

표상적 심리론의 이론적 특징과 그 전망*

김 광 수

(서울대 철학과 대학원)

1. 들어가기 : 논의의 배경

인간의 행동을 믿음이나 욕구 같은 지향적 심리 상태를 원인 삼아 설명하는 방식을 상식 심리학(folk psychology)이라 한다. 이를테면 철수가 민주당에 투표한 행동이 세금이 내렸으면 하는 바람과 민주당이 세금을 내릴 것이라는 믿음에 의해 야기되었다는 식의 설명, 또는 영희가 아이스크림을 먹지 않는 이유가 살찌고 싶지 않은 소망과 아이스크림이 살찌게 한다는 믿음 때문이라는 식의 설명이바로 상식 심리학인 것이다.

이같은 상식 심리학은 우리의 일상 생활에 너무나 깊게 뿌리 박고 있어서, 우리는 그것에 어떤 이론적 전제가 있다는 사실을 의식하지 못하고 있다. 하지만 모든 이론적 설명이 그러하듯 인간의 사고와 행동에 관한 상식적 설명의 배후에도 암묵적이거나 이론적 전제가 관여하고 있다. 포더에 따를 것 같으면, 상식 심리학 배후에 깔려있는 전제는 크게 다음의 두 가지로 요약되어질 수 있다.¹⁾

첫째로 상식 심리학은 믿음이나 바람 같은 지향적 심리 상태들이 의미론적 속성을 가지고 있다고 전제하고 있다. 예를 들어 우리가 아이스크림을 먹지 않는 영희의 행동을 설명하려고 한다면, 영희의 심리 상태를 『불 같은 욕구』나 『확고 부동한 믿음』 식으로 명세화해서

* 이 글은 1995년 8월에 발표된 글쓴이의 석사논문 「사고와 행위에 대한 상식적 설명 방식의 옹호 가능성」의 일부이다.

1) Fodor(1994) pp.3-29

는 않된다. 상식 심리학적 설명이 요구하는 명세화는 『P라는 믿음』 『Q라는 욕구』²⁾ 식으로 명제적 종속절에 의거해 이루어지는 명세화이기 때문이다. 여기서 'that P'나 'that Q' 등이 나타나고 있는 바를 지향적 심리 상태의 내용(content)이라고 부른다. 상식 심리학은 바로 이런 명제적 내용이 해당 심리 상태의 의미론적 속성을 표현하고 있다고 본다.

자연 언어의 경우, 한 문장의 의미론적 속성은 진리 조건을 통해 표현된다. 그리고 그런 진리 조건의 만족 여부에 따라 한 문장이 참이나 거짓이나 하는 의미론적 평가(semantic evaluation)가 이루어진다. 이와 마찬가지로, 지향적 심리 상태의 의미론적 속성 역시 진리 조건과 유사한 만족 조건(satisfaction condition)을 통해서 표현되어지고 또 평가되어진다. 다시 말해, 『비가 올 것이라는 믿음』은 실제로 비가 올 때 오직 그 때에만(iff) 맞는 믿음이 되고, 또 『스파게티를 먹고 싶다는 바램』은 실제로 스파게티를 먹을 때 오직 그 때에만(iff) 만족된 바램이 되는 것이다. 그러므로 상식 심리학의 첫 번째 전제를 다른 식으로 표현한다면, 모든 지향적 심리 상태들에 대해 만족 조건을 부여할 수 있고, 또 그것을 통해 해당 심리 상태의 의미론적 속성을 평가할 수 있다는 말이 된다.

다음으로, 상식 심리학의 두 번째 전제는, 심리 상태와 심리 상태 사이에 또 심리 상태와 행동 사이에 의사 결정 이론(decision theory) 규칙이나 추론 규칙 같은 규칙들이 존재하고 있어서, 그런 규칙에 따라 한 심리 상태의 예화(tokening)가 다른 심리 상태나 행동의 예화를 유발하게 된다는 전제이다. 이점을 예를 들어 설명하면, 가령 철수가 점심 때 스파게티 먹기를 바라고, 또 이탈리아 식당에 가면 스파게티를 먹을 수 있다고 믿는다면, 철수는 다른 문제가 없는 한 점심 때 이탈리아 식당으로 가리라 예상된다. 또 영희가 워싱턴이 미국의 도시 가운데 하나라고 믿고 있고, 동시에 백악관이 워싱턴 안에 있다는 사실을 믿고 있다면, 그녀는 백악관이 미국의 어느 한 도시 안에 있다는 사실 역시도 당연히 믿고 있을 것으로 추정된다. 이

2) 이 글에서는 '『...』'을 심리 상태나 표상을 표시하는 기호법으로 사용하겠다.

같은 추정이나 설명이 가능한 이유는, 사람들이 $A \rightarrow B$ 를 믿고 있고 B가 되기를 소망할 경우 이번이 없는 한 A를 행하고자 한다는 것이나, 사람들이 $A \rightarrow B$ 를 믿고 동시에 A를 믿을 경우 이번이 없는 한 B역시 믿게 된다는 것 등이 상식 심리학에 의해 전제되고 있기 때문이다.

이상의 두 전제에 기초한 상식 심리학은 일상 생활에서 요긴하게 쓰이고 있지만 사고와 행동의 정확한 예측에는 성공하지 못하고 있다. 그래서 그동안 사고와 행위의 정확한 예측을 가능케 해 줄 보다 유물론적인 설명에 대한 추구가 계속되어 왔다. 최근 신경 과학 등에서 성취되는 성과들은 그러한 노력의 결과라 할 수 있다. 아무튼 이처럼 인간의 사고와 행동에 대해 유물론적인 설명이 가능하다는 사실이 차츰 가시화되면서, 최근에는 그간 전혀 의심 받아본 적이 없었던 상식 심리학의 이론으로서의 실재성이 의심받기 시작하였다. 특히 제거적 유물론이라고 불리는 일단의 연구 노선은 상식 심리학을 일종의 허구로 간주하여 그것을 과학의 무대에서 제거해야 한다고 주장하기까지 하고 있다. 이런 사태에 직면하여, 상식 심리학의 이론적 실재성을 옹호하고 그것을 유물론이라는 현대 사조의 대전제에 부합되도록 정교화하려는 시도가 바로 표상적 심리론(representational theory of mind)이다.

상식 심리학을 유물론의 틀 안에 포섭하기 위해서는 크게 보아 두 가지 작업이 요구된다. 우선 상식 심리학의 첫 번째 전제인 심리 상태의 의미론에 관한 전제가 유물론적으로도 모순없는 주장임을 보여야 한다. 그 다음으로 상식 심리학의 두 번째 전제인 구속 규칙에 대한 전제가 유물론적으로 구현 가능한 것임을 보여야 한다. 이 두 가지 작업을 위해 표상적 심리론이 동원하는 장치가 바로 표상(representation)이라는 이론적 존재자와 컴퓨터 인지 모형이다.

그 가운데 먼저 표상에 대해 살펴보자. 표상적 심리론이 표상이라는 이론적 존재자를 상정하는 이유는, 표상을 통해서 그간 애를 먹어왔던 지향적 심리 상태의 유물론적 환원 가능성을 확보하고자 함이다.

표상적 심리론 이전에도 상식 심리학을 유물론과 연결시키려는 시

도는 여러번 있어 왔다. 동일론이나 행동주의 같은 낙관적 환원주의자들의 시도가 좋은 예가 될 것이다. 그러나 그들은 믿음이나 바램 같은 지향적 심리 상태의 환원 문제에 부딪쳐 좌초하고 말았다. 『P라는 믿음』 『Q라는 바램』 같은 지향적 심리 상태들은 아픔이나 쾌감 같은 발생적 심리 상태와 중요한 차이점을 지니고 있다. 다름 아니라 종속절 P자리나 Q자리를 채울 문장들의 수효가 무한한 탓으로, 지향적 심리 상태는 원칙상 무한가지로 분류된다는 점이다. 그런데 한 이론을 다른 이론으로 환원하기 위해서는 두 이론의 용어들 간에 교량 법칙(bridge law)을 통한 상호 번역이 가능해야만 한다. 그러므로 상식 심리학의 핵심이 되는 지향적 심리 상태들의 수효가 무한정이라면 그것들을 번역할 교량 법칙은 확보할 수가 없고, 결과적으로 상식 심리학의 유물론적 환원은 불가능하게 되는 것이다.

이 문제를 해결하기 위해, 표상적 심리론은 『P라는 믿음』 『Q라는 바램』 등을 <『표상 P』와 맺고 있는 믿음의 관계>, 또는 <『표상 Q』와 맺고 있는 바램의 관계> 같은 식으로 분해하여 이해할 것을 제안한다.³⁾ 이렇게 할 경우, 『표상 P』나 『표상 Q』는 유한수의 정신 언어적 개념들이 회귀적으로 조합된 결과라고 가정할 수 있게 된다. 따라서 이제 내적 표상의 빌딩 블록이 되는 정신 언어적 개념들을 유물론적으로 해명하기만 하면 상식 심리학의 환원 가능성이 확보되는 것이다.

다음으로, 표상적 심리론이 동원하고 있는 장치는 컴퓨터라는 인지 모형이다. 표상의 존재론이 지향적 심리 상태로 하여금 명제적 내용을 지니면서도 유물론적으로 환원 가능하도록 하기 위해 도입되었다면, 컴퓨터 유비는 그런 지향적 심리 상태가 어떻게 서로 서로에 대해서 또 행동에 대해서 인과적 효력을 미치는지를 설명하기 위해 도입되고 있다. 앞서도 보았던 것처럼, 지향적 심리 상태들 간에 또 그것과 행동간에 잘 짜여진 규속 규칙이 존재한다는 것이 상식 심리학의 전제였다. 컴퓨터는 그러한 규속 규칙이 어떻게 작동하는지를 보여줄 좋은 모형을 제공하고 있기에, 표상적 심리론은 이 모형을 통해

3) Fodor(1981) p.244

상식 심리학을 유물론적으로 정당화하고자 한다.

지금까지 본 것처럼 표상적 심리론은 표상의 존재론과 컴퓨터 유비라는 두 가지 개념 장치를 통해 그 특징을 규정할 수 있다. 그 가운데 이 글에서는 주로 표상에 대한 논의를 중심으로 표상적 심리론의 성격을 정리하여 보고자 한다. 좀더 자세히 말하자면, 2 장에서는 우선 표상적 심리론의 대표자 격인 포더의 이론을 중심으로, 표상적 심리론이 과연 어떤 주장인지 대략적으로 스케치하여 볼 것이다. 그런 다음 3 장에서는 그와 같은 표상적 심리론이 상식 심리학의 첫 번째 전제인 지향적 심리 상태의 의미론에 대해서 어떤 식의 해명을 제공하고 있는지 요약하여 볼 것이다. 4 장에서는 표상적 심리론의 강력한 반대자인 퍼트남의 주장을 통해, 표상의 의미론적 해명 시도에 대한 비판을 살펴볼 것이다. 끝으로 5 장에서는 논의를 종합하여 표상적 심리론의 전망에 대해 평가를 내려 볼 것이다.

2. 표상적 심리론의 이론적 윤곽

2.1. 표상적 심리론의 특징과 배경

표상적 심리론의 골자는 정신 활동을 표상에 대한 조작으로 이해하자는 데 있다. 여기서 표상은 하나의 물리적 실재로 간주되며, 인과적 관계에 의해 두뇌 밖의 대상들과 관련을 맺고, 그 덕분에 두뇌 밖의 대상들을 대표한다고 가정된다.

그러나 이와 같은 발상은 철학사에서 새로운 것이 아니다. 기억을 이미지를 다루는 것이라고 보았던 아리스토텔레스나 정신 활동을 오성의 관념 조작이라고 보았던 로크, 또는 정신 현상을 지향적 내재자들과 맺는 관계로 설명하려 하였던 브렌타노 등의 예처럼 철학사 곳곳에서 이와 유사한 생각들을 찾아볼 수 있다.⁴⁾ 그렇다면 현대의 표상적 심리론은 어떤 점에서 과거의 발상들과 구분될 수 있을까.

현대의 표상적 심리론은 설명의 과정에서 보다 자연화된 방식을

4) 정성호(1994)

따르고자 한다는 점에서 과거의 생각들과 구분된다. 그런 자연화된 설명을 위해 동원되는 것이 컴퓨터라는 구체적 인지 모형이다. 한 가지 예를 들어 이 점을 분명히 하여보자.

관념에 대해 근대 경험론자들이 지녔던 전형적인 생각은, 그것이 마음 밖의 사물과 유사성을 가지고 있기 때문에 그 사물을 대표할 수 있다는 것이었다. 이를테면 붉음의 관념은 현실의 붉음과 어떤 유사성을 가지고 있으므로 붉음의 관념일 수 있고, 사과와 관념은 현실의 사과와 어떤 유사성이 있으므로 사과의 관념일 수 있다고 보았던 것이다. 경험론자들은 이같은 생각을 바탕에 깔고, 오성이 이 관념과 저 관념을 비교하고 그 유사성에 따라 둘을 연결한다는 식으로 이론을 전개하였다.

하지만 외적 사물과의 유사성 관계를 통해 관념의 속성을 설명하는 것은, 유사성 관계가 지닌 본질적 애매함 때문에 아무런 설명의 역할도 하지 못한다.” 그 이유는 모든 사물은 다른 사물과 어떤 점에서는 유사하다고 볼 수 있기 때문이다. 예를 들어, 앞집 강아지 복돌이는 형제 강아지 순돌이 하고만 유사한 것이 아니라, 뒷집 고양이 야옹이 하고도 포유류라는 점에서는 유사하고, 심지어는 하늘에 떠 있는 천왕성파도 부피를 가지고 있다는 점에서는 유사할 수 있다. 결국 A가 B와 유사하기 때문에 A가 B를 대표할 수 있다고 주장하는 것은 A와 B의 유사성을 판별하는 작업 역시 또 하나의 정신 작용이라는 점을 간과하는 것이 된다. 그런 숨겨진 정신 활동의 도움 없이도 하나의 관념과 다른 관념이 유사성에 따라 비교된다든가 연결된다면, 그것은 일종의 마술일 수밖에 없다. 그러므로 경험론자들의 탐구는 인간 정신 활동의 문제를 설명한 것이 아니라, 오히려 또다른 종류의 숨겨진 정신 활동을 가정함으로써 문제를 한걸음 퇴행시킨 것에 불과하고 할 수 있다.

이에 반해, 현대의 표상적 심리론은 표상을 물리적/신경학적 실재로 간주하고, 표상의 의미를 그것이 가지는 인과적/기능적 속성에 귀인시킴으로써 설명의 무한 퇴행을 방지한다. 여기서 표상의 인과적/

5) Sterelny(1990) pp.112-114

기능적 속성을 명료화하기 위해 이용되는 것이 컴퓨터라는 인지 모형이다.

컴퓨터의 기호 조작 과정은 유효 절차, 즉 알고리즘(algorithm)에 의해 분명하게 규정되어 진다. 그러므로 컴퓨터 유비를 통해 정신 활동을 설명하게 되면, 앞서 경험론에서 본 것 같은 마술적 기능을 상정하는 일은 피할 수 있게 되는 것이다. 이와 같은 설명의 자연주의적 성격이 현대의 표상적 심리론을 과거의 사유들과 구별되게 만드는 점이다.

2.2. 표상의 존재론을 수용하는 경험적 이유

하지만 굳이 옴의 면도날을 이야기하지 않더라도 이론의 존재론이 단순할수록 좋다는 점은 자명하다. 하나의 존재론을 수용하게 되면, 그에 따라 그런 존재자들이 어떻게 개별화(individuation)되는지, 또 그런 존재자의 판별 기준(criterion)은 무엇인지와 같은 악명 높은 문제들을 해결해야 하기 때문이다. 이는 표상의 경우에도 마찬가지이다. 앞 절에서 지향적 심리 상태의 유물론적 환원을 위해 표상의 존재론이 요구된다고 하였지만, 만일 표상의 존재론을 수용할 필요성이 그 정도에서 그친다면, 차라리 불편을 감수하면서라도 존재론적 검약의 원칙을 따르는 것이 나을지 모른다.

그러나 표상의 존재론을 받아들여야 하는 이유는 그와 같은 이론적 필요성에 그치지 않는다. 거기에는 보다 현실적이고 적극적인 이유들이 있다. 그 가운데 무엇보다 중요한 것은 내적 표상의 존재를 상정하지 않고서는 인간을 포함한 자연계 유기체들의 지성적 행동(intelligent behavior)을 설명할 수 없다는 사실이다.

유기체의 지성적인 행동은 물에 떠다니는 나무 토막의 움직임과는 근본적으로 차이가 있다. 뿐만 아니라 유전적으로 프로그램 된 본능적 행동과도 구분된다. 스티어르니는 유기체의 지성적 행동을 본능적 행동이나 물체의 운동과 구별해주는 특징으로 두 가지를 들고 있는데, 바로 목표 지향성과 환경에 대한 민감성이다.⁶⁾

6) 앞의 책 pp.19-20

그 가운데 목표 지향성이란, 끊임 없이 유동하는 환경 가운데서도 적절한 피드백을 통해 고정된 목표 상태에 도달한다는 지성적 행동의 특징을 말하는 것이다. 반면 환경 민감성이란, 주어진 환경 조건을 세밀하게 변별하여 그것에 맞는 적절한 반응을 산출하는 특징을 말한다. 예를 들어 냇물을 따라 흐르는 나무 토막의 경우, 냇물의 흐름이 변화하면 그것의 흘러가는 방향도 달라지게 된다. 그러나 미로 상자에 들어있는 배고픈 생쥐는 미로 상자의 환경에 변화가 생겨도 굶어 죽지 않는다. 환경 변화에 따라 민감하게 행동을 수정해가며, 마침내는 음식을 먹는다는 목표를 달성하고야 말기 때문이다.

이처럼 지성적 행동은 환경 변화에 민감하면서도 목표를 달성하기 위한 유연한 융통성(flexibility)을 지니고 있다. 그런데 이런 행동이 가능하려면, 유기체는 목표 상태와 환경 조건을 표상하면서 그 갯을 효율적으로 조정할 수 있어야만 한다. 따라서 유기체가 지성적 행동을 할 수 있다는 사실은 결국 유기체 내부에 어떤 표상 체계가 존재함을 증명하게 되는 것이다.⁷⁾

인간의 경우에는 사태가 보다 분명하다. 인간들이 보여주는 다양한 종류의 의사 결정 행동과 문제 풀이 행동 또 학습 등은 내적 표상의 도움이 없이는 불가능한 듯 보이기 때문이다. 인간에게 내적 표상 체계가 있음을 주장하며 포더가 드는 예는 체스 경기이다. 포더에 따르면, 체스 경기를 할 때 복잡한 수를 마음 속으로 이리 저리 두어볼 수 있다는 사실은 우리에게 체스 경기를 표현할 수 있는 표상 체계가 존재함을 증명하는 것이다. 더 나아가 포더는 학습을 위해서도 내적 표상이 요구됨을 지적한다.⁸⁾ 학습에는 마음 속으로 가설을 형성하고 수정하는 작업이 요구되는데, 내적 표상이 없다면 그런 가설을 표현할 수 없기 때문이다.

이상과 같은 일련의 이유들 때문에, 표상적 심리론자들은 존재론적 검약의 원칙을 어겨가면서까지 표상의 존재론을 받아들이고 있는 것이다.

7) 앞의 책 p.23

8) Fodor(1981) p.401

2.3. 사고 언어 가설

표상의 존재론을 수용하고 나면, 이제 풀어야 문제는 그같은 표상이 과연 어떤 구조를 가지고 있는가 하는 것이다. 이 문제에 대한 표상적 심리론의 표준적 견해는 표상의 구조가 마치 자연 언어 같다는 것이다. 이른바 사고 언어 가설(language of thought hypothesis)로 불려지는 이같은 생각은 포더에 의해 최초로 제안되었으며, 오늘날에는 대부분의 표상적 심리론자들에 의해 어떤 형태로든 수용되고 있다.

그렇다면 표상이 자연 언어와 같다는 것은 과연 무슨 뜻일까. 자연 언어는 일반적으로 다음의 특징들을 가지고 있다. 그러므로 표상이 언어와 유사하다는 주장은 결국 자연 언어의 이런 특징들이 표상에서도 똑같이 발견된다는 주장이 된다.

우선 구문론적으로 보자면, 자연 언어는 유한한 수효의 단어들과 그것들을 결합하는 회귀적(recursive) 문법으로 이루어져 있다. 자연 언어는 이런 회귀적 문법의 도움으로 무한 수의 문장들을 만들어낼 수 있는 생산성을 가진다. 또 자연 언어 문법의 구조적 특징 때문에, 자연 언어 문장들간에는 이른바 체계적 관계가 성립한다. 이를테면 “여우가 토끼를 잡아먹었다”는 자연 언어 문장이 참이면 “토끼는 여우에게 잡아먹혔다”는 문장도 참이 되는 것이다. 이런 점들이 자연 언어에서 일반적으로 찾아볼 수 있는 구문론적 특징들이다.

다음으로 의미론적 특징들을 살펴보면, 자연 언어에는 그 의미 수준에 있어 차이가 나는, 단어와 문장이라는 서로 다른 종류의 의미 단위들이 있다. 이 가운데 단어의 의미는 지시 관계나 동의성 등에 의해 자체적으로 설명되는 반면, 문장의 의미는 그런 단어들의 의미를 통하여 환원적으로 설명된다고 여겨진다. 이른바 빌딩 블록(building block) 의미 이론이라고도 불리는 이러한 생각은, 문제가 없는 것은 아니지만, 그럼에도 자연 언어에 대한 가장 일반적이고 표준적인 생각으로 받아들여지고 있다.⁹⁾

표상적 심리론은 내적 표상 체계가 이상과 같은 자연 언어의 특징

9) 빌딩 블록 의미론에 대한 비판은 Davidson(1984) p.220 참조

들을 똑같이 갖추고 있다고 가정한다. 먼저 구문론적으로 보자면, 표상에도 자연 언어의 단어에 해당되는 개념들이 있고, 또 문장에 해당하는 명제 내용들이 있다고 가정된다. 그리고 자연 언어와 마찬가지로, 개념들의 수는 제한되어 있지만, 그것들을 결합하는 회귀적 규칙이 존재하는 덕분에 우리가 생각할 수 있는 명제 내용의 수는 무한하다고 가정된다. 다시 말해, 우리의 생각 역시도 생산적이라는 것이다. 뿐만 아니라, 표상의 경우에도 구문론적 규칙 덕분에 명제 내용들간의 체계적 관계가 성립한다고 가정된다. 예컨데 만일 누군가가 『여우가 토끼를 잡아먹었다』는 생각을 가지고 있다면, 그는 당연히 『토끼는 여우에게 잡아먹혔다』는 생각도 가진다는 것이다.¹⁰⁾

의미론의 문제는 자연 언어의 경우나 표상의 경우 모두 마찬가지로 아직까지 설명되지 않은 부분이 많다. 그러나 개략적으로 말한다면, 앞서 언급한 빌딩 블록 의미 이론이 내적 표상에 대해서 그래도 가장 일반적으로 받아들여지고 있는 견해라 할 수 있다. 포더, 드레츠키, 스티어르니 등 다수의 표상적 심리론자들은, 사고 언어의 개념들이 어떤 식으로든 독자적으로 의미를 가지게 되며, 명제 내용의 의미는 그런 개념들의 의미를 통해 환원적으로 설명될 수 있다는 견해를 지지하고 있다.

이렇게 보면, 표상적 심리론이 사고 언어 가설을 채택함으로 부딪치게 되는 난제들이 어떤 것일지 대략 짐작이 된다. 자연 언어 의미론에서 골치거리인 의미의 전체성 문제나 지시의 유동성 문제가 고스란히 표상적 심리론의 문제로 전이되는 것이다. 그렇다면 이처럼 풀기 어려운 문제까지 끌어 안으면서 내적 표상 체계가 굳이 언어적 구조를 지닌다고 가정해야 할 필요성은 어디있는 것일까. 표상적 심리론자들은 이에 대해 몇가지 이유를 제시하고 있다. 다음 절에서는 그런 이유들에 대해서 살펴보기로 한다.

2.4. 사고 언어 가설을 지지하는 일상 언어적 근거

표상적 심리론은 상식 심리학을 정교화하려는 시도이다. 여기서 상

10) 최훈(1994) 참조

식 심리학이란 우리의 일상 언어적 표현들 가운데 녹아 있는 사고와 행위에 대한 상식적 견해를 말하는 것이다. 때문에 포더는 상식 심리학의 정교화가 왜 언어적 형태의 표상을 통해 이루어져야 하는지에 대해 몇가지 일상 언어적 이유들을 제시하고 있다. 이런 포더의 논증을 좀더 자세히 살펴보도록 하자.¹¹⁾

포더는 일상 언어적 직관에 부합하는 적절한 명제 태도 이론이라면 아래에서 보게 될 네가지 조건들을 반드시 만족시켜야 한다고 본다. 그 네가지 조건을 포더는 이항성 조건, 벤들러(Vendler) 조건, 프레게 조건, 아리스토텔레스 조건이라 부른다.

이 가운데 이항성 조건이란, ‘믿는다’ ‘바란다’ 등의 명제 태도 동사들이 대부분의 자연 언어에서 이항 관계 술어로 취급됨에 주목해야 한다는 조건이다. 영어의 경우를 예로 들면, “John believes that it’s raining”이라는 문장은 그 형식상 “John has a dog”이라는 이항 술어 문장과 동일하다. ‘believes’라는 동사는 ‘has’라는 동사와 마찬가지로 이항 술어의 하나로 간주되고 있는 것이다. 때문에 포더는 만일 명제 태도 문장을 일항 술어로 분석하려 한다면, “John believes that it’s raining”라는 문장을 “John is purple” 처럼 취급하게 되는 어리석은 결과를 낳는다고 경고하고 있다.

그러나 그럼에도 불구하고 일부 이론가들은 앞서와 같은 명제 태도 문장을 ‘believes-that-it’s-raining’이라는 일항 술어를 통해 분석하려고 한다. 이런 분석은 대개 ‘that’ 이하 종속절의 내포적 불투명성을 분석에 반영시키고자 하는 동기에서 비롯된다. 그러나 포더가 보기에, 그와 같은 일항적 분석은 내포적 불투명성을 분석에 반영하는 댓가로 또다른 중요한 사실을 반영하는 데 실패하고 만다. 다시 말해, 그런 분석은 “John believes that it’s raining”라는 문장과 “it’s raining”이라는 문장의 관계를 반영할 수 없게 되는 것이다.

여기서 “it’s raining”이라는 문장은 존의 믿음 내용을 만족 조건의 형태로 표현하고 있다. 그러므로 이런 대응 문장이 달라짐에 따라 표현하고자 하는 존의 믿음 내용도 달라지게 된다. 중요한 사실은 이처

11) 이하의 논의는 Fodor(1981) pp.239-274 참조

럼 믿음 내용을 표현하기 위해 채택될 수 있는 대응 문장의 수효가, 자연 언어의 일반 문장이 그러하듯, 무한하다는 점이다. 이런 사실은 존의 믿음 내용 자체도 이처럼 무한가지일 수 있어야 함을 보여주는 것이다.

한편 두 번째 조건인 벤들러 조건은, 많은 자연 언어에서 명제 태도 동사들과 언표 동사들이 보여주는 평행성을 해명해야 한다는 조건이다. 명제 태도 동사와 언표 동사 사이의 평행성이란, 이를테면 'that P'가 철수에 의해서 말해지거나, 선언되어지거나, 질문되어지는 방식과 마찬가지로, 철수에 의해 믿어지거나, 바라지거나, 기대되어진다는 사실을 말하는 것이다.

이런 평행성이 주목받는 것은 연접이나 선접으로 결합된 내용들을 믿거나 말하는 경우 때문이다. 즉 철수가 'that P & Q'를 말했다면 이것은 철수가 'that P'를 말했고 또(&) 철수가 'that Q'를 말했다는 사실을 함축한다. 이와 마찬가지로 철수가 'that P & Q'를 믿는다는 사실 역시 철수가 'that P'를 믿고 또(&) 'that Q'를 믿는다는 사실을 함축한다. 결국 이런 평행성은 우리가 믿고, 바라고, 두려워하고, 기대하는 것들이 우리가 말하고, 약속하고, 질문하고, 선언하는 것들과 마찬가지로 그 구조에 있어 체계적이라는 사실을 보여주는 것이다.

포더가 제시하는 세 번째 조건은 프레게 조건이다. 이 조건은 명제 태도 문장의 종속절에 나타나는 지시적 불투명성이 해명되어야 한다는 조건이다. 일반적으로 명제 태도 문장의 종속절에 출현하는 술어들은 동일 외연의 다른 술어들로 대체될 수 없다. 그같은 대체가 명제 태도 문장의 진리치를 종종 변경하여 놓기 때문이다. 그래서 예컨대 “철수가 새벽별이 금성이라고 믿는다”는 문장과 “철수가 저녁별이 금성이라고 믿는다”는 문장은 구별되어야 하는 것이다. 이같은 사실은 철수의 내적 표상 체계 안에서 새벽별에 대한 생각과 저녁별에 대한 생각이 분명히 구별될 수 있어야 함을 보여주는 것이다.

이에 반해 아리스토텔레스 조건이란, 명제 태도 종속절의 그와 같은 불투명성에도 불구하고 어떻게 해서 명제적 내용을 통한 비형식적 추론이 가능한지 밝혀져야 한다는 조건이다. 아리스토텔레스의 실천 삼단 논법에 의할 것 같으면, 사람들은 Q를 바라고 $P \rightarrow Q$ 를 믿을 때,

이변이 없는 한 P를 행한다. 이것은 불투명한 맥락의 명제적 내용을 가지고도 개략적인 추론은 가능하다는 사실을 보여주는 것이다.

포더는 우리의 내적 표상 체계가 이상을 조건을 모두 만족시키려면 자연 언어와 상당히 유사해야만 한다고 주장한다. 즉 이항성 조건에서 지적된 것처럼 생산적이며, 밴들러 조건이 말하는 것처럼 체계적이고, 프레게 조건이 요구하는 것처럼 개념들을 풍부하게 구분해 줄 수 있으며, 아리스토텔레스 조건이 요구하는 것처럼 구분론적 특질까지 갖춘 그러한 것이 우리의 내적 표상이라면, 그것은 이미 자연 언어의 주요 특징들을 모두 갖추고 있는 셈이기 때문이다.

2.5. 사고 언어 가설과 최선의 설명

표상적 심리론자들이 사고 언어를 가정하는 데에는 그것이 일상 언어적 직관과 부합한다는 이유 말고도 또다른 이유가 있다. 그 가운데 포더는 특히 사고 언어 가설이 인간의 언어 습득과 의사 소통의 핵심적 국면을 설명하기 위한 일종의 「최선의 설명」으로서 요구된다는 점을 강조한다.

물론 포더의 이런 주장은 사고 언어가 생득적이라는 그의 또다른 주장과 맞물려 있기 때문에, 사고 언어의 생득성을 인정하지 않으려는 다른 표상적 심리론자들의 견해와는 구별된다.¹²⁾ 그러나 그럼에도 불구하고, 이같은 포더의 주장은 현재까지 제시된 사고 언어 가설의 지지 이유 가운데 가장 강력한 경험적 기반을 가지고 있기 때문에 음미의 가치가 있다. 그러므로 이번 절에서는 사고 언어 가설이 언어 습득과 의사 소통을 설명하기 위한 최선의 설명이라는 포더의 주장에 대해 살펴보고, 더불어서 이 주장과 불가분의 관련을 맺고 있는 사고 언어의 생득성에 대한 포더의 견해를 개괄하여 보겠다.

포더 자신도 인정하는 것처럼, 사고 언어 가설의 기본 발상은 상당 부분 촘스키의 생득 언어학에서 비롯되고 있다. 잘 알려진대로, 촘스키 생득 언어학의 목표는 언어 습득에 있어서 내적 정신 구조의 능동

12) 사고 언어의 생득성을 인정하지 않는 표상주의자들은 사고 언어가 자연언어를 습득함으로써 부차적으로 형성된다고 가정한다.

적 역할을 밝히는 것이다. 그런 점에서, 촘스키의 생득 언어학은 스키너 등으로 대표되는 행동주의 언어 학습 이론과 대조를 이룬다고 할 수 있다.¹³⁾

스키너 등의 행동주의자들은, 어린이의 언어 학습이 지적(ostention)을 통한 자극 제시와 그에 대한 어린이의 언어적 반응을 강화함으로써 이루어진다고 설명하였다. 하지만 인간이 말할 수 있는 문장은 원리상 무한하기 때문에 자극/반응의 연합으로는 결코 인간이 발화 가능한 모든 문장의 학습을 마칠 수 없다는 점을 들어, 촘스키는 그런 행동주의식 언어 학습 이론에 의문을 제기하였다. 그리고 행동주의 이론을 대신해서, 언어를 배우는 것이 언어의 구조, 곧 문법을 익히는 것이라는 새로운 주장을 펼쳤다.

그러나 언어 학습을 문법의 습득이라고 보게 되면, 새로운 문제가 발생한다. 그것은 어린이가 언어를 배우는 과정에서 제시받는 자극이 자연 언어의 복잡한 문법을 형성하기에 충분치 못하다는 사실이다. 촘스키는 이 문제를 해결하는 최선의 방법이 우리의 정신 내부에 일종의 보편 문법이 생득적으로 주어져 있다고 가정하는 길뿐이라고 주장하였다. 다시 말해, 어린이가 것처럼 빈곤한 자극으로부터 자연 언어의 복잡한 문법을 형성해낸다는 사실을 설명하려면, 불가피하게 생득적 정신 구조를 가정해야 한다는 것이다.¹⁴⁾

포더도 사고 언어 가설을 입증하기 위해 촘스키와 비슷한 논리를 편다.¹⁵⁾ 그러나 포더가 가정하고 있는 사고 언어는 촘스키가 가정하고 있는 생득적 문법보다 더 포괄적인 개념이다. 촘스키의 경우에 생득적으로 주어진 것은 표층 언어의 생성과 변환에 관련된 구문론적 조작 기제에 불과한 반면, 포더의 사고 언어 가설에서는 구문론적 조작 기제 외에도 개념 내지 개념들까지 생득적으로 주어져야 하기 때문이다.

포더는 이같이 대담한 가설을 입증하기 위해, 대상 지각과 관련된 일종의 자극 빈곤 논변(Poverty of stimulation argument)을 펼친

13) Chomsky(1975) 참조

14) Ramsy & Stich(1990) 참조

15) 이하 논의는 Fodor(1975) 및 Fodor(1981)의 10장 참조

다. 포더의 자극 빈곤 논변도 촘스키의 자극 빈곤 논변과 마찬가지로 주어진 원초적 자극이 그 결과로 나타나는 반응의 풍부함을 산출하는데 충분하지 못하다는 점을 지적하고, 그 부족분이 정신의 선제적 구조에 의해서 채워져야 함을 주장하는 식으로 전개된다. 다만 포더의 논변에서는 촘스키의 경우와 달리 주어지는 자극이 언어적 자극이 아니라 말초적 감각들이고, 그에 따라 산출되는 결과도 문법에 대한 지식이 아니라 개념에 대한 지식이 된다. 이런 포더의 논변을 좀더 자세히 살펴보자.

포더에 의하면 지각 과정이란 감각과 표상의 매개 역할을 하는 중간 단계이다. 인간의 감각 기관에 주어진 무질서한 자극들은 이 단계를 거치면서 표상의 기초 단위인 개념들을 예화시키게 된다. 그런데 이런 지각 과정에서 포더가 주목하는 사실은, 우리에게 주어지는 감각 자극들이 대부분의 경우 개념들을 구분해내기에는 충분치 못하다는 사실이다.

이같은 불충분성을 보여주는 좋은 예가 콰인이 말한 원초적 번역 상황일 것이다. 그 예에서 '가바가이'의 정확한 번역이 어려운 이유는, 주어진 감각 자극들을 대상으로 개별화하기 위해 경험적 자료 외에도 별도의 추가 정보가 요구되기 때문이다.

그런데 보다 심각한 문제는 어린이가 말을 배우는 상황도 사실 원초적 번역 상황과 다를 바가 없다는 점이다. 콰인은 모국어를 배우는 경우에는 배경 지식이 공유되기 때문에 문제될 것이 없다고 주장한 바 있지만, 이런 주장은 데이비슨이 지적하는 바대로 일종의 논점 선취에 불과하다.¹⁶⁾ 논의의 관건이 바로 그런 배경 지식이 어린이에게 어떻게 습득되느냐는 문제이기 때문이다.

감각 자극과 언어적 개념 사이에 존재하는 이와 같은 미달 결정성의 문제를 표상적 심리론에서는 「으로서 문제(qua problem)」라고 한다.¹⁷⁾ 이 「으로서 문제」는 두 가지 종류로 좀더 세분해 볼 수 있다. 하나는 이른바 「넓이 문제(breath ambiguity)」로서, 이를테면 앞에

16) Davidson(1984) pp.227-241

17) Sterelny(1990) p.113

있는 토끼를 보고 『토끼』라는 개념을 생각해야 하는지, 아니면 『포유류』라는 개념을 생각해야 하는지, 아니면 그도 저도 아니고 『물체』라는 보다 포괄적인 개념을 생각해야 하는지 결정할 수 없다고 하는 개념적 위계의 미달 결정성(underdetermination) 문제이다. 또 하나는 「깊이 문제(depth ambiguity)」로서, 앞에 있는 토끼를 보고 『토끼의 형태』를 생각해야 하는지, 아니면 『토끼』 자체를 생각해야 하는지, 그것도 아니고 『토끼 다음』을 생각해야 하는지 결정할 수 없다고 하는 추상화 정도의 미달 결정성 문제이다.

포더는 이런 미달 결정성이 존재함에도 불구하고 인간이 언어를 배우고 또 문제 없이 의사 소통한다는 사실에 주목하자고 한다. 그리고 이같은 현실을 설명하기 위해서는, 진화론적 기제 등에 의하여 지각 과정에 생득적이고 비교적 보편적인 개념 분류 체계가 자리잡게 되었다고 가정할 필요가 있음을 주장한다.¹⁸⁾ 이것은 달리 말해, 배경적 개념틀이 생득적 기원을 갖는다는 주장이 된다.

포더는 이런 주장을 뒷받침하기 위해 발달 심리학의 최근 연구 결과 한 가지를 인용한다.¹⁹⁾ 그것은 어린 아이들이 말을 배울 때 중간 크기의 범주들을 먼저 배운다는 관찰 결과이다. 즉 『가구』 같은 거시 개념이나 『안락 의자』라는 미시 개념을 배우기 전에, 어린이들은 먼저 『의자』라는 중간 크기의 개념을 배운다는 것이다. 포더는 이같은 관찰 결과를 인간에게 원초적 개념들의 생득적 체계, 즉 생득적 사고 언어가 존재한다는 증거의 하나로 간주한다.

이 외에도 포더는 생득적 사고 언어 가설을 뒷받침하기 위해 보다 논리적인 근거를 한 가지 더 든다. 바로 개념의 학습이다. 가령 푸르거나 삼각형 무늬를 가진 카드들을 ‘파카스롭다’²⁰⁾라는 새로운 어휘로 정의한다고 했을 때, 포더는 우리가 이런 신종 개념을 어떻게 학습하는지 생각해보자고 한다. 자연적 학습 상황과 유사한 실험실의 학습 과정은 다음과 같을 것이다. 우선 피험자는 카드를 하나씩 보고 그것

18) Fodor(1981) pp.353-365

19) 앞의 책 p.410

20) 앞의 책 p.353 참조. 본문에서는 FLURGR = ‘... 이 초록색이거나 사각형이다.’

이 파악스러운지 아닌지를 대답하게 된다. 그리고 선생인 실험자는 피험자가 정답을 말하였는지 아닌지를 알려준다. 피험자는 이런 단순한 절차를 거치면서 궁극적으로는 '파악스럽다'는 말이 무슨 말인지 알게 된다. 그런데 중요한 점은, 이런 깨달음의 과정에서 피험자에게 일종의 가설이 요구된다는 점이다. 즉 피험자는 처음에, 이를테면 『'파악스럽다'가 파란 동그라미를 말하는 것 아닐까』 하는 가설을 세우고 시험해보지만 곧이어 틀렸음을 확인하고, 그래서 또 새로운 가설을 만들어 시험해보고, 하는 일련의 과정 끝에 하나의 개념을 학습한다는 것이다. 한데 이처럼 가설을 형성하고 그것을 수정하는 것이 가능하기 위해서는, 그에 앞서 이미 가설을 표현할 수 있게 해주는 내적 표상 체계가 선재해 있어야 한다. 때문에 포더는, 이같은 논증이 생득적 사고 언어를 가정함 없이 자연 언어를 배울 수 있다는 주장에 대한 귀류법적 반박이 된다고 본다.

한편 포더는 개념들을 산출하는 지각 과정이 진정으로 생득적이고 보편적인 기제이기 위해서는, 그 작용이 배경 지식의 영향으로부터 격리된 일종의 모듈 체계(modular system)여야 한다고 주장한다.

여기서 모듈(module)이란, 포더가 상정하는 인지 처리의 기본 단위로써, 두 가지 특징을 통해 그 성격이 규정된다. 하나는 그 처리 과정에서 활용 가능한 지식의 범위가 지극히 국한되어 있다는 점이고,²¹⁾ 다른 하나는 그럼에도 불구하고 단순한 조건 반사와는 달리 제한된 범위에서나마 일종의 추론을 수행한다는 점이다.²²⁾ 즉 모듈은 의식적

-
- 21) 모듈은 정보처리에 있어 상호 독립적이다. 이점을 보여주는 심리학적 사례 한 가지를 보자. 만약 왼쪽 귀에 음악을 들려주면서 오른쪽 귀에 문장을 들려줄 경우 나중에 음악과 문장을 모두 회상해 낼 수 있다. 반면 두쪽 귀 모두에 문장을 들려줄 경우, 그 가운데 어느 한쪽은 나중에 회상되지 못한다. 앞의 경우에는 음악 정보를 처리하는 모듈과 문장 정보를 처리하는 모듈이 서로 다른 모듈이라서 상호 간섭이 일어나지 않지만, 뒤의 경우에는 단일한 모듈 혼자서 두 가지 정보를 처리해야 하므로 어느 한쪽은 포기할 수밖에 없게 되는 탓이다.
- 22) 모듈의 추론 능력을 보여주는 사례 한 가지. 홀러 쓴 13과 영어 필기체 B는 그 모양만으로는 구분이 안된다. 하지만 우리는 책을 보면서 아무런 의식적 노력 없이도 13과 B를 구분해 낸다. 이런 구분의 배후에는 앞뒤의 맥락을 통해 그 글자가 무엇일지 추론해 내는 무의식적 과정이 자리하고 있다. 바로 이런 무의식적 추론을 담당하는 기제가 모듈이다.

추론처럼 배경 지식으로부터 영향을 받지 않으면서도 단순한 조건 반사는 아닌, 일종의 준사고(semi-cognition)를 수행하는 기제라 할 수 있다. 그리고 모듈 체계는 이렇게 준사고를 수행하는 모듈들이 집합적으로 모여서 중층적 위계를 구성하고 있는 체계로 보면 된다.²³⁾

모듈 체계의 처리 과정은 강제적이면서 자동적인 상승식 처리 과정(mandatory automatic bottom-up process)이다. 이러한 모듈 체계의 처리 방식을 잘 보여주는 예가 데이비드 마(Marr)의 3차원 시각 과정에 대한 이론이다.²⁴⁾ 마는 시각 시스템이 망막에 맺힌 2차원 화면으로부터 어떻게 3차원적 대상을 지각하게 되는지를 몇 단계의 위계로 나누어 설명하는 모형을 제안한 바 있다. 이 모형에서 시각 정보 처리의 각 위계에는 윤곽선 파악 모듈, 꼭지점 파악 모듈, 명암 파악 모듈 등과 같은 업무 전담 체계가 자리잡고 있다. 그리고 각 모듈은 자신의 업무 처리에 필요한 자료를 보다 낮은 위계의 하위 모듈들로부터 공급 받지만, 같은 위계의 다른 모듈들이나 보다 상급 위계 모듈들의 처리 결과로부터는 전혀 영향을 받지 않는다.

포더는 지각 체계가 그와 같은 모듈 체계라고 보고 있다. 때문에 일단 동일한 말초 감각 자료들이 주어지게 되면, 그것으로부터 예화되는 개념과 믿음 내용은 배경 지식의 영향을 받지 않고 강제적/자동적으로 결정된다고 주장하고 있다.²⁵⁾ 이같은 지각 과정의 강제성/자동성을 보여주는 예로 포더가 즐겨 인용하는 것이 물러/라이어 착시 현상이다.

아다시피 물러/라이어 착시란 화살표의 꺾임 방향이 달라짐에 따라 동일한 길이의 두 선분이 실제보다 크거나 작게 보이는 현상을 말한다. 이런 현상과 관련하여 포더가 주목하는 점은, 관찰자가 아무리 물러/라이어 화살표들이 길이가 같다는 배경 지식을 갖는다 하더라도, 그런 배경 지식의 습득이 화살표의 길이 지각에 전혀 영향을 주지 못한다는 사실이다. 포더는 이같은 물러/라이어 착시 현상을 들어, 지각 과정, 곧 말초 감각에 의해 표상이 산출되는 과정이 배경 지식의 영

23) Fodor(1992) pp.195-230

24) Marr의 이론에 대한 설명은 이인식 저 「사람과 컴퓨터(까치)」 3장 참조

25) 앞의 책 pp.231-263

향으로부터 배제되어 있다는 주장을 한다.

이상에서 살펴본 바를 정리하자면 다음과 같다. 말초 감각 기관에 주어지는 자극들은 무질서하고 종종 불충분하다. 따라서 그것으로부터 개념들을 예화하기 위해서는 어느 정도의 능동적 추론이 그 부족분을 채워주어야 한다. 그런데 이 능동적 추론 과정은 진화론적 기제 등에 의해서 생득적으로 고정되어 있다. 뿐만 아니라 후천적으로 습득되는 배경 지식의 영향을 받지 않는다. 결과적으로 하나의 개념이 감각 자료들과 연결되는 방식은 생득적이고 보편적이다.

이와 같은 주장이 개념의 생득설이라는 조금은 플라톤주의적 명칭으로 불리는 포더 견해의 전모이다. 글쓴이는 포더의 이같은 견해가 그 명칭에서 암시되는 것과는 달리 상당히 자연주의적이고 또 경험적 설득력도 갖춘 견해라고 생각한다. 다만 글쓴이가 지적하고 싶은 것은 생득적 지각 과정에 배경 지식의 영향력이 전혀 미치지 못한다는 포더의 주장이 너무 강하지 않은가 하는 점이다. 개념의 가소성(plasticity)을 주장하는 처치랜드 역시 바로 그런 이유 때문에 포더의 이론을 강하게 비판하고 있다.²⁶⁾

3 표상의 의미론 해명

앞 장에서는 내적 표상 체계가 언어적 구조를 가져야 한다는 표상적 심리론의 주장을 정리하여 보았다. 이번 장에서는 그런 언어적 표상 체계를 상정할 때, 상식 심리학의 전제 가운데 하나인 심리 상태의 의미론에 대한 가정이 어떻게 정당화되는지 살펴보고자 한다.

이미 지적한 것처럼, 표상적 심리론은 지향적 심리 상태의 속성 전반을 표상의 속성을 통해 간접적으로 설명한다. 때문에 표상적 심리론이 상식 심리학의 전제인 지향적 심리 상태의 의미론 가정을 성공적으로 정당화하기 위해서는 먼저 표상의 의미론에 관해 유물론적 해명을 제시할 수 있어야 한다.

26) Chirchland(1989) 8장 및 12장 참조

그러나 의미의 의미라는 말이 있을 정도로 의미론의 해명이라는 말 자체가 설명을 요하는 모호한 개념이다. 더욱이 표상은 자연 언어가 아니므로, 그것의 의미론 해명이 충족시켜야 할 조건도 자연 언어의 경우와 차이가 난다. 그러므로 본격적인 논의에 앞서, 표상의 의미론에 대한 유물론적 해명이란 도대체 어떤 것이어야 할지, 그 제약 조건들에 대해 간략하게 짚고 넘어가 보자.

첫째로, 표상은 자연 언어가 아니라 두뇌 속의 물리적 실재이므로, 그 의미론에 대한 해명은 화자의 의도나 사회적 규약 같은 것에 의존하여 이루어질 수 없다. 대신 표상의 의미는 표상과 물질적 세계와의 직접적 관계를 밝힘으로써 이루어져야 한다.

둘째로, 물질적 세계와의 직접적 관계라고 해도, 철저히 자연적인 관계들 만이 의미론 해명에서 고려 대상이 된다. 이를테면 경험론자들이 이용하였던 관념과 사물의 유사성 관계 같은 것은, 물질적 세계와의 직접적 관계임에도 불구하고, 물리적 기준을 통해 규정될 수 있는 관계가 아니므로 표상의 의미론 해명을 위한 고려 대상에서 제외된다. 물리적 실재인 표상이 세계와 맺을 수 있는 직접적이고도 자연적인 관계란 인과적 관계 외에는 없다. 따라서 표상적 심리론은 표상과 세계의 인과적 관계를 통해서 그것의 의미론적 속성을 해명하려 시도한다.

셋째로, 표상의 의미론적 해명은 불가피하게 원자론적 형태를 띠지 않을 수 없다. 이미 지적한대로 표상적 심리론은 표상을 물리적 실재라고 가정하며, 동시에 표상의 의미론적 속성 역시 물리적 실재의 인과적/기능적 속성의 하나라고 간주한다. 그런데 여기서 물리적 상태의 안과적/기능적 속성이라는 것은, 요소의 속성이 결합하여 복합체의 속성을 이루는 구성적 구조(compositionality)를 가지지 않을 수 없다. 때문에 표상의 의미론적 속성을 물리적 상태의 인과적/기능적 속성이라고 볼 경우, 표상의 의미론적 구조 역시 마찬가지로 구성적 구조여야 한다.

자연 언어의 경우, 의미의 구성적 구조에 대한 가정을 바탕으로 전개되는 의미론이 이른바 빌딩 블록(building block) 의미론이다. 빌딩 블록 의미론의 기본 발상은, 개별 단어들이 자체로 확정된 의미를 가

지고 있고, 그런 단어들이 결합하여 요소 문장의 의미를 구성하고, 또 그런 요소 문장들이 결합하여 복합 문장의 의미를 구성한다는 것이다. 그러므로 표상의 의미론적 구조가 구성적이라면, 그것의 의미론을 해명하는 작업 역시 자연 언어의 빌딩 블록 의미론에 준하여 원자론적으로 수행되어야 할 것이다.²⁷⁾

이상과 같은 제약 조건만 보아도, 표상의 의미론을 해명하기 위해 넘어야 될 난관이 얼마나 심각한지 드러난다. 의미가 전체론적으로 결정된다고 보는 모든 종류의 전체론, 추상적 대상이나 부재하는 대상에 관한 지시 문제, 지시가 사회적으로 결정된다는 주장, 그리고 이론바 쌍둥이 문제로 불리는 지시의 지소성(indexicality) 문제 등 다양한 종류의 문제들이 위의 세 가지 조건들을 쉽게 충족되기 어려운 것으로 만들고 있다.

그러나 이런 어려움을 살펴보기에 앞서, 우선 표상의 의미론을 해명하려는 시도로 어떤 작업들이 현재 진행되고 있는지 살펴볼 필요가 있을 것이다. 글쓴이의 견해로는 크게 보아 네가지 정도의 접근이 주목받고 있는 것 같다. 그러므로 지금부터는 이들 네가지 접근들에 대해 차례로 살펴보자.

3.1. 드레츠키의 정보 이론

표상의 의미론을 해명하는 작업에서 답해야 하는 문제는 두 가지이다. 먼저 유기체의 내부 상태인 표상이 현실 세계의 어떤 사태를 의미한다고 할 때, 그런 <... 이 ... 을 의미함>의 관계가 도대체 어떤 관계를 말하는지 답하여야 한다. 다음으로는 그같은 의미론적 관계가 어떻게 유기체의 내부 상태와 세계 사이에 형성되게 되었는지 보여주어야 한다.

첫 번째 문제에 대해 드레츠키 이론이 제시하는 대답은 <... 이 ... 를 의미함>이 <... 이 ... 에 관해 정보를 담지함>이라는 것이다. 즉 드레츠키는 두뇌의 한 물리적 상태가 현실 세계의 어떤 사태를 의미한다는 말을, 그 두뇌 상태가 현실 세계의 해당 사태에 관한 정보를 담지

27) 최훈(1994) 참조

하고 있다는 말로 규정하고 있는 것이다.

여기서 정보라는 말은 보다 명세화될 필요가 있다. 드레츠키에 따르면, 한 두뇌 상태가 현실 세계의 어떤 사태에 관한 정보를 담지할 수 있으려면, 문제되는 두뇌 상태의 예화(tokening)와 현실 세계의 해당 사태 발생 사이에 법칙적 상관 관계(nomic relation)가 성립되어야만 한다. 그래서 드레츠키의 이론은 다음과 같은 식으로 정식화된다.²⁸⁾

D1: 기호의 예화인 S 사건이 의미론적 내용 P를 표현하는 것은, “현실 세계의 사태 P가 기호 예화 사건 S를 유발한다”라는 것이 법칙적 일반화일 경우에 한한다: S-events(tokenings of symbols) express the semantic property P if the generalization “Ps cause Ss” is counterfactual supporting.²⁹⁾

한편 이와 같은 정보 담지의 관계, 즉 표상과 현실 사태간의 법칙적 공변(covariance) 관계가 형성되게 되는 과정을, 드레츠키는 유기체 생애 초기의 학습에 의해서 설명한다. 즉 우리의 개념 『강아지』의 예화와 현실적 강아지의 출현이 법칙적 공변 관계에 있게 되는 것은, 개념 『강아지』와 현실적 강아지를 연결시키도록 우리 생애 초기에 훈련 받은 결과라는 것이다. 이것이 앞서 의미 이론의 두 번째 질문에 대한 드레츠키의 답변이다.

행동주의의 자극/반응 학습이론 같은 념새를 질게 풍기는 드레츠키의 이같은 답변은, 그런 점에서 개념 구조의 생득적 선재를 주장하는 포더의 견해와 대조를 이룬다. 요컨대 포더는 현실적 강아지의 출현과 개념 『강아지』의 예화가 공변 관계에 있는 것이 진화가 지각 기체에 부여한 어떤 선재적 형식 덕분이라고 주장하고 있는 반면, 드레츠키는 현실적 강아지의 출현과 개념 『강아지』의 예화가 공변하는 것이 유기체 탄생 이후 학습의 결과라고 보고 있는 것이다.

28) 이하의 논의는 Dretske(1983) 참조

29) 드레츠키의 정식화가 아니고 포더의 재정식화이다. 뒤의 D2도 마찬가지이다. 글쓴이에게는 포더의 정식화가 보다 명료해 보인다.

그러나 이렇게 강아지와 『강아지』의 공변 관계가 유기체의 탄생 이후 학습의 결과라고 보게 되면, 앞서의 「으로서 문제」 외에도 여러 가지 어려운 난점들에 부딪치게 된다. 이를테면 우리가 『너구리』라는 개념을 학습하는 과정은 실제 너구리를 접함으로서 이루어지는 것이 아니라 너구리 인형, 너구리 만화, ‘너구리’라는 말 자체, 심지어는 너구리와 비슷한 족제비를 관찰함으로써 이루어진다. 그리고 『너구리』 개념의 예화도 너구리를 지각하였을 때만 발생하는 것이 아니라 너구리와 흡사한 족제비, 오소리, 수달 등을 지각할 때도 실수로 발생하는 수가 있다. 그렇다면 왜 우리의 『너구리』 개념은 그같은 실수까지 모두 고려하여 『너구리 또는 족제비 또는 오소리 또는 수달』 개념이라고 하지 않고 단지 『너구리』 개념이라고 간주되는가. 결국 이러한 사례는 드레츠키의 의미론적 정식이 잘못되었음을 보여주는 것 아닌가.

이와 같은 문제가 소위 오류 표상의 문제 또는 선접의 문제(disjunction problem)라고 불리는 것이다. 이런 오류 표상의 문제 또는 선접의 문제는, 드레츠키가 단순히 생애 이후의 학습 과정이라고 표현한 개념과 사태의 공변 형성 과정이 충분히 명세화되지 못했음을 보여준다.³⁰⁾

그래서 드레츠키는 자신의 이론을 보다 정교화하여, 개념의 학습이 이루어지는 시기가 생애의 특정 기간으로 한정되어 있다는 주장을 펼친다. 즉 그같은 결정적 개념 학습 시기에 정확한 공변 관계가 이루어지지 못하는 경우, 이후의 생애에서 오류 표상이 유발된다는 것이다. 구체적인 예를 들어서 이 점을 좀더 자세히 살펴보자.

드레츠키는 이를테면 개념 『오렌지』가 오렌지 아닌 자몽의 출현 시에도 예화되는 이유가, 『오렌지』라는 개념을 배울 때, 자몽이라는 유사한 사례를 접할 기회가 없었기 때문이라고 본다. 만일 일찍이 『오렌지』 개념을 학습하는 시기에 자몽을 보고 그것이 오렌지와 다르다는 가르침을 받을 수 있었다라면, 그와 같은 오류 표상은 없었을 것이라는 설명이다. 결과적으로 이것은 개념 『오렌지』의 오류 표상은 그 개념을 학습하는 과정이 정상적인 환경이 못되었기 때문에 발생한

30) Fodor(1992) pp.62-64

다는 말이다. 이와 같이 전개되는 드레츠키의 입장을 『오렌지』를 예로 들어 정리해보면 다음과 같을 것이다.

D2: 개념 『오렌지』가 오렌지 또는 자몽이 아니라 오렌지를 의미할 수 있는 이유는, 정상적인 학습 상황에서 개념 『오렌지』와 공변 관계를 형성하게 될 대상이 오렌지 또는 자몽이 아니라 오렌지이기 때문이다.

그러나 이렇게 정상적인 학습 상황이라는 모호한 양상 개념까지 동원하게 되어서는 드레츠키 이론의 취지를 보존할 수가 없다. 다시 말해, 드레츠키 이론의 핵심은 표상의 의미를 그것이 담지하는 환경에 대한 정보로 간주하자는 것인데, 만일 그 정보의 신뢰성이 유기체에게 직접 알려질 수는 없고 다양한 가능 세계들까지 고려해야 비로소 확보되는 것이라면, 그것은 이미 정보라고 할 수 없기 때문이다.

뿐만 아니라, 보다 심각한 문제는, 개념의 학습 시기가 드레츠키가 말하는 것처럼 그렇게 특정 기간으로 제한될 수 있는지도 의심스럽다는 점이다. 이는 앞서 지적한 개념의 가소성(plasticity) 문제를 다시 한번 상기하게 만든다. 개념이 생득적으로 확정된다고 보기도 어렵지만, 그렇다고 개념의 습득 시기가 생애 초기의 일정 기간으로 제한되어 있다고 생각하는 것 역시 그럴듯해 보이지 않기 때문이다.

3.2. 밀리칸의 고유 기능 이론

밀리칸의 경우, 표상의 의미는 그것의 고유 기능(proper function)을 통해 규정된다. 그리고 그런 고유 기능의 형성은 표상 산출 기제의 자연 선택을 통해 이루어진다.

고유 기능이라는 개념은 보다 상세한 설명을 요하는 개념이다. 대략적으로 정의하자면, 고유 기능이란 한 생물학적 범주가 수행하는 다양한 기능들 가운데서, 자연 선택을 통한 그 범주의 존속과 번영을 인과적으로 설명해 줄 수 있는 기능을 말한다.³¹⁾ 예를 들어, 인간의

31) Milikan(1984) p.24

심장이라는 생물학적 기제가 지금까지 존속하는 것은, 다른 어떤 기능 보다는 혈액을 순환시키는 기능 덕분이라고 할 수 있다. 만일 혈액을 순환시키는 기능을 수행하지 못하였더라면, 지금과 같은 형태의 심장은 진화의 과정에서 이미 제거되어졌을 것이기 때문이다. 그러므로 밀리칸의 정의에 따르자면, 심장의 고유 기능은 바로 혈액을 순환시키는 것이 된다.

밀리칸이 이처럼 복잡한 고유 기능 개념을 도입하고 있는 이유는, 생물학적 범주의 본질적 기능을 자연적이면서도 비자의적인 방식으로 규정하기 위해서이다.³²⁾ 예를 들어 인간의 치아를 보면, 그것은 음식을 씹는 기능 말고도 병마개를 열고, 휘파람 소리를 내고, 심지어 싸움의 무기로도 이용되는 등 다양한 기능들을 수행한다. 그런가하면 어떤 경우에는, 충치 때문에 정작 음식을 씹는다는 기능은 수행하지 못하는 수도 있다. 그러므로 치아가 현실적으로 수행하고 있는 이런 다양한 기능들 가운데서 치아라는 범주의 진정한 생물학적 기능을 규정하기 위해서는, 우연적이고 비본질적인 기능들을 배제하고, 치아가 당연히 수행하도록 기대되는 본질적 기능 만을 골라낼 필요가 있다. 하지만 그와 같은 골라냄이 관습이나 인간의 주관에 의해 비자의적인 방식으로 이루어져서는 의미가 없다. 바로 이런 문제에 직면해서, 밀리칸의 고유 기능 개념은 생물학적 범주들의 본질적 기능이 무엇인지를 자연적이고도 비자의적인 방식으로 규정할 수 있게 해주는 것이다.

물론 밀리칸이 치아의 기능을 설명하려고 고유 기능 개념을 도입하는 것은 아니다. 밀리칸은 생물학적 범주라는 개념을 심장이나 간장 같은 신체 기관뿐만이 아니라, 언어적 표현들, 내적 표상들, 심지어는 일상 도구들까지 포함해서 생존적 가치의 관점에서 평가가 가능한 모든 종류의 범주들을 포괄하는 넓은 개념으로 사용한다. 그러므로 고유 기능 개념을 도입함으로써 밀리칸이 성취하고자 하는 것은, 언어적 표현들이나 내적 표상들을 생물학적 범주의 하나로 분류하여, 그것의 본질적인 기능을 생존적 가치의 관점에서 규정하고, 그런 기능 규정을 통해 언어적 표현들이나 내적 표상들의 의미를 해명하는

32) 앞의 책 pp.1-2

것이다.

언어적 표현의 의미론으로만 국한할 경우, 이와 같은 밀리칸의 시도는 그라이스의 의미 이론과 대조를 이룬다. 그라이스도 언어적 표현의 의미를 그것이 수행하는 실용적 기능을 통해 해명하려 한다는 점에서 밀리칸과 공통되는 측면이 있다.

그러나 하나의 언어적 표현은 문자적 의미를 전달한다는 기능 말고도 현실적으로 다양한 기능들을 수행한다. 그라이스는 그런 다양한 기능들 가운데서 한 언어적 표현의 의미가 되는 기능이 무엇인지는 사용자의 의도에 따라 결정된다고 주장한다. 하지만 이처럼 사용자의 의도를 끌어들이기 때문에, 그라이스의 이론은 언어적 표현의 의미를 자연적인 방식으로 규정하는데 실패하고 만다. 뿐만 아니라 언어적 표현이 사용되는 맥락의 다양성 때문에, 그같은 사용자의 의도를 통해서 사회적으로 통용되는 안정적인 표준적인 의미를 가려 뽑아낼 수 있는지도 의문스럽다.

이에 반해 밀리칸의 접근에서는, 한 언어 표현의 의미가 그 표현이 사람들의 입을 통해 계속적으로 재생산되며 사용되게끔 만들어주는 고유 기능을 통해서 규정된다. 그리고 그와 같은 고유 기능은 사회 성원들에 의해 보편적으로 이해되고 기대되는 표준적이고 안정적인 기능으로 필연적으로 수렴하게 되어 있다. 그 의도가 이해되지 못하는 언어적 표현들은, 사람들의 입을 통해 재생산적으로 전수되지 않으므로 결국은 사멸하고 말 것이기 때문이다.³³⁾

예를 들어 “지금 비가 온다”라는 직설법 문장은, 그것의 기능이 지금 비가 온다는 믿음을 산출하는 것이라고 사회 성원들에 의해서 일반적으로 받아들여지고 있기 때문에, 구체적 상황에서 지금 비가 온다는 믿음을 산출하기 위해 사용되어질 수 있다. 그리고 이같은 의사 전달의 성공 덕분에, “지금 비가 온다”는 문장은 계속적으로 사람들의 입을 통해 재생산되어진다. 결국 “비가 온다”는 문장의 고유 기능은 「청자에게 지금 비가 온다는 믿음을 산출하는 것」으로 규정되는 것이다.

33) 앞의 책 pp.51-71

이와 같은 방식으로, 고유 기능 개념은 한 언어적 표현의 의미를 자연적이고 비 자의적인 방식으로 해명하여 주면서도, 그것을 사회적으로 받아들여지고 기대되는 규범적인 기능과 연결할 수 있게 해 준다.

밀리칸은 이런 고유 기능 개념을 내적 표상들의 의미론에도 적용한다. 그러나 내적 표상의 경우, 그 고유 기능은 사람들 사이의 상호 작용을 매개하는 것이 아니라 유기체와 환경 사이의 상호 작용을 매개하는 것이 된다. 그리고 그와 같은 유기체와 환경 사이의 상호 작용을 제대로 매개하지 못하는 표상 산출 기제는 자연 선택의 과정을 통해서 제거되게 된다.³⁴⁾

예를 들어, 토기가 여우를 발견할 때 『여우가 있다』라는 표상을 예화하게 되는 것은, 그런 예화에 실패한 토기들이 자연 선택을 통해 제거되었기 때문이다. 그러므로 이런 경우, 토기들 머리 속에 존재하는 『여우가 있다』라는 표상의 존속과 번성은 여우가 있음을 알려준다는 그것의 기능에 의해 인과적으로 설명되어진다. 따라서 『여우가 있다』라는 표상의 고유 기능은 다름 아니라 「여우가 있다는 사실을 알려줌」으로 규정되게 된다.

이런 식으로 고유 기능은 표상들의 의미 역시도 자연적이고 비자의적인 방식으로 규정할 수 있게 해 준다.

한편 밀리칸 식으로 표상의 의미를 진화의 역사와 관련시키게 되면, 드레츠키를 괴롭혔던 오류 표상이 문제는 쉽게 해결된다. 즉 밀리칸이 어떤 표상이나 심리 상태의 고유 기능이라고 간주하는 것은, 그 표상이나 심리 상태가 현실적으로 수행하고 있는 기능이 아니라 진화의 역사를 통해 그 심리 상태나 표상의 존속과 재생산에 결정적 역할을 하였던 표준적(Normal)³⁵⁾ 기능이다. 그러므로 어떤 내적 표상이 외부 세계를 정확하게 지시하는 경우는, 그런 지시가 일어나는 환경이 진화의 역사에서 문제의 표상을 산출하는 기제가 선택되도록 만든

34) Milikan(1990) pp.123-133

35) 밀리칸이 사용하는 'Normal'과 'normal'은 그 의미가 다르다. 간략하게 설명하자면, 앞의 'Normal'은 한 종의 진화사라는 통시적 배경에 비추어 보았을 때 정상적임(곧 할일을 제대로 함)을 뜻하고, 이에 비해 뒤의 'normal'은 일상적인 의미대로 현재의 조건 하에서 정상적임(곧 할일을 제대로 함)을 뜻한다.

환경, 즉 표준적(Normal) 환경과 일치하는 경우뿐이다. 표준적 환경이 아닌 환경에서 표상이 예화되는 경우, 그런 표상은 오류 표상이 될 수밖에 없다. 이를테면 알에서 갓 태어난 오리 새끼들이 장난감 오리를 보고 『어미 오리』라는 표상을 예화시키는 것은, 오리 새끼들의 현재 환경이 진화 역사에서 오리의 표상 산출 체계로 하여금 『오리』라는 표상을 예화하도록 강요한 표준적 환경과 다르기 때문이다. 이런 식으로, 밀리칸의 고유 기능 이론은 어떻게 표상이 현실 세계에 대한 정보를 담지하게 되는지를 설명하면서도, 그런 설명 안에 오류 표상의 여지를 허용하고 있다.

글쎄이는 밀리칸의 접근이 타당성과 설득력을 가지고 있다고 본다. 다만 한 가지 문제점은, 스티어르니가 지적하는 것처럼, 밀리칸의 이론에서 자연 선택이 감당하고 있는 설명의 역할이 너무 크다는 점이다. 자연 선택은 어떤 것이든 설명할 수 있는 만병 통치약이 아니라는 말이다. 밀리칸 이론의 이런 문제점을 보다 분명히 파악하려면, 밀리칸처럼 진화론적으로 표상의 의미에 접근하면서도 세부적인 사항에서는 미묘한 차이를 보이고 있는 파피뉴의 이론을 살펴보는 것이 필요하다.

표상에 대한 파피뉴 주장의 골자를 추리하면,³⁶⁾ 하나의 믿음이 가진 의미론적 기능은 과거 진화의 역사에서 선조 유기체들이 그와 같은 믿음을 가졌을 때 처해 있던 환경을 지시하는 기능이라는 것이다. 그러므로 여기까지는 파피뉴의 주장도 밀리칸의 주장과 유사하다. 파피뉴와 밀리칸의 차이는, 표상으로 하여금 그런 기능을 갖게 만드는 자연 선택을 과연 무엇을 선택하는 과정이라고 보는지의 차이이다. 밀리칸은 자연 선택을 표상을 산출하는 기제의 선택 과정이라고 보는 반면, 파피뉴는 자연 선택을 통해 표상 산출 기제가 아니라 표상들이 직접 선택된다고 보기 때문이다.

이런 파피뉴의 주장은, 눈이라는 시각 시스템이 자연 선택을 통해 진화한 것이 아니라, 눈을 통해 보게 되는 시각 장면들이 진화하였다는 주장과도 같다. 하지만 글쎄이가 지금 눈으로 보고 있는 시각 장

36) Papineau(1993) p.59

면(본 논문의 한 페이지를 영사하고 있는 컴퓨터 화면)은 과거 글쓴이의 조상들이 전혀 접해보지 못하였던 시각 장면이다. 그러므로 고드프리드 스미스가 지적하는 것처럼, 역사상 최초로 발생하게 되는 시각 장면들이 자연 선택을 통해 결정될 수는 없다. 결국 표상들의 의미가 자연 선택을 통해 확정된다는 파피뉴의 주장은 설득력을 갖지 못하는 것이다.³⁷⁾

이런 어려움을 피하기 위해, 밀리칸은 생물학적 기제에 대한 자연 선택과 그런 기제가 산출하는 상태들에 대한 자연 선택을 분명히 구분한다. 그리고 자신은 표상의 의미 확정에 관한 한, 후자의 자연 선택을 거부한다. 대신 표상을 산출하는 생물학적 기제에 대한 자연 선택을 통해, 기제의 각 상태들까지도 적응적 기능을 가지게 된다고 주장한다. 즉 기제 자체가 가지는 기능이 직접적 고유 기능이라면, 그런 기제의 상태들은 기제로부터 파생되는 일종의 파생적 고유 기능을 갖게 된다는 것이다.³⁸⁾

하지만 이런 식의 설명은 정작 가장 중요한 부분, 즉 도대체 어떤 방식으로 하나의 표상이 세계를 지시하게 되는지에 대해서는 알려주는 것이 없다. 직접적 고유 기능과 파생적 고유 기능 사이의 넓은 설명적 공백이 그대로 남아 있는 것이다. 밀리칸의 이론이 말하고 있는 것은, 어떤 알지 못하는 방식에 의해서 표상이 세계를 지시할 수 있고, 그런 알지 못할 방식은 진화를 통해 결정되었다는 것 뿐이다. 결국 밀리칸의 이론은, 의미론에 관한 한 발견적 효용은 없고 단지 정당화의 효용만을 가지게 된다.

그러나 글쓴이는 밀리칸 이론의 이런 약점이 포더식의 이론과 제휴함으로써 상당 정도 완화된 수 있을 것으로 본다. 밀리칸의 이론이 일종의 통시적 이론으로서, 표상 산출 기제의 형성 과정에 관심을 가지고 있다면, 포더의 이론은 일종의 공시적 이론으로서, 표상 산출 기제가 현재 어떤 식으로 작동하는지에 보다 많은 관심을 가지기 때문이다.

또 한 가지 밀리칸 이론의 어려움은, 포더의 개념 생득론에서도 문

37) Sterelny 앞의 책 p.131

38) Milikan(1984) pp.39-49

제가 되었던 개념 구조의 가소성 문제이다. 밀리칸은 표상 산출 기계가 진화를 통해 선택된다는 점만을 강조하다보니, 우리가 생애를 통해 새로운 개념들을 배우고 또 그것들을 정교화한다는 사실을 소홀히 다루고 있다. 그러나 유기체 가운데 최소한 인간과 같은 고등 동물들은, 탄생 이후에 경험을 통해서 새로운 개념들을 배울 수 있고, 또 그렇게 획득된 개념들을 경험과 연결시키는 방식까지도 어느 정도 변경시킬 수 있다고 보는 것이 경험적으로 보다 그럴듯할 것이다.

3.3. 포더가 제안한 표지(indication)의 비대칭적 의존성

앞에서 포더의 개념 생득성 논제를 살펴보았다. 그렇지만 개념 생득성 논제는 표상과 세계 사이의 의미론적 관계가 어떻게 형성되는지에 대한 설명일 뿐이므로, 포더 역시도 하나의 표상이 무언가를 의미한다는 것이 도대체 어떤 관계를 말하는가 하는 보다 원론적인 질문에 대해서는 또다른 대답을 준비해야만 한다.

포더는 퍼트남 등과 달리, 표상의 의미론적 속성이 그것의 지시 관계를 통해 결정된다고 보지 않는다. 즉 표상의 의미와 그것의 외연적 지시를 구분하고 있다. 소위 좁은 내용(narrow content) 논제³⁹⁾라고 불리는 포더의 이런 주장은, 사실상 표상적 심리론 전체의 운명이 그것의 진리성 여부에 달려 있다고 할 정도로 심각한 중요성을 가지고 있다. 그러므로 이에 대해서는 별도로 한 절을 할애해서 다루기로 하고, 여기서는 다만 포더가 사용하는 의미 개념이 지시와는 관련이 없다는 점 만을 지적하고 넘어가겠다.

그러나 포더가 표상의 의미를 지시와 상관 없는 것으로 본다고 해서, 그것이 현실 세계와 전적으로 무관하다고 주장하는 것은 아니다. 포더 역시 앞서 밀리칸이나 드레츠키처럼, <표상이 무언가를 의미한다>는 말을 <표상이 현실 세계의 어떤 사태를 표지(indicate)한다>는 말로 이해하고 있기 때문이다. 물론 여기서의 표지 관계는 지시 관계와 달리 오류의 여지를 허용하는 관계이다. 그러므로 포더의 이

39) 좁은 내용 논제에 대해서는 블록과 공동 집필한 논문 “심리 상태가 아닌 것” Fodor(1981) 3장 참조.

론도 앞에서 살펴본 선접 문제에 대해 어떤 식으로든 대답을 해야 할 것이다.

선접 문제의 요점을 다시 한번 정리해보자면, 우리의 『너구리』 표상이 너구리를 보았을 경우에만 예화되는 것이 아니라, 쪽제비를 보았을 때에도 실수로 예화되는 수가 있는데, 왜 우리는 그 표상을 『너구리』 또는 쪽제비』 표상이라고 하지 않고 단지 『너구리』 표상이라고 하는가 하는 것이다.

이와 같은 선접 문제의 해결은, 하나의 표상이 너구리의 표상이기 위한 조건이 무엇인지를 명시적으로 밝힘으로써 이루어져야 한다. 포더는 표지 관계의 비대칭적 의존성이라는 개념을 동원하여 그같은 조건을 제시하고자 한다. 포더의 제안은 이렇다.

암소가 『암소』 토큰을 예화하고 또 고양이도 『암소』 토큰을 예화한다. 그러나 『암소』가 암소를 의미하고 암소 또는 고양이를 의미하지 않는 이유는 고양이에 의해 『암소』가 예화된다는 사실이 암소에 의해 『암소』가 예화된다는 사실에 의존하고 있기 때문이다.⁴⁰⁾

결국 포더의 논점은 고양이가 『암소』를 예화시키는 사건은 암소가 『암소』를 예화시키는 사건에 의존하는 반면, 암소가 『암소』를 예화시키는 사건은 고양이가 『암소』를 예화시키는 사건에 의존하지 않는다는 것이다. 이는 다시 말하자면, 고양이에 의해서 야기된 『암소』 토큰은 암소에 의해서 야기된 『암소』 토큰에 비대칭적으로 의존하고 있다는 말이 된다. 포더는 이와 같은 표지 관계의 비대칭적 의존성(asymmetrical dependence)이 『암소』가 암소 또는 고양이가 아니라 암소를 의미하는 표상이기 위한 조건이 된다고 본다.

3.4. 불록의 개념적 역할 의미론

지금까지 살펴본 이론들은 모두 현실 세계와의 관련을 통해 표상

40) Fodor(1992) p.91

의 의미론적 속성을 밝혀보려는 시도들이었다. 하지만 블록은 표상의 의미론적 속성을 그것이 다른 표상들과 맺고 있는 구문론적 관계를 통해 규정하려고 한다. 즉 한 개념의 의미는 정의나 추론에서 그것이 담당하는 개념적 역할(conceptual role)이라는 것이다.⁴¹⁾

블록이 표상의 의미를 이처럼 개념적 역할을 통해 규정하고자 하는 동기는 크게 두 가지 정도로 볼 수 있다. 첫 번째 이유는 표상의 경우에도 프레게가 말한 뜻과 지시체의 구분 같은 것이 요구된다는 점이다. 예를 들면, 오이디프의 머리 속에 있는 『조카스터와 결혼하고자 한다』는 표상과 『어머니와 결혼하고자 한다』는 표상은 분명히 구분되어야 한다. 그러나 오이디프스의 내적 표상 『조카스터』와 『어머니』는 그런 의미 구분의 필요성에도 불구하고 그 외연이 동일하다. 결국 이런 구분을 가능케 하기 위해서는, 표상의 의미를 그 외연을 통해서 규정해서는 안되고, 그것이 다른 표상들과 맺고 있는 관계를 통해 규정해야 할 것으로 여겨지는 것이다.

두 번째 이유는 우리가 가진 표상 가운데 많은 것들이 현실 세계에 그 지시 대상을 갖지 않는 추상적 개념이나 허구적 개념들이라는 점이다. 이를테면 『마녀』나 『유니콘』 또는 『시대 정신』, 『프로지스톤』 같은 개념들이 그런 예일 것이다. 이같은 표상들은 어쩔 수 없이 외연이 아닌 개념적 역할을 통해 의미를 부여받을 수밖에 없게 된다.

그러나 블록의 경우에도, 표상의 의미를 외연적 지시와 전적으로 무관하다고는 보지 않는다. 그렇기 때문에 블록은 포더처럼 표상의 의미를 좁은 내용으로 규정하며, 이런 좁은 내용과 맥락 조건의 결합을 통해 표상의 외연적 지시가 결정된다고 주장한다.

그렇지만 좀더 자세히 살펴보면 포더의 좁은 내용과 블록의 좁은 내용은 분명히 구분된다. 포더의 좁은 내용 논제는, 감각적 한계에 기인한 표상의 외연 구분 불가능성을 인정하는 정도에서 그친다. 그래서 쌍둥이 지구와 같이 특이한 상황이 아니라, 인간의 감각을 통해서 정상적으로 구분이 가능한 대상들 사이에서는, 좁은 내용을 통해서도 충분히 외연적 지시가 이루어질 수 있다고 포더는 보는 것이다. 이에

41) 블록의 이론에 대한 비판은 Fodor & Lepore(1992) 6장 참조

반해, 블록의 좁은 내용 논제는 표상의 의미를 개념적 역할로만 규정하고 있기 때문에, 표상의 내용은 단지 외연뿐만이 아니라 감각적 경험과도 무관하게 된다. 결국 블록 식의 좁은 내용은 순수하게 구문론적 규정만을 가지게 되는 것이다. 그러므로 글쓴이의 생각으로는, 포더 식으로 이해된 내용을 좁은 내용이라고 부른다면 블록 식의 내용은 아마 좁디 좁은 내용 정도로 불러야만 할 것 같다.

이와 같이 보자면, 블록의 이론 같은 이중 요소적(two-factor) 접근⁴²⁾이 성공하기 위해서는 개념적 역할과 맥락의 결합이 어떻게 표상의 외연을 결정하게 되는지 분명히 밝혀져야 함을 알 수 있다. 그러나 블록은 이에 대해 어떤 이론적 개입도 보여주지 않는다. 단지 맥락 조건과 개념 역할의 결합을 통해 외연이 고정될 수 있다는 주장에 그칠 뿐이다.

글쓴이의 생각으로는, 블록과 유사한 데빌의 견해를 통해 이러한 블록의 약점이 보완될 수 있을 것 같다. 데빌은 블록과 마찬가지로 표상의 의미를 그 외연적 지시와 개념적 역할의 두 요소로 분석하면서도, 블록이 간과하고 있는 두 요소 사이의 관계를 해명하려고 하기 때문이다.⁴³⁾

데빌의 생각을 간략히 말하자면, 개념들이 세계와의 경험적 관련 정도에 있어서 하나의 위계를 이루고 있다는 것이다. 다시 말해, 모든 개념들의 의미가 개념적 역할과 외연적 지시라는 두 요소를 공히 가지고 있는 것이 아니고, 그 둘의 비율에 있어 차이가 나는 다양한 개념들의 위계가 있다고 보자는 것이다. 예를 들어, 감각적인 경험에 밀접한 기초 개념들은 외연적인 지시에 의해서 의미가 결정되는 반면, 보다 추상적인 개념들은 외연 보다는 기초 개념들과 맺고 있는 정의 내지 추론 관계 상의 개념적 역할을 통하여 의미가 상당 부분 결정된다는 것이다.⁴⁴⁾

글쓴이는 이러한 데빌의 제안이 상당히 건전하다고 본다. 그러나 표상적 심리론이 데빌의 이같은 제안을 수용하기 위해서는, 개념의

42) 개념의 추론적 역할과 지시적 속성을 구분하려는 입장

43) Lycan(1990) pp.371-398

44) 앞의 책 p.376

가소성이라는 문제와 의미의 전체론이라는 문제를 어떤 식으로든 처리해야만 할 것이다. 표상의 의미론을 해명하고자 하는 대부분의 시도들이 표상과 세계와의 직접적인 인과 관계를 통해서 그 해명을 도모하고 있는 데에는, 그런 의미론만이 개념 가소성 문제나 의미 전체론의 문제를 피해갈 수 있다는 사실이 상당 부분 이유로 작용하고 있기 때문이다. 다음 절에서는, 이러한 사항들까지도 고려하면서 표상의 의미론에 대한 비판적 시각을 검토하여 보겠다.

4. 표상의 의미론 해명 작업에 대한 비판

지금까지는 표상적 심리론에서 어떤 식으로 표상의 의미론에 대한 해명이 시도되고 있는지를 살펴보았다. 이번 장에서는 그같은 시도에 대한 비판적 견해들을 검토하여 보면서, 과연 표상의 의미론에 대한 해명 작업이 성공할 수 있는 것인지에 대해 논의하여 보도록 하자.

표상의 의미론에 대한 해명 작업이 성공할 수 없다는 비판적 견해는 크게 두 가지 방향에서 제시되고 있다. 하나는 처치랜드 부부 등의 제거론 진영에서 제기되는 견해로서, 지금까지 제안된 이론들의 경험적 기반이 취약하다는 점을 주로 지적하고 있다. 또다른 비판론은 퍼트남 등에 의해 제기되는 것으로서, 이것은 표상의 의미론에 대한 해명 작업이 일종의 잘못 설정된 과제(ill posed problem)이기 때문에 성공할 수 없다는 보다 원론적인 비판을 담고 있다. 그 두 종류의 비판론 가운데 여기서는 철학적으로 보다 신중한 검토를 요하는 퍼트남의 견해에 대해 살펴볼 것이다.

4.1. 퍼트남의 비판에 깔려있는 비개체주의적 전제

퍼트남은 「표상과 실재」⁴⁵⁾의 초반부 3 장을 통해 표상적 심리론이 표상의 의미론 해명에 성공할 수 없다는 비판을 강한 논조로 제기하고 있다. 그 글에는 의미 전체론 문제로부터 지시의 사회성 문제까지,

45) Putnam(1988)

표상적 심리론에 대해 제기될 수 있는 모든 문제거리가 백화점식으로 나열되고 있는데, 그러나 그럼에도 불구하고 정작 퍼트남이 제시하는 문제들이 왜 표상적 심리론의 시도를 반박하게 되는가는 분명하게 드러나지 않는다. 그 이유는 무엇일까. 퍼트남이 열거하는 문제가 표상적 심리론에 대해 진정한 반박이 되기 위해서는 그에 앞서 별도의 배경 전제가 타당한 것으로 논증되어야만 하는데, 퍼트남은 그 글에서 이같은 논증을 제시하고 있지 않기 때문이다. 여기서 별도의 논증이 요구되는 전제란 바로 심리 상태의 내용이 비개체주의적으로 분류되어야 한다는 전제를 말한다.

개체주의와 비개체주의(individualism vs nonindividualism)의 문제는 <좁은 내용 문제> 또는 <유아론 문제> 등의 다른 이름으로 불리우며, 표상적 심리론의 타당성을 둘러싼 논쟁에서 항상 논란의 핵심이 되어왔다. 이 문제의 골자를 간단히 말하자면, 심리 상태의 내용을 분류하고 그것의 정체를 규정하는데 있어, 소설론에서 이른바 1인칭 주인공 시점이라고 불리우는 것을 취해야 하는지, 아니면 3인칭 전지적 관찰자 시점이라고 불리는 것을 취해야 하는지의 문제라 할 수 있다. 한 가지 예를 들어 이 점을 보다 자세히 살펴보자.

철수가 만일 브로콜리라는 채소를 파슬리로 잘못 알고 있다면, 철수가 가진 『파슬리가 먹고 싶다』는 욕구는 사실 파슬리에 관한 것이 아니라 브로콜리에 관한 것이다. 그러므로 이 경우, 철수의 욕구 내용은 어떤 관점을 취하느냐에 따라 두 가지 다른 방식으로 분류되어 기술될 수 있을 것이다. 먼저 철수의 관점에 따라 그것을 『파슬리가 먹고 싶다』는 욕구로 분류할 수 있을 터인데, 이런 식의 분류를 지지하는 입장이 개체주의이다. 반면 철수의 욕구가 사실은 브로콜리에 관한 것이므로, 그것을 『브로콜리가 먹고 싶다』는 욕구로 분류해야 한다고 볼 수도 있을 터인데, 이런 입장이 바로 비개체주의의이다.

퍼트남은 그 가운데 비개체주의적 입장을 취한다. 그런데 표상적 심리론은 표상을 두뇌 속의 물리적 실재로 가정하고 있고, 또 그 의미론적 속성이라는 것 역시 그런 물리적 실재의 인과적/기능적 속성으로 가정하고 있으므로, 표상적 심리론이 취하는 심리 상태의 분류 기준은 개체 내 표상의 인과적/기능적 속성이라는 개체주의적 기준이

아니면 앎될 것이다. 그런 개체주의적 분류 기준은 퍼트남이 취하는 비개체주의적 관점에서 볼 때는 심리 상태를 제대로 분류하지 못한다. 때문에 개체주의를 바탕에 깔고 있는 퍼트남으로서는 표상적 심리론이 잘못된 이론이라고 결론내리게 되는 것이다.

전형적인 예로 퍼트남의 쌍둥이 지구 논변을 보면, 쌍둥이 지구에서 물을 바라보는 쌍둥이 퍼트남의 심리 상태나 진짜 지구에서 물을 바라보는 진짜 퍼트남의 심리 상태나, 두뇌의 인과적/기능적 속성이라는 기준으로 보자면 동일해야 한다. 하지만 전지적 관찰자 시점에서 보자면, 분명히 하나는 H_2O 에 대한 생각이고 다른 하나는 XYZ에 대한 생각이다. 그러므로 인과적/기능적 속성을 기준으로 표상을 분류하려는 표상적 심리론의 시도는 비개체주의적 관점에서 볼 때 잘못이라는 것이다.

이와 유사한 비판들이 「표상과 실재」에서도 이어지고 있다. 그러나 정작 그런 비판들이 진정한 비판이기 위해 요구되는 추가의 논증, 즉 비개체주의에 대한 옹호 논증은 퍼트남에 의해 수행되고 있지 않다. 아마도 퍼트남은 그것을 별도의 논변이 필요 없는 당연한 사실로 받아들이고 있는 것 같은데, 글쓴이의 견해로는 그것이 그리 당연한 사실만은 아닌 것 같다.

개체주의를 비판하고 비개체주의의 타당성을 옹호하는 논변은 퍼트남 보다는 주로 버지에 의해서 이루어져 왔다. 그러므로 다음 절에서는 개체주의를 공격하고 비개체주의를 옹호하는 버지의 주장에 대해 정리하여 보면서, 그런 버지의 주장에 어떤 문제점이 있는지를 살펴볼 것이다. 만일 이와 같은 고찰을 통해서 버지의 주장이 설득력이 부족한 것으로 판명된다면, 버지의 비개체주의를 바탕에 깔고 전개되는 퍼트남의 표상적 심리론 비판은 상당 정도 그 힘을 상실하게 될 것이다.⁴⁶⁾

4.2. 버지의 비개체주의 옹호론

개체주의는 심리 상태를 개체 내부의 상태, 특히 두뇌 상태를 기준

46) 그러나 나중에 다시 언급되었지만, 퍼트남 비판의 또다른 칼날인 의미 전체론은 버지의 개체주의와 관계 없이 표상적 심리론에 여전히 위협이 된다.

으로 분류하려는 입장이다. 개체주의자들이 이처럼 유아론적 방식으로 심리 상태를 분류하고자 하는 이유는, 개체의 행동에 대한 인과적 설명이 개체 내부의 상태만 가지고도 완료(exhausted)된다고 보기 때문이다.

이와 같은 개체주의의 골자를 버지는 다음과 같이 요약한다.⁴⁷⁾

전제 1 : 내부 상태가 동일한 두 개체의 행동은 환경적 맥락에 관계 없이 항상 동일하다.

전제 2 : 심리학은 행동에 관한 과학이다. 그러므로 심리학에서 문제가 되는 것은 행동의 원인이 되는 개체 내부의 상태뿐이다.

결론 : 따라서 심리학을 위한 심리 상태의 분류는 개체 내부의 상태를 기준으로 이루어져야 한다.

여기에 대해 버지는 전제 1과 전제 2가 첫눈으로는 타당해보여도 사실 것처럼 쉽게 받아들일 수는 없는 생각이라고 주장한다. 먼저 첫 번째 전제에 대한 버지의 비판을 살펴보자.

첫 번째 전제를 비판하기 위해 버지는 우선 행동이라는 것이 단순한 신체의 움직임과는 구분된다는 점을 지적한다. 행동의 동일성을 판별하는 기준이 신체 움직임은 아니라는 것이다. 버지는 대신 어떤 행동을 어떤 행동과 같다고 볼 것인지의 행동 분류 기준이 사회적으로 확립되어 있다고 주장한다. 그리고 그와 같은 사회적 기준은 행동을 사회적 관계라는 관점에서 분류한다는 점을 지적한다.⁴⁸⁾

이처럼 행동이 사회적 관계를 고려하여 분류될 경우, 그것은 신체의 움직임과는 전혀 다른 이질적 분류 체계를 구성하게 된다. 예를 들어 '애정 표시'라는 행동 집합은 그 분류에 있어 다양한 사회적, 자연적 환경 맥락들을 고려해야 하기 때문에, 특정한 종류의 신체 움직임의 집합으로는 규정될 수가 없게 되는 것이다. 만일 '애정 표시'라는 행동 집합이 신체 움직임의 집합으로 규정되지 못한다면, 그것을

47) Burge(1986)

48) 앞의 글 pp.15-16

개체 내부의 상태를 기준으로 규정하는 것은 더더욱 불가능하다. 결국 행동 기술과 내부 상태의 기술은 서로 비투사적인 관계에 놓이게 되는 것이다. 뿐만 아니라 더 나아가 개체의 행동에 대한 기술과 개체의 내부 상태에 대한 기술을 연결하는 법칙적 관계도 있을 수 없게 된다. 이는 달리 말하자면, 동일한 내부 상태에서부터 유발된 행동이라 하더라도 서로 다른 행동으로 간주될 수 있다는 말이 된다. 결국 이상과 같은 이유 때문에 전제 1의 타당성이 의심스럽다고 버지는 주장한다.

이런 식으로 전제 1을 비판한 버지는 이어 전제 2에 대해서도 비판을 제기한다. 현실적으로 전개되고 있는 심리학적 연구들도 개체주의의 방식을 따르지 않는다는 것이다. 그런데 그가 전제 2를 공격하기 위해 들고 있는 심리학적 연구 사례는 다름 아니라 데이비드 마의 시각 연구이다. 마의 연구에서 시각 시스템에 주어지는 입력들은 물론 개체주의적으로 기술되지 않고 비개체주의적으로 기술된다. 시각 시스템에 주어지는 입력 자료들은, 그것이 시각 주체에게 어떻게 보이는지에 따라서가 아니라, 객관적으로 그것이 어떤 것에 관한 자료인지에 따라서 분류된다는 것이다. 이를 근거로 버지는 현행 심리학 연구들이 심리 상태를 비개체주의적 방식으로 분류하고 있다고 주장한다.⁴⁹⁾

하지만 글쓴이가 볼 때, 마의 연구 사례는 버지의 주장을 단지 사소하게 지지하고 있을 뿐이다. 심리 상태의 개체주의적 분류와 비개체주의적 분류가 문제되는 경우는 불투명 맥락이 형성될 수 있는 표상 수준의 분류인데 비해, 버지가 들고 있는 마의 예는 표상이 아니라 감각 수준의 처리에 관련되어 있기 때문이다. 감각 자료의 처리는 비교적 보편적인 생물학적 기체에 의해서 진행되는 탓으로 개체들간의 차이가 그다지 문제되지 않는다. 그러므로 굳이 개체주의적 관점을 필요로 하지 않는 것이다. 하지만 이와 같은 보편성이 표상 수준에서는 유지되지 않는다. 즉 개체들마다 각자의 배경 지식이 다르고 가진 욕구들이 다를 수 있기 때문에, 표상 수준의 연구에서는 개체들

49) 앞의 글 p.29

간의 차이를 보다 중시하는 방식으로 심리 상태가 분류되어야 한다고 볼 충분한 이유가 있는 것이다.

하지만 버지의 비개체주의 옹호론에서 발견되는 보다 심각한 문제점은, 그것이 인식론적 문제를 존재론적 문제와 결합시키고 있다는 점일 것이다. 즉 개체주의와 비개체주의 사이의 핵심적 대립점은, 인과적 효력을 지닌 심리 상태의 분류가 개체주의적인 기준을 따른 분류인가 아니면 비개체주의적 기준을 따른 분류냐 하는 존재론적 문제인데 비해, 버지가 비개체주의의 근거로 들고 나오는 논점은, 우리가 확보할 수 있는 행동에 대한 법칙이 개체주의적으로 분류된 심리 상태와 행동간의 법칙일 수 없다는 인식론적 난점이기 때문이다.

개체주의적으로 분류된 심리 상태와 행동간의 법칙화가 어려운 이유는, 우리가 다른 사람들의 행동을 그 의도에 따라 정확하게 분류할 수 있게 해주는 사회적으로 확립된 분류 체계를 가지고 있지 못하기 때문이다. 하지만 글쓴이가 보기에, 그와 같은 인식론적 난점이 개체의 행동이 개체의 내부 상태에 의해 야기된다는 존재론적 사실을 반박하지는 못할 것으로 여겨진다.

4.3. 퍼트남 비판의 요지

퍼트남은 표상의 존재를 가정함으로써 완수해야 하는 목표가 자연 언어의 의미론 해명이라고 본다. 즉 표상의 존재론이 요구되는 이유가 자연 언어의 의미를 내적 표상과 연계시켜 설명하기 위함이라고 보는 것이다.⁵⁰⁾

이처럼 표상의 용도가 자연 언어의 의미론 해명이라고 보게 되면, 여기서 요구되는 표상이 사회적 의사 소통을 설명해 줄 수 있는 비개체주의적으로 분류된 표상이어야만 하는 것은 당연할 것이다. 그러므로 퍼트남은 표상적 심리론의 타당성 여부를, 그것이 제시하는 표상 분류 기준이 공공적 의미 해명을 위한 비개체주의적 표상 분류를 가능하게 하는지 여부에 의해 평가하려고 한다. 그런데 표상적 심리론

50) Putnam(1986) p.66 “정신주의의 전체 목표가 단어의 의미를 단어 사용법을 아는 화자의 두뇌 속에 있는 어떤 것과 동일시하는 것이기 때문이다.”

이 가정하는 분류의 기준은 바로 두뇌 속 실재인 표상의 인과적/기능적 속성이다. 따라서 퍼트남 식으로 보자면, 표상적 심리론에 대한 평가는 표상의 인과적/기능적 속성을 기준으로 삼아 공공적 의미 해명을 위한 비개체주의적 표상 분류가 이루어질 수 있는지 여부에 의해 이루어진다.

인과적/기능적 속성을 기준으로 분류되는 표상들이 비개체주의적 관점에서 합당한 분류이기 위해서는, 동일한 개념과 연결된 표상들 사이에 인과적/기능적 속성의 어떤 공통점, 다시 말해 본질이 있어야 할 것이다. 그러나 퍼트남은 동일한 표상 유형에 의해 공유되는 그와 같은 본질적인 인과적/기능적 속성이란 존재하지 않음을 지적한다. 다시 말해 『옆집 고양이와 어제 강아지와 싸웠다는 믿음』으로 분류된 모든 심리 상태들이 공유하고 있는 초개체적이고 초시간적인 인과적/기능적 공통 속성 같은 것은 없다는 것이다.⁵¹⁾

퍼트남은 이같은 인과적/기능적 공통 속성의 부재가 표상적 심리론이 잘못 정초되었음을 함축하게 된다고 본다. 그래서 표상적 심리론의 비판을 위해 그런 공통 속성의 부재를 증명하려 한다.

우선 퍼트남은 포더가 말하는 생득적 보편 개념들에 대한 검토부터 시작한다. 앞에서 살펴본 대로, 포더의 개념 생득성 가설은 감각으로부터 개념을 매개하는 과정이 보편적이고 생득적인 지각 기제, 즉 모듈 체계에 의해 지배된다고 본다. 그러므로 그같은 가설이 옳다면, 하나의 개념을 예화시킬 수 있는 감각 입력들은 개체를 넘어서서 보편적이고 안정적인 방식으로 이미 결정되어 있어야만 할 것이다. 따라서 이 경우, 동일한 표상 유형들은 그것을 유발시킨 어떤 공통적인 인과적/기능적 속성을 가진다고 말할 수 있게 된다.

그러나 퍼트남은 그와 같이 어떤 개념을 필연적으로 예화시키게 되는 보편적이고 불변적인 감각 입력들의 집합이란 존재할 수 없다고 주장한다. 그리고 이같은 사실을 입증하기 위해 세 가지 근거를 제시한다.⁵²⁾

51) 앞의 책 p.28

52) 앞의 책 pp.37-51

퍼트남이 제시하는 첫 번째 근거는, 의미가 전체론적으로 결정된다는 사실이다. 이와 같은 의미 전체론은 그 자체로는 개념 생득성 논제에 위협이 되지 않는다. 다만 그런 전체론적 의미 그물망이 시간에 따라 유동적이고 사람들에 따라 다르다는 점 때문에 문제가 되는 것이다. 즉 의미 그물망이 시간에 따라 달라질 수 있다면, 하나의 개념이 예화되기 위해 요구되는 감각 입력들의 종류도 그때마다 달라져야 한다. 뿐만 아니라, 그렇게 유동적인 의미 그물망이 사람들 사이에서 보편성을 유지하기도 힘들므로, 사람마다 감각 입력을 개념과 연결하는 방식에 있어 차이가 나게 된다. 결국 의미 전체론이 사실이라면, 하나의 개념에 본질적으로 연결된 감각 입력들의 집합이란 존재할 수 없게 되는 것이다.

퍼트남이 제시하는 두 번째 근거는, 단어의 의미가 원칙적으로 자비의 원리와 같은 합리성의 규범을 따라 결정된다는 사실이다. 이를테면 ‘전자’ 같은 말의 의미는 그 말이 포함된 이론을 최대한 합리화해 주는 방식으로 결정된다. 그래서 1900년 보어가 사용한 단어 ‘전자’와 그가 상보성 원리를 발견한 1934년 이후에 사용한 단어 ‘전자’가 이론적으로는 분명히 일치될 수 없음에도 불구하고, 기존 이론을 최대한 합리화 해주는 방식으로 ‘전자’의 지시 대상이 재해석되기 때문에, 여전히 동일한 의미를 가진다고 간주되는 것이다. 만일 이처럼 한 개념의 의미가 이론 전체의 합리화를 위해 유동적으로 재조정될 수 있다면, 그런 개념에 연결된 보편적이고 불변적인 감각 입력들의 집합이란 있을 수가 없다.

퍼트남의 세 번째 근거는, 진화에 의해 미리 예측되지 못하였을 『카부레타』나 『양자 잠재력』 같은 개념들의 경우, 그 의미가 생득적으로 고정되어 있을 수 없다는 점이다. 인간의 지각 기체 내에 『카부레타』 같은 신종 개념들을 예화시키기 위해 어떤 감각 입력들이 요구되는지에 대한 정보가 이미 생득적으로 내장되어 있다고 보는 것은 아무래도 그럴듯하지 못하기 때문이다. 요컨대 『카부레타』 같은 신종 개념들은 그것과 연결된 보편적이고 불변적인 감각 입력들의 집합을 가질 수 없는 것이다.⁵³⁾

이상과 같이 포더의 생득적 사고 언어 가설을 공격한 퍼트남은, 이

어서 포터뿐만이 아니라 표상적 심리론 전반에 적용될 수 있는 보다 포괄적인 비판을 제기한다. 그 비판은 자연종 명사의 지시 관계가 지소적 성격(indexicality)을 가진다는 점에 근거한다.⁵³⁾

자연종 명사의 지시 관계가 지소적 성격을 가진다는 말은, 달리 말하면, 자연종 명사에 대한 우리의 표상이 ‘이것’ ‘저것’이라는 지소사들처럼 맥락 의존적인 방식으로 대상과 연결된다는 말이다.

예를 들어 단어 ‘금’의 지시 관계에 대해 생각해 보자. 단어 ‘금’은 어떤 경우에는 금속 금을 지시할 것이지만, 그런 금을 확인하는 사회적 기준은 지역에 따라 또는 시대에 따라 수시로 달라져 왔다. 그런데 우리가 ‘금’이라는 말을 들을 때 머리 속에 예화시키게 되는 금의 표상은 그와 같이 상대적이고 수시로 변하는 사회적 판별 기준에 의해 영향받지 않을 수가 없다. 그러므로 사람들의 머리 속에 있는 금에 대한 표상은 금속 금과 아무런 항구적 관계도 갖지 않으며, 단지 시대적 상황적 맥락에 의해 우연히 금속 금을 지시한다고 보아야 하는 것이다.

퍼트남은 더 나아가 그와 같은 지시 대상의 확인 방법이 사회의 모든 성원들에게 알려져 있지 않고 대개 일부 전문가 집단에게만 알려져 있다는 점을 지적한다. 만일 것처럼 지시 대상의 판별을 전문가 집단에게 의존해야 한다면, 일반인들이 가진 금에 대한 표상을 금속 금과 관계지우는 것은 더욱더 불가능한 일이 될 것이다.⁵⁴⁾

지시의 지소적 성격에 근거한 이상과 같은 퍼트남의 비판은, 쌍둥이 지구 논변이라는 그의 유명한 논변을 통해 보다 극명한 형태로 표현된다. 그 논변을 다시한번 요약하자면 다음과 같은 식이다.

언어, 풍습, 환경 등이 모두 지구와 동일한 쌍둥이 지구에 쌍둥이 퍼트남이 살고 있다고 가정하자. 그런 쌍둥이 지구의 쌍둥이 퍼트남이 쌍둥이 물을 생각할 때 머리 속에 가지게 되는 표상의 인과적/기능적 속성은, 진짜 퍼트남이 진짜 지구의 진짜 물을 생각할 때 가지

53) 사실 포터도 그런 개념들까지 생득적이라고 주장하지는 않는다. Fodor(1981) p.382 참조

54) Putnam 앞의 책 p.65

55) 앞의 책 p.64

게 되는 표상의 인과적/기능적 속성과 다를 수 없을 것이다. 따라서 그 두 표상은 인과적/기능적 기준에 따를 때 동일한 표상으로 봐야 한다. 그런데 쌍둥이 지구의 쌍둥이 물과 진짜 지구의 진짜 물이 그 현상적 속성의 유사성에도 불구하고 화학적 조성에 있어서는 차이가 난다고 가정해 보자. 즉 쌍둥이 물은 진짜 물 H_2O 와 미세한 차이가 나는 XYZ이라는 분자 구조를 가지고 있다고 가정해 보자. 이 경우, 쌍둥이 퍼트남의 표상은 XYZ를 지시하는 표상이라고 봐야 한다. 반면 지구 퍼트남의 표상은 H_2O 를 지시한다. 결국 인과적 기능적 속성을 기준으로 분류되는 표상은 대상과 지시 관계를 맺을 수 없다.⁵⁶⁾

5. 결론: 표상적 심리론의 평가와 전망

퍼트남이 인과적 기능적 속성이라는 개체 내부적 기준으로 표상의 의미론을 해명할 수 없다고 보게 되는 이유는 그가 잡고 있는 의미론 해명의 목표가 너무 높기 때문이다. 그는 표상의 의미론 해명이 궁극적으로 자연 언어 의미론 해명의 연장선 상에서 요구된다고 생각하고 있다. 자연 언어의 의미론 해명은 단어의 지시 관계나 문장의 진리 조건을 중심으로 전개되는데, 여기서 지시나 진리 조건은 언어 사용자들 간에 공유되는 사회적 규약조차 넘어서 있는 것들이다. 따라서 그와 같은 지시나 진리 조건을 규명하기 위해 도입되는 것이 표상의 의미론이라면, 그런 의미론을 두뇌 속의 물리적 실재인 표상의 인과적/기능적 속성을 통해 해명할 수 없다는 것은 당연하다.

하지만 글쓴이는, 표상적 심리론에 표상의 존재론이 도입된 이유가 자연 언어의 의미론을 해명하기 위해서라기보다는 오히려 행위와 사고의 설명이라는 상식 심리학 본래의 목적 때문이라고 본다. 표상적 심리론은 글자 그대로 심리론의 한 형태이지 의미론은 아니라는 말이다. 그리고 간혹 포더의 경우처럼 표상을 통해 자연 언어의 의미론 해명을 시도한다 하더라도, 그것은 다만 사람들의 의사 소통 행위를

56) 앞의 책 pp.78-86 참조

설명한다는 제한된 목적을 위해서라고 보아야 할 것이다.

사람들의 행위와 사고를 설명한다는 목적을 위해서는 개체주의적 표상 분류가 충분히 적합성을 가지고 있다. 그리고 인과적/기능적 속성이라는 표상적 심리론의 분류 기준은 그런 개체주의적 표상 분류를 위해 아무런 문제 없이 사용될 수 있다. 개체주의적 표상 분류가 행위와 사고의 설명을 위해 보다 적합하다는 점은, 쌍둥이 지구와 같은 상상적 사례가 아니라 포더가 들고 있는 다음과 같이 보다 현실적인 사례를 살펴보면 분명히 드러나게 된다.⁵⁷⁾

개구리의 시각 시스템은 인간의 시각 시스템과는 다르기 때문에 사물의 윤곽이나 색채를 파악하지 못한다고 한다. 대신 개구리의 시각 시스템에는 움직이는 작고 검은 물체를 감지하는 예민한 시각 기구가 마련되어 있다. 이는 물론 파리와 같은 날아다니는 곤충들을 잡아 먹기 위해서이다. 그러나 그렇기 때문에 개구리는 주위에 아무리 많은 파리가 있어도 날아다니지 않는 한 감지할 수가 없어 잡아먹지 못한다. 반대로 개구리 주위에 검은 딱충알 같은 것을 날리면, 이번에는 즉시로 그것을 향해 혀를 낼름 내밀게 된다고 한다. 이제 이와 같은 개구리의 행동을 설명하는 일련의 법칙을 만든다고 해보자. 혀를 내미는 개구리의 행동과 법칙적으로 연결될 수 있는 개구리 심리 상태의 분류와 기술은 개체주의적으로 이루어져야 할 것인가, 아니면 비개체주의적으로 이루어져야 할 것인가.

비개체주의자들의 견해를 따르게 되면, 개구리의 혀내미는 행동은 『파리』에 대한 생각뿐만이 아니라 『검은 딱충알』에 대한 생각 또는 『검정 쿵』에 대한 생각, 심지어는 『검은 점의 홀로그램』에 대한 생각 등 이루 다 열거할 수 없는 무수한 종류의 심리 상태로부터 촉발된다고 보아야만 한다. 그러나 그와 같은 비개체주의적 표상 기술이 과연 개구리의 혀내미는 행동을 설명하는 문제에 관련성을 가지고 있는지의심스러워 보인다.

오히려 개구리의 혀내미는 행동을 설명하기 위해서는, 개구리의 사고 내용을 개구리의 관점에 따라 개체주의적으로 분류, 기술해야 할

57) Fodor(1992) p.73

것으로 보인다. 즉 파리를 지각할 때 개구리의 생각이나, 검은 딱총알을 지각할 때 개구리의 생각이나 모두 동일하게 『음식』⁵⁸⁾에 대한 생각 정도로 기술해야 한다는 말이다. 이렇게 개체주의적으로 심리 상태를 분류해야만 개구리의 심리 상태와 개구리의 행동을 연결하는 적절한 법칙의 확보가 가능해진다. 이러한 사실은 결국 개체주의적 표상 분류 방식이 행동의 설명을 위해 보다 적절하다는 점을 보여주는 것이다.

한편 인간의 경우, 다른 유기체와 달리 언어 사용이 행위의 중요한 부분을 이루고 있기 때문에, 이런 언어 사용 행위를 설명하지 못한다면 표상적 심리론은 행위 설명에 성공한다고 말할 수 없다.

그렇지만 유념할 사항은 인간의 언어 사용 행위라는 것이 크게 보아 의사 소통 행위의 하나라는 점이다. 이런 의사 소통 행위를 설명하기 위해서 표상의 존재론이 요구된다면, 표상의 의미는 퍼트남이 주문하는 것 같은 절대적이고 불변적인 지시 관계나 진리 조건을 통해 규정될 필요가 없다. 즉 ‘금’이라는 말을 사용하는 화자들의 의사 소통 행위를 설명하기 위해서라면 화자들의 머리 속에 있는 금의 표상이 실제 금과 절대적으로 연결되어 있을 필요는 없다는 말이다. 대신 사람들의 금에 대한 표상 가운데 사회적 의사 소통을 보장할 정도의 최소한도의 인과적/기능적 보편성만 존재하면 충분할 것이다.⁵⁹⁾

그런데 다행스럽게도 표상적 심리론은 그와 같은 최소한의 인과적/

58) 사실 포더의 입장에 충실하자면, 이 개념을 인간의 언어를 통해 ‘음식’의 개념이라고 불러서는 안된다. 이 논변은 원래 밀리칸을 비판하기 위해 제시된 것인데, 밀리칸은 표상적 심리론자로서는 드물게 비개체주의를 지지하는 입장이다. 밀리칸의 논지는, 환경을 객관적으로 표지하지 못하고 유아론적으로 표상하는 표상 산출 기제는 진화의 과정에서 대부분이 도태되어 버렸기 때문에, 표상을 개체주의적으로 분류하기보다는 오히려 그것이 마땅히 표지해야 하는 객관적 환경을 통해서 비개체주의적으로 분류해야 한다는 것이다. 하지만 포더는 다음과 같은 한마디로 그런 밀리칸의 주장을 일축한다. “Darwin cares how many flies you eat, but not what description you eat them under. (진화 기제는 내가 얼마나 많은 파리를 잡아먹는지에만 관심을 두지, 그 파리들을 무엇이라 부르면서 잡아먹는지에는 관심을 두지 않는다.)” 즉 진화론적 선택 기제는 표상의 기술(description) 문제와 무관하다는 것이다.

59) 역설적이지만, 사회 전체의 통일된 착각은 오히려 의사 소통의 충분 조건일지 모른다.

기능적 공통성이 어떤 식으로 사람들 사이에 공유될 수 있는지 설명해 줄 수 있다. 즉 앞에서 포더가 제안한 지각적 모듈 체계 이론은, 감각 경험과 연관된 기초 개념들의 표상 가운데 사회적 의사 소통의 기반이 될 인과적/기능적 보편성이 존재한다는 점을 보여주고 있다.

더욱이 이런 지각적 모듈 체계 이론은 「으로서 문제」의 해결을 가능하게 해 준다. 그리고 모듈 체계를 밀리칸 식으로 자연 선택의 결과라고 간주할 경우, 드레츠키와 같은 오류 표상의 어려움에 빠지지 않으면서도, 그것이 가지는 환경 표지의 기능을 설명할 수 있다.

물론 포더는 우리가 가진 개념 가운데 상당 부분이 지각적 모듈 체계에 의해 예화되는 기초 개념들이라고 보고 있지만, 지각적 모듈 체계에 대한 이론을 받아들이기 위해 이런 강한 생득론까지 수용해야 하는 것은 아니다. 오히려 블록이나 데빌이 제안하는 것처럼, 개념들이 다른 개념들과의 정의 관계 내지 추론 관계 같은 개념적 역할을 가지면서 동시에 감각 입력들과도 위계적으로 연결되어 있다고 생각해 볼 수 있을 것이다. 이렇게 볼 경우, 지시 대상을 갖지 않은 개념의 문제도 해결될 수 있다. 또 전문가에게 지시를 의존해야 하는 일부 개념들에 대해서도 설명이 가능해진다. 즉 그와 같은 개념들은, 개념적 역할에 의해 의미의 보다 많은 부분이 결정되고, 감각 입력과의 연결 관계는 상대적으로 덜 중요한, 위계적으로 상위에 위치한 개념들이라고 볼 수 있는 것이다.

그러나 표상 존재론의 도입 이유를 이처럼 의사 소통을 포함하는 행위 설명에 국한한다 하더라도, 몇가지 이유 때문에 퍼트남이 지적하는 의미 전체론 문제는 여전히 표상적 심리론에 부담이 되고 있다.

그 이유는 첫째로, 의미 전체론은 감각 입력과 개념이 연결되는 방식 조차 배경 지식에 따라 변할 수 있다는 점을 시사하고 있기 때문이다. 만일 것처럼 감각 입력과 개념이 연결되는 과정에 배경 지식의 영향이 개입할 수 있다면, 모듈 체계의 안정성은 무너지고, 기초 개념에 있어서 조차도 사람들이 가진 표상들 사이에 인과적/기능적 공통성이 발견되기 힘들게 된다.

두번째로, 의미의 전체론이 시사하는 대규모적이고 신속한 의미 구조의 변동이 두뇌 속 표상들의 구조에 어떻게 것처럼 신속하게 반영

될 수 있는지 설명이 어렵다는 점 때문이다. 만일 개념 구조의 그런 신속한 가소성이 물리적인 방식으로 어떻게 구현 가능한지 대답될 수 없다면, 표상을 통한 마음의 설명도 또하나의 마술이 되고 만다.

세 번째로, 의미 전체론이 함축하는 개념 체계의 전반적 상대성이, 개체의 지향적 심리 상태를 자연 언어를 통해서만 기술할 수 없도록 만든다는 점 때문이다. 즉 의미의 전체론이 타당하다면, 감각 입력과 밀접하게 연관된 일부 기초 개념들의 경우에는 문제가 다르지만, 추상 개념을 포함한 대부분의 개념들에 대해서는 초 개인적인 어떤 기술도 불가능하게 될 것이다. 예를 들어, 앞에서 언급한 철수의 『파슬리』 개념 같은 것도 엄밀히 따져보면 ‘파슬리’라는 자연 언어를 가지고는 기술될 수 없는 것이다. 왜냐하면 자연 언어 ‘파슬리’가 가진 지시적 속성과 개념적 역할이, 철수의 개념 『파슬리』가 가진 지시적 속성 및 개념적 역할과 일치하지 않기 때문이다.

스티치는 내적 표상의 자연 언어적 기술에 수반되는 이런 근본적 어려움 때문에, 표상의 의미론적 속성 자체를 부인하고 대신 그것이 순수히 구문론적 속성만을 가지고 있다고 볼 것을 주장한다. 이같은 주장을 지지하기 위해 스티치가 드는 유명한 예는, 맥켄리 대통령의 암살에 대한 Mrs T의 사고 내용이다.⁶⁰⁾

Mrs T는 젊은 시절 맥켄리 대통령의 암살을 인상 깊게 체험하였지만, 이제는 노쇠하여 그가 암살 당했다는 단편적 사실 외에 그것과 연관된 다른 기억들은 서서히 사라져가고 있다. 그러므로 스티치와 Mrs T가 대화를 나눈다면 다음과 같은 식의 대화가 될 것이다.

스티치 : 맥켄리 대통령에게 무슨 일이 일어났는지 말해주세요?

Mrs T : 오! 맥켄리 대통령은 암살 당했지.

스티치 : 그가 지금 어디 살고 있지요?

Mrs T : 잘 모르겠는데.

스티치 : 제 말은 그가 지금 살아있느냐 아니면 죽었느냐는 거예요?

Mrs T : 누가?

60) Stich(1986) p.56

스티치 : 맥켄리 대통령 말이예요.

Mrs T : 글썄, 잘 모르겠는 걸.

스티치 : 암살 당한 사람은 죽은 사람 아닌가요?

Mrs T : 글썄, 그것도 잘 기억이 나지 않는구만.

스티치 : 그렇지만 방금 맥킨리 대통령이 암살 당했다고 말씀하
셨잖아요?

Mrs T : 그래 맞아, 맥켄리 대통령은 암살 당했어. 분명해.

이와 같은 예에서, Mrs T의 믿음을 과연 『맥켄리 대통령이 암살 당했다는 믿음』으로 기술할 수 있는지 의문스럽다. Mrs T는 분명히 맥켄리 대통령이 암살 당했다는 사실 자체는 인정하지만, 그럼에도 불구하고, 그녀가 가진 믿음의 내용은 일상적으로 “맥켄리 대통령이 암살 당했다”라는 자연 언어 문장이 의미하는 바와 전혀 관련이 없는 듯 보이기 때문이다. 심지어 그녀는 암살 당한 사람이 살아있는지 죽었는지조차도 구분하지 못하고 있다. 이렇게 볼 때, Mrs T의 믿음은 공공적 자연 언어를 통해서 어떤 식으로도 기술이 불가능한 듯 보인다.

포더는 이런 문제를 표상의 번역에 대한 별도의 가정으로 간주하여, 표상이 의미를 가진다는 가정과는 분리시키려 한다. 즉 표상적 심리론이 개입하는 가정은 단지 표상이 의미론적 속성을 갖는다는 가정 뿐이고, 그런 의미론적 속성이 자연 언어로 완벽하게 번역/기술될 수 있다는 별도의 보다 강한 가정은 표상적 심리론과 관련되지 않는다고 보는 것이다. 다음의 말은 포더의 이런 생각을 잘 보여준다.

좁은 내용은 근본적으로 표현이 불가능하다 ... 우리가 내용을 기술하기 위해 문장을 사용하는 경우(즉 문장을 명제 태도 진술의 종속절로 삽입하는 경우), 우리가 할 수 있는 최선의 방도는(이것은 원리적으로도 우리에게 유일한 방도이지만) 안착된 문장(sentences qua anchored)을 가지고 내용을 기술하는 것 뿐이다.⁶¹⁾

61) Fodor(1988) p.50

여기서 포더가 사용하는 ‘안착되었다(anchored)’는 말은 다름 아니라 대화의 공통 지반이 되는 배경 지식에 안착되었다는 말이다. 따라서 그런 배경 지식의 공유를 전제할 때에만, 표상은 공공 언어로 표현 가능하다. 스티치가 들고 있는 극단적인 예에서는, 스티치와 Mrs T 사이에 배경 지식이 거의 공유되지 못하고 있다. 그러므로 Mrs T의 사고 내용을 표현하는데는 당연히 어려움이 따르게 되는 것이다.

하지만 글쓴이는, 포더의 견해를 따른다 해도 여전히 문제 해결에는 미흡한 부분이 있다고 본다. 포더는 대부분의 개념들이 생득적이라고 가정하고 있으므로, 번역의 문제를 고도로 추상적인 일부 개념들에 국한된 특수한 문제로 간주할 수 있다. 뿐만 아니라 광범위한 생득적 기초 개념들을 토대로 삼아 여타의 표상들을 공공 언어로 기술하는 것이 개략적이거나 가능하다고 주장할 수도 있다. 그렇지만 문제는 포더가 생각하는 것처럼 생득적 개념이 전체 개념 가운데서 차지하는 비중이 큰가 하는 점이다. 만일 그렇지 않다면 표상을 공공 언어로 번역하기 어렵다는 문제는 포더가 생각하는 것보다 훨씬 커다란 위협이 될 것이다.

글쓴이는 이같은 문제의 해결 단서가, 밀리칸이 말하는 의미의 사회적 수렴에서 찾아질 수 있다고 본다. 앞서서도 언급한 것처럼, 밀리칸은 자연 언어 표현들이 사회적으로 표준화된 안정적 의미를 갖게 되는 것을, 개인간 의사 소통 과정을 통해 이루어지는 언어 체계의 차별적 선택 때문이라고 설명하고 있다. 곧 언어 체계가 의사 소통 요구의 압력으로 인하여 안정적이고 수렴적인 체계로 진화해 간다는 말이다. 그런데 사회적 선택 기제를 통한 이와 같은 진화는 언어 표현만이 아니라 개인들이 가지고 있는 개념 체계에 있어서도 충분히 일어날 수 있을 것이다. 즉 사회적 의사 소통의 필요성이 개인들의 개념 체계에도 압력으로 작용하여 그것들을 진화시키고, 그런 진화의 결과로 개인들의 개념 체계간에 존재하는 편차가 차츰 좁혀져 전반적 일치점으로 수렴하게 될 수 있다는 말이다.

위노그라드는 기본적으로 이와 동일한 발상 위에서, 개인적 개념 체계의 사회적 수렴을 이야기하고 있다.⁶²⁾ 위노그라드가 보기에, 개인간에 벌어지는 의사 소통은 객관적으로 존재하는 세계에 관한 것이

아니라 대화 참여자들에 의해 공유되는 해석된 세계에 관한 것이다. 다시 말해, 언어 사용의 궁극 목적은 세계에 대해 객관적인 기술을 확보하는 것이 아니라, 각자의 세계 해석을 상호 교류하는 것이란 말이다.⁶²⁾ 그러므로 이런 이해의 목적 하에서 진행되는 개인들간의 대화는, 파손(break down)을 최소화하는 방향으로 개인들의 표상 체계를 조정하는 효과를 가져오게 된다. 여기서 위노그라드가 사용하는 파손이라는 말은, 개념 체계의 상충에서 비롯되는 개체들간의 의사 소통의 단절 현상을 뜻하는 것이다. 그러므로 이런 파손의 극복 과정, 위노그라드가 애용하는 해석학의 표현을 빌자면 개인들 간에 벌어지는 이해의 지평 융합 과정을 통해, 개인적 개념 체계의 사회적 구조 조율(structural coupling)이 이루어지게 된다.⁶³⁾

글쓴이는, 우리가 가지게 되는 개념 체계의 공통성을 설명하는데 있어서, 위노그라드가 제안하는 사회적 구조 조율을 통한 설명이 포더가 제안하는 전반적 개념 체계의 생득성을 통한 설명보다 경험적으로 더욱 설득력이 있다고 본다. 그러나 위노그라드 식으로 사회적 조율을 통해 배경적 개념들의 공유를 설명하려면, 우리 개념 체계의 구조가 의사 소통의 파손을 반영하여, 수시로 조정될 수 있는 가소성을 지녀야만 할 것이다. 만일 그런 가소성을 갖춘 개념 체계가 우리의 머리 속에 물리적으로 실현될 수 없다고 한다면, 위노그라드의 설명 역시도 다시 한번 마술적 설명으로 전락하고 만다.

이제 이상과 같은 사실들을 종합하여 표상적 심리론의 전망에 대한 종합적 결산을 해보기로 하자.

1. 우선 표상의 의미론 해명에 대한 퍼트남의 비판론은, 근본적으로

62) Winograd(1986) pp.48-52

63) 앞의 책 pp.27-37

64) 사회적 구조 조율의 좋은 유비는 물리학에서 말하는 동조 현상이다. 예를 들어 벽시계들의 시계추 길이가 모두 같고 또 벽면에 잘 밀착되어 있다면, 그것들은 처음에 주기가 어떻게 시작되었든 상관없이, 서로에게 간섭하여 차츰 주기가 가까워지다, 나중에는 마침내 모든 시계추가 일치하여 똑딱거리게 된다. 이와 같은 현상을 동조라고 하는데, 이런 동조 현상은 시계가 많을 때 오히려 더욱 잘 일어난다.

로 그가 표상의 존재론이 도입되는 이유를 달리보았다는 점에서 비롯된다. 퍼트남과 달리 표상의 존재론이 행위와 사고의 설명을 위해 도입되는 것으로 이해할 경우, 표상적 심리론이 제안하는 이른바 좁은 내용으로서의 의미가 적합성을 가진다. 그리고 이런 좁은 내용으로서의 의미는 표상적 심리론이 제안하는 방식을 따라 충분히 유물론의 틀 안으로 포섭되어질 수 있다.

2. 좁은 내용으로서의 의미는 개념과 감각 입력과의 연결 관계 그리고 개념과 개념간의 연결 관계라는 이중적 차원을 가질 수 있다. 그리고 생존을 위해 보다 중요한 일부 개념들의 경우, 감각 입력과의 연결 관계는 포더가 말하는 모듈 체계 같은 방식에 의해서 비교적 안정적이고 보편적인 방식으로 고정되어 있다고 가정해 볼 수 있다. 이런 안정적/보편적 개념들을 기초 개념이라고 간주할 때, 그와 같은 기초 개념들을 토대로 어린이의 언어 학습이나 원초적 번역 같은 사회적 의사 소통이 이루어진다고 설명할 수 있다.

3. 그러나 기초 개념이 보편적이고 안정적이라는 가정을 유지하기 위해서는, 의미의 전체론 논제가 함축하는 개념 체계의 유동성이 기초 개념들에 대해서는 적용되지 않는 것으로 입증되어야만 할 것이다.

4. 기초 개념을 제외한 나머지 추상적 개념들의 경우, 그런 추상적 표상들이 의미 전체론 논제에도 불구하고 어떤 의미를 지니기 위해서는 우리 개념 체계의 구조가 가소적이어야 한다. 여기에는 두 가지 이유가 있다. 첫째로, 의미 전체론이 함축하는 대규모적이고 신속한 이론 변화를 반영하기 위해서 개념 체계의 구조는 가소적일 필요가 있다. 둘째로, 의사 소통을 통한 의미의 사회적 수렴이 가능하도록 하기 위해서도 개념 체계의 구조는 가소적이어야 한다.

위의 결론들 가운데 1과 2는 문제가 없지만, 3과 4는 표상적 심리론에 풀어야 할 숙제를 던져준다. 즉 결론 3이 요구하는 것처럼, 기초

개념과 감각 입력의 연결 관계가 전체적 의미 그물망의 변화에 의해 영향받지 않는다는 사실을 표상적 심리론은 아직까지 입증하지 못하고 있다. 또 결론 4가 요구하는 가소적인 구조의 개념 체계가 물리적으로 어떻게 구현될 수 있는지 표상적 심리론은 아직까지 대답하지 못하고 있다.⁶⁵⁾ 이 두 가지 이론적 도전에 대해 표상적 심리론이 자체의 이론적 자원을 가지고 성공적으로 응답할 수 있을 때, 비로소 그것은 표상의 의미론을 완전히 해명하였다고 말할 수 있을 것이다.

65) 글쓴이는 연결주의 인지 모형이 바로 그와 같은 가소적 인지 구조의 물리적 구현 가능성을 보여준다고 생각한다. 하지만 연결주의 인지 모형은, 지금까지 표상적 심리론을 지지하기보다는 제거적 유물론자들에 의해 전유되어 오며 표상적 심리론에 대한 공격의 칼로 간주되어 왔다. 때문에 글쓴이의 석사 학위 논문은, 연결주의 인지 모형에 대해 제거적 유물론자들의 시각과는 별도의 해석을 줌으로써, 그것을 통해 표상적 심리론과 제거적 유물론의 조화와 절충을 시도하고 있다.

참고문헌

- 김영정(1989), “기능주의의 여러 유형과 감각질 문제” 『인지 과학』
서울: 민음사
- _____ (1992), “내재적 실재론에 있어서의 칸트적 요소” 『실재론과
관념론』 서울: 서광사
- _____ (1994), “인지 과학과 가족 유사성” 『철학연구』 봄호 34집
- 김재권(1994), 『수반과 심리철학』 서울: 철학과 현실사
- 민찬홍(1994), “비 환원적 유물론과 인과적 힘” 『수반의 형이상학』
서울: 철학과 현실사
- 박영태(1994), “과학적 실재론에 관한 고찰” 서울대학교 박사학위 논문
- 소홍렬(1993), 『자연주의적 유신론』, 서울: 서광사
- 이영의·이초식(1993), “정보 가변 추리의 특징과 논리적 구조”, 『'93
인지 과학회 춘계 학술대회 발표 논문집』
- 이초식(1993), 『인공지능의 철학』 서울: 고려대학교 출판부
- 정성호(1994), “대상 지칭: 현시와 표상” 『수반의 형이상학』 서울:
철학과 현실사
- 최훈(1994), “연결주의와 구성적 구조: 포더와 그의 동조자들에게 보
내는 응답”, 서울 대학교 석사학위 논문
- Bechtel, W.(1988), Philosophy of Science: An Overview for
Cognitive Science. LEA
- Bechtel, W. and Abrahamsen, A.(1991), Connectionism and the
Mind: An Introduction to Paralell Processing in
Networks, Basil Blackwell.
- Block, N.(1983), “Mental Picture and Cognitive Science” in W.
Lycan(ed) (1990), Mind and Cognition: A Reader,
Oxford: Basil Blackwell 577-606
- Carnap, R.(1966), Introduction to the Philosophy of Science, 『과

- 학 철학 입문(윤용택 역)』 서울: 서광사
- Chomsky, N.(1987), "Language and Inateness" in W. Lycan(ed) (1990), *Mind and Cognition: A Reader*, Oxford: Basil Blackwell 627-645
- Churchland, P.M.(1988), *Mattrer and Consciousness* Cambridge, 『물질과 의식(석봉래 역)』 서울: 서광사
- _____ (1989), *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*, cambridge MA: MIT Press/Bradford Books
- Churchland, P.S.(1986), *Neurophilosophy: Toward a Unified Understanding of the Structure of Science*, Cambridge MA: MIT Press/Bradford Books
- Davidson, D.(1980), *Essays on Action and Events*, Oxford
- _____ (1984) *Inquires into Truth and Interpretation*, Oxford
- _____ (1981), "True Believer: The Intentional Strategy and Why it Works" in W. Lycan(ed) (1990), *Mind and Cognition: A Reader*, Oxford: Basil Blackwell 150-166
- Dennett, D.(1987), "Three Kind of Intentional Psychology" in *Intentional Stance*: 43-83
- Devitt, M.(1990), "A Narrow Representational Theory of the Mind" in W. Lycan(ed) (1990), *Mind and Cognition: A Reader*, Oxford: b? 372-398
- Field, H.(1978), "Mental Represantation" in Block, N.(ed)(1981), *Readings in Philosophy of Psychology vol.2*: 78-116
- Fodor, J.(1981), *Representation: Philosophical Essays on the Foundations of Cognitive Science*, 『표상(정성호 역)』 서울: 민음사
- _____ (1983), *The Modularity of Mind: An Essay on Faculty Psychology*, cambridge MA: MIT Press/Bradford Books

- _____ (1985), *The Language of Thought*, New York: Thomas Y. Crowell.
- _____ (1987), *Psychosemantics: The Problem Of Meaning in Philosophy of Mind*, Cambridge MA: MIT Press/Bradford Books
- _____ (1992) *A Theory of Content and Other Essays*, Cambridge MA: MIT Press/Bradford Books
- Fodor, J. and Pylyshyn, Z. (1988), "Connectionism and Cognitive Architecture: A Critical Analysis" *Cognition* 28: 3-71
- Fodor, J. and Lepore, E. (1992) *Holism: A Shopper's Guide*, Basil Blackwell
- Glymore, C. (1992), *Thinking Things Through: An Introduction to the Philosophical Issues and Achievements*, Cambridge MA: MIT Press/Bradford Books
- Lewis, D. (1972), "Psychophysical and Theoretical Identification" in Block, N. (ed) (1981), *Readings in Philosophy of Psychology* vol. 2: 207-215
- McClelland, L. and Rumelhart, D. and the PDP Research Group (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Cambridge MA: MIT Press/Bradford Books
- Milikan, R. (1984), *Language, Thought and Other Biological Categories*, Cambridge MA: MIT Press/Bradford Books
- _____ (1990), "Compare and Contrast Dretske, Fodor and Milikan on Teleosemantics" reprinted in *White Queen Psychology and Other Essays for Alice*, Cambridge MA: MIT Press/Bradford Books: 123-133
- Nagel, E. (1961), *The Structure of Science*, Harcourt, Brace
- Papineau, D. (1993), *Philosophical Naturalism*, 56-101 Basil Blackwell:
- Putnam, H. (1967), "The Nature of Mental State" in W.

- Lycan(ed) (1990), *Mind and Cognition: A Reader*, Oxford: Basil Blackwell 37-56
- _____ (1981), *Reason, Truth and History*, 『이성, 진리, 역사(김효명 역)』 서울: 민음사
- _____ (1988), *Representation and Reality*, 『표상과 실재(김영정 역)』 서울: 이화여대 출판부
- Ramsey, W. and Stich, S.(1990), "Connectionism and Three Levels of Nativism" *Synthesis* 82: 177-205
- Ramsey, W. and Stich, S. and Garon, J.(1991), "Connectionism, rla Eliminitivism, and the Future of Folk Psychology" in Greenwood, J.(ed)(1991) *The Future of Folk Psychology: Intentionality and Cognitive Science*: 93-119
- Ryle, G.(1949), *The Concept of Mind*, Lomdon, H.C.
- Smolensky, P.(1988), "On the Proper Treatment of Connectionism" *Behavioral and Brain Sciences* 11: 1-74
- _____ (1991), "Connectionism Constituency and the Language of Thought" in Loewer, B. and Rey, G.(ed)(1991), *Meaning in Mind: Fodor and his Critics* Basil Blackwell: 201-228
- _____ (1990), "Tensor Product Variable Binding and the Representation of Symbolic Structures in Connectionist Systems", *Artificial Itelligence* 46: 159-216
- _____ (1993), "The Constituent Structure of Connectionist Mental State: A Reply to Fodor and Pylyshyn" in Horgan, T. and Tienson, J.(ed)(1991) *Connectionism and the Philosophy of Mind*, London: K.A.P.
- Sterelny, K.(1990), *Representational Theory of Mind*, Basil

Blackwell

Stich, S.(1983), From Folk Psychology to Cognitive Science:
The Case Against Belief cambridge MA: MIT
Press/Bradford Books

Tye, M.(1991), "Representation in Pictorialism and
Connectionism", in Horgan, T. and Tienson,
J.(ed)(1991) Connectionism and the Philosophy of
Mind, London: K.A.P.

Winograd, T. and Flores, F. (1986), Understanding Computer
and Cognition, New york: Addison-Wesley P.C.