 사용하는 논문을 혐오논문이라고 결론 내릴 수 있을 것이다. 이 논증은 다음과 같이 정식화될 수 있다.

- P1. 임의의 집단X에 대해, X에 대한 집단혐오용어를 사용하는 논문은 X에 대한 혐오 논문이다.
- P2. 윤지선 (2019)는 한국남성과 남자아이에 대한 집단혐오용어를 사용한다.
- C. 그러므로 윤지선 (2019)는 한국남성과 남자아이에 대한 혐오 논문이다.

먼저 P1에 대해 살펴보자. 어떤 논문이 살인 범죄에 대해 다루면서 “살인

20) “윤지선(2019)”라는 표현은 사람 윤지선 또는 윤지선이 2019년에 쓴 논문을 가리킬 수 있다. 나는 이 표현을 주로 논문을 가리키기 위해 사용할 것이다.

피의자 중 흑인의 비율이  $n\%$ 이다”라고 통계적으로 주장할 수 있다. 하지만 어떤 논문이 “살인 피의자 중 니거의 비율이  $n\%$ 이다” 혹은 나아가서 “니거는 살인자다”라고 주장한다면, 이 논문은 통계적 사실에 관계없이 인종차별 주의적인 혐오 논문으로 받아들여질 것이다 (여기서 한 가지 주의할 점은 어떤 논문이 “살인 피의자 중 니거의 비율이  $n\%$ 이다” 혹은 “니거는 살인자다”라고 주장한다면, 이 논문은 살인 범죄에 대한 논문임과 동시에 흑인에 대한 혐오논문이라는 것이다. 이 둘은 양립불가능한 것이 아니다). 마찬가지로 어떤 논문이 한국의 성매매 실태에 대해 다루면서 “성매매종사자 중 여성의 비율이  $n\%$ 이다”라고 통계적으로 주장할 수 있다. 하지만 어떤 논문이 “성매매종사자 중 김치녀의 비율이  $n\%$ 이다” 혹은 “김치녀는 창녀이다”라고 주장한다면, 이 논문은 통계적 사실에 관계없이 성차별주의적인 혐오 논문으로 받아들여질 것이다.

P2는 윤지선 (2019)이 “한남충”, “한남유충”이라는 용어들을 사용하고 있으며 이 용어들이 한국남성과 남자아이에 대한 집단혐오용어라는 사실에 의해 뒷받침된다.<sup>21)</sup> 이에 대해 두 가지 정도 예상가능한 반론이 있다. (1) 우선 저자는 “한남충”, “한남유충”이 전문적 용어(technical term)이기에 혐오표현이 아니라고 주장할 수 있다. 그렇지만 이는 논문의 여러 부분과 충돌한다. 첫째, 어떤 용어가 전문적 용어라면 그 용어가 처음 도입될 때 명확한 정의가 제시되어야 한다. 그렇지만 이 용어들에 대한 정확한 정의는 논

21) 다음은 “한남충”, “한남유충”이라는 용어가 사용된 몇 가지 사례들이다.

“이를 통해 본 논문은 한남유충에서 한남충으로의 변태(metamorphosis) 과정의 추이가 ‘관음충’의 지수(factor)를 통해 구체적으로 어떠한 방식으로 단계적으로 진화 가능한지를 첨예하게 드러낼 것이다.” (p.261)

“필자는 본 논의에서 한국남아가 한남유충에서 한남(성)충으로 성장, 진화하는 과정을 측정, 규명하는 주요인자로서 관음충 되기의 지수를 활용할 것이며 왜 관음충이 개체적 발생이 아닌 군집구성체적 발생인지에 대해 고찰하고자 한다.” (p.262)

“한남유충”이란 용어의 고안은 20대 여성들이 던진 다음과 같은 질문으로부터 추동되었다. 과연 한남충은 어떻게 탄생하는가? 현재의 한남(성)충들은 성인이가 도래하자 갑자기 출현한 군집체인가, 아니면 한국여성들이 미성년자일 때부터 학교 교실에서, 학원에서 포진된 채 여성혐오놀이에 심취해있던 그 남아 군집체들의 연장성인가?” (p.263)

문 어디에도 제시되어 있지 않다. 둘째, 3.2절에서 보았듯이 저자는 “한남충”과 “한남유충”이 혐오용어라는 것을 인지하고 이의 사용을 정당화하려고 시도한다.<sup>22)</sup> 만약 이 두 용어가 중립적인 용어라면 이러한 시도는 이해하기 힘들 것이다.

(2) 다음으로 저자는 논문의 맥락 안에서 “한남충”, “한남유충”이라는 용어들은 남성 불법촬영 범죄자만을 위한 혐오표현이라고 주장할 수도 있다. 그렇지만 “한남충”은 관습적으로 한국 남성 일반을 지시하는 집단혐오용어이고 “한남유충”은 관습적으로 한국 남자아이 일반을 지시하는 집단혐오용어이다. 이는 마치 살인 범죄에 관한 논문에서 “니거”라는 표현을 쓰면서 사실은 “나는 이 논문에서 “니거”라는 말로 흑인 범죄자만 지시하는 거야”라고 주장하는 것과 같다. 또는 성매매 실태에 관한 논문에서 “김치녀”라는 표현을 쓰면서 사실은 “나는 이 논문에서 “김치녀”라는 말로 성매매 종사여성만 가리키는 거야”라고 주장하는 것과 같다. 상식적인 사람이라면 이런 주장을 받아들이기 힘들 것이다.

다음으로 3절의 논의의 함축에 대해 살펴보자. 3절에서 나는 비대칭성 논제가 지탱될 수 없다고 논증하였다. 만약 이러한 입장이 옳다면, 대부분의 경우 집단혐오용어를 사용하는 것은 도덕적으로 정당화될 수 없을 것이다. 그리고 그 이유는 집단혐오용어가 무고한 사람들의 공격받지 않을 권리를

22) 다음 인용문들로 볼 때, 저자의 태도는 이 두 용어가 혐오표현이지만 저항의 수단으로 사용되기 때문에 사용이 정당하다는 입장이라고 볼 수 있다.

“우선 필자는 ‘한남유충과 한남충, 관음충’이라는 용어가 방사하고 있는 다면체적 효과에 대해 논의하고자 한다. 왜냐하면 위의 용어들이 기존의 사회적 소수자들을 표적하고 있는 ‘~충’의 혐오용어들과 동일한 선상에서 놓일 수 있는지, 아니면 이와는 다른 반경효과를 함축하고 있는지에 대해 논증해야 할 필요성이 제기되기 때문이다. ‘한남유충과 한남충, 관음충’은 대한민국 남성에 대한 혐오용어로서 금기시되어야 하는가? 그렇다면 이때껏 가부장제 남성중심주의 사회에서 여성을 표적으로 했던 수많은 혐오용어들-김치녀, 된장녀, 꽃뱀, 맘충, 룸나무-의 고안자와 사용자는 누구였는가?” (pp.262-3)

“‘한남유충’이란 용어가 단순히 남아혐오용어로 그 의미와 효과가 한정되는 것이 아니라 한국 여자들이 교실·학원·SNS에서 사용하는 ‘한남유충’이란 용어의 기능이란 남아 군집체들이 일상 안에서 공격적으로 내뿜는 여성혐오용어들의 효과를 무력화시킬 수 있는 맞대응 전략의 지점에서 파생되었다.” (p.264)

침해하기 때문이다. 이 논증은 다음과 같이 정식화될 수 있다.

- P1. 윤지선 (2019)는 한국남성 및 남자아이에 대한 혐오 논문이다.
- P2. 한국남성 및 남자아이에 대한 혐오논문은 무고한 사람의 권리를 침해한다.
- P3. 무고한 사람의 권리를 침해하는 논문은 비도덕적인 논문이다.
- C. 그러므로 윤지선 (2019)은 비도덕적인 논문이다.

P1은 이전 논증의 결론이다. 그러므로 이전 논증을 받아들인다면 P1을 받아들여야 할 것이다. 3절의 논의에 따르면, P2는 참이다. 왜냐하면 한국남성 및 남자아이 혐오논문은 이들 집단에 속한 모든 이들에 대한 무차별적 공격이며 따라서 이들 집단에 속한 이들 가운데 무고한 사람들의 공격받지 않을 권리를 침해하기 때문이다. P3 역시 참이다. 왜냐하면 무고한 사람은 공격받지 않을 권리를 가진 사람이며, 그러한 권리를 가진 사람을 물리적, 정서적 혹은 언어적으로 공격한다면 이는 그 사람의 권리를 침해한 것이기 때문이다.

지금까지 제시된 두 논증을 모두 받아들이면 우리는 윤지선 (2019)을 비도덕적인 혐오논문 이라고 결론 내릴 수 있다. 그렇다면 비도덕적인 혐오논문도 표현의 자유(freedom of speech)에 의해 보호받아야 하는가? 밀의 해악 원칙에 따르면, 표현의 자유는 오직 어떤 표현이 임박한 위험(imminent risk)을 야기할 때에만 제한될 수 있는 것 같다. 그런데 혐오 논문은 어떤 임박한 위험도 야기하지 않는다. 따라서 밀의 해악원칙에 따르면 비도덕적인 혐오논문도 표현의 자유에 의해 보호되어야 하는 것 같다.<sup>23)</sup>

그렇지만 이것이 혐오논문에 대해 아무런 대응도 할 수 없다는 것을 의미하는 것은 아니다. 혐오논문에 대해 우리는 크게 두 가지 방식으로 대응할

23) 이에 대해 어떤 이들은 밀의 해악원칙 외에 표현의 자유가 제한될 수 있는 다른 추가적인 조건들이 있다고 주장할 수 있다 (Mill 2021). 그러나 나는 여기서 우리가 표현의 자유에 대해 가장 자유로운 입장 (즉, 해악원칙만이 표현의 자유에 대한 유일한 제약이라는 입장)을 채택하더라도 여전히 혐오논문에 대해 대응할 방법이 있다는 것을 보이고자 한다.

수 있다. 첫째, 우리는 혐오논문은 학문의 자유(academic freedom)에 의해 보호받지 못한다고 주장할 수 있다. 흔히 표현의 자유와 학문의 자유의 목적은 진리(= 참인 명제들)의 발견 및 증진이라고 한다 (Lackey 2018; Mill 2011; Mill 2021). 그러나 표현의 자유를 지배하는 규범과 학문의 자유를 지배하는 규범은 구분될 수 있다. 만약 표현의 자유가 오직 밀의 해악 원칙에 의해서만 제한될 수 있다면, 이는 표현의 자유가 진리의 발견 및 증진이라는 목적 외에도 정치 권력에 대한 비판, 시민으로서의 의사 표현 및 자율성 함양 등 다양한 목적을 위함일 것이다. 이에 반해 학문의 자유를 보장하는 유일한 혹은 가장 중요한 목적은 진리의 발견 및 증진이다 (Dewey 1902). 즉, 대학의 가장 중요한 목적은 진리의 발견 및 증진이고 이를 위해 제재나 보복 등에 대한 걱정 없이 어떤 주제에 관해서도 가르치고 연구할 자유가 필요하다는 것이다. 만약 그렇다면, 진리 발견 및 증진이라는 목적에 기여하지 않는 행동들은 학문의 자유에 의해 보호될 필요가 없을 것이다.

만약 이러한 입장이 옳다면, 최소한 집단혐오용어를 사용하는 논문들은 학문의 자유에 의해 보호될 필요가 없다. 이를 다음 사례를 통해 살펴보자.

(1) a. 살인범죄 피의자의 대부분은 흑인이다.

b. 살인범죄 피의자의 대부분은 니거이다.

집단혐오용어에 대해 한 가지 널리 받아들여지는 가정에 따르면, 집단혐오용어는 많은 경우 중립적인 대응표현(neutral counterpart)을 가진다. 그리고 집단혐오용어와 그에 대응하는 중립적인 대응 표현은 정확히 같은 의미론적 의미(= 진리치에의 기여)를 가진다 (Anderson & Lepore 2013; Bolinger 2017; Jeshion 2013). 가령, “흑인”은 “니거”의, “한국 남성”은 “한남”의, “한국 남자아이”는 “한남유충”의 중립적인 대응 표현이다. 그리고 이들은 정확히 같은 의미론적 기여를 가진다는 것이다. 이는 (1a)와 (1b)가 모든 맥락에서 같은 진리치를 가진다는 것을 의미한다. 만약 그렇다면, 우리는 적어도 학문의 영역에서 (1b)를 금지한다고 해도 아무 것도 잃을 것이 없을 것이다. (1b)를 대체할 수 있는 표현 (= (1a))이 항상 존재하기 때문이다. 게다가 (1b)는 다른 사람의 권리를 침해하지만 (1a)는 그렇지 않다. 비유적으로 말

하자면 이는 사상의 자유시장에 두 가지 똑같은 상품이 있는 것과 같다. 유일한 차이는 하나는 다른 사람의 권리를 침해하지만 다른 하나는 그렇지 않다는 것이다. 그렇다면 우리는 굳이 첫 번째 상품을 사상의 자유시장에 남겨둘 필요가 없을 것이다. 이러한 입장에 따르면, 윤지선(2019)과 같은 비도덕적인 혐오논문은 설사 표현의 자유에 의해 보호받을 수 있을지라도 (따라서 시위, 개인 블로그, 사적 모임 등 진리 발견 및 증진이라는 목적 외 다른 목적도 관여할 수 있는 경우에는 허용될 지라도) 학문의 자유에 의해서는 보호받지 못한다. 즉, 연구 및 수업과 같이 진리의 발견, 증진 및 전달이 유일한 목적인 맥락에서는 무고한 사람의 권리를 침해하는 비도덕적인 혐오 논문은 허용되지 않을 것이다.<sup>24)</sup>

두 번째는 혐오 논문에 대해 더 많은 표현으로 대응하는 것이다. 이는 단순히 더 많이, 혹은 더 크게 말함으로써 혐오표현을 침묵(silencing)시키자는 이야기가 아니다. 사상의 자유시장에서의 자정작용을 통해 편견 및 오류를 드러냄으로써 이러한 편견 및 오류가 더 이상 확산되는 것을 막자는 의미이다. 예를 들어, 윤지선(2019)의 주된 목적은 “한국에서 태어난 남아가 어떻게 ‘관음충’으로 집단적으로 성장, 진화하는가를 분석”(p.260)하는 것이라고 하며 이러한 경험적 주장에 대한 유일한 통계적 근거는 “불법촬영범죄 피의자의 97%가 남성”(각주 15)이라는 것이다. 이러한 점은 다음 인용문에서 잘 드러난다.

“대한민국 사회에서 불법촬영물을 중심으로 집단적으로 서식하는 관음충의 성별성은 97% 이상이 남성이며 그것의 연령 분포도도 매우 다양하다고 한다. 그러하기에 필자는 한국에서 태어난 남아가 어떻게 ‘관음충’으로 집단적으로 성장, 진화하는가를 분석하고자 …” (pp. 266-67)

24) 위 논증은 연구뿐만 아니라 수업에도 적용될 수 있다. 가령, 어떤 교수가 범죄에 관한 수업에서 (1a)를 발화한다면 이 발화는 학문의 자유에 의해 보호될 수 있을 것이다. 그러나 그가 (1b)를 발화한다면 이는 지식 전달이라는 수업의 목적을 벗어난 것으로 학문의 자유에 의해 보호될 필요가 없을 것이다.

저자의 추론은 (2)와 같이 정식화될 수 있다. 그러나 이 추론은 매우 명백한 논리적 오류를 범하고 있다. 만약 (2)의 추론이 타당하다면, (3), (4), (5)의 추론도 타당해야 한다. 그러나 그렇지 못하다.

(2) 불법촬영범죄 피의자의 대부분은 남성이다.

⇒ 남자 아이는 집단적으로 관음충(= 불법촬영 가해자)으로 성장, 진화한다

⇒ 남성의 대부분은 관음충(= 불법촬영 가해자)이다.

⇒ 남성은 관음충(= 불법촬영 가해자)이다.

(3) 살인범죄 피의자의 대부분은 흑인이다. (T)

⇒ 흑인 아이는 집단적으로 살인자로 성장, 진화한다. (F)

⇒ 흑인의 대부분은 살인자이다. (F)

⇒ 흑인은 살인자이다. (F)

(4) 성매매업에 종사하는 사람의 대부분은 여성이다. (T)

⇒ 여자 아이는 집단적으로 매춘부로 성장, 진화한다. (F)

⇒ 여성의 대부분은 매춘부이다. (F)

⇒ 여성은 매춘부이다. (F)

(5) 성범죄 무고범의 대부분은 여성이다. (T)

⇒ 여자 아이는 집단적으로 성범죄 무고범으로 성장, 진화한다. (F)

⇒ 여성의 대부분은 성범죄 무고범이다. (F)

⇒ 여성은 성범죄 무고범이다. (F)

(6) 초등학교 교사의 대부분은 여성이다. (T)

⇒ 한국사회에서 여자아이는 집단적으로 초등학교 교사로 성장한다. (F)

⇒ 여성의 대부분은 초등학교 교사이다. (F)

⇒ 여성은 초등학교 교사이다. (F)

윤지선 논문이 분석하고자 하는 대상은 “한국에서 태어난 남아가 어떻게 관음충으로 집단적으로 성장, 진화하는가”인데 이는 저자가 제시한 통계적 근거에 의해 따라나오지 않는다.<sup>25)</sup> 이는 저자가 분석하고자 하는 현상이 존재하는지조차 제대로 확립되지 않았다는 것을 보여준다. (2)-(6)의 첫 번째 문장은 각각 통계적 사실일 수 있다. 그러나 왜 그러한 통계적 사실이 나타나게 되었는가는 조심히 다루어져야 할 문제이다. 하나의 통계적 사실은 다양한 요소들, 즉 생물학적 (e.g., 신체의 차이, 타고난 재능 및 선호의 차이, 위험(risk)에 대한 성향 차이 등), 사회적 (e.g., 사회계층, 본인의 교육 수준, 부모의 교육 수준, 본인의 형제 내 순서, 거주 지역 등), 문화적 (e.g., 성역할, 가족 문화 등) 요소들이 복합적으로 작용한 결과일 수 있기 때문이다. 이러한 세심한 고려 없이 위의 추론들을 받아들이는 것은 인종차별적인, 성차별적인 편견을 무비판적으로 받아들이는 것일 뿐이다. 이 외에도 윤지선 논문의 내용에 여러 문제가 있으나 이는 본 논문의 직접적인 목적과 범위를 넘어서기에 더 이상 다루지 않도록 하겠다.

## 6. 결론

우리는 지금까지 대한민국 언론 및 페미니즘 분야에서만 독특하게 나타나는 혐오표현에 관한 한 가지 도덕적 논제(= 비대칭성 논제: 여성(혹은 소수자/약자)에 대한 혐오 표현은 도덕적으로 정당하지 않지만 남성(혹은 다수자/강자)에 대한 혐오 표현은 도덕적으로 정당하다)에 대해 살펴보았다. 비대칭성 논제를 위한 논증들은 크게 사회적 지위에 근거한 논증과 부정의에 대한 저항에 근거한 논증으로 나뉘어질 수 있다. 각각의 논증은 혐오표

25) 간단한 모델로 오직 백인과 흑인으로만 구성된 나라가 있다고 가정해보자. 이 나라에서 1000명당 범죄자의 수는 백인의 경우 5명, 흑인의 경우 50명이고 백인의 인구는 흑인보다 5배 더 많다고 가정해보자. 이러한 상황에서 “범죄자의 대부분은 흑인이다”와 “흑인의 인당 범죄율은 백인보다 10배 더 높다”는 각각 참일 것이다. 그러나 이로부터 “흑인의 대부분은 범죄자이다” 혹은 “흑인은 범죄자이다”가 따라나오지는 않는다. 왜냐하면 여전히 대부분의 흑인(95퍼센트)은 범죄자가 아니기 때문이다.



현의 도덕적 정당성이 사람들의 사회적 지위 혹은 누가 먼저 공격하였는가에 달려있다고 주장한다. 그러나 비대칭성 논제를 위한 논증들은 모두 심각한 반론에 직면한다. 나는 혐오표현의 도덕적 정당성은 무고한 사람에 대한 공격을 포함하느냐 여부에 달려 있다고 논증하였다. 다시 말해, 혐오표현은 무고한 사람에 대한 공격을 포함할 때 도덕적 잘못이다. 이러한 관점에 따르면, 여성에 대한 혐오표현뿐만 아니라 남성에 대한 혐오표현 역시 무고한 사람에 대한 공격을 포함하기에 도덕적으로 정당화될 수 없다.

## 참고문헌

### 신문기사 및 칼럼

- 곽재훈. “‘위마드’ 논란, ‘남성혐오’는 허구다”, 프레시안 2017.12.01.
- 남정미. “김치녀는 혐오표현이고 김치男은 아니다?”, 조선일보 2019.03.13.
- 박경신. “넥슨 여성운동 탄압사태, 눈치들 보지 마라”, 경향신문 2016.08.01.
- 서정호, 강재연, 강승민. “홍성수 교수가 말하는 ‘혐오표현’이란?” YTN 2019.06.26.
- 손국희. “김치남·백형 혐오표현 아니다? 학교서 가르치라는 여가부”, 중앙일보 2019.03.15.
- 정희진. “‘남성혐오’는 가능한가: 여성주의와 양성평등”, 인물과 사상 2015.10: 110-124.
- 홍성수. “핵심은 의도 아닌 ‘효과’다”, 한겨레 2020.08.20.

### 자료

- 여성가족부 (2019). 초중고 성평등 교수 학습 지도안 사례집.

### 논문 및 저서

- 박대아 (2021). “남성혐오표현과 여성혐오표현의 상호작용 연구: 반사형 혐오표현을 중심으로”, 『한국학연구』 76, pp. 95-118.
- 윤지선 (2019). “관음충의 발생학: 한국남성성의 불완전변태과정의 추이에 대한 신물질주의적 분석”, 『철학연구』 127, pp. 259-88.
- 윤지영 (2015). “전복적 반사경으로서의 메갈리안 논쟁: 남성혐오는 가능한가”, 『한국여성철학』 24, pp. 5-79.
- 홍성수 (2018). 『말이 칼이 될 때: 혐오표현은 무엇이고 왜 문제인가?』, 어크로스.
- Anderson, L., & Lepore, E. (2013). “Slurring words”, *Nous* 47(1), pp. 25-48.

- Austin, J. (1975). *How to do Things with Words*. Oxford: Clarendon.
- Bolinger, R. (2017). “The pragmatics of slurs”, *Nous* 51(3), pp. 439-462.
- Dewey, J. (1902). “Academic freedom”, *Educational Review* 23, pp. 1-14.
- DiFranco, R. (2017). “Appropriate slurs”, *Acta Analytica* 32(3), pp. 371-384.
- Jeshion, R. (2013). “Slurs and stereotypes”, *Analytic Philosophy* 54(3), pp. 314-329.
- Lackey, J. (2018). “Academic freedom”, In J. Lackey (ed.) *Academic Freedom*. Oxford: Oxford University Press.
- Mill, J. S. (2011). *On Liberty*. Cambridge: Cambridge University Press. [originally published in 1859].
- Popa-Wyatt, M., & Wyatt, J. (2018). “Slurs, roles, and power”, *Philosophical Studies* 175, pp. 2879 - 2906.
- Primoratz, I. (2018). “Terrorism”, *The Stanford Encyclopedia of Philosophy*. (<https://plato.stanford.edu/entries/terrorism/>)
- van Mill, D. (2021). “Freedom of speech”, *The Stanford Encyclopedia of Philosophy*. (<https://plato.stanford.edu/entries/freedom-speech/>)
- Walzer, M. (2004). *Arguing about War*. New Haven: Yale University Press.
- Walzer, M. (2006). *Just and Unjust Wars: A Moral Argument with Historical Illustrations*. 4<sup>th</sup> edition, New York: Basic Books.

University of Southern California Ph.D candidate  
junhyole@usc.edu



## Feminism and Terrorism

Junhyo Lee

It is widely held among contemporary Korean feminists that while using a slur for women (or subordinate/minority groups) is morally impermissible, using a slur for men (or dominant/majority groups) is morally permissible. I call this *the asymmetry thesis*. The main aim of this paper is to argue that the asymmetry thesis is not tenable. One of the main arguments for the asymmetry thesis is that since at least some men attacked women first by using a slur for women, women are morally justified in attacking men by using a slur for men. However, this argument for the asymmetry thesis is logically in parallel with Osama bin Laden's argument for terrorism, so they are a package deal. After criticizing the main arguments for the asymmetry thesis, I argue that what is morally relevant to the use of a slur is whether using a slur violates innocent people's right not to be attacked. On this view, using a slur for men, as well as using a slur for women, is morally impermissible because it violates innocent people's right not to be attacked.

**[Subject]** applied philosophy of language, normative ethics, feminism

**[Keywords]** feminism, terrorism, the asymmetry thesis, hate speech, slurs