

## 11. Vorlesung

Ich möchte Ihnen, bevor wir das Thema wechseln und Autoren diskutieren, die irgendeine Variante des Realismus verteidigen, noch einen wichtigen zeitgenössischen Philosophen vorstellen, der einer idealistisch-phänomenalistischen Position zuneigt: Hector-Neri Castañeda. Castañeda ist am 13. Dezember 1924 in San Vicente, Guatemala, geboren und starb am 7. September 1991 in Bloomington/ Indiana als naturalisierter US-Amerikaner. Er war bereits ein bekannter Linguist, hatte u. a. eine Grammatologie, eine Theorie spanischer Grammatiken, *Al margen de la gramática tradicional*, veröffentlicht und die spanische Philologie in Guatemala in Lehre und Forschung neu gestaltet, bevor er 1949 begann, bei Wilfrid Sellars und Mary Brodbeck an der University of Michigan Philosophie, Linguistik, Griechisch und Mathematik zu studieren. Nach dem Abschluss eines von Wilfrid Sellars betreuten Ph. D. an der Universität von Minnesota im Jahre 1954 arbeitete er von 1955-56 mit einem Stipendium an der Universität Oxford. Ab 1974 wirkte er als Mahlon Powell Professor für Philosophie an der Indiana University in Bloomington, wo er die Zeitschrift *Noûs* herausgab, die er 1966 gegründet hatte. Als seine wichtigsten Theorien sah er an: Eine Metaphysik des Praktischen und allgemeine Handlungstheorie, welche die (bisher) umfassendste deontische Logik enthält, die auf der Autonomie des praktischen Denkens gründet. Bekannter wurde Castañeda in den letzten Jahren bei uns aber vor allem, wo nicht ausschließlich, durch seine sprach- und bewusstseinsphilosophischen Arbeiten zur Indexikalität und zum Selbstbewusstsein.

In diesen Arbeiten gibt es viel außerordentlich Interessantes zu lesen und zu lernen – aber das wäre Gegenstand einer anderen Vorlesung. Für uns ist

folgende Konstellation von Bedeutung: Castañeda vertritt eine These, die wir im Grunde von Kant und Fichte kennen (auch wenn er sie im analytischen Vokabular formuliert): nämlich dass das ‚ich‘ des Selbstbewusstseins nicht reduzierbar sei auf das, worauf wir mit Demonstrativpronomina, mit Namenwörtern oder mit (Frege’schen) Kennzeichnungen (Russellschen ‚definite descriptions‘) Bezug nehmen. Man kann diese These vorderhand so resümieren: Selbstbewusstsein (als das, worauf wir mit dem Fürwort der ersten Person Singular Bezug nehmen) ist nicht analysierbar in Begriffen einer Sprache über Gegenstände und (objektive) Ereignisse. Daran knüpft Castañeda aber eine weitere These, die den Zusammenhang mit der kantischen Transzendentalphilosophie noch deutlicher macht: nämlich dass die Bezugnahme auf Gegenstände (Ereignisse, Sachverhalte usw.) der Welt nur *vermittels* von Selbstbewusstsein möglich ist. Und so formuliert, bekommt seine These sofort einen idealistischen touch. Ich will versuchen, sie Ihnen so knapp und doch so deutlich wie möglich vorzustellen.

Zunächst müssen wir die wesentlichen Gründe für die erste These in den Blick bringen. Noch einmal: es ist eine These über die Irreduzibilität von Selbstbewusstsein auf (wahrnehmungsvermitteltes) Gegenstandsbewusstsein. Castañeda hat sie zuerst 1966<sup>1</sup> vorgestellt und begründet: in einem bahnbrechenden Aufsatz mit dem schlichten Titel: „He’: A Study in the Logic of Self-Consciousness“ (zit.: „He’“). Die Argumentation dieses kurzen Textes ist außerordentlich subtil. Aber ohne alle Schritte durch Formeln an der Tafel (oder Folien) zu kontrollieren (und sich entsprechend Zeit zu nehmen), können wir sie nicht genau ermessen. Ich gebe darum nur einige Resultate.

---

<sup>1</sup> In: *Ratio* 8, Oxford 1966, 130-157.

Castañedas Aufsatz heißt „He“. Spricht er also wirklich vom Selbstbewusstsein? Allerdings. Nur nimmt er, statt geradewegs vom Pronomen der ersten Person Singular auszugehen, den Umweg über das der dritten Person (singularis). Und von ihm will er wissen, wie wir vermittels seiner anderen Personen Selbstbewusstsein zusprechen können. Dass hier überhaupt ein Problem liegt, ist Ihnen vielleicht nicht sofort einsichtig. Man kann sich davon aber rasch überzeugen, wenn man einige Beispielsätze untersucht. Diese Sätze formulieren grundsätzlich intentionale Aussagen, also solche mit einem Hauptsatz, der einer Person ein Intentionsverb zulegt, z. B. : ‚Paul glaubt (denkt, weiß, hofft, wünscht usw.)‘. Der davon abhängige Nebensatz drückt dann den ‚propositionalen Gehalt‘ des Intendierten (Geglaubten, Gewussten, Gehofften, Gewünschten usw.) aus. Also etwa: ‚Paul glaubt, dass Marie glücklich ist‘. Hier haben wir einen so genannten Glauben *de dicto*. Paul glaubt nämlich das *dictum* (das Gesagte, die Proposition) ‚dass Marie glücklich ist‘. Aber Paul kann auch einen Sach-Glauben haben, also (wie die Scholastiker das nannten) einen Glauben *de re*. Dann hat er eben eine Überzeugung *über etwas* (oder *jemanden*), wie sie sich etwa in dem Satz artikuliert: ‚Paul glaubt *von Marie*, dass sie glücklich ist‘. Zwischen beiden Formulierungen ist ein großer Unterschied. Das zeigt sich, wenn man Implikationen und Wahrheitsbedingungen der beiden Formulierungen (*de dicto* versus *de re*) untersucht. So kann Paul glauben, Marie sei glücklich, ohne dass Marie überhaupt existiert (er kann jemand anders gemeint und sich im Namen vertan haben; er kann auch einfach halluzinieren, träumen, aufs Hörensagen urteilen oder was weiß ich). Dagegen ist Pauls Glauben *über Marie*, wonach sie glücklich sei, ein solcher, der Maries Existenz impliziert. Denn ‚über Marie‘ steht vor dem Nebensatz, der den (auch falsch sein könnenden) Glaubensinhalt spezifiziert. Wenn Paul etwas *von jemandem* (hier: von Marie) glaubt, so muss die Existenz der Person, wel-

cher etwas zugeschrieben wird, unabhängig vom Inhalt der Zuschreibung gesichert sein.

Noch anders ist die Situation, wenn wir Marie dasselbe über sich selbst glauben lassen. Dabei haben wir wieder mehrere Möglichkeiten. Sie lassen sich so formulieren:

(a) Marie glaubt, dass Marie glücklich ist.

Oder:

(b) Marie glaubt von Marie, dass sie glücklich ist.

Oder:

(c) Marie glaubt, dass sie selbst glücklich sei.

(a) ist wieder ein Beispiel für Glauben *de dicto*, (b) wieder ein Beispiel für Glauben *de re*. Und (c) hat David Lewis vorgeschlagen, als Glauben *de se* zu fassen.<sup>2</sup> Mit (a) referiert Marie auf das dictum ‚dass Marie glücklich ist‘. Für *uns* ist damit klar, dass sie auf sich selbst referiert. Aber für Marie selbst muss das nicht evident sein: Sie kann den Nebensatz für wahr halten, ohne zu wissen, dass von ihr selbst die Rede ist. (Sie kann eine andere Marie meinen; oder sie kann als Waisenkind bei Pflegeeltern aufgewachsen sein, die sie ‚Berta‘ nannten, oder was weiß ich.) Zum Selbstbewusstsein ist aber mehr erforderlich als dies, dass eine(r) auf sich selbst Bezug nimmt. Das tun selbstregulative Strukturen wie rekursive Sätze oder lebende Organismen oder sich beim Filmen (etwa über einen Spiegel) filmende Kameras schließlich auch, ohne dass wir ihnen darum Selbstbewusstsein zusprechen. Beim Selbstbewusstsein muss das Bewusstsein also nicht nur denselben Bezugs-

---

<sup>2</sup> „Attitudes De Dicto and De Se“, in: D. L., *Philosophical Papers*, Volume I, Oxford University Press 1983, 133-159.

gegenstand haben, der es selbst ist, sondern es muss dies auch *als denselben* Bezugsgegenstand wissen.

Aus diesem Grunde reicht auch die Formulierung (b) *de re nicht aus*, um Marie ein Selbstbewusstsein zuzusprechen. Zwar hat sie, wenn sie *von Marie* etwas glaubt, ein Bewusstsein *von sich*. Aber das ist nur *uns* bekannt. Sie selbst muss nicht wissen, dass der Gegenstand ihres Glaubens sie selbst ist – jedenfalls gibt es dazu logisch keine Notwendigkeit. Es mag sogar so sein, dass Marie die Person, über die sie spricht, richtig identifiziert hat und auch meint, dass das ‚sie‘ im Nebensatz für (die so identifizierte) Marie steht. Dennoch bleibt es logisch möglich, dass sie sich mit dieser Person nicht *selbst* meint, also mit ihrer Bezugnahme kein wirkliches (für sie selbst bestehendes) *Selbstbewusstsein* ausdrückt. So meint Narziss von Narziss, dass er wunderschön sei – und er vertut sich auch weder in der Identifikation des Gegenstandes seiner Verliebtheit noch in der Zuordnung des Pronomens im Nebensatz. Dennoch glaubt er (jedenfalls zunächst) nicht, dass er *von sich selbst* spricht, wenn er entzückt ausruft: ‚wie schön du bist!‘ Ihm fehlt also bei der *Selbstreferenz* das *Selbstbewusstsein* (also ein Bewusstsein von sich *als* von sich). Genau dieses explizite Selbstbewusstsein drückt aber die (c)-Formulierung aus: Wenn Marie (oder Narziss) ‚von sich glauben‘, dass sie soundso sind (oder, in einer Formulierung *de dicto*: glauben, ‚dass sie selbst schön sind‘), dann glauben sie etwas von sich *als* von sich. Denn in dieser Wendung stehen sie gar nicht in der logischen Möglichkeit, sich im Bezugsgegenstand ihres Glaubens zu vertun. Sie haben sich nämlich direkt auf das logische Subjekt der Selbstzuschreibung bezogen, ganz egal, wie dies in einer objektivierenden Beschreibung näherhin spezifiziert wird. Also auch wenn sie nicht wissen, wer oder was das ist, worauf sie sich beziehen oder ob das Geglaubte auch wahr ist, auch dann haben sie mit *car-*

tesianischer Gewissheit einen Glauben über sich selbst. Dieses ‚er/sie sich selbst‘ der direkten epistemischen Selbstzuschreibung symbolisiert Castañeda durch einen Asterisk: ‚er/sie\*‘. (Er spricht von dem mit einem Sternchen indizierten Pronomen, um es vom gewöhnlichen Indexwort ohne garantiertes Selbstbewusstsein zu unterscheiden, als von einem ‚Quasi-Indikator‘; aber das muss uns jetzt nicht interessieren.)

Hinter dieser etwas mühsamen linguistischen Übung verbirgt sich eine kardinale Entdeckung mit weitreichenden Konsequenzen. Wittgenstein hatte sie früher als den Subjektgebrauch von ‚ich‘ (bzw. ‚er\*‘) bezeichnet.<sup>3</sup> Beispiele dafür sind ‚Ich habe Zahnweh‘ oder ‚Ich fühle mich müde‘. Das Wesentliche an solchen Gedanken ist, dass sie ihren Bezugsgegenstand (mich selbst) notwendig direkt erreichen, egal wer ich bin. Ich mag ein Hirn sein, das in der Experimentation eines sadistischen Neurobiologen trostlos in einer Nährflüssigkeit schwimmt (und dessen Nervenenden mit einem Super-Computer verbunden sind). Auch dann, wenn ich gar nichts mehr über mein körperliches Aussehen weiß (weil ich mich in dem Zustand nie gesehen habe noch je zu sehen oder anderswie wahrzunehmen bekomme), – auch dann kann ich immer noch den Gedanken fassen: ‚Mein Gott, so etwas möchte ich nie wieder erleben (I won’t let this happen again)!‘<sup>4</sup> Und das ‚ich‘ in diesem Gedanken verweist todsicher auf *mich*, auch wenn ich keine Wahrnehmung und keine Beschreibung von mir habe. Anders verhält es sich beim Objekt-Gebrauch von ‚ich‘. Er kommt in Sätzen wie ‚Ich habe eine Beule an der Stirn‘ oder ‚Ich stehe neben dem Tisch‘. In beiden Fällen referiert ‚ich‘ nicht auf das logische Subjekt des Gedankens, sondern auf

<sup>3</sup> *Das Blaue Buch*, in: Suhrkamp-Werkausgabe, Frankfurt/M. 1984, Bd. 5, 106 (im Kontext). Dazu Sydney Shoemaker, „Self-reference and self-awareness“, in: ders., *Identity, Cause, and Mind. Philosophical Essays*, Cambridge University Press 1984, 6-18, hier: 7 ff.

<sup>4</sup> Vgl. Elizabeth Anscombe, „The First Person“, in: *The Collected Philosophical Papers of G. E. M. Anscombe*, Vol. II: Metaphysics and the Philosophy of Mind, Oxford 1981, 21-36, hier: 31,2).

meinen Körper (oder einen Teil desselben oder auf dessen Lokalisierung usw.) – aber *dass* das so ist, muss ich nicht wissen. Die Beule muss ich nicht spüren, und den Tisch kann ich nicht gesehen haben. Diese Informationen gewinne ich nur über die Wahrnehmung. Und im Falle des direkten epistemischen Selbstbezugs gibt es nichts, dessen Wahrnehmung mir entgehen oder in dessen Identifikation ich mich vertun könnte. In Castañedas Worten:

The crucial weapon in the refutation of claim (iii) is the fact, which philosophers (especially Hume and Kant) have known all along, that there is no object of experience that one could perceive as the self that is doing the perceiving. [Claim (iii) hatte behauptet, daß sich der emphatische Selbstbezug<sup>5</sup> von ‚er/sie\*‘ auf den des gewöhnlichen Pronomens ‚er/sie‘ reduzieren lasse.]<sup>6</sup> However it is that one identifies / an object of experience as oneself, whenever one does, one identifies an object in experience with a thing which is not part of the experience, and this thing is the one to which the person in question will refer to by ‚I‘ (or its translation in other languages), and another person will refer to by ‚he\*‘, or ‚he himself‘ in the special Self-ascription-Juse („He““, 142 f.).

Kurz: Emphatisches Selbstbewusstsein ist nicht wahrnehmungsvermittelt. Und da uns Gegenstände nur vermittelt durch Wahrnehmungen gegeben werden, dürfen wir sagen: Die cartesianische Gewissheit des emphatischen Selbstbewusstseins gründet in seiner radikalen Ungegenständlichkeit. Wir haben uns im Selbstbewusstsein als Subjekte, nicht als Gegenstände. Wie sollten wir auch – ohne zusätzliche Evidenz – von einem Gegenstande außer uns lernen, dass *wir* das sind?

Klar: daraus allein folgt noch nicht, dass wir nicht aus einer anderen, eben einer Fremd- oder Wahrnehmungs-Perspektive, auch Gegenstände wären.

<sup>5</sup> So nennt ihn Roderick Chisholm in *The First Person. An Essay on Reference and Intentionality*, Brighton/Sussex: The Harvester Press, 1981, 24. ‚Emphatisch reflexiv‘ sind die ‚he, himself locutions‘ (17 ff.).

<sup>6</sup> Genauer besagt der „claim“: „if one asserts truly, X believes (knows, thinks) that he is  $\phi$ “ using ‚he‘ demonstratively to refer to the person X and the person X is in a position to take one’s use of ‚he‘ as a demonstrative usw., then X believes (knows, thinks) that he\* is  $\phi$ “ („He““, 139 u.).

Aber wir haben hier ein echtes erkenntnistheoretisches Problem. Ein eingefleischter Verifikationist kann sich auf den Standpunkt stellen, nur solche Sätze seien sinnvoll, die dasjenige, was sie als bestehend ausgeben, auch einem möglichen Bewusstsein zugänglich machen können (das war z. B. Ayers empiristisches Sinnkriterium). Ein solcher wird dann aber auch ontologische Konsequenzen aus der epistemischen Asymmetrie von wahrnehmungsfreier, also kriterienloser Selbstzuschreibung und wahrnehmungsgestützter Zuschreibung ziehen. Er wird sagen, der Vorrang von ‚ich‘ (bzw. ‚er/sie\*‘) vor dem durch Eigennamen oder Demonstrativa oder ‚bestimmte Beschreibungen‘ Identifizierten sei nicht nur eine ‚referentielle Priorität‘. Diese Priorität sei damit auch erkenntnistheoretisch, ja ontologisch (gelte also nicht nur für unser Erkennen der Gegenstände, sondern für diese gleich mit). Und genau das sagt Castañeda in der Tat („He“, 144 f.).

Zunächst ein paar Worte zu *referentiellen Priorität* von ‚ich‘ (bzw. ‚er/sie‘: Demonstrative, namentliche oder kennzeichnungs-vermittelte Bezugnahme kann bekanntlich fehlschlagen („dies da“ kann verschiedenes meinen, mehrere Gegenstände könnten denselben Namen tragen oder dieselbe bestimmte Beschreibung erfüllen). Unmöglich aber kann ‚ich‘ (korrekt gebraucht) zu einer solchen ‚infelicity‘ führen. ‚Ich‘ ist ferner *ontologisch* vor Namen, allen anderen Zeigewörtern und Kennzeichnungen durch die Existenz-Garantie des Referenten ausgezeichnet. Sie sind vor der Möglichkeit geschützt, dass der Bezugsgegenstand vielleicht nicht der ist, den wir meine, oder – wie im Falle von optischen Täuschungen, Halluzinationen oder Einbildung *nicht* ist (während es sehr wohl möglich ist, dass Gegenstände, auf die Namen oder Kennzeichnungen Bezug zu nehmen scheinen, nicht existieren). Und schließlich ist ‚ich‘ *erkenntnistheoretisch* vor Namen, definiten Beschreibungen und Zeigewörtern ausgezeichnet; seine Unanalysierbarkeit



durch Namen, Kennzeichnungen, aber auch alle anderen Indikatoren (selbst ‚hier‘ und ‚jetzt‘) ist sogar ein „fundamentales Faktum, das der Idee des transzendentalen Subjekts zugrunde liegt“.<sup>7</sup> Vom Subjekt (und seinem Selbstbewusstsein) her wird die indikatoren- und begriffsvermittelte Kenntnis aufgebaut, die Heidegger das In-der-Welt-Sein nennt. Nicht aber ist ‚ich‘ aus der Semantik des In-der-Welt-Seins, und auch nicht „durch die Demonstrativpronomen der dritten Person“ zu analysieren.

Sage ich z. B. ‚Dies da ist blau‘, so bin ich – später – nicht unmittelbar berechtigt zu schließen ‚dies da *war* blau‘; sondern ich muss das Demonstrativpronomen z. B. durch eine bestimmte Beschreibung ersetzen und nun etwa sagen: ‚Das Stück Gletschereis, in das ich damals und dort eine Eis-schraube bohrte, war blau‘. D. h., Demonstrativa sind für den Benutzer grundsätzlich eliminierbar. Dagegen habe ich genau dieses Recht der Ersetzung des Präsens durch die Vergangenheitsform bei Sätzen wie ‚Ich bin  $\phi$ ‘, auf die ich später durch ‚Ich war  $\phi$ ‘ zurückkomme. Hier muss ich ‚ich‘ nicht durch eine Beschreibung ersetzen, ja, ich könnte das gar nicht, ohne in die Beschreibung wieder ein ‚ich‘ einzufügen und dessen Unersetzbarkeit in direkter Rede so erneut zu bewähren (‚ich bin derjenige, der zum Zeitpunkt  $t_1$  das und das getan hat oder in dem und dem Zustand gewesen ist‘). Zwischen ‚ich bin  $\phi$ ‘ und ‚ich war  $\phi$ ‘ waltet eine epistemisch zugängliche garantierte Identität über Zeit (es gibt Kontinuität von Selbstbewusstsein). An ihr hängt die Wiederidentifizierbarkeit von früher wahrgenommenen (und demonstrativ aufgezeigten) Objekten.<sup>8</sup> Anders: Dass eine demonstrative

---

<sup>7</sup> SE, 167,1.

<sup>8</sup> Diese These hat Castañeda später aufgegeben, um sich Russells Idee des Kurzzeit-Selbst anzuschließen. Es drückt sich aus im dem Präfix „Ich hier jetzt“ und geht mit dem Jetzt-Punkt unter: „It is crucial, however, to hold on to the fact that remembering having done some action A, or having had an experience E, is not to remember a self or an I, in itself, so to speak, nakedly, doing A or experiencing E. Past selves – as Hume claimed for present selves – are nowhere to be found. What one remembers is a situation from a certain perspective, even if it is the merely intellectual perspective of a hierarchi-

Bezugnahme in der Erinnerung (in Abwesenheit des zeigbaren Objekts) funktioniert, hängt am seidenen Faden fortwährender epistemischer Selbstreferenz: ‚Ich‘ gibt dem ‚dies‘ (auch dem erinnerten) allererst seine Referenz und seinen Sinn.

Ich möchte *einen* Aspekt der These vom erkenntnistheoretischen Vorrang von ‚ich‘ noch etwas genauer unter die Lupe nehmen. Es ist dieser: Nur wer schon über den Subjekt-Gebrauch von ‚ich‘ direkt mit sich bekannt war, kann demonstrativ, namentlich oder über Kennzeichnungen Angezieltes gegebenenfalls erfolgreich auf sich selbst beziehen. ‚Erkenntnistheoretisch‘ darf dieser Vorrang darum heißen, weil jede andere bewusste Bezugnahme auf Gegenstände an einer doppelten Voraussetzung hängt: 1. sich (allem Gegenstandsbezug zuvor) ungegenständlich selbst erfassen zu können und 2. die Kenntnis von *Objekten* (auf die ich gegebenenfalls durch ‚dies da‘ oder Eigennamen oder Beschreibungen referieren kann) eben als durch die *Kenntnis* von Objekten vermittelt zu begreifen.<sup>9</sup> Und erst diese Möglichkeit macht verständlich, wie ein selbstbewusstes Subjekt ‚hier‘- oder ‚jetzt‘-Gedanken haben kann. Denn wie könnte ‚hier‘ oder ‚jetzt‘ oder auch ‚dies‘ anders eingeführt werden als über Wendungen, die schon ein ‚ich‘ enthalten (‚der Ort, an dem *ich* stehe‘; ‚die Zeit, zu der *ich* diesen Gedanken fasse‘, ‚diese Finger, mit denen *ich* die Tasten meine Computers drücke‘ usw.)?<sup>10</sup> Dies ist, wie wir eben hörten, nach Castañedas Ansicht „ein fundamentales Faktum, das der Idee des transzendentalen Subjekts zugrunde-

---

cally organized set of beliefs“ (*The Self and the I-Guises, Empirical and Transcendental*, in: *Theorie der Subjektivität*, hg. von K. Cramer, H. F. Fulda, R.-P. Horstmann und U. Pothast, Frankfurt/M. 1989, 105-140, hier: 137).

<sup>9</sup> „In order to keep knowledge or belief, or in order merely to rethink. of the objects originally apprehended by means of demonstratives one must reformulate one’s knowledge or belief, or thought, of those objects. One must replace each purely demonstrative reference by a reference in terms of [...] the ‚I“ („He“, 145,2).

<sup>10</sup> H.-N. Castañeda, „Indikatoren und Quasi-Indikatoren“, in: *Sprache und Erfahrung. Texte zu einer neuen Ontologie*, Frankfurt/M. 1982, 166 f.

liegt“,<sup>11</sup> Damit ist der Gedanke gemeint, dass jede mittelbare Bezugnahme auf anderes vermittelt ist durch unmittelbares Selbstbewusstsein. ‚Transzendental‘ aber darf diese Vermitteltheit heißen gemäß einer schwachen Definition von ‚transzendental‘: was nicht geradehin auf Gegenstände sich bezieht, sondern auf Voraussetzungen der Bezugnahme auf Gegenstände, *sofern diese eine seiner selbst unmittelbar gewisses Subjekt zur Voraussetzung hat.*

---

<sup>11</sup> L. c., 167,1.

## Neueinsatz

Um zu resümieren: Zwei Nachweise hat Castañedas Aufsatz „He“ zu erbringen versucht: 1. den, dass es einen irreduziblen ‚er (sich selbst)‘-Gebrauch gibt, durch den wir anderen Subjekten Selbstbewusstsein zuschreiben; 2. den, dass ein solcher ‚er (sich selbst)‘-Gebrauch genau dann gerechtfertigt ist, wenn ein Hörer ausdrücken will, was ein Sprecher behauptet, wenn er ein uneliminierbares ‚ich‘ in indirekter Rede verwendet, das heißt, wenn es sich (selbst) egologisches Selbstbewusstsein zuschreibt. Anders: mit ‚er/sie\*‘-Wendungen bringe ich im Blick auf andere Subjekte und in indirekter Rede die Überzeugung zum Ausdruck, die diese sich selbst (implizit oder explizit) in direkter Rede zuschreiben: nämlich selbstbewusste Subjekte von Intentionen (und anderen mentalen Zuständen oder Dispositionen) zu sein.

Was „He“ sicher nicht zu leisten vorhat, ist die Relativierung der ‚ich‘-Perspektive als solcher. Sie wird von dem Quasi-Indikator ‚er/sie\*‘ ja nur aus der Fremdperspektive auf mentale Selbstzuschreibungen aufgegriffen und geht allem Pronominal-Gebrauch, ja jeder Verwendung singularer Termini als deren transzendente Bedingung voraus. So begegnet Castañeda (in jüngeren Publikationen) dem kantischen ‚Ich denke‘ wieder, welches alle meine Vorstellungen muss begleiten können und die ganze Erscheinungswelt in einen Kontext analog zu dem der indirekten Rede verwandelt (‚Ich denke, dass *a* und dass *b* und dass *c* und dass ... *n*‘).<sup>12</sup> So ist das Ich nicht

---

<sup>12</sup> *The Self and the I-Guises*, l. c., 118. Dort werden zwei Prinzipien aufgestellt. Das erste besagt: „Every utterance expressing a thought (whether believed or not by the speaker) is in principle and at bottom subordinated to an implicit *I think that* (or, even, to an *I say that*)“; das zweite sei eine Konsequenz des ersten und besage: „Every statement lies implicitly, or explicitly, in *oratio obliqua* (indirect speech), and the only true or genuine *oratio recta* (direct speech) is the unspoken *I think* – this is Kant, but I wish to add – or, rather, *I think here now*.“

nur am Ursprung des Systems der deiktischen Ausdrücke (deren Bedeutungen immer von der jeweiligen Verwendungs-Situation oder, wie Castañeda sagt, ‚Hier-jetzt-Perspektive‘ des Sprechers abhängen), sondern ‚am Nullpunkt seines ganzen referentiellen Koordinatensystems‘.<sup>13</sup> Insofern unterschreibt auch Castañeda – freilich mit bedeutenden Vorbehalten, die ich hier nicht entwickeln kann – Fichtes erkenntnistheoretische These, jedes Bewusstsein *von* etwas (also: jedes referentielle Bewusstsein) hänge vom vorgängigen Selbstbewusstsein ab, und eröffnet Durchblicke auf den un vermuteten Einstand des idealistischen Grundgedankens in der jüngsten analytischen Philosophie.

*Exkurs über die Irreduzibilität von ‚ich‘:* Gegen die These von der Irreduzibilität des Pronomens der 1. Person Singular hat es immer schon Einwände gegeben. In neuerer Zeit hat z. B. Ernst Tugendhat geltend gemacht, dass ‚ich‘ nur lernen könne, wer gleich mitlerne, dass der Gegenstand von ‚ich‘ aus der Fremdperspektive mit ‚er/sie‘ oder auch ‚du‘ bezeichnet werde. Ein noch stärkerer Ausdruck für diese These ist die Formulierung, zur regelmäßigen Beherrschung des Pronomens ‚ich‘ gehöre Kenntnis, dass andere den Bezugsgegenstand des Pronomens durch ‚er‘ oder ‚sie‘ „aufgreifen“ können. So ist die ‚er/sie‘- in die ‚ich‘-Perspektive logisch ‚eingeschlossen‘. Danach muss ich mich zwar im einsamen epistemischen Selbstbezug nicht als raumzeitliches (und mithin wahrnehmbares) Objekt identifizieren, wie andere das in Bezug auf mich tun, aber a priori wissen, dass ich für sie (und für mich selbst als Beobachter) ein solcher neutral wahrnehmbarer Gegenstand bin. Anders gesagt: Selbstbewusstsein „impliziert“ die Kenntnis sozialer Regeln, die die Konvertierbarkeit der Sprecherstandpunkte besorgen,

---

<sup>13</sup> L. c., Kap. 14, 132-134.

und ist nur von ihnen her aufzuklären.<sup>14</sup> Diese Kenntnis charakterisiert Tugendhat durch die Prädikate ‚vorgängig‘, ‚allgemein‘ und ‚notwendig‘. Sie ist also Kenntnis-a-priori.<sup>15</sup>

Aber Castañeda hatte uns längst vor Tugendhats hilflosem Versuch zeigen können, dass diese Erklärung auf einem Zirkel beruht. Erstens ist ‚ich‘ kein Pronomen (da es für kein Nomen steht), wie wir gesehen haben: ‚Ich denke‘, von Ludwig Wittgenstein geäußert, bedeutet nicht ‚Ludwig Wittgenstein denkt‘ (vgl. „Das blaue Buch“, WA, 107: „Das Wort ‚ich‘ bedeutet [überhaupt] nicht dasselbe wie ‚L. W.‘, selbst wenn ich L. W. bin, noch bedeutet es dasselbe wie der Ausdruck ‚die Person, die jetzt spricht“). Zweitens steht ‚ich‘ nicht für einen auch unabhängig identifizierbaren Einzelgegenstand (sonst müsste die *de-re*-Erklärung auch für ‚ich‘ zutreffend sein). Das wäre der Fall, wenn ich mich in ‚ich‘-Sätzen erst dann auf mich beziehen könnte, wenn ich mich zuvor mit fremden Augen hätte betrachten, also für jemand hätte halten müssen, auf den ich selbst – wie allen anderen – mit ‚der da‘ oder ‚du‘ oder ‚er‘ oder Bezug nehmen. Der offenkundige Zirkel in dieser Beschreibung besteht darin, dass ich mit dem Referenten von ‚ich‘ vor aller gegenständlichen Referenten bekannt sein musste. Denn nichts an einem noch so vollständig repräsentierten Körper verrät mir, dass *ich* dieser

<sup>14</sup> Ernst Tugendhat, *Selbstbewußtsein und Selbstbestimmung. Sprachanalytische Interpretationen*, Frankfurt a. Main 1979 [hinfert zit.: *SuS*], 87. Tugendhat sagt: „[...] diese Möglichkeit der Intersubjektivität ist ihrerseits eine notwendige, weil sie aus der Bedeutung von ‚ich‘ folgt“ (88). Vorher hatte er behauptet, „daß es von vornherein und allgemein zur Verwendungsweise von ‚ich‘ gehört, daß er, der es verwendet, es so verwendet, daß er weiß, daß ein anderer seine Rede so aufnehmen kann, daß er auf dieselbe Person, auf die er mit ‚ich‘ bezugnimmt, mit ‚er‘ bezugnimmt“ (84). – Dass seine Selbstbewußtseins-Theorie unhaltbar ist, hat Tugendhat inzwischen in Reaktion auf Dieter Henrichs Einwände („Noch einmal in Zirkeln. Eine Kritik von Ernst Tugendhats semantischer Erklärung von Selbstbewußtsein“, in: Clemens Bellut & Ulrich Müller-Schöll [Hgg.], *Mensch und Moderne. Beiträge zur philosophischen Anthropologie und Gesellschaftskritik* [= FS Helmut Fahrenbach], Würzburg 1989, 93-132) zugegeben.

<sup>15</sup> Und zwar offenbar ein Fall applizierter Regel-Kennntnis. Neuere analytische Arbeiten zur Bewusstseins-Theorie haben aber überzeugend dargetan, dass epistemisches Selbstbewusstsein nicht als Fall erfolgreicher Sprachkonditionierung – etwa aus der Beherrschung der Verwendungsregeln Pronomens der 1. Person singularis verständlich gemacht werden kann. Diese Regeln kann vielmehr nur lernen, wer mit dem Referenten von ‚ich‘ oder ‚mein‘ durch eine direkte Kenntnis vertraut ist.

Körper bin – es sei denn, das war mir schon vorher bekannt und ich kann *deshalb* die Beschreibung oder Demonstration auf *mich* beziehen. Daraus folgt, dass ‚ich‘ nicht ‚er‘ impliziert – jedenfalls nicht in dem Sinne, dass ich mich erst dann erfolgreich mit ‚ich‘ auf mich beziehen könnte, wenn ich zuvor den mit ‚er‘ identifizierten Körper ausgemacht habe, mit dessen Fingern ich taste, durch dessen Augen ich sehe, mit dessen Ohren ich höre usw.

Niemand hat das klarer gesehen als Sydney Shoemaker.

Ich habe darauf hingewiesen, daß es zum Wesen bestimmter mentaler Zustände gehört, ein solches [de-se-]Wissen von sich selbst hervorzubringen. Was dies von anderen Fällen unterscheidet, in denen die Tatsache, daß so-und-so der Fall ist, das Wissen, daß so-und-so der Fall ist, hervorbringt, ist die Unmittelbarkeit der Verbindung. Sie ist nicht durch das Hervorbringen anderer Überzeugungen vermittelt, von denen die abschließende Überzeugung abgeleitet ist (so wie ein historisches Ereignis eine seine Existenz betreffende Überzeugung dadurch hervorbringt, daß es in der Gegenwart eine Evidenz für seine Existenz hervorbringt), und sie ist nicht durch irgendetwas, was einem Sinneseindruck analog wäre, vermittelt. Der letzte Punkt ist besonders wichtig. Er zeigt, daß das fragliche Wissen radikal verschieden von einem auf Wahrnehmung beruhenden Wissen ist. Der Grund dafür, daß man sich selbst im Selbstbewußtsein nicht „als ein Objekt“ gegeben ist, liegt darin, daß Selbstbewußtsein kein auf Wahrnehmung beruhendes Bewußtsein ist, d. h. nicht die Sorte von Bewußtsein, in der Objekte gegeben sind. Es handelt sich um ein Bewußtsein von Tatsachen, das nicht durch ein Bewußtsein von Objekten vermittelt ist. Aber man sollte festhalten, daß man in der Erklärung des Wissens von sich selbst nicht sehr weit kommt, wenn man sich seiner selbst als eines Objektes in solchen Fällen bewußt wäre (wie man sich ja tatsächlich seiner selbst als eines Objektes bewußt ist, wenn man sich im Spiegel sieht). Denn das Bewußtsein, daß das gegebene Objekt  $\phi$  war, würde einem nicht mitteilen, daß man selbst  $\phi$  war, es sei denn, man hätte das Objekt als sich selbst identifiziert, und dies könnte man nur dann, wenn man schon irgendein Wissen von sich selbst hat, nämlich das Wissen, daß man der alleinige Besitzer irgendeiner Reihe von Eigenschaften des gegebenen Objekts ist, das man herausgriff, um zu zeigen, daß man selbst es ist. Wissen von sich selbst, das auf Wahrnehmung beruht, setzt ein Wissen von sich selbst voraus, das nicht auf Wahrnehmung beruht, also kann nicht das gesamte Wissen von sich selbst auf Wahrnehmung beruhen. Macht man sich diese Tatsachen klar, dann sollte dies genügen, um die Vorstellung zu zerstreuen, die Natur des Wissens von sich selbst unterstütze die Cartesianische Auffassung, das Selbst sei eine beson-

dere Sorte von Objekten, oder die Hume'sche Auffassung, das Selbst sei überhaupt kein irgendwie geartetes Objekt (Sydney Shoemaker, „A Materialist's Account“, in: ders. & Richard Swinburne, *Personal Identity*, Oxford: Basil Blackwell, 1984, 104 f.).

Wir haben eben gehört: Selbst Wittgenstein glaubt nicht, der Ausdruck ‚ich‘ sei ein Pronomen (also ein Ausdruck, der einen Eigennamen aufgreift). Die Sätze ‚Ludwig Wittgenstein spricht‘ und ‚ich spreche‘ (von Ludwig Wittgenstein gebraucht) haben nicht dieselbe Bedeutung. Der Vorschlag, ‚ich‘ meine (im Munde der äussernden Person) ‚der/die gerade Sprechende‘, stammt von Hans Reichenbach.<sup>16</sup> David Kaplan nennt sie „ingenious“. Er benützt für die Übersetzung von ‚ich‘ in ‚der/die gerade Sprechende‘ das erfundene Demonstrativ-Pronomen ‚dthat‘.<sup>17</sup> Aber Castañeda hatte Reichenbachs reduktionistische Position in *Indicators and Quasi-Indicators* durch ein einziges Gegenbeispiel aus dem Sattel heben können: Offenbar ist ja die Äußerung ‚Ich äußere (gerade) nichts‘ nicht übersetzbar in die ‚Die Person, die dieses Zeichenereignis äußert, äußert gerade nichts‘. Ersteres kann, wörtlich genommen, zufällig falsch sein (man beginge einen *performativen* Widerspruch), letzteres aber ist *logisch* selbstwidersprüchlich.

Das meint:

Wenn eine Person diesen Satz äußert, so falsifiziert sie die entsprechende Behauptung, doch die Behauptung *hätte*, sogar in einem solchen Fall, wahr sein können. Dagegen sind alle Behauptungen, die ausgedrückt werden können durch ‚Die Person, die dieses Zeichenereignis [token] äußert, äußert nichts‘ selbst-widersprüchlich: sogar wenn sie niemand behauptet, können sie nicht wahr sein.<sup>18</sup>

<sup>16</sup> *Elements of Symbolic Logic*, New York 1947, 284. Dazu Kaplan, l. c., 519 ff.

<sup>17</sup> „Demonstratives“, in: *Themes from Kaplan*, ed. by Joseph Almog, John Perry, and Howard Wettstein, New York-Oxford: Oxford University Press, 1989, 481-614, hier: 519-521.

<sup>18</sup> Castañeda, *Sprache und Erfahrung*, l. c., 166. Linne Rudder Baker hat seither versucht, Castañedas Einwurf zu entkräften: „On Making and Attributing Demonstrative Reference“, in: *Synthese*, Vol. 49, 1989, 245-273, hier: 267 f. Sie tut das durch Einführung einer nicht-selbstwidersprüchlichen Lesart von ‚Die Person, die dieses Zeichenereignis äußert, äußert nichts‘. Das könnte statt ‚Es gibt jemanden, der dies Zeichenereignis äußert und auch nicht äußert‘ auch so verstanden werden: ‚Es stimmt nicht, daß der Sprecher dieses token der Sprecher dieses token ist‘. Die letztere Formulierung sei aus den gleichen bloß kontingenten Gründen falsch wie der ‚ich‘-Satz, aber nicht selbst-widersprüchlich. Denn,



Das hängt aber, meint Kaplan (der sich nicht ausdrücklich mit Castañeda auseinandersetzt), an der Lesart von ‚dthat‘-Sätzen: Ich kann ‚dthat‘ als Ersatz für ein Demonstrativ-Pronomen („demonstrative surrogate“) nehmen, und dann meint es: ‚die Person, die gerade dies Zeichenereignis äußert (und die Subjekt eines singulären Aussagesatzes sein kann)‘; oder ich kann es als „rigidifier“<sup>19</sup> (im Sinne Kripkes) auffassen, und dann haben ‚dthat‘ und ‚ich‘ nicht mehr dieselben semantischen Eigenschaften, denn nun ist ‚dthat‘ ein genereller Terminus, und ‚ich‘ ist ein singulärer. ‚Dthat‘ ist dann eine Kennzeichnung von der Art ‚Quadratwurzel aus 4‘ und referiert nicht mehr direkt, ist aber doch faktisch ein ‚rigider Designator‘ in dem Sinne, dass er den einzigen Gegenstand, der in allen möglichen Welten (oder Umständen) unter diesen Begriff fällt, ‚obstinat starr‘ bezeichnet. Die Bewertung eines Satzes mit ‚dthat‘ in mehreren Umständen (möglichen Welten) führt nicht in diese Welten einen Referenten ein, sagt Kaplan, sondern gibt Anweisungen an die Hand, die uns erlauben, aus irgendeiner möglichen Welt wieder nachhause (in die unsere) heimzukehren und den Einzelgegenstand herauszupicken, der der Beschreibung ‚der dieses Zeichenereignis äußert‘ befriedigt.<sup>20</sup> Wenn also – nach der ersten (Kaplan’schen) Lesart – ‚ich‘ für ‚syn-

---

gestützt auf Kaplans Unterscheidung zwischen Äußerungs-Kontexten und Bewertungs-Umständen, könnte man geltend machen, dass hienieden ‚Ich sage gerade nichts‘ allemal falsch ist, dass es aber Umstände (mögliche Welten) geben könnte, wo (der Referent von) ‚ich‘ wirklich nichts sagt. Damit widerspricht man dann freilich Kaplans aktualistischer Maxime und behandelt ‚ich‘, synonym mit ‚der dieses Zeichenereignis äußert‘, als starre Designatoren quer Beet über alle möglichen Welten. ‚Ich‘ referierte dann auf dieselbe Person in allen denkbaren Umständen – worunter auch solche, in denen sie tatsächlich nichts sagte.

Eros Corazza diskutiert Für und Wider des Einwandes ausführlich in *Dire ‚je‘*, 17/18 ff. Vgl. auch Kaplans *Afterthoughts*, l. c., 579 ff. („Two interpretations of the syntax and semantics of ‚dthat‘“).

<sup>19</sup> Kaplan, *Afterthoughts*, 580.

<sup>20</sup> L. c.: „In this case the proposition would not carry the individual itself into a possible world but rather would carry instruction to run back home and get the individual who there satisfies certain specifications. The complete dthat-term would then be a rigid description which induces a complex ‚representation‘ of the referent into the content; it would not be directly referential. The operator ‚dthat‘ might still be regarded als *involving* direct reference, though its own referent would not be the individual denoted by the complete dthat-term, but, like that of all operators, would be an abstract, higher-order functional type.“

onym mit *dthat*' erklärt wird, dann kann es nicht als ‚rigid designator‘, sondern muss (um seine Direkt-Referenz wiederzugewinnen) als Ersatz für ein Demonstrativum aufgefasst werden. Darin gründet die Synonymie beider: „For two words or phrases to be synonyms, they must have the same content in every context. This is because indexicals are directly referential, and the compound phrases which can be used to give their reference (‘the person who is speaking’, ‘the individual being demonstrated’, etc.) are not.“<sup>21</sup>

Damit ist nun zwar die Synonymie von ‚*dthat*‘ als starrer Designator, nicht aber die von ‚*dthat*‘ als Demonstrativ-Surrogat mit ‚*ich*‘ geleugnet. Castañeda These hatte aber auch das letztere geleugnet und ‚*ich*‘ für ursprünglich und nicht weiter analysierbar erklärt. Kaplan gibt übrigens zu, dass seine Deutung von ‚*ich*‘ als Synonym mit ‚*der gerade dieses Zeichen benützt*‘ „subtile Komplikationen“ in sich birgt.<sup>22</sup>

Offenbar sind diese Komplikationen nicht nur feinsinniger Natur, sondern schlicht unauflösbar. ‚*Ich*‘ kann auf keine Weise von Demonstrativa her verstanden und mit ihnen auf eine Ebene gebracht werden – darin ist Elizabeth Anscombe und Hector-Neri Castañeda nach wie vor beizupflichten. Kaplan behandelt ‚*dthat*‘ aber als Surrogat eines ‚echtes Demonstrativs‘ (und die es vervollständigende Beschreibung als Ersatz der ergänzenden Demonstration).<sup>23</sup> Und das, obwohl er im X. Kapitel („Reichenbach on Token Reflexives“) subtile Einwände gegen die Ansicht geäußert hatte, ‚*ich*‘

---

<sup>21</sup> „Demonstratives“, 521.

<sup>22</sup> L. c., 522: „The fact that this alleged synonymy is cast in the theory of utterances rather than occurrences introduces some subtle complications, which have been discussed by Reichenbach.“ Auf diese subtilen Komplikationen konzentriert sich Eros Corazza, *Dire 'je'*, 20 ff., seinerseits gestützt auf Felicia Ackerman, „Content, Character, and Nondescriptive Meaning“, im selben von J. Almog et al. edierten Sammelband *Themes from Kaplan*, in dem auch „Demonstratives“ abgedruckt ist, 5-21, hier: 8 f.

<sup>23</sup> L. c., 581.

*bedeute* dasselbe wie ‚die Person, die dieses Zeichen(vorkommnis) äußert‘. Dass beide Ausdrücke *nicht synonym* sind, zeigt Kaplan an folgenden drei Beispiel-Sätzen:<sup>24</sup>

- ( $\alpha$ ) I am the person who utters this token.
- ( $\beta$ ) If no one were to utter this token, I would not exist.
- ( $\gamma$ ) If no one were to utter this token (in S), the person who utters this token (in S) would not exist.

( $\alpha$ ) und ( $\gamma$ ) sind offenbar wahr, während ( $\beta$ ) falsch ist. ( $\beta$ ) müsste aber wahr sein, wenn „I’ means the same as ‚the person who utters this token“ im Sinne einer strengen Synonymie verstanden und dann ‚the person who utters the token‘ eingesetzt würde anstelle von ‚ich‘. ([ $\gamma$ ] wird nur angeführt, um eine sinnvolle Interpretation von [ $\beta$ ] zu liefern.) Aber, meint Kaplan, die Tatsache, dass eine (Reichenbach’sche) Synonymie zwischen ‚ich‘ und ‚die Person die dies äußert‘ ausgeschlossen werden kann, hindert nicht, dass ‚ich‘ in jeder Verwendungssituation auf die Person, die ‚ich‘ äußert, *refert*.<sup>25</sup> Und insofern könne allerdings ‚ich‘ durch ‚that‘ (mit entsprechender deskriptiver Ergänzung: nämlich ‚die Person, die gerade „ich“ äußert‘) ersetzt werden.

Tatsächlich ist der Reichenbachsche Vorschlag, ‚ich‘ und ‚die Person, die dies Token äußert‘, für synonym zu halten, regressbedroht. Er läuft darauf hinaus, ein so genanntes ‚reines Indexikal‘ (wie ‚ich‘, ‚hier‘ und ‚jetzt‘) durch ein Demonstrativ(-Surrogat) zu ersetzen. Interpretiere ich „this token“ durch die bestimmte Beschreibung „the token pointed at by a gesture accompanying the utterance of the token“, so habe ich stillschweigend „jetzt“ impliziert (denn Begleitungen sind dem Begleiteten immer gegen-

---

<sup>24</sup> L. c., 519 f.

<sup>25</sup> L. c., 520 (D2).

wärtig). Nun ist ‚jetzt‘ ein reines Indexikal; und mithin ist der Anspruch, reine Indexwörter durch Demonstrativ(-Surrogate) zu ersetzen, gescheitert, oder vielmehr: es ist eingetreten, was Castañeda so formuliert hatte: Analysen von Quasi-Indikatoren können nur durch Ausdrücke gelingen, die selber wieder (explizit oder implizit) Quasi-Indikatoren enthalten. – Nun leugnet Kaplan zwar die starke Synonymie-These und meint, ‚ich‘ *bedeute* nicht, sondern *referiere* nur auf ‚die Person, die „ich“ äußert‘. Da wir nun wissen, dass ‚ich‘ ein reines Indexwort ist und reine Indexwörter ihren Referenten allein aufgrund des ‚Character‘ (der sprachlichen Bedeutung von ‚ich‘) determinieren, taucht der Zirkel in der Definition von ‚dthat‘ (‚die Person, die gerade „ich“ äußert‘) auch bei der Referenz-Fixierung auf. So muss der Versuch, ‚ich‘ ohne Rekurs auf reine Indexwörter rein durch Demonstrativ(-Surrogate) zu analysieren, für gescheitert gelten.

Die von Reichenbach so genannte ‚Zeichen-Reflexivität‘ von ‚ich‘ besteht doch eben darin, dass es der Sprecher/die Sprecherin *selbst* sein muss, die mit diesem Wort auf sich Bezug nimmt. Die deskriptive Paraphrase von ‚dthat‘ lautet aber ‚der/die da gerade spricht‘ und kann auch auf eine fremde Person angewandt werden – mithin haben ‚ich‘ und ‚dthat‘ weder dieselbe Bedeutung noch fixieren sie notwendig denselben Referenten. Jedenfalls gilt, dass ‚ich‘ zwar ‚dthat‘ logisch impliziert (denn ‚ich‘ meint immer die gerade sprechende Person), während das umgekehrte nicht gilt. Mithin ist ‚ich‘ logisch auf ein – wie immer kennzeichnungs-unterstütztes – Demonstrativum unzurückführbar.

Ein ähnliches Argument zugunsten der Irreduzibilität findet sich bei John Perry.<sup>26</sup> Er meint, ‚ich‘ könne nicht ohne Einbuße seiner „explanatory force“ durch einen anderen Ausdruck mit gleicher Referenz ersetzt werden. Sage ich: ‚Ich glaube, da richte ich eine ganz schöne Schweinerei an‘, so meine ich nicht, dass jemand, der gerade ‚ich‘ sagt und den ich im Spiegel unauffällig observiere (oder den ich im Traum im Supermarkt verfolge), die Beschwerung anrichtet, sondern dass *ich selbst* es bin.<sup>27</sup> Indem Perry den mentalen Zustand des ‚ich‘-Sprechers (wie er in der ‚ich‘-Perspektive erlebt wird) für prinzipiell unersetzbar hält (John Perry könnte im obigen Beispiel sehr wohl nicht wissen, was seine staatsbürgerliche Identität ist, ja worauf ‚ich‘ überhaupt referiert), plädiert er für die Anerkennung des singulären ontologischen Status derjenigen Realitäten, auf die mit ‚ich‘ verwiesen wird. Nimmt man sich nur ein wenig Zeit und befragt seine eigene Phantasie (oder die Weltliteratur), so wird man rasch ungezählte Beispiele finden, die Kaplans Reduktionismus als die scholastische Engstirnigkeit diskreditieren, die sie ist. In einem Gedankenexperiment kann man sich jedenfalls leicht tausende von Fällen vorstellen, in denen ein Sprecher, der ‚Da richte ich einen ganz schönen Unfug an‘ für sich als zutreffend annimmt, damit aber noch nicht den Satz ‚Dthat (die eben dies Zeichenereignis äußernde Person) richtet eine schönen Unfug an‘ für damit äquivalent hält. Denn das Geheimnis der cartesianischen Gewissheit besteht u. a. darin, dass sie sich eines Denkers nur dann bemächtigt, wenn der eine Formulierung der ersten Person singularis als Zugangsweg wählt („ego cogito“), nicht auch dann, wenn er einen Denker sich über sein Denker äußern hört.

---

<sup>26</sup> „The Problem of the Essential Indexical“, in: *Propositions and Attitudes*, ed. by Nathan Salmon and Scott Soames, Oxford University Press 1988, 83-101. (Der Aufsatz erschien zuerst 1979 in *Noûs*, Vol. XIII, No. 1, 3-21.)

<sup>27</sup> L. c., 92: „Saying that I believed of John Perry that he was making a mess leaves out the crucial change, that I came to think of the messy shopper not merely as the shopper with the torn bag or the man in the mirror, but as *me*.“

Es gibt noch einen dritten Einwand gegen die Gleichsetzung von ‚ich‘ mit ‚der/die gerade das Zeichen Äußernde‘. Er findet sich bei Elizabeth Anscombe. In einem klassischen Text („The First Person“) hat sie aporetische Konsequenzen gezogen aus der Annahme, das Pronomen der ersten Person singularis sei entweder ein Eigename oder ein Demonstrativpronomen und habe, wie alle anderen seinesgleichen, mithin einen Referenten. Wer diese Annahme teilt, muss angeben können, wer oder was dieser Bezugsgegenstand ist.<sup>28</sup>

Nehmen wir an, ‚ich‘ sei ein Name. Syntaktisch könnte das einleuchten, denn ‚ich‘ tritt in Sätzen an die Stelle, die sonst von Namen (oder anderen singulären Termini) eingenommen werden kann: es ist ein Für-Wort, ein Pro-nomen. Aber dann müsste der Bezugsgegenstand – wie der aller singulären Termini – ein Gegenstand sein; und der müsste sich – nach Frege – in eine Reihe von Kennzeichnungen auflösen lassen (in generelle Termini – Begriffe – oder in Arten und Weisen seines Gegebenseins, wie Frege sagt). Anders gesagt: jeder Gegenstand kann auf beliebig viele Weisen (Frege’sche ‚Sinne‘) spezifiziert werden. Aber ein Gegenstand, der unter gar keiner Beschreibung existiert, wäre kein Gegenstand. Wie ist uns aber der Gegenstand von ‚ich‘ gegeben? Als Selbstbewusstsein oder als cartesianisches *cogito*: Nur so könnte laut Elizabeth Anscombe die Antwort lauten. Denn ‚ich‘ funktioniert auch in Unkenntnis aller deskriptiven Merkmale der ‚ich‘ sagenden Person; mithin meinen wir mit ‚ich‘ anderes, als was wir mit Körpereigenschaften bezeichnen können. Also wäre der (Frege’sche) Sinn des Referenten von ‚ich‘ etwas Körperloses wie Selbstbewusstsein.

---

<sup>28</sup> Ich stütze mich im Folgenden zuweilen auf einen glänzend informierten und scharfsinnigen Text von Eros Corazza, „Dire ‘je‘“, Université de Genève 1989 (unpubliziert).

Nimmt man dagegen, um diese unerwünschte Konsequenz zu vermeiden, an, ‚ich‘ sei ein Demonstrativum, so stößt man sich an Freges Ausschluss der Demonstrativ-Pronomen aus der wahrheitsfähigen Rede – denn Demonstrativa bezeichnen ihre Gegenstände geradehin, und nicht unter Beschreibungen (Begriffen, Kennzeichnungen, einem Wie des Gegebenseins). Wir hätten dann zwar einen Referenten, von dem wir aber gar nicht sagen könnten, *was* er ist oder womit wir zu tun hätten (er fiel unter keinen Begriff, nicht einmal ein Sortal). Das Problem in der Deutung von ‚ich‘ als Namen-Wort würde sich so noch verschärfen.

Aber Elizabeth Anscombe weist die Annahme, ‚ich‘ sei ein Demonstrativum, auch noch mit dem zusätzlichen Einwand zurück, dass Demonstrativa wie ‚dies‘ oder ‚jenes‘ auch in Abwesenheit (oder Unkenntnis) ihres Bezugsobjekts funktionieren können. Als Beispiel wählt sie jemanden, der mit einem Behältnis des Weges kommt und sagt: „Das ist alles, was von dem armen Jones übrig geblieben ist.“ ‚Das was?‘ möchte man zurückfragen und wird erwarten, der erfragte Gegenstand werde durch die Kennzeichnung ‚dies Häuflein Asche‘ spezifiziert. In der Tat aber war (ohne dass der Sprecher das wusste) die Urne leer.<sup>29</sup> So kann von einer generell ‚garantierten Referenz‘ bei Demonstrativa keine Rede sein, obwohl sie manchmal gegeben ist (z. B. wenn ich sage ‚Dies Sirren und Schwirren in meinen Ohren ist unerträglich‘).

Dagegen macht es keinen Sinn, einen solchen Verweisungsirrtum auch für ‚ich‘ zu argwöhnen; denn ‚ich‘ ist mit sich bekannt, gleich welches die Be-

---

<sup>29</sup> „Die erste Person“, *Collected Philosophical Papers*, II, 28; deutsche Übersetzung bei Bieri (Hg.), I. c. 230.

schreibung (der Freges'che ‚Sinn‘) ist, unter der es sich zu fassen bekommt. ‚Ich‘, sagt Elizabeth Anscombe, ist keine Chimäre, sondern „etwas Wirkliches, obwohl noch immer Unerklärtes“<sup>30</sup>. Sein Referent (wenn es denn einen gibt) kann nicht identisch sein mit etwas durch Beobachtungsprädikate oder Kennzeichnungen Spezifizierbares. Denn wenn John Smith von dem (in einem Testament so spezifizierten) John Horatio Auberon Smith spricht, muss er nicht wissen, dass in diesem Testament von *ihm selbst* die Rede ist. Und wenn Descartes (der vielleicht Gassendi hieß und sich selbst blitzgescheite Einwände zuschickte) an der Existenz von Descartes (als Namens-träger und als raum-zeitlicher Person) zweifelt, so zweifelt er keineswegs an *sich* als dem diesen Zweifel äußernden denkenden Subjekt (res cogitans). Wir meinen eben mit ‚ich‘ nicht einfach das, was ein observierbares und spezifizierbares Objekt sein könnte und hinsichtlich dessen Identifikations- (oder Spezifikations-)Irrtümer auftreten könnten. Observiert wird Körperliches. In direkten oder indirekten Reflexions-Vollzügen treffen wir dagegen nur auf Bewusstsein; und wenn dieses Bewusstsein seinerseits ein Körper-Bewusstsein ist – z. B. ein Fühlen –, so ist uns die physische Komponente darin ihrerseits nur mental – eben als Gefühltes – erschlossen. Eben darum könnten wir den Gebrauch von ‚ich‘ auch nicht einfach so erklären, dass mit ihm ein jeder sich selbst bezeichnet. Dies träfe für den Ausdruck ‚A‘ in der Sprache der ‚A-Benutzer‘ auch zu, und doch ist ‚ich‘ durch ‚A‘ nicht zu ersetzen. Ich zitiere die entscheidenden Passus mal ganz:

Man frage sich: ist es wirklich wahr, daß ‚ich‘ nur deshalb nicht ein Eigenname genannt wird, weil jedermann es ausschließlich dazu gebraucht, um sich auf sich selbst zu beziehen? Wir wollen einen eindeutigen Fall genau eines solchen Namens konstruieren. Man stelle sich eine Gesellschaft vor, in der jeder mit zwei Namen beschriftet ist. Der eine erscheint auf dem Rücken und auf dem oberen Ende des Brustkorbes. Diese Namen, die von ihren Trägern nicht gesehen werden können, sind für die Träger je verschie-

---

<sup>30</sup> „It is something real, though as yet unexplained“ (l. c., 26, dt. 227).



dene: sagen wir von „B“ bis „Z“. Der andere, „A“, ist auf der Innenseite ihrer Handgelenke aufgedruckt und der gleiche für jeden. Jeder, der von Handlungen der Leute berichtet, gebraucht die Namen auf ihrem Brustkorb und Rücken, wenn er sie sehen kann oder daran gewöhnt ist, sie zu sehen. Außerdem lernt jeder auf Äußerungen des Namens auf seinem eigenen Brustkorb und Rücken zu antworten, und zwar so, wie wir auf Äußerungen unserer Namen zu antworten neigen. Berichte von eigenen Handlungen, die man direkt aufgrund von Beobachtung gibt, werden durch Verwendung des Namens auf der Brust gegeben. Derartige Berichte werden nicht allein aufgrund von Beobachtungen, sondern ebenso aufgrund von Schlussfolgerungen und Zeugenaussagen oder aufgrund anderer Informationen gegeben. „B“, zum Beispiel, zieht Schlussfolgerungen, ausgedrückt durch Sätze mit „A“ als Subjekt, aus Aussagen anderer Leute, die „B“ als Subjekt gebrauchen.

Man mag fragen: Was meint man mit „Berichte von jemandes eigenen Handlungen“? Wir wollen festlegen, daß darunter zum Beispiel Berichte zu verstehen sind, die von B's Mund über Handlungen von B ausgehen. Das heißt: Berichte aus B's Mund, die besagen, daß A das und das tat, werden prima facie verifiziert, indem man in Erfahrung bringt, daß B es tat, und endgültig falsifiziert, indem man herausfindet, daß er es nicht tat.

Somit gibt es für jede Person *eine* Person, von der sie charakteristisch eingeschränkte, aber auch charakteristisch privilegierte Anschauungen hat. Außer in Spiegeln sieht sie nie die ganze Person, und kann nur eher spezielle Anschauungen von dem, was sie sieht, gewinnen. Einige von ihnen sind eigentümlich gut, andere eigentümlich schlecht. Natürlich mag ein Mensch B bisweilen einen Fehler machen, indem er den Namen „A“ auf dem Handgelenk eines anderen erblickt und nicht realisiert, daß es das Handgelenk eines Menschen ist, dessen anderer Name für B schließlich nicht in der speziellen Weise unzugänglich ist wie sein eigener Name („B“) es für ihn ist.

(Es mag für die Vorstellungskraft einiger hilfreich sein, wenn wir das Beispiel ändern: statt dieser eher unmenschlichen Leute nehmen wir mit Abtastgeräten ausgestattete Maschinen an, die genauso, wie die Leute in meiner Geschichte mit ihren Namen, mit Zeichen markiert und dafür programmiert sind, das in Berichte zu übersetzen, was auf den Bildschirmen ihrer Abtaster erscheint.)

In meiner Geschichte liegt uns die Spezifikation eines Zeichens als eines für alle gleichen Namens vor, der aber von jedem nur dazu gebraucht wird, um von sich selber zu sprechen. Wie läßt sich das mit „ich“ vergleichen? – Das erste, was es zu beachten gilt, ist, daß unsere Beschreibung kein Selbstbewußtsein seitens der Leute einschließt, die den Namen „A“ so, wie ich es beschrieben habe, gebrauchen. Sie haben vielleicht kein Selbstbewußtsein, obgleich jeder eine Menge von dem Objekt weiß, das er (tatsächlich) ist, und einen Namen hat – den gleichen, wie jeder andere – und zwar einen Namen, den er bei Berichten über das Objekt, das er (tatsächlich) ist, gebraucht.

Die These, daß diese Leute kein Selbstbewußtsein haben – mag einem gerade aus diesem Grund nicht richtig vorkommen. *B* hat Bewußtsein von, d. h. er beobachtet einige von *B*'s – will sagen, seine eigenen Aktivitäten. Er gebraucht den Namen „*A*“, wie es jeder andere tut, um auf sich selbst zu referieren. Also ist er sich seiner selbst bewußt. Also hat er Selbstbewußtsein.

Wenn wir jedoch von Selbstbewußtsein sprechen, meinen wir nicht dies. Wir meinen etwas, das sich durch den Gebrauch von „ich“ im Gegensatz zu „*A*“ kundtut (Elizabeth Anscombe, „The First Person“, in: *The Collected Philosophical Papers of G. E. L. Anscombe*, Vol II: „Metaphysics and the Philosophy of Mind“, Oxford 1981, 21-36, hier: 24).

Somit müssen wir unseren Logiker nötigen, eine „garantierte“ Referenz von „ich“ anzunehmen. Gesteht er dies zu, so gibt es drei Argumentationsebenen dessen, was er behaupten kann:

(1) Er kann sagen, daß der Verwender von „ich“ natürlich existieren muß, sonst würde er nicht „ich“ verwenden. Da er ja der Referent *ist*, ist es dies, worauf „garantierte“ Referenz hinausläuft. Hinsichtlich einer derartigen garantierten Referenz, mag er hinzufügen, wird es keine Differenz zwischen „ich“ und „*A*“ geben. Aber die Frage ist, warum von „ich“ gesagt wurde, daß es auf den Benutzer von „ich“ referiert. Unser Logiker hielt dafür, daß „ich“ logisch gesehen ein Eigenname sei – ein singulärer Terminus, dessen Rolle es ist, zu referieren – und das aus zwei Gründen: Der eine ist der, daß „ich“ die gleiche syntaktische Position wie solche Ausdrücke hat. Und der andere ist der, daß es *salva veritate* durch einen (gewöhnlicheren) Namen von *X* ersetzt werden kann, wenn es in Subjektposition in von *X* aufgestellten Behauptungen vorkommt. Wenn er das sagt, glaubt er sich selbst zweifelsohne auf keine Ansichten über den Sinn von „ich“ oder darauf, was der „Ich“-Benutzer mit „ich“ meint, verpflichtet. Aber sein zweiter Grund läuft auf folgendes hinaus: Jemand, der eine Behauptung mit „ich“ als Subjekt hört oder liest, muß wissen, wessen Behauptung es ist, wenn er wissen will, auf wen das Prädikat zutrifft, falls die Behauptung wahr ist. Nun, diese Bedingung könnte durch das Aufblinkenlassen eines Grünlichts in Verbindung mit dem Prädikat oder vielleicht durch Anfügen eines Endungs-„O“ an das Prädikat signalisiert werden. (Ich entschuldige mich bei jedem, der diesen Vorschlag im Ganzen genommen allzu phantasievoll findet, und bitte ihn, Zweifel zu unterdrücken.) Was würde ein solches Signal oder Suffix zu einem referentiellen Ausdruck machen? Das entscheidende Argument kann nicht ein Argument zurück von der Syntax zur Referenz sein, denn solch ein Argument hinge allein von der Form des Satzes ab und wäre absurd. (Z. B. glaubt niemand, daß „Es“regnet“ einen referentiellen Ausdruck enthält, und zwar „Es“. Es scheint so zu sein, daß unser Logiker nicht darauf verzichten kann, sich mit dem Sinn von „ich“ zu beschäftigen, oder zumindest mit dem, was der „Ich“-Benutzer meinen muß.

(2) Also muß der „Ich“-Benutzer beabsichtigen, auf etwas zu referieren, wenn „ich“ ein referentieller Ausdruck ist. Nun gibt es hier aber zwei ver-

schiedene Dinge, die unter „garantierter Referenz“ verstanden werden können: Es kann (2a) die garantierte Existenz des Objektes, das von dem Benutzer gemeint ist, bedeuten. D. h., dasjenige Objekt muß existieren, von dem er etwas meint, wenn er den Ausdruck in Verbindung mit ihm benutzt. Wenn ich somit annehme, jemanden namens „X“ zu kennen und ich nenne etwas „X“ mit der Intention, auf diese Person zu referieren, so wäre die Garantie von Referenz in diesem Sinn eine Garantie dafür, daß es ein solches Ding wie X gibt. Der Name „A“, den ich erfand, hätte diese Art garantierter Referenz. Der Benutzer von „A“ meint von einem bestimmten menschlichen Wesen zu reden, einem, das in seinen Wahrnehmungshorizont auf ziemlich eigentümliche Weise fällt. Diese Person ist er selbst, und folglich kann er, vorausgesetzt, daß er den Gebrauch von „A“ begriffen hat, nur von einer wirklichen Person sprechen.

Wenn unser Logiker dies für eine angemessene Darstellung der „garantierten Referenz“ hält, dann wird er einräumen müssen, daß es einen dritten Typ „garantierter Referenz“ gibt, welchen „ich“ *nicht* hat. Garantierte Referenz für jenen Namen „X“ in diesem zusätzlichen Sinne (2b) schließt eine Garantie mit ihm nicht nur dafür ein, daß es solch ein Ding wie X gibt, sondern auch dafür, daß das, was ich für X halte, *X ist*. Wir sahen, daß der Benutzer von A nicht immer immun gegen die irrtümliche Identifikation eines anderen als „A“ war. Wird das auch bei „ich“ der Fall sein?

Die Annahme scheint absurd zu sein. Es scheint klar zu sein, daß, wenn „ich“ überhaupt ein ‚referentieller Ausdruck‘ ist, ihm beide Typen garantierter Referenz zugeschrieben werden können. Ein Objekt, das ein Benutzer von „ich“ meint, muß existieren, solange wie er „ich“ benutzt. Ebenso wenig kann er das falsche Objekt für das Objekt halten, welches er mit „ich“ meint. (Der Bischof mag das Knie der Dame für sein eigenes halten, aber könnte er denn die Dame selbst für sich halten?)

Stellen wir die Frage nach der Bedeutung von „ich“ zurück und fragen *nur*, wie Referenz auf das richtige Objekt garantiert werden könnte. (Das ist ein angemessenes Verfahren, denn sicherlich hat man hier immer die Vorstellung eines Typs reiner direkter Referenz, bei dem man zuerst einfach etwas meint und anschließend auf ein Objekt vor einem referiert.) Es scheint ferner so zu sein, daß diese Referenz nur dann bombensicher sein könnte, wenn der Referent von „ich“ zum einen mit jedem Gebrauch von „ich“ neu definiert würde, und zum anderen solange im Blick bliebe, wie etwas für *Ich* gehalten würde. Aber auch hier wird vorausgesetzt, daß etwas anderes seinen Platz nicht erschlichen einnimmt. Vielleicht sollten wir sagen: Eine derartige Voraussetzung ist im Falle von „ich“ außerordentlich zuverlässig. Und es wäre aufs Ganze gesehen ein Auswuchs von Skeptizismus, sie in Zweifel zu ziehen! Folglich akzeptieren wir die Voraussetzung, und es scheint zu folgen, daß das, wofür „ich“ steht, ein cartesianisches Ego sein muß (l. c., 29-31).

Kurz: es hat mit ‚ich‘ eine gegenüber Nomina und Demonstrativa eigene Bewandtnis. Ich kann mir nämlich schlechterdings keine Situation vorstellen, in der ‚ich‘ keinen Referenten hätte oder sich im Verweisungsbezug täuschen könnte (wie John Smith über seine Identität mit John Horatio Auberon Smith oder der Träger der Urne über die Asche des armen Jones). ‚Ich‘ wäre damit so etwas wie ein Russellscher ‚logischer Eigenname‘ oder ein Name im Sinne des Wittgensteinschen *Tractatus* oder ein Kripkescher ‚rigid designator‘ – d. h. ein solcher, der – regelkonform verwendet – seinen Referenten unabhängig von allen Beschreibungen notwendig, weil direkt, erreicht.

Das ist, wie wir schon hörten, auch die Ansicht von Hector Neri Castañeda:

The pronoun ‚I‘ has an *ontological priority* over all names and descriptions. A correct use of ‚I‘ cannot fail to refer to the object it purports to refer.<sup>31</sup>

\*

Dem Wink auf die transzendentalphilosophischen Konsequenzen seiner frühen Selbstbewusstseins-Theorie ist Castañeda in seinen letzten Publikationen nachgegangen. Z. B. in *The Self and the I-Guises, Empirical and Transcendental*,<sup>32</sup> charakteristischerweise seinem Beitrag zur Festschrift für Dieter Henrich. Ihm ist der Aufsatz gewidmet mit den Worten: „Here is a mere prolegomenon to a general theory of selfconsciousness – dedicated TO DIETER HENRICH with gratitude and with admiration for his illumining

<sup>31</sup> ‚He‘: A Study in the Logic of Self-consciousness, 144; dt.

<sup>32</sup> In: *Theorie der Subjektivität* (=FS für Dieter Henrich), hg. von K. Cramer, H.F. Fulda, R.-P. Horstmann und Ul. Pothast, Frankfurt/M. 1989, 105-149 (zit.: SIG).

contributions to our understanding of the nature of consciousness, selfhood, and selfconsciousness“ (105).

Castañeda hat seine späte Theorie ‚Guise Theory‘ genannt (vgl. 105 f.), Helmut Pape hat das durch ‚Theorie der Gestaltungen‘ übersetzt. Als erste Annäherung soll uns genügen, dass Castañeda – wider den herrschenden realistischen Trend (vgl. Anm. 2 auf S. 111) – an einer Abart der Frege-Semantik festhält. Von ihr ergibt sich für ihn ein organischer Weg in eine Art Idealismus der ‚guises‘, sagen wir vorläufig: der Frege-Sinne. Sie erinnern sich: Frege hatte die ‚Bedeutung‘ von Ausdrücken dasjenige genannt, worauf die Ausdrücke Bezug nehmen. Bei singulären Termini sind das Einzelgegenstände, bei ganzen Sätzen ist es der Wahrheitswert. Unter ‚Sinn‘ verstand er dagegen die ‚Art und Weise‘, wie dieser Gegenstand gegeben ist. So ist der Sinn ein Weg zur Identifikation des Gegenstandes (neben anderen). Das berühmte Beispiel, das Frege gibt, ist das der Venus, die sich am Abendhimmel als Abend- und am Morgenhimmel als Morgenstern zeigt. Sätze wie ‚O du mein holder Abendstern‘ und ‚O du mein holder Morgenstern‘ haben daher verschiedenen Sinn (oder drücken, wie Frege sagt, verschiedene ‚Gedanken‘ aus). Sie haben aber denselben Wahrheitswert, weil die beiden singulären Termini in den Sätzen für denselben Gegenstand stehen. So verhält sich’s auch mit ‚Gaurisankar‘ und ‚Mount Everest‘ oder ‚gleichschenkliges‘ und ‚gleichwinkliges Dreieck‘ usw. Frege konnte mit dieser Unterscheidung u. a. eine für Castañeda (wie wir gleich sehen werden) wichtige Differenzierung erklären: dass nämlich zwei Personen (oder eine Person zu verschiedenen Zeiten) von zwei Sätzen den einen glauben und den anderen für falsch halten können, weil beide verschiedene Gedanken ausdrücken, obwohl sie beide dieselbe Bedeutung haben. Obwohl Frege damit überhaupt keine Annäherung an die Perspektive des Erkennenden

Subjekts im Sinn hatte, kommt doch (bei geeigneter Erweiterung) der erkenntnistheoretische Aspekt des Guise-vermittelten Zugangs zu den Gegenständen in seiner Theorie auf seine Kosten. Darum durfte ich vorderhand ‚Guise‘ mit ‚Frege-Sinn‘ übersetzen: mit dem über Weisen des Gegebenseins vermittelten Zugang zu den Gegenständen oder Tatsachen. Den Schritt zu einer idealistischen Lesart vollzieht Castañeda, wenn er den Gegenstand in den Erscheinungen mit Haut und Haar aufgehen lässt, in denen er sich uns zeigt. Damit wird die Realität zu einem Konstrukt aus subjektiven Gegebenheitsweisen.

Wie das geschieht, müssen wir uns aber etwas geduldiger anschauen – denn in der Philosophie zählen nie die Überzeugungen, sondern die argumentativen Wege, durch die wir zu ihnen gelangen. Zunächst erinnert uns Castañeda (im Zentrum von dessen Werk ja die Analyse von indexikalischen Ausdrücken steht) an eine Eigentümlichkeit derselben. Indikatoren sind Fürwörter wie ‚ich‘, ‚du‘, ‚dies‘ oder ‚jenes‘; und indexikalische Kennzeichnungen sind dann Ausdrücke wie ‚dies hübsche Gesicht‘, ‚diese Felswand‘ (SIG 107). Indexikalische Ausdrücke sind eine Unterklasse singulärer Termini, also von Ausdrücken, die für Einzelgegenstände (nicht für Klassen oder Eigenschaften) stehen. Nun ist das Entscheidende beim Gebrauch von indexikalischen Ausdrücken ihre Kontextabhängigkeit. ‚Hier‘, ‚dies‘, ‚jetzt‘ oder ‚ich‘ bedeuten von Verwendung zu Verwendung, von Zeitpunkt zu Zeitpunkt, von Ort zu Ort, ja von Person zu Person etwas Verschiedenes. Und darum bedürfen sie zu ihrer referentiellen Erfüllung (das ist eigentlich was Triviales, aber Castañeda gibt diesem Umstand eine überraschende Wendung) einer Ergänzung. Für Castañeda ist das (wie für Ayer) die konkrete (sinnliche) Erfahrung: „Briefly, a token T of an indexical singular term gains a referent by virtue of an actual *experience* lived through the tokening

in context C of some sentence containing the token T" (107 u.). Ein Gegenstand ist eben nicht einfach *dieser* oder *du* durch zufällige Nachbarschaft zum Sprecher. Er wird aus vielen anderen herausgefiltert durch eine konkrete und intentionale Begegnung (*encounter* [108 o.]). Und diese Begegnung wird durch *Erfahrung* vermittelt („Indexical reference is executive and experiential“). Wir sahen schon, dass das Ich der Erfahrung zwar selbst nicht Gegenstand einer Wahrnehmung ist. Aber auch die Bedeutung von ‚ich‘ wird durch das aktuelle Erfahrungserlebnis festgelegt, in dem es auftritt („[...] an *I* is the subject of an experience conceived *qua* subject of the experience while the experience lasts“). Castañeda vertritt entsprechend, wie z. B. vor ihm Russell, eine Theorie, nach der das Subjekt ein ‚short term self‘ ist, also nur das flüchtige („evanescent“) Subjekt dieser einzelnen Erfahrung. So ist auch das Ich ‚guised‘, also immer das konkrete Subjekt dieser Einzelerfahrung: „An I-guise is the subject of an experience intimately linked to the present and presented time and the present space of experience, i. e. to a Now-guise and a Here-guise. For these there *esse is to be involved in an experiencing*“ (106,<sub>1</sub>; 138). (Damit stellen sich natürlich Fragen nach der überzeitlichen Einheit des Selbst; sie ähneln denen von Prousts ‚mois successifs‘.)

Dies geklärt, bereitet Castañeda (vorerst dunkel) eine wichtige Unterscheidung vor: die zwischen dem semantisch-pragmatischen Bezugsgegenstand eines singulären Terminus und seinen doxastischen ‚co-denotata‘ (108,<sub>2</sub>). Ich nehme durch einen Ausdruck Bezug auf etwas, indem ich die semantischen Regeln der Referenz, aber auch den (pragmatischen) Kontext beachte, in dem der Referent gegeben wird. Die Bezugnahme ist aber immer ‚eine Denk-Episode der betreffenden Person‘, also etwas, das sich durch ein vor-

geschaltetes ‚Ich bin’s, der das denkt, wahrnimmt, hofft usw.’<sup>33</sup> ergänzen lässt und diese Ergänzung in Wahrheit immer voraussetzt (ohne diese Voraussetzung bliebe die Prämisse außer Kraft, dass alle Kontexte ihre letzte Bestimmung erst aus der Selbstsytuierung des erfahrenden Subjekts gewinnen). ‚Doxastisch‘ (von griechisch δοξάζω: meinen, glauben) bedeutet die Filterung dieser Referenzen durch das, was der Sprecher für dasselbe oder nicht dasselbe hält: So hätte ein Sprecher zur Zeit des Aristoteles den Satz ‚der Abendstern ist derselbe Planet wie der Morgenstern‘ nicht für wahr gehalten. Den Übergang von der direkten Referenz zur doxastischen Mit-Vermeinung bereitet Castañeda geschickt dadurch vor, dass er die Referenz selbst in eine Art doxastischen Raum einbettet: jede Referenz ist ja selbst (in letzter Instanz) ein ‚Ich glaube, daß dort und jetzt ein soundso‘. Dennoch gibt es Unterschiede zwischen Glauben und Referieren. So werden wir sagen, dass auch schon die alten Babylonier, wenn sie schmachend auf den Abendstern guckten (und ihn für den Abendstern hielten – und sonst nichts), in Wahrheit auf die Venus Bezug genommen haben. Wir werden aber, als eingefleischte Verifikationisten, hinzufügen, dass sich der Abendstern seither auf eine selbst erfahrungsgestützte Weise als die Venus *herausgestellt* hat. So ist auch die Korrektur an der ehemaligen Bezugnahme eingebettet in ein System von für wahr gehaltenem, also in einen letztlich doxastischen Kontext.

Das wird manifest an dem, was Castañeda das ‚Paradox der Referenz‘ nennt (110 ff.). Es stellt ein Grundproblem der Philosophie dar. Sie alle wissen,

---

<sup>33</sup> Castañeda macht natürlich einen Unterschied zwischen Erfahren und Denken (SIG 109): Denken ist „das Rückgrat der Erfahrung“, weil indifferent gegen den jeweiligen Erfahrungstyp. So kann es zum Organ der „Einheit der Erfahrung“ (quer durch die Vielfalt der Einzelerfahrungen) werden: „thinking is intrinsically the same whatever the type of the experience the flow of thinking constitutes, whatever it is a thinking of existents, or of fictions, or of hypotheticals, or of hypostases, whether the thinker realizes or judges what kind of thinking it is“ (l. c.).



dass es möglich ist, Gegenstände für verschieden zu halten, die in Wirklichkeit identisch sind. Betrachtet man nun die entsprechenden Sätze, also einen Glaubenssatz und einen Behauptungssatz in direkter Rede, dann entsteht das Paradox genau dann, wenn man die Denotate (also die bezeichneten Gegenstände) der singulären Termini in beiden Sätzen für identisch hält und andererseits die im Behauptungssatz direkter Rede ausgesagte Identität der singulären Termini für eine Identität im strengen Sinne hält. In diesem Fall wären nämlich aufgrund von Leibniz' Gesetz die singulären Termini, deren Denotat im Behauptungssatz direkter Rede identifiziert wird, auch im Glaubenssatz *salva veritate* austauschbar, obwohl dieser gerade deren Verschiedenheit besagt. Es würde sich für den Glaubenssatz ein logischer Widerspruch durch die Substitution ergeben, wenn in ihm zwei Gegenstände für nicht identisch gehalten würden, die tatsächlich identisch sind. Nehmen wir das bekannte Beispiel, das Castañeda in Form der folgenden 7 Sätze anführt (und das zahlreiche Autoren, z. B. Gareth Evans, von ihm übernommen haben):

(1) Zur Zeit der Pest glaubte Iokaste sowohl, daß Ödipus' Vater mit Ödipus' Vater identisch ist, als auch, daß der frühere König von Theben nicht mit Ödipus' Vater identisch ist.

(2) Der frühere König von Theben war dieselbe Person wie Ödipus' Vater.

(3) Es war nicht der Fall, daß zur Zeit der Pest Iokaste sowohl glaubte, daß Ödipus' Vater mit Ödipus' Vater identisch und Ödipus' Vater nicht mit Ödipus' Vater identisch ist.

(T1) Für jeden Einzelgegenstand (for any individuals)  $x$  und  $y$  gilt: wenn  $x$  (im eigentlichen oder strengen Sinne) mit  $y$  identisch ist, dann trifft alles, was auf  $x$  zutrifft, auch auf  $y$  zu und umgekehrt [= Leibniz' Gesetz].

(T2) Der Stammsatz (sentential matrix) von (1) und (3), nämlich „zur Zeit der Pest glaubte Iokaste sowohl, daß Ödipus' Vater mit Ödipus' Vater identisch und nicht mit Ödipus' Vater identisch ist“, drückt etwas Wahres aus über (eine Eigenschaft des) Einzelgegenstand(s), der durch den singulären

Terminus bezeichnet (denoted) wird, [etwas,] das durch Ausfüllung der Leerstelle (blank) im Stamm (matrix) einen wahren Satz produziert.

(T3) Der Ausdruck ‚ist der/dieselbe wie‘ in (2) drückt echte (genuine) Identität aus.

(T4) Die singulären Termini ‚Ödipus‘ Vater‘ und ‚der frühere König von Theben‘ haben in den Prämissen (1)-(3) durchgängig dieselbe Bezeichnung ([denselben Gegenstand: denotation], und zwar sowohl in direkter wie in indirekter Rede. (110 f.)

„Der Widerspruch entspringt aus einer Spannung zwischen Selbigkeit und Unterschiedenheit. In Iokastes Geist oder doxastischem System gibt es einen *Unterschied* zwischen Ödipus‘ Vater und dem früheren König von Theben. Dieser Unterschied erhellt aus (1) und (3) zusammengenommen. Andererseits herrscht *Selbigkeit* zwischen den beiden Entitäten, wie sie in Prämisse (2) zum Ausdruck kommt. Diese Selbigkeit gilt (holds) in der Wirklichkeit oder zumindest im Geist eines jeden, der Prämisse (2) für richtig hält (believes)“ (111).

Frege hat das Paradox zu lösen gesucht, indem er das Denotat der singulären Termini (‚Ödipus‘ Vater‘ und ‚der frühere König von Theben‘) mit ihrem Sinn gleichsetzt. Dagegen plädiert Castañeda für eine alternative Lösung: Danach soll die Identität im Behauptungssatz direkter Rede nicht als strikte Identität verstanden werden. Darum wäre das hier Identifizierte nicht *salva veritate* im Glaubenssatz substituierbar. So glaubt Castañeda, eine Annahme vermeiden zu können, die Frege machen muss: dass nämlich das Denotat von singulären Termini in direkten und indirekten Kontexten verschieden ist (112 f.).<sup>34</sup>

<sup>34</sup> wie Frege das annahm: ihm zufolge referiert ein Eigenname (singular term) in indirekter Rede nicht auf den Bezugsgegenstand, sondern auf seinen Sinn. Frege verwirft (T4) und behauptet, (T1) sei unanwendbar auf (1) & (2). Trotzdem hält er an (T1) fest, dem so genannten Leibniz-Gesetz, fest, meint aber, dies sei in psychologischen Kontexten unanwendbar. Dessen ungeachtet hält er an einer Version von (T2) fest, läßt also psychologische Eigenschaften wie die in (T2) genannten zu. Diese Ei-

Das kann Castañeda darum tun, weil er den Unterschied zwischen objektiver Identität und geglaubter Identität auflöst. Identität ist nie strikt, sondern kontingent. Ein Gegenstand besteht aus ‚konsubstantiierten‘ Bestandteilen. Und was da jeweils zusammengestellt wird, um einen einigen Gegenstand zu bilden, ist grundsätzlich relativ auf das Glaubenssystem einer Person. Anders gesagt: was relativ zu ihrem doxastischen System von einer Person (zu einem Zeitpunkt) für dasselbe gehalten wird, das und nichts anderes macht ein individuelles Guise (ein komplexes Eigenschafts- oder Erscheinungs-Bündel) aus (116 Anm. und 117). Alles andere, was von der Person nicht für identisch gehalten wird, es aber dennoch (in primärer Referenz) ist (nämlich für andere Personen oder für einen idealen Beobachter), wird als kontingente Klassenbildung aus diesen individuellen Bestandteilen verstanden. Es ist klar, sagt Castañeda, „daß jedes Paar von singulären Termini, die gewöhnlich für ko-referentiell gehalten werden, in einem Differenzierungstest durchgegangen werden kann (wie es der ist, der Iokastes Überzeugungen und Annahmen enthält: (T1)-(T4)). Die Überzeugungen (beliefs) einer Person sind somit Aussiebeverfahren (sifting devices) zur Unterscheidung von Identität und Konsubstantiation und zwischen strikt identischen und verschiedenen (distinct) individuellen Guises, die in irgendeiner Hinsicht dieselben sind“ (114 u.). Anderswo: Subjektive Einstellungen (attitudes) „sind *ontologische Prismen*, die massive Glaubensobjekte in denkbare Guises (auf)brechen (diffract)“ (115).

---

genschaften stellen sich aber als Relationen zwischen Denkern und Sinnen, nicht als solche zwischen Denkern und primären Referenten heraus. Andererseits akzeptiert Frege (T3), hält also an einem Begriff von Selbigkeit fest. Außerdem hat Frege eine schlichte Ansicht von Prädikation. Trotzdem bleibt unklar, wie individuelle Sinne einerseits auf Eigenschaften, andererseits auf die primären Referenten sich beziehen. Eine andere Form der Prädikation scheint erfordert, um Sinne und Eigenschaften zu verknüpfen, zumindest wenn sie in indirekter Rede vorkommen. Das versucht Castañeda (SIG 113).

Die ontologische Analyse, die die Bestandteile der wirklichen Welt ermittelt und diese dann sekundär als ein Konstrukt aus diesen Bestandteilen auffasst, hat ihr Prinzip und Kriterium also im Überzeugungssystem eines bzw. aller Sprecher. Dass die Realität aufgrund dieser Analyse als Konstrukt aus subjektiven Gegebenheitsweisen aufgefasst werden muss, wird deutlicher, wenn man sich überlegt, dass sich das Überzeugungssystem eines Sprechers wandelt und verschiedene Sprecher offensichtlich unterschiedliche Überzeugungssysteme haben. So kann etwa ein Sprecher annehmen, dass Ödipus' Vater mit dem vormaligen König von Theben identisch ist, und ein anderer (nämlich Ödipus und Iokaste), dass es da um verschiedene Personen handelt. Daraus ergibt sich, dass im ersten Fall ein Guise die Identität von Ödipus' Vater mit dem früheren thebanischen König zum Inhalt hat, während im zweiten Fall zwei unabhängige Guises angenommen würden. Die Bestandteile der Wirklichkeit sind danach abhängig von den jeweils zugrunde gelegten Überzeugungssystemen. An sich existierende Identitäten sind ausgeschlossen. Der Unterschied zweier Überzeugungen ist eben nur ein solcher: Wenn Columbus auf Kuba landete, glaubte er nicht, Fidel Castros Land zu betreten – und doch ist gerade dies in einer anderen glaubensabhängigen Gestaltung (der unseren) der Fall. Was nie der Fall ist, ist ein Realitäts-Vorkommnis, das niemandes Überzeugung wäre – also ein sinnfrei gegebener Referent.

Hier ist natürlich das Absprungbrett zu einer Art von Idealismus. Er teilt mit dem klassisch-idealistischen sogar die eigentümliche Verknüpfung der Auffassung von der Welt als Erscheinung und ihrer Abhängigkeit vom Selbstbewusstsein. Wir sahen schon, dass Castañeda jeden expliziten Erfahrungsbezug für übersetzbar hält in ein implizites ‚Ich bin's, der diese Erfahrung macht'. Das ‚ich' in einem explizit gemachten Überzeugungssystem

steht immer für das Subjekt der Erfahrung. Da das Subjekt aber je nur das Subjekt einer einzelnen Erfahrung ist, muss die Person (einer in der Zeit sich erstreckenden Lebensgeschichte) als Konsubstanziation aus vielen Gegebenheitsweisen mit einem I-Guise verstanden werden. An sich ist das Subjekt sich aber völlig unabhängig von einer solchen Gemeinschaft mit deskriptiven Bestandteilen gegeben, wie Kants ‚Apperzeption‘: *„The essence and the substance of an I is just to conceive itself as a subject qua subject“* (118). Und diese Unabhängigkeit des Subjekts von allen Objektbestimmungen ist ihm selbst in der Entdeckung des I-Guise ‚selbst-gegeben‘. (*„An I-guise is the I of an experience. An I, a Here and a Now constitute the inner framework of an experience. They are mutually irreducible and irreducible to the contents of experience“* [139].)

Diese Entdeckung wird dann zum Angelpunkt für Castañedas weitere Überlegungen. Er versucht nämlich Kants berühmten Satz (aus dem § 16 der *KrV*), dass das Ich denke alle meine Vorstellungen muss begleiten können, linguistisch zu transformieren. Kants Gedanke lasse sich als das Prinzip (I-OO) rekonstruieren, wonach jede Äußerung eines Sprechers, die einen vollständigen Gedanken ausdrückt (ob von ihm geglaubt oder nicht), implizit von einem „Ich denke, dass“ abhängig ist, womit der jeweilige Sprecher auf sich selbst Bezug nimmt (118). Und daraus ergibt sich das zweite Prinzip (I-OR), dass alle Aussagen sich als Formen indirekter Rede analysieren lassen, die von dem „Ich denke“ abhängig sind. Castañeda sucht im Folgenden zu zeigen, dass der I-Guise den Status eines transzendentalen Subjekts einnehmen kann. Da ‚transzendental‘ die Bedingung heißt, unter der wir allererst auf Erscheinungen *als* auf Objekte Bezug nehmen können, darf die Abhängigkeit jedes Objektbezugs von dem impliziten ‚Ich denke, dass‘ in der Tat ‚transzendental‘ heißen. Castañeda nennt das ‚Ich denke, dass‘ das

‚transzendente Präfix‘ (121) und das Gesamt der Erscheinungen, die die Vorschaltung dieses Präfixes *seine* Gedanken werden, den „BALLOON“ (119 f.): sozusagen die Sprechblase im Munde des transzendentalen Subjekts. Aller Glaube *de re* wird also auf Glauben *de dicto* zurückgeführt (auf einen Glauben über den Balloon); und dieser Glaube wiederum gründet in einem vorgängigen epistemischen Sich-selbst-Gegebensein, einer Evidenz *de se* sozusagen.

Diese Verschmelzung einer Theorie vergegenständlichter (und von ihren Referenten abgelöster) Frege-Sinne mit dem kantischen Phänomenalismus erfordert eine Kombination aus beiden. Castañeda spricht von der C-Kant-Frege-These (122), wobei das ‚C‘ für ‚combined‘ steht. Sie besagt (ausformuliert), dass alle Aussagen als indirekte Rede in Abhängigkeit vom Ich-denke des jeweiligen Sprecher-Subjekt verstanden werden müssen. Damit wird nun Freges These verbunden, wonach das Denotat singulärer Termini in indirekten Kontexten ihr Sinn ist. Eine Konsequenz dieser kombinierten These ist, dass die Bedeutung im Sinne Freges für die Erklärung der Semantik singulärer Termini keine Rolle mehr spielt. Bedeutung existiert demnach überhaupt nur noch als ‚individual guise‘ (vgl. 124,1), als subjektive epistemische Überschreitung von Gegebenem auf seinen Sinn hin. Für Frege hatte die Rede vom Sinn durchaus nicht diese epistemische oder gar individualistische Komponente: sie war etwas durchaus Allgemeines und keineswegs relativ auf das Subjekt, das einen Gedanken (wie allerdings auch Frege sich ausdrückt) ‚erfasst‘. So wie Ayer physische Gegenstände als logische Konstruktionen aus einzelnen ‚sense-qualia‘ beschreibt, so sucht Castañeda die Referenz als Konstruktion aus Einzel-Sinnen bzw. „individual guises“ zu deuten.

Dazu bedarf's eine Theorie indexikalischer Referenz (124 ff.). Castañedas Überlegungen dazu lassen sich so zusammenfassen:

1. Indexikalische Referenz verweist auf den jeweiligen I-Guise des Sprechers zurück und ist durch ihn festgelegt.
2. Referenz überhaupt läßt sich auf indexikalische Referenz zurückführen und ist deshalb jeweils nur innerhalb eines Referenzsystems bestimmt, dessen Nullpunkt durch den jeweiligen I-Guise festgelegt ist. Alle Referenzen möglicher Bezugnahmen sind deshalb jeweils vom einem I-Guise abhängige „individual guises“ (123 f.).

Nun wäre natürlich noch allerhand darüber zu sagen, wie denn aus lauter Individualperspektiven eine intersubjektiv zugängliche, also eine so genannte objektive Realität sich konstruieren lässt. Castañeda muss hier statt mit Perspektive-unabhängigen Identitäten mit Ähnlichkeits-Relationen zwischen ‚individual guises‘ verschiedener (durch jeweils individuelle I-guises festgelegter) Referenzsysteme arbeiten<sup>35</sup> – aber das ist grundsätzlich

<sup>35</sup> also z. B. auch annehmen, dass indexikalische Referenzen eines anderen Sprechers letztinstanzlich in je meinem Referenzsystem gedeutet werden müssen (und umgekehrt). Sehen wir, wie Castañeda den folgenden Beispielsatz analysiert: (21) „Mark McPherran believes that NOW I am working on Plato's views of relational facts in THIS DIALOGUE OF HIS MIDDLE AGE“ (SIG 125). Die hier in Kapitälchen gesetzten indexikalischen Ausdrücke „NOW“, „I“ und „THIS“ gehören zwar grammatisch zur indirekten Rede, die von McPherran Redeperspektive abhängt. Doch wird ihre Referenz tatsächlich durch die Perspektive des Sprechers (der zu sich „ich“ sagt) bestimmt.

Akzeptiert man diese Analyse, dann folgt zwangsläufig, dass durch die indexikalische Sprecherreferenz nicht festgelegt ist, wie (d. h. mit welchem ‚individual guise‘) Mark McPherran auf die Gegenstände bezugnimmt, auf die der Sprecher (der zu sich ‚ich‘ sagt) mit „I“, „NOW“ und „THIS“ referiert. Wir können unter Zugrundelegung von Castañedas Guise-Theorie sogar weitergehend sagen, dass die Denotate der Bezug nehmenden Ausdrücke aus der Perspektive McPherrans mit den Denotaten aus der Perspektive des Sprechers nicht identisch sein können, weil sie jeweils abhängig sind von der Perspektive der intentionalen Referenz-Festlegung.

Die so genannte Alpha-Analyse formalisiert diese Analyse:

(21.A) „There is a person *alpha*, who is the same as I, and a book *beta*, which is the same as THIS DIALOGUE PLATO WROTE WHEN HE WAS MIDDLE AGE, and a time *gamma* which is the same as NOW, such that Mark McPherran believes that *alpha* is at time *gamma* working on Plato's view of relational facts in *beta*“ (128).

In dieser Analyse ist nach Castañeda von entscheidender Bedeutung, dass die Werte der Variablen nicht Substanzen sind (auch keine Locke'schen unbestimmten Träger), sondern Klassen von ‚individual guises‘, also jeweils private Gegebenheitsweisen. Welche Gegebenheitsweisen auf Seiten McPherrans in die Variablen eingehen, bleibt aus der Perspektive des Sprechers unbestimmt. Direkt referieren nur die indexikalischen Ausdrücke, und ihre Referenz ist vom jeweiligen I-Guise des Sprechers abhängig.

Castañeda verallgemeinert diesen Fall nun, indem er behauptet, dass erstens alle indexikalische Referenz (und dazu gehört auch die auf andere Sprecher, von denen wir vernehmen, dass sie ‚ich‘ zu sich sagen) nur jeweils innerhalb eines Referenz-Systems bestimmt ist, das seinen Ursprung im jeweiligen I-Guise (*I believe here now* [129]) hat; und indem er zweitens behauptet, dass die gesamte Erfahrungswelt (einschließlich aller nicht-indexikalischen Fakten) sich nur innerhalb eines indexikalischen Referenz-Systems erfahren lasse.

nicht problematischer als die Frage, wie wir uns verstehen können, obwohl doch jeder von uns in letzter Instanz nur durch individuelle Hypothesen auf den Sinn der verwandten Ausdrücke kommt (das sprachphilosophische Pendant von Castañedas Phänomenalismus).<sup>36</sup> Solch individuelle Hypothesen bewähren sich eben (pragmatisch) oder bewähren sich nicht in interindividueller Interaktion: das war schon Schleiermachers und Humboldts Überzeugung. Intersubjektive Geltung kann ohne Solipsismus-Gefahr vom Einzelsubjekt oder vom Einzelsprecher her eingeführt werden (vielleicht unter Zuhilfenahme einiger spieltheoretischer Annahmen). Dies alles kann ich hier und heute nicht mehr entwickeln.

---

Castañeda will aber seinem Selbstverständnis nach nicht bei einer solipsistischen Position ankommen, die sich auf der Grundlage der vorangehenden Analyse anzubieten scheint. Deshalb weist er darauf hin, dass die Alpha-Analyse selbst deutlich macht, dass der I-Guise jeweils nur eine relative Perspektive der Erfahrung bedingt. Sie enthält nämlich Variablen als Elemente, die nicht durch die jeweiligen ‚individual guises‘ der einzelnen Perspektive vollständig beschreibbar sind, sondern Leerstellen enthalten, die erst aus anderen Perspektiven komplettiert werden können.

Mit dieser Auffassung stellt sich Castañeda offensichtlich (und das Stichwort fällt auch) in die Tradition des Phänomenalismus, wie er von Russell, Carnap und Ayer vertreten wurde. Dieser Phänomenalismus behauptet nicht die Subjektivität aller Erfahrungen (wie der klassische Empirismus), sondern versteht die objektive Welt als Konstrukt (oder Bündel) aus perspektivischen Erscheinungsweise. (Aus einem Referat von Thomas Grundmann, S. 14 f.)

<sup>36</sup> Vgl. dazu meinen Aufriss von Castañedas Auffassung in *Stil in der Philosophie*, Stuttgart: Reclam, 1992, 41 ff.



## 12. Vorlesung

Nachdem wir in den vergangenen Wochen dem Idealismus ein breites Forum zur Aussprache gewährt haben, müssen wir nun allmählich aus Gründen der Gerechtigkeit mit dem Realismus ebenso verfahren. Zwar war auch bisher der Realismus nie völlig abwesend – und zwar, verabredungsgemäß, vor allem in der Person Hilary Putnams. Da er die bedeutendste Gestalt des zeitgenössischen Realismus verkörpert und seine Konzeption des Realismus außerdem eine besonders lehrreiche Entwicklung durchlaufen hat, werde ich in den folgenden Wochen den Akzent stark auf sein Werk setzen.

Aber um Sie brüsk in den realistischen Kontext zu versetzen, werde ich mich heute zunächst an einen anderen Autor halten, nämlich an Colin McGinn. McGinn hat nämlich 1979 „An A Priori Argument for Realism“<sup>37</sup> vorgetragen – und was wäre gründlicher als ein Argument, das den Realismus aus logischen Gründen (also solchen, die aller Erfahrung zuvor als verbindlich einleuchten) zwingend macht? Zwar könnte schon zu Beginn die Frage Befremden hervorbringen, ob mit dem Apriori nicht besser ein Beweis für die Unumgänglichkeit des Idealismus geleistet werden könnte. Denn apriorische Wahrheiten nennt man doch, seit Leibniz, auch ‚Vernunftwahrheiten‘. Und wenn es in letzter Instanz die Vernunft ist, die für die Richtigkeit des Realismus plädiert, so bleibt der so begründete Realismus eben auf die Vernunft begründet – kaum anders, als das in Maimons oder Fichtes idealistischer Ableitung der relativen Berechtigung des realistischen Standpunkts der Fall war. Aber stellen wir diese unsere Bedenken zurück, bis wir Colin McGinn sich haben aussprechen hören.

---

<sup>37</sup> In: *The Journal of Philosophy*, Volume LXXXVI, No. 3, March 1979, 113-133.

Colin McGinn ist übrigens ein relativ junger Oxforder Philosoph, Schüler u. a. des berühmten Anti-Realisten Michael Dummett. Da die Psychologie doch auch ihren Anteil an der Ausbildung einer philosophischen Position hat, ist klar, dass der Schüler dem Meister widersprechen und ein ‚apriorischer Realist‘ werden muss. Colin McGinn hat aber vor allem über geistphilosophische Themen gearbeitet, und zwar auf eine besonders anregende Weise. Publiziert hat er u. a. *The Character of Mind* (1981), *The Subjective View* (1982), *Wittgenstein on Meaning* (1984), *Mental Content* (1989) und *The Problem of Consciousness* (1990). Vor seinem Wechsel an die Rutgers University in New Jersey war Colin McGinn Wilde Reader in Mental Philosophy an der University of Oxford. Diesen Wechsel begründet er im Vorwort *The Problem of Consciousness* ergötzlich mit den Auswirkungen der Thatcher’schen Politik im allgemeinen und britischen Kulturpolitik im besonderen – eine Politik, die so viele kluge jüngere englische Philosophen in die neue Welt fast wie in die Emigration vertrieben hat (fast nur noch die alten Strawson und Dummett selbst, beide schon jenseits der Emeritierungsschwelle, der erste schon fast jenseits von Gut und Böse, halten einsam die Fahne in Oxford hoch). Colin McGinn schreibt also:

On the negative side, I should like to record my contempt for the educational policies of the present British government, which have done so much harm to philosophy in Britain – as to much else. I must also express my regret at the treatment accorded to me by the General Board of Oxford University (1989), which greatly impeded the completion of the present work (l. c., viii).

Colin McGinn lebt seither in Manhattan.

Wie sein Lehrer Michael Dummett hält Colin McGinn Realismus und Anti-Realismus für semantische Thesen. Dabei folgt er näherhin der Intuition, die Dummett vom frühen Wittgenstein übernommen hat, dass Bedeutung

etwas mit den Wahrheitsbedingungen von Sätzen zu tun hat: Einen Satz verstehen, heißt dann: wissen, unter welchen Bedingungen er wahr ist (vgl. *Tractatus* 4.024: „Einen Satz verstehen, heißt wissen, was der Fall ist, wenn er wahr ist.“). Nach dieser Sprachregelung kann man sagen: Der Realist nimmt an, dass die Wahrheitsbedingungen als solche die Bedingungen der Kenntnis transzendieren, *wie* ich den fraglichen Satz denn verifizieren oder falsifizieren kann („An A Priori Argument for Realism“, 128,2). Aus dieser Wissens-Transzendenz der realistisch gedeuteten Wahrheit zieht der Skeptiker seinen Zweifel. Er fragt: wie, wenn ich auf diese Weise die Wahrheit, wie sie an sich ist, niemals in meinem Bewusstsein darstellen könnte? Aber um Skeptiker zu sein, meint McGinn (mit Thomas Nagel), muss ich mich bereits innerhalb einer realistischen Semantik bewegen: Denn wenn ich die Bedeutung meiner Sätze gleich auf das mir Bekannte (als ihre Wahrheitsbedingungen) einschränke, kann ich keine skeptischen Fragen stellen. Denn es macht dann keinen Sinn, das mir Bekannte von einer mir prinzipiell unbekanntem Instanz her in Zweifel zu ziehen (128 f.). Anders als sein Schüler McGinn, hatte Dummett die Einstellung des Anti-Realisten für die einzig einleuchtende erklärt (und den Skeptizismus abgewiesen). Seiner Meinung nach macht es keinen Sinn, die Wahrheit über die Bedingungen ihrer (An)erkennbarkeit hinaus zu verlegen. „Dummett suggests instead that we might do better to adopt a verificationist or assertability-conditions theory of meaning, thus relinquishing (as he thinks) realism about the subject matter of the relevant statements“ (130,1). Colin McGinn möchte dagegen zeigen, dass auch der Anti-Realist – ähnlich dem Skeptiker – die Wahrheit des Realismus je immer schon anerkannt haben muss, wenn er sein eigenes Geschäft erfolgreich durchführen will. Das ergebe sich schon aus der wohlverstandenen „Definition“ der beiden Standpunkte. Da man ‚a priori‘ Sätze nennt, deren Wahrheit aus dem Verständnis ihrer bloßen Definition

fließen, hält Colin McGinn sein Argument gegen den Anti-Realismus mit hin für ein apriorisches (129).

Wieder im Einklang mit Dummett befindet sich McGinn dagegen, wenn er die Rede von *dem* Realismus oder *dem* Anti-Realismus für sinnlos erklärt. Realist (bzw. Antirealist) ist man jeweils hinsichtlich eines bestimmten ontologischen Bereichs – also etwa Realist in Bezug auf die unabhängige Existenz von Bewusstsein und Antirealist hinsichtlich des Seins der so genannten Außenwelt. Oder man ist Realist hinsichtlich theoretisch angenommener Entitäten (wie Wellen oder Elektronen) und zugleich Realist hinsichtlich makroskopischer physischer Körper (wie Tische und Felsen). Oder man kann Realist in Bezug auf Zahlen *und* auf abstrakte Gegenstände sein und gleichzeitig semantischer und mentaler Antirealist. Wenn wir mit dem Durchprobieren solcher Paarbildungen fortfahren, werden wir schnell merken, dass sie in genau zueinander passenden Entsprechungen bestehen. So stoßen wir endlich auf eine Logik wechselseitiger Abhängigkeiten zwischen Typen von Realismen und Antirealismen. Und in diese Abhängigkeiten lässt sich eine Ordnung ausmachen. Gehen wir nämlich davon aus, dass man jeweils Realist bzw. Antirealist hinsichtlich der materiellen Außenwelt bzw. der inneren Welt geistiger Zustände und Ereignisse sein kann, so haben wir eine Vierheit möglicher Kombinationen. Aussagen über Psychisch-Mentales nennt McGinn einfachheitshalber P-statements (sagen wir: Psycho-Aussagen), solche über materielle Gegenstände M-statements (nennen wir sie: Materi[al]e-Aussagen). Dann haben wir die folgenden vier Kombinationen: (i) Antirealismus sowohl hinsichtlich M- als auch hinsichtlich P-Aussagen (eine Position, die offenbar absurd ist); (ii) Antirealismus bezüglich M-Aussagen verbunden mit Realismus bezüglich P-Aussagen (etwa die Position Berkeley oder Fichtes); (iii) Realismus hinsichtlich M-

Aussagen kombiniert mit Anti-Realismus hinsichtlich P-Aussagen (die Position Sartres); schließlich (iv) Realismus sowohl in Bezug auf M- wie auf P-Aussagen (McGinns eigene Position, auch die von Thomas Nagel [114]).

Die Strategie von McGinns Text besteht nun darin, (iv) als triftig zu erweisen, aber nicht direkt, sondern durch den Nachweis, dass die drei Alternativkombinationen inkonsistent sind. Da (i) bis (iv) alle denkbaren Möglichkeiten erschöpfen, würde dann folgen, dass (iv) als die einzige sinnvollerweise beziehbare Position übrigbleibt. McGinn sagt:

Das Argument ist mithin indirekt, wie ein Beweis durch *reductio [ad absurdum]*. Denn wir stellen einen kombinierten Realismus auf, indem wir dessen Negation erwägen und so auf drei Möglichkeiten kommen, deren jeder wir eine Art von Inkonsistenz nachweisen. Es wird kein direktes Argument für den Realismus in beiden Bereichen gesucht (obwohl wir, denke ich, sehen werden, daß das mutmaßliche Ergebnis intuitiv stark einleuchtet): vielmehr werden gewisse wechselseitige Abhängigkeiten (interconnections) gefordert und systematisch ausgebeutet. Der Resultat ist (gemeint als) ein A-priori-Argument für den Realismus über M- und P-statements (114).

Dem Argument selbst schickt McGinn eine Vorbemerkung zu den Standpunkten des Realismus und Anti-Realismus voraus. Beide, sagt er (unter Berufung auf Dummett), sind konträre Einstellungen zur Beziehung zwischen Wahrsein und Als-wahr-Erkennen. Der Realist verlangt nur Wahrheits-Bedingungen, der Anti-Realist will auch Verifikations-Bedingungen kennen: also Weisen, wie er sich epistemisch von der Wahrheit einer Behauptung überzeugen kann (das sind, nach Deweys bekannter Wortschöpfung, „assertability or verification conditions“ [115]). Wir wissen schon: der Kern der realistischen Sicht ist, dass die Wahrheitsbedingungen in einem bestimmten Sinn die Annehmbarkeits-Bedingungen transzendieren, also darauf nicht reduziert werden können. Dagegen nimmt der Anti-Realist an, die Sache, um die es zu tun ist, dürfe nie die Gründe transzendieren, kraft

deren ich sie allererst als richtig einsehen (also anerkennen) kann – es wäre sonst, wie wenn man mich um Zustimmung zu einer These bitten würde, die ich grundsätzlich nicht (und nie) überprüfen kann (oder als ob ich Behauptungen aufstellen dürfte, für deren Begründetheit ich mich nicht stark machen müsste). So fordert die realistische Unterstellung erkenntnis-transzendenter Wahrheiten die anti-realistische Skepsis geradezu heraus (diesen Punkt hatte, wie Sie sich vielleicht erinnern, schon Thomas Nagel geltend gemacht). Angewandt auf unsere vier Kombinationen: Realismus bezüglich Aussagen über M[ateriales] besteht in der These, dass die Wahrheitsbedingungen dieser Aussagen die *Erfahrungsgrundlagen*, kraft deren die Behauptung erging, transzendieren – und genau das leugnet der entsprechende Anti-Realismus. Und Realismus bezüglich P[sycho]-Aussagen ist die These, wonach die Wahrheitsbedingungen die *behavioralen* Gründe transzendieren, kraft deren die Behauptung erging – und genau das leugnet der zugeordnete Anti-Realismus (115 f.). „Die realistischen Thesen implizieren so eine Art von Erkenntnis-Transzendenz bezüglich der Wahrheit der fraglichen Aussagen; und genau wegen dieser epistemischen Kluft zwischen Ausweis (evidence) und Wahrheit kommen die skeptischen Probleme hinsichtlich Existenz und Natur der Außenwelt und anderer Geister auf“ (116).

Nach diesen begrifflichen Vorverständigungen können wir uns an die schrittweise Prüfung der vier kombinierten Positionen machen. Die erste besteht, wie Sie sich erinnern, im kombinierten mentalen und physischen Anti-Realismus. Es ist schwer, diese Position auch nur ernst zu nehmen – und indem wir das sagen, haben wir eigentlich schon vorab zugestanden, wohin Colin McGinn uns allererst bringen will: nämlich dass es absurd (selbstwidersprüchlich) wäre, nicht *irgendeine* Form von Realismus zu ver-

treten (das hatten wir schon ganz zu Beginn dieser Vorlesung zugegeben). Es wäre die Position eines, der sein ‚Sach‘ buchstäblich auf Nichts gestellt hätte – wie vielleicht der Sophist Gorgias, der ja vertreten haben soll, es gebe nichts; wenn es etwas gäbe, wäre es nicht erkennbar; wäre es erkennbar, wäre es nicht mitteilbar. Sage ich nämlich: weder Geist noch Materie existieren unabhängig, so muss ich das Sein des einen im anderen begründen. Entbehren beide der unabhängigen Existenz, so hebt sich die Position auf.

Colin McGinn macht mit ihr weniger kurzen Prozess. Er identifiziert die beiden Anti-Realismen zunächst als kombinierten Phänomenalismus und Behaviorismus (116). Phänomenalismus – das kennen wir – ist die Position, wonach die physische (materielle) Welt nur in Erscheinungen besteht (bzw. auf der subjektiven Seite, ausschließlich durch Sinnes-Erfahrungen verifiziert werden kann); und Behaviorismus ist die Position, wonach Geistig-Psychisches ausschließlich an körperlich sinnenfälligem Verhalten verifiziert werden kann. Colin McGinn wählt wieder eine Abkürzung und nennt die Sätze über behaviour (also sinnenfälliges Verhalten) B-statements und solche über Erfahrungen (experiences) E-statements. (Im mündlichen Vortrag ist es geschickter, die Abkürzungen auszusprechen; ich sage also ‚Erfahrungssätze‘ und ‚Verhaltens‘- oder ‚Behaviour-Sätze‘). Die These des doppelten Anti-Realismus (i) besagt dann, dass Aussagen über Materielles (M-statements) bzw. Psychisches (P-statements) wahr gemacht werden durch solche über Erfahrungs- bzw. Verhaltens Befunde (E[rfahrungs]- bzw. B[ehaviour]-statements). Schließlich (noch einmal) ist der Anti-Realist einer, der Wahrheitsbedingungen durch Verifikationsbedingungen ersetzt und mithin Realitäten wie Geistiges oder Physisches auf Zugangsweisen zu Geistigem oder Physischem reduziert. Und diese Zugangsweisen werden eben durch B- bzw. E-Sätze ausgedrückt. (Sie werden vielleicht fragen:

wieso durch Sätze? Sind Zugangsweisen nicht zunächst mal Bewusstseinszustände? Ich bin ganz Ihrer Meinung – aber Colin McGinn ist nun mal ein Sprachanalytiker. Da er auch die erkenntnistranszendente ‚Wahrheit‘ durch ‚statements‘ wiedergibt, kann man hier natürlich erst recht fragen, ob Sprache sinnvoll als was Außermentales und gar noch als was unserer Auffassung Transzendentes beschrieben werden kann. Aber ich will diesen Einwand hier nicht vertiefen.) Kurz: der Anti-Realist *reduziert* Klassen von Sätzen auf andere Klassen von Sätzen – genau das ist sein Job. McGinn sagt: „For both anti-realisms, then, the important point is that the truth of a statement of a given class consists in nothing other than the truth of some statement of the reducing class“ (117,<sub>1</sub> u).

Sätze über Geistiges werden also auf Verhaltensbeobachtung und Sätze über Körperliches auf sinnliche Erfahrungen begründet (116, Anm. 3).<sup>38</sup> Bevor er an der Nachweis der Unhaltbarkeit dieses Beginnens sich macht, empfiehlt McGinn die Unterscheidung dreier Weisen, wie diese Reduzierbarkeits-Beziehung verstanden werden kann. Die reduzierende Klasse kann 1. das Bivalenzprinzip aufrechterhalten: dann wird die Wahrheit von Sätzen einer Klasse auf die Wahrheit von Sätzen einer anderen reduziert. Sie kann 2. neutral sein hinsichtlich des Bivalenzprinzips (also weder die Wahrheit noch die Falschheit der zu reduzierenden Sätze verbürgen). Und sie kann schließlich (3.) das Bivalenzprinzip auf die Weise respektieren, dass sie Sätze, die somit unentscheidbar hinsichtlich ihrer Wahrheit oder Falschheit sind, für bedeutungslos erklärt. Die Wahl hängt davon ab, für welchen Typ von Anti-Realismus man sich entscheidet und vor allem: wie man die Logik der reduzierenden Klasse bestimmt. Klassischerweise, sagt

---

<sup>38</sup> Verwirrenderweise sagt McGinn, Erfahrungssätze seien selbst eine Unterklasse von Sätzen über Physisches (116 u.), so wie Verhaltenssätze eine Unterklasse von Sätzen über Geistiges seien (117 o.). Aber Erfahrungen sind doch selbst geistig, und Verhalten wird doch selbst sinnlich wahrgenommen.



McGinn, enthalten die reduzierenden Sätze von Phänomenalismus und Behaviorismus kontrafaktische Konditionale, also Sätze wie: ‚Wäre *a* nicht der Fall gewesen, so wäre auch *b* nicht eingetreten (oder wahr)‘. Denn hielt man sich allein an *wirkliche* (aktuelle) Sinneserfahrungen, würde der Wahrheitswert der von ihnen reduzierten M-Sätze nicht immer genügend festgelegt (determiniert) – und entsprechendes gilt für das Verhältnis von Sätzen über Körperliches und B-Sätze.

Nun kommt McGinns Einwand. Er lautet auf Zirkularität im Verfahren. Denn die Konsequentien der Konditionale seien nur aufrechtzuerhalten, wenn die Antezedentien Bedingungen genau der Art einführen, die als reduzierbar angenommen wurden – also je nachdem Bedingungen für physische Gegenstände oder für Geistiges (118<sub>1</sub>). Und dann trete folgende Situation ein: Da keines der beiden Glieder des Konditionals eine unabhängige (also keine im realistischen Sinne sich selbst tragende) Wahrheit hat,<sup>39</sup> sondern erst durch die Sätze der reduzierenden Klasse wahr gemacht werde, komme keinem eine unabhängige Wahrheit zu. Es ergeht ihnen also ungefähr wie in Heines Gedicht den beiden Polen Crapülinski und Waschlapski: „Speisten in derselben Kneipe,/ Und da keiner wollte leiden,/ Daß der andre für ihn zahle,/ Zahlte keiner von den beiden“ (*Zwei Ritter*). Anders gesagt: die aufgestellten Konditionale sind nicht geradehin (realistisch) wahr, sondern bedürfen zu ihrer Bewährung genau der (kategorischen, also nicht-konjunktivisch formulierten) M[aterie]- und P[sycho]-statements, die zu reduzieren sie bestimmt sind (118<sub>1</sub>).

---

<sup>39</sup> D. h.: nicht „*barely* true“ im Sinne Dummetts ist. ‚Geradehin wahr‘ ist, was nicht durch einen anderen (grundlegenderen) Satz wahr gemacht wird (118, Anm. 7).

Ich gebe Ihnen Colin McGinns Konsequenz aus dieser Beobachtung in Übersetzung:

Ist einmal die allgemeine Form des phänomenalistischen und des behavioristischen Anti-Realismus klar herausgearbeitet, dann wird ziemlich evident, warum die beiden Lehren nicht gemeinsam vertreten werden können. Der Grund ist einfach, daß sie konkurrierende Vorschläge darüber machen, aus welchen statements die Grundwahrheiten (basal truths) bestehen sollen: Der Phänomenalismus nimmt Erfahrungssätze (E[rfahrungs]-statements), also eine Unterklasse von P[psychologischen]-statements, als grundlegend, während der Behaviorismus B[ehaviour]-statements, also eine Unterklasse von M[ateriellen]-statements, als grundlegend ansetzt. Daraus ergibt sich folgendes: während eine der beiden reduktiven Thesen ein statement als nicht selbst eine Reduktionsoperation erfordern darstellt, besteht die andere darauf, daß eine solche Reduktion durchgeführt wird. Da ein statement, das für den einen der beiden Anti-Realismen grundlegend ist, für den anderen abgeleitet (derivative) ist, ist es klar genug, daß durch die Verknüpfung der beiden Lehren ein fehlerhafter Regreß ausgelöst wird; und das frustriert in der Tat den Reduktionsehrgeiz, der in beiden anti-realistischen Thesen steckt. Genauer: Angenommen, wir haben ein M[aterielles]-statement *s*. Dann wird der Phänomenalismus als das es reduzierende statement eine E[rfahrungs]-Wahrheit *s'* liefern. Da aber *s'* seinerseits ein P[psychologisches]-statement ist, bietet der Behaviorismus ein B[ehaviorales]-statement *s''* auf für dessen Reduktion. Nun ist *s''* selbst ein M[ateriales]-statement; also verlangt es vom Phänomenalismus zu seiner Reduktion ein weiteres E[rfahrungs]-statement *s'''*. Und so beginnt der Kreis aufs neue, *ad infinitum*. Mithin widersprechen einander die beiden Lehren im Blick auf die Frage, welche statements welche anderen statements wahr machen. Daraus folgt, daß Position (i) nicht eingenommen werden kann. Und allein das schon legt eine Einschränkung jedes globalen Möchtegern-Anti-Realismus nahe; es sieht so aus, als sperre sich da *etwas* irreduzibel Realistisches in unserer Sprache und in unserem Weltsystem (118 f.).

Ich glaube: die Widerlegungs-Strategie von Colin McGinn ist trotz einigen technischen Raffinessen ziemlich klar. Und wir haben ja unsererseits schon zu Vorlesungsbeginn eingeräumt, dass ein totaler Anti-Realismus im Wortsinne bodenlos, also unbegründbar sein würde. Er müsste einen ontischen Bereich in einem anderen fundieren, der selbst des unabhängigen Seins entbehrte: eine Absurdität. *Etwas* also, das können wir Colin McGinn gern

zugeben, muss existieren, damit ich auch nur ein anderes darauf reduzieren wollen kann.

Aber warum sollte das nicht das Psychisch-Geistige sein, wie Berkeley oder Fichte das annehmen? Dann wäre doch der Zirkel vermieden. Dafür spricht die Beobachtung, dass Behaviour-Sätze ihrerseits im Grunde Erfahrungssätze sind. Sage ich (wie McGinn es tut), dass körperliches Verhalten, durch welches der Behaviorist psychische Zustände verifiziert, seinerseits durch Erfahrung (also was Psychisches) verifiziert wird, dann scheint doch das Psychische als die letzte reduzierende Klasse übrig zu bleiben. Das war z. B. Carnaps Auffassung, den McGinn in einer Fußnote zitiert (119, Anm. 9). Und natürlich war es die Position des frühen Ayer und (irgendwie auch) des späten Castañeda.

Wir müssen daraus aber nicht vorschnell einen Einwand gegen McGinn machen. Denn gerade die Positionen des kombinierten mentalen Realismus und physischen Anti-Realismus (und umgekehrt) werden ja nun ausführlich von ihm selbst diskutiert.

Charakteristisch für diese Positionen ist also (gegenüber der ersten) die Kombination von Realismus und Anti-Realismus. Der materialistische Realist besteht darauf, dass die Wahrheit von Sätzen über Materiales völlig unabhängig ist von der von Erfahrungssätzen. Und umgekehrt wird der mentale Realist die Verhaltensbeobachtung nicht als Wahrheitskriterium anerkennen (120). Anders gesagt: Das Sein von materiellen Tatsachen ist nicht ihr Wahrgenommenwerden; und psychische Tatsachen haben ihr Sein nicht im beobachtbaren Verhalten. Man kann auch noch genauer werden und sagen: Fürs Vorliegen einer materiellen Tatsache sind Erfahrungsbefunde

keine *hinreichende* Bedingung; und umgekehrt ist die Wahrheit von Sätzen über Materielles keine *notwendige* Bedingung dafür, dass ein entsprechender Erfahrungssatz wahr ist. Die typische Erklärung, die der Realist für die Beziehung beider ontischer Bereiche bereit hält, ist die, dass die Erfahrung durch die materielle Wirklichkeit *verursacht* wird (oder umgekehrt dass Mentales – z. B. eine EntschlieÙung – Physisches herbeiführt [120 f.]). Verursacher und Verursachtes sind aber „ontologisch distinkt“ (121 o.). Wäre die anti-realistische Ansicht die richtige, bestünde also die Wahrheit von materialen Sätzen in der von Erfahrungssätzen (und die von psychischen Sätzen in Verhaltensbeobachtungs-Sätzen), dann sähe man nicht mehr, wie Sätze der ersten Art solche der zweiten *erklären* könnten, meint McGinn (121,1 u.). (Dieser Einwand überzeugt aber natürlich nur den, der – anders als Putnam – die Kausalität für eine der physischen Welt realiter eingebildete Struktur hält, was sie nicht ist.)

Wie dem auch sei: Wenn wir die realistische Unabhängigkeits-These (verbunden mit der Kausaltheorie) zugrunde legen, sind materielle Körper von Sinneserfahrungen (durch die sie verifiziert werden) wie Ursache von Wirkung unterschieden. Und das entsprechende gilt für das Verhältnis von Psychischem und dem Verhalten, wodurch es (sinnenfällig) verifiziert wird. Dann müssen beide Bereiche aber ‚symmetrisch unterschieden‘, also auch ontisch voneinander unabhängig sein. Und das heißt: sie müssen beide unabhängig voneinander real sein (122). Ist eins das Explanandum des anderen, so kann dieses mithin aufs Explanans nicht reduziert werden. Und daraus allein scheint schon zu folgen, dass wir einen doppelten (mental und physikalischen) Realismus vertreten müssen. „In short, the formulation of each realism in terms of independence seems, on the face of it, to imply an

equally realist interpretation of the statements that comprise the assertability conditions of our given statements“ (122,2 u.).

Dies Argument geht aber ein bisschen zu schnell. Vor allem übersieht es ein entscheidendes Manöver, das dem partiellen Realisten (und also partiellen Anti-Realisten) noch offen steht (122 f.). Er wird nämlich folgendes Rezept empfehlen: ‚Fang an mit einer unabhängig-realistischen Formulierung über Materielles, Kausalität und Erklärung, wobei du die dazu erfordernten E[rfahrungs-]Sätze einführst. Beziehe dann diesen gegenüber einen anti-realistischen Standpunkt, und verlange, daß sie als wahr nur kraft geeigneter B[ehaviour-]Sätze konstruiert werden – so daß die für die realistische Formulierung der M[aterie-]Sätze gebrauchten E[rfahrungs-]Sätze ihrerseits auf eine Unterklasse von M[aterie-]Sätzen reduziert werden [Behaviour ist ja was Physisches]‘ (123,1). Entsprechend soll der mentale Realist verfahren. Er wird zunächst betonen, dass die Realität seines Seelenlebens unabhängig von behavioristischer Beobachtung feststeht. Dann aber wird er die eingeführten B[ehaviour-]Sätze einer phänomenalistischen Reduktion unterziehen, „thus rendering them true in virtue of appropriate E-statements. The first philosopher is a behavioristic realist about the external world, the second a phenomenalist realist about the internal world. Are these not perfectly consistent positions? I think that ultimately they are not, but the reason is somewhat subtle“ (l. c.).

\*

Untersuchen wir nun zunächst die Position (ii). Es ist die des materialistischen Realisten und psychologischen Antirealisten. Er reduziert Psychisches auf Sätze über Verhaltensbeobachtung. Gleichzeitig muss er konse-

quenterweise die Unabhängigkeit von Verhalten (das ja ebenfalls materiell ist) von Erfahrungssätzen behaupten (die ja psychisch sind). Um das tun zu können, muss er auf einen Bereich bislang noch nicht reduzierter Sätze zurückgreifen, und das sind eben die Erfahrungssätze. Da Erfahrungen aber wieder psychisch sind, scheint jetzt das Spiel von vorne loszugehen, und er scheint sich im Zirkel zu bewegen. Dem partiellen Realisten des Seelenlebens geschieht das gleiche, nur andersherum. Er wählt Erfahrungssätze als reduzierende Klasse und hält sie (weil sie ja psychisch sind) für unabhängig von Materie, also für in sich real. Konsequenterweise muss er sie auch für unabhängig von B[ehaviour]-Sätzen erklären (denn Behaviour ist physisch). Diese verlangen also ihrerseits nach phänomenalistischer Reduktion, und so kehrt der Zirkel mit seiner unabschließbaren Abfolge von B- und E-statements wieder. Aber dieser Zirkel, sagt Colin McGinn, scheint nicht unbedingt fehlerhaft (vitiös), wie es der *circulus des doppelten Antirealismus* war. Denn die gemischten Positionen (ii) und (iii) sind vor Selbstaufhebung geschützt dadurch, dass sie jeweils ein Standbein auf festem realistischen Boden haben (124<sub>1</sub>). Und der ist ja kraft seiner Unabhängigkeit von der reduzierenden Klasse vor Vernichtung geschützt, während er seinerseits die Sätze der anti-realistisch angesetzten Klasse auf die eigenen reduzieren kann (also etwa im Falle des mentalen Realisten Physisches auf Sinneserlebnisse). Trotzdem, meint McGinn, kann der Regress-Verdacht durch ein Zusatzargument erhärtet werden. Dies führt dann zur Widerlegung dieser Mischpositionen.

Sie geht aus von der Beobachtung der für jeden Realismus charakteristischen Kluft zwischen Wahrheits- und Verifikations-Bedingungen. Die Wahrheit ist für den Realisten erkenntnistranszendent, also kognitiv unverfügbar. Kraft dieser Transzendenz kann der Skeptiker jederzeit den Fuß in

die realistische Türe stellen. Er kann etwa sagen: gerade weil die Dinge, wie sie an sich sind, in der Weise, wie wir sie in unserer Erfahrung auffassen, nicht ankommen, kann ich sie nur als Gesamt von Erscheinungen erfassen. Umgekehrt kann der mentale Realist sagen: wie es in seinem Inneren wirklich aussehe, das könne der Behaviorist von außen nie sicher oder nie angemessen wissen; so erklärt sich z. B. die Möglichkeit eines Skeptizismus hinsichtlich fremder Subjekte. Die Frage ist nun: gibt es einen Test, der zeigt, dass solche antirealistischen Umformulierungen das realistische Programm noch angemessen zum Ausdruck bringen? Hier ist er: wir negieren die so umformulierte realistische Ansicht und sehen dann zu, ob die sich daraus ergebenden Sätze die Bedingungen einer angemessenen Formulierung des *Anti-Realismus* für die beiden Bereiche erfüllen (124<sub>2</sub>).

Die Antwort lautet (absehbarerweise) ‚nein‘. Sie ergibt sich aus dem Durchspielen der beiden Alternativmöglichkeiten. Die erste ist der Antirealismus hinsichtlich von M[atérielle-Welt-]Sätzen (M-statements). Ihm zufolge reduziert sich Materielles auf (evidente) Erfahrungen (und damit scheint unsere Kenntnis der Außenwelt skeptis-frei gesichert). Aber behaviorale Tatsachen sind unter den materiellen Tatsachen der Außenwelt in keiner Weise ausgezeichnet und können so die erkenntnistheoretisch unproblematische Fundierung nicht leisten, die der Anti-Realismus uns versprochen hatte. Die andere Alternativmöglichkeit ist der mentale Anti-Realismus. Ihm zufolge haben wir Kenntnis fremder Subjekte nur aufgrund öffentlich zugänglicher Verhaltens-Tatsachen. Wenn aber behaviorale Tatsachen auf Erfahrung-Tatsachen reduziert werden, dann ist der vom Anti-Realisten gewünschte erkenntnistheoretische Fortschritt übers Mentale nicht besonders eindrucksvoll. Denn diese Erfahrungen sind gerade so privat und unzugänglich wie die P[sycho-]Sätze, die sie reduzieren sollten (125). So führt

die Negation der vorgeschlagenen Umformulierungen des mentalen und materialen Realismus nicht zu konsistenten Formulierungen der entsprechenden Anti-Realismen. Aber die Umformulierungen stellen auch die Realismen selbst schief dar. Denn wenn wir uns einen unproblematischen erkenntnistheoretischen Zugang zu Verhaltens-Tatsachen zuschreiben (als Ersatz für die unmittelbare Erlebnisevidenz der inneren Erfahrung), müssen wir die Frage nach den Bedingungen unserer Kenntnis der Außenwelt *schon gelöst haben* – sonst berauben wir uns einfach der Grundlage zum Stellen dieser Frage (l. c.). Und wieder umgekehrt: Wenn wir versuchen, alle Verhaltens-Tatsachen als letztlich erfahrungsfundiert zu konstruieren, dann verfehlen wir die für den mentalen Realismus entscheidende Idee, dass psychische Tatsachen die *öffentlich* zugänglichen behavioralen Tatsachen transzendieren und von ihnen verschieden sind. Denn unter solch einer Reduktion stellen sich die Annehmbarkeits-Bedingungen („assertibility conditions“) für P[sycho-]Sätze selbst als genauso privat und subjektiv heraus wie die Sätze, deren Wahrheit sie bezeugen sollen. Ich zitiere McGinns Konklusion:

Dem Realismus der Außenwelt zufolge transzendieren M[ateriale-]Tatsachen ihre Annehmbarkeits-Bedingungen so, daß Information über das Gesamt der E[rfahrungs-]Wahrheiten nicht letztinstanzlich festlegen kann, welche M-Sätze wahr sind, oder auch nur, ob es überhaupt welche sind. Jemand, der solche Information hat, kann immer noch skeptisch gegenüber M-Sätzen bleiben. Ähnlich steht es um die Kenntnis / des Gesamt der B[ehavioralen-]Wahrheiten. Sie kann die skeptische Frage nicht verhindern, ob es da überhaupt mentale Zustände gibt, die das beobachtbare Verhalten verursachen. Selbst ein idealer Beobachter, jemand in einer Art göttlicher Stellung, könnte solcherlei skeptische Fragen aufwerfen, wenn Information nur über Annehmbarkeits-Bedingungen erfolgt. Wenn wir aber E[rfahrungs-]Tatsachen auf B[ehaviorale] Tatsachen reduzieren, gilt das nicht länger. Denn Kenntnis von E[rfahrungs-]Tatsachen wird eben in Kenntnis entsprechender B[ehavioraler-]Tatsachen *bestehen* und so den unproblematischen Zugang zum Bereich M[aterieller-]Tatsachen voraussetzen. Ähnlich stünde es, wenn unser idealer Beobachter im Besitz aller B[ehavioralen-]Tatsachen und diese auf die entsprechenden E[rfahrungs-



]Tatsachen reduziert wären. Dann hätte er nämlich die Frage schon entschieden, ob die Welt mentale Zustände über materielle Gegenstände hinaus enthält. Somit hätte keine der beiden Reduktionen die gewünschte Konsequenz, daß die Wahrheitsbedingungen der M- oder P-statements wirklich ihre üblichen Annehmbarkeitsbedingungen transzendieren. Darum gibt keine von ihnen eine adäquate Charakterisierung dessen, worin die Realität von M- und P-Tatsachen besteht. Die einfache Wahrheit ist dies – und dazu werde ich später noch was zu sagen haben –, daß eine echt realistische Aussage über materielle Körper die unreduzierte Anerkennung von Erfahrungen als demjenigen voraussetzt, von dem die Außenwelt per definitionem unabhängig ist; und daß ein echtes Verständnis des mentalen Realismus die unreduzierte Anerkennung von Verhaltens-Episoden als dem voraussetzt, wovon die Innenwelt unabhängig ist. Dieser Punkt springt noch lebhafter in die Augen, wenn man die erkenntnistheoretischen Konsequenzen aufspürt, die von einer angemessenen Formulierung des Realismus beider Bereiche erfordert wird. So denke ich, daß am Ende unser anfängliches krudes Argument aus der Unabhängigkeit anerkannt und die Positionen (ii) und (iii) als uneinnehmbar erklärt werden müssen (125 f.).

Die Auffassung dieser verschiedenen Alternativ-Kombinationen verlangt eine gewisse Aufmerksamkeit, obwohl sie im Grunde nicht schwierig ist. Ich gebe darum ein ganz umgangssprachliches Resümee. Der Realist setzt auf die Ausgangsthese, dass etwas unabhängig davon real ist, ob ein anderes sich von seiner Realität auch subjektiv überzeugen kann. So meint er, dass diese Felsnadeln an meiner Rechten auch dann existieren, wenn keiner sie sieht. Wenn er auch seelische Tatsachen für real hält, wird er das entsprechende auch von ihnen fordern: ich habe Liebeskummer auch dann, wenn andere das meinem Körperverhalten nicht ansehen können. Diese Formulierungen scheinen die Mischpositionen aus Realismus und Anti-Realismus als vernünftig nahezu legen. Aber Colin McGinn macht geltend, dass es keinen Sinn macht, von ‚Unabhängigkeit‘ (oder ‚Transzendenz‘) zu reden, wenn das, wovon ein Seinsbereich unabhängig ist, seinerseits nicht unabhängig existiert. Man kann das auch so ausdrücken: Wenn die Aussagen-Klasse, *auf die* ich eine andere reduzieren will, in sich nicht unabhängig besteht (oder seinerseits nicht geradehin wahr ist), so würde ich ja unab-

hängiges Sein auf ein relatives Nicht-Sein gründen. Das wäre, als wollte ich sagen: Ein Seiendes hat sein Sein in anderem, das aber selber keins hat und darum in einem dritten suchen muss – und so *ad infinitum*.

Damit bleibt nur noch die IV. Position übrig: die des doppelten (mental und physikalischen) Realismus. Wir sahen, daß Colin McGinn diese Position für die richtige hält. Die Eleganz seines Arguments besteht darin, dass er die idealistischen bzw. materialistischen Intuitionen der beiden zugeordneten anti-realistischen Konkurrenten in seine Formel einbauen kann. Denn er leugnet ja nicht, im Gegenteil: er behauptet selbst, *dass* es seelisch-geistige Phänomene *gibt*. Und ebenso wenig leugnet er, dass Seelisches sich umgekehrt materiell manifestieren muss, ja dass Materie unabhängig davon besteht, ob jemand sie wahrnimmt. Der mentale Realist und materialistische Anti-Realist kann in dieser Formel sehr gut seine Intuition unterbringen, 1. dass Geistiges *existiert*, und 2. dass ihm Materielles nur durch Erfahrungen hindurch *bekannt* wird. Nur widerspricht McGinn der antirealistischen Teil-These, dass es Materielles also als ein vom Subjekt Unabhängiges nicht gibt.

Um der Intuition des (partiellen) Anti-Realisten entgegenzukommen, lehnt McGinn die Formel des metaphysischen Realisten ab. Der will kurzen Prozess machen und empfiehlt, die Realität geradehin als Unabhängigkeit von Erfahrungs- und Verhaltens-Sätzen zu formulieren (126 f.). Er wird also näherhin sagen, die Sätze einer realistischen Theorie gehorchten dem Bivalenz-Prinzip; und ferner sei leicht einzusehen, dass Sätze, die geradehin gelten („are barely true“), das nicht tun können kraft Verifikation durch eine (von der ersten unterschiedene) Klasse anderer Sätze (127,2). Aber, meint

Colin McGinn, wir brauchen diese anderen Sätze einfach, und zwar zu einer intuitiv einleuchtenden Formulierung des Realismus selbst (127,1).

So kontert er dem eliminativen Realisten (der heißt eliminativ, weil er die Möglichkeit und Notwendigkeit von Verifikation einfach eliminiert) in zwei Schritten (127,2): 1. verfehlt die eliminative Formulierung gerade das Charakteristikum der realistischen Unabhängigkeitsthese. Etwas kann unabhängig nur sein von etwas anderem, das dann mitgenannt werden muss. Anders gesagt: Es muss die Kluft zwischen Wahrheits- und Anerkennungsbedingungen wirklich geben (und also auch beide Teilmglieder derselben), wenn der Realismus konsistent soll formuliert werden können. Und 2. gilt: auch wenn die gegebene Klasse nicht imstande sein sollte, geradehin Wahrheit aufzuweisen, folgt daraus noch nicht die Richtigkeit der antirealistischen Interpretation. „Nehmen Sie z. B. Physikalismus hinsichtlich mentaler Zustände und Elementarteilchen-Reduktion hinsichtlich makroskopischer materieller Körper. Ob wahr oder falsch, solche Thesen sind intuitiv realistisch für P[sycho-] und M[aterie-]Aussagen; nicht jeder Reduktionismus ist gleich anti-realistisch. Ich würde sagen, diese Lehren sind intuitiv realistisch, weil sie den Schnitt zwischen Wahrheit und Kenntnis (recognition) von Wahrheit nicht leugnen. Anders als bei Phänomenalismus und Behaviorismus vollzieht sich mit ihnen keine behauptete Reduktion auf assertability conditions. Eben darum, weil nackte Wahrheit (bare truth) keine notwendige Bedingung für eine realistische Interpretation ist, ist Unabhängigkeit von *allen* anderen Aussage-Klassen (classes of statements) eine gar zu strenge Forderung an eine solche Interpretation“ (l. c.). Und es wäre eine zu schwache, wenn der Realist diese Bedingung kurzerhand aufgeben wollte. Damit hätte er sich über die berechtigten Interessen der Erkenntnistheorie hinweggesetzt; aber auch ohnedies ist dieser Standpunkt

zu schwach. Bloßer Respekt vor dem Bivalenz-Prinzip macht, wie vor allem Dummett gezeigt hat, noch keinen Realisten (die kontrafaktische Formulierung, die dem Antirealisten offensteht, hält das Bivalenz-Prinzip ja gerade aufrecht). Der Unterschied liegt also anderswo: Der Antirealist leugnet und der Realist behauptet die Wichtigkeit der logischen Unabhängigkeit, der Kausalbeziehungen und der erklärenden Kraft (explanatory role), wie sie für die realistische Sicht von M[aterie-] und P[sycho-]statements charakteristisch ist.

This leads me to conclude that my initial formulation in terms of these features uniquely captures the content of the realist conception of the two types of statement. But if the argument that starts from that style of formulation is sound, we are left with global realism as the only really viable position (128<sub>1</sub>).

Die ‚Ausgangsformulierung‘ – erinnern wir uns kurz – hatte gelautet: Die für den Realismus konstitutive Überzeugung von der Unabhängigkeit der Wahrheit (von der Verifikation) bzw. der Gegenstände (von ihrer Erkenntnis) impliziere Realismus für Aussagen über Materielles *wie* für Psychisches (122). Die Unterschiedenen seien ‚symmetrisch‘ unterschieden. Und wenn ich das eine (das Explanandum) durch das andere (als sein Explanans) erklären wolle, müssen *beide unabhängig* voneinander existieren: „In short, the formulation of each realism in terms of independence seems, on the face of it, to imply an equally realist interpretation of the statements that comprise the assertibility conditions of our given statement“ (l. c.).

Diese Conclusio komme schließlich auch unseren dualistischen Grund-Intuitionen entgegen, meint McGinn. Wir meinen, es gebe eine objektive Welt von Gegenständen, die ein göttlicher Verstand so erfasse, wie sie an ihr selbst ist, und ihr gegenüber eine subjektive Sicht: eine Weise, in der einem irgendwie zumute ist und durch die sich die objektive (bewusst-

seinsunabhängige) Welt in eine solche von (bewusstseinszugewandten) Erscheinungen verwandle (131 f.). An dieser Sicht ist was dran – aber sie duldet keine anti-realistische Deutung. Subjektives kann ich nämlich nur im Kontrast zum Objektiven erfassen – und umgekehrt (132,2). Was soll der Ausdruck ‚subjektiv erscheinen‘ bedeuten, wenn die mitgedachte Alternative nicht lautet: ‚im Gegensatz dazu, wie es an sich *ist*‘? Aber auch das Subjektive muss irgendwie *sein*, und mehr sein als nur das, was von anderen Personen an meinem Körperverhalten abgelesen werden kann. Und umgekehrt muss die physische Welt mehr sein, als was in meine subjektive Wahrnehmung fällt. Aber das bedeutet nicht, dass Erfahrungen und Behaviour (also die Verifikationsbedingungen für Materie- bzw. Psycho-Sätze) *nicht existierten*. Es bedeutet im Gegenteil, dass sie wirklich die Verifikationsbedingungen der ihnen realistisch zugeordneten Wahrheiten *sind* (133: „it is an a priori truth that E- and B-statements constitute assertibility conditions for M- and P-statements“); und dass wir ontogenetisch wirklich durch Wahrnehmung und Behaviour *lernen*, was materielle Körper und fremde Seelenregungen *sind*. „It is this fact“, schließt McGinn, „that captures the element of truth in phenomenalism and behaviorism; but this alone“ (l. c.).

Wie sollen wir abschließend McGinn Apriori-Argument für den Realismus beurteilen? Es gibt in einem seiner anderen Bücher, in *Mental Content* (Oxford: Basil Blackwell 1989, 21991), einen charakteristischen Hinweis auf Jean-Paul Sartres *L'être et le néant* (22, Anm. 31). McGinn nennt Sartre dort „den vielleicht extremsten Externalisten in der Philosophie“. Externalismus ist die Ansicht, dass die Außenwelt einseitig die Inhalte unserer Gedanken bestimmt, egal, wie viel von dieser Bestimmung in unserem expliziten Bewusstsein ankommt. Wenn Aristoteles den Abendstern betrachtete und glaubte, er sei an eine bewegliche Schale geheftet, so war die Bedeutung

seines Glaubensgehalts trotzdem ‚die Venus‘ (auch wenn er noch gar nicht wusste, dass sie mit dem Abendstern identisch war). So weit geht Sartre nicht. Er unterscheidet aber das Sein-an-sich vom Für-sich-Sein oder Bewusstsein, das durch Nichtigkeit (nothingness) ausgezeichnet ist. Es ist eine substanzlose Leere, wie er sagt. Der Preis, den es entrichtet, um das Sein zur Erscheinung zu bringen, ist: selbst *nicht zu sein*. Was wäre aber radikaler dem Sein entgegengesetzt als das Nicht-Sein?

The nothingness of intentional consciousness, as Sartre understands it, makes consciousness a radically different kind of ‚being‘ from the world of objects; in effect, he is an anti-substantialist about the mind. His view is, indeed, strongly analogous to relationalism about space, since he conceives mental states as nothing (n.b.) over and above intentional relations to things in the world.

Die Beschreibung ist treffend. Sie kennen wahrscheinlich Sartre berühmtes ‚ontologisches Argument‘: Das zu den Evidenzen des Bewusstseins (seit Descartes) das Sein gehört, das Bewusstsein aber ein ‚unsubstantielles Leeres‘ ist, muss es sich auf dasjenige Sein hin überschreiten, das nicht ist. So ist es wesentlich Intentionalität (oder, mit einem Wort des Novalis: ‚Sein außer dem Sein beim Sein‘). Sartre ist also mentales Antirealist und ontologischer Realist.

So weit geht McGinn aber gerade nicht. Er glaubt: beide Sphären müssten aus logischen Gründen existieren. Diese Konsequenz scheint mir zu weit zu gehen. Es genügt in der Tat (wie in einem umgekehrten Sartreanismus), dass eine der beiden Sphären existiert und die andere kontrastiv von sich ausschließt: als Nicht-Seiendes (oder Erscheinung oder Nicht-Ich). Es scheint mir ein Fehlschluss zu sein, die richtige Beobachtung, dass Aussagen über Materielles und solche über Geistiges einander wechselseitig fordern, so zu überdehnen, als folge die Existenz des jeweils anderen *logisch*.

Es genügt, dass eines seine *Bestimmtheit* nur haben kann in Abhebung gegen das andere („omnis determinato est negatio“). Aber Bestimmtheit ist was Logisches und hat nichts mit Existenz zu tun. Auch Einhörner, die nicht existieren, haben eine Bestimmtheit und also einen konstitutiven Gegensatz. McGinns Apriori-Argument ähnelt dem ontologischen Gottesbeweis in der Version des Heiligen Anselm oder Descartes', wie Kant ihn ein für allemal widerlegt hat: Existenz ist kein ‚reales Prädikat‘ und folgt mithin nicht aus logischen Bestimmungen. Aber richtig ist und bleibt doch, dass der totale Anti-Realismus ein auch logisch unmögliches Projekt ist. Denn ich kann – wenn ich mit Existenz einmal a posteriori bekannt geworden bin – a priori einsehen, dass ich es nicht auf ein Nicht-Seiendes begründen kann. Und genau das müsste der doppelte Anti-Realismus (bei seinem Doppelverzicht auf ein unabhängiges Sein) tun.

### 13. Vorlesung

Colin McGinn vertritt – wie wir sahen – eine ziemlich harte Gangart von Realismus: einen solchen, den Hilary Putnam als ‚metaphysisch‘ bezeichnet hätte. Aber immerhin ist McGinns Realismus nicht eliminativ: Er hält Geistig-Seelisches für ebenso real für Materiell-Körperliches (und er zwingt uns in seinem Aufsatz nicht, *einen* ontischen Bereich auf den *anderen* zu reduzieren). Hierin kommt seine Position der ebenfalls geistfreundlichen seines Kollegen Thomas Nagel recht nahe. Und darum kann man diese Position nicht einfach als ‚Materialismus‘ etikettieren. Materialismus ist (wie wir früher sahen) ohnehin nicht dasselbe wie Realismus. Wenn wir unter ‚Materialismus‘ (oder, was dasselbe ist, ‚Physikalismus‘) die Überzeugung verstehen, dass nur die von den empirischen Naturwissenschaften benutzten Verfahren zu echten Erkenntnissen führen, dann können wir uns ganz leicht vorstellen, dass gerade auch eine Anti-Realist (wie Richard Rorty) sie teilt. Er wird ihr dann einfach die folgende pragmatische Wendung geben: Die Verfahren der Naturwissenschaft sagen uns nicht, was wahr oder falsch ist (das ist uns sowieso ziemlich egal). Sie geben uns die besten Werkzeuge zum Umgang mit der Wirklichkeit an die Hand (they are the best tools for handling reality); sie liefern die größt-mögliche Prognose-Sicherheit. Putnam nennt diese Position ‚Instrumentalismus‘, weil ihr zufolge wissenschaftliche Erkenntnis am instrumentellen oder praktischen Nutzen fürs Überleben der Gattung gemessen wird. Denkbar wäre durchaus, dass eine in diesem Sinne nützliche Theorie an sich falsch wäre. Aber, wie gesagt, der anti-realistische Materialist interessiert sich für ‚wahr‘ und ‚falsch‘ nur insoweit, als Sätze, die wir so charakterisieren, eine Theorie insgesamt pragmatisch ‚bewähren‘. Da aber Theorien unendlich viele Voraussagen möglich machen, ist schon daraus klar, dass sie sich nicht überprüfen lassen.



Und darum kann für den Realisten Putnam die Äquivalenz-Behauptung ‚Eine Theorie ist genau dann wahr, wenn alle mit Hilfe der Randbedingungen deduzierbaren Voraussagen wahr sind‘ nicht als Definition des Prädikats ‚wahre Theorie‘ in Frage kommen. Vielmehr spricht Putnam, bevor er die Wende zu einem schwächeren, einem so genannten ‚internen‘ Realismus vollzieht, von einer ‚externen‘ Wahrheit, ohne allerdings je genauer zu bestimmen, was externe Wahrheit *ist*. – Jedenfalls ist eine solch pragmatistische Wahrheitstheorie Colin McGinns Standpunkt genauso wenig wie der von Hilary Putnam. Beide sind Realisten. D. h.: beide vertreten eine These über die *Existenz* von materiellen *und* mentalen Gegenständen/Ereignissen und über die Verschiedenheit beider Bereiche (eine Verschiedenheit, die die Reduktion eines aufs andere und mithin die Position des globalen Materialismus ausschließt; Putnam hat der These im 3. Band seiner *Philosophical Papers*, Cambridge University Press, 1983, 1992, 229-247 einen Aufsatz mit dem Titel „Why reason can’t be naturalized“ gewidmet).

‚Metaphysisch‘ kann McGinns (und Putnams früher) Realismus heißen, weil er eine Einstellung zum Gesamt der Wirklichkeit bekundet, die als solche weder bewiesen noch widerlegt werden kann. Aber das ist – wohlbemerkt – per se noch kein Einwand: Die Sätze, die wir beweisen können, setzen allemal eine als verbindlich unterstellte Theorie voraus, innerhalb deren wir uns orientieren und die nicht abermals als die richtige bewiesen werden kann (und die Erfolgserprobtheit dieser Theorie scheidet, wie wir sehen, aus, weil sie ein anti-realistisches Kriterium ins Spiel bringt). Heidegger hatte diesen Rahmen, in dem wir unsere Welt interpretieren, ‚Erschlossenheit‘ genannt und von derjenigen Wahrheit unterschieden, kraft deren wir ‚innerweltliche Entdeckungen‘ machen. Bevor wir nämlich Sätze wahrheitsfunktional positiv oder negativ auf Sachverhalte *in* der Welt be-

ziehen, mussten wir schon eine voll interpretierte Sprache haben (von der es keinen Sinn macht, sie im gleichen Sinn ‚wahr‘ zu nennen wie die *kraft* ihrer möglichen Entdeckungen). „Nur in einer Sprache“, sagt Wittgenstein im genau gleichen Sinne, „kann ich etwas mit etwas meinen“ (*PhU*, in: *WA I*, 260 u.) Putnam spricht vom Realismus als von einem ‚Modell‘ oder ‚Bild‘ der Wirklichkeit: einem (im Wortsinne) Weltbild.<sup>40</sup>

*Metaphysical realism [...] is less an empirical theory than a model [...]. It is, or purports to be, a model of the relation of any correct theory to all parts of THE WORLD. I have come to the / conclusion that this model is incoherent. This is what I want to share with you (RaR 123 f.).*

Diese Äußerung findet sich in *Realism and Reason*, jenem Aufsatz von 1976 (gedruckt 78), in dem Putnam seine Wende vom metaphysischen zum (wie er jetzt sagt) internen (oder pragmatischen) Realismus öffentlich gemacht hat. Es wäre verführerisch, die Auseinandersetzung mit Putnams Standpunkt gleich mit dieser Wende zu beginnen – schon darum, weil die Texte seither so viel angenehmer zu lesen sind als die früheren, in denen er noch einen ziemlich knochenharten (metaphysischen) Realismus vertreten hatte. Aber wie soll man eine Wende beurteilen können, wenn man nicht weiß, von was weg und auf was zu sich orientiert?

Ich wähle als Ausgangspunkt (schwerpunktmäßig, wenn auch nicht nur) einen frühen Text (von 1960): „Do True Assertions Correspond to Reality?“<sup>41</sup> Der Text ist von einer trockenen, ungeschwätigen und manchmal arroganten Eleganz, wie sie für Putnams frühere Phase insgesamt charakteristisch ist – aber er ist eben (darum wähle ich ihn) charakteristisch für Putnams früheren Realismus insgesamt. Ich werde versuchen, das Skelett des

<sup>40</sup> „Metaphysical realism was only a *picture* anyway“, sagt Putnam. In: „Realism and Reason“, in: *Meaning and the Moral Sciences* (zit.: *RaR*), London: Routledge & Kegan Paul, 1978, 123-140, hier: 130.

<sup>41</sup> In: *Mind, Language and Reality. Philosophical Papers*. Volume 2, Cambridge University Press 1975, First paperback edition 1979, <sup>8</sup>1992. 70-84.

Gedankens freizulegen, ohne in den technischen *Détails* zu lange zu verweilen.

Wir hatten ganz zu Beginn dieser Vorlesung eine Reihe von notwendigen Voraussetzungen für die Rede von ‚Realismus‘ aufgestellt. Die wichtigste besteht in der Transzendenz (oder: Unabhängigkeit) der Welt gegenüber/von dem Bewusstsein, das wir von ihr haben. Auf der wissenschaftslogischen Ebene wiederholt sich dies Verhältnis in der Transzendenz (Unabhängigkeit) der Wahrheit gegenüber (bzw. von) unseren Verifikationsmethoden (Stichwort, noch einmal, ‚Anti-Instrumentalismus‘). Ein zweites Postulat garantiert die Existenz einer unabhängig von unserem Erkenntnisvermögen strukturierten Welt, einer „ready-made world“, wie Putnam später spöttisch übersetzt. (Einen Schritt weiter in dieselbe Richtung tut das Postulat, die Welt weise an ihr selber – also „mind-independent[ly]“ und unabhängig von den Klassifikationen einer Theorie – natürliche Arten auf [vgl. Nola, l. c., 5].) Die dritte Voraussetzung wäre eine Auffassung von Wahrheit als Bivalenz. Sie ist der instrumentalistischen Ansicht entgegengesetzt.<sup>42</sup> Denn nach den Instrumentalisten haben theoretische Sätze keinen Wahrheitswert und die darin enthaltenen theoretischen Ausdrücke (wie ‚kinetische Energie‘ oder ‚Elektron‘) keine Referenz. Sie dienen nur als Instrumente, als Hilfsmittel zur Gewinnung lebenserhaltender Voraussagen. Diese Voraussagen beziehen sich – je nachdem, ob der Instrumentalist hinsichtlich der beobachtbaren Gegenstände materialistisch oder phänomenalistisch denkt – auf zukünftige Ereignisse der materiellen Welt (der Alltagsobjekte) oder auf Regelmäßigkeiten in der menschlichen Erfahrung („regularities in human experience“, sagt Putnam in *Mathematics, Matter and Method*, vii). Dem stellt er – als ein wesentliches realistisches Credo – die

---

<sup>42</sup> Vgl. zum folgenden Alex Burri, *Realisms und Semantik bei Putnam*, Diss. Bern 1991, 28 ff.

Überzeugung entgegen, dass in einer wissenschaftlichen Theorie alle Sätze einen Wahrheitswert aufweisen, also entweder wahr oder falsch sein müssen (l. c.). Putnam gibt dem eine besondere Wendung im Blick auf die Gesetzmäßigkeit von Theorien: „Laws are statements that we hope to be *true*“ (l. c., 255). Putnam begründet die These (unter Berufung auf Richard Boyd) wie folgt: (1) In einer reifen (adult) Wissenschaft referieren die Ausdrücke normalerweise (typically); (2) die Gesetze einer Theorie, die zu einer reifen Wissenschaft gehören, sind gewöhnlich (typically) approximativ wahr (*Mind, Language and Reality*, 290). Wir müssen hier nicht auf Boyds Begründung eingehen. Dass Putnam sie anführt, zeigt jedenfalls, dass er mit der Dekretierung eines schlichten realistischen Credo nicht zufrieden ist, sondern die Bivalenz-These für Gesetze auf eine Begründungs- und Argumentations-Kette stützen möchte.

Es gibt noch zwei weitere (nur rasch in Erinnerung zu bringende) Postulate des (metaphysischen) Realismus: dass eine ideale Theorie die Wirklichkeitsstrukturen unserem Erkennen ganz oder in beträchtlichem Maße zugänglich machen kann und dass das Verhältnis von Erkennen zur Wirklichkeit als ein solches der Eins-zu-eins-Entsprechung von Sachverhalten und Sätzen zu verstehen ist (die berühmte Korrespondenz-Theorie der Wahrheit).

Diese Postulate hat der frühe Putnam selbst unterschrieben, besonders die These von der Bewusstseins-Transzendenz der Wirklichkeit. So nennt er in einem kleineren Text „A philosopher looks at quantum mechanics“<sup>43</sup> den idealistisch-phänomenalistischen Standpunkt gegenüber der Welt abwegig:

---

<sup>43</sup> In: *Matter, Mathematics and Method. Philosophical Papers*. Volume I, Cambridge University Press, 1975. Second edition, with additional chapter, 1979, <sup>6</sup>1993, 130-158, hier: 147.

To define a measurement as the apprehension of a fact by a human consciousness [...] would be to interpret quantum mechanics as asserting a dependence of what exists on what human beings happen to be conscious of, which is absurd.

Sie werden aber sofort die Spannung bemerkt haben, die zwischen der These der Bewusstseins-Jenseitigkeit der Welt und der anderen ihrer vollständigen Abbildbarkeit in einer idealen Theorie besteht. (Denn wenn die Realität als ganze einer Theorie sich erschließt, hat gleich der Anti-Realist den Fuß in der Tür; er wird fragen, warum er die Totalität aller bewährten Sätze noch um die Totalität aller Tatsachen verdoppeln muss, da doch die Erkenntnisse alles sind, was wir brauchen. Ich komme darauf nächste Woche zurück.<sup>44</sup>) Jedenfalls steht die These des Realisten auf dem Boden der Korrespondenz-Theorie der Wahrheit. Sie scheint eine unentbehrliche Voraussetzung des Realisten zu sein – und zugleich das schwächste Glied in der Kette. So verwundert es nicht, daß Putnam seine wachsenden Zweifel am metaphysischen Realismus von der folgenden Frage aus entwickelt (ich übersetze sie frei):

Das Problem, mit dem die vom metaphysischen (oder, wie Kant sagt, ‚transzendentalen‘) Realismus Überzeugten von jeher konfrontiert waren, ist der Begriff der ‚Korrespondenz‘. Es gibt viele verschiedene Weisen, Zeichen einer Sprache und Dinge einer Menge  $M$  miteinander in Entsprechung (correspondence) zu bringen (und eine sehr große endliche Zahl, wenn  $M$  eine große endliche Menge ist). Selbst wenn ‚Korrespondenz‘ eine Referenz-Beziehung sein und wir spezifizieren sollen, welche Sätze welchen *wirklich bestehenden Sachverhalten* entsprechen sollen, folgt aus dem Theorem der Modelltheorie, daß es / immer noch unendlich viele Weisen gibt, wie sich solch eine Korrespondieren *spezifizieren* läßt. Wie können wir irgendeine Entsprechung (correspondence) zwischen unseren Worten (oder Gedanken) und den vermeintlich geistunabhängigen Dingen herausgreifen, *wenn wir keinen direkten Zugang zu den geistunabhängigen Dingen haben?* (Die deutsche Philosophie begann fast immer mit einer besonderen Antwort auf diese Frage – nämlich der Antwort ‚wir können nicht‘ – nach Kant.) („Why there isn’t a ready-made world“, in: *Philosophical Papers II*, 206 f.).

---

<sup>44</sup> Vgl. Putnam, „Realism and Reason“, in: *Meaning and the Moral Sciences*, London 1978, 125 f.

Hier hat die Problematisierung schon zur Abweisung der Korrespondenz-Theorie geführt. (Wir müssten, um zu wissen, ob ein Satz mit der Wirklichkeit korrespondiert, sowohl innerhalb als auch außerhalb der Sprache stehen können, um beide Ansichten zu vergleichen.) In „Do true assertions correspond to reality“ ist er davon noch weit entfernt, auch wenn die Abweichung vom Korrespondenz-Dogma hier erste Symptome zeigt. In diesem frühen Aufsatz unternimmt Putnam nämlich den Versuch, die Schwächen der traditionellen Korrespondenz-Theorie durch (immanente) Verbesserungsvorschläge zu reparieren. Was die Lektüre so mühsam macht, ist, dass das nur ganz skizzenhaft geschieht. Man legt den Aufsatz etwas verwirrt aus der Hand, ohne sich das Urteil zu getrauen, ob der Versuch gelungen ist.<sup>45</sup> Putnam beschränkt sich im Wesentlichen auf zweierlei: Erstens argumentiert er gegen die (etwa von Popper vertretene) Annahme, dass Tarskis berühmte semantische Wahrheitstheorie eine Korrespondenztheorie sei. Und zweitens bemüht er sich um eine genauere Bestimmung des für die Korrespondenztheorie zentralen Entsprechungs-Begriffs.

Der erste Feldzug ist ein bißchen technisch und kompliziert; darum versuche ich, möglichst knapp und schlicht zu sein. Alfred Tarski hatte ja in den 40er Jahren, unter expliziter Berufung auf die alte aristotelische Definition der Wahrheit als Übereinstimmung der Gedanken und der Wirklichkeit, eine so genannte semantische Wahrheits-Definition vorgelegt. Viele haben in dieser Definition (die eben wirklich nur eine Definition sein will, keine Wahrheits-Theorie) eine Option auf die Korrespondenz-Theorie sehen wollen – und mit dieser Deutung setzt sich Putnam hier (und übrigens noch in vielen anderen Aufsätzen) auseinander. Nach Tarski wird ein Satz

---

<sup>45</sup> Vgl. Alex Burri, l. c., 32 ff.

einer bestimmten Sprache (z. B. ‚Schnee ist weiß‘), auf einen Sachverhalt abgebildet (oder vielmehr: in eine Äquivalenz-Beziehung ‚gdw‘ gebracht), der ihn wahr macht: hier den Sachverhalt, dass Schnee [wirklich] weiß ist. Der erste Satz (in Anführungszeichen) ist in der Objekt-Sprache, der zweite (ohne Anführungszeichen) in der so genannten Meta-Sprache formuliert. Aber wenn man genau hinguckt, handelt es sich beide Male um *Sätze* in je einer Sprache (der eine *in*, der andere *ohne* Anführungszeichen), und nicht – wie die echte Korrespondenz-Theorie verlangen würde – um eine Bewährung eines Satzes durch eine außer-sprachliche Realität (vgl. Putnam, l. c., 70). So bleibt die Definition von Wahrheit ohne alle Stütze durch die real existierende Wirklichkeit (71 o.). Gesagt wird nur, dass die Bedeutung des Wortes ‚wahr‘ darin besteht, dass Sätze auf der linken Seite ihr Äquivalent in solchen der rechten haben müssen, so, dass, wer das eine zugibt, auch das andere zugeben muss. So klärt Tarski gerade nicht die Beziehung zwischen Sätzen und außersprachlichen Tatsachen – wie es doch der Anspruch der Korrespondenz-Theorie ist. Tarski hat nur erklärt, was unter dem Ausdruck ‚wahr‘ zu verstehen ist. Putnams Kritik an der Auffassung, Tarski habe selbst einen Beitrag zur Bestimmung der Korrespondenz von Wirklichkeit und Sprache geleistet, ist viel komplizierter und raffinierter – aber ich darf es hier bei diesen vagen Andeutungen bewenden lassen (vgl. l. c., 72 f.).

Neben dieser Kritik geht es Putnam in „Do True Assertions Correspond to Reality?“ um eine Präzisierung dessen, was – Tarski hin oder her – unter ‚Korrespondenz‘ denn eigentlich zu verstehen ist. Die Präzisierung besteht (sehr grob gesagt) darin, dass die Entsprechung als die einzige ‚natürliche Relation‘ bestimmt wird, die zwischen Sätzen und möglichen Welten bestehen kann. Während Putnam den Begriff der möglichen Welt mit Hilfe von möglichen Massen- und Ladungsverteilungen in der vierdimen-

sionalen Raumzeit zu erklären versucht (also einer Art physikalistischer Elementar-Formel: 80 f.), probiert er, dem ausgesprochen vagen Ausdruck ‚natürliche Relation‘ durch eine Liste von Einschränkungen beizukommen, die eine als ‚natürlich‘ geltende Relation zu erfüllen hätte (77-80). Die unbefriedigende und (wie man fühlt) eher lustlose Ausführung dieses Rettungsversuchs lässt schwer erkennen, worauf es bei der (auch von Putnam als unentbehrliche Grundlage des Realismus proklamierten) Korrespondenz-Theorie im einzelnen ankommt. Eine Bedingung ist sicher klar: Sie muss Wahrheit von Sätzen/Gedanken als etwas bestimmen, das von der Außenwelt abhängt (Vgl. *The Refutation of Conventionalism*, in: l. c., 178). Was aber außer einer allgemeinen Satz/Gedanke-Tatsache- oder Theorie-Realität-Beziehung sonst noch erforderlich ist, damit sinnvollerweise von einer Korrespondenz gesprochen werden kann, bleibt unerläutert. (Immerhin verrät Putnam ausgeprägtes Bewusstsein davon, dass hier ein besonderes *Problem* besteht.)

Auf einen Zug möchte ich aber besonders hinweisen, weil er nach Putnams Wende größere Bedeutung erhalten soll. Es ist der Vorschlag, die Entsprechungs-Beziehung nicht als dyadisch, sondern als triadische Struktur zu verstehen (82, 1 u.). Zweigliedrig wäre sie, wenn (wie der Titel des Aufsatzes es formuliert) wahre Sätze mit der Wirklichkeit übereinzustimmen hätten, Punkt (hier wäre ‚Satzwahrheit‘ das eine, ‚Wirklichkeit‘ das andere Glied). Diese Beziehung interpretiert Putnam als Abbildung (mapping onto) einer gegebenen Sprache auf „sets of valuations over *mass, charge*“ (81). Ein Satz der gegebenen Sprache ist genau dann wahr, wenn er durch diese Abbildung erzeugt wurde. Der „philosophisch interessante Punkt“ dieser unscheinbar klingenden Erklärung, sagt Putnam, besteht darin, dass er zwar keine Erklärung des Wortes ‚wahr‘ liefert (wie das die Tarskische Theorie versucht habe).



Wohl aber würde sie einen interessanten *synthetischen* Sinn liefern, in welchem *tatsächlich* die wahren Sätze gerade die wären, die eine bestimmte nicht-triviale Beziehung zur aktuellen Welt unterhielten. Denn die Behauptung, daß es eine einzige Abbildung der [eben] beschriebenen Art gibt und daß ein gegebener Satz *S* auf ein Paar von Funktionen  $f_1/f_2$  abgebildet wird, das mit dem Paar *Masse, Ladung* identisch ist, wäre eine nicht-triviale Behauptung über *S* und wahr genau dann, wenn *S* wahr wäre. Mehr noch: diese Behauptung würde zu ihrer Wahrheit sowohl von den ‚Tatsachen der Welt‘ (der Art und Weise, wie Masse und Ladung über Raum-Zeit-Punkte verteilt sind) abhängen als auch von der Grammatik von *L* und der Art und Weise, wie Sprecher sich aktuell und potentiell verhalten gegenüber von Sätzen von *L*. In diesem Sinne würde sie die Existenz ‚einer Beziehung zwischen *S* und außer-sprachlichen (und sprachlichen) Tatsachen‘ behaupten. Vielleicht liegt der Hauptgrund für die Falschheit der (traditionellen) Korrespondenz-Theorie darin, daß sie das in Klammern gesetzte ‚und sprachlichen‘ wegläßt – es ist sehr wichtig [zu sehen], daß ein wahrer Satz *nicht* ein solcher ist, der eine bestimmte Beziehung zu außer-sprachlichen Tatsachen unterhält, sondern ein solcher, der eine bestimmte Beziehung zu außer-sprachlichen Tatsachen *und zum Rest der Sprache* unterhält. (Die ‚Korrespondenz‘ ist eher triadisch als dyadisch.) (81 f.).

Damit sei nun zwar nicht die Bedeutung des Wortes ‚wahr‘ aufgeklärt, aber die ‚Natur‘ der Wahrheit selbst, nämlich durch eine solche dreigliedrige Entsprechung von Sätzen zur Welt einerseits, andererseits aber zu dem Verständigungssystem, wie es die Sprache ist, in der die Sätze formuliert werden. Damit macht Putnam erstmals ein zaghaftes Zugeständnis an die hermeneutische Dimension desjenigen Verständigungs-Rahmens, in welchem überhaupt erst so etwas wie ein Bezugnehmen auf Tatsachen der so genannten außersprachlichen Wirklichkeit erfolgen kann. Und damit deutet sich auch ein erster Schimmer der späteren Einsicht an, dass die Rede von einer Wirklichkeit, ‚so wie sie an ihr selbst ist, ohne menschliche Zutat‘ (wie Friedrich Engels sagt), reiner Unsinn ist: Wir modellieren sie allemal nach den Vorgaben einer Theorie (in letzter Instanz einer Sprache), obwohl Putnam jederzeit darauf besteht, dass damit keine anti-realistische Metaphysik eingekauft werden muss. Wenn sich die Wirklichkeit durch die Brillen ver-

schiedener Theorien verschieden modelliert, handeln doch – meint er – alle Theorien von der einen und selben Wirklichkeit.

\*

Neben dem Korrespondenzbegriff spielt auch das Merkmal der Annäherung (,approximation') eine wichtige Rolle in Putnams frühem Wahrheitsbegriff.<sup>46</sup> Einerseits kommt im popperisch zugeschnittenen Schlagwort ,science as approximation to truth' (*Mathematics, Matter, and Method*, vii) der Approximationsbegriff zur Charakterisierung der wissenschaftlichen (Vorwärts-)Entwicklung und der Theoriendynamik zum Zuge. Andererseits verwendet ihn Putnam auch zur Beschreibung einzelner Theorien oder Aussagen. So spricht er z. B. davon, dass eine Theorie oder ein von einer wissenschaftlichen Theorie postuliertes Naturgesetz ,approximately true' sei (*Mind, Language, and Reality*, 209; und *Meaning and the Moral Sciences*, 61).

Es ist völlig klar, was es meint: eine Theorie sei annähernd wahr [...]; das meint, daß die von der Theorie geforderten Beziehungen (relationships) nicht genau gelten (hold), sondern nur zu einem bestimmten spezifizierbaren Grad an Irrtum. Kurz, es meint, die Theorie ist nicht wahr, aber eine bestimmte *logische Konsequenz* der Theorie *ist* wahr, z. B. eine, zu der wir kommen, wenn wir ,ist gleich' durch ,ist gleich plus oder minus Delta' ersetzen. Solch eine logische Konsequenz kann man eine *Approximations-Theorie* nennen (*Mind, Language, and Reality*, 118).

An einer anderen Stelle bemerkt Putnam (und das ist ja auch einleuchtend), dass die Semantik der approximativen Wahrheit die Semantik der Wahrheit voraussetzt (l. c., x, Anm.). In den John-Locke-Lectures von 1976 lässt er zudem erkennen, was wir unter ,hold not exactly' zu verstehen haben. Er sagt da:

---

<sup>46</sup> Zum Folgenden vgl. wieder Alex Burri, 33 ff.

Wenn wir sagen, die Erklärungsmodelle, die die Physik für dergleichen natürliche Arten wie ‚Wasserstoff-Atom‘ liefert, seien nicht annähernd richtig, sollten wir die Tatsache nicht aus dem Auge verlieren, daß für die meisten Zwecke (purposes) der Annäherungsgrad in der Tat sehr hoch sein kann. (Im Falle der Wasserstoff-Atome können Erscheinungen in der Quanten-Elektrodynamik *bis auf acht Dezimalstellen vorausgesagt* werden [Meaning and the Social Sciences, 62].)

Das lässt sich offenbar so verstehen, dass sich der Grad der approximativen Wahrheit an der zahlenmäßigen Genauigkeit (Stellen hinter dem Komma) ihrer Voraussagen bemisst. Mit den Begriffen ‚wahr‘, ‚wahrscheinlich wahr‘, ‚annähernd wahr‘ und ‚wahrscheinlich annähernd wahr‘ und diesem rein quantitativen Begriff der annäherungsweisen Wahrheit führt Putnam also offensichtlich eine Gradabstufung der Wahrheit ein. Er wäre sinnlos, stünde dahinter nicht das metaphysisch-realistische Konzept einer Welt an sich – unabhängig von aller Beschreibung, aber so, dass die Beschreibung sich ihr anzunähern versucht, ohne sie ganz zu erreichen. Der Abstand zwischen der Wirklichkeit-selbst und der ‚best available theory‘ ist gerade eine Grundvoraussetzung des Realismus, der ja die Unabhängigkeit der Welt von unseren Beschreibungen dogmatisch voraussetzt.

Darum ist noch wichtiger als der Annäherungsaspekt von Putnams Wahrheitskonzeption der externe Charakter von Wahrheit. Im Gegensatz zu Tarskis semantischer Wahrheitstheorie, die die Wahrheit von *einzelnen Sätzen* in einer *von vornherein gegebenen Sprache* (in der der referentielle Bezug aller denotierenden Ausdrücke *feststeht*) formal korrekt definiert, findet Putnams Wahrheitsbegriff auch bezüglich *ganzer Theorien* Anwendung und ist deshalb extern. Das heißt, dass es nicht um die Verifikation einfacher sprachlicher Gebilde (Sätze) geht, sondern um die Verifikation komplexer sprachlicher Gebilde (Theorien) anhand von etwas gänzlich Unsprachli-

chem, nämlich der Wirklichkeit, wie sie an sich ist. Gegenüber Tarskis Theorie hat dieser Ansatz den großen Nachteil, nie in präziser Form vorgelegt worden zu sein. Trotzdem kann Putnam zwei verwandte Argumente dafür präsentieren, warum ein externer, auf ganze Theorien bezogener Wahrheitsbegriff seinem positivistischen und verifikationistischen Ersatz – dem Ausdruck ‚führt zu erfolgreichen Voraussagen‘ – vorzuziehen ist.

Das erste Argument läuft darauf hinaus, dass externe Wahrheit und Falschheit unentbehrliche Hilfsmittel für den rationalen Vergleich und die rationale Kritik von konkurrierenden wissenschaftlichen Theorien verkörpert. Denn wenn – so räsontiert Putnam – die Bedeutungen der Ausdrücke (einer Theorie) tatsächlich, wie die Positivisten und Verifikationisten annehmen, von der sie enthaltenden Theorie abhängen, kann Tarskis sprachrelativer Wahrheitsbegriff gar nicht dazu verwendet werden, um – wie das doch faktisch dauernd geschieht – zwei Theorien mit ‚wahr‘ und ‚falsch‘ gegeneinander abzuwägen. Putnam geht sogar so weit zu behaupten, es sei im Rahmen der positivistischen und verifikationistischen Grundvoraussetzungen streng genommen nicht einmal erlaubt, eine von einer Forschergruppe verifizierte Aussage mit einer Aussage, die von einer anderen Forschergruppe verifiziert worden ist, zu einer Konjunktion zu verbinden (*Mind, Language, and Reality*, x f. mit Anm.).

Putnams zweites Argument ist ähnlich:

Wenn ein realistisch eingestellter (minded) Wissenschaftler [...] eine Theorie akzeptiert, so akzeptiert er sie als wahr (oder wahrscheinlich wahr, oder annähernd-wahr oder wahrscheinlich annähernd wahr). Da er auch die *Logik* akzeptiert, kennt er gewisse Züge (moves), die *die Wahrheit erhalten*. Wenn er z. B. eine Theorie  $T_1$  als wahr annimmt und ebenso eine andere Theorie  $T_2$ , dann weiß er, daß  $T_1$  und  $T_2$  – die *Konjunktion* von  $T_1$  und  $T_2$  – aus logischen Gründen ebenfalls wahr ist, und so akzeptiert er  $T_1$  &  $T_2$ . [...]

Ich möchte nun die Meinung vertreten, daß der Übergang (move) von dieser Annahme zu der Annahme der Konjunktion eine solche ist, zu der man nach der positivistischen Wissenschafts-Philosophie nicht berechtigt ist. [...] / Die Schwierigkeit ist sehr einfach. Die Annahme von  $T_1$  bedeutet für den Positivisten die Annahme des Kalküls  $T_1$  als zu erfolgreichen Voraussagen führend. [...] Und so entsprechend die Annahme von  $T_2$ . Aber von der Tatsache, daß  $T_1$  ebenso wie  $T_2$  zu erfolgreichen Voraussagen führt, folgt überhaupt nicht, daß auch die Konjunktion  $T_1 \& T_2$  zu erfolgreichen Voraussagen führt. Die Schwierigkeit ist, in nuce, daß das Prädikat, welches die *Rolle* der Wahrheit spielt – nämlich das Prädikat ‚führt zu erfolgreichen Voraussagen‘ – nicht die *Eigenschaften* der Wahrheit hat (*Mind, Language, and Reality*, 210 f.).

Da Putnam hier gar nicht darauf eingeht, ob die (konjunktive) Vereinigung zweier wahrer Theorien zu Inkonsistenzen führen könnte, müssen wir annehmen, dass sein früher Realismus im Gegensatz zu Quines Position erkenntnistheoretisch monistisch ist. Diese Annahme wird durch eine Kritik an Quines Unterbestimmtheitstheese vorerst erhärtet.

Diese These ist für die Realismus-Diskussion sehr wichtig und ich muss sie kurz in Erinnerung bringen. Quine hat seine berühmte These von der unüberwindlichen Unbestimmtheit (indeterminacy) der Übersetzung (die bekanntlich nicht erst zwischen verschiedenen Sprachen, sondern schon „at home“ auftritt) in Radikalisierung seiner These von der Unterbestimmtheit (underdetermination) aller empirischen Theorien entwickelt. Nur in sie geht es uns hier. Die Unterbestimmtheits-These hat Quine aus der Kombination zweier Annahmen begründet: 1. Peirce's Vorschlag, die Bedeutung eines Satzes als einen unser Interpretationsvermögen einbeziehenden, bewährungsabhängigen Schlussprozess aufzufassen (die Bedeutung eines Satzes ist, was durch das Gesamt der seine Wahrheit begründenden Aussagen gestützt wird): Dieser Prozess der Zeicheninterpretation ist empirisch unabschließbar; aus dieser Annahme folgt, dass Be-

deutung (meaning) und Information (information), Sprache (language) und Theorie (theory), sich nicht trennen lassen. Als 2. Annahme fungiert der so genannte Holismus, der möchte, dass Theorien nur als ganze, nicht in einzelnen Sätzen an der Wirklichkeit erprobt (tested) werden. Ist sprachlich artikulierte Wirklichkeitsauffassung insgesamt theorieimprägniert (mithin deutungsabhängig) – anders gesagt: macht es überhaupt keinen Sinn mehr, eine Unterscheidung einzuführen zwischen ‚der Art und Weise, wie die Welt wirklich ist‘ und ‚unserer Weise, über sie zu sprechen‘<sup>1</sup> – , dann kann die Möglichkeit nicht ausgeschlossen werden, dass ein und derselbe, durch Experiment sichergestellte, Erfahrungsbefund durch mehrere, untereinander inkompatible Theorien treffend interpretiert werden kann. „In a word, they can be logically incompatible and empirically equivalent“ (Quine, „On the reasons for indeterminacy of translation“, in: *The Journal of Philosophy* 67 [1970], 179). Nach welchen Kriterien wird man dann Theorien/Sprachen auswählen? Quine lässt nur den pragmatischen Gesichtspunkt größtmöglicher Brauchbarkeit in lebenspraktischer Absicht zu: Theorien sind nicht Re-präsentation theorieunabhängiger Wirklichkeit, sondern ‚tools for handling reality‘. Ein solcher pragmatischer Gesichtspunkt ist z. B. die Einfachheit und Übersichtlichkeit einer Theorie.

Quine ist nun zwar, wie der frühe Putnam, ein acharnierter Naturalist – aber eben – anders als Putnam – kein Realist. Darum ist zu erwarten, dass Putnam Quines Unterbestimmtheits-These angreifen wird. Das läge auch auf der Linie der eben gegebenen Zitate. Putnam meint in der Tat, dass es bei beobachtungsmäßig äquivalenten Theorien im Sinne Quines eben

---

<sup>1</sup> Richard Rorty, „Non-Reductive Physicalism“, in: *Theorie der Subjektivität*, Ffm. 1987, 282.

doch eine empirische<sup>47</sup> Tatsache geben könnte, die die eine als falsch und die andere als richtig ausweist (*Mind, Language, and Reality*, 180 f.). Auf der anderen Seite sagt er (noch vor dem „turn in my thinking“):

Ein Realist im 20. Jahrhundert *kann nicht* das Bestehen äquivalenter Beschreibungen ignorieren: der Realismus ist nicht zu dem Credo verpflichtet, daß eine (und *nur* eine) Theorie die wahre ist. [...] Anzunehmen, es gebe da eine ‚Tatsache‘, die die Vergabe des Wahrheitsprädikats imperativ vorschreibt, *wenn immer* wir zwei intuitiv verschiedene Theorien vor uns haben, ist naiv (*Meaning and the Moral Sciences*, 50).

Zwischen den beiden Zitaten gibt es eine offenbare Spannung. Aber sie ist doch nur scheinbar. Denn wenn Putnam einen Theorien-Pluralismus anerkennt (viele Modelle führen zur Wirklichkeit), so doch keinesfalls die Zulässigkeit der Annahme intertheoretischer Inkonsistenzen. So etwas könnte nur ein Antirealist zugeben.

Damit ist nun ein weiterer wichtiger Punkt von Putnams frühem Realismus verbunden. Da er die Unabhängigkeits-These nur unter der Bedingung vertreten kann, dass die Gegenstände der real existierenden Wirklichkeit eben keine bloßen Theorie-Konstrukte sind, muss er stabile theorieunabhängige Wahrheits- und Referenzbedingungen unterstellen. Was soll das heißen, und was folgt aus der These?

Offenbar muss Putnam annehmen, dass die Realien, von denen die Theorien in der Abfolge ihrer Geschichte handeln, so sind, wie sie sind, egal wie die Theorie sie beschreibt. Anders würde nämlich nicht erklärbar sein (meint er), wie es im wissenschaftlichen Erkennen eine *Konvergenz*, ein *Hinstreben zu immer größerer Annäherung an die Wahrheit*, kurz: so etwas

---

<sup>47</sup> An dieser Stelle kann natürlich der Phänomenalist wieder den Fuß in die realistische Tür zu stellen versuchen. Er wird sagen: Empirische Tatsachen sind immer Wahrnehmungsbefunde, und deren Elemente sind Sinnes-Daten, also was Mentales.

wie Wissenschaftsfortschritt geben kann (vgl. bes. *Meaning and the Moral Sciences*, 20 ff.). Und genau diese Konvergenz ist von Thomas Kuhn, erst recht von Paul Feyerabend mit weitreichenden Wirkungen bestritten worden. In *The Structure of Scientific Revolution* suggeriert Kuhn (wie Sie wissen), dass zwei paradigmatisch differierende Theorien trotz gleichem Vokabular nicht über dieselben Gegenstände quantifizieren. Die gleichen Ausdrücke beider Theorien haben also unterschiedene Bedeutungen und Bezugsgegenstände. Deshalb referiert z. B. ‚Elektron‘ in der Theorie von Bohr und Rutherford nicht auf das, was wir heute, nach der Entwicklung der Quantenmechanik, als ‚Elektron‘ bezeichnen. Da sich aber ‚Elektron‘ im Kontext von Bohr und Rutherford *aus heutiger Sicht* nicht auf Elektronen bezieht und die einzigen Referenzobjekte, die für ‚Elektron‘ bei Bohr und Rutherford überhaupt in Frage gekommen wären, Elektronen sind, referiert ‚Elektron‘ bei Bohr und Rutherford *aus heutiger Sicht* gar nicht. Es gibt gar keine Entitäten mit den von Bohr und Rutherford beschriebenen Eigenschaften. Diese Überlegung führt Putnam zu einer karikierenden Meta-Induktion: So, wie sich gezeigt hat, dass kein einziger *vor einem halben Jahrhundert* (oder wie viel Jahren auch immer) in der Wissenschaft verwendeter Terminus wirklich referiert, so wird sich herausstellen, dass kein einziger *jetzt* in der Wissenschaft verwendeter Terminus referiert.

Das erscheint Putnam als eine ganz unannehmbare Folgerung. Um ihr einen Riegel vorzuschieben, bedient er sich eines Nachsichtigkeits-Grundsatzes: des so genannten *Principle of Benefit of the Doubt*, also zum Grundsatz, im Zweifelsfall zu jemandes Gunsten zu entscheiden. Dank dieses Prinzips wird die Behauptung, dass sich der Ausdruck ‚Elektron‘ bei Bohr-Rutherford und in der Quantenmechanik sehr wohl auf einen und denselben Gegenstand beziehe, zu einem *Postulat*. Der Begriff der theorie-



unabhängigen Entität setzt den der verifikationsunabhängigen Existenz voraus. Diese ist gerade so, wie sie ist: ein an sich Seiendes wie das Ding an sich. Die Unterschiede seiner Beschreibung in verschiedenen, nacheinander von der Menschheit aufgestellten Theorien beruhen darauf, dass physikalische Größen in einer bestimmten Theorie *nicht definiert* sind, sondern nur mehr oder weniger genau gemessen werden können (hier tritt wieder das Annäherungs-Prinzip in seine Rechte). Also sind Elektronen bei Bohr/Rutherford nicht ultimativ definiert, sondern nur annähernd genau beschrieben. Ein Elektron darf also nicht als diejenige Entität aufgefasst werden, die ihrer theoretischen Beschreibung voll und ganz entspricht – sonst würde nämlich überhaupt keine betreffende Entität existieren, wenn sich diese theoretische Beschreibung *im Lichte einer späteren Theoriebildung* als unangemessen erweist. Putnam ist also (ähnlich wie Kripke, ähnlich selbst wie die Externalisten) der Meinung, ein Eigenname (wie ‚Elektron‘) referiere auf seinen Gegenstand unabhängig von der Triftigkeit oder Untriftigkeit der Beschreibungen, durch die ich ihn fasse. Ich muss einfach nur sicher sein, dass er über die Zeit hinweg in einem ähnlichen Sinne ‚derselbe‘ bleibt wie ein Mensch, der zwischen Geburt und Altersverwitterung die meisten seiner Eigenschaften wechselt – dessen Lebensgeschichte aber durch eine zusammenhängende Kausalkette zusammengeschweißt wird. Ebenso ist die referentielle Semantik von ‚Zeit‘ oder ‚Elektron‘ unabhängig von den Eigenschaften, die eine bestimmte Theorie der Zeit und den Elektronen gerade zuweist.

Andererseits, meint Putnam, erklärt nur die Theorie-Unabhängigkeit der Welt das Phänomen der wissenschaftlichen Konvergenz. Warum? Nun: wenn die positivistisch-verifikationistische Wissenschafts-Auffassung der Tatsache Rechnung tragen will, dass Nachfolgetheorien gewöhnlich *besser*

sind als ihre Vorgängerinnen, muss sie annehmen, eine Nachfolgetheorie habe so viele Beobachtungssätze zu implizieren wie möglich. Faktisch versuchen aber die Wissenschaftler etwas anderes. Sie versuchen, so viele Gesetze der Vorgängerin wie möglich als approximativ wahr in die neue Theorie hinüberzuretten; die Beobachtungssätze der Vorgängerin spielen dabei keine Rolle. Konvergenz hat also nichts mit einem Maximum an gemeinsamen Beobachtungssätzen zu tun, sondern damit, dass sich eine alte Theorie auf der Ebene der Gesetze als Grenzfall (limiting case) der neuen herausstellt (*Meaning and the Moral Sciences*, 20). Putnam stützt sich in diesem Zusammenhang immer wieder auf zwei – schon zitierte – Prinzipien, die sein fellow realist Richard Boyd formuliert hatte: 1., dass in einer ausgewachsenen (adult) Wissenschaft Ausdrücke normalerweise referieren; 2. dass die Gesetze einer Theorie, die einer ausgewachsenen Wissenschaft angehören, normalerweise approximativ wahr sind (l. c. f.). Mit dem ersten Prinzip ist die Meta-Induktion von vornherein ausgeschlossen, während das zweite Prinzip den Grenzfall-Charakter der wissenschaftlichen Entwicklung erklärt. Den Zusammenhang zwischen den Principle of Benefit of the Doubt und den beiden Boyd-Prinzipien sieht Putnam darin, dass das erste ohne das zweite *unvernünftig* wäre:

If there is no convergence, if later scientific theories cease having earlier theories as ‚limiting cases‘, if Boyd’s principles (1) and (2) are clearly false from the point of view of future science, then Benefit of the Doubt will always turn out to be unreasonable (l. c., 25).

Mit ‚unvernünftig (unreasonable)‘ meint Putnam: Wenn die Gültigkeit der Nachsichtigkeits-Grundsatzes nicht durch die Boydschen Prinzipien gerechtfertigt und eingeschränkt würde, müssten wir konsequenterweise auch annehmen, dass Termini wie ‚Phlogiston‘ oder ‚Äther‘ referieren. D. h. wir könnten den Fortschritt in der Annäherung-an-die-Wirklichkeit in Theo-

rien, die das wirklich tun, nicht wirklich abgrenzen von Theorien, die mit leeren Termini arbeiten.

Das typische Argument des Realisten gegen den Idealisten (oder, in neuerer Zeit, gegen den phänomenalistischen Instrumentalisten) beruht nach Putnam eben darauf, dass dieser den *Erfolg der Wissenschaft* nicht zu erklären vermag. Warum gelingt es – um mit den Instrumentalisten zu sprechen – den Elektronen-Kalkülen, Raum-Zeit-Kalkülen und DNA-Kalkülen, beobachtbare Phänomene richtig vorherzusagen, wo doch (angeblich) weder Elektronen noch DNA noch die gekrümmte Raum-Zeit Theorie-unabhängig existieren? Der Realist kann dagegen erklären, weshalb diese Theorien so erfolgreich sind (weil sie nämlich teilweise-wahre Darstellungen des Verhaltens von Elektronen bzw. von DNA sind) und wie das zeitliche Einander-Ablösen von Theorien erfolgt (nämlich dadurch, dass eine teilweise wahre, teilweise falsche Beschreibung eines theoretischen Gegenstandes durch eine *bessere* Beschreibung *desselben* Objekts ersetzt wird). Und weil der Realismus für die empirisch nachweisbaren Erfolge der Wissenschaft eine *Erklärung* liefert, muss er selbst so etwas wie eine empirische Hypothese sein (l. c., 18 f.). (Das wäre dann aber selbst schon fast so etwas wie eine anti-realistische, nämlich pragmatische Begründung des Realismus: eine Art performativen Widerspruchs für den metaphysischen Realisten.)

Um zu resümieren: Putnams früher Realismus darf insgesamt als eine sehr starke Form von Realismus bezeichnet werden. Sowohl im Bereich der Ontologie, wo sich Putnam als Befürworter einer Quantifikation 2. Stufe zu einem starken Artenrealismus (d. h. einem Realismus hinsichtlich physischer Größen) verpflichtet, als auch im Bereich der semantischen Theorie-Interpretation, wo er mit der Bivalenz-These und den transverifikationisti-

schen Existenzannahmen einen betont antiinstrumentalistischen Kurs steuert, nimmt Putnam die jeweils stärkste mögliche Position ein. Obwohl er nirgends explizit dazu Stellung nimmt, ob sein wissenschaftlicher Realismus reduktionistisch ist oder nicht, lässt sein funktionalistischer Standpunkt in der Philosophie des Geistes vermuten, das er auch vor dem „turn in my thinking“ einem konsequenten Reduktionismus nicht zugestimmt hätte. In Bezug auf die erkenntnistheoretischen Realismus verfiht Putnam eine etwas gemäßigtere Auffassung. Seine ausdrückliche Zurückweisung des Theorien-Monismus macht ihn trotz seinem Glauben an wissenschaftliche Konvergenz (und – damit verbunden – an einen externalistischen, approximativen Wahrheitsbegriff) und trotz seiner Ablehnung intertheoretischer Inkonsistenzen à la Quine zu einem eindeutigen erkenntnistheoretischen Pluralisten. Darin liegt auch der Unterschied zwischen Putnams frühem Standpunkt und dem knochenartigen metaphysischen Realismus, den er seit 1976 immer wieder kritisiert: Während die beiden Positionen hinsichtlich der Wahrheitsmerkmale Korrespondenz, Bivalenz und auch Unabhängigkeit (vgl. l. c., 34 f.) übereinstimmen, gehört im Gegensatz zu Putnams frühem Realismus auch Einzigkeit (es kann nicht mehr als eine vollständige und wahre Beschreibung der Welt geben) als konstitutives Element zum metaphysischen Realismus. Für den frühen Putnam ist zudem bezeichnend, dass er seinen Realismus als Erklärung und damit auch als empirisch-revidierbare Position verstanden wissen möchte. Das verleiht seinem Realismus eine zusätzliche Komponente, mit der er die Realismus-Diskussion paradoxerweise aus dem Bereich der Philosophie zu verbannen versucht.