

탈인간중심주의적 의미론으로서의 리터럴리즘

노 형 래*

순천대학교

심리적 개념은 불변하는 보편적 범주가 아니다. 예컨대, ‘통증’ 개념이 적용되는 대상의 영역은 인간과 유사한 신경생리적 기제를 공유하는 동물에서부터 그와 다른 기제를 통해 통증을 느끼는 어류로 확장되어왔다. 이러한 예는 과학의 발전에 따라 인간 고유의 생물학적 속성과 무관하게 심리적 범주가 탈인간중심주의적으로 확장될 수 있음을 보여준다. 미래사회는 오늘날보다 더 빠르고 거대한 과학적 발전을 마주할 것이므로, 우리는 그에 대응하여 탈인간중심주의적으로 개념의 의미를 수정할 수 있어야 한다. 본 논문은 리터럴리즘이 그러한 역할을 수행할 수 있음을 논증한다. 리터럴리즘은 심리적 개념이 과학적 맥락에 사용되는 것을 근거로 그 개념의 외연을 결정한다. 예를 들어, 인지과학자들은 특정한 수학적 모델을 통해 인간의 경제적 행동과 초파리의 지각적 행동 모두를 예측한다. 이때, 그러한 수학적 모델의 의미는 ‘결정하다’와 같은 심리적 술어를 통해 해석된다. 리터럴리즘에 따르면, “초파리는 동쪽으로 날아가기로 결정했다”와 “철수는 주식에 투자하기로 결정했다”에서 초파리와 철수의 행동은 동일한 수학적 모델로 설명되므로 두 문장에 등장하는 ‘결정하다’라는 심리적 술어는 동일한 의미를 갖는다. 리터럴리즘은 과학의 발전에 따른 심리적 범주의 탈인간중심주의적 확장에 대한 이론적 정당성을 마련한다.

주요어 : 철학, 포스트휴머니즘, 탈인간중심주의, 리터럴리즘, 심리적 개념, 수학적 모델, 의미론

* 주저자: 노형래/순천대학교 철학과 조교수/전남 순천시 중앙로 255(석현동) 인문예술대학 517호
/Tel: 061-750-3487/E-mail: hyungrae-noh@scnu.ac.kr

I. 서론

1. 연구의 필요성

도덕성과 마음에 관한 개념들은 불변하는 보편적 범주가 아니다. 요컨대, 불과 한 세기 이전만 하더라도 비인간 동물에게 쓸데없는 고통을 주는 것이 도덕적으로 나쁘다는 믿음은 보편적이지 않았다. 기존에는 생물학적 인간만을 도덕적 고려 대상으로 여겼으나, 사회적 변화에 따라 연관된 도덕적 범주에 비인간 동물을 포함한 것이다. 심리적 범주는 과학의 발전에 따라 수정되어 왔다. 예를 들어, 어류도 포유류처럼 고통을 느낀다는 연구 결과는 고통의 범주에 어류도 포함된다는 것을 알려준다 (Sneddon, 2015, p. 967). 인공지능 로봇의 발전은 그러한 로봇을 마음을 가진 존재로 인정해야 하는가에 관한 철학적 논의에 불을 붙였다 (Prescott and Robillard, 2021, p. 1). 이렇듯 심리적 범주는 과학적 발전에 따라 생물학적 인간 너머로 확장되어 왔다.

인간의 생물학적 속성은 도덕적 범주와 심리적 범주의 확장, 즉 연관된 개념들의 외연 확장과 연관이 없는 것으로 보인다.¹⁾ 공리주의에 따라 고통을 느끼는 모든 대상은 쓸데없는 고통을 당하지 않을 권리인 피동적 도덕성을 지닌다고 가정하자. 이러한 가정에 따르면, 인간의 생물학적 속성만을 근거로 ‘피동적 도덕성’ 개념의 외연을 결정할 수 없다. 어류의 예를 살펴보자. 어류가 고통을 느끼는 신경생리적 기제는 인간 같은 포유류의 그것과 다르다 (Sneddon, 2015, p. 969).²⁾ 어떤 대상과 인간의 생물학적 거리는 그 대상에게 피동적 도덕성을 부여할 수 있는가의 유일한 판단 근거가 아니다. 과학의 발전과 그에 따른 사회적 변화는 전통적으로 생물학

적 인간에게만 (또는 인간과 생물학적으로 유사한 대상들에게만) 적용되어온 개념들의 외연 확장으로 이어질 수 있다. 달리 말하자면, 도덕성과 마음에 관한 개념들이 적용되는 대상의 영역이 과학의 발전에 따라 미래에는 우리가 예측할 수 없을 정도로 넓어질 수 있는 것이다. 본 논문은 마음에 관한 개념들의 외연이 확장되는 방식에 집중한다.

미래사회는 오늘날보다 더 빠르고 거대한 과학적 발전을 마주할 것이므로, 그에 대응하여 심리적 개념의 의미를 적절히 수정할 수 있는 이론이 요구된다. 본 논문은 그러한 이론으로서 리터럴리즘을 제시한다. 리터럴리즘에 의하면, 심리적 범주는 연관된 개념이 과학적 맥락에 등장하는 것을 참고하여 결정될 수 있다. 예를 들어, 인지과학자들은 특정한 수학적 모델을 통해 인간의 경제적 행동과 초파리의 지각적 행동 모두를 예측한다. 이때, 그러한 수학적 모델의 의미는 ‘결정하다’와 같은 심리적 술어를 통해 해석된다. 리터럴리즘에 따르면, “초파리는 동쪽으로 날아가기로 결정했다”와 “철수는 주식에 투자하기로 결정했다”에서 초파리와 철수의 행동은 동일한 수학적 모델로 설명되므로 두 문장에 등장하는 ‘결정하다’라는 심리적 술어는 동일한 의미를 갖는다. 이렇듯 리터럴리즘은 과학적 발전에 따라 심리적 범주가 생물학적 인간 너머로 확장된다는 것의 이론적 정당성을 마련한다.

첫 번째로 나는 심리적 개념의 외연을 탈인간중심주의적으로 확장하는 이론이 갖추어야 할 조건이 무엇인지 설명할 것이다. 이를 설명하기 위해 우선 포스트휴머니즘이 제안하는 탈인간중심주의가 무엇인지 살필 것이다. 그리고 그러한 탈인간중심주의적 제안에 따라 개념의 외연을 확장한다는 것이 무엇인지 ‘피동적 도덕성’ 개념의 예시를 통해 설명할 것이다. 이와 같은 예시를 분석하여 심리적 개념의

1) 본문에서 ‘범주’와 ‘개념의 외연’은 동의어이다.

2) 포유류-고통의 신경생리적 기제는 주로 C 신경섬유로 구성되지만 어류-고통의 신경생리적 기제는 주로 A-델타 신경섬유로 구성된다 (Sneddon, 2015, p. 969).

의연을 탈인간중심주의적으로 확장하는 데 따르는 조건이 무엇인지 명확히 할 것이다. 마지막으로 리터럴리즘이 그러한 조건을 잘 만족하는 의미론임을 논증할 것이다.

II. 탈인간중심주의적 의미론이란 무엇인가?

1. 포스트휴머니즘의 탈인간중심주의

강영안과 이상현에 따르면, 포스트휴머니즘은 데카르트적 휴머니즘에 대한 반대에서 출발한다 (강영안, 이상현, 2013, pp. 154-156). 데카르트적 휴머니즘에 의하면, 인간을 가장 인간답게 만들어주는 것은 환원 불가능한 인간의 본질로서의 자기의식이다. 이와 같은 데카르트적 휴머니즘이 참이라면, 자기의식을 갖는 인간과 자기의식을 갖지 않는 비인간 사이의 이분법이 성립한다. 이러한 이분법은 비인간으로 구성된 자연을 인간의 욕구 충족을 위한 수단으로 삼는 근거가 되었다. 그 결과로 데카르트적 휴머니즘은 “전체로부터의 개체의 소외, 전통과 문화와 역사를 통해 형성된 의미로부터의 단절, 자원 고갈과 환경 훼손, 그리고 보다 철저하고 조직적인 사회통제 등을 초래하였다”(강영안 1991, p. 239; 강영안, 이상현, 2013, p. 155에서 재인용). 하지만 포스트휴머니즘에 따르면, 환원 불가능한 인간의 본질 같은 것은 없으므로 이 세계를 인간과 비인간으로 구분하는 이분법은 성립하지 않는다.

포스트휴머니즘에 의하면, 인간과 비인간의 이분법이 성립하지 않는 것처럼 몸과 마음의 이원론도 성립하지 않는다. 이원론은 ‘대중적’ 포스트휴머니즘과 ‘비판적’ 포스트휴머니즘의 구분을 통해 이원론이 성립하지 않는 것을 설명한다(이원봉, 2018, pp. 66-70). 우선 전자를 살펴보자. ‘대중적’ 포스트

휴머니즘은 육체와 두뇌의 한계로부터 자유로운 인간 형태를 추구하는 사상이다. 이에 따르면 컴퓨터나 온라인에 업로드된 인간의 형태야말로 ‘포스트휴먼’에 적합하다(이원봉, 2018, p. 65). 이와 같은 ‘대중적’ 포스트휴머니즘은 육체와 대비되는 존재로서의 모종의 이상적인 인간 본질을 전제한다. 달리 말하자면, ‘대중적’ 포스트휴머니즘은 환원 불가능한 자기의식과 육체를 구분하고 나아가 육체로부터 완전히 자유로운 자기의식으로서의 인간 형태를 추구한다.

육체에서 완전히 분리된 자기의식이 가능하다는 ‘대중적’ 포스트휴머니즘은 데카르트적 휴머니즘의 이분법을 답습한다(이원봉, 2018, p. 67). 데카르트적 휴머니즘이 인간중심주의적 이분법을 통해 인간을 자연으로부터 분리한 것처럼, 인간을 일종의 신경망의 정보패턴으로 보는 ‘대중적’ 포스트휴머니즘도 인간과 육체를 분리하고 있다. 즉, ‘대중적’ 포스트휴머니즘의 이분법은 육체로부터 독립적인 인간 본질을 전제한다. 요컨대, 신경망이 없는 식물 같은 것들은 인간과 본질적으로 다르다는 것이다.

‘비판적’ 포스트휴머니즘은 마음이 몸에 의존한다는 체화된 인지를 전제한다. 이에 따르면, 개별적 인간이 자연 독립적으로 존재할 수 없듯, 육체와 분리된 ‘나’란 존재할 수 없다(이원봉, 2018, p. 68). 몸과 마음 사이에 명확한 경계 따윈 없다는 주장은 마음을 가진 존재와 그렇지 않은 것처럼 보이는 존재 사이에도 명확한 심리학적 경계가 없다는 것을 함축한다. 체화된 인지, 즉 인간의 지각이 신체적 특성에 의존하는 방식으로 작동한다고 가정하자(Shapiro and Spaulding, 2021). 이때, 인간이 체화된 지각으로 사물을 인식하는 것과 덩굴식물이 줄기 끝 감각 수용체로 벽을 타고 올라가는 것 사이에 본질적 차이란 존재하지 않게 된다(Calvo Garzón, 2007, p. 208). 체화된 인지를 받아들인다면, 마음이 없는 것처럼 보이는 돌맹이, “어느 정도” 마음이 있는 식물, 우리에게 친숙한 마음을 가

진 존재들로(예를 들어, 침팬지) 이어지는 심리학적 연속성이 있을 뿐 이들 사이를 가르는 명확한 심리학적 경계란 없다.

‘비판적’ 포스트휴머니즘에 따르면, 인간만이 소유하는 불변하는 본질적인 속성 따윈 존재하지 않는다(이원봉, 2018, p. 69). 인간과 유사한 신경생리적 기제를 공유하지 않음에도 불구하고 통증을 느끼는 어류, 인간처럼 지각하는 식물 등은 ‘비판적’ 포스트휴머니즘의 증거이다. 전통적으로 인간 고유(또는 인간과 생물학적으로 유사한 대상들만의) 것이라고 여겨졌던 마음에 관한 개념들의 외연이 과학의 발전에 따라 확장되고 있는 것이다.

포스트휴머니즘은 몸과 마음이 분리될 수 없다는 것에서 한발 더 나아가 탈인간중심주의를 추구한다.³⁾ 신상규에 의하면, 포스트휴머니즘은 ‘인간,’ ‘기계,’ ‘생명’에 대한 새로운 이해의 패러다임이나 언어 문법을 모색한다는 점에서 탈인간중심주의적이다(신상규, 2020). 포스트휴머니즘은 이전에는 생물학적 인간에게만(또는 특정한 계급/유형의 생물학적 인간에게만) 적용되어온 개념들, 특히 도덕적 개념을 노예나 장애인뿐만 아니라 식물과 동물 같은 생명체 그리고 사이보그나 인공지능 로봇 같은 기술적 존재에게까지 적용한다. 달리 말하자면, 포스트휴머니즘은 생물학적 종으로서의 인간을(또는 특정한 계급/유형의 생물학적 인간을) 중심으로 만들어진 다양한 도덕적 그리고 심리적 위계를 해체한다.

2. ‘피동적 도덕성’ 개념으로 설명하는 탈인간중심주의적 의미론

개념이 심적 표상이라면, 개념의 외연은 표상되

는 대상 또는 개별자들의 유형으로 결정된다.⁴⁾ 예컨대, 데카르트적 휴머니즘에 따른 도덕성 관련 개념의 외연은 자기의식적 존재 또는 그러한 속성을 갖는 개별자들로만 구성된다. 반면, 탈인간중심주의적 도덕성 개념은 생물학적 인간, 나무, 초파리, 사이보그, 인공지능 로봇과 같은 개별자들을 적어도 원리상 외연 안에 포함할 수 있어야 한다.

정윤경에 따르면, 포스트휴머니즘의 탈인간중심주의는 다음의 요건을 갖춰야 한다(정윤영 2019, p. 134). 첫째, 생물학적 인간의 자기-지시적 속성을 도덕적 그리고 심리적 개념들 외연의 표준으로 삼으면 안 된다. 예컨대, 인간종의 유전적 특성이나 신경생리적 특성 같은 것은 표준이 될 수 없다. 둘째, 연관된 개념들 외연의 중심에 생물학적 인간을 위치시키고 그에 따라 존재론적 위계를 설정하면 안 된다. 예컨대, 인간과 유사한 생김새를 갖는 동물들, 예를 들어 침팬지에게는 “높은” 도덕적 지위를 주고 그렇지 않은 곤충들에게는 “낮은” 도덕적 지위를 줄 수 없다.

연관된 개념들의 외연을 탈인간중심주의적으로 결정하는 방법은 크게 내재론적 방법과 관계론적 방법으로 구분된다. 간략하게 설명하자면, 전자의 방법은 대상이 가진 특정한 속성/능력/기능 또는 속성/능력/기능의 특정한 집합을 통해 개념의 외연을 결정하는 방법이고, 후자의 방법은 언어적 관습과 같은 개념의 실제 사용 방식에 호소하여 개념의 외연을 결정하는 방법이다.⁵⁾ ‘피동적 도덕성’ 개념의 예시로 두 방법의 차이를 구체적으로 설명하겠다.

데카르트적 휴머니즘은 인간의 자유의지를 행사할 수 있는 능력에 호소하여 ‘도덕성’ 개념의 외연을 결정한다. 이와 같은 내재론적 방법은 기계적 세계의 법칙에 전적으로 종속된 (또는 적어도 그런

3) 이후 본문에 나오는 포스트휴머니즘은 모두 ‘비판적’ 포스트휴머니즘을 의미한다.

4) 논의를 위해 본문에서 개념에 대한 표상주의를 언급했지만, 본 논문은 개념을 심적 표상으로 간주하는 표상주의적 존재론을 전제하지는 않는다. 개념에 관한 존재론적 논의에 대해서는 Margolis and Laurence(2019)를 참조하라.

5) 앞으로 본문에서 속성, 능력, 기능 세 가지 용어를 구분하지 않고 사용하겠다.

것처럼 보이는) 본능적인 존재인 동물이 어찌서 도덕적 고려의 대상이 아닌가를 직관적으로 보여준다(신상규, 2019, p. 247). 반면, 벤담(Jeremy Bentham)으로부터 시작하여 싱어(Peter Singer)에 이르는 공리주의자들은 인간과 비인간 동물 모두 갖는 고통이나 쾌락을 느낄 수 있는 감수 능력(sentiment)에 호소하여 ‘도덕성’ 개념의 외연을 비인간 동물들에게까지 확장한다. 도덕적 행위를 할 수 있으며 도덕적 책임의 주체인 행위자(agent)와 그러한 행위자적 도덕성은 갖지 않지만 감수 능력을 갖는, 즉 도덕적 행위의 대상이 되는 피동자(patient)를 구분하는 것이다. 이러한 구분을 통해 공리주의는 동물에게 비록 스스로는 도덕 행위의 주체가 될 수 없지만 다른 도덕 행위자의 행동으로 권리가 침해될 수 있는 제한된 도덕성을 부여한다(신상규, 2019, p. 248).

신상규는 공리주의적으로 ‘도덕성’ 개념의 외연이 탈인간중심주의적으로 확장된 것을 다음과 같이 요약한다.

“... 우리 인간이 도덕적으로 중요한 것은 우리가 이성적이기 때문이 아니라, 스스로 생명의 주체(subject-of-a-life)임을 경험하는 존재이기 때문이다. ... 만약 이것이 인간에게 내재적 가치를 부여하고 존중받을 도덕적 권리를 부여하는 근거라면, 우리는 그와 일관되게 생명-주체로서의 지위를 갖는 비인간 동물들에 대해서도 내재적 가치와 함께 단지 수단으로서가 아니라 목적으로 대우받을 도덕적 권리를 부여해야만 한다. ... 이제 도덕적으로 유의미하고 중요한 존재는 단지 나와 다른 인간들로만 국한되지 않는다. 그리고 어떤 기준을 채택하느냐에 따라, 동물들 또한 그 도덕적 이해관계를 고려해야만 하는 도덕공동체의 정당한 일원으로 인정될 수 있다”(신상규, 2019, p. 250).

공리주의에서 말하는 감수 능력은 생물학적 인간과의 유사성으로 결정되지 않는다. 감수 능력의 존재 여부는 고통이나 쾌락의 기능을 실현하는 (하지만 인간의 신경 구조와는 다른 유형의) 신경 구조, 자극에 따른 행동 반응 등을 통해 파악할 수 있기 때문이다. 앞서 언급한 어류의 통증에 관한 과학적 연구는 이에 대한 좋은 근거가 된다.

감수 능력의 존재 여부를 현상적이며 주관적인 경험 주체만이 접근 가능한 의식 상태에 호소하여 결정할 수는 없다. 의식 상태는 개념적으로 의식 주체만이 접근 가능한 상태이므로, 타인의 현상적인 고통에 접근하는 것은 원천적으로 봉쇄된다. 즉, 고통의 의식적 상태에 호소하는 방식은 경험 주체인 나를 제외한 모든 존재의 감수 능력에 대한 인식론적 회의론으로 이어진다.

정리하자면, 공리주의는 기능적 또는 행동적으로 정의된 감수 능력을 가진 대상 모두를 ‘피동적 도덕성’ 개념의 외연에 포함한다. 달리 말하자면, 공리주의는 감수 능력이라는 내재론적 속성에 호소하여 ‘피동적 도덕성’ 개념의 외연을 확장한다.

목광수에 의하면, 공리주의가 내재론적 방법에 따라 ‘피동적 도덕성’ 개념의 외연을 결정하는 것은 다음과 같은 문제점을 갖는다(목광수, 2017, pp. 193-196). 첫째, 내재론적 방법은 배제의 문제를 야기한다. 감수 능력의 존재 여부가 불투명한 개별자들, 예컨대 새우나 혼수상태 인간에게 피동적 도덕성을 귀속할 수 있는지의 문제가 발생하는 것이다. 둘째, 현재의 인공지능 로봇 중 기능적 또는 행동적으로 정의된 감수 능력을 가진 로봇은 존재하지 않는다. 즉, 지금까지 개발된 인공지능 로봇 중 고통이나 쾌락의 기능을 온전히 실현하는 로봇은 없다. 하지만 사람들이 일상에서 인공지능 로봇을 대하는 방식에 관한 연구에 따르면, 사람들은 마치 인공지능 로봇이 감정이나 의도를 가진 것처럼 대하며 나아가 인공지능 로봇에게 피동적 도덕성을 귀속한다(신홍임, 2021, p. 204). 공리주의의 내재론적 방법은

인공지능 로봇의 피동적 도덕성에 관한 대중의 탈인간중심주의적 직관을 잘 반영하지 못한다(신상규, 2019, pp. 260-261). 정리하자면, 내재론적 방법은 특정한 개별자에게 감수 능력을 귀속할 수 있는지에 관한 인식론적 문제와 인공지능 로봇의 피동적 도덕성에 관한 탈인간중심주의적 직관을 충실히 반영하지 못한다는 문제를 야기한다.

내재론적 방법과 달리 관계론적 방법은 감수 능력이라는 내재론적 속성에 호소하지 않는다. 신상규에 의하면, 관계론적으로 피동적 도덕성을 귀속한다는 것은 다음과 같다.

“... 관계론은 도덕적 지위의 문제를 도덕과 관련한 속성과 관련된 형이상학이나 존재론의 문제가 아니라, 일상적인 경험이나 실천의 맥락 속에서 우리가 실제로 그것들을 어떻게 대우하고 상호작용하는가에 초점을 맞춘 현상학이나 해석학의 문제로 접근할 것을 제안한다. 이러한 접근은 “관계”에 입각하여 세계에 대한 더 “올바른” 견해로서의 새로운 존재론을 구성하려는 시도가 아니다. 대신에 이는 주체와 객체 사이의 구체적 관계성, 즉 우리가 인식적, 도덕적으로 AI로봇들과 어떻게 관계 맺는가에 초점을 맞추어 그것들의 도덕적 지위 문제에 접근해야 한다는 제안이다. ... 도덕적 지위는 ... 인간과 다른 존재들이 맺고 있는 관계로부터 창발하는 것이다. ... 도덕적 지위의 변화도 새로운 삶의 양식의 발전(자라남)과 동시에 혹은 그것에 수반하여 일어나는 관계의 사람으로 이해될 수 있다”(신상규, 2019, pp. 261-262).

위에서 언급했듯, 사람들은 마치 인공지능 로봇이 마음을 가진 것처럼 대하고 나아가 인공지능 로봇에게 피동적 도덕성을 귀속한다. 이와 같은 현상은, 관계론에 따르면, ‘피동적 도덕성’ 개념의 외연

이 탈인간중심주의적으로 확장되어야 할 근거가 된다. 목광수도 일종의 관계론적 방법에 따라 인공지능 로봇에게 피동적 도덕성을 귀속할 것을 주장한다. 목광수에 의하면, AI 시대에 적합한 ‘피동적 도덕성’ 개념이란 사람들이 대상을 도덕적 피동자로 인정(recognize)하는 기준에 따른 개념이어야 한다(목광수, 2017, pp. 198-202). 목광수의 인정(recognition)에 근거한 ‘피동적 도덕성’ 개념은 도덕 공동체의 대화와 토론을 통해 그 외연을 인공지능 로봇을 포함하는 방식으로 확장할 수 있다는 점에서 탈인간중심주의적이다(목광수, 2017, p. 203).

관계론적 방법의 문제점은 대중의 직관에 호소하는 방식으로 외연을 확장하거나 축소할 수 없는 개념이 존재한다는 것이다. ‘피동적 도덕성’ 개념의 외연을 결정하려는 경우, 대중이 그러한 개념을 어떻게 사용하는지 파악하여 대중의 직관을 이론에 반영하는 것은 정당한 철학적 작업인 것으로 보인다. 왜냐하면 대중의 직관이 철학적 이론에 반영되어야 한다는 전제 아래에서 이루어지는 대표적인 철학적 작업인 실험철학은 실제로 ‘도덕성’이나 ‘자유의지’ 같은 개념을 주로 다루기 때문이다(Knobe et al. 2012, pp. 81-99). 하지만 모든 개념의 외연이 대중의 직관에 따라 결정될 수 있는 것은 아니다. ‘금’ 개념의 외연이 결정되는 방식을 살펴보자. 먼 과거에는 ‘금’ 개념의 외연이 통속 물리학(folk physically)으로 결정되었다면, 현대의 ‘금’ 개념의 외연은 물리학과 화학을 통한 과학적 방식으로 결정된다(Figdor, 2017, p. 4290). 전자는 “무르고 번쩍이는 광택을 갖는 금속이 곧 금이다”와 같은 금에 대한 대중의 직관적 이해에 의존한다는 점에서 엄격한 방식이 아니다. 황동과 금을 엄격히 구분해야 할 다양한 이유가 있다는 것을 상기한다면, ‘금’ 개념의 외연이 후자의 방식에 따라 엄격하게 결정되어야 하는 것은 당연하다. 즉, ‘금’ 개념의 외연을 과학적으로 결정할 수 있는 방법이 주어진 이상 우리는 대중의 직관에 의존하여 ‘금’ 개념의 외연을 결정할

수 없다. 포스트휴머니즘에 의하면, 도덕성이나 마음에 관한 개념들의 외연은 확정적이지 않다. 하지만 그 외연이 어떤 방식으로 확장되거나 축소되어야 하는가에 대한 합의된 이론 따위는 존재하지 않는다. 연관된 모든 개념의 외연이 온전히 ‘피동적 도덕성’ 개념의 경우처럼 대중의 직관으로 결정될 수 있는 것인지, 아니면 그중에는 ‘금’ 개념 같은 것도 있는지에 대한 심각한 논의가 이루어지지 않은 것이다.

관계론적 방법의 두 번째 문제는 특정 개념에 관한 보편적인 대중적 직관이 존재하지 않는 것으로 보인다. 예를 들어, 알렉산더는 자유의지의 본성과 도덕적 책임에 관한 실험철학 실험 결과를 분석하여 연관된 개념에 관한 보편적인 대중의 직관이란 존재하지 않음을 보였다(알렉산더, 2015, pp. 60-73). 이 세계가 인과적으로 결정되어있더라도 인간은 스스로의 행동에 대해 도덕적 책임을 갖는다는 양립가능론적 직관 그리고 인과적 결정론이 참이라면 도덕적 책임도 존재하지 않는다는 양립불가능론적 직관 중 우세한 직관 따위는 없다는 것이다. 이처럼, 대중의 직관은 개별자, 문화, 성별 등에 따라 다르게 나타난다. 이렇듯 대중의 직관에 호소하여 개념의 외연을 결정하는 것은 개념의 외연이 맥락에 따라 바뀐다는 규범적 상대주의에 노출되어 있다(알렉산더, 2015, pp. 72-73). 하지만 대상의 도덕적 지위가 대상이 처한 맥락에 따라 바뀐다는 입장, 그리고 대상의 심리적 상태가 그 맥락에 따라 바뀐다는 입장 모두 그다지 매력적이지 않다.

지금까지 포스트휴머니즘에 따라 탈인간중심주의적으로 도덕성과 마음 관련 개념들의 외연을 결정한다는 것이 무엇인지 설명했다. 내재론적 방법에 따른 탈인간중심주의는 생물학적 인간의 자기-지시적 속성이 아닌 모종의 내재적 속성에 호소하여 연관된 개념의 외연을 확장하려 한다. 이와 같은 방법은 대상이 특정한 속성을 갖는지의 여부를 알기 어렵다는 인식론적 문제와 탈인간중심주의적 직관을

충실히 반영하지 못한다는 문제를 야기한다. 관계론적 방법에 따른 탈인간중심주의는 사람들이 실제로 비인간 개별자들을 마치 사람처럼 대우한다는 사실에 근거하여 연관된 개념들의 외연이 확장되어야 한다고 주장한다. 이와 같은 방법은 대중의 직관만으로 외연을 결정할 수 없는 개념이 존재한다는 점에서 그리고 대중의 직관은 상대적이라는 점에서 급진적이다.

다음 절에서는 내재론적 측면과 관계론적 측면을 모두 갖는 탈인간중심주의적 의미론인 리터럴리즘(Literalism)을 소개할 것이다. 리터럴리즘은 형식화된 심리적 속성에 호소하여 심리적 개념의 외연을 결정한다는 점에서 내재론적이다. 동시에 리터럴리즘은 과학자 공동체의 전문적 담론에 호소하여 심리적 속성이 귀속되는 대상을 결정한다는 점에서 관계론적이다. 나는 리터럴리즘은 내재론과 관계론 각각이 갖는 문제를 야기하지 않음을 보일 것이다. 달리 말하자면, 나는 두 방법의 장점만 취하고 단점으로부터는 자유로운 리터럴리즘에 따라 심리적 개념의 외연을 확장하는 것이야말로 포스트휴머니즘이 요청하는 탈인간중심주의에 적합함을 논증할 것이다.

Ⅲ. 내재론적이며 관계론적인 탈인간중심주의적 의미론 리터럴리즘

1. 리터럴리즘이란 무엇인가?

심리적 개념의 외연을 탈인간중심주의적으로 확장하는 한 가지 방법은 인간과 비인간 동물 사이의 유사성에 호소하는 것이다. 위에서 논의했듯, 포스트휴머니즘이 요구하는 탈인간중심주의는 생물학적 인간 또는 생물학적 인간의 자기-지시적 속성을 중심에 놓고 그것과의 유사성을 통해 비인간을 심리적 개념의 외연에 포함하면 안 된다. 하지만 생물학

적 인간의 심리적 속성과 전혀 무관한 속성, 예컨대 삼각형의 기하학적 속성 같은 것을 통해 심리적 개념의 외연을 확장할 수는 없다. 탈인간중심주의적으로 심리적 개념의 외연을 확장하려면 인간의 생물학적 속성과 무관한 방식으로 인간과 비인간의 심리적 유사성에 호소해야 한다. 리터럴리즘은 그러한 조건을 만족하는 의미론이다.

케리 피그돌의 리터럴리즘은 심리적 개념의 외연을 생명과학의 발견을 통해 수정하려는 시도이다(Figdor, 2018a).⁶⁾ 다음은 피그돌의 리터럴리즘 요약이다.

“과학자들에 의하면 초파리나 식물은 이리저리한 결정을 내리고(decide) 박테리아들은 언어적으로 소통하고(linguistically communicate) 신경세포는 무언가를 선호하는데(prefer), 리터럴리즘에 의하면 이와 같은 심리적 술어의 예기치 않은 사용에 대한 가장 그럴듯한 해석은 바로 과학자들이 그러한 술어를 문자 그대로의 의미로 사용한다는 것이다”(Figdor, 2018a, p. 1).

심리적 용어에 관한 자연주의 이론 일반에 따르면, 과학자들이 “초파리가 바나나 위에 앉기로 결정했다”라고 말할 때 원리상 그러한 문장을 ‘결정’ 같은 심리적 용어를 사용하지 않는 방식으로 다시 설명할 수 있어야 한다.⁷⁾ 예컨대, 폴 처칠랜드는 ‘결정’ 같은 심리적 용어를 사용하지 않고, 좀 더 정확히 말하자면 신경과학 같은 비심리적 용어만으로

파리의 행동을 설명하는 것이 가능하며 그러한 비심리적 설명만이 진정한 과학적 설명이라고 주장한다(Churchland, 1981, pp. 84-90). 하지만 리터럴리즘에 따르면, 심리적 용어에 관한 자연주의가 반드시 비심리적 용어로의 대체 또는 심리적 용어의 제거를 함축하지는 않는다. 리터럴리즘은 인간이나 비인간 동물의 행동에 대한 설명에 담긴 심리적 용어를 그대로 보존하면서 동시에 그러한 설명에 대한 자연주의적 (또는 과학적인) 근거를 제공한다. 나아가 리터럴리즘에 의하면, 예컨대, 인간의 행동과 초파리의 행동에 관한 과학적 설명 각각에 ‘결정하다’라는 심리적 용어가 사용되었다면, 그와 같은 심리적 용어의 사용은 ‘결정하다’라는 심리적 술어의 외연에 초파리를 포함해야 한다는 좋은 근거가 된다. 리터럴리즘에 따르면, 과학자들의 공인된 과학적 담론은 심리적 술어의 외연을 결정하기 위한 좋은 근거이다.

리터럴리즘은 두 부분으로 구성된다(Figdor, 2018a, p. 5). 첫째, 리터럴리즘에 의하면 심리적 술어의 외연은 과학적 설명의 일종인 수학적 모델로 결정될 수 있으며 그러한 수학적 모델 중에는 인간과 비인간의 구분 없이 적용되는 것이 있다.⁸⁾ 둘째, 특정한 수학적 모델이 인간과 비인간 구분 없이 적용된 경우, 그와 같은 과학적 사실은 그러한 모델과 연관된 심리적 술어의 외연이 탈인간중심주의적으로 확장되어야 한다는 좋은 근거가 된다.

리터럴리즘은 심리적 술어, 즉 태도 귀속(attitude ascription)에 관한 이론이지 내용 귀속(content ascription)에 관한 이론은 아니다(Figdor, 2018a,

6) 리터럴리즘(Literalism)을 우리말로 번역하자면 ‘문자 그대로 주의’가 될 것이다. 하지만 이는 자연스러운 번역이 아니기 때문에 본문에서는 Literalism을 원문 그대로를 우리말로 읽은 ‘리터럴리즘’이라고 표기했다.
 7) 자연주의를 느슨하게 해석하자면, 어떤 현상에 대한 자연주의적 설명은 우리가 이미 이해하는, 예컨대 과학적 용어로만 구성된 설명이다. 심리철학에서의 자연주의란 주로 심리적 행동을 비심리적 용어, 특히 생물학적 용어로 설명하는 것이 가능하다는 입장이다. 심리철학적 자연주의에 대한 자세한 설명은 Dretske(2002)를 참조하라.
 8) 피그돌은 이와 같은 첫 번째 측면을 반-예외주의(Anti-Exceptionalism)라는 리터럴리즘과 구분되는 (또는 리터럴리즘의 형이상학적 측면으로서의) 이름으로 부른다. 본문에서는 반-예외주의와 리터럴리즘을 구분하지 않을 것이다(Figdor, 2018a, p. 5).

p.11). 인간의 행동을 설명하기 위해 명제 태도를 귀속하는 경우를 고려해보자. 예를 들어, 우산을 챙겨 외출하는 철수의 행동에 대해 영희가 “철수는 비가 올 것이라고 믿는다”라는 통속 심리적 설명을 했다고 가정하자. 이와 같은 영희의 설명은 명제적 내용을 통한 태도 귀속, 즉 명제 태도 귀속에 해당한다.⁹⁾ 리터럴리즘이 이와 같은 인간에 대한 명제 태도 귀속을 부정하는 것은 아니며, 리터럴리즘으로부터 초파리에 대한 명제 태도 귀속이 가능하다는 것이 따라 나오지도 않는다. 리터럴리즘에 따르면, “초파리가 바나나 위에 앉기로 결심했다”라는 과학적 설명은 (만약 그러한 과학적 설명이 주어진다면) ‘결심하다’라는 심리적 술어가 초파리에게 귀속되는 방식과 인간에게 귀속되는 방식에는 아무런 차이가 없다는 것의 증거가 된다. 하지만 리터럴리즘이 초파리의 심적 태도의 대상이 명제임을 주장하는 것은 아니다.¹⁰⁾ 리터럴리즘은 내용 이론(theory of content)이 아니므로, 리터럴리즘은 초파리의 마음이 어떤 종류의 내용을 갖는지 설명하지 않는다. 즉, 리터럴리즘은 초파리와 인간이 동일한 결심-능력을 갖는다고 주장할 뿐 그 둘이 동일한 내용적 상태를 갖는다고 주장하지 않는다.¹¹⁾

리터럴리즘을 완전하게 이해하려면 우선 수학적 모델이 무엇인지 이해한 후 그러한 모델이 심리적 설명에 등장한다는 것의 함축을 이해해야 한다. 전자는 ‘결심하다’와 같은 심적 태도에 관한 자연주의적 설명이고, 후자는 그러한 자연주의적 설명이 갖

는 탈인간중심주의적 함축에 관한 것이다. II-2와 II-3에서는 전자를 설명할 것이고, II-4와 II-5에서는 후자를 설명할 것이다.

2. 수학적 모델이란 무엇인가?

마이클 와이스버그에 따르면, 모델이란 실제-세계의 현상을 표상하는 이상화된 구조를 통한 설명 방식으로, 수학적 모델은 어떤 현상의 수학적 표상과 잠재적인 연관 관계에 놓일 수 있는 속성을 갖는 추상적 구조를 통한 설명 방식을 뜻한다(Weisberg, 2013: 8). 피그돌은 이와 같은 와이스버그 식의 모델 이론을 받아들인다(Figdor, 2018a, pp. 34-35).

모델은 일반적으로 구조(structure)와 해석(interpretation 또는 construal)을 기본적 요소로 갖는다. 모델의 구조란 모델이 기술되는 방식에 따라 결정된다. 예를 들어, 의류 디자이너가 그린 옷의 도면은 평면의 단단한(concrete) 구조를 가지고, 보잉 747 비행기의 외형을 본뜬 소형 플라스틱 비행기는 삼차원의 단단한(concrete) 구조를 갖는다(Figdor, 2018a, pp. 34-35). 모델을 만든 사람은 모델의 구조가 무엇을 표상하는지, 즉 그 모델이 무엇의 모델인지 결정해야 하는데, 이를 모델의 해석이라 한다(Weisberg, 2013, p. 45). 이때, 해석 안에는 설명하고자 하는 대상과 모델 사이의 관계 설정도 포함된다. 그러한 관계 설정을 통해 모델이 표상 또

9) 본문의 예를 좀 더 자세히 설명하자면, 영희에 의하면, 철수는 “비가 온다”라는 명제에 대한 ‘믿는다’라는 심적 태도를 취하고 있기 때문에 우산을 챙겨 외출한 것이다. 철수의 입장에서 철수의 심적 상태는 “나는 비가 온다고 믿는다”가 될 것이다.

10) 본문의 예를 좀 더 자세히 설명하자면, 과학자가 “초파리가 바나나 위에 앉기로 결심했다”라는 문장으로 초파리의 행동을 설명할 때, 리터럴리즘에 따르면, 과학자의 이러한 설명은 “초파리는 ‘결심했다’라는 심적 태도를 취하고 있다.”를 함축한다. 하지만 리터럴리즘은 초파리의 심적 태도가 어떤 종류의 내용에 관한 것인지에 대해서는 말해주지 않는다. 리터럴리즘은 심리적 술어 같은 심적 태도에 관한 자연주의 이론인지 명제나 확률적 정보 같은 내용에 관한 자연주의 이론이 아니기 때문이다.

11) 내용 귀속은 주로 심적 내용 또는 지향성에 관한 자연주의 이론에서 다룬다. 자연주의 이론 중에는 명제적 내용을 요구하지 않는 방식으로, 예컨대 확률적 정보 같은 것으로 지향성을 자연화하려는 이론도 있다. 자세한 설명은 Shea(2013)을 참조하라.

는 지시 하도록 고안된 실제-세계의 시스템이 무엇인지 고정되고 나아가 그러한 모델을 만든 사람의 연구 목적과 연관된 대상 시스템의 측면이 무엇인지가 고정된다. 앞의 플라스틱 소형 비행기 모델을 고려해보자. 이 모델은 보잉 747 비행기라는 실제 대상을 표상하도록 만들어졌다. 이 모델의 해석에 따르면, 이 모델이 표상하는 대상의 범위(scope)는 실제 대상의 공학적 측면이 아닌 실제 대상의 외적 형태와 도색에 한정된다. 즉, 이 모델은 단단한 구조를 통해 실제 보잉 747 비행기를 표상하며 그러한 표상 방식은 실제 대상의 외형적 측면에 특화되어 있다.

모델의 해석이 무엇이냐에 따라 모델과 타겟이 되는 실제-세계의 관계에 관한 충실도 기준(fidelity criteria)의 정도가 결정된다(Figdor, 2018a, p. 35). 충실도 기준이란, 역동성(dynamical)과 표상성(representational)을 통해 모델이 타겟에 얼마나 잘 들어맞는지 평가하는 기준이다. 역동성 충실도 기준은 모델이 만들어내는 결과값(output)과 (또는 모델이 만들어내는 예측값(prediction)과) 실제 시스템이 만들어내는 결과값의 차이가 얼마만큼 벌어져도 여전히 그 모델을 타겟에 관한 모델이라고 여길 수 있는가를 결정하는 기준이다(Figdor, 2018a, p. 35). 예컨대, 기상 변화를 예측하는 컴퓨터 프로그램 모델의 가치는 그 모델이 얼마나 높은 역동성 충실도 기준을 만족시키느냐에 따라 결정된다. 표상성 충실도 기준은, 모델의 해석에 따른 모델의 표상 범위가 주어졌을 때, 모델이 실제 시스템에 얼마나 잘 들어맞는지(maps onto) 결정하는 기준이다 (Figdor, 2018a, p. 35). 위의 플라스틱 소형 비행기 모델을 다시 고려해보자. 이 모델의 역동성 충실도 기준은 거의 0에 가까운데, 그 이유는 이 모델은 실제 보잉 747 비행기의 행동 따위를 예측하기 위해 제작된

모델이 아니기 때문이다. 즉, 이 모델은 실제 보잉 747 비행기가 작동하는 방식을 알고 싶은 공학자에게는 아무런 쓸모가 없다. 하지만 이 모델은 실제 보잉 747 비행기의 외형에 관한 높은 표상성 충실도 기준을 만족시키기 위해 제작되었다. 즉, 이 모델은 실제 보잉 747 비행기의 외형에 관심 있는 장난감 수집가에게 유용하다. 정리하자면, 플라스틱 소형 비행기 모델은 타겟의 외형에 관한 높은 충실도 기준만을 만족시킬 뿐, 그 외의 충실도 기준을 만족시키도록 고안된 모델이 아니다(Weisberg, 2013, p. 43).

물리적으로 단단한 구조를 갖는 모델과 달리 수학적 모델은 높은 (역동성 또는/그리고 표상성) 충실도 기준을 만족시키면서 동시에 매우 넓은 범위의 실제 세계 현상들을 설명하는데 사용될 수 있다(Figdor, 2018a, pp. 35-36). 매우 낮은 표상성 충실도 기준이란 마치 공깃돌 하나를 두고 “이것을 한 반도라고 치자”라고 말하는 것과 같다. 충실도 기준이 낮다면 모델을 만든 사람의 의도에 따라 그 공깃돌은 실제-세계 그 무엇이든 표상할 수 있다는 것이다. 플라스틱 소형 비행기 모델은 실제 보잉 747 비행기의 외형에 관한 높은 충실도 기준을 만족하지만, 오직 실제 보잉 747 비행기의 외형만을 표상한다는 점에서 이 모델의 적용성 범위는 매우 좁다. 즉, 플라스틱 소형 비행기 모델이 높은 표상성 충실도 기준을 만족시킨다는 것은 곧 이 모델을 통해 C-130 수송기, HH-60 헬리콥터 등과 같은 것들을 설명하기 어렵다는 것을 뜻한다.¹²⁾ 반면, 수학적 모델은 특정한 수학적 구조를 통해 물리적으로 구분 가능한 다양한 종류의 실제-세계 현상들에 적용될 수 있는데, 이때 수학적 모델은 동일한 (역동성 또는/그리고 표상성) 충실도 기준을 만족하면서 다양한 현상에 적용된다. 예를 들어, 진동에 관한

12) 물론 플라스틱 소형 비행기 모델은 일반적인 헬리콥터의 외형보다 일반적인 비행기 외형을 설명하는 데 더 적합할 것이다. 즉, 표상성 충실도 기준을 낮춘다면 플라스틱 소형 비행기 모델은 비행기 일반의 외형을 설명하기에 적합한 모델이 될 것이다.

수학적 모델은 펜듈럼(pendulum)¹³⁾, 진동하는 입자, 물의 움직임 같은 다양한 물리적 현상들을 설명하는 데 동일한 충실도 기준을 만족하면서 사용된다(Figdor, 2018b, p. 53). 정리하자면, 수학적 모델은 높은 충실도 기준을 유지하는 동시에 다양한 종류의 실제-세계 현상들에 적용 가능하다는 점에서 유연하다(flexible).

3. 로트카-볼테라 모델로 살펴보는 수학적 모델을 통한 개념의 외연 확장

수학적 모델의 유연성을 잘 보여주는 예시로 로트카-볼테라(Lotka-Volterra) 모델을 살펴보자. 로트카-볼테라 모델은 본래 아드리아해에 서식하는 피식자와 포식자 물고기 개체군, 예를 들어 상어-개체군과 대구-개체군의 크기 사이에 성립하는 역동적인 관계를 모델화하기 위해 만들어졌다(Weisberg, 2007, p. 210).

$$\frac{dV}{dt} = rV - (aV)P, \quad (1)$$

$$\frac{dP}{dt} = b(aV)P - mP. \quad (2)$$

본래 목적에 따라, 위 모델은 두 이상화된 물고기 개체군과 그러한 개체군의 상대적 크기에 관한 관계성을 수학적으로 기술한 것으로 구성된다. 이때, 위 모델의 해석에 따르면, 이 모델의 타겟은(또는 이 모델이 설명하려는 실제-세계 현상은) 두 종류의 실제-세계 물고기 개체군과 두 개체군 사이에 성립하는 특정한 실제-세계 관계이다(Figdor, 2018a, pp. 36-37).

다음은 위 모델의 의미를 해석하는 것에 관한 피그돌의 설명을 그대로 옮긴 것이다. 로트카-볼테라 모델의 본래 목적에 따라, V는 대구(피식자 물고

기)-개체군의 크기를 표상하고 P는 상어(포식자 물고기)-개체군의 크기를 표상한다고 하자. 매개 변수 r은 대구-개체군의 내재적(비상대적) 성장률을 표상하고, m은 상어-개체군의 내재적인 생존률, a는 대구의 피식률, b는 상어의 출산율을 각각 표상한다고 하자. 수식 (1)은 시간의 경과에 따른 대구-개체군의 크기 변화, 즉 좌 항이 대구-개체군의 성장률과 대구의 상어에 대한 피식률 사이의 차이, 즉 우항과 같다는 등식이다. 수식 (2)는 시간의 경과에 따른 상어-개체군의 크기 변화, 즉 좌 항이 대구의 상어에 대한 피식률에 비례한 상어-개체군의 출산율과 상어-개체군의 사망률 사이의 차이, 즉 우항과 같다는 등식이다. 수식 (1)과 (2)는 대구의 피식률을 통해 서로 연결된다. 즉, 수식 (2)에서 대구-피식률의 상어-개체군 출산률로의 변환은 수식 (1)에서의 대구의 상어에 대한 피식률의 함수이다. 간략히 설명하자면, 위 모델에 따르면 상어-개체군의 크기 증가는 대구-개체군의 크기 증가에 비례한다. 로트카-볼테라 모델은 상업적 어획이 중단된 1차 세계 대전 동안 아드리아해에서 발생한 피식자와 포식자 물고기 개체군의 크기 변화를 잘 설명했다는 점에서 경험적으로 증명되었다(Figdor, 2018a, p. 37).

피그돌에 따르면, 로트카-볼테라 모델이 물고기에 관한 모델인 이유는 다음과 같다(Figdor, 2018a, p.37). (i) 이 모델은 물고기에 관한 설명을 하도록 만들어졌다. (ii) 이 모델을 만든 사람에 의하면, 수식에 사용된 기호, 예를 들어, V는 피식자-물고기를 나타낸다. (iii) 물고기가 무엇인지는 누구나 잘 알고 있으며, (iv) 우리는 ‘물고기’라는 개념을 통해 물고기를 지시한다. 정리하자면, 로트카-볼테라 모델 같은 수학적 모델의 의미론적 성격은 자연 언어의 의미론적 성격과 별반 다르지 않다.

흥미로운 점은 ‘피동적 도덕성’ 같은 일상적 개념

13) 펜듈럼이란 시계의 진자나 흔들리는 추를 의미한다.

의 외연에 관한 의미론적 문제가 발생하는 원인과 수학적 모델의 외연에 관한 의미론적 문제가 발생하는 원인이 별반 다르지 않다는 것이다.

“의미론적 문제는 (만약 그런 문제가 발생한다면) 우리가 어떤 모델이 적용되는 방식을 그러한 모델이 본래 적용되었던 대상과 별다른 유사성이 없는 새로운 대상들에게까지 확장하려 할 때 발생하거나, 또는 (달리 말하자면) 어떤 모델의 용어들이 그 모델의 본래 해석에 따라 특정하는(specified) 실제-세계 관계들과는 전-이론적(pretheoretic) 관점에서 (이러한 관점은 종종 질적 유사성(qualitatively-based)에 근거한다) 전혀 이질적인 대상의 집합에 대해 그 모델을 적용할 때 발생한다”(Figdor, 2018a, p.38).

수학적 모델의 외연에 관한 의미론적 문제는 어떤 모델의 본래 해석이 특정하는 대상 또는 대상들의 관계와 직관적으로 다른 대상 또는 대상들의 관계에까지 그 모델의 적용성을 확장하려 할 때 발생한다. 이와 같은 의미론적 문제는 본래 지시하는 대상 또는 대상들의 관계와 새로운 대상 또는 대상들의 관계 사이의 이질성에 따라 그 심각성이 결정된다(Figdor, 2018a, pp. 38-39). 예컨대, 로트카-볼테라 모델의 외연은 여우-개체군과 토끼-개체군 사이의 관계, 늑대-개체군과 무스-개체군 사이의 관계, 식물-개체군들 사이의 관계 등 생물학적 ‘포식(predation)’과 연관된 다양한 관계를 포함하는 방식으로 확장된다. 이와 같은 확장은 물고기-개체군 사이의 포식-관계와 육지동물-개체군 사이의 포식-관계가 (또는 식물-개체군 사이의 포식-관계가) 직관적으로 별반 다를 것 없어 보인다는 점에서 자연스럽다. 이번에는, P가 노동소득분배율(labor share of national income)을 표상하고 V가 고용률(employment rate)을 표상하는 경우를 고려해보자. 노동소득분배

율과 고용률은 생물학적 ‘포식’과 직관적으로 아무런 연관성이 없어 보이지만, 로트카-볼테라 모델은 이러한 이질적인 대상들의 관계에 대해서도 성공적으로 적용된다(Arora and Boer, 2006, p. 1). ‘피동적 도덕성’ 개념의 외연에 강아지를 포함하는 것보다 초파리를 포함하는 것이 더 반직관적이므로, 후자의 경우 상대적으로 심각한 의미론적 문제가 발생한다. 마찬가지로, 로트카-볼테라 모델의 외연에 여우와 토끼를 (또는 식물들) 포함하는 것보다 노동소득분배율과 고용률을 포함하는 것이 더 반직관적이므로, 후자의 경우 상대적으로 심각한 의미론적 문제가 발생한다.

로트카-볼테라 모델의 반직관적 유연성은 생물학적 ‘포식자’와 연관된 대상들과 경제적 대상들 사이에 (로트카-볼테라 모델을 통해 드러난) 형식적 유사성이 성립한다는 것을 뜻한다. 피그들에 의하면, 이와 같은 형식적 유사성은 위에서 논의한 경제적 대상들도 ‘포식자’ 개념 안에 포함해야 한다는, 즉 ‘포식’ 개념의 외연이 그러한 경제적 대상들을 포함하는 방식으로 확장되어야 한다는 주장의 근거가 된다(Figdor, 2018a, pp. 40-41). 위의 인용문에서 언급했듯, 로트카-볼테라 모델의 의미론적 문제가 발생하는 이유는 새로운 대상들이 본래 지시되던 대상들과 전-이론적인 관점에서 다르기 때문이다. 즉, ‘포식’에 대한 직관적 이해에 의하면 그러한 개념은 생물학적인 대상들만을 지시해야 하는 것처럼 보이는데, 경제학적 대상은 생물학적 대상이 아니므로 의미론적 문제가 발생한다는 것이다. 하지만 이와 같은 전-이론적인 직관은 별다른 근거 없이 우리에게 주어진 것이다. 반면, 로트카-볼테라 모델의 (반직관적) 유연성은 형식적 유사성이란 (경험적으로 증명된) 과학적 근거를 갖는다. 로트카-볼테라 모델이 동물, 식물, 경제학적 대상 등에 관한 넓은 범위의 실제-세계 현상을 설명하는데 사용된다는 사실은 로트카-볼테라 모델의 의미론적 함축을 반영한다. 달리 말하자면, 위의 논의는 우리가 ‘포식’

개념에 관한 전-이론적 직관을 폐기하고 로트카-볼테라 모델의 과학적이고 형식적인 유연성에 따라 ‘포식’ 개념의 외연을 확장해야 함을 말해준다.

4. 드리프트-디퓨전 모델에 따른 ‘결정’ 개념의 외연 확장

이번에는 인간과 비인간 동물의 행동에 관한 심리학적이며 동시에 수학적 모델인 드리프트-디퓨전 모델(DDM)을 살펴보자. DDM은 본래 인간 피실험자가 두 선택지 중 단 한 번의 시도를 통해 어느 것을 고를지 1, 2초 안에 결정하는 인지적 과정의 결정-속도와 결정-정확성과 관련된 여러 인지적 요인들 사이의 트레이드오프(tradeoff)를 측정하는 모델로 고안되었다(Figdor, 2018a, p. 43). 연관된 실험에서 피실험자는 결정을 내리는 데 필요한 증거, 즉 반응 기준(response criteria)을 만족하는 증거가 충분히 축적되면 결정을 내리고 그에 따른 행동을 하도록 지시받는다(Figdor, 2018a, p. 44). 다음은 연관된 실험의 예시이다. 피실험자에게 표본(sample) 시각-자극으로 집이 찍힌 사진을 보여주었다. 약간의 시간이 지난 후 피실험자에게 실험(test) 시각-자극을 주고 그 실험 시각-자극이 표본 시각-자극과 같은지 빠르고 정확하게 버튼을 눌러 결정하라고 지시했다. 즉, 피실험자는 실험 시각-자극으로부터의 증거를 기반으로 최대한 빠르고 가능한 한 정확하게 그러한 실험 시각-자극이 자신의 기억에 저장된 표본 시각-자극과 일치하는지 판단하고 그 결정을 행동으로 옮겨야 한다. 표본 시각-자극으로부터 주어지는 증거의 누적률을 드리프트율(drift rate)이라고 하고, 피실험자가 결정을 내리는데 주어진 시간 제한을 결정 역치(decision threshold)라고 한다. 결정 역치와 드리프트율은 실험자의 의도에 따라 조작할 수 있다. 후자의 경우, 실험자는 실험 시각-자극의 질을 조절하여 드리프트율을 조작한다. 실험 시각-자극의 질은 실험의 난이도를 결정한다. 낮은

질(quality), 즉, 선명도가 떨어지는 또는 잡음(noise)이 섞인 자극은 실험의 난이도를 높이고, 높은 질, 즉 선명하고 잡음이 섞이지 않은 자극은 난이도를 낮춘다. 난이도가 낮아질수록 드리프트율은 올라가고, 난이도가 높아질수록 드리프트율은 내려간다. 일반적으로 난이도가 높아질수록 평균 결정-시간은 증가하고 (즉, 결정-속도가 느려지고) 결정-정확성은 낮아진다. DDM은 이와 같은 트레이드오프를 예측하는 모델이다. 즉, DDM은 실험 시각-자극의 질적 변화에 따른 실험의 난이도에 상응하는 결정-속도와 결정-정확성의 수준(level)과 분포도를 예측한다 (Figdor, 2018a, p. 44). 예시의 실험 결과의 데이터와 DDM의 예측이 어느 정도 일치한다는 것은, 이 수학적 모델이 데이터에 들어맞거나(fit) 데이터를 수용한다는(accommodate) 것으로 표현된다.

DDM이 실험 결과의 데이터에 들어맞는지 결정하기 위해서는 실험 결과로부터 결정-시간을 알아낼 수 있어야 한다. 실험 결과가 일차적으로 알려주는 것은 피실험자가 버튼을 누르는 것 따위의 행동을 포함하는 반응-시간이다. 하지만 반응-시간은 결정의 인지적 과정과 관련 없는 요소를 포함한다. 반응-시간의 구성 요소를 형식적으로 표현하자면 다음과 같다(Figdor, 2018a, p. 45).

반응-시간(reaction time, 줄임말로 RT)은
엔코딩-시간(encoding time, 줄임말로 u), 결정-
시간(decision time, 줄임말로 d), 그리고 반응-
출력-시간(response output time, 줄임말로 w)
으로 구성된다.

$$RT: u + d + w.$$

u와 w는 결정의 인지적 과정과 관련 없는, 즉 비-결정 구성 요소이므로, 실험 데이터에 DDM을 적용하려면 RT에서 d를 추출해야 한다. 이러한 추출 작업은 드리프트율과 반응역치를 조작하여 u, d, w

각각의 값의 변화를 끌어내는 것을 통해 이루어진다.14) 이를 통해 최종적으로 DDM이 d를 기반으로 한 데이터에 들어맞는지 살필 수 있다.

이제 DDM으로 초파리의 결정 과정 중 발생하는 트레이드오프를 예측하는 실험을 살펴보자. FoxP 유전자는 인간의 언어 능력과 말하기 능력의 장애와 연관이 있는 유전자이다(Figdor, 2018a, p. 45). FoxP 결함이 있는 사람이 결정을 내릴 때의 인지적 과정과 건강한 사람의 그것 사이에는 차이가 존재하는데, 그러한 차이에 관한 데이터는 DDM에 잘 들어 맞는다(Figdor, 2018a, p. 45). 피그들은 인위적으로 만들어낸 FoxP 변이(mutant) 초파리와 야생형(wild-type) 초파리의 행동을 비교한 실험을 소개하고 있다(Figdor, 2018a, pp. 46-47). 다음은 연관된 실험에 대한 피그들의 설명을 그대로 옮긴 것이다. 이 실험은 메틸시클로헥사놀(4-methylcyclohexanol)의 특정한 냄새 농도에 (즉, 표본 화학-자극에) 반응하여 회피 행동을 보이도록 조건화된(conditioned) FoxP 변이 초파리와 야생형 초파리가 실험 화학-자극에 반응하는 것을 측정하는 실험이다. 즉, 실험에서 초파리는 실험 화학-자극의 농도가 표본 화학-자극의 농도에 도달했는지 결정하고 그러한 결정에 따라 회피 행동을 실행해야 한다. 실험자는 실험 화학-자극의 질을 조절하여 실험의 난이도를 결정했다. 초파리는 냄새가 유입되는 구멍과 빠져나가는 구멍이 있는 방(chamber)에 들어있으며, RT(반응-시간)는 초파리가 방 안에 마련된 특정한 결정-영역 안으로 들어가고 나오는 시간을 통해 측정된다. 실험 결과의 데이터는 아래의 DDM에 잘 들어맞는 것으로 밝혀졌다.

$$T(x) = \frac{A}{kx} \tanh(Akx) + t_{residual} \quad (1)$$

$$F(x) = F_{\max} \frac{1}{1 + e^{-2Akx}}. \quad (2)$$

T(x)는 자극-강도의 함수로서의 RT이고, F(x)는 자극-강도의 함수로서의 정확성이다. 두 함수 모두 x에 관한 함수라는 점에서 연결되며, 이때 x는 (실험의 난이도를 결정하는) 냄새 농도, 즉 자극-강도 (또는 실험 화학-자극)이다. k는 드리프트율로, x를 조작하여 난이도를 올리면 k가 내려가고 x를 조작하여 난이도를 내리면 k가 올라간다. A는 결정 바운더리(decision boundary), 즉 초파리의 결정이 회피-행동으로 이어졌는지 아니면 비회피-행동으로 이어졌는지를 나타낸다. Fmax는 실험의 난이도가 가장 낮을 때 (즉 실험 화학-자극의 질이 가장 높을 때) 성공적인 결정 횟수의 비율을 실제 실험을 통해 얻은 값이다.15) 첫 번째 등식(RT)의 우항은 두 변항으로 구성되는데, 좌측은 결정-시간이고 우측(residual)은 비-결정 구성 요소의 집합이다. 실제 RT 데이터를 T(x)에 대입하면 A와 k 각각의 값을 얻을 수 있는데, 이러한 값을 F(x)에 대입하면 정확성을 얻을 수 있다.

위 실험에 의하면, FoxP 변이 초파리가 결정을 내릴 때의 인지적 과정과 야생형 초파리의 그것 사이에는 차이가 존재하며 그러한 차이에 관한 데이터는 DDM에 잘 들어맞는다. 달리 말하자면, 위 실험은 (실험 화학-자극의 질에 따른) 결정의 난이도와 (그러한 난이도에 따른 올바른 선택 횟수의 비율로서의) 파리의 행동 사이의 관계가 DDM에 잘 들어맞는다는 사실을 증명한다. 위 실험의 DDM이 본래 인간이 결정을 내릴 때의 인지적 과정을 예측하기 위해 만들어졌다는 사실을 상기하자. 이러한 사실은, 위의 로트카-볼테라 모델이 물고기와 경제 적 대상 모두에게 적용되는 것처럼, DDM도 인간과

14) 예를 들어, 실험 시각-자극의 질을 높이거나 낮추는 것은 결정-시간에는 영향을 미치겠지만 (결정에 대한 인지적 과정이 끝난 후 버튼을 누르는 것에 관한) 반응-출력-시간에는 영향을 미치지 않을 것이다.

15) Fmax에 관한 전체 실험이 10일 때 성공 횟수가 5라면, 본문의 비율은 0.2가 된다.

초파리 모두에게 적용됨을 함축한다. 피그돌에 의하면, DDM의 이와 같은 과학적이고 형식적인 유연성은 ‘결정’ 개념의 외연이 초파리를 포함하는 방식으로 확장되어야 한다는 의미론적 함축을 갖는다. 달리 말하자면, 과학자들이 “실험에서 FoxP 변이 초파리가 “결정”을 내리는 시간은 야생형 초파리의 그것에 비해 좀 더 길다”라고 보고했을 때 그들이 사용한 “결정”이라는 심리적 용어의 의미론은 자연주의적으로 보존되며 (또는 자연주의적으로 비심리적 용어로 대체되거나 제거될 필요가 없으며), 그러므로 우리는 과학자들의 보고에서 겁따옴표를 제거해야 한다(Figdor, 2018a, pp. 47-48).

위에서 논의했듯, 객관적이고 정량적으로 접근 가능한 속성에 호소하여 심리적 개념의 외연을 확장하는 것은 인간의 생물학적 속성에 호소하지 않는다는 점에서 탈인간중심주의적이다. 수학적 모델인 DDM은 (실험 결과의 데이터에 따라) 대상에게 ‘결정하다’라는 심적 태도를 귀속할 수 있는 과학적인, 즉 객관적이고 정량적인 근거를 제공한다. 이와 같은 과학적인 근거는 ‘결정하다’에 관한 우리의 전-이론적인 직관보다 인식론적으로 우월하다.

“실제 관계와 단순한 유사성 관계 사이의 형이상학적 차이는 어떤 술어의 특정한 외연은 자연스럽고 다른 외연은 자연스럽지 않다는 느낌을 통해 결정되지 않는다. 질적인 유비와는 달리 정량적 유비(quantitative analogy)는 심리적 술어의 실제 외연에 관한 우리의 전통적 이해를 수정해야 할 새로운 그리고 강한 인식론적 이유를 제공한다. 모델의 유연성과 표준적인 모델적 설명의 양상을 고려할 때, 정량적 유비는 심리적 용어에 대한 우리의 전통적인 인간중심주의적 의미론에 직접적으

로 도전한다”(Figdor, 2018a, pp. 42-43).

우리의 인간중심주의적 직관에 따르면, 인간은 결정하지만 초파리는 “결정”한다. (또는, 초파리는 마치 결정하는 것처럼 보이기기는 하지만 그러한 초파리의 행동은 진짜 결정-행동은 아니다.) 하지만 이와 같은 전-이론적인 인간중심주의적 직관이 어디에 근거를 두고 있는지는 확실하지 않다. 반면, DDM의 과학적이고 형식적인 유연성은 인간의 결정과 초파리의 결정이 같은 인지적 과정이라는 것에 대한 과학적 증거가 된다.¹⁶⁾ 결론적으로 우리는 ‘결정하다’에 관한 인간중심주의적 직관을 폐기하고 그 외연을 초파리 같은 비인간 동물을 포함하는 방식으로 확장해야 한다. 달리 말하자면, ‘결정하다’와 같은 심적 태도가 인간의 전유물이 아니라는 결론은 곧 연관된 심리적 개념의 외연을 탈인간중심주의적으로 확장해야 한다는 좋은 근거가 된다.

5. 리터럴리즘의 내재론적이고 관계론적인 심리적 개념 외연 확장

DDM과 같은 수학적 모델에 호소하여 심리적 개념의 외연을 확장하는 것은 위에서 논의한 내재론적 방법의 문제점을 야기하지 않는다. 위의 논의에 따르면, 내재론적 방법은 대상이 특정한 속성을 갖는지의 여부를 알기 어렵다는 인식론적 문제와 탈인간중심주의적 직관을 충실히 반영하지 못한다는 문제를 야기한다. 바로 위에서 논의했듯, DDM과 같은 수학적 모델은 인간의 생물학적 속성과 무관하다는 점에서 탈인간중심주의적이다. 나아가, 수학적 모델은 추상화된 구조를 통해 객관적인 방식으로 실제-세계 현상에 적용된다는 점에서 DDM에 호소하는 것은 인식론적 문제를 야기하지 않는다.

16) DDM 같은 수학적 모델이 포착하는 속성의 형이상학적 본성이 무엇인지는 본 논문의 주제와 별다른 관련이 없다. 예를 들어, DDM은 대상의 내적인 표상적 상태를 포착하는 것일 수도 있고 일종의 인지적 과정으로서의 활동(activity)을 포착하는 것일 수도 있다(Figdor, 2018a, p. 108, 각주 4).

달리 말해, DDM은 대상이 ‘결정하다’라는 심적 태도를 갖는지의 여부를 과학의 객관적이고 정량적인 방식으로 알려준다. 종합하자면, DDM과 같은 수학적 모델은 포스트휴머니즘의 탈인간중심주의에 적합한 내재론적 방법이다.

위에서 논의했듯, 관계론적 방법에 따른 탈인간중심주의적 개념 외연 확장 방법에 의하면 심리적 개념 외연의 확장은 사람들이 실제로 비인간 개별자들을 마치 사람처럼 대우한다는 사실에 근거할 수 있다. 달리 말하자면, 관계론적 방법은 심리적 개념이 실제로 사용되는 맥락에 근거하여 그 외연을 확장하는 방법이다. 이러한 관계론적 방법은 검증되지 않은 그리고 맥락에 따라 달라지는 대중의 직관에만 의존한다는 점에서 문제를 야기한다. 하지만, 지금까지 살펴본 수학적 모델을 통한 탈인간중심주의적 개념 외연 확장은 전문가들의 담론에 호소한다는 점에서 이러한 문제를 야기하지 않는다.

“비록 내가 강조하지는 않았지만, 지금까지 논의한 [수학적 모델에 관한] 모든 사례는 동료 과학자를 일차적인 독자로 삼는 공개-심사(peer-reviewed) 학술지에 실린 논문에 보고된 사례들이다. ... 리터럴리즘에 의하면, 글자 그대로의 의미를 갖는 심리적 술어는 인간과 비인간 영역(domain) 모두에 걸쳐 동일하게 적용되는 과학적으로 발견된 구조를 가리키기 위해(pick up) 사용된다. ... 이것이 함축하는 바는, 지금까지 우리가 당연하게 받아들여 온 지시체의 동일성에 관한 인간중심적 기준에 대한 지시체적 수정(referential revision)이 이루어졌다는 것이다. ... “결정-내리기”와 같은 심리적 술어의 지시체적 수정은 우리가 그러한 용어를 통해 지시하고자 하는 본래(original) 의도를 좀 더 정확하게 만들어준다. 상당수의 심리적 술어를 통해 우리가 본래 지시하고자 하는 의도는 어쨌든 전혀 결정적이지 않은 것

이다”(Figdor, 2018a, pp. 60-62).

리터럴리즘은 과학자 공동체의 전문적 담론에 해당하는 권위 있는 과학 학술지에 근거하여 심리적 술어의 외연에 대한 확장을 시도한다. 즉, 리터럴리즘은 심적 개념이 실제로 사용되는 과학적 맥락에 근거하여 (관계론적으로) 심적 개념의 외연을 확장하려 한다. 이와 같은 리터럴리즘은 안락의자에서 이루어지는 철학적 작업이 아니다. 리터럴리즘은 과학자 공동체의 심사를 통과한 여러 논문에 등장하는 다양한 사례들에 대한 분석에 근거한 탈인간중심주의이다. 예컨대, DDM에 관한 다양한 연구에 따르면 수학적 모델인 DDM은 인간과 초파리 모두에게 적용되는데, 이러한 사실은 ‘결정하다’라는 심리적 술어의 외연이 탈인간중심주의적으로 확장되어야 할 과학적 근거가 된다는 것이다. 마치 ‘금’ 개념의 의미가 ‘무르고 광택이 나는 금속’과 같은 일반인의 직관이 아닌 ‘원자 번호 79인 원소 Au’와 같은 화학적 발견에 따라 수정되어온 것처럼, ‘결정하다’ 개념의 의미도 인간과 초파리에 관한 직관적 차이가 아닌 DDM에 관한 과학적 담론에 근거하여 수정되어야 한다고 피그들은 주장하고 있다.

종합하자면, 리터럴리즘은 수학적 모델의 인간과 비인간 영역을 관통하는 형식적 유연성을 통해 인간과 비인간 동물 모두에게 동일한 심리적 술어를 적용할 객관적인 근거를 마련한다. 이와 같은 근거는 과학자 공동체의 전문적 담론을 통해 검증되었다는 점에서 우리의 인간중심적인 전-이론적 직관보다 인식론적으로 우월하다. 결론적으로 리터럴리즘으로 개념의 외연을 결정하는 방식은 내재론적 방식과 관계론적인 방식의 장점만을 모두 취하고 있다.

IV. 결론

서론에서 제시한 문제에 따르면, 과학의 발전에

따라 심리적 개념의 외연이 확장되고 있으며 미래 사회는 더 빠르고 거대한 과학적 발전을 마주할 것이므로 그에 적절한 탈인간중심주의적 의미론이 요구된다. 본 논문을 통해 나는 리터럴리즘이 바로 그러한 의미론임을 논증했다. 포스트휴머니즘이 요청하는 탈인간중심주의란 생물학적 인간의 자기-지시적 속성으로서의 본질을 중심으로 위계화된 인간과 비인간 구분의 부정이다. 리터럴리즘은 심리적 개념의 외연을 탈인간중심주의적으로 확장한다. 리터럴리즘은 비인간 영역에 관한 과학 논문에서 사용되는 심리적 술어의 의미를 보존해야 하는 이유를 알려준다. 리터럴리즘에 따르면, 과학자들이 초파리는 결정하고 신경세포는 선호(prefer)한다고 말할 때 우리는 그러한 심리적 술어를 문자 그대로(literally) 받아들여야 한다.¹⁷⁾ 즉, 리터럴리즘은 과학적 발전에 따른 연관된 심리적 개념 외연의 탈인간중심주의적 확장의 정당성을 마련한다.

리터럴리즘이 단순히 내재론적이고 관계론적인 장점만을 갖는 것은 아니다. 리터럴리즘의 또 다른 장점 중 하나는 생물학적 인간의 예외성의 여지를 남기는 동시에 생물학적 인간의 영역과 비인간 영역의 경계를 허문다는 점이다. 리터럴리즘은 심적 태도에 관한 이론이지 심적 내용에 관한 이론은 아니다. 예를 들어, “초파리가 이리저리한 내용에 관해 결정했다”라는 과학적 진술과 “철수가 우산을 챙겨서 외출하기로 결정했다”라는 일상적 진술에서 리터럴리즘에 따라 보존되는 것은 ‘결정했다’라는 심리적 술어의 의미일 뿐이다. 리터럴리즘은 초파리의 심적 태도가 인간과 같은 명제적 내용을 갖는다고 주장하지 않는다. 인간의 명제적으로 사고하는 능력은 (만약 명제주의적 심리철학이 옳다면) 인간 고유의 영역으로 남을 수 있다.¹⁸⁾

물론 DDM을 통한 ‘결정한다’라는 심리적 개념의

외연 확장이 인공지능 같은 기술적 존재에까지 적용될 수 있을지는 확실하지 않다. DDM이 인간의 경제적 행동과 초파리의 지각적 행동 모두를 설명할 수 있는 이유는 어쩌면 진화론적 유사성에 있을지도 모르기 때문이다(노형래, 2021, p. 1). 즉, 리터럴리즘의 탈인간중심주의가 모든 형태의 비인간 개별자에게 적용되는지는 아직 알 수 없다.

리터럴리즘의 개념 외연 확장 방식은 탈인간중심주의적 개념 확장 방식의 하나일 뿐이다. 인공지능 로봇과 같은 기술적 존재를 심리적 개념의 외연 안에 포함할 것인가의 여부는 그에 맞는 탈인간중심주의적 의미론으로 결정될 수 있다. 예컨대, 코켈버그는 연관된 개념이 일상적으로 사용되는 방식에 호소하여, 즉 대중의 직관에 따른 관계론에 호소하여 연관된 개념의 외연을 인공지능 로봇에게까지 확장한다(Coeckelbergh, 2014, p. 61).

본 논문을 통해 나는 리터럴리즘이 포스트휴머니즘이 요청하는 탈인간중심주의의 기본적 요건을 만족시키므로, 우리는 초파리 같은 개별자도 심리적 외연에 포함해야 한다는 것을 보였다. 이다음 연구는 인공지능 로봇과 같은 기술적 존재자들을 도덕적 그리고 심리적 개념들의 외연에 포함할 수 있는 탈인간중심주의적 의미론을 찾는 작업이 될 것이다.

참고문헌

- 강영안(1991). 현대철학의 반데카르트적 경향. 철학과 현실, 8, 239.
- 강영안, 이상현(2013). 포스트휴머니즘에 관한 철학적 성찰. 지식의 지평, 15, 150-171.
- 노형래(2021). 진화론적 기능 개념에 따른 복수실현 문제. 철학적 분석, 45, 1-31.

17) 신경세포의 사례에 관해서는 Figdor (2018a, pp. 48-55)를 참조하라.

18) 물론 리터럴리즘이 초파리는 명제적으로 사고할 수 없다고 주장하는 것도 아니다. 리터럴리즘은, 적어도 피그들의 주장에 따르면, 내용 이론 일반과 양립 가능하다.

- 목광수(2017). 인공지능 시대에 적합한 인격 개념. *철학논총*, 90(4), 187-208.
- 신상규(2019). 인공지능의 도덕적 지위와 관계론적 접근. *철학연구*, 49, 243-273.
- 신상규(2020, 1, 23). 포스트휴먼과 포스트휴머니즘, 그리고 삶의 재발명. *Horizon*, <https://horizon.kias.re.kr/12689/>.
- 신홍임(2021). 누가 마음을 갖고 있는가? 로봇에 대한 마음지각과 도덕 판단. *사회과학연구*, 32(1), 195-213.
- 이원봉(2018). 포스트휴머니즘은 휴머니즘이 될 수 있는가? 포스트휴머니즘 논쟁을 통해 본 휴머니즘의 의미와 한계. *인간연구*, 37, 57-83.
- 정윤영(2019). 포스트휴머니즘과 휴머니즘에 기반한 교육 재고. *교육철학연구*, 41(3), 117-147.
- Arora, V. K. & Boer, G. J. (2006). Simulating Competition and Coexistence between Plant Functional Types in a Dynamic Vegetation Model. *Earth Interactions* 10(10), 1-30.
- Calvo Garzón, Francisco. (2007). The Quest for Cognition in Plant Neurobiology. *Plant Signaling & Behavior*, 2(4), 208-211.
- Churchland, P. M. (1981). Eliminative Materialism and the Propositional Attitudes. *Journal of Philosophy*, 78(2), 67-90.
- Coeckelbergh, M. (2014). The Moral Standing of Machines: Towards a Relational and Non-Cartesian Moral Hermeneutics. *Philosophy of Technology*, 27, 61-77.
- Dretske, F. (2002). A Recipe for Thought. Chalmers, David(Eds.), *Philosophy of mind: classical and contemporary readings*(491-499). Oxford University Press.
- Figdor, C. (2017). On the proper domain of psychological predicates. *Synthese*, 194, 4289-4310.
- Figdor, C. (2018a). *Pieces of Mind: The Proper Domain of Psychological Predicates*. Oxford University Press.
- Figdor, C. (2018b). The Fallacy of the Homuncular Fallacy. *Belgrade Philosophical Annual*, 31, 41-56.
- Knobe, J., Buckwalter, W., Nichols, S., Robbins, P., Sarkissian, H., & Sommer, H. (2012). Experimental Philosophy. *Annual Reviews of Psychology*, 63, 81-99.
- Margolis, E., & Laurence, S. (2019, 6, 17). Concepts. *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/entries/concepts/>.
- Prescott, Tony J., & Robillard, Julie M. (2021). Are friends electric? The benefits and risks of human-robot relationships. *iScience*, 24(101993), 1-14.
- Shapiro, Lawrence., & Spaulding, Shannon. (2021, 6, 25). Embodied Cognition. *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/entries/embodied-cognition/>.
- Shea, Nicholas. (2013). Naturalizing Representational Content. *Philosophy Compass*, 8(5), 496-509.
- Sneddon, Lynne U. (2015). Pain in aquatic animals. *The Journal of Experimental Biology*, 218, 967-976.
- Weisberg, M. (2007). Who is a Modeler?. *British Journal of Philosophy of Science*, 58, 207-233.
- Weisberg, M. (2013). *Simulation and Similarity: Using Models to Understand the World*. Oxford University Press.

투고일자 : 2022. 05. 15.

심사일자 : 2022. 06. 01.

게재확정일자 : 2022. 06. 30.

Literalism as Non-anthropocentric Semantics

Hyungrae Noh

Sunchon National University

Psychological concepts are not immutable categories of universality. For example, the proper extension of the concept of pain has been expanded from animals with human-like neural mechanisms to fish whose neural mechanistic bases of pain are different from those of human. This example shows that proper extensions of psychological concepts can be scientifically expanded in ways that are irrelevant to biological properties of human. The future society will face scientific progresses that are much more faster and greater than those of present, and we should be able to revise proper extensions of the relevant concepts accordingly. I argue that Literalism can play the role of semantic revision. According to Literalism, psychological categories are determined in terms of the relevant scientific contexts where psychological concepts are used. For instance, cognitive scientists predict economic behavior of human beings and perceptual behavior of fruit flies by the same mathematical model. Consider the statements “the fruit fly decided to fly east” and “Charles decided to invest in stocks”. Provided that the same mathematical model is used to explain behaviors of the fruit fly and Charles, Literalism holds that the ‘decide’ in the former statement is semantically identical to the ‘decide’ in the latter statement. Literalism provides us with a theoretical justification of the non-anthropocentric, scientific expansion of proper extensions of psychological concepts.

Keywords: Philosophy, Posthumanism, Non-anthropocentrism, Literalism, Psychological Concepts, Mathematical Models, Semantics