

Universität Konstanz  
Philosophische Fakultät  
FG Philosophie  
Lehrstuhl Prof. Dr. Spohn  
WS 1998/99

Stefan Alexander Hohenadel

## INTERNALISTISCHE BEDEUTUNGSTHEORIEN

ÜBER THEORIE UND GESCHICHTE OBJEKTIVER UND SUBJEKTIVER  
BEDEUTUNGSBEGRIFFE VON FREGE BIS ZUR GEGENWART

Abschlußarbeit zur Erlangung des akademischen Grades des Magister Artium  
an der Philosophischen Fakultät der Universität Konstanz

# Inhaltsverzeichnis

<b>Vorwort</b>	<b>4</b>
<b>Zu Terminologie und Übersetzung</b>	<b>9</b>
<b>I Die Kennzeichnungstheorie der Bedeutung: Von Frege bis Searle</b>	<b>12</b>
I.1 Bedeutung und Kennzeichnungen . . . . .	12
I.2 Freges Unterscheidung von Sinn und Bedeutung . . . . .	13
I.3 Funktionalitätsprinzip, Substitutionsprinzip und die ungerade Rede . . . . .	15
I.4 Kognitive Signifikanz und Gedanken bei Frege . . . . .	18
I.5 Wieso Namen nichts bedeuten: Russell . . . . .	23
I.6 Der klassische formale Rahmen: Carnap . . . . .	28
I.7 Zur Kontextabhängigkeit von Äußerungen: Strawson . . . . .	35
I.8 Die Bündeltheorie der Bedeutung von Namen: Searle . . . . .	39
I.9 Das Ende der Carnapschen Semantik: Donnellan . . . . .	42
I.10 Was ist erreicht? . . . . .	46
<b>II Die Wende: Kripke</b>	<b>48</b>
II.1 Die Unzulänglichkeit der Bündeltheorie . . . . .	48
II.2 Apriorität und Notwendigkeit . . . . .	51
II.3 Das Prinzip der starren Designation . . . . .	54
II.4 Name und Notwendigkeit . . . . .	56
II.5 Eine Bestandsaufnahme . . . . .	59
<b>III Eine Theorie des Indexikalischen: Kaplan</b>	<b>61</b>
III.1 Einige wichtige Unterscheidungen . . . . .	61
III.2 Direkte Referenz, Äußerungskontexte und Auswertungssituationen . . . . .	62
III.3 Die doppelte Abhängigkeit der Indexikale . . . . .	64
III.4 Die zwei Arten der Bedeutung: Intension und Charakter . . . . .	65
III.5 Kripke und Kaplan . . . . .	69
III.6 Frege und Kaplan . . . . .	70
<b>IV Der semantische Externalismus: Putnam und Burge</b>	<b>72</b>
IV.1 Das internalistische Konzept der Bedeutung und seine Begrifflichkeiten . . . . .	72
IV.2 Die zwei Grundlagen des internalistischen Bedeutungsbegriffs . . . . .	75
IV.3 Wie die internalistische Bedeutungstheorie entwickelt wird . . . . .	76
IV.4 Erster Schritt der Widerlegung: Bedeutungen sind nicht im Kopf . . . . .	77
IV.5 Zweiter Schritt: Bedeutung, Realismus und versteckte Indexikalität . . . . .	81

IV.6	Dritter Schritt: die Sprachgemeinschaft und die Welt bestimmen die Extension	83
IV.7	Vierter Schritt: die untergeordnete Rolle der individuellen Sprachkompetenz	85
IV.8	Fünfter Schritt: der Normalformtyp des Bedeutungsvektors . . . . .	86
IV.9	Und nun? . . . . .	87
<b>V</b>	<b>Der „Enge Inhalt“ als Verteidigung gegen den Externalismus</b>	<b>89</b>
V.1	Wie man Putnam begegnen könnte . . . . .	89
V.2	Der Inhalt einer Einstellung ist nicht propositional festgelegt: Loar . . . . .	90
V.3	Inhalt in formalem Gewand: White . . . . .	97
V.4	Wie man Einstellungen doch individualistisch verstehen könnte: Fodor . . . . .	103
V.5	Fodors Zuordnungstheorie des engen Inhalts und ihre Herleitung . . . . .	107
V.6	Warum enger Inhalt überflüssig ist: Fodor in der Kritik von Schiffer . . . . .	111
V.7	Enger Inhalt zwischen Syntax und Holismus: Fodor in der Kritik von Block . . . . .	113
V.8	Blocks Versuch des holistischen engen Inhalts . . . . .	119
V.9	Epilog zum engen Inhalt . . . . .	120
<b>VI</b>	<b>Zur Semantik der Überzeugungen: Stalnaker</b>	<b>123</b>
VI.1	Relationalität und Intransitivität . . . . .	123
VI.2	Behauptungen und Propositionalkonzepte . . . . .	128
VI.3	Kaplan und Stalnaker . . . . .	132
VI.4	Voraussetzungen für eine Theorie der Überzeugung . . . . .	133
VI.5	1. Stufe: die Semantik der Behauptungen . . . . .	134
VI.6	2. Stufe: die Semantik der Überzeugungen . . . . .	136
VI.7	Frege und Stalnaker . . . . .	137
	<b>Nachwort</b>	<b>139</b>
	<b>Literaturverzeichnis</b>	<b>141</b>

# Vorwort

Diese Arbeit unternimmt den Versuch der systematischen Besprechung eines wichtigen Zusammenhangs in der modernen Philosophie der Semantik von ihren Anfängen bei Frege bis zu zeitgenössischen Auffassungen, wie sie zum Beispiel von White oder auch von Stalnaker vertreten werden. Der Titel mag vor dem Hintergrund dieses Anspruchs verwundern: schließlich sind keineswegs alle in dieser Schrift besprochenen Bedeutungstheorien internalistisch. Bevor ich jedoch genauer darauf zu sprechen komme, was ich von dieser Arbeit geleistet sehen möchte, will ich erst erklären, was unter einer „internalistischen Bedeutungstheorie“ zu verstehen ist.

Bereits seit die Philosophie die Sprache als einen der zentralen Gegenstände ihres Interesses identifiziert hat (man kann streiten, wann oder durch wen), gilt ihr besonderes Augenmerk der Frage, auf welche Weise Wörter Dinge bedeuten.

Unter einer Theorie der Bedeutung verstehe ich eine Theorie, die Angaben darüber macht, wie den sprachlichen Zeichen und Ausdrücken einer gegebenen Sprache ihre Bedeutung zugeordnet wird.

Mit Frege setzt die systematische philosophische Frage nach der Beschaffenheit dessen ein, was wir intuitiv als „Bedeutung“ bezeichnen. Insbesondere seine beiden Aufsätze „Über Sinn und Bedeutung“ (= [Fr94]) und „Der Gedanke“ (= [Fr93]) können als Grundriß der ersten philosophischen Semantik gelten.

Die Semantik ist – neben der Syntax, der Phonologie, der Morphologie und der Pragmatik – bekanntlich ein prominentes Teilgebiet der modernen Sprachwissenschaft. Die Pragmatik und die Semantik sind zugleich auch die beiden Teilgebiete, die in besonderem Maße philosophisch relevant sind. Die Semantik hat als Gegenstand die *Bedeutung* der sprachlichen Zeichen, das heißt die Beschaffenheit ihres Bezuges zur Welt. Die Pragmatik setzt schließlich die bedeutungstragenden und interpretierten sprachlichen Zeichen in Beziehung zu den *Sprechern*. (Wir werden noch sehen, daß diese Unterscheidung allenfalls als Arbeitshypothese fungieren kann, denn eine sehr starke These der neueren Sprachphilosophie besagt, daß man auch Semantik nicht losgelöst von den Sprechern und deren Sprachgemeinschaft betreiben kann. Zu dieser These werden wir uns durcharbeiten.)

Die Semantik einer gegebenen Sprache – nennen wir sie *L* – beschreibt also, wie den sprachlichen Ausdrücken von *L* Bedeutung zugeordnet wird. Sie gibt die Regeln an, nach denen sich die Bedeutung eines komplexen Ausdrucks von *L* aus den Bedeutungen der in diesem Ausdruck enthaltenen Teilausdrücke zusammensetzt – und dies geht „hinunter“ bis zu den kleinsten bedeutungstragenden Einheiten von *L*. Das Interessante daran ist folgendes: mit einer natürlichen Sprache lassen sich nahezu unendlich viele Ausdrücke formen. Die Semantik von *L* gibt uns aber nur *endlich* viele Regeln an, mit deren Hilfe wir die semantischen Werte eines komplexen Ausdrucks aus den semantischen Werten von dessen Teilausdrücken zusammensetzen können (es müssen endlich viele Regeln sein, weil kein

Sprecher eine Sprache mit unendlich vielen Regeln kompetent beherrschen kann). Es ist deshalb möglich, eine Sprache mit unendlich vielen Ausdrucksmöglichkeiten in endlich vielen, angebbaren Regeln zu beschreiben, weil sich der semantische Wert eines komplexen Ausdrucks aus den semantischen Werten seiner Teilausdrücke insofern zusammensetzt, als daß der Wert des Ganzen *eine Funktion* der Werte der Teile ist. Das heißt, ich kann für *L* Regeln angeben: 1.)... , 2.)... , 3.)... , die beschreiben, wie ich einfache Ausdrücke zu komplexen Ausdrücken zusammensetzen darf. Die Semantik erklärt dabei unter Rückbezug auf die Teilausdrücke *schrittweise*, wie der Wert des ganzen Ausdrucks zusammengesetzt wird. Man nennt eine solche Semantik daher *rekursiv*. (Wie eine rekursive Semantik aussehen kann, werden wir in Abschnitt 6 von Kapitel I erfahren. In Kapitel V werden unter anderem einige Erfordernisse deutlich, die wissenschaftliches Sprechen an die Beschaffenheit einer rekursiven Semantik stellt – meistens jedoch wird es in dieser Arbeit um gesprochene Sprache gehen.)

Nun können wir auch etwas genauer sagen, was im Rahmen dieses Papiers als Bedeutungstheorie verstanden werden soll: eine Bedeutungstheorie im engen Sinne ist eine Theorie, die eine rekursive Semantik auf theoretischer Ebene zu realisieren versucht – was bedeutet, daß es eben nicht nur eine rekursive Semantik des Deutschen oder des Englischen oder des Französischen undsofort sein soll, sondern ein theoretisches Modell wie zum Beispiel bei Carnap und Kaplan. Eine Bedeutungstheorie im weiten Sinne ist dann eine Theorie, die lediglich Bestandteile eines solchen Projektes formuliert, wie es beispielsweise bei Donnellan und Strawson geschieht. Zum Vergleich: in Kapitel I, III und VI geht es zumeist um Bedeutungstheorien im engen Sinne, während die in Kapitel II, IV und V vorgestellten Theorien sich in eher allgemeiner Weise auf den Begriff der Bedeutung beziehen und ich daher dazu tendiere, sie als Bedeutungstheorien im weiten Sinne aufzufassen.

Nun bedarf das Wort „internalistisch“ noch einer Erklärung: ‚Internalistisch‘ also soll im hier besprochenen Zusammenhang die Auffassung heißen, daß Bedeutung – was immer auch eine jeweilige Theorie darunter versteht – etwas ist, das insofern individualistisch beschrieben werden kann und muß, als daß die Bedeutung eines sprachlichen Ausdrucks allein von der individuellen Sprachpraxis eines gegebenen Individuums abhängt, der Inhalt und die Referenz der Wörter also letztlich dadurch bestimmt werden, was der Sprecher „im Kopf“ hat<sup>1</sup>.

Der Versuch einer Theorie der Wortbedeutungen hat, von Frege ausgehend, über Russell und Strawson bis zu Searle und Donnellan, sich immer auf diese fundamentale Annahme gestützt. Bei Kripke zeigen sich erste Ansätze einer fundamentalen Kritik dieser Annahme und in aller Schärfe wird diese von Putnam und Burge in den in Kapitel IV besprochenen Texten herausgearbeitet.

In Kapitel IV wird die – oben erst vage formulierte internalistische These – auch weitere Differenzierungen erfahren. Was ich hier als „Internalismus“ bezeichne, nennt Putnam (in Anlehnung an Descartes) „methodologischen Solipsismus“, Burge bezeichnet es als „In-

---

<sup>1</sup>Am deutlichsten hebt Burge diesen Punkt hervor, als er Russells Idee der *acquaintance* kritisiert: „The model tends to encourage individualistic treatments of the mental. For it suggests that what a person thinks depends on what occurs or “appears” within his mind. Demythologized, what a person thinks depends on the power and extent of his comprehension and on his internal dispositions toward the comprehended contents.“ [Bu79] p. 103

dividualismus“ und Stephen Stich nennt es „Prinzip der Autonomie“. Viele Rezensenten – Kritiker wie Proponenten – dieser drei Philosophen haben zwischen diesen drei Begriffen subtile Unterschiede gesehen, die allerdings nicht Gegenstand dieser Arbeit sein werden.

Die präzisierte These, die allen drei Spielarten gemeinsam ist, kann wie folgt ausgedrückt werden: bestimmte Überzeugungszustände – seien es linguistische Überzeugungen oder Überzeugungen im üblichen Sinne – sind (in Form psychischer Zustände) intrinsische Eigenschaften des jeweiligen Trägers dieser Überzeugungen.

Dabei wird zwischen zwei Arten von psychischen Zuständen unterschieden: solchen, die die Existenz von anderen Individuen voraussetzen ( $x$  ist eifersüchtig auf  $y$ ,  $x$  liebt  $y$ ) und solchen, die lediglich die Existenz einer einzelnen Bewußtseinsinstanz voraussetzen ( $x$  glaubt, Wasser sei naß). Die präzisierte These des Internalismus bezieht sich auf diese letzte Art von psychischen Zuständen, die bei Putnam (wie wir später sehen werden) psychische Zustände im engeren Sinne heißen.

(Was hier so kurz und prosaisch formuliert ist, bezeichnet in Wirklichkeit eine mächtige und sehr alte philosophische Tradition, die spätestens mit Platons Ideenlehre beginnt, sich über die gesamte Logik und Philosophie des Mittelalters, die Ausführungen Descartes' zum Leib-Seele-Dualismus und die kritische Philosophie Kants bis zum logischen Empirismus des zwanzigsten Jahrhunderts und dem durch Putnam selbst aufgebrauchten Funktionalismus in der Philosophie des Geistes der Sechziger und Siebziger Jahre fortsetzt und Grundlage fast jeder Theorie der Ethik ist. Die Auffassung, daß Gedanken und Überzeugungen intrinsische Überzeugungen des Einzelnen seien, könnte bei jedem dieser Philosophen und in jeder der angeführten Richtungen nachgewiesen werden, was ich hier aber nicht tun werde. Für unsere Zwecke kann die knappe und präzisierte Form, die oben angegeben ist, allemal genügen.)

Der logische Empirismus als die weitaus dominanteste Wissenschaftstheorie unseres Jahrhunderts hat, wie Putnam behauptet, diese internalistische Sichtweise als eine ihrer fundamentalen Annahmen verinnerlicht. Putnam jedoch verwirft in seinem an die Philosophie S.A. Kripkes anschließenden Aufsatz „The Meaning of ‚Meaning““ (= [Put75]) nicht nur die seit Frege klassische Zweiteilung des Bedeutungsbegriffs in Intension und Extension, sondern auch andere fundamentale Annahmen des logischen Empirismus zur Natur des Bedeutungsbegriffs, unter anderem die internalistische Kernthese, daß das Verstehen eines Begriffs ein psychischer Zustand im engen Sinne sei.

Kripkes scharfe Trennung der Kategorien Apriorität, Analytizität und Notwendigkeit und die Vorstellungen von Bedeutung und Wahrheit, die er im Zusammenhang mit dem Begriff der Querweltrelation entwickelt, schufen die theoretischen Voraussetzungen, um den Bedeutungsbegriff und die bisherige Forschungstradition aus dieser ungewöhnlichen Richtung zu kritisieren und dem Internalismus zumindest auf der Ebene der Bedeutungstheorie jene Theorie entgegenzusetzen, die als semantischer Externalismus bekannt geworden ist und vor allem in [Put75] und [Bu79] zur Blüte gekommen ist. Die Kernthese des Externalismus ist, daß das Verstehen von Bedeutung nicht nur nicht mit psychischen Zuständen (ieS) zusammenfällt, sondern daß die eigentliche Determinante der Semantik die Sprachgemeinschaft (und bei Putnam auch die physikalische Welt) ist, daß Gedanken demnach „nicht im Kopf“ sind.

So kann man, um den Bogen zurück zum Titel dieser Arbeit zu schlagen, zwei Arten von Bedeutungstheorien unterscheiden: internalistische Bedeutungstheorien und solche, die mit

der internalistischen Tradition brechen, indem sie den Bedeutungsbegriff partiell oder ganzheitlich ent-individualisieren, indem sie nämlich die Abhängigkeit des Bedeutungsbegriffs von Faktoren behaupten, die nicht in mentalen Konfigurationen des Sprechers bestehen. (Die Kapitel I-III beschäftigen sich mit den implizit oder unbewußt internalistischen Theorien der Bedeutung während Kapitel V die prominenteste internalistische Verteidigungsstrategie gegen Putnams Argument nachzuzeichnen versucht. In Kapitel VI kommt es zu einer Synthese der zusammengetragenen Argumente.)

Die Internalismus/Externalismusdebatte gibt der Diskussion um Bedeutungsbegriffe gegen Mitte der Siebziger Jahre eine entscheidende Wende: bisher stand (seit Frege) das Projekt der Realisierung des Modells einer rekursiven Semantik im Mittelpunkt der philosophischen Diskussion. Mit der Formulierung des semantischen Externalismus jedoch tritt ein vielleicht noch wesentlicherer Punkt hinzu, nämlich die Frage nach „Ort“ und Beschaffenheit der Bedeutungen sprachlicher Ausdrücke.

War bisher der einzelne Sprecher als Inhaber aller Bedeutung angenommen worden, so stellt sich nun die Frage, inwiefern die Sprachgemeinschaft Bedeutungen kollektiv entwickelt und gebraucht, während der Einzelne bei vielen Wörtern nur Stereotypen als Sprachkompetenz zur Verfügung hat. Das heißt, daß pragmatische Überlegungen die Philosophie der Semantik immer stärker beeinflussen: Geht man mit Burge davon aus, daß ein Sprecher einen Ausdruck nicht vollständig verstehen muß (was immer das heißt), um ihn erfolgreich verwenden zu können, dann zeichnet sich ab, daß Bedeutungen vielleicht nicht – oder: nicht *nur* – im Kopf sind, sondern daß wir woanders suchen müssen.

Damit ergibt sich jedoch die Frage nach dem, was im Kopf ist: was haben wir erworben, wenn wir ein Wort verstanden haben? Mit anderen Worten: was ist die *kognitive Signifikanz* eines Wortes? Diese Frage ist es, deren Historie ich anhand der Geschichte der Bedeutungstheorie nachgehen will und die ich bei Frege anfangend bis zu den neuesten Versuchen von Fodor, Stalnaker und anderen verfolgen will. Es bietet sich an, diese Frage an der Externalismusdebatte aufzuhängen, weil der semantische Externalismus der einzige Komplex von Bedeutungstheorien ist, der sich in der Interpretation kognitiver Signifikanz geschlossen gegen kartesianische, empiristische und funktionalistische Interpretationen des Internalismus abhebt und in einer Besprechung der Materie mit den älteren und neueren internalistischen Theorien kontrastiert werden muß. Ich habe hier eine eher historische Herangehensweise gewählt, doch auch eine systematische Einführung müßte dem Rechnung tragen.

Nun aber ist es auch leicht einzusehen, daß das Projekt einer internalistischen oder externalistischen Beschreibung der kognitiven Signifikanz sprachlicher Ausdrücke nicht losgelöst vom Projekt einer rekursiven Semantik gesehen und verstanden werden kann. Man kann sich leicht ausmalen, daß der Wegfall der Prämisse „Bedeutungen sind im Kopf“ eine Änderung des Bedeutungsbegriffs auslöst, denn schließlich muß eine rekursive Semantik auch Sätzen, die Individuen Überzeugungen zuschreiben, systematisch Bedeutung zuordnen (vgl. oben). Der herkömmliche Bedeutungsbegriff, den wir von Frege ausgehend kennenlernen werden, verortet an Wörtern eine Intension (einen Sinn) und eine Extension (ein Denotat). In Kapitel IV habe ich transparent zu machen versucht, wie eine externalistische Erklärung kognitiver Signifikanz (soweit Putnam diese *wirklich* gibt) in direkter Linie zur Widerlegung dieses klassischen Bedeutungsbegriffs führt. Es ist klar, daß das externalistische Argument für das Projekt einer rekursiven Semantik eine große Herausforderung bedeutet.

Zwar eröffnen sich hier Fragestellungen, die den Rahmen dieser Arbeit sprengen werden, doch ich werde dieser Abhängigkeit des Projekts der kognitiven Signifikanz vom Projekt der rekursiven Semantik dadurch Rechnung tragen, daß ich zeigen werde, wie sich jenes aus diesem entwickelt. Dies ist das vorrangige Ziel dieser Arbeit.

Ich werde in Kapitel I den Komplex der sogenannten Kennzeichnungstheorien der Bedeutung der Namen darlegen, wobei es insbesondere Russells elegante Theorie war, die aufgrund der Gleichsetzung der Namensbedeutung mit der Bedeutung einer entsprechenden Kennzeichnung diese Namensgebung inspiriert hat. Wir werden in diesem Kapitel die impliziten (sozusagen noch unbewußten) Internalisten von Frege bis Donnellan besprechen. Donnellan markiert meines Erachtens bereits den Niedergang der Kennzeichnungstheorie, weil er zeigt, daß Kennzeichnungen eine eigene semantische Funktionsweise haben, die in keiner der beiden Verwendungsweisen, die er aufzeigt, mit der von Namen gleichgesetzt werden kann – obwohl sich ähnliche Thesen für Namen formulieren lassen.

In Kapitel II werden wir die beeindruckende Zerstörung der Fregeschen Tradition durch Kripke erleben. In diesem Kapitel wird aufgezeigt, daß die Kennzeichnungstheorie deshalb fehlschlug, weil sie erstens von falschen Auffassungen hinsichtlich der Modalitäten Notwendigkeit und Möglichkeit ausging und zweitens von der grundfalschen These bezüglich der Substituierbarkeit von Namen durch Kennzeichnungen.

Kapitel III ist ganz der Theorie der Demonstrativa und der indexikalischen Wörter gewidmet, die Davin Kaplan vorgelegt hat. Wir werden darin sehen, wie sich die kognitive Signifikanz von indexikalischen Wörtern zumindest auf logischer Ebene eindrucksvoll und effizient abbilden läßt und wie man einen Teil von Freges Auffassungen gegen Kripke verteidigen kann – und wie der universale Geltungsanspruch der Fregeschen These fallen muß, daß die Intension eines Ausdrucks dessen Extension bestimme.

In Kapitel IV wird dann der semantische Externalismus anhand der Gedankenexperimente von Putnam und Burge dargestellt, seine Herleitung habe ich in Form eines 5-Stufen-Programms beschrieben. Das bisher Gesagte wird mit den Argumenten dieser beiden Philosophen kontrastiert und wir werden feststellen, daß die Frage nach dem Ort kognitiver Signifikanz bisher sträflich vernachlässigt wurde.

Kapitel V nimmt eine der dunkelsten Diskussionen auf, die es in der neueren Philosophie der Bedeutung gab: der Verteidigungsstrategie Jerry Fodors gegen den semantischen Externalismus. Fodor (und andere) bieten den Begriff des „engen Inhalts“ als Möglichkeit an, behaupten zu können, daß Bedeutung – entgegen Putnam – doch „im Kopf“ sei. Das Konzept des engen Inhalts litt von Anfang an unter gewissen Unklarheiten. Anhand der Positionen von Loar, White und Fodor wird ein exemplarischer Begriff des engen Inhalts erarbeitet und anschließend der Kritik durch Schiffer und Block ausgesetzt.

Kapitel VI konzentriert sich auf neuere Versuche zur Beschreibung kognitiver Signifikanz von Namen und anderen Wortarten anhand einiger Texte von John Perry, Ulrike Haas-Spohn und Robert Stalnaker. Wir werden darin sehen, wie die in Kapitel IV aufgezeigten Probleme innerhalb des Projekts einer rekursiven Semantik in Angriff genommen werden.

Bevor wir nun wirklich anfangen, noch einige methodische Anmerkungen:

1. Die Geschichte der Theorie der Bedeutung ist in weiten Teilen eine Geschichte der Theorie der Bedeutung von Namen. Eigennamen haben auf die Philosophie spätes-



tens seit Mill eine große Faszination ausgeübt. Wahrscheinlich liegt es, wie Ursula Wolf sagt<sup>2</sup>, daran, daß Namen in einer Weise, die wir intuitiv als „direkt“ empfinden, mit Gegenständen verknüpft sind. Man kann auf einen Gegenstand zeigen und ihm einen Namen geben: fortan steht dieser Name für diesen Gegenstand. Die Philosophie mag deshalb davon so fasziniert gewesen sein, weil dies die einfachste Art zu sein scheint, sich mit Sprache auf die Welt zu beziehen. Kann man diese Art des Bezugs erklären, so die Hoffnung der Philosophie, kann man daraus vielleicht das Modell für die Erklärung aller anderen Bezugsarten gewinnen. Dies scheint sich jedoch als falsch herauszustellen. Noch dazu wird sich im Laufe unserer Besprechung zeigen, daß selbst die normalen Namen vermutlich nicht nach diesem Schema funktionieren. Doch das ändert nichts am Bezug der verarbeiteten Literatur: Gegenstandsbezeichnungen (oder Designatoren, wie man sie auch nennt) im gemeinen und Namen im besonderen ist im ersten Kapitel dieser Arbeit daher ganz besondere Aufmerksamkeit gewidmet.

2. Es ist völlig illusorisch, „die“ Bedeutungsdiskussion darstellen zu wollen, weil es diese nicht gibt. Ich glaube, einen – wenigstens in der hier gewählten Perspektive – halbwegs repräsentativen Querschnitt durch die neuere Philosophie der Bedeutung gegeben zu haben und denke, daß man nach der Lektüre dieses Papiers eine Vorstellung davon haben kann, an welchem Punkt die philosophische Arbeit hier angelangt ist.
3. Ich bitte ich den Leser, keine runde Sache zu erwarten, wo es keine runde Sache geben kann. Insbesondere die auf Putnam und Burge antwortende Literatur krankt an einer gewissen Begriffsverwirrung, auf die ich noch zu sprechen kommen werde. Oftmals muß man sich für das synoptische Resümee auf den kommentierenden Teil am Schluß des Kapitels gedulden. Insbesondere in den letzten beiden Kapiteln liegt dies in der Natur der noch offenen Diskussion.

## Zu Terminologie und Übersetzung

Nahezu alle hier besprochenen Texte sind in englischer Sprache verfaßt worden. Die getreue Übersetzung bestimmter englischer Fachtermini führt im Deutschen in einigen Fällen zu eigenartigen Formulierungen. Übersetzungen, die als problematisch empfunden werden könnten, habe ich durch die Angabe des kursiv gedruckten Originalbegriffs in eckigen Klammern ergänzt. In den wenigen Fällen, wo es verschiedene, koextensive Originalbegriffe gibt, sind diese durch Schrägstriche getrennt angegeben.

Ich verwende die Ausdrücke „Gegenstandsbezeichnung“ und „Designator“ austauschbar. Sie orientieren sich an der Definition des Designators aus [Ho91]. Eine Gegenstandsbezeichnung, die einen Gegenstand durch die Angabe einer Eigenschaft dieses Gegenstandes herausgreift, nenne ich in Übereinstimmung mit der Tradition eine „Kennzeichnung“ [*definite description*] – eine Kennzeichnung kann zum Beispiel sein: „der Mann, der zuviel wußte“ oder „der letzte Kaiser von China“. Kennzeichnende Ausdrücke, die nicht einen Gegenstand herausgreifen, wie zum Beispiel „einer der Angeklagten von Nürnberg“ heißen in der an-

---

<sup>2</sup>vgl. [Wo85]

gelsächsischen Philosophie „indefinite description“. Ich bezeichne sie als „unbestimmte Beschreibung“.

Den sperrigen Ausdruck „natural kind term“ habe ich durchweg mit „natürliches Prädikat“ übersetzt. Dagegen kann man argumentieren, daß Wörter wie „Wasser“ Designatoren und keine Prädikate sind, doch Putnam, durch dessen Auffassung dieser Begriff in diesem Papier eingeführt wird, verwendet den Begriff in einer Weise, die sich mit „ist Wasser“ übersetzen lassen würde, so daß eigentlich nur das Prädikat und nicht der Designator gemeint sein kann. Außerdem geht es Putnam um das Problem, was „Wasser“ genannt werden kann – auf welche Gegenstände also das Prädikat „ist Wasser“ zutrifft.

In Kapitel III geht es um indexikalische Wörter, die Kaplan als „indexicals“ bezeichnet. Ich habe diesen Sprachgebrauch übernommen und bezeichne diese Wörter als „Indexikale“. Die Indexikale kann man in Übereinstimmung mit Kaplan demnach in die reinen Indexikale [*pure indexicals*] und die Demonstrativa [*demonstratives*] unterteilen.

Mit Kapitel IV richtet sich die Aufmerksamkeit auf die „propositional attitudes“, die ich als „propositionale Einstellungen“ oder kurz „Einstellungen“ bezeichne. Die Phrase „to individuate a propositional attitude“ müßte man getreuerweise als „die Identitätskriterien einer propositionalen Einstellung angeben“ übersetzen. Ich werde stattdessen verkürzend davon sprechen, daß propositionale Einstellungen „identifiziert“ werden. Eine propositionale Einstellung selbst ist streng von dem Satz zu unterscheiden, den wir benutzen, um diese Einstellung einem Individuum zuzuschreiben, beispielsweise „Heike glaubt, daß Geld nicht glücklich macht“. Einen Satz, der einem Individuum einen Glaubenszustand zuschreibt, bezeichnet die angelsächsische Literatur in der Regel als „ascription of belief“ oder „belief attribution“. Ich verwende beide Ausdrücke austauschbar und spreche von „Überzeugungszuschreibungen“ (wobei es richtiger „Überzeugungszustandszuschreibung“ heißen müßte).

Die meisten Schwierigkeiten bei der Vereinheitlichung der Sprache machte der Stoff aus Kapitel V. Die englischen Begriffe „notion“ und „concept“ sehe ich als austauschbar an und übersetze beide als „Begriff“. Die Ausdrücke „social content“, „wide content“ und „broad content“ sind in Kapitel V synonym verwendet und heißen bei mir „weiter Inhalt“. Auf das Zusammenfallen von weitem Inhalt und Wahrheitsbedingungen [*truth conditions*] werde ich in Kapitel V eingehen und ab und zu an die Austauschbarkeit dieser beiden Ausdrücke erinnern.

In Kapitel V kommt die Formulierung „oblique occurrence“ vor, die von Burge und Loar so intendiert ist, daß sie bezeichnen soll, was Frege einen „ungeraden Gebrauch“ nennt. Der Begriff der „oblique occurrence“ wird von Burge und Loar (und anderen) mit der populären Unterscheidung zwischen de-dicto- und de-re-Überzeugungszuschreibungen in einen theoretischen Zusammenhang gebracht. Zugunsten einer gewissen begrifflichen Bündigkeit und weil ich das Thema dieser Arbeit eng halten möchte, verzichte ich auf die Einführung dieser (wichtigen) Unterscheidung. Ich spreche anstelle der „oblique occurrence“ von einem „nichtextensionalen Vorkommen“. („Intensionales Vorkommen“ wäre bereits zu eng, denn nach Carnap sind die Einstellungskontexte, um die es bei Burge und Loar geht, *nicht* intensional.)

Eigene Wortschöpfungen und Bezeichnungen, die nicht von den jeweils besprochenen Autoren verwendet werden, habe ich im Text als solche kenntlich gemacht.

Die Arbeit ist in Kapitel unterteilt, die ihrerseits in Abschnitte unterteilt sind. Wird inner-

halb eines Kapitels von Abschnitt  $x$  gesprochen, so ist natürlich der betreffende Abschnitt dieses Kapitels gemeint. Querverweise auf spätere oder vorangehende Kapitel sind im Text meistens in der Form *Kapitelnummer.Abschnittnummer* angegeben, zum Beispiel VI.4. Beispielsätze sind durch die ganze Arbeit fortlaufend als „S. . .“ numeriert.

# I Die Kennzeichnungstheorie der Bedeutung: Von Frege bis Searle

## I.1 Bedeutung und Kennzeichnungen

Dieses erste Kapitel widmet sich einem Komplex von verwandten und aufeinander aufbauenden Bedeutungstheorien, die sich vor allem auf die Bedeutung von Namen und deren Verhältnis zu kennzeichnenden Ausdrücken beziehen. Dabei wird von den „alten“ Semantikern Frege und insbesondere Russell angenommen, daß Namen mit Kennzeichnungen synonym sind, während spätere Ansätze, wie der von Searle, diese Theorie dahingehend erweitern, daß die Bedeutung, mindestens jedoch die Referenz eines Namens durch ein Bündel von Kennzeichnungen festgelegt sei, daß viele Gegenstandsbezeichnungen also Bündelbegriffe [*cluster concepts*] seien. Eigennamen und Kennzeichnungen werden jedoch bei unserer Betrachtung der Bedeutungstheorien als die prominentesten Gruppen von Ausdrücken figurieren, so daß ich eine Theorie, die sie in eine Substituierbarkeitsbeziehung zueinander setzt, guten Gewissens als „Kennzeichnungstheorie“ bezeichnen werde, so wie alle Ansätze, die auf solchen Auffassungen implizit oder explizit aufbauen. Searles Ansatz nenne ich in Übereinstimmung mit Kripke die „Bündeltheorie“. Nach dieser Nomenklatur gehört auch die Bündeltheorie zu den Kennzeichnungstheorien.

Im Verlauf dieses Kapitels geht es mir darum, das Aufkommen der modernen Bedeutungstheorie anhand einer ihrer gedanklichen Hauptlinien nachzuzeichnen, nämlich eben der Kontroverse um die Bedeutung von Namen und Kennzeichnungen. Das entstehende Bild wird dabei natürlich roh und skizzenhaft bleiben müssen, denn uns interessieren in dieser Arbeit nicht alle semantisch-logischen Aspekte der Bedeutungstheorie, sondern, wie bereits im Vorwort angedeutet, vor allem ihre philosophischen Konsequenzen für das Projekt des Internalismus.

Ich werde also die Bedeutungstheorie Freges schildern, dann kurz auf die Kritik Russells und die Formalisierung Carnaps eingehen, bevor ich abschließend knapp die Weiterentwicklungen von Strawson und Searle darstelle. Wenn der Leser nach der Lektüre dieses Kapitels eine grundrißartige Vorstellung von der Kennzeichnungstheorie der Bedeutung hat und ihre Konsequenzen für die kognitive Rolle der Bedeutung von Ausdrücken verstehen kann, so hat dieses Kapitel bereits seinen Zweck erfüllt.

Die erste und für dieses Kapitel wichtigste philosophische Position ist die von Gottlob Frege, der als der Gründervater der modernen Bedeutungstheorie und der analytischen Philosophie schlechthin gesehen wird. In seinen beiden Aufsätzen „Über Sinn und Bedeutung“ (= [Fr94]<sup>1</sup>) und „Der Gedanke“ (= [Fr93]<sup>2</sup>) wie auch an anderen Stellen in seinem

<sup>1</sup>ursprünglich 1892, vgl. Literaturverzeichnis

<sup>2</sup>ursprünglich 1918, vgl. Literaturverzeichnis

Werk entwickelt Frege eine Theorie der Bedeutung sprachlicher Ausdrücke, die bis Ende der Sechziger Jahre unseres Jahrhunderts als Grundlage aller Philosophie der Semantik galt und trotz der Angriffe durch Kripke und den semantischen Externalismus das philosophische Sprechen über Bedeutung bis heute nachhaltig beeinflusst. Ein Teil von Freges Gedankengut erlebt insbesondere durch die Arbeiten David Kaplans seit Anfang der Achtziger Jahre eine gewisse Renaissance. Leider kann ich Freges Theorie hier nur sehr gerafft wiedergeben: der Schwerpunkt wird auf der Besprechung der beiden angegebenen Aufsätze liegen, das heißt, ich werde zum Verständnis von Freges Bedeutungstheorie die Unterscheidung von Sinn und Bedeutung und ihre Konsequenzen nachvollziehen und Freges Konzept des „Gedankens“ erläutern.

## I.2 Freges Unterscheidung von Sinn und Bedeutung

Bevor wir wichtige Begriffe einführen, wollen wir zunächst das Problem schildern, das für Frege den Anlaß gab, über eine Differenzierung des Bedeutungsbegriffs nachzudenken. Betrachten wir den Satz: „Der Abendstern ist mit dem Morgenstern identisch“. Wir können auch vereinfachend sagen:

S1) Hesperus = Phosphorus.

Die Ausdrücke „Hesperus“, „Phosphorus“, „Abendstern“ und „Morgenstern“ bezeichnen alle dasselbe Ding. Also unterscheidet sich die mit S1 gemachte Aussage nicht von der Aussage, die ich mit folgendem Satz mache:

S2) Hesperus = Hesperus.

Offensichtlich ist es aber so, daß S1 eine informative Aussage enthält, S2 dagegen lediglich die triviale Identität eines Gegenstandes mit sich selbst ausdrückt. S2 wissen wir daher a priori, eben „von vornherein“, S1 dagegen könnte das Ergebnis langer und aufwendiger Beobachtungs- und Forschungsarbeiten sein. Wie könnte eine Semantik aussehen, die diesem Unterschied Rechnung trägt?

Freges Antwort lautet, daß jeder Eigenname – jedes sprachliche Zeichen überhaupt – mit zwei semantisch relevanten Komponenten verbunden ist.

„Es liegt nun nahe, mit einem Zeichen (Namen, Wortverbindung, Schriftzeichen) außer dem Bezeichneten, was die Bedeutung des Zeichens heißen möge, noch das verbunden zu denken, was ich den Sinn des Zeichens nennen möchte, worin die Art des Gegebenseins enthalten ist. [...] Es würde die Bedeutung von „Abendstern“ und „Morgenstern“ dieselbe sein, doch nicht der Sinn.“ [[Fr94] p.41]

Da ist also zunächst das Referenzobjekt, das durch diesen Namen bezeichnet wird. Frege nennt das, was ein Name bezeichnet, seine Bedeutung. Die Fregesche Bedeutung des Namens „Aristoteles“ ist demnach Aristoteles. Wir werden diesen kontraintuitiven Gebrauch

des Wortes „Bedeutung“ nicht übernehmen, sondern weiterhin von Bezeichnetem, Referenzobjekt oder allenfalls von Fregescher Bedeutung sprechen.

Außerdem hat ein sprachliches Zeichen eine Art Inhalt, einen „Sinn“, der „die Art des Gegebenseins“ enthält. Das ist etwas dunkel und Frege sagt wenig, was diese Formulierung erhellen könnte, doch mir scheint der Begriff „Inhalt“ in diesem Zusammenhang erhellend zu sein, denn der Inhalt eines Zeichens kann als die Art gesehen werden, in der das Zeichen uns sein Bezeichnetes „gibt“.

Die Beziehung zwischen Zeichen, Sinn und Bezeichnetem besteht nach Frege darin, daß ein sprachlicher Ausdruck sein Bezeichnetes (seine Fregesche „Bedeutung“) bedeutet, das heißt *bezeichnet*, und seinen Sinn *ausdrückt*.

Diese Auffassung ist gegen die bis dahin populäre Theorie von John Stuart Mill gerichtet, der in Buch I, Kapitel 2 seines „A System of Logic“ argumentiert, Eigennamen hätten zwar Referenzobjekte, doch in diesem Sinne keinen „Inhalt“. Dementsprechend sieht Mill Namen nur als eine Art Etikett, deren einzige Funktion es ist, das, woran das Etikett angebracht ist, zu bezeichnen. Was Frege „Bedeutung“ nennt, heißt bei Mill „Denotation“; Freges „Sinn“ läßt sich ungefähr mit dem bei Mill verwendeten Ausdruck „Konnotation“ übersetzen.

Dem Zeichen entspricht nach Frege ein Sinn und diesem Sinn ein Referenzobjekt oder eine Klasse von Referenzobjekten (die Fregesche „Bedeutung“). Diesem Referenzobjekt jedoch entspricht nicht nur *ein* Zeichen. Ich kann auf wohl alle Dinge auf vielfältige Weise Bezug nehmen. Zum Beispiel kann ich auf „Napoleon“ auch durch den Ausdruck „Bonaparte“ referieren. Das heißt natürlich auch, daß in der gesprochenen Sprache derselbe Sinn auf mehrere Weisen sprachlich ausgedrückt werden kann.

Mehr noch: ein einziges Referenzobjekt kann durch Ausdrücke mit sehr unterschiedlichem Sinn eindeutig herausgegriffen werden. Aristoteles kann zum Beispiel von einer Person durch folgenden Ausdruck herausgegriffen werden: der Schüler Platons und Lehrer Alexanders des Großen. Ein anderer Sprecher jedoch verbindet mit dem Namen „Aristoteles“ den Sinn: der aus Stageira gebürtige Lehrer Alexanders des Großen. Für eine gegebene Fregesche Bedeutung wird es also im Normalfall immer „Schwankungen des Sinns“, wie Frege sie nennt, geben. Hier wird auch klar, wieso ich Frege als einen Vertreter der Kennzeichnungstheorie der Bedeutung bezeichnet habe: wir können den Sinn eines Namens als eine Kennzeichnung verstehen, wie im Beispiel von Aristoteles. Wir *müssen* den Sinn sogar so auffassen, denn wie sollten wir „die Weise des Gegebenseins“ sonst darstellen? (Bei Russell wird dieser Zusammenhang freilich noch deutlicher.)

Ein Ausdruck kann durchaus auch einen Sinn haben, ohne ein Referenzobjekt zu haben, wie zum Beispiel bei „Pegasus“ oder „Phönix“, woraus auch folgt, daß das Erfassen des Sinns noch nicht zu einer Fregeschen Bedeutung führt. Das scheint allerdings problematisch zu sein, denn die beiden vorgenannten Eigennamen sind in diesem Sinne *leer*, sie haben kein existierendes Referenzobjekt. Wenn wir also über Pegasus und Phönix sprechen, so sprechen wir in beiden Fällen über nichts und es will scheinen, als sprächen wir demzufolge über dasselbe. Dieses Problem wird uns später noch beschäftigen. Freges Lösung hierfür ist einfach: Leere Namen können nicht Bestandteil wahrheitswertfähiger Sätze sein, weil sie keine Fregesche Bedeutung haben. Doch Wahrheit ist, wonach die Wissenschaft sucht. Daher sind solche Sätze in wissenschaftlichem Gebrauch verboten.

In der Dichtung dagegen spielt Wahrheit keine Rolle. Hier ist es erlaubt, leere Namen in

Sätzen zu verwenden, denn beim Lesen eines Buches oder dem Verfolgen eines Theaterstücks kommt es uns ausschließlich auf den Sinn des Satzes an, eine Bedeutung brauchen wir zum Verstehen nicht, nur zum Urteilen. (Das ist ein Vorgriff: wir werden in den nächsten beiden Abschnitten sehen, wie es sich mit dem Urteilen und dem Sinn von Sätzen verhält.)

Sinn ist bei Frege eine kognitive Kategorie – was uns in Abschnitt 4 beschäftigen wird. Der Sinn ist das, was wir über das Referenzobjekt eines sprachlichen Ausdrucks wissen, wenn wir diesen Ausdruck verstanden haben. Es ist daher im Falle der Eigennamen durchaus legitim, den Sinn eines Namens als Methode der Identifizierung seines Referenten anzusehen. Doch bisher ist bei Frege unklar, was es bedeutet, einen Sinn verstanden zu haben. Auch hierauf werden wir in I.4 zu sprechen kommen.

Es ist zu beachten, daß der Sinn immer die Fregesche Bedeutung festlegt. Das heißt, zwei Ausdrücke können zwar unterschiedliche Sinne und die gleiche Bedeutung haben (wie zum Beispiel die beiden Ausdrücke „Abendstern“ und „Morgenstern“), doch das Umgekehrte ist nicht möglich: aus der Identität der Sinne folgt zwingend die Identität der Fregeschen Bedeutungen. Dies ist eine entscheidende Grundannahme der Philosophie der Bedeutung, deren Tragweite an dieser Stelle noch nicht erklärt werden kann. In Kapitel IV werden wir sehen, wie Putnam mit dieser Prämisse umgeht und wie schwer es ist, diese scheinbar offensichtliche These zu verteidigen. Doch dazu kommen wir später.

### I.3 Funktionalitätsprinzip, Substitutionsprinzip und die ungerade Rede

Wie sieht es mit Sinn und Bedeutung von Sätzen aus? (Wenn im folgenden von Sätzen die Rede ist, so sind dabei immer Aussagesätze gemeint, Fragen und Imperativsätze sind ausgeklammert.)

Frege sagt, daß mit einem Aussagesatz (auf jeden Fall mit einem Hauptsatz, auch mit vielen Arten von Nebensätzen) ein Gedanke verbunden sei. Ein Gedanke ist etwas, das als wahr oder falsch beurteilt werden kann. Gedanken sind – anders ausgedrückt – Wahrheitsbedingungen von Sätzen, sie geben die Bedingung an, unter denen der Satz wahr wird<sup>3</sup>. Gedanken wären gute Aspiranten, um als Bedeutungen von Sätzen zu gelten, doch Frege verneint dies aufgrund der Geltung des *Funktionalitätsprinzips für Bedeutungen*: **die Fregesche Bedeutung eines komplexen Ausdrucks ist eine Funktion der Fregeschen Bedeutungen seiner Teile**, das heißt Freges Semantik ist rekursiv. Aus diesem Prinzip ergibt sich das *Substitutionsprinzip für Bedeutungen*, welches lautet: **Wenn in einem komplexen Ausdruck zwei Ausdrücke mit identischer Fregescher Bedeutung füreinander substituiert werden, so ändert sich die Fregesche Bedeutung des ursprünglichen Ausdrucks dadurch nicht<sup>4</sup>.**

<sup>3</sup>Bei dem, was wir später als propositionale Einstellung kennenlernen werden, ist dies der eingebettete daß-Satz. Vergleiche zum Beispiel: ‚Ich glaube, daß ich dieses Kapitel heute fertig bekomme.‘ Dabei darf allerdings der daß-Satz nicht mit der Wahrheitsbedingung des Einstellungssatzes verwechselt werden: ob ich an etwas glaube, hängt nicht davon ab, ob das so ist. Der obige Satz kann auch wahr sein, ohne daß ich dieses Kapitel jemals fertig schreibe.

<sup>4</sup>Frege selbst verwendet die Ausdrücke „Funktionalitätsprinzip“ und „Substitutionsprinzip“ nicht. Ich habe sie hier eingeführt, um später auftauchende Paralleltäten bei der Rekursivität und Substitution von Sinn und

Wäre ein Gedanke die Bedeutung eines Satzes, dann dürfte er sich demnach nicht ändern, sofern wir innerhalb des Satzes einen Ausdruck durch einen anderen Ausdruck mit identischer Fregescher Bedeutung ersetzen. Das geschieht aber immer dann, wenn wir sinnverschiedene Ausdrücke in einem Satz gegeneinander austauschen: „Duncan wurde von Macbeth ermordet“ drückt einen anderen Gedanken aus als „Der König von Schottland wurde von Macbeth ermordet“. Also, so Frege, ist ein Gedanke nicht als Fregesche Bedeutung, sondern als Sinn eines Satzes aufzufassen: er wird durch diesen ausgedrückt, nicht bezeichnet.

Was aber ist dann die Fregesche Bedeutung eines Satzes? Nach Frege ist das einzige, was hierfür in Frage kommt, der Wahrheitswert eines Satzes. (Auch diese Lösung mutet allerdings seltsam an, weil wir dann einräumen müssen, daß die Sätze „Napoleon war Franzose“ und „Sirius ist der hellste Stern im Sternbild des Canis majoris“ dasselbe *bedeuten*.)

Die Theorie von Sinn und Bedeutung für Prädikate – also für diejenigen Satzteile, die übrigbleiben, wenn man das logische Subjekt aus dem Satz entfernt – ist bei Frege nicht sehr deutlich ausgeführt, besonders bezüglich des Sinns von Prädikaten äußert er sich fast nicht. Die Bedeutung eines Prädikats ist laut Frege ein Begriff. Ein Begriff ist eine metaphysische Entität, die einen Begriffsumfang hat, das ist die Menge aller Gegenstände, die unter diesen Begriff fallen. Der Sinn eines Prädikats könnte dann die Eigenschaft sein, die es ausdrückt, doch das führt zu Schwierigkeiten hinsichtlich der Trennschärfe der Begriffe „Begriff“ und „Eigenschaft“. Ich will das hier nicht weiter verfolgen, weil es für unsere Zwecke keine große Rolle spielt.

Worüber man spricht, wenn man sprachliche Ausdrücke benutzt, sind die Referenzobjekte der Ausdrücke. Wenn ich also „jener griechische Philosoph, der Lehrer Alexanders war“ sage, dann beziehe ich mich auf das Referenzobjekt dieses Ausdrucks, nämlich auf Aristoteles. Diesen Fall der Rede nennt Frege die „gerade Rede“. Doch es gibt einen anderen Fall, nämlich die sogenannte *ungerade Rede*. Diese liegt vor, wenn ein Ausdruck nicht, wie in der geraden Rede, seine Fregesche Bedeutung bezeichnet, sondern seinen Sinn. Zum Beispiel kommt in dem Satz „Es ist möglich, daß die Zahl der Planeten kleiner als 7 ist“ der Ausdruck „Zahl der Planeten“ in ungerader Rede vor. Ein typisches Kennzeichen ungerader Rede ist es, daß wir einen Ausdruck, der in ungerader Rede vorkommt, nicht unter sicherem Erhalt des Wahrheitswerts – *salva veritate* sagt der Logiker – gegen einen Ausdruck mit identischer Fregescher Bedeutung ersetzen können, was bei gerader Rede möglich ist. Beispielsweise können wir in dem gerade angeführten Satz den Ausdruck „Zahl der Planeten“ nicht *salva veritate* durch den Ausdruck „9“ ersetzen, obwohl diese beiden Ausdrücke identische Referenzobjekte haben.

In ungerader Rede hat ein Ausdruck als Referenzobjekt nicht seine gerade Fregesche Bedeutung – also seine normale Referenz – sondern seinen Sinn. Die ungerade Fregesche Bedeutung eines Ausdrucks ist also sein gerader Sinn. Demnach bezeichnet ein ungerade gebrauchter Ausdruck nicht sein Referenzobjekt, sondern seinen Sinn. Die Frage, die bei Frege keine Antwort findet, ist, was ein so gebrauchter Ausdruck dann anstelle seines geraden Sinns ausdrückt. Es müßte eine Art „ungerader Sinn“ sein, doch Frege sagt dazu nichts.

Frege versucht nun, das Problem der Substituierbarkeit zu lösen, indem er analog zum

---

Bedeutung besser beschreiben zu können.



Funktionalitätsprinzip und Substitutionsprinzip auf der Ebene der Fregeschen *Bedeutungen* ein Funktionalitätsprinzip und ein Substitutionsprinzip für Fregesche *Sinne* postuliert. (Dies ergibt sich sowieso, denn in ungerader Rede *ist* der Sinn eines Ausdrucks seine Fregesche Bedeutung und für Fregesche Bedeutungen gelten die beiden Prinzipien.) In ungerader Rede findet also eine Verschiebung der Referenz des Ausdrucks von seiner Fregeschen Bedeutung zu seinem Sinn (=seiner ungeraden Fregeschen Bedeutung) statt. Man kann in solchen ungeraden Kontexten laut Frege also aufgrund des Substitutionsprinzips für Bedeutungen sinngleiche Ausdrücke *salva veritate* substituieren.

Allerdings müßte Frege erklären, wie Kriterien für Sinngleichheit aussehen könnten. Die Intuition liefert hier zunächst keine greifbaren Ergebnisse. (Entweder wir haben dieselbe Kennzeichnung vor uns, dann ist es zweimal derselbe Sinn und nicht zwei gleiche Sinne – oder es sind zwei verschiedene Sinne, nämlich wenn es nicht dieselben Kennzeichnungen sind. Frege müßte hier genauer werden.)

Außerdem ergibt sich ein Problem mit mehrfach ungeraden Kontexten. In dem Satz „Lady Macbeth fürchtet, daß Prinz Malcolm vermutet, Macbeth könne der Mörder König Duncans sein“ kommt der Eigenname „Macbeth“ in doppelt ungeradem Kontext vor. In einem solchen doppelt ungeraden Kontext bezeichnet ein Ausdruck nach Frege seinen einfach ungeraden Sinn und drückt seinen doppelt ungeraden Sinn aus, was immer dieser sein mag. In einem dreifach ungeraden Kontext bezeichnet ein Ausdruck seinen doppelt ungeraden Sinn und drückt seinen dreifach ungeraden Sinn aus - undsofort. Wir kommen auf diese Weise in eine unendliche hierarchische Schachtelung von Sinnen, was intuitiv seltsam anmutet.

Wir haben jetzt mehr als genug an theoretischer Substanz zusammengetragen, um Freges Antwort auf das Problem mit den Sätzen S1 und S2 geben zu können:

Die Lösung für das Problem der Informativität von Identitätssätzen wird durch die Einführung des Sinnbegriffs ermöglicht: in S1 ist uns dasselbe Objekt auf verschiedene Weisen gegeben, in S2 auf dieselbe Weise. Also ist die Fregesche Bedeutung der beiden Ausdrücke in S1 und S2 zwar jeweils dieselbe, doch differieren die Ausdrücke links und rechts des Gleichheitszeichens in S1 hinsichtlich ihres Sinns, in S2 nicht. Es liegt auf der Hand, daß diese Auffassung ein starkes Argument gegen die oben vorgetragene Auffassung von Mill beinhaltet, die keine Erklärung für das Problem der kognitiven Differenz von S1 und S2 hat. Allerdings ist der Sinnbegriff bei Frege – wenn auch intuitiv recht klar – noch wenig präzise ausgeführt.

Ein Satz ist informativ, wenn wir ihn nicht von vornherein – also a priori – wissen können. Man könnte das Problem auch so ausdrücken, daß wir S2 für a priori wahr ansehen, S1 nicht, obwohl beide Sätze denselben Sachverhalt behaupten. Eigentlich ist das Problem also nicht semantischer, sondern epistemischer, das heißt kognitiver Natur. Wenn Frege nach einer semantischen Kategorie sucht, mit deren Hilfe dieses Problem lösbar wird, so wird diese semantische Kategorie auch den Unterschied in der kognitiven Signifikanz der beiden Sätze abbilden und vielleicht einen Schlüssel zu deren Beschreibung liefern. Dieser Punkt ist sehr wichtig, weil sich hier bereits die Verbindung zwischen dem Projekt einer rekursiven Semantik und dem Projekt der Beschreibung kognitiver Signifikanz sprachlicher Ausdrücke zeigt: Frege sucht eine *semantische* Kategorie, die *kognitive* Aspekte abbilden kann. Es wird damit deutlich, daß wir einen Bedeutungsbegriff brauchen, der sowohl objektive wie subjektive Dimensionen reflektieren kann, um solche epistemischen Unterschiede semantisch

abzubilden. Diese wichtige Einsicht wird im weiteren Verlauf der Arbeit eine zentrale Rolle spielen und alle folgenden Theorien müssen sich daran messen lassen, inwiefern sie in der Lage sind, dieser Forderung, die in I.6 noch präzisiert werden wird, zu genügen. In Kapitel VI werden wir eine Theorie kennenlernen, die das kann.

## I.4 Kognitive Signifikanz und Gedanken bei Frege

Mit Frege läßt sich die kognitive Signifikanz eines sprachlichen Ausdrucks nur als sein Sinn auffassen. Wir haben über den Sinn bei Eigennamen und anderen Gegenstandsbezeichnungen gesprochen und wollen in diesem Abschnitt noch etwas über den Sinn von Sätzen sagen, um eine Vorstellung von der kognitiven Signifikanz von Sätzen zu bekommen.

Der Sinn eines sprachlichen Ausdrucks ist scharf abzugrenzen von dem, was Frege eine „Vorstellung“ nennt. Die Verknüpfung einer Fregeschen Bedeutung mit den Sinnen der sie bezeichnenden Ausdrücke ist, darauf liegt Freges besondere Betonung, etwas Objektives. (Auch dies heißt natürlich, daß wir, sofern Sinne durch Kennzeichnungen angegeben werden, den Sinn eines Ausdrucks oder Namens als synonym mit diesen Kennzeichnungen auffassen müssen.) Für die Referenzobjekte eines Sprachzeichens ist die Objektivitätsforderung leicht einzusehen, wenigstens im Falle der Eigennamen und Kennzeichnungen, doch auch der Sinn eines Ausdrucks ist von der *Vorstellung* zu unterscheiden, die sich jemand von diesem Ausdruck oder seinen Referenzobjekten macht.

Eine Vorstellung ist subjektiv und gehört ins Gebiet der Psychologie. Eine Vorstellung von etwas bedarf eines Trägers, der diese Vorstellung hat. Meine Vorstellung von der Schönheit der Alpen kann nicht losgelöst von mir existieren.

Außerdem ist eine Vorstellung nicht übertragbar: keine zwei Sprecher haben von etwas dieselbe Vorstellung. Selbst wenn zwei Leute nebeneinander spazieren gehen, sehen sie nicht dasselbe, erleben nicht dasselbe und machen sich nicht dieselbe Vorstellung von dem, was sie sehen und erleben.

Das ist ein wichtiger Unterschied zum Begriff des Sinns, denn ein Sinn kann von vielen Leuten erfaßt werden und ist in dieser Hinsicht nicht subjektiv. Daß Aristoteles der Lehrer Alexanders war, läßt sich mitteilen und kann der geistige Besitz vieler sein.

Doch da mehrere Sprecher mit einem Ausdruck verschiedene Sinne verbinden können, ist es auch nicht so, daß für eine gegebene Fregesche Bedeutung ein bestimmter Sinn eindeutig festgelegt wäre, wie wir schon gesehen haben. Welchen Sinn ein Sprecher mit einem Ausdruck verbindet, hängt von den kognitiven Dispositionen des Sprechers ab – und nur von diesen! Insofern ist der Sinnbegriff ein internalistischer Bedeutungsbegriff! Und er ist in dieser Hinsicht subjektiv: Aus der Klasse der mit einem gegebenen Ausdruck assoziierten Sinne kann der Sprecher sich sozusagen einen „aussuchen“. Insofern hat nicht der eigentliche Begriff des Sinns, wohl aber seine intraindividuelle Verknüpfung mit der Fregeschen Bedeutung ein Moment, das der Subjektivität einen gewissen Spielraum läßt, wohingegen die Fregesche Bedeutung selbst eine streng objektive Sache ist. Frege sagt dazu:

„Die Bedeutung eines Eigennamens ist der Gegenstand selbst, den wir damit bezeichnen; die Vorstellung, welche wir dabei haben, ist ganz subjektiv; dazwischen liegt der

Sinn, der zwar nicht mehr subjektiv, wie die Vorstellung, aber doch auch nicht der Gegenstand selbst ist.“ [[Fr94] p.44]

Der Sinn eines Ausdrucks hat also sowohl subjektive wie objektive Dimensionen. Objektiv ist er insofern, als daß wir Sinne, die zu einem gegebenen Ausdruck gehören, im Normalfall von Sinnen unterscheiden können, die nicht zu diesem Ausdruck gehören. Und subjektiv ist der Sinnbegriff insofern, als daß zwei Sprecher, die sich über die Referenz, also die Frege'sche Bedeutung, eines gegebenen Ausdrucks einig sind, durchaus uneinig über den Sinn des Ausdrucks sind und wir unter Umständen beiden ein korrektes Verständnis des Ausdrucks zuschreiben können. In dieser Hinsicht ist der Sinn eine subjektive Kategorie, was ihn aber noch nicht zu einer psychologischen Kategorie macht, wie es die Vorstellung ist.

Wie wir bereits wissen, ist der Sinn eines Satzes der Gedanke, den er ausdrückt, und ein Gedanke, das wissen wir auch schon, ist etwas, das wahr oder falsch sein kann. Damit ist jedoch über Gedanken und auch über das Denken nicht genug gesagt, um eine Vorstellung von der kognitiven Signifikanz von Sätzen zu gewinnen.

Ein Gedanke fällt mit dem Inhalt eines Aussagesatzes zusammen – man könnte auch sagen, der Gedanke entspricht der durch diesen Satz ausgedrückten Proposition. Eine Proposition kann durch mehrere Sätze zum Ausdruck gebracht werden und deren grammatischen Unterschieden muß kein Unterschied in der ausgedrückten Proposition entsprechen. Die Proposition ist sozusagen die Wahrheitsbedingung des Satzes. Wir werden jedoch in I.6 sehen, wie diese Definition noch genauer gefaßt wird.

Es ist zu beachten, daß ein Hauptsatz immer einen vollständigen Gedanken ausdrückt. Ein Nebensatz tut das in der Regel auch, doch gibt es hier zahlreiche Ausnahmen, auf die ich nicht eingehen kann (vgl. [Fr94] p.51-65, wo nur darüber gesprochen wird). Jedenfalls werden wir uns im folgenden auf Aussagehauptsätze beschränken.

Von diesem propositionalen Charakter des Gedankens ausgehend, kann man drei Phasen des Umgehens mit Gedanken unterscheiden:

- 1.) Das Denken: das Fassen des Gedankens,
- 2.) Das Urteilen: das Anerkennen des Wahrheitswertes eines Gedankens,
- 3.) Das Behaupten: das Kundgeben des Urteils.

Bezüglich Schritt 1.) sind Gedanken wieder von Vorstellungen abzugrenzen: Vorstellungen werden *gehabt*, Gedanken werden *gefaßt*. Dies entspricht dem Unterschied, den wir eingangs besprochen haben: eine Vorstellung bedarf eines Trägers, von dem ihre Existenz abhängig ist. Diese Vorstellung *habe* ich. Ein Gedanke bedarf keines Trägers, um existieren oder einen Wahrheitswert annehmen zu können. Der Gedanke, daß im rechtwinkligen Dreieck die Summe der Kathetenquadrate gleich dem Hypothenusenquadrat ist, bedarf keines Trägers, um wahr zu sein. Dieser Gedanke wird lediglich *gefaßt*, nicht *gehabt*, er wird vom Bewußtsein des Sprechers aufgegriffen und kann von diesem Bewußtsein durch Sprache mitgeteilt werden, ohne aber von ihm abhängig zu sein.

Bezüglich Schritt 2.) muß man erklärend hinzufügen, daß das Urteilen dem Fassen eines Gedankens und der Kenntnis seines Wahrheitswertes beziehungsweise der Meinung über

seinen Wahrheitswert entspricht. (Natürlich können wir falsch urteilen: wenn wir über den Wahrheitswert im Irrtum sind.) Insofern markiert das Urteilen bei Frege, was wir uns gemeinhin unter Erkenntnis vorzustellen haben.

Wir haben nun die ganze Zeit zwischen dem Bewußtsein des einzelnen Sprechers und dem Bereich unterschieden, in dem sich die Gegenstände befinden. Frege spricht hier von Innenwelt und Außenwelt. Nimmt man als die Innenwelt des Sprechers dessen Bewußtsein an und als die Außenwelt den Bereich der Gegenstände und der physikalischen Welt, wohin gehören dann die Gedanken, welches ist ihr „Ort“? Freges Antwort mag überraschen: die Gedanken sind in ihrem eigenen, einem „dritten“ Reich angesiedelt. Gedanken unterscheiden sich von Vorstellungen, die zur Innenwelt gehören dadurch, daß sie Besitz vieler sein können und daß sie keines Trägers bedürfen. Von den Gegenständen unterscheiden sie sich jedoch dadurch, daß sie nicht sinnlich wahrnehmbar sind. Außerdem sind Gedanken, anders als Gegenstände, ewig und unveränderlich.

	Vorstellung	Gedanke	Gegenstand
bedarf eines Trägers	Ja	Nein	Nein
kann Besitz vieler sein	Nein	Ja	Ja
ist sinnlich	Nein	Nein	Ja

Der Tabelle kann man entnehmen, daß dem „Ort“ eines Fregeschen Gedankens – wir können auch sagen: einer Proposition – weder das Bewußtsein des Sprechers noch die Außenwelt entspricht. Daß Gedanken nicht zur Außenwelt gehören, läßt sich leicht einsehen, doch was spricht dagegen, sie dem Bewußtsein des Sprechers – dessen Innenwelt – zuzurechnen?

Die Antwort ist: das Vorhandensein einer kumulativen Naturwissenschaft. Nehmen wir an, Gedanken gehörten zur Innenwelt des Sprechers, dann wären sie subjektiv und individuenabhängig. Es gäbe dann meinen Satz des Pythagoras, Julias Satz des Pythagoras, Heikes Satz des Pythagoras undsofort. Die Mathematik – und überhaupt alle Wissenschaft, sofern es sie gäbe – wäre demnach von der Psychologie dominiert und von dieser abhängig. Das Projekt einer Ansammlung und Weitergabe von methodisch gewonnenem Wissen über die Generationen hinweg wäre praktisch unmöglich, da jeder nur über die Inhalte seines Bewußtseins redete und jede Objektivität Illusion wäre. Gedanken können also nicht als der Inhalt des Bewußtseins aufgefaßt werden.

Der Schlüssel zum Verstehen dieser eigenwilligen Konzeption liegt, wie Frege in „Der Gedanke“ weiter darlegt, darin, zwischen den *Gegenständen des Denkens* und den *Inhalten des Bewußtseins* zu unterscheiden. Bisher haben wir den Begriff ‚Bewußtsein‘ in einem weiten Sinne verwandt, doch bei Frege bezeichnet das Bewußtsein nur den Bereich der Vorstellungen. Vorstellungen sind Bewußtseinsinhalte, Gedanken in diesem Sinne von Bewußtsein nicht; Gedanken sind die Gegenstände des Denkens. Das heißt, jeder Sprecher hat auf dieses „dritte Reich“ der Gedanken einen Zugriff, dessen Art und Bereich von seinen individuellen psychischen Dispositionen abhängt. *Welche* Gedanken er allerdings faßt und wie er den Sinn der darin vorkommenden Ausdrücke expliziert, hängt von seinem – und nur

seinem – Denken ab. Insofern ist bei Frege nicht nur der Begriff des Sinns (als intuitiver Bedeutungsbegriff) internalistisch, sondern demzufolge auch der Begriff des Gedankens. Hat eine Bewußtseinsinstanz einen Gedanken gefaßt, so spricht mit Frege nichts dagegen, das Gefaßthaben dieses Gedankens als intrinsische Eigenschaft dieser Bewußtseinsinstanz anzusehen. Nur müssen wir Gedanken als intersubjektive, kognitive Kategorien qualitativ von Vorstellungen als rein subjektiven, hochindividuellen Kategorien unterscheiden.

Auf diese Weise können wir die Kumulativität des Wissens anerkennen und der Psychologie dennoch ihren verdienten Platz im Kanon der Wissenschaften einräumen.

Der Grundriß von Freges Theorie der Bedeutung, die, wie wir gesehen haben, in eine wenn nicht unproblematische, so doch unabhängige und originelle Philosophie des Geistes mündet, steht nun fertig vor uns. Wir wollen rekapitulieren, was wir alles haben: Freges Semantik kann mit der Informativität von Identitätssätzen umgehen. Informativität läßt sich, wie wir gesehen haben, mit Nichtapriorität gleichsetzen. Obwohl Frege nicht über die seit Kant zum philosophischen Standardsprachschatz gehörenden Modalkategorien wie Apriorität und Notwendigkeit spricht, ist mindestens die Kategorie der Apriorität in seinem Denken über Identitätssätze bereits angelegt.

Freges Lösungsversuch für das Problem der leeren Namen funktioniert allerdings nicht so recht. Freges Funktionalitätsprinzip für Fregesche Bedeutungen zufolge ist der Satz „Persephone wurde von Hades entführt“ weder wahr noch falsch. Zu behaupten, daß ein Satz, der einen leeren Namen enthält, nicht wahrheitswertfähig ist, verstößt auf sehr offensichtliche Weise gegen den Satz vom ausgeschlossenen Dritten, wonach „A oder  $\neg A$ “ (in jeder zweiwertigen Logik) für jeden Satz A wahr sein muß. Doch das Funktionalitätsprinzip wollen wir auch nicht aufgeben. Wollen wir die zweiwertige Logik aufrecht erhalten, wird hier also weitere Arbeit anfallen, die in I.6 geleistet wird.

Die Unterscheidung von Sinn und Bedeutung hat auch insofern wissenschaftliche Relevanz, als daß sie uns die Charakterisierung wissenschaftlichen Sprechens ermöglicht: die Fregesche Bedeutung ist für die Wissenschaft deshalb wichtig, weil einem Satz, der Bestandteile ohne Bedeutung enthält, kein Wahrheitswert zugesprochen werden kann. In der Dichtung dagegen kommt es lediglich auf den *Sinn* der Sätze an, nach Bedeutung oder Wahrheit zu fragen würde heißen, die Absicht der Dichtung zu verkennen.

Allerdings schweigt Frege sich an zu vielen Stellen – zum Beispiel bezüglich von Sinn und Bedeutung von Prädikaten – über wichtige Fragen aus, weswegen vieles an seiner Theorie skizzenhaft bleibt.

Die Gleichsetzung von Eigennamen mit Kennzeichnungen erfolgt implizit, indem Frege den Sinn von Eigennamen immer durch kennzeichnende Ausdrücke wiedergibt, wie im Beispiel des Namens „Aristoteles“ (vgl. [Fr94] p.42 Fn 2). Demnach verfügt jeder Sprecher in seinem Denken über eine Kennzeichnung, die er als Sinn mit dem Namen „Aristoteles“ verknüpft denkt. Verschiedene Sprecher können dabei verschiedene Kennzeichnungen mit „Aristoteles“ assoziieren. Frege räumt diese Schwierigkeit der „Sinnschwankung“ ein, doch er liest sich an dieser Stelle (ibid.), als halte er dies für eine Unsauberkeit in der gesprochenen Sprache, die das Projekt einer formalen Semantik einer Idealsprache nicht berührt und deshalb getrost ignoriert werden kann. Damit hat er das Problem unterschätzt, wie wir noch sehen werden.

Frege räumt ein, daß der Sinn eines Ausdrucks nichts rein Subjektives ist, weil der Sinn

eben zu jenem Ausdruck gehört, wie „der Lehrer Alexanders“ als korrekte Sinnangabe zu dem Namen „Aristoteles“ gehört; das hat mit dem Subjekt nichts zu tun. Subjektiv ist dagegen, welchen möglichen Sinn eines gegebenen Ausdrucks der Sprecher mit diesem verknüpft; auf diese Weise entstehen Sinnschwankungen (vgl. *ibid.*), aufgrund derer wir einräumen müssen, daß ein Sinn auch nichts rein Objektives sein kann. Hier besteht bei Frege eine gewisse Spannung, deren Aufhebung weitere Präzisierung erfordert.

Frege hat bemerkt, daß koextensive Ausdrücke nicht immer *salva veritate* ersetzbar sind. In der Unterscheidung zwischen gerader und ungerader Rede ist die Unterscheidung zwischen extensionaler und intensionaler Logik bereits vorgezeichnet.

Außerdem nimmt Freges Auffassung von Gedanken bereits den Hauptteil dessen vorweg, was heute Grundlage des philosophischen Sprechens über Gedanken ist. Freges Begriff des Gedankens ist deshalb so folgenreich, weil er einem Gedanken propositionalen Charakter zuspricht – mit anderen Worten: wir können Gedanken mithilfe von Aussagesätzen wiedergeben und infolgedessen mit ihnen ähnliche Operationen ausführen, die wir auch mit Sätzen und sprachlichen Zeichen ausführen können. Der Fregesche Gedanke ist – im heutigen Sprachgebrauch – der daß-Satz in dem, was wir in Kapitel IV als propositionale Einstellung kennenlernen werden, zum Beispiel: Julia glaubt, daß Napoleon Franzose war. Doch dazu kommen wir später.

Auch das Problem einer Rechtfertigung des Internalismus, an dem diese Arbeit aufgehängt ist, läßt sich bereits ablesen: wir müssen gegen unsere Intuition zwischen Denken und Bewußtsein unterscheiden, wobei das Denken nicht einmal Teil des Bewußtseins sein darf, weil Vorstellungen und Gedanken nicht denselben Bereich teilen dürfen. Frege bietet hier eine Unterscheidung *per definitionem*, doch eine echte Lösung des Problems durch logische Explikation steht noch aus. Auch sagt Frege nichts darüber, wie dieses „dritte Reich“ der Gedanken beschaffen sein soll, wie die Psychologie damit umzugehen habe und „wo“ es hingehört, wenn nicht ins Bewußtsein und nicht in die Außenwelt.

Frege geht von zwei Grundlagen aus, die in Kapitel IV ins Zentrum der Diskussion geraten:

1. Das Fassen eines Gedankens muß bei Frege eine Art mentale Disposition des Sprechers sein. Anders lassen sich die von Frege eingeräumten Sinnschwankungen nicht erklären. Frege kann nicht einerseits zugeben, daß ein gegebener Ausdruck objektiv mehrere Sinne haben kann und andererseits zugestehen, daß das Verstehen eines Sinns oder Fassen eines Gedankens dem Herausgreifen eines dieser Sinne gleichkommt, ohne auch zu behaupten, daß die Verknüpfung eines Ausdrucks mit einem Sinn ein mentaler Vorgang ist, der eine intrinsische Eigenschaft der entsprechenden Bewußtseinsinstanz ist. Frege muß in dieser Hinsicht Internalist sein.
2. Der Sinn eines Ausdrucks bestimmt die Fregesche Bedeutung. Gleichheit der Fregeschen Bedeutungen zweier Ausdrücke heißt nicht Gleichheit der Sinne, doch aus der Gleichheit der Sinne folgt die Identität der Fregeschen Bedeutungen.

Die erste Annahme kann als internalistische Rahmgebung für das Projekt zur Verortung der kognitiven Signifikanz von sprachlichen Ausdrücken gelesen werden, die zweite Annahme als theoretische Grundlage zu einer rekursiven Semantik. Wir werden in Kapi-

tel IV sehen, daß diese beiden Annahmen, erhält man sie beide aufrecht, auf ungeahnte Schwierigkeiten stoßen.

Freges Rolle in der neueren Philosophiegeschichte besteht unter anderem darin, für die gesamte analytische Sprachphilosophie wenigstens des Zwanzigsten Jahrhunderts die Agenda geschrieben zu haben. All diese aufgezählten Motive werden uns nämlich wieder begegnen. Wie wir gesehen haben, steckt bereits in diesen wenigen Seiten Frege mehr gedankliches Material, als es zunächst den Anschein hat und das meiste davon wird Gegenstand dieser Arbeit bleiben. Im Laufe dieses Kapitels wird die semantische Seite an Freges Theoriegebäude und dessen Ergänzungen zunächst stärker in den Vordergrund treten, doch wir werden das Problem des Internalismus nicht aus den Augen verlieren.

## **1.5 Wieso Namen nichts bedeuten: Russell**

Der folgende Absatz setzt sich mit der Theorie der Kennzeichnungen und Eigennamen auseinander, die von Bertrand Russell 1905 in dem Aufsatz „On Denoting“ aufgestellt wurde (= [Rus05]). Es ist für unsere Zwecke nicht nötig, den Russellschen Ansatz in ganzer Breite zu besprechen, daher werde ich auch hier nur einen gerafften Überblick geben. Zwar kritisiert Russell die Theorie Freges, doch in den grundlegenden Eigenschaften sind sich die beiden Theorien sehr ähnlich, insbesondere, was das Verhältnis zwischen Eigennamen und Kennzeichnungen angeht. (Als reine Theorie der Namen gelesen, unterscheiden sich Russell und Frege natürlich diametral, wie wir sehen werden!)

Wir hatten bei Frege zwei Probleme geortet, eines konnte seine Theorie lösen, das andere nicht:

Erstens war da das Problem der Substitution in ungeraden Kontexten. Wie kann es sein, daß wir in S1 den Ausdruck „Phosphorus“ nicht durch den bedeutungsgleichen Ausdruck „Hesperus“ ersetzen können, ohne die kognitive Signifikanz des Satzes zu verändern? Freges Lösung dafür lautete: weil die Ausdrücke sinnverschieden sind. Ähnlich, nur noch etwas deutlicher ist das in dem Satz „Jakob will wissen, ob Hesperus mit Phosphorus identisch ist“. Wenn wir hier den Ausdruck „Phosphorus“ durch den bedeutungsgleichen Ausdruck „Hesperus“ ersetzen, dann resultiert ein falscher Satz. Das Informativitätsproblem, mit dem Freges Behandlung des Sinnbegriffs ansetzt, mündet unmittelbar in dieses Substitutionsproblem.

Zweitens besteht ein Problem mit dem Funktionalitätsprinzip der Bedeutung und dem Satz vom ausgeschlossenen Dritten: ein Satz, der eine leere Gegenstandsbezeichnung enthält, ist nach Frege nicht wahrheitswertfähig, weil der Wahrheitswert – also die Fregesche Bedeutung – des ganzen Satzes sich aus den Bedeutungen der Satzteile zusammensetzt. Kommt darin ein bedeutungsloser Teil vor, so hat auch das Ganze keine Bedeutung. Dieses Problem konnte Frege nicht zufriedenstellend lösen.

Mit Russell lernen wir nun ein drittes Problem kennen, das wir jedoch auch aus Freges Abhandlungen bereits hätten ableiten können: das Problem negativer Existenzsätze. Ein Satz wie „Die Skylla existiert nicht“ kann nicht wahr sein – aus dem gleichen Grund wie im zweiten Problem. Der Ausdruck „Skylla“ hat keine Fregesche Bedeutung, also kann auch der ganze Satz keine Bedeutung haben. Wenn aber der Ausdruck „Skylla“ eine Bedeutung hätte, dann wäre der Satz falsch – denn dann gäbe es die Skylla. Auch für diese Aporie muß eine

Lösung gefunden werden.

Diese drei Probleme treten im Zusammenhang mit Gegenstandsbezeichnungen auf. Russells Ansatz nun konzentriert sich auf die Semantik von Gegenstandsbezeichnungen. Seine Strategie in diesem Zusammenhang ist ebenso unorthodox wie elegant: Gegenstandsbezeichnungen haben gar keine eigenständige Bedeutung, weshalb man sie durch logische Analyse aus jedem sprachlichen Ausdruck restlos eliminieren kann – mit Ausnahme der sogenannten *logischen Eigennamen*, auf deren Konzeption ich hier nicht eingehen kann.

Die Logik unterscheidet zwei Arten von Ausdrücken: einerseits solche, die für sich selbst genommen eine Bedeutung haben. Diese Ausdrücke heißen *kategorematisch*. Ausdrücke, die für sich selbst genommen keine Bedeutung haben, heißen *synkategorematisch*. Diese Unterscheidung ist anerkannt und bereits sehr alt – sie bildete bereits die Grundlage für die mittelalterliche Logik und Semantik<sup>5</sup>. Was allerdings nicht so klar ist, ist, welche Ausdrücke nun *kategorematisch* und welche *synkategorematisch* sind. Dieser Streit ist so alt wie die Einteilung selbst. Russells These nun ist es, daß Gegenstandsbezeichnungen wie zum Beispiel Kennzeichnungen, unbestimmte Beschreibungen („ein  $F$ “, „irgendein  $F$ “), Quantoren („alle  $F$ “, „jeder  $F$ “, „kein  $F$ “) und natürlich Namen im Gegensatz zur bisherigen Auffassung *synkategorematische* Ausdrücke sind. Wenn dies der Fall ist, dann müssen sie sich aus jedem sprachlichen Ausdruck systematisch eliminieren lassen. Diese Strategie sieht so aus:

Ein Ausdruck  $C$ , der auf eine Variable  $x$  zutrifft, bildet mit dieser eine Proposition  $C(x)$ . Die grundlegenden Kennzeichnungen „alles“, „einiges“ und „nichts“ können dann folgendermaßen aufgefaßt werden:

„ $C(\text{alles})$ “ bedeutet: „ $C(x)$  ist immer wahr“.

„ $C(\text{einiges})$ “ bedeutet: „Es ist falsch, daß ‚ $C(x)$  ist falsch‘ immer wahr ist“.

„ $C(\text{nichts})$ “ bedeutet: „‚ $C(x)$  ist falsch‘ ist immer wahr“.

Dann kann man alle oben angeführten Gegenstandsbezeichnungen systematisch und restlos eliminieren. Bei meiner Darstellung hier will ich mich auf ein Beispiel mit einer Kennzeichnung beschränken<sup>6</sup>.

Aus „Der  $F$  ist ein  $G$ “ wird in Russellscher Analyse: „Mindestens ein  $x$  ist  $F$  und höchstens ein  $x$  ist  $F$ , und was immer  $F$  ist, ist  $G$ “. In moderner Prädikatenlogik 1. Stufe mit Identität heißt das:

$$\exists xFx \wedge \forall y(Fy \rightarrow y = x) \wedge \forall x(Fx \rightarrow Gx)$$

Das ist nun doch etwas abstrakt. Versuchen wir es, indem wir für die Formel „der  $F$ “ etwas Konkretes einsetzen. Sehen wir uns den Satz an:

„Der Vater von Charles II. ist hingerichtet worden.“

---

<sup>5</sup>Vergleiche zum Beispiel die *Summa Logica* des William von Ockham oder auch *De Puritate Artis Logicae* von Walter Burley, um nur die prominentesten zu nennen. Die Einteilung in *kategorematische* und *synkategorematische* Termini findet sich jedoch in jedem scholastischen Logik-Lehrbuch.

<sup>6</sup>Für die anderen Gegenstandsbezeichnungen vgl. [Rus05].



In Russellscher Analyse wird daraus:

„Mindestens ein  $x$  zeugte Charles II. und höchstens ein  $x$  zeugte Charles II. und wer immer Charles II. zeugte, wurde hingerichtet.“

Es resultiert dieselbe prädikatenlogische Formel, wie oben, wenn wir für  $F$  „Vater von Charles II.“ einsetzen und für  $G$  „ist hingerichtet worden“.

Man kann hier gut sehen, daß nach Russell ein Satz, der einen kennzeichnenden Ausdruck enthält, in drei Behauptungen zerfällt:

1. eine Existenzbehauptung, dargestellt durch das Konjunktionsglied  $\exists xFx$ ,
2. ein Eindeutigkeitsprädikat, dargestellt durch das Konjunktionsglied  $\forall y(Fy \rightarrow y = x)$ ,
3. eine inhaltliche Behauptung, dargestellt durch das Konjunktionsglied  $\forall x(Fx \rightarrow Gx)$ .

Mithilfe dieser Methode können wir – das ist Russells Pointe – jeden Satz, in dem eine Kennzeichnung vorkommt, *salva veritate* so umformen, daß sie darin nicht mehr vorkommt. Also entfällt die Notwendigkeit einer separaten Theorie.

Russell vertritt die Auffassung, daß Eigennamen Abkürzungen für Kennzeichnungen sind. Demnach läßt sich das Vorkommen eines Eigennamens in einem Satz durch die entsprechende Kennzeichnung und diese durch das ersetzen, was Russell eine Propositionsfunktion nennt. (Diese Propositionsfunktionen sind zum Beispiel: „ $x$  ist ein  $F$ “ oder „ $x$  ist ein  $F$  wenn  $x$  ein  $G$  ist“. Wir brauchen diesen Begriff hier nicht zu erörtern.)

Diese These der systematischen Eliminierbarkeit zeigt, daß Russell der erste Kulminationspunkt dessen ist, was ich als Kennzeichnungstheorie bezeichnet habe: die Auffassung, Eigennamen seien eine Art Abkürzung für Kennzeichnungen, bedeutet, daß ich zum Beispiel den Ausdruck „Aristoteles“ in dem Satz „Aristoteles war Philosoph“ *salva veritate* durch die entsprechende Kennzeichnung ersetzen kann, zum Beispiel: „Der Lehrer Alexanders war Philosoph“. Ein Eigenname ist daher nichts anderes als eine Kennzeichnung, die wir auszusprechen uns nicht die Mühe gemacht haben.

Was bewirkt nun diese Methode der systematischen Elimination in Bezug auf unsere drei Probleme?

### **I. Das Problem negativer Existenzsätze:** Ein Satz wie

S3) Skylla existiert nicht

scheint nicht wahr sein zu können, egal, ob „Skylla“ etwas bezeichnet oder nicht. Da Russell von der Ersetzbarkeit von Namen durch Kennzeichnungen ausgeht, wird S3 zu

S4) Das sechsköpfige Ungeheuer existiert nicht.

Daraus wird:

S5a) Es ist nicht der Fall, daß mindestens ein  $x$  sechsköpfig und ein Ungeheuer ist und höchstens ein  $x$  sechsköpfig und ein Ungeheuer ist.

S5b) Die Propositionsfunktion „ $x$  ist ein Ungeheuer und hat sechs Köpfe“ ist niemals oder mehrmals wahr.

S5b ist dabei nur eine kondensierte Formulierung von S5a. Die Aussagen S5a und S5b sind ersichtlich sowohl sinnvoll als auch wahr. Wann immer wir also auf einen Existenzsatz stoßen, in dem ein Eigenname vorkommt, können wir ihn nach Russell so transformieren, daß er zu einer Aussage über den Status der Erfülltheit einer bestimmten Satzfunktion wird. Damit ist das Problem negativer Existenzsätze beseitigt.

## **II. Das Problem mit dem Satz vom ausgeschlossenen Dritten:**

S6) Der gegenwärtige König von Frankreich ist nicht kahlköpfig

S6 hat keinen Wahrheitswert, denn der Ausdruck „der gegenwärtige König von Frankreich“ bezeichnet nichts, weil Frankreich keinen König hat. Hier ist der Fall etwas schwieriger. Russell unterscheidet zwischen primärem und sekundärem Vorkommen von Kennzeichnungen in Sätzen. Ein primäres Vorkommen liegt vor, wenn der kennzeichnende Ausdruck den weitesten Bereich (=Skopus) hat – gibt es einen Ausdruck, dessen Bereich weiter ist, so liegt ein sekundäres Vorkommen vor. Das heißt, wir können S6 auf zwei Weisen lesen, zum einen als S7a und zum anderen als S7b.

S7a) Es gibt genau einen König von Frankreich und dieser ist nicht kahlköpfig.

S7b) Es ist nicht der Fall, daß es genau einen König von Frankreich gibt und dieser kahlköpfig ist.

In S7a kommt der Ausdruck „der König von Frankreich“ primär vor, in S7b sekundär. S7a ist falsch, denn es gibt keinen König von Frankreich. S7b jedoch ist offensichtlich wahr. Der Wahrheitswert von S6 hängt also von der logischen Lesart ab, die wir an ihn anlegen. Doch je nach Lesart hat der Ausdruck wenigstens einen Wahrheitswert. Der Satz „Der gegenwärtige König von Frankreich ist kahlköpfig“ erlaubt nur eine primäre Lesart und ist daher falsch. Nun können wir mithilfe der üblichen Junktoren beliebig solche Sätze mit beliebig vielen Gliedern bilden und sie werden teilweise ambig sein, doch in jeder möglichen Lesart werden sie einen Wahrheitswert haben.

## **III. Das dritte und wichtigste Problem war das der Substituierbarkeit von Ausdrücken in ungeraden Kontexten:**

S8) Sven will wissen, ob Hesperus mit Phosphorus identisch ist.

In diesem Satz haben wir ein solches Problem, denn wenn wir hier den Ausdruck „Phos-

phorus“ durch den Ausdruck „Hesperus“ ersetzen, so ergibt sich:

S9) Sven will wissen, ob Hesperus Hesperus ist.

S8 ist wahr und S9 ist falsch. Da Hesperus Phosphorus ist, gerät dieses Beispiel in Konflikt mit dem Substitutionsprinzip der Bedeutung, demzufolge koreferentielle Ausdrücke, also Ausdrücke, die denselben Gegenstand bezeichnen, *salva veritate* substituiert werden können. Russells Argument ist nun, daß „Hesperus“ und „Phosphorus“ synkategorematische Ausdrücke sind, die für sich genommen nichts bezeichnen, weswegen wir bei ihnen auch nicht von Gleichheit der Referenz sprechen können.

S1 können wir nach Russell transformieren zu

S10)  $\exists x (x \text{ ist morgens in der-und-der Position zu sehen}$   
 $\wedge \forall y (y \text{ ist morgens in der-und-der Position zu sehen} \rightarrow y = x)$   
 $\wedge x = \text{Hesperus})$

Nun ist S8 zu analysieren, indem man darin den Bestandteil, den man als S1 paraphrasieren kann, durch S10 ersetzt. Dadurch, daß der Ausdruck „Phosphorus“ eliminiert wurde, wird deutlich, daß es sich nicht um eine echte Identitätsaussage der Form  $a = a$  handelt. Ohne, daß „Phosphorus“ in dem Satz vorkommt, können wir jedoch auch „Phosphorus“ nicht mehr durch „Hesperus“ ersetzen, wodurch die Transformation von S8 in S9 verhindert wird.

Die Eleganz, mit der Russell die angegebenen Probleme löst, ist einnehmend. Dennoch ist ihm die Sprachphilosophie nicht in der kontraintuitiven Auffassung gefolgt, daß Namen nichts bedeuteten.

Bezüglich des Verhaltens der Russellschen Theorie zu der von Frege läßt sich folgendes feststellen:

Erstens kommt Russells Theorie ohne die Unterscheidung von Sinn und Bedeutung aus. Die Einfachheit seines logischen Instrumentariums verleiht seinem Ansatz ebenfalls eine gewisse Eleganz. Leider hat er uns infolgedessen auch nichts wirklich Neues über kognitive Signifikanz zu sagen: in Bezug auf Eigennamen fällt die kognitive Signifikanz mit der Kennzeichnung zusammen, durch die wir diesen Eigennamen ersetzen. Die Klasse der Kennzeichnungen, durch die ein Name ersetzbar ist (bei Frege ist dies die Klasse der Sinne) gibt damit die Möglichkeiten und Grenzen seiner kognitiven Auslegung an. Auf epistemische Kategorien nimmt Russell keinen Bezug.

Im Gegensatz zu Freges Theorie kann die Russells mit leeren Namen umgehen. Freges Lösung ist in der Form, die er selbst vorgelegt hat, unzureichend. Die Menge der synkategorematischen Ausdrücke ist bei Frege viel kleiner als bei Russell: bei Russell nämlich haben Namen und Kennzeichnungen einfach keine Referenz, weswegen leere Namen sich nicht wirklich von „normalen“ Namen unterscheiden, denn auch diese bedeuten nichts.

Auch Russells Lösung für das Problem der Substitution in ungeraden Kontexten unterscheidet sich von derjenigen Freges: Frege löst es, indem er eine Verschiebung der Referenz von der Fregeschen Bedeutung hin zum Sinn behauptet. Russell spricht den Ausdrücken Referenz ab und blockiert dadurch die Anwendung des Substitutionsprinzips, so daß kein

Konflikt entstehen kann.

Im Gegensatz zu Frege, aus dessen Ansatz sich die Substituierbarkeit von Kennzeichnungen durch Namen nur ableiten ließ, behauptet Russell sie explizit. Die beiden Theorien sind sich damit ähnlicher als es vielleicht zunächst den Anschein haben mag. Oft findet man in der Literatur die Rede von der „Frege/Russell-Auffassung“.

## I.6 Der klassische formale Rahmen: Carnap

Die bedeutendste Weiterentwicklung von Freges Theorie wurde 1947 durch Rudolf Carnap geleistet, in dessen Buch „Meaning and Necessity“ (= [Ca47]) er eine formale Semantik der Sprache  $L$  entwirft, für die er das Projekt einer rekursiven Semantik realisiert. Seine in [Ca47] entwickelte Theorie war so erfolgreich, daß praktisch jede Arbeit zur formalen Semantik auf seinem Begriffspaar von Intension und Extension aufbaut.

Es wird im folgenden darum gehen, die Weiterentwicklungen, die Carnap für Freges Theorie vorschlägt, vorzustellen und ihre Bedeutung für die Entwicklung der Kennzeichnungstheorie der Namen transparent zu machen. Außerdem werden wir einen Fortschritt hinsichtlich der Formulierung der kognitiven Signifikanz sprachlicher Ausdrücke und der Präzisierung des intuitiven Bedeutungsbegriffs erleben.

Carnaps semantische Weiterentwicklungen beruhen auf zwei wesentlichen Verbesserungen von Freges Theorie:

Erstens hat Carnap den Begriff dessen, was Frege „Sinn“ nannte, präzisiert und auf diese Weise einer formalen Handhabung und Analyse zugänglich gemacht.

Zweitens ist es ihm gelungen, die Prinzipien der Funktionalität und Substitution, die bei Frege noch zu Problemen führen, in problemfreier und formal präziser Weise zu formulieren.

Widmen wir uns zunächst dem ersten Aspekt: der neuralgische Punkt von Freges Semantik war das Begriffspaar von Sinn und Bedeutung. Der Sinn eines Ausdrucks war dabei die Weise, in der uns der Ausdruck seinen Gegenstand gibt, die Bedeutung eines Ausdrucks war das Referenzobjekt des Ausdrucks.

Bei Carnap geht es nun formaler zu:

Wie bei Frege mit jedem sprachlichen Ausdruck ein Sinn und eine Bedeutung verknüpft ist, so ist auch bei Carnap jeder Ausdruck der Sprache  $L$  mit einer Intension und einer Extension verknüpft. Die Sprache  $L$  besteht aus Gegenstandsbezeichnungen, Prädikaten und Aussagesätzen. Diese Ausdrücke kommen für den Besitz einer Intension und Extension in Frage.

Der Begriff der Extension ersetzt Freges kontraintuitiven Bedeutungsbegriff. Die Extension eines Ausdrucks besteht – vereinfacht gesagt – in einer Klasse von Gegenständen, die diesen Ausdruck in angemessener Weise syntaktisch „erfüllen“: die Extension einer Gegenstandsbezeichnung ist der Gegenstand, den dieser Ausdruck bezeichnet oder die Menge der Gegenstände, die er bezeichnet. Die Extension eines Prädikats ist die Klasse aller Gegenstände, auf die dieses Prädikat zutrifft und die Extension eines Satzes ist sein Wahrheitswert. Die Ähnlichkeit mit der Fregeschen Bedeutung ist offensichtlich.

Die Intension eines Ausdrucks entspricht bei Carnap dem, was wir intuitiv als „Bedeutung“

bezeichnen und was nichts mit dem Fregeschen Bedeutungsbegriff zu tun hat. In Carnaps Theorie ersetzt der Intensionsbegriff den Fregeschen Begriff des Sinns. Die Intension eines Begriffs ist eine Funktion, die eine Zuordnung aus dem Bereich möglicher Situationen in den Bereich adäquater Extensionen für diesen Ausdruck in der jeweiligen Situation festlegt. Das müssen wir genauer erklären:

Im algebraischen Sinne ist eine Funktion  $f$  von der Menge  $A$  in die Menge  $B$  eine Abbildung (d.h.: Zuordnung) von  $A$  nach  $B$  derart, daß jedes  $x \in A$  auf genau ein  $y = f(x) \in B$  abgebildet wird. Einfacher ausgedrückt heißt das: eine Funktion ist ein Verknüpfungsmodus zweier Mengen, der *jedem* Element aus  $A$  *genau ein* Element aus  $B$  eindeutig zuordnet. Eine Funktion enthält mindestens eine Argumentstelle, das heißt, ihr muß ein Argument – eben ein Wert – übergeben werden. Dieses Argument  $x$  verknüpft sie dann nach den Regeln, die in ihr festgelegt sind, mit einem Funktionswert  $y = f(x)$  aus der Menge  $B$ . (Eine solche Regel kann zum Beispiel sein, daß für jedes  $x$  aus der Menge der natürlichen Zahlen  $f(x) = x^2$  ist.) Die Menge  $A$  ist in unserem Beispiel die sogenannte Definitionsmenge von  $f$ , das heißt, der Bereich, für den diese Funktion definiert ist und der alle in Frage kommenden Argumente umfaßt. Die Menge  $B$  ist die Wertemenge von  $f$ , das heißt, der Bereich, der alle in Frage kommenden Funktionswerte umfaßt. Wenn im Laufe meiner Abhandlung von Funktionen von  $A$  in  $B$  die Rede ist, so sind Funktionen in diesem algebraischen Sinne gemeint.

Der Definitionsbereich der Intensionsfunktion ist immer die Menge möglicher Situationen. Ich möchte aus Gründen der Einfachheit und Konstanz der Darstellung im folgenden den populären, doch von Carnap noch nicht verwendeten Begriff der „möglichen Welt“ verwenden. Wenn in Zukunft von einer möglichen Welt die Rede ist, so ist einfach eine kontrafaktische Situation gemeint, die angibt, wie es auch hätte sein oder passieren können. Insofern ist eine mögliche Welt etwas, das unter Umständen durch recht wenige Angaben festgelegt werden kann. Die Beschreibung einer möglichen Welt für die Verteilung eines Pokerblattes kann beispielsweise aus nur fünf Variablen bestehen, deren Werte die Kartenwerte abbilden. Eine mögliche Welt, in der ich vier Asse und einen Joker habe, kann durch diesen Begriff ebenso abgebildet werden, wie ein Blatt, das nur aus Luschen besteht. Der Begriff der möglichen Welt wird hier nur in diesem transparenten Sinne gebraucht, in dem eine mögliche Welt in der Angabe relevanter Eigenschaften einer kontrafaktischen Situation besteht. Es gibt andere Auffassungen dieses Begriffs, die möglichen Welten eine qualitative Natur unterstellen, doch diese sollen uns hier nicht interessieren. ‚Mögliche Welt‘ heißt für uns nur: kontrafaktische Situation.

Der Wertebereich der Intensionsfunktion besteht aus syntaktisch geeigneten Extensionen für den Ausdruck, um dessen Intension es geht. Der Wertebereich für die Intensionen von Gegenstandsbezeichnungen sind die Gegenstände oder die Gegenstandsklassen, die von diesen Ausdrücken bezeichnet werden können. Der Wertebereich für die Intensionen von Prädikaten sind Klassen von Gegenständen, auf die dieses Prädikat zutreffen kann und der Wertebereich für die Intensionen von Sätzen sind Wahrheitswerte.

Beispielsweise ist nun die Intension der Gegenstandsbezeichnung „der erste Bundeskanzler der Bundesrepublik Deutschland“ diejenige Funktion, die *unserer* Welt Konrad Adenauer als Extension zuordnet. In anderen möglichen Welten könnte dieser Ausdruck jedoch aufgrund derselben Intension eine andere Extension haben, zum Beispiel wenn nicht Adenauer, sondern Erhardt der erste Bundeskanzler gewesen wäre. In einer solchen Welt wäre die Ex-

tension des Ausdrucks „der erste Bundeskanzler. . .“ nicht Adenauer, sondern Erhardt.

Die Intension des Prädikats „schwarz“ zum Beispiel ordnet diesem die Klasse aller schwarzen Dinge zu<sup>7</sup> – in unserer Welt fällt daher die Porta nigra in die Extensionsklasse des Prädikats „schwarz“ (wenn man es hinreichend weit versteht). In einer anderen möglichen Welt gibt es möglicherweise keine Porta nigra, sondern beispielsweise eine Porta alba, die nicht in die Extensionsklasse des Prädikats „schwarz“ fällt. Dasselbe Wort kann in verschiedenen Welten verschiedene Extensionen haben – dies ist ein Tatbestand, der mit Carnap ins Licht der allgemeinen Aufmerksamkeit rückt.

Leicht läßt sich dies an der Intension von Sätzen sehen: Der Satz „Heike hat am Samstag chinesisch gegessen“ ist in unserer Welt falsch. Es läßt sich aber eine mögliche Welt festlegen, in der Heike nicht letzten Freitag, sondern eben am Samstag chinesisch essen gegangen ist. Dieser Welt würde die Intensionsfunktion dieses Satzes den Wahrheitswert „wahr“ zuordnen.

In diesem Zusammenhang ist nun auch die in I.4 bereits angedeutete Beziehung zwischen Propositionen und Sätzen klar: ein Satz ist eine Instanz einer Proposition und die Proposition selbst können wir mit ihrer Intension identifizieren, also ist eine Proposition eine Funktion möglicher Welten in Wahrheitswerte.

Diese Explikation des Begriffs der intuitiven Bedeutung – bei Frege durch den „Sinn“ dargestellt – ist ein erheblicher Fortschritt. Nehmen wir den Sinn als kognitive Signifikanz eines sprachlichen Ausdrucks, so ist auch hier der Begriff der Intension wesentlich attraktiver, weil er einer formalen Analyse zugänglich ist. Dabei ist natürlich die Intension lediglich eine Darstellungsform der kognitiven Signifikanz eines Ausdrucks. Der Intensionsbegriff berücksichtigt jedoch keine epistemischen Dimensionen. Gehen wir weiterhin von Freges Philosophie des Geistes aus, so müssen wir annehmen, daß Intensionen von einer denkenden Bewußtseinsinstanz gelernt werden und dann „im Kopf“ sind. Die Intension ist darüber hinaus ein objektiver Bedeutungsbegriff, das heißt, sie orientiert sich an dem ahistorischen semantischen Zustand einer gegebenen Sprache und bildet diesen ab. Sie erklärt keine abweichenden Phänomene, zum Beispiel, wie jemand redet, der Wörter falsch versteht. Zwar läßt Carnaps Semantik Raum für kognitive Theorien, doch sie ist selbst keine Theorie, die zur epistemischen Qualität sprachlicher Ausdrücke Stellung nimmt.

Die zweite wesentliche Verbesserung von Freges Theorie liegt in der Neuformulierung der Prinzipien der Funktionalität und Substitution für Intension und Extension.

Carnap überträgt das Funktionalitätsprinzip der rekursiven Zusammensetzung von Freges Sinnen auf den Intensionsbegriff, das heißt, die Intensionen komplexer Ausdrücke entstehen nach angebbaren Regeln aus der Zusammensetzung der Intensionen ihrer Teilausdrücke. Auch hier ist klarer als bei Frege, was wir uns unter Rekursion vorzustellen haben. In der Darstellung Freges wird nicht transparent, was es heißt, daß der Sinn eines komplexen Ausdruck in dieser Weise zusammengesetzt sein soll. Betrachten wir das Beispiel eines prädikativen Satzes: „ $Fa$ “. Nach Carnap setzt sich dessen Intension aus den Intensionen des Prädikats  $F$  und der Gegenstandsbezeichnung  $a$  zusammen. Die Intension von  $Fa$  ist diejenige Funktion, die jeder Welt genau dann den Wahrheitswert „wahr“ zuordnet, wenn der

---

<sup>7</sup>Farbprädikate sind knifflige Beispiele, denen die Philosophen von jeher besondere Aufmerksamkeit geschenkt haben. Bei solchen Prädikaten bestehen ähnliche Schwierigkeiten, wie zum Beispiel mit Worten wie „schön“.

Gegenstand, der dieser Welt durch die Intension von „*a*“ zugeordnet wird, in derjenigen Menge liegt, die dieser Welt durch die Intension von „*F*“ zugeordnet wird.

In analoger Weise gelten die Prinzipien der Funktionalität und Substitution für den Extensionsbegriff.

In diesem Zusammenhang muß nun darauf hingewiesen werden, daß Carnap zwei Arten sprachlicher Kontexte – das heißt, zwei<sup>8</sup> Arten von Vorkommnissen von Ausdrücken in Sätzen – unterscheidet. Wie Frege zwischen geraden und ungeraden Kontexten unterscheidet, so unterscheidet Carnap zwischen extensionalen und intensionalen Kontexten. Wie müssen wir uns das vorstellen? In ungeraden Kontexten schien das Substitutionsprinzip und das Funktionalitätsprinzip der Bedeutung nicht zu gelten, doch Frege wollte an der universalen Gültigkeit dieser Prinzipien festhalten, weswegen er eine Art Verschiebung der Referenz von Fregescher Bedeutung in Sinn konstatierte: Bei Frege bezeichnete ein Ausdruck in einem geraden Kontext seine gerade Bedeutung und drückte seinen geraden Sinn aus. In einem ungeraden Kontext bezeichnete er seine ungerade Bedeutung (d.i.: sein gerader Sinn) und drückte seinen ungeraden Sinn (was immer man darunter verstehen mag) aus. Diese Theorie führt, wie wir bereits besprochen haben, zu intolerablen Nebeneffekten.

Carnap löst dieses Problem, indem er Extension und Intension nicht die gleiche Abhängigkeit vom Satzkontext zugesteht, wie Frege es bei Sinn und Bedeutung tat. Außerdem schränkt Carnap die Gültigkeit von Funktionalitäts- und Substitutionsprinzip ein. Vielmehr besteht das Substitutionsprinzip für Extensionen bei Carnap in der Substituierbarkeit von koextensionalen Ausdrücken (solchen mit identischer Extension) in allen extensionalen Kontexten unter Erhalt der Extension des Gesamtausdrucks. Diese Substituierbarkeit *salva extensione* heißt im Falle von Sätzen: *salva veritate*, denn die Extensionen von Sätzen sind Wahrheitswerte. Ein Kontext, in dem diese Substitutionsmöglichkeit besteht, heißt extensional. In einem extensionalen Kontext gilt auch das Funktionalitätsprinzip für Extensionen: die Extension des Gesamtausdrucks setzt sich aus den Extensionen der Teilausdrücke zusammen. Extensionale Kontexte sind in diesem Sinne extensionsfunktional.

Ein Kontext heißt dagegen intensional, wenn er (1) nicht extensional und (2) intensionsfunktional ist. Intensionsfunktionalität bedeutet analog, daß die Gesamtintension sich aus den Intensionen der Teilausdrücke zusammensetzt. Es gibt noch eine dritte Art von Kontexten, für die ich hier das Wort hyperintensional verwenden möchte, das Carnap noch nicht gebraucht: dies sind Kontexte, die weder extensional noch intensional sind. Solche Kontexte lassen sich mit Carnaps Theorie nicht analysieren. Ich komme gleich darauf zurück.

Bis zu diesem Punkt besteht Carnaps Theorie nur aus Definitionen: erkenntnisträchtig wäre eine Aussage darüber, welche Kontexte extensional sind und welche nicht. Ein prominentes Beispiel für intensionale Kontexte sind bei Carnap die Modalkontexte, von denen schon die Rede war. In dem Satz „Es ist notwendig, daß 9 größer als 7 ist“ kann man den Ausdruck „9“ nicht durch den extensionsgleichen Ausdruck „die Anzahl der Planeten“ ersetzen, ohne den Wahrheitswert zu verändern. Man kann den Ausdruck „9“ jedoch durch einen der intensionsgleichen Ausdrücke „27 durch 3“ oder „Quadratwurzel aus 81“ ersetzen, ohne daß der Wahrheitswert sich ändert. Zu den intensionalen Kontexten gehören außerdem, wie sich leicht sehen läßt, das umgangssprachliche „wenn *x*, dann *y*“, sowie alle Formen von

---

<sup>8</sup>Wie wir sehen werden, sind es eigentlich mindestens drei.

konditionalen und kausalen Aussagen, auch kontrafaktische Aussagen.

Was hyperintensionale Kontexte angeht, so sind die Zitierkontexte hier das Paradebeispiel. Doch im Gegensatz zu Frege behandelt Carnap auch Einstellungskontexte als hyperintensional, während diese bei Frege im normalen Sinne „ungerade“ waren und man sie seiner Konzeption nach als intensional behandeln müßte.

Betrachten wir uns die beiden folgenden Sätze:

a) Jakob glaubt, daß *A*.

b) Jakob glaubt, daß *B*.

Es ist ohne weiteres vorstellbar, daß a) falsch ist und b) wahr. Wären a) und b) intensionsfunktionale Kontexte, so müßte man aber, sofern *A* und *B* intensionsgleich sind, in a) den Ausdruck *A* durch den Ausdruck *B salva intensione* (und das heißt hier: *salva veritate*) ersetzen können. (Man kann sich das auch anhand des ausgearbeiteten Beispiels in V.2 verdeutlichen.) a) und b) bilden somit keine intensionalen Kontexte. (Carnap hat damit sehr weit gesehen: dieser Aspekt spielt erst wieder in Kapitel V im Zusammenhang mit [Kri79] und seiner Deutung durch Loar eine Rolle. Einen Lösungsansatz für dieses Problem findet sich erst in Kapitel VI bei Stalnaker.)

Nun fügt sich Carnaps Theorie zu folgendem Bild: der Fregesche Bedeutungsbegriff wurde ohne größere theoretische Einschnitte in den Extensionsbegriff transformiert, doch der Sinnbegriff hat eine entscheidende Weiterentwicklung in Form seiner Präzisierung zum Begriff der Intension erfahren. Aus diesen Ausdrücken kann man mithilfe des Funktionalitätsprinzips und des Substitutionsprinzips, die sowohl für Intensionen als auch für Extensionen in extensionsfunktionalen beziehungsweise intensionsfunktionalen Kontexten gelten, eine rekursive Semantik formen. Diese Semantik ist so präzise und weitreichend, daß wir alle in diesem Kapitel desweiteren besprochenen Kennzeichnungstheorien der Namen mit ihrer Hilfe werden abbilden können – erst Donnellans Theorien der unterschiedlichen Gebrauchsweisen von Kennzeichnungen setzt dem ein Ende.

Zwar ist der intuitive Begriff der Bedeutung durch seine Objektivierung zum Intensionsbegriff wesentlich genauer gefaßt, als bisher – und wir haben, was die Angabe der kognitiven Signifikanz angeht, jetzt immerhin schon ein Ordnungsprinzip in Form einer Funktion im algebraischen Sinne an der Hand.

Doch wie sieht es mit dem Problem der Subjektrelativität von Bedeutungen aus? Frege löst es, indem er zwischen Bewußtseinsinhalten und Gegenständen des Denkens trennt. Bei Carnaps formalem Projekt geht es, wie schon gesagt, nicht um philosophische Probleme auf dieser Ebene. Doch ein paar Schlußfolgerungen zu diesem Gebiet unseres Interesses können wir durchaus ziehen:

In [Fr93] diskutiert Frege ausführlich den Sinn des Eigennamens „Dr. Lauben“ und sagt, daß der Sinn dieses Namens für jeden Sprecher ein anderer sei, eben weil Dr. Lauben jedem auf andere Weise gegeben sei. Dazu treten die Bemerkungen über die Sinnschwankungen bei Namen wie „Aristoteles“, die sich in [Fr94] (p.42 Fn 2) finden. Freges Sinnbegriff bleibt subjektrelativ, auch wenn er nichts rein Subjektives ist, sondern, wie oben beschrieben, sich auf der Grenze zwischen Subjektivität und Objektivität befindet.



Von Carnap hören wir in diesem Punkt nicht viel anderes. Die mit einem Eigennamen assoziierten Intensionen können auch hier von Individuum zu Individuum variieren und jede Form von Abweichung zur lexikalischen Definition eines Wortes genießt formal gesehen den gleichen semantischen Status wie die kanonische Bestimmung durch die im Lexikon fixierte Sprachkonvention selbst. Auch durch die Formalisierungen der Sprache  $L$  ist es nicht gelungen, Subjektivität aus der Semantik zu eliminieren und es gibt einen schwerwiegenden Grund, der dieses Ziel unerreichbar zu machen scheint:

Die Realisierung einer rekursiven Semantik fordert eine Theorie der Bedeutung, die auch Sätzen der Form „ $x$  glaubt, daß  $a = b$ “ oder „ $x$  glaubt, daß  $a$  ein  $F$  ist“ systematisch Bedeutungen zuordnet. Für die Wahrheit und die Bedeutung solcher Sätze – wir werden sie später als Überzeugungszuschreibungen kennenlernen – kommt es nicht auf die objektive Bedeutung, ja nicht einmal auf die Wahrheit des daß-Satzes an, sondern darauf, in welcher Beziehung das Individuum  $x$  zu der durch den daß-Satz repräsentierten Proposition steht. In diese Beziehung ist ein Subjekt involviert und die Wahrheitsbedingung solcher Einstellungssätze hängt eben von der Einstellung des epistemischen Subjekts ab, nicht des objektiven Sachverhalts, der im Komplementsatz formuliert wird. Abgesehen davon kann ein Subjekt einen Satz oder Teile eines Satzes auch falsch verstehen und ihn dennoch als Repräsentation einer Proposition auffassen. Das alles heißt aber, daß eine Theorie der Bedeutung, die solchen Sätzen systematisch objektive Bedeutung zuweisen will, nicht ohne eine ergänzende Trägertheorie *subjektiver* Bedeutungen und des subjektiven Verständnisses von Sätzen auskommen kann. Ich will diese Einsicht hier als die Forderung zur Subjektivitätskompetenz der Bedeutungstheorie bezeichnen. Diese Forderung wird in diesem Kapitel nicht eingelöst werden. Wie eine solche Semantik der Überzeugung aussehen kann, werden wir erst bei der Besprechung von Stalnakers Theorie in Kapitel VI erleben.

Man kann anhand der angegebenen Theorien zur Subjektivität der Fregeschen Sinne und der Carnapschen Intensionen auch leicht sehen, wo das Problem für die Kennzeichnungstheorie der Namen sich abzuzeichnen beginnt. Wenn der Kennzeichnungstheoretiker behaupten will, daß die Bedeutung von Namen mit der Bedeutung einer bestimmten Kennzeichnung für den Träger dieses Namens zusammenfällt, dann muß er zugestehen, daß diese Verknüpfung zwischen Name und Kennzeichnung erstens von der jeweiligen möglichen Welt und zweitens vom Subjekt abhängig ist.

Die Subjektabhängigkeit von Namen besteht in den „Schwankungen“ des Sinns, die Frege in der in Abschnitt 2 dieses Kapitels dargestellten Bemerkung über den Namen „Aristoteles“ einräumt. Zwei Sprecher können demnach den gleichen Namen mit verschiedenen Kennzeichnungen assoziieren, wodurch der Sinnbegriff eine subjektive Färbung erfährt. Diese Art von Subjektivität sähen wir gerne ausgeräumt.

Doch die Gleichsetzung von Namen und Kennzeichnungen sieht sich einem weiteren Problem gegenüber: nach Carnaps präzisiertem Intensionsbegriff ist die Bedeutung eines Namens zwar mit der Bedeutung einer bestimmten Kennzeichnung gleichzusetzen, doch nur relativ zum gegebenen Kontext. Das kann man leicht daran sehen, daß es für jeden Namen eine gewisse Anzahl von Kennzeichnungen gibt, deren Intensionen unserer gegebenen Welt dieselbe Extension zuordnen, wie die Intension des Namens. Allerdings heißt das noch nicht, daß der Name und alle diese Ausdrücke intensionsgleich sind. Die Intension der Kennzeichnung „der Verfasser der Kritik der reinen Vernunft“ ordnet unserer Welt das Individuum zu,

das wir außerdem durch den Namen „Immanuel Kant“ bezeichnen. Das heißt aber nicht, daß diese beiden Ausdrücke intensionsgleich sind. Leicht ist zu sehen, daß es mögliche Welten gibt, in denen nicht Kant, sondern vielleicht Moses Mendelssohn oder Herder die Kritik der reinen Vernunft verfaßt haben. Dann würden die Ausdrücke „Kant“ und „der Verfasser der Kritik der reinen Vernunft“ nicht dasselbe Individuum zurückliefern. Welche Kennzeichnungen für welchen Namen stehen, ist somit von der jeweiligen möglichen Welt abhängig.

Mit anderen Worten: die Koreferenz dieser Kennzeichnungen mit Eigennamen ist kontingent. Das ist aber ein gewaltiges Problem, wenn wir sagen wollen, worin die objektive kognitive Signifikanz von Eigennamen besteht, denn nach unserem jetzigen Wissensstand müßten wir sagen: in der Intension. Dann aber müssen wir einräumen, daß Namen und Kennzeichnungen nicht im intuitiven Sinne bedeutungsgleich sein können. Der Kennzeichnungstheoretiker müßte nämlich, will er an der Gleichsetzung festhalten, behaupten, daß die mögliche Welt, in der Herder die Kritik der reinen Vernunft verfaßt hat, eine ist, in der Herder Kant ist. Dieses überaus wichtige Problem findet eine vorläufige Lösung in der Bündeltheorie von Searle, zu der wir bald kommen, doch in aller analytischen Schärfe wird es erst von Kripke aufgegriffen und wie wir sehen werden, wird man es mit den Mitteln Freges, Carnaps und Searles nicht lösen können.

Doch außerhalb dieses Problems der Kennzeichnungstheorie der Namen tritt eine weitere Eigenschaft von Carnaps Theorie in den Vordergrund: *L* markiert – überspitzt formuliert – die Semantik einer quasisolipsistischen Privatsprache. Als einzig relevante Instanz zur Formung von Bedeutungsbeziehungen wird nämlich der Sprecher angenommen, andere Faktoren kommen schlicht nicht vor. Diese Auffassung führt zu Problemen, die sich bereits in Kapitel II andeutungsweise zeigen werden und in Kapitel IV im Rahmen von Putnams Externalismus zur vollen Entfaltung kommen. Zwar weichen Anhänger von Carnaps Semantik normalerweise auf den Standpunkt aus, daß *L* eine ideale Sprache sei, die natürlich eine Exaktheit aufweisen könne, die die vage, natürliche Sprache nicht erreiche. Doch selbst wenn wir von der Idealisierung ausgehen, daß jeder Ausdruck (speziell jedes Prädikat, denn um diese wird es später bei Putnam gehen) nur *eine* Intension hat und insofern nicht vage ist, hat Carnap ebensowenig wie seine Vorgänger etwas dazu gesagt, was es heißt, wenn ein Sprecher eine Intension *erfaßt*. Hat er dann eine Funktion im Kopf? Das mutet reichlich seltsam an. Wir hatten oben schon bemerkt, daß der Intensionsbegriff für die Wiedergabe kognitiver Signifikanz nur eingeschränkt funktionieren kann, was natürlich in dieser Idealisierung seine Ursache hat.

Über die kognitive Rolle der semantischen Begriffe schweigt Carnap sich aus, weshalb wir mit der Frage nach der kognitiven Signifikanz oder dem „Ort“ von Bedeutungen an demselben Punkt stehen, an dem wir mit Frege bereits waren. Wir wissen nach wie vor nicht, wie die Beziehung zwischen Sprecher und objektiver Bedeutung zu explizieren wäre, weil die bisherigen Theorien implizit annehmen, daß Bedeutung internalistisch ist, das heißt, daß die Kenntnis von Bedeutungen ebenso wie Überzeugungseinstellungen und Urteile intrinsische Eigenschaften des Sprechers sind. Hier verläßt uns auch der Begriff der Intension: dieser kann unter anderem deshalb keine Explikation der Bezeichnungsrelation sein, weil die Bezeichnungsrelation ja gerade mithilfe des Intensionsbegriffs definiert worden ist<sup>9</sup>. Es

---

<sup>9</sup>Vergleiche hierzu auch Putnams Kritik der „kalifornischen Semantik“ in [Put75].

muß hier einfach eine andere Möglichkeit geben.

Das alles sind Probleme, die für die kognitive Seite der Bedeutungstheorie nach wie vor schwerwiegend sind, auch wenn wir ihre semantische Seite zunehmend präziser fassen können. Doch bevor wir diese wirklich harten Probleme in Angriff nehmen, wollen wir der Kennzeichnungstheorie noch einmal Chancen einräumen. Wir werden jetzt sehen, wie Strawson sie ergänzt und wie Searle sie so modifiziert, daß sie mit der oben angegebenen Schwierigkeit fertig zu werden scheint.

## **I.7 Zur Kontextabhängigkeit von Äußerungen: Strawson**

In seinem Aufsatz „On Referring“ (= [Str50]) kritisiert P. F. Strawson die Kennzeichnungstheorie Russells. Er führt dabei zwei wichtige neue Aspekte in die Diskussion um Bedeutungsbegriffe ein, auf die ich hier eingehen möchte, weil sie meines Erachtens der Diskussion eine wichtige neue Wendung geben, erstens die Berücksichtigung pragmatischer Elemente für die Bedeutungstheorie, zweitens die Unterscheidung zwischen Inhalt beziehungsweise Implikationen von Äußerungen einerseits und deren Präsuppositionen andererseits.

Der erste wichtige Aspekt liegt in der Berücksichtigung von sprachpragmatischen Überlegungen zur Relevanz von Bedeutungsbegriffen. Ich hatte in meiner Bemerkung zu Carnaps Semantik darauf hingewiesen, daß die drei bisher besprochenen Theoretiker ein idealsprachliches Programm verfolgen: Carnaps Semantik orientiert sich an wissenschaftlichen Verwendungsmöglichkeiten von Sprache, Frege will im Rahmen seines Logizismusprogramms eine Sprache des reinen Denkens schaffen und auch Russell verfolgt dieses Ziel, wobei „On Denoting“ auch in Beziehung zu seiner Philosophie des logischen Atomismus zu sehen ist. Dieses Idealsprachenprogramm endet mit Strawson, der die normale, gesprochene Sprache als Gegenstand der Bedeutungstheorie ansieht.

Diese Berücksichtigung der gesprochenen Sprache führt Strawson zu der Erkenntnis, daß die Referenz vieler Gegenstandsbezeichnungen durchaus von sprachpragmatischen Komponenten abhängen kann – eben vom Kontext, in dem diese Äußerung gemacht wird. Unter einem Kontext ist dabei, nicht wie bisher, einfach der grammatische beziehungsweise logische Zusammenhang des Satzes zu verstehen, sondern der Realitätszusammenhang, in dem eine Äußerung gemacht wird. Typische Kontextaspekte sind zum Beispiel der Sprecher, der Ort, die Zeit und sofort. Wenn ich zum Beispiel sage: „Der Mieter, der im Parterre wohnt, holt gerade seine Post herein“, so ist, wen ich mit dem kennzeichnenden Ausdruck „der Mieter, der im Parterre wohnt“ meine, davon abhängig, wer dort wohnt. Dies wiederum kann sich über die Jahre ändern. Äußere ich diesen Satz 1996, so meine ich vielleicht Frau Schmidt, ein Vierteljahr später kann mit dem gleichen Satz Herr Kugaczinsky gemeint sein. Die Zeitbestimmung, die zur Festlegung der Referenz entscheidend ist, muß unter Zuhilfenahme des Kontexts ergänzt werden.

Oder noch deutlicher: Meine Schwester kommt nach Hause und ich frage: „War’s gut?“ und sie antwortet „Ja, war toll.“ Nur wir beide wissen, daß sie auf der Geburtstagsfeier einer Freundin war und der Ausdruck „es“, den ich verwendete, ihre Teilnahme an dieser Festivität bezeichnet. Ebenso könnte ich jedoch mit dem gleichen Satz fragen, ob sie ihre Zwischenprüfung bestanden oder eine Stelle bekommen hat. All dies hängt vom jeweiligen

Gebrauchskontext des sprachlichen Ausdrucks ab.

Der Ausdruck „es“ ist – genauso wie „ich“, „gestern“, „hier“ und andere – offenkundig indexikalisch, das heißt, ihre Referenz ist bei jeder Verwendung vom Kontext abhängig. Um diese Wörter wird es in Kapitel III gehen. Doch auch die Referenz von Gegenstandsbezeichnungen kann, wie die obigen Beispiele zeigen, kontextabhängig sein.

Strawson trägt der Kontextabhängigkeit der Referenz (in unserem Beispiel: bei Gegenstandsbezeichnungen) Rechnung, indem er drei Dinge unterscheidet:

- a) den Satz
- b) die Verwendung beziehungsweise den Gebrauchsfall eines Satzes
- c) die Äußerung eines Satzes.

Eine analoge Unterscheidung trifft er für Ausdrücke. Der Witz dieser Unterscheidung liegt darin, daß ein Satz viele Verwendungen haben kann und eine Verwendung in vielen Äußerungen vorkommen kann. Zum Beispiel kann ich sagen: „Ich mag meinen Klassenlehrer nicht“. Wenn ich das in der fünften Klasse sage, dann unterscheidet sich diese Gebrauchsweise von dem Gebrauch, den ich mit dieser Äußerung mache, wenn ich in der achten Klasse bin. Die Referenz eines Ausdrucks kommt demnach nicht durch die syntaktische Konstruktion des Satzes allein zustande, sondern erst, indem wir diesen Satz in bestimmter Weise gebrauchen. Das bedeutet eine grundlegende Korrektur des bisherigen Programms der rekursiven Semantik. Wir müssen nämlich nun umdenken und sagen, daß auch Wahrheit und Falschheit keine Eigenschaften von Sätzen, sondern von Gebrauchsfällen, das heißt einzelnen Verwendungen von Sätzen sind. Der Satz selbst ist nicht wahr oder falsch, sondern wir *gebrauchen* ihn, um eine Aussage zu machen, die zusammen mit ihrem Kontext, der den Bezug festlegt, wahr oder falsch wird.

Daß diese recht einleuchtend klingenden Dinge erst hier eingeführt werden, hängt mit der idealsprachlichen Perspektive der bisher besprochenen Positionen zusammen. Mit Strawson verschiebt sich der Schwerpunkt der Diskussion um Bedeutungen endgültig von der idealsemantischen Theorie auf die gesprochene Sprache, deswegen werden diese Aspekte jetzt *wichtig*.

Mit dieser perspektivischen Korrektur verbindet sich jedoch auch eine Veränderung des Bedeutungsbegriffs selbst: für Strawson ist die Bedeutung eines gegebenen Ausdrucks nicht mehr eine Eigenschaft, die ausschließlich durch eine Intension ausgedrückt wird, sondern eine Menge von sprachlichen Regeln und Konventionen, die den korrekten Gebrauch des Ausdrucks festlegen. In Bezug auf unser permanentes Beispiel der Gegenstandsbezeichnungen heißt das, daß wir, um die Bedeutung eines solchen Ausdrucks angeben zu können, angeben müssen, wie man diesen Ausdruck korrekt verwendet, um sich auf ein Objekt zu beziehen<sup>10</sup>. Natürlich heißt dies auch, daß wir verunsichert werden, was die kognitive Signifikanz von sprachlichen Ausdrücken angeht, denn diese bestünde dann in der Beherrschung der Gebrauchsregeln. Die Intensionsfunktion müßte also im Sinne der Kenntnis der sprachlichen Konventionen ergänzt werden, um die kognitive Signifikanz von Ausdrücken

<sup>10</sup>Eine ähnliche Auffassung findet sich auch an sehr prominenter Stelle, nämlich bei Wittgenstein in den „Philosophischen Untersuchungen“, vgl. [Wit95].

abbilden zu können. Bei Strawson scheint die kognitive Signifikanz eines Ausdrucks in der Kenntnis seiner korrekten Gebrauchsmöglichkeiten zu liegen – eine Position, die dem späten Wittgenstein nahesteht.

Strawson faßt leider nicht genauer, wie wir uns die Regeln und Konventionen vorzustellen haben, die die korrekten Gebrauchsmöglichkeiten festlegen, so daß diese Theorie etwas vage dasteht, doch wir werden in Kapitel III eine Vorstellung erarbeiten, wie solche Regeln jeweils auszusehen haben.

Der zweite wesentliche Aspekt, den Strawson in die Bedeutungsdebatte einführt, ist die Unterscheidung zwischen der logischen Implikation eines Satzes einerseits und andererseits dem, was ich im folgenden als Präsupposition eines Satzes bezeichnen möchte. Dieser Aspekt, auf den ich gleich genauer eingehen werde, ist ein Resultat von Strawsons Auseinandersetzung mit Russells Problem vom ausgeschlossenen Dritten. Wir erinnern uns: Sätze mit leeren Gegenstandsbezeichnungen geraten in Konflikt mit dem Satz vom ausgeschlossenen Dritten, weil sie weder wahr noch falsch sein zu können scheinen. Russell löste dies, indem er ganz einfach behauptete, daß solche Sätze eigentlich Existenzaussagen sind und ein Satz der Form „es gibt ein  $x$ “ immer wahr oder falsch ist, während wir, falls wir anerkennen, daß es sich um eine Subjekt-Prädikat-Aussage handelt, der angegebenen Schwierigkeit ins Auge sehen.

Strawson stützt Freges Argument, daß Sätze mit leeren Gegenstandsbezeichnungen nicht wahrheitswertfähig sind. Nach Strawson wäre dies ein Gebrauchsfall eines Satzes, durch den keine Aussage beziehungsweise Behauptung gelingt – eben weil es das, worüber gesprochen wird, nicht gibt. Das heißt aber, Strawson muß, wenn er mit dieser Auffassung Erfolg haben will, Russells Ansicht, es handle sich bei Prädikationen um verkappte Existenzsätze, zurückweisen. In diesem Zusammenhang nun ist die Unterscheidung zwischen Implikation und Präsupposition wichtig.

Satz  $A$  ist eine (logische) Implikation des Satzes  $B$  genau dann, wenn  $B$  nicht wahr sein kann, ohne daß auch  $A$  wahr ist. Hingegen ist Satz  $A$  eine Präsupposition von  $B$  genau dann, wenn  $A$  wahr sein muß, damit  $B$  überhaupt wahrheitswertfähig wird – man sagt auch,  $B$  präsupponiere  $A$ . Diese Unterscheidung kann man auf Sätze ebenso wie auf Gebrauchsfälle von Sätzen anwenden.

Präsuppositionen sind nicht schwer vorzustellen: beispielsweise präsupponiert der Satz „Kohl hat aufgehört zu arbeiten“, daß Kohl irgendwann einmal gearbeitet hat – und er präsupponiert, daß es den gibt, den ich als „Kohl“ bezeichne. Doch behauptet der Satz selbst die Existenz einer Entität namens Kohl?

Contra Russell antwortet Strawson, daß der Gebrauch einer Gegenstandsbezeichnung in einem Satz den Bezug dieser Gegenstandsbezeichnung auf genau einen Gegenstand eben präsupponiert und nicht impliziert, das heißt, daß die Wahrheitswertfähigkeit, nicht die Wahrheit des Satzes, in dem dieser Gebrauch vorkommt, von dieser vorgeschalteten Existenzbehauptung abhängig ist.

Wie Frege, so nimmt auch Strawson mit der Einführung des Präsuppositionsbegriffs in Kauf, daß der Rahmen der klassischen, zweiwertigen Logik durch seine Bedeutungstheorie gesprengt wird. Das heißt natürlich, daß sich hier ein weites Feld von Arbeit öffnet, doch sich darauf umzusehen, ist hier nicht unsere Aufgabe. Wie wir gesehen haben, löst Strawson das Problem vom ausgeschlossenen Dritten ähnlich wie Frege: es wird negiert.

Das Problem negativer Existenzsätze behandelt Strawson in „On Referring“ nicht. Dafür hat er aber einen sehr transparenten Lösungsansatz für das Problem der Nichtapriorität von Identitätssätzen anzubieten, die mindestens eine Kennzeichnung enthalten:

Nehmen wir an, ich sage: „Der Mann da ist der, der von der Polizei gesucht wird“. Frege würde das Wort „ist“ in diesem Satz als das Symbol „=“ verstehen und den Satz als einen der Form „ $a = b$ “ auffassen, was zu den bekannten Problemen führt. Strawson macht in diesem Zusammenhang einen Vorschlag, der ebenso einfach wie verblüffend ist:

„Der Mann da ist der, der von der Polizei gesucht wird“ soll nicht als Identitätssatz, sondern als Prädikation gelesen werden, also als: „Der Mann da wird von der Polizei gesucht“. Diese beiden Sätze haben nach Strawson die gleiche intuitive Bedeutung und die gleichen Wahrheitsbedingungen. Der Unterschied liegt nur in ihren Präsuppositionen: der erste Satz präsupponiert, daß genau einer von der Polizei gesucht wird, der zweite Satz präsupponiert das nicht – und beide Sätze sind (nicht nur für die Kripo) informativ. Damit ist das Nichtaprioritätsproblem verschwunden. Diese Lösung funktioniert allerdings nur für Identitätssätze mit Kennzeichnungen, nicht für solche, die nur Eigennamen enthalten.

Wie kann man nun Strawsons Vorschlag in Carnaps System darstellen? Wir erinnern uns an den Carnapschen Begriff der Intension als logische Transformation des intuitiven Bedeutungsbegriffs. Eine Intension ist eine Funktion von möglichen Welten in Extensionen.

Die Carnapsche Semantik übersteht den Wegfall der zweiwertigen Logik: Strawsons Vorschlag für die Wahrheitswertlosigkeit von Sätzen mit leeren Gegenstandsbezeichnungen könnte man abbilden, indem man den Intensionsbegriff von Gegenstandsbezeichnungen verändert. Wir erinnern uns, daß die Carnapsche Intension eines Begriffs für jeweils alle Welten definiert ist, das heißt, sie ordnet jeder möglichen Welt für den betreffenden Begriff eine Extension zu. Sofern nun aber, wie Strawson sagt, eine Aussage, die eine Gegenstandsbezeichnung enthält, mit einer Existenzpräsupposition verbunden ist, müßte man einräumen, daß die Intension dieser Gegenstandsbezeichnung für alle Welten, in denen diese Existenzpräsupposition nicht erfüllt ist, der Gegenstandsbezeichnung auch keine Extension zuordnen kann, denn diese gibt es ja nicht. Also kann man Strawsons Argumente in Carnaps Theorie einbauen, indem man die Intensionen von Gegenstandsbezeichnungen nicht wie Carnap als total, sondern als partiell definiert auffaßt: für alle Welten, in denen die Existenzpräsupposition(en) nicht erfüllt ist (sind), liefert die Intensionsfunktion keinen Wert, das heißt, für diese Welten ist sie nicht definiert. Entsprechendes gilt für die Intension von Sätzen: zwar ordnet die Intensionsfunktion eines Satzes diesem nun nach wie vor nur die beiden Wahrheitswerte „wahr“ oder „falsch“ zu, doch sie ordnet eben nicht *jedem* Satz einen Wahrheitswert zu.

In gewisser Hinsicht hat durch Strawson eine Bewegung stattgefunden: der Schwerpunkt der Betrachtung hat sich von dem einseitigen Interesse an formalen Systemen fort zur gesprochenen Sprache hin verlagert. Die Unterscheidung von Sätzen, deren Gebrauch und deren Äußerung dokumentiert dies.

Die semantischen Fortschritte, die wir mit Strawson gemacht haben, sind dagegen gering. Auch was die Frage nach der kognitiven Signifikanz von Wörtern – oder wenigstens von Namen – angeht, hat sich, außer daß wir den Definitionsbereich der Intensionsfunktion beschnitten haben, nicht viel getan. Doch hier tritt nun die Bündeltheorie auf den Plan.

## I.8 Die Bündeltheorie der Bedeutung von Namen: Searle

In diesem Abschnitt werden wir noch einmal zu der grundsätzlichen Frage zurückkehren, ob Gegenstandsbezeichnungen – hier geht es jetzt speziell um Namen – Sinn haben und falls ja, wie dieser Sinn aussehen könnte. Frege hatte diesen Begriff eingeführt, Russell hatte ihn wieder abgeschafft, Carnap hat ihn durch eine formale Konstruktion ersetzt und Strawson hat (sozusagen) seine Gültigkeit wieder eingeschränkt. Wir haben uns aber entschlossen, die Konstruktion der Intension, die aus Freges Sinnbegriff erwuchs, vorerst einmal mit der kognitiven Signifikanz von sprachlichen Ausdrücken gleichzusetzen. Wie wir aber (gerade) feststellen mußten, wissen wir immer noch erstaunlich wenig über diesen seltsamen und bemerkenswerten Aspekt von Sprache. Dies ändert sich aber mit der Theorie, die ich jetzt darstellen möchte, denn diese hat es sich zum Ziel gesetzt, die Beschaffenheit des Sinns von Eigennamen zu erklären.

Die Kennzeichnungstheorie der Namen kulminiert (historisch letztmalig) in einer philosophischen Position, die als Bündeltheorie des Sinns von Eigennamen bekannt geworden ist und ihren locus classicus in dem Aufsatz „Proper Names“ (= [Sea58]) von John Searle gefunden hat.

Zunächst: Wie kommt ein Eigenname zu seinem Gegenstandsbezug? Die Antwort auf diese fundamentale Frage muß auch den Startpunkt zum Projekt der rekursiven Semantik geben, denn wenn wir wissen, wie das Ding und sein Name verknüpft sind, dann wissen wir auch, was der Sprecher weiß, wenn er einen Namen *kennt*, und wie diese Verknüpfung Modell für eine Theorie der Bedeutung anderer Namen und der rekursiven Zusammensetzung von Bedeutungen sein kann. Wie also bezieht sich ein Name genau auf das Objekt, als dessen Name er verwendet wird? Searle sieht drei Möglichkeiten:

Erstens: Der Bezug wird über andere Namen für denselben Gegenstand hergestellt. Als Lösung taugt diese These nicht, denn selbst wenn es so wäre, ist damit erst nicht erklärt, wie der Bezug von Namen aussieht. Das Problem wird nur verschoben.

Zweitens: Der Bezug eines Namens zu dem von ihm bezeichneten Gegenstand wird durch eine Ostension oder Demonstration hergestellt, indem wir auf den Gegenstand zeigen und so den Bezug des gebrauchten Namens klar machen. Auch dies bringt uns nicht weiter, denn wir benennen sehr, sehr viele Dinge, auf die wir aus vielen Gründen nicht zeigen können, zum Beispiel eine Sonate, eine Utopie, einen Trocknungsvorgang, eine Beschleunigung, einen Zeitpunkt undsofort. Und selbst die Dinge, auf die wir zeigen *könnten*, sind seltenst gerade greifbar (Napoleon?).

Für Searle besteht die Lösung zu dieser wichtigen Frage in einem dritten Punkt, den wir bereits kennen: der Bezug eines Namens auf einen Gegenstand wird durch Kennzeichnungen für diesen Gegenstand ermöglicht. Dadurch, daß wir auf einen Gegenstand durch einen ihn kennzeichnenden Ausdruck referieren, identifizieren wir ihn anhand einer seiner Eigenschaften. Bei Russell war es noch so, daß eine einzige Kennzeichnung hier entscheidend ist, und der Name diese Kennzeichnung sozusagen abkürzt. Jedenfalls führt die These, der Bezug des Namens werde über Kennzeichnungen hergestellt, zu dem Ergebnis, daß der Name infolgedessen unauflösbar mit den Eigenschaften des bezeichneten Objektes verknüpft ist. Die Tatsache, daß diese Eigenschaften mit dem Namen assoziiert sind, ermöglicht erst den Gegenstandsbezug. Searle argumentiert, daß sich auf diese Weise schließen lasse, daß

Namen aus mehr bestehen, als nur dem Bezug auf ihr Denotat: offenbar haben sie darüber hinaus das, was wir intuitiv eine „Bedeutung“ nennen und was bei Frege „Sinn“ hieß.

Wir müssen hier aufpassen: wenn die Bedeutung eines Namens durch Kennzeichnungen gegeben ist, darf man dennoch nicht in die Falle tappen, zu behaupten, daß diese Kennzeichnungen insofern synonym mit dem Namen verwendet werden, als daß sie Bestandteil des Inhalts von Behauptungen wären, die mit diesem Namen gemacht werden. Zwar mag es solche Fälle geben, doch denken wir an die intensionalen Kontexte: Ludwig will wissen, ob der Abendstern der Morgenstern ist. Das muß nicht heißen, daß Ludwig wissen will, ob der Stern, der in der-und-der Position abends zu sehen ist, der Morgenstern ist. Vielleicht hat er nur über den Abendstern gelesen und ist nicht in der Lage, ihn am Abendhimmel zu identifizieren. Falls doch, so will er aber doch ganz sicher deshalb nicht wissen, ob die Venus die Venus ist. Nein, soweit waren wir schon.

Searle umgeht dieses Problem unter Bezugnahme auf die Begrifflichkeiten Strawsons. Daß die Kennzeichnungen oder Beschreibungen, durch die die Bedeutung eines Namens gegeben sein soll, auf den Namensträger zutreffen, ist für Aussagen mit diesem Namen keine Implikation sondern eine Präsupposition. Die Implikationen solcher Aussagen beziehen sich ausschließlich auf den Träger des Namens, von dem präsupponiert wird, daß er sich in bestimmter Weise beschreiben läßt.

Und wie läßt sich nun für die Bedeutung eines solchen Namens eine positive Charakterisierung geben? Es ist inakzeptabel anzunehmen, daß ein Name mit der Gesamtheit der Kennzeichnungen, die auf seinen Träger zutreffen, synonym sei. In diesem Fall gehörte diese Gesamtheit der Kennzeichnungen zur Präsupposition der Aussage, die den Namen enthält. Das würde bedeuten, daß wir gar nichts Informatives mehr über den Namensträger aussagen könnten, da alle wahren Aussagen über ihn analytisch und deshalb a priori wären.

Beschränkt man die Verknüpfung auf ein Minimum, so hieße dies, der Name stehe nur für etwas, auf das zutrifft, daß es so genannt wird. Eine Aussage, die den Namen „Aristoteles“ enthält, präsupponiert also nur, daß es einen gibt, der so genannt wird oder wurde. Das ist auch unattraktiv, denn es kann viele Träger eines Namens geben – man stelle sich die Präsupposition einer Aussage über Herrn oder Frau Schmidt vor.

Die Funktion der Bedeutung eines Namens ist es, so Searle, uns Kriterien zur Identifizierung des Namensträgers zu geben. Es liegt an der Vagheit und ungenauen Bestimmung dieser Kriterien, daß die Bedeutung eines Namens so schwer zu charakterisieren ist. Wie identifizieren wir Aristoteles? Er muß nicht der Lehrer Alexanders gewesen sein. Das könnte sich als falsch herausstellen und dennoch würden wir nicht sagen, daß es Aristoteles nicht gegeben habe. Für jede Kennzeichnung, die solche Eigenschaften bezeichnet, läßt sich argumentieren, daß es akzidentelle, das heißt *kontingente*, nicht notwendige Eigenschaften sind, die der Träger auch nicht hätte haben können<sup>11</sup>. Also kann, wie ich schon in I.6 angedeutet habe, auf jeden Fall nicht *eine* solche Kennzeichnung die Bedeutung eines Namens

---

<sup>11</sup>Es sei erwähnt, daß an diesem Punkt die Bedeutungstheorie in Ontologie umschlägt, denn hier müssen wir fragen, ob ein Gegenstand überhaupt essentielle Eigenschaften hat und falls ja, welche seiner Eigenschaften essentiell für ihn sind und welche nicht. Zu diesem Gebiet gibt es eine breite und interessante Diskussion unter den Philosophen, auf die wir hier nicht eingehen können, doch es ist interessant, festzuhalten, daß überall da, wo Modalkategorien angewendet werden, diese Verbindung von Semantik und Ontologie sich zeigt.



einfangen.

Andererseits müssen wir zugeben, daß wir, wenn nun alle Kennzeichnungen, die sich auf kontingente Eigenschaften von Aristoteles beziehen, sich nach und nach als falsch herausstellten – daß er Bücher schrieb, daß er der Schüler Platons war, der Lehrer Alexanders undsofort – geneigt wären, zu sagen, daß uns die Bedeutung von „Aristoteles“ verloren geht. Trifft etwa ein großer Teil der Kennzeichnungen auf jemand anders zu und stellen wir fest, daß etwa in Wirklichkeit Antisthenes oder Diogenes (überraschenderweise!) das meiste oder vieles von dem gemacht haben, was Aristoteles zugeschrieben wird, dann müssen wir wohl einräumen, daß es den historischen Aristoteles, wie wir ihn uns denken, nicht gab.

Dem trägt Searle Rechnung, indem er Namen als Bündelbegriffe [*cluster concepts*] bezeichnet. Die Bedeutung eines Namens liegt in der Mehrheit der Kennzeichnungen und Annahmen über den Namensträger. Es ist dabei weder ausgeschlossen, daß manche Kennzeichnungen in dieser Hinsicht schwerer wiegen als andere, noch läßt sich sagen, wie groß die Mehrheit der wahren Kennzeichnungssätze über den Träger sein muß, um die Bedeutung für fixiert halten zu können. All das bleibt bei Searle sehr vage, doch diese Vagheit, so sein Argument, liegt in der Natur der Sache und läßt sich nicht weiter eliminieren. Da wir oft sehr unbestimmte Kriterien für die Identifizierung von Gegenständen haben, brauchen wir für diese vagen Bündel Etiketten, die sie abkürzen. (Wenn wir ein eindeutiges Kriterium zur Identifikation hätten, würden wir dieses benutzen.) Auch hier also finden wir uns in bester Tradition der Kennzeichnungstheorie der Bedeutung.

Die Kernthese der Bündeltheorie ist also: Namen, beziehungsweise Aussagesätze, in denen solche Namen vorkommen (ausgenommen Zitate etc.) präsupponieren, daß das meiste, was wir über den Namensträger annehmen und womit wir ihn kennzeichnen, tatsächlich auf ihn zutrifft.

Übertragen in die Carnapsche Semantik heißt das: die Bedeutung eines Namens „*N*“ ist die partiell definierte Intensionsfunktion, die jeder möglichen Welt *w* denjenigen Gegenstand zuordnet, der in *w* die meisten unserer Annahmen der Form „*N* ist der *F*“, „*N* ist ein *G*“, „*N* ist *p*“ undsofort erfüllt, sofern es in *w* genau einen solchen Gegenstand gibt, und die für alle anderen möglichen Welten nicht definiert ist.

Man beachte, daß es „die meisten *unserer* Annahmen“ heißt, nicht: „*meiner* Annahmen“. Gemeint ist hier das „wir“ der Sprachgemeinschaft, die diesen Namen benutzt. Searle will auf diese Weise gewährleisten, daß die leidige Tendenz, Bedeutungen zu subjektivieren, von seiner Theorie vermieden wird (dies gelingt zwar leidlich, doch steckt sie dafür voller Vagheiten). Wenn *meine* Annahmen über Aristoteles, den ich zum Beispiel für einen neapolitanischen Pizzabäcker halten könnte, der in den Achtzigern zumachen mußte, sich als falsch herausstellen, dann heißt das für die Bedeutung des Namens „Aristoteles“ noch gar nichts. Erst wenn die *kollektiven* Annahmen über Aristoteles sich als entscheidend von der Realität abweichend herausstellen, kommt die etablierte Bedeutung zu Fall.

Wir haben die Intension eines Wortes als Abbildung der kognitiven Signifikanz ihrer intuitiven Bedeutung verstanden. Jetzt sehen wir, daß die Sprachgemeinschaft offenbar bei der Festlegung der Bedeutung eine Rolle spielt. Ist das am Ende eine Kehrtwende in der Bedeutungstheorie?

Nein, denn nach wie vor setzen sich die kollektiven Annahmen der Sprachgemeinschaft aus den Annahmen Einzelner zusammen. Searle verallgemeinert – man könnte sogar sagen:

abstrahiert – den Internalismus insofern, doch er weist ihn nicht zurück. Jedes Mitglied der Sprachgemeinschaft kann zwar Namen falsch verstehen oder anderen Anwendungen unterliegen, die es vom korrekten Gebrauch des Namens ausschließen – doch letztendlich kommt die Bedeutung aus den Köpfen der Sprecher, in Form von Annahmen und Beschreibungen, die sie über den Namensträger erwerben oder lernen. Die Tatsache, daß die Bedeutung kollektives Eigentum ist, heißt nur, daß sie sich in allen Köpfen und nicht nur in einem oder einigen ändern muß, um als geändert gelten zu können. Das ist dieselbe internalistische Theorie, die wir die ganze Zeit schon verfolgen, nur ist es eine Art kollektiver Internalismus. (Eine genauere Besprechung der Bündeltheorie – wahrscheinlich die genaueste, die es gibt – werden wir in Kapitel II erleben, vorgestellt von S. A. Kripke, der diese Theorie buchstäblich „auseinander nimmt“.)

Wir haben Carnaps Semantik gut zur Abbildung der Ergänzungen gebrauchen können, die die Bedeutungstheorie seit [Ca47] erfahren hat. Sie hat deshalb so gut funktioniert, weil wir bisher immer von einer Privatsprache ausgegangen sind und selbst bei Searle mußten wir am formalen Instrumentarium nichts ändern. Zwar sind wir mit Strawson von der Idealsprache zur „Normalsprache“ übergegangen, doch das ändert nichts daran, daß alles, was bisher gesagt wurde, auch auf eine Sprache zutreffen könnte, die nur ein einziger Mensch spricht – Searles Theorie würde dies als einen Sonderfall betrachten, in der die Anzahl  $n$  der Sprecher einer Sprachgemeinschaft eben 1 wäre. Doch behandeln ließe sich ein solcher Fall ohne Probleme, was bedeutet, daß der einzelne Sprecher, respektive der Kopf des einzelnen Sprechers bisher das einzige war, in dem sich Bedeutung abspielte. Beziehungen kausaler Art zwischen Sprechern, zum Beispiel die Tatsache, daß einer Verwendung eines Satzes ein *intentionales* Motiv zugrundeliegt, wurden von jeder Betrachtung ausgeblendet. In dem Maße, in dem die Bedeutungstheorie sich immer mehr der gesprochenen Sprache zuwendet, wird diese Prämisse immer obskurer. Bevor ich jedoch zeige, wie sie letztlich unter Beschuß gerät, will ich im nächsten Abschnitt noch etwas über die Unterscheidung von Verwendungsarten von Kennzeichnungen sagen.

Im nächsten Abschnitt werden wir uns auch von Carnaps Semantik verabschieden, die die im folgenden dargestellten Veränderungen nicht mehr abbilden kann. Dieser letzte Abschnitt des Kapitels bespricht die Unterscheidung von Verwendungsarten von Kennzeichnungen, die von Keith S. Donnellan in [Don66] eingeführt wurde, und der Theorie, die wir in Kapitel II besprechen, bereits den Boden bereitet hat, während der Stern der Kennzeichnungstheorie zu sinken beginnt.

## I.9 Das Ende der Carnapschen Semantik: Donnellan

In seinem Artikel „Reference and Definite Descriptions“ geht Donnellan über Strawsons Kritik an Russell und die Möglichkeiten des formalen Rahmens der Carnapschen Semantik hinaus, indem er zwei Verwendungsweisen kennzeichnender Ausdrücke unterscheidet. Ich möchte betonen, daß es bei Donnellan ausschließlich um Kennzeichnungen geht, genauer gesagt: um *Verwendungen* von Kennzeichnungen.

Nach Donnellan kann eine Kennzeichnung auf zwei Arten verwendet werden: zum einen in *attributiver* Weise, sofern etwas über den Träger der Eigenschaft ausgesagt werden soll,

die der kennzeichnende Ausdruck beschreibt, egal wer dieser Träger ist, zum anderen in *referentieller* Weise, wenn der Sprecher sich mithilfe des kennzeichnenden Ausdrucks auf einen ganz bestimmten Gegenstand beziehen möchte, von dem etwas ausgesagt werden soll.

Wie hat man sich das vorzustellen? Man denke an eine Schulklasse, in deren Klassenraum eine Scheibe zersprungen ist. Die Lehrerin sagt:

S11) „Der Schüler, der die Scheibe eingeworfen hat, soll sich melden.“

In diesem Fall liegt klarerweise eine attributive Verwendung des kennzeichnenden Ausdrucks „der Schüler, der die Scheibe eingeworfen hat“ vor. Die Lehrerin will sich mit diesem Ausdruck nicht auf ein einzelnes, bestimmtes Individuum in der Klasse beziehen, auf das sie auch zeigen könnte. Vielmehr will sie – deswegen finde ich das Beispiel so gut – feststellen, welcher Schüler Träger der genannten Eigenschaft ist und wer immer diese Eigenschaft erfüllt, für den und für keinen anderen gilt, daß die Lehrerin will, daß er sich meldet.

Nehmen wir weiterhin an, der betreffende Schüler (oder die Schülerin?) hat Charakter und tritt vor. Die Lehrerin erkennt in dem Übeltäter das Problemkind der Klasse und sagt *zu ihm oder ihr*:

S12) „Also das Kind, das hier einfach eine Scheibe einwirft, wird sich jetzt vor dem Direktor zu verantworten haben.“

In diesem Fall verwendet sie den kennzeichnenden Ausdruck „das Kind, das hier einfach eine Scheibe einwirft“ in *referentieller* Weise, denn sie bezieht sich auf ein ganz bestimmtes Individuum, von dem sie etwas aussagen will und der Bezug zu dieser einzelnen Gegenstandsinstanz wird durch die *referentiell* verwandte Kennzeichnung hergestellt.

Um dieses Beispiel nicht falsch zu verstehen, sei gesagt, was auch Donnellan besonders wichtig ist, nämlich, daß es in Bezug auf *Sätze* keinerlei *semantische* Merkmale gibt, anhand derer sich bemerken, zeigen oder beweisen ließe, ob die eine oder die andere Verwendungsweise vorliegt. Man sieht es einem Satz, dessen Äußerungskontext man nicht kennt, nicht an, ob er *referentielle* oder *attributive* Verwendung gefunden hat.

Dadurch wird Strawsons Betonung der Wichtigkeit pragmatischer Komponenten bei der Festlegung der Bedeutung von sprachlichen Ausdrücken gestützt und verstärkt. In der Tat kann man nur durch Kenntnis des Äußerungskontexts des Satzes sehen, ob eine *referentielle* oder *attributive* Verwendung vorliegt. Erst der Satz *in* seinem Kontext ermöglicht es uns, diese Unterscheidung vorzunehmen.

Zur Unterscheidung zwischen *attributiver* und *referentieller* Verwendung von Kennzeichnungen gibt Donnellan vier Merkmale an:

1) Die Wahrheitsbedingungen eines Satzes, in dem eine Kennzeichnung *referentiell* verwendet wird, unterscheiden sich von denen eines syntaktisch identischen Satzes, in dem dieser Ausdruck *attributiv* verwendet wird. In der *referentiellen* Lesart des Satzes „der *F* ist *G*“ wird von einem bestimmten Gegenstand gesagt, er sei *G*, nämlich jener, den wir durch „der *F*“ kennzeichnen. Der Satz kann in dieser Verwendungsweise sogar wahr sein, wenn

das Referenzobjekt gar kein  $F$  oder nicht der einzige  $F$  ist. Entscheidend ist die vom Sprecher intendierte Referenz. Wenn er einen bestimmten Gegenstand herausgreift, über den er etwas aussagt, so glückt die Referenz. Mithin geht in die Wahrheitsbedingung des Satzes der Gegenstand ein, den wir meinen. In attributiver Lesart ist die Sache anders: hier wird nicht von einem bestimmten Gegenstand, sondern von dem, der  $F$  ist, etwas ausgesagt, welcher auch immer es sei.

Nehmen wir an, der Schüler, der vorgetreten ist, hat die Scheibe gar nicht eingeworfen, sondern will einen Freund decken, dessen Eltern nicht haftpflichtversichert sind. In diesem Fall ist die Äußerung S12 *wahr*, wenn sie *referentiell* interpretiert wird: die Lehrerin meint, daß genau dieser vor ihr stehende Schüler, den sie situationsgebunden als den (vermeintlichen) Scheibeneinwerfer herausgreift, sich vor dem Direktor wird verantworten müssen. Liest man den Satz jedoch *attributiv*, so ist er *falsch*, denn wer immer die Scheibe eingeworfen hat, sitzt unerkannt auf seinem Platz und wird sich demzufolge auch nicht vor dem Direktor verantworten müssen, wenigstens nicht jetzt.

2) Wenn in einem Satz des Typs „der  $F$  ist  $G$ “ die Kennzeichnung der  $F$  auf keinen Gegenstand zutrifft, also leer ist, so sind die eintretenden Effekte je nach Verwendung unterschiedlich. In attributiver Lesart ist der Satz dann über nichts. In referentieller Lesart jedoch kann die Referenz dennoch glücken.

Nehmen wir statt der Erwägung in Bedingung 1) an, es war keiner der Schüler, der die Scheibe eingeworfen hat, sondern sie war nicht richtig eingesetzt und ist, nachdem die Sonne eine Weile herein geschienen hat, unter dem Druck der physikalischen Ausdehnung zerborsten. (Der Schüler, der sich meldet, wollte das Fenster gerade zumachen und hält sich daher für schuldig, obwohl niemand Schuld hat.) In diesem Fall ist in Äußerung S11, die wir attributiv interpretieren *müssen*, die Kennzeichnung „der Schüler, der die Scheibe eingeworfen hat“ leer. Diese Aussage bezieht sich auf nichts, die Prädikation geht ins Leere. In S12 jedoch glückt die Referenz, denn wenn man diese Äußerung referentiell liest (sie anders zu verstehen wäre nun wirklich kontraintuitiv), dann bezieht sich die Lehrerin auf den Delinquenten, der vor ihr steht, wofür sie den kennzeichnenden Ausdruck benutzt.

Durch eine referentiell gebrauchte Kennzeichnung kann also auch dann ein Gegenstandsbezug hergestellt werden, wenn diese leer ist; bei attributiver Verwendung ist dies dagegen nicht möglich.

3) Der attributive und referentielle Gebrauch von Kennzeichnungen unterscheidet sich hinsichtlich seiner jeweiligen Präsuppositionen. Zwar präsupponieren beide Verwendungsweisen, daß es etwas gibt, das die Kennzeichnung erfüllt. Doch zusätzlich gehört es zur Präsupposition der referentiell verwendeten Kennzeichnung, daß es ein bestimmter Gegenstand sein muß, der sie erfüllt. In der Urform des Scheibenbeispiels präsupponiert S11 lediglich, daß jemand die Scheibe eingeworfen hat. S12 präsupponiert dies ebenfalls und zusätzlich präsupponiert die Lehrerin dadurch, daß der vor ihr stehende Schüler dies war.

4) Zwar präsupponieren beide Gebrauchsweisen von Kennzeichnungen die Existenz von etwas, das die Kennzeichnung erfüllt – doch geschieht dies aus unterschiedlichen Gründen. Bei attributivem Gebrauch einer Kennzeichnung ist dies deshalb der Fall, weil sonst keine

wahrheitswertfähige Aussage gemacht werden könnte, denn dann tritt Bedingung 2) für attributiven Gebrauch in Kraft. Die Lehrerin präsupponiert in S11, daß es jemanden gibt, auf den zutrifft, daß er die Scheibe eingeworfen hat, weil sie denkt, daß die Scheibe eingeworfen wurde. Ohne diese Präsupposition wäre die Aussage nicht wahrheitswertfähig. Bei referentiellen Gebrauch entsteht die Existenzpräsupposition dadurch, daß ein Gegenstand herausgegriffen und diesem Gegenstand dann etwas zugeschrieben werden soll. Schließlich präsupponiert die Lehrerin (als Kartesianerin), daß es den vor ihr stehenden Schüler *gibt*.

Mit diesen Unterscheidungen führt Donnellan etwas ein, was sich nicht von Gebrauchs-fällen von Sätzen auf Sätze selbst übertragen läßt. Mithin können wir, da kein semantisches Merkmal diese Unterscheidung bestimmt, sie auch nicht in Carnaps Semantik darstellen.

Es ist leicht einzusehen, daß die Auffassung Donnellans grundlegend von den Auffassungen Strawsons und insbesondere Russells abweicht. Russell war der Meinung, daß jeder Satz, der einen kennzeichnenden Ausdruck enthält, eine Existenzbehauptung einschließt. Strawson *und* Donnellan halten dagegen, daß es sich dabei nicht um eine logische Implikation, sondern um eine Präsupposition handelt. Donnellans Distinktion zufolge ist die Erfüllung dieser Präsupposition jedoch nur in attributivem Gebrauch notwendig, in referentiellen Gebrauch kann auch ohne sie eine wahrheitswertfähige Aussage gemacht werden. Das aber heißt, daß Russells ganze Theorie nur auf attributiv verwandte Kennzeichnungen angewendet werden kann, den referentiellen Gebrauch hat er einfach nicht gesehen. (Auch Strawson äußert sich zu einer Unterscheidung zweier Verwendungsweisen von Kennzeichnungen, die ich hier nicht besprechen möchte. Diese fängt jedoch den Unterschied, den Donnellan gesehen hat, noch nicht ausreichend ein.)

Donnellan macht ein semantisch relevantes Merkmal einer Äußerung, nämlich Referenz, von nichtsemantischen Eigenschaften dieser Äußerung abhängig. Schon allein darin zeigt sich seine Unterscheidung gegenüber Carnaps Semantik angriffslustig.

Es ist leicht zu sehen, daß das Vorhandensein einer referentiellen Lesart von Kennzeichnungen den Wunsch eines wie auch immer gearteten Zusammenfallens der Bedeutung von Namen und bestimmten Kennzeichnungen unrealistisch aussehen läßt. Wie kann die Bedeutung eines Namens durch ein Bündel von Kennzeichnungen gegeben sein, wenn ich diese Kennzeichnungen verwenden kann, um mich auf jemand anders als den Träger dieses Namens zu beziehen – noch dazu mit Erfolg? Möglicherweise ließe sich dieses Problem mit einer ausgefeilteren Semantik als der Carnaps wieder richten, doch die Kennzeichnungstheorie muß sich mit der Auffassung anfreunden, daß sie lange Zeit Referenzmöglichkeiten übersehen hat, die semantisch hochgradig relevant sind und die man in die kennzeichnungstheoretische Auffassung integrieren muß. In diesem Sinne ist Donnellan bereits kein Kennzeichnungstheoretiker mehr.

Für unsere Frage nach kognitiver Signifikanz ist zwar Donnellans Unterscheidung selbst nicht so wichtig, doch man kann wohl sehen, daß es mit der Rolle der Kennzeichnungsbündel für die kognitive Signifikanz von Eigennamen nicht allzuweit her sein kann, wenn ich mithilfe dieser Kennzeichnungsbündel in referentieller Lesart eine korrekte Referenz herstellen kann, ohne daß das Referenzobjekt die meisten dieser Kennzeichnungen erfüllen muß. Insofern liefert Donnellan ein starkes Indiz dafür, daß Searle etwas übersehen hat.

Und es gibt den Ansatz einer positiven These, die Donnellan freilich noch nicht expli-

zit macht: Wir haben gesehen, daß Donnellans Unterscheidung sich in Carnaps Semantik nicht darstellen läßt. Der Grund für diese Inkompatibilität: Donnellan berücksichtigt bei der Referenzfestlegung eines Ausdrucks die Intentionen des Sprechers, die Sprache in bestimmter Weise zu verwenden. Für diesen, bei Donnellan noch schwach entwickelten Aspekt des Sprachgebrauchs, finden sich in Carnaps Semantik keine Darstellungsmöglichkeiten, weil die Darstellung von Intentionalität voraussetzen würde, daß diese überhaupt besteht. Carnaps Theorie ist aber keine Theorie der Sprachverwendung. Wo, wie bei Strawson, pragmatische Merkmale mithilfe semantischer Merkmale wie zum Beispiel Partialisierung abgebildet werden konnten, versagte Carnaps Semantik nicht – doch für Donnellans Unterscheidung der Lesarten von Kennzeichnungen gibt es keinerlei syntaktische Merkmale, die man dem Satz selbst entnehmen könnte. Daher müssen wir, wenn wir den Unterschied in den Referenzmodi der beiden Lesarten abbilden wollen, eine semantische Kategorie konstruieren, die nicht nur die Referenz zwischen möglichen Welten differenzieren kann, sondern auch in möglichen Welten. Wir werden in Kapitel III mehr dazu hören.

## I.10 Was ist erreicht?

Die kognitive Signifikanz von Gegenstandsbezeichnungen wird von Frege als die Weise angesehen, in der ein Ausdruck uns seinen Gegenstand „gibt“. Frege nennt dies den Sinn des Ausdrucks. Der Sinn eines Namens und manch anderer Gegenstandsbezeichnungen kann durch einen kennzeichnenden Ausdruck wiedergegeben werden, woraus sich ableiten läßt, daß die kognitive Signifikanz von Gegenstandsbezeichnungen – speziell von Eigennamen – auf der Synonymie von Eigennamen mit Kennzeichnungen beruht.

Wir stoßen allerdings auf das Problem, daß die Aussagen, die diese Kennzeichnungen über den Namensträger machen, nicht immer Teil von Aussagen sind, die mithilfe dieses Namens gemacht werden. In den Kontexten, die wir als intensional bezeichnen, ist dies zum Beispiel nicht so. Für die rekursive Semantik hat Carnap dieses Problem vorläufig gelöst, in dem er die rekursiven Prinzipien der Funktionalität und der Substitution auf entsprechende Kontexte beschränkte.

Laut Carnap kann die kognitive Signifikanz von Gegenstandsbezeichnungen durch eine sogenannte Intension angegeben werden. Die Intension eines Begriffs ist eine Funktion, die jeder Gebrauchssituation dieses Ausdrucks eine entsprechende Extension als ein Bezeichnetes zuordnet. Die Kennzeichnungstheorie der Bedeutung von Namen nimmt daher nun an, daß es für jeden Namen Kennzeichnungen gibt, die mit diesem Namen intensionsgleich sind. Von Searle wurde dann die These aufgebracht, daß es ein ganzes Bündel von Kennzeichnungen braucht, um die Bedeutung eines Namens festzulegen. Dabei stoßen wir auf modale Probleme: Kennzeichnungen geben in aller Regel kontingente Eigenschaften des Namensträgers an und wie soll man einen Namen für etwas anhand von dessen kontingenten Eigenschaften so festlegen, daß die Methode der Intension und Extension noch funktioniert? Das Ergebnis wäre, daß wir zugestehen müßten, daß sich in die kognitive Signifikanz von Eigennamen akzidentelle Elemente als bedeutungsfestlegende Faktoren mischen, was unserem intuitiven Begriff von Bedeutung widerspricht. Searles Antwort darauf ist, daß die Disjunktion aller kennzeichnenden Aussagen auf den Namensträger notwendig zutrifft. Da-

für fehlt aber ein gutes Argument.

Außerdem gibt es ein Argument *gegen* die Synonymie von Kennzeichnungen mit Namen, das aus Donnellans These extrahiert werden kann, nämlich daß Kennzeichnungen auch *nicht* synonym mit Namen verwendet werden können und trotzdem eine Referenz auf einen Namensträger stattfinden kann, den die verwendete Kennzeichnung gar nicht kennzeichnet. Auf diese beiden Schwierigkeiten – die erste, modale und die zweite, pragmatische – findet die Kennzeichnungstheorie keine Antwort. Als Theorie der kognitiven Signifikanz von Eigennamen wird sie daher unzulänglich.

Bezüglich der Verortung der kognitiven Signifikanz von Eigennamen können wir eine internalistische Basistheorie ausmachen, die wesentlich von Frege begründet wurde: die Auffassung nämlich, daß die Kenntnis von Bedeutungen heißt, daß unser Bewußtsein das, was wir als Bedeutung bezeichnen, irgendwie aufgreift und dann mit sich herumträgt, so daß jedwede Form von Überzeugungen, Urteilen und Begriffen intrinsische Eigenschaften des Sprechers sind. Diese Annahme wurde als unproblematisch akzeptiert, so daß man glaubte, die eigentlichen Probleme lägen nur auf der formal-semantischen Seite der Bedeutungstheorie. Wie sich aber seit I.6 andeutet, muß eine rekursive Semantik, die wirklichen Anspruch auf Erklärung erhebt, *subjektivitätskompetent* sein. Diese Forderung ist von der internalistischen Theorielinie bisher nicht berücksichtigt worden, obwohl sich seit Carnap eine Ahnung davon andeutet. Die Kennzeichnungstheorie jedenfalls hat die Forderung zur Subjektivitätskompetenz bisher nicht entsprechend berücksichtigt.

Wir werden im nächsten Abschnitt zunächst sehen, wie die Kennzeichnungstheorie von Kripke abgeschafft wird, ohne bei ihm zunächst auf einen positiven Gegenvorschlag zu stoßen. Vorschläge zur Überwindung beziehungsweise Weiterentwicklung der Theorie der Designatoren finden sich dann in den Kapiteln III und IV.

## II Die Wende: Kripke

### II.1 Die Unzulänglichkeit der Bündeltheorie

Im Januar 1970 hielt Saul Kripke an der Princeton University einen dreiteiligen Vortrag mit dem Titel „Naming and Necessity“. Dieser Vortrag, veröffentlicht als [Kri72], markiert einen entscheidenden und irreversiblen Wendepunkt in der Geschichte der Bedeutungstheorie (und der analytischen Philosophie schlechthin). Zwar widmet Kripke der Philosophie der Namen besondere Aufmerksamkeit, doch seine Erwägungen sind von entscheidender Relevanz für alle Bedeutungsbegriffe.

In „Naming and Necessity“ kann man drei Hauptmotive identifizieren, die erst in Relation zueinander, sozusagen „konzertant“ die eigentliche Wendemarke bestimmen, von der ich gesprochen habe.

Erstens weist Kripke nach, daß bereits der Ansatz von Frege und Russell einem fundamentalen Fehlverständnis der Bedeutung von Namen unterliegt. Die Idee, Namen seien als synonym mit einer Beschreibung oder, wie später bei Searle, einem Bündel von beschreibenden Kennzeichnungen anzusehen, weist Kripke entschieden zurück und setzt dem die Lehre von den Namen als sogenannten starren Designatoren entgegen. Zweitens führt Kripke bezüglich des Notwendigkeitsbegriffs Neuerungen von gewaltiger Tragweite ein: er ist der erste Philosoph, der unbeirrbar strikt zwischen der epistemischen Modalität der Apriorität und der metaphysischen Modalität der Notwendigkeit unterscheidet und diese Unterscheidung auch konsequent in die eigene Philosophie einfügt. Drittens ergänzt Kripke diese Überlegungen noch um die Vorstellung einer kausalen Theorie der Referenz, wonach die Relation des Nennens und nichts sonst die Referenz eines Namens konstituiert. Dieser Name wird mitsamt seiner Referenz wie in einer kausalen Kette [*causal chain*] von Sprecher zu Sprecher weitergegeben.

Außerdem werden wir sehen, daß sich bei Kripke bereits erste Elemente des Theoriegebäudes finden, das später von Putnam und Burge konstruiert werden wird und das unumkehrbar von der Grundannahme des Internalismus wegführt, Überzeugungszustände seien intrinsische Eigenschaften des Individuums.

Kommen wir zum ersten Motiv von Kripkes Theorie: der Zurückweisung der Bündeltheorie der Bedeutung und der Lehre von der starren Designation.

Nach Frege und Russell ist es möglich, den Sinn eines Namens durch eine Kennzeichnung anzugeben, wobei verschiedene Sprecher mit demselben Namen korrekterweise verschiedene Sinne assoziieren können. Wie wir gesehen haben, meint Frege, daß diese „Schwankungen des Sinns“ (vgl. [Fr94] p.42, Fn 2) tolerabel seien, solange nur die Extension dieselbe sei, wobei diese Schwankungen nur als Verunreinigungen der gesprochenen Sprache hinnehmbar seien, nicht jedoch für eine Sprache strenger Wissenschaft.



Searle bietet für dieses Problem einen Ausweg an, indem er einräumt, daß die Bedeutung eines Namens (beziehungsweise in einer schwächeren Lesart der Bündeltheorie: die Referenz) durch mehrere Kennzeichnungen – eben ein ganzes Bündel – fixiert würde<sup>1</sup>. Das Individuum, auf das die meisten der Kennzeichnungen aus dem Bündel zuträfen, sei dann der Referent, egal, was sonst noch auf ihn zuträfe.

Kripke gibt sich mit der Darstellung dieser Bündeltheorie viel Mühe. Letztenendes kann man sie, seiner Explikation folgend, in 6 Thesen und einer Bedingung darstellen:

- (1) Jedem Designator „X“ entspricht ein Bündel von Eigenschaften, nämlich die Familie der Eigenschaften  $\phi$ , für die gilt: A glaubt „ $\phi X$ “.
- (2) A glaubt, daß eine der Eigenschaften oder einige Eigenschaften zusammen einen bestimmten individuellen Gegenstand als einzigen herausgreifen.
- (3) Wenn die meisten oder eine ausschlaggebende Menge von Elementen aus  $\phi_1 \dots \phi_n$  von einem einzigen Gegenstand  $y$  erfüllt werden, dann ist  $y$  der Referent von  $X$ .
- (4) Wenn die Klassifizierung durch das Bündel  $\phi$  nicht einen einzigen Gegenstand liefert, dann referiert „X“ nicht.
- (5) Die Aussage „Wenn  $X$  existiert, dann hat  $X$  die meisten der Elemente aus  $\phi$ “ weiß der Sprecher a priori.
- (6) Die Aussage „Wenn  $X$  existiert, dann hat  $X$  die meisten der Elemente aus  $\phi$ “ drückt im Idiolekt des Sprechers eine notwendige Wahrheit aus.

[B] Die Eigenschaften, die in  $\phi$  vorkommen, dürfen den Begriff der Referenz nicht ihrerseits in einer Weise enthalten, der seine Eliminierung letztlich unmöglich macht. Mit anderen Worten: Für die Bündeltheorie gilt wie für jede wissenschaftliche Theorie, daß sie nicht zirkulär sein darf.

Die Crux mit der Bündeltheorie ist nun, daß sich keine Version formulieren läßt, in denen (1)-(6) wahr sind und [B] erfüllt ist. Vielmehr ist jede der sechs Thesen bereits für sich genommen zweifelhaft. Ich kann die Gegenbeispiele hier nicht in der Breite besprechen, wie Kripke es tut, doch ich will im folgenden die sechs Thesen kurz durchgehen:

These (1) formuliert lediglich die Bündeltheorie in Form einer Zusammenfassung. Sie ist eine Definition und wir brauchen dazu nichts zu sagen.

These (2) bedeutet, daß ich, um einen Gegenstand als einzigen herauszugreifen, seinen Namen mit einem Bündel von Eigenschaften korrelieren muß, was ich in einer Art mentalen Zeremonie tue, indem ich das Bündel festlege. Doch wenn ich einen Gegenstand als einzigen herausgreifen möchte, dann ist dies auch ohne das Bündel zu benutzen möglich. Zum

---

<sup>1</sup>Eigentlich war Searle selbst, den auch Kripke als locus classicus für die Bündeltheorie einräumt, nicht wirklich der Meinung, daß es so einfach sei, daß die Bedeutung eines Namens einfach synonym mit der Disjunktion gewisser Kennzeichnungen sei. Vielmehr muß man, denke ich, Searle so verstehen, daß er glaubte, es gehöre zur Präsupposition des Namens, daß der bezeichnete Gegenstand sich durch eben dieses Bündel beschreiben lasse.

Beispiel kann ich einen Satz über Aristoteles bilden, ohne wirklich eine Ahnung zu haben, wer er ist oder was er getan hat. Wenn ich sage: „Ich kenne Aristoteles nicht“, dann greife ich nichtsdestoweniger Aristoteles heraus, obwohl ich über ihn nur weiß, daß er Aristoteles heißt. Andererseits könnte (2) erfüllt sein und dennoch gegen die Bedingung der Nonzirkularität verstoßen, zum Beispiel, wenn ich Einstein herausgreifen will und sage: „Einstein ist derjenige, der die Relativitätstheorie aufgestellt hat.“ Dann werde ich gefragt, was die Relativitätstheorie denn besage, und da ich als unbedarfter Mensch nie von Überlegungen über den Zusammenhang zwischen den Naturgesetzmäßigkeiten von Inertialsystemen und relativer Bewegung gehört habe, sage ich: „Das ist die Theorie, die von Einstein aufgestellt wurde.“ Das ist offensichtlich nicht im Sinne von (2) verwendbar.

These (3) hat zur Folge, daß die Beschreibung durch das Bündel die Identität des Referenten festlegt. Das entspricht aber weder unseren Mustern zu denken, noch der Logik, der Fälle von Beschreibung und Identifikation unterliegen. Beispielsweise war die naheliegendste Art, auf Peano zu referieren die, ihn (vereinfacht) als den Begründer der Theorie der natürlichen Zahlen zu beschreiben. Also wird man These (3) folgend sagen, der Name „Peano“ referiere auf den Begründer der Theorie der natürlichen Zahlen, wer immer das auch sein mag. Liest man Peano genau, wird man entdecken, daß er selbst anmerkt, stark von Richard Dedekind beeinflusst zu sein, der der eigentliche Begründer der Theorie ist, die die Folge der natürlichen Zahlen charakterisiert. Dann muß ich mit (3) annehmen, daß ich mit „Peano“ auf Dedekind referiert habe und das ist nicht wahr. Ich habe eine falsche Meinung über Peano und nicht eine richtige über Dedekind. These (3) greift nicht (cf. hierzu [Kri72] p. 293-295).

Auch These (4) macht uns unglücklich: Es könnte erstens sein, daß  $\phi$  mehr als *einen* Gegenstand liefert. Wir haben in dem Aristoteles-Beispiel zu These (2) bereits gesehen, daß in Bezug auf bestimmte Namen das charakterisierende Bündel so dünn sein kann, daß der Gegenstand nicht wirklich von anderen unterschieden wird. Wenn ich zum Beispiel weiß, daß Aristoteles ein griechischer Philosoph ist, und ich weiß, daß Platon ein griechischer Philosoph ist, dann kann ich auf keine Folge  $\phi_i - \phi_j$  hinweisen, die es mir ermöglichte, Aristoteles von Platon zu unterscheiden. Dennoch bezeichne ich Aristoteles, wenn ich über ihn rede, durch „Aristoteles“ und Platon durch „Platon“. Zweitens könnte es sein, daß  $\phi$  die leere Menge zurückliefert. Nehmen wir das Beispiel der biblischen Figur Jona. Niemand wurde je von einem Walfisch verschluckt und lebendig wieder ausgespien, und sehr wahrscheinlich ist Jona auch nicht nach Ninive gegangen, um dort zu predigen. Also muß  $\phi$  für „Jona“ die leere Menge liefern. Dadurch, daß Jona an anderen Stellen in der Bibel erwähnt wird, weiß man aber, daß er eine historische Figur ist, jemand, der tatsächlich gelebt hat. Und wenn ich Sätze bilde, in denen ich „Jona“ benutze, um Jona zu bezeichnen, dann ist damit eine völlig korrekte Referenz konstruiert. Also funktioniert (4) nicht.

These (5) kann selbst in Fällen, in denen die Thesen (3) und (4) aus empirischem Zufall wahr sind, falsch sein. Der Sprecher wird das Ausgedrückte kaum im Sinne der Theorie a priori wissen. Wenn ich lese, daß Peano der Begründer der Theorie der natürlichen Zahlen war, dann ist das kein apriorisches Wissen und kann sich im nachhinein als falsch herausstellen: vielleicht hat Peano ja auch alles gefälscht, obwohl man davon wohl nicht ausgehen kann. Ein schönes Beispiel hierfür ist das folgende, das Kripke nicht anführt, das ich aber für passend halte: gewisse Literaturwissenschaftler glauben, Shakespeare habe nie existiert.

Stattdessen habe eine Geheimorganisation von Schriftstellern aus politischen Gründen vorgegeben, die Werke, die sie verfaßt haben, seien von der fiktiven Person Shakespeare. Ihre Geschichtsfälschung war dabei so perfekt, daß niemals jemand etwas merkte. Wenn diese Theorie sich als wahr herausstellt, dann kann mein vermeintliches Wissen über Shakespeare, daß er der Verfasser all der Dramen und Sonette oder wenigstens der meisten dieser Werke sei, wohl nicht apriorisch gewesen sein. Und woher soll ich wissen, welche vermeintlich apriorischen Urteile über Personen oder Orte der Geschichte sonst noch Humbug sind? Praktisch alles kann sich als falsch herausstellen.

These (6) ist die von Kripke am breitesten diskutierte These. Bereits das Shakespeare-Beispiel aus (5) zeigt, daß hier Probleme bestehen. Doch ich will im Rahmen der Widerlegung von These (6) überleiten zu dem Gegenentwurf, den Kripke anbietet.

These (6) sagt, daß (im Idiolekt des Sprechers) die Aussage, daß  $X$ , sofern es existiert, die meisten der Elemente aus  $\phi$  habe, eine notwendige Wahrheit ausdrücke. (Wir erinnern uns: Searle räumt ein, daß jede einzelne der Kennzeichnung in dem identifizierenden Bündel eine kontingente Eigenschaft beschreibe. Das Zutreffen der Disjunktion all dieser Eigenschaften auf den Träger des Namens jedoch sah er als notwendig an.) Bevor wir nun etwas über den Wert dieser These in Erfahrung bringen können, müssen wir, um Kripke verstehen zu können, in einem kurzen Exkurs begreifen, was Notwendigkeit bedeutet. Dies ist die Überleitung zu dem zweiten Motiv seiner Untersuchung, der Trennung von Apriorität und Notwendigkeit.

## II.2 Apriorität und Notwendigkeit

Die Philosophie, so Kripke, wirft ungeachtet der definatorischen Unterscheidung durch Kant die drei Kategorien Apriorität, Analytizität und Notwendigkeit als koextensiv in einen Topf. Alles, was a priori erkannt werden könne, sei analytisch wahr und daher notwendig. Dies ist jedoch nicht wahr. Insbesondere die beiden Modalitäten Apriorität und Notwendigkeit sind, so Kripke, scharf zu trennen und haben miteinander nichts zu tun.

Apriorität, so sagt Kripke unter Rückgriff auf Kant, ist eine erkenntnistheoretische, das heißt epistemische Modalität. Ein Satz ist dann a priori als wahr oder falsch erkennbar, wenn wir seine Wahrheit oder Falschheit unabhängig von jeder Erfahrung hinsichtlich der Beschaffenheit der Welt erkennen können.

Das heißt aber, daß wir von einem a priori erkennbaren Satz lediglich sagen können, daß man ihn a priori erkennen *kann* – wohingegen die traditionelle Philosophie Apriorität sehr oft so interpretierte, daß man einen a priori erkennbaren Satz a priori erkennen *muß*. Das ist aber offensichtlich falsch, denn selbstverständlich kann ich etwas auch empirisch erkennen und stelle hinterher fest, daß ich kein Erfahrungswissen hätte anwenden müssen, um zu dieser Erkenntnis zu kommen, sondern daß ich, was ich jetzt weiß, von vornherein hätte wissen können. So kann zum Beispiel mein Computer für mich ausrechnen, ob die Zahl  $g$  eine Primzahl ist. Wenn ich aufgrund der Meldung des Computers glaube, daß  $g$  eine Primzahl sei, dann glaube ich das kraft meines Wissens über die Bauweise des Computers und aufgrund der empirischen Entdeckungen der Naturgesetze, auf der die Funktionsweise dieses Computers basiert. Ich glaube, daß der Computer deshalb richtig liegt. Dennoch hätte

ich  $g$ , sofern es keine zu große Zahl ist, auch ohne Rekurs auf Erfahrungswissen dahin prüfen können, ob es eine Primzahl ist. „Kann a priori gewußt werden“ ist also nicht koextensiv mit „muß a priori gewußt werden“.

Doch anders als für Kant ist Apriorität für Kripke ein Begriff, der von der Beschaffenheit des epistemischen Agenten abhängig ist. Ein Satz kann nicht für sich genommen a priori erkennbar sein, sondern nur für ein erkennendes Subjekt. Apriorität ist also ein Begriff, der letztendlich eher eine Aussage über das erkennende Subjekt als über den Gegenstand der Erkenntnis beinhaltet.

Um zu beweisen, daß Apriorität nicht mit Notwendigkeit zusammenfällt, braucht es ein Beispiel einer kontingenten Wahrheit a priori. Das in dieser Hinsicht prominenteste Beispiel ist Kripkes Beispiel vom Urmeter.

Die Referenz des Ausdrucks „ein Meter“ ist durch die Länge des Urmeters festgelegt, eines Stabes, der sich in Paris befindet. Nennen wir diesen Stab  $S$ . Die Referenz wird festgelegt, in dem man festsetzt:

S13) Der Ausdruck „ein Meter“ bezeichnet die Länge des Stabes  $S$  zum Zeitpunkt  $t_0$ .

S13 ist nur kontingenterweise wahr, denn die Länge von  $S$  zum Zeitpunkt  $t_0$  ist eine völlig akzidentelle Größe und eine lediglich kontingente Eigenschaft des Stabes  $S$ . Wir hätten auch jede andere Länge auf diese Weise herausgreifen können und das hätten wir auch getan, wenn  $S$  sie zum Zeitpunkt  $t_0$  gehabt hätte. Es hätte also sein können, daß  $S$  zum Zeitpunkt  $t_0$  eine andere Länge gehabt hätte, als die, die wir heute als „ein Meter“ bezeichnen und in diesem Sinne ist es kontingenterweise wahr, daß  $S$  zu  $t_0$  genau einen Meter lang ist. Doch *daß* die Länge, die  $S$  zum Zeitpunkt  $t_0$  hat, einem Meter entspricht, *das* wissen wir a priori, weil wir es nämlich so festgesetzt haben. Insofern muß etwas a priori Erkennbares nicht notwendig wahr sein.

Notwendigkeit ist im Gegensatz zu Apriorität eine metaphysische Modalität. Kripke gebraucht den Notwendigkeitsbegriff in einem recht landläufigen Sinn, wonach all das notwendig wahr ist, was nicht hätte anders sein können, das heißt, wofür wir keine mögliche Welt festlegen können, in der es nicht wahr ist (vgl. hierzu [Kri72] p. 261.)

Notwendigkeit hängt nicht davon ab, wann oder wodurch etwas erkannt werden kann, weil sie von der Natur des epistemischen Agenten unabhängig ist. Ob etwas notwendig wahr ist oder nicht, hängt nicht vom erkennenden Subjekt ab und es hat nichts damit zu tun, ob etwas a priori erkennbar ist oder nicht. Es gibt daher Beispiele von Sätzen, von denen wir wissen, daß ihnen ihr Wahrheitswert notwendig zukommt, ohne diesen Wahrheitswert zu kennen. Diese Sätze sind notwendig wahr und nur a posteriori erkennbar.

Kripkes Beispiel hierfür ist die Goldbachsche Vermutung: sie besagt, daß jede gerade Zahl, die größer ist als 2, die Summe zweier Primzahlen sei. Bis heute ist für diese These kein Beweis gefunden, aber wir wissen, daß der Wahrheitswert dieser These ihr mit Notwendigkeit zukommen muß. Der Beweis muß eine arithmetische Berechnung sein und man kann davon ausgehen, daß arithmetische Berechnungen ihre Ergebnisse notwendig haben. Es ist zwar denkbar, daß der Satz „ $2 + 2 = 4$ “ eine andere Proposition repräsentiert als die, die wir meinen. Es ist hingegen nicht denkbar, daß eine beliebige Repräsentation der Proposition, die wir durch „ $2 + 2 = 4$ “ ausdrücken, je falsch wird. Also wird man zugestehen, daß die

Goldbachsche Vermutung nur notwendig wahr oder notwendig falsch sein kann. Doch wir kennen ihren Wahrheitswert nicht. Vielleicht *könnte* man ihn kennen, doch solange man einen begrenzten Geist als epistemischen Agenten annimmt, ist nicht klar, wie diese Bewußtseinsinstanz a priori erkennen soll, ob Goldbach recht hat oder nicht, weil sie nicht alle Zahlen durchgehen kann. Wir wissen nicht einmal, ob sich der Wahrheitswert der Vermutung überhaupt durch einen Beweis fixieren läßt. Was wir hingegen wissen, ist, daß es einen Wahrheitswert gibt und daß er der These notwendig zukommt.

Es gibt außerdem Beispiele, die uns sehr vertraut sind: nämlich alle informativen Identitätsaussagen zwischen Eigennamen, wie schon Frege sie beschrieb. Der Satz „Der Abendstern ist mit dem Morgenstern identisch“ ist, wie wir wissen, eine nichtapriorische Tatsache. Dennoch drückt er eine notwendige Wahrheit aus. Doch mehr zu diesem Beispiel später.

Nachdem Kripke die beiden Modalitäten Notwendigkeit und Apriorität auf diese Art eingeführt hat, definiert er Analytizität als das Zusammenfallen von Notwendigkeit und Apriorität.

Diese Differenzierungen ermöglichen nun ein ganz anderes Sprechen über kontrafaktische Situationen. Wir können, indem wir kontrafaktische Situationen – bleiben wir bei dem Terminus „mögliche Welten“ - festlegen, zwischen kontingenten und notwendigen Sachverhalten unterscheiden. Dies ist für die Rede über Bedeutung und Referenz, auf die wir zusteuern, unerlässlich.

Was in diesem bisher nicht ganz zwei Seiten langen Abschnitt zusammengefaßt ist, ist für die Philosophie – speziell für die analytische Philosophie – nichts weniger als ein Erdbeben, eine grundsätzliche Kurskorrektur. Wo bisher jede Analyse, die auf die Begriffe von Notwendigkeit, Apriorität und Analytizität zurückgriff, sich durch deren ungenügende Trennschärfe in unentrinnbare Schwierigkeiten brachte, dort öffnet die Trennung und Präzisierung der besprochenen Modalitäten den Weg zu einem philosophischen Sprechen in neuer Klarheit.

Um nun eine Eigenschaft *P* dahingehend zu untersuchen, ob sie ihrem Träger notwendig zukommt, müssen wir uns andere mögliche Welten ansehen. Wir werden uns erinnern, daß bereits Searle einräumt, daß Aristoteles auch nicht der Lehrer Alexanders hätte sein können. Wie aber sollen wir Aristoteles in einer anderen möglichen Welt erkennen, wenn wir nicht wissen, welche unserer Kennzeichnungen er dort erfüllt? Wie sollen wir den Träger eines Namens in anderen möglichen Welten identifizieren?

Kripke antwortet, daß diese Frage falsch gestellt ist: man kann über mögliche Welten nicht so reden, als würde man sie erst entdecken müssen. Kontrafaktische Situationen werden *festgelegt* und nicht *entdeckt*. Ich weiß bereits, wer in *unserer* Welt Aristoteles ist. Dieses Wissen genügt völlig, um festzulegen, wer in einer kontrafaktischen Situation Aristoteles *hätte sein können*.

Dies ist eine wichtige methodologische Voraussetzung über das Reden von möglichen Welten, die Kripke ausführlich und deutlichst heraushebt, die aber oft vergessen oder übergangen wird: mögliche Welten sind uns nicht *qualitativ* gegeben.

„‘Possible worlds’ are *stipulated*, not *discovered* by powerful telescopes.“ [[Kri72] p.267]

Der Gedanke, mögliche Welten könnten erforscht werden, ist sinnlos, weil wir nur das

entdecken werden, was wir zuvor selbst festgesetzt haben. Was wir nicht festgelegt haben, werden wir nicht herausbekommen, und was wir festgelegt haben, wissen wir, ohne es herausbekommen zu müssen. Dennoch ist eine Analyse kontrafaktischer Situationen in der Regel erhellend, weil sie uns viel über den Status gewisser Eigenschaften und Gegenstände verraten kann. Ich brauche Aristoteles nicht in einer anderen möglichen Welt zu suchen, denn entweder, ich habe festgelegt, wer es ist, dann weiß ich es; oder ich habe über Aristoteles nichts festgelegt, dann gibt es auch nichts, was ich über ihn herausbekommen könnte.

Man stelle sich einen Bäcker vor, der einen Kuchen nach eigenem Rezept backt. Zwar weiß er nicht vorher, wie der Kuchen schmecken wird. Das erfährt er nur, indem er ihn rezeptgetreu backt und ein Stück probiert. Auch bei der modalen Analyse wissen wir nicht vorher, was sich als kontingent oder notwendig herausstellen wird. Aber auch wenn der Bäcker von vornherein nicht alles über den Kuchen weiß, den er backt, so wäre es Unsinn, einen rezeptgetreu gebackenen Kuchen *im nachhinein* wissenschaftlich auf die in ihm vorkommenden Substanzen hin zu analysieren, denn diese stammen aus dem Rezept, das schon *vorher* bekannt gewesen sein muß, denn sonst gäbe es keinen Kuchen. Insofern wurde der Kuchen vom Bäcker vorher „festgelegt“: qualitativ interessant ist, welche *Eigenschaften* er aufgrund dieser Festlegung hat oder entwickelt, aber nicht, woraus er *besteht*. Er besteht aus dem, was der Bäcker hineingetan hat. Interessant ist für den Bäcker, welche Speise das Rezept ergibt; das Rezept selbst aber braucht er, nachdem er es festgelegt hat, nicht mehr daraufhin zu analysieren, was es enthält.

Diese Überlegungen sollten wir parat haben, wenn es nachher um die modale Bedeutungsanalyse bei Putnam und Burge geht.

Nach diesen zentralen methodologischen Überlegungen kehren wir nun zum ersten Motiv Kripkes zurück und ich komme auf seine eigentliche „Theorie“ zu sprechen, wobei wir, jetzt gut ausgerüstet, auch sehen werden, wie es sich mit These (6) verhält.

### II.3 Das Prinzip der starren Designation

These (6) der Bündeltheorie drückt aus, daß jeder Gegenstand, der durch ein Bündel von Kennzeichnungen  $\phi$  herausgegriffen wird, die logische Gesamtheit der in  $\phi$  beschriebenen Eigenschaften notwendig erfüllt.

Das genau *ist* der Fehler von These (6). Wir *wissen*, so Kripkes Argument, daß in keinem intuitiven Sinne von Notwendigkeit es notwendig ist, daß Aristoteles, um Aristoteles zu sein, irgendetwas von dem getan haben muß, was er faktisch getan hat. Kontrafaktisch nämlich kann man annehmen, daß er keine einzige Schrift verfaßt und nie Alexander unterrichtet hat, zum Beispiel weil er in jungen Jahren einer Kinderkrankheit erlegen ist. Also darf man die Modalität der Notwendigkeit nicht auf Aussagen ausdehnen wie „Aristoteles war der Lehrer Alexanders“, auch nicht auf deren Gesamtheit, weil faktisch nicht die meisten der Eigenschaften, die wir mit Aristoteles assoziieren, und auch nicht deren Gesamtheit zutreffen muß, damit Aristoteles Aristoteles ist. Aristoteles würde auch in einer kontrafaktischen Situation Aristoteles gewesen sein, in der Stageira Jahrzehnte vor seiner Geburt ausgelöscht worden ist und er selbst, nachdem seine Eltern ihm einen völlig anderen Namen gegeben haben, im Alter von vier Jahren erkrankte – das alles könnten wir festlegen.

Kripke setzt der Bündeltheorie aufgrund dieser Argumentation das entgegen, was in die Geschichte der analytischen Philosophie als das Prinzip der starren Designation [*rigid designation*] eingegangen ist<sup>2</sup>. Eine Gegenstandsbezeichnung (das heißt ein Designator oder Bezeichner wie ich diese Ausdrücke auch genannt werden) ist *starr* genau dann, wenn er in jeder Welt, in der er überhaupt etwas bezeichnet, denselben Gegenstand bezeichnet. Eigennamen sind demnach offensichtlich *alle* starr (vgl. [Kri72] p. 270).

Wenn wir also die Person, die wir Aristoteles genannt haben, starr mit dem Namen „Aristoteles“ bezeichnen, dann folgt daraus noch nicht, daß Aristoteles irgendwelche bestimmten Kennzeichnungen erfüllen müßte. Vielmehr bezeichnet der Name „Aristoteles“ fortan in jeder möglichen Welt, die wir festlegen, in der es Aristoteles gegeben hat, auch *den* Aristoteles, den wir in unserer Welt faktisch meinen. Mit Kennzeichnungen hat das nichts zu tun. These (6) muß falsch sein: Wir können eine mögliche Welt festlegen, in der Aristoteles nicht der Lehrer Alexanders ist, aber wir können keine mögliche Welt festlegen, in der Aristoteles nicht Aristoteles ist. Die Selbstidentität eines Gegenstandes kann nicht aufgegeben werden, weil sie auf jeden Fall eine notwendige Eigenschaft eines jeden Gegenstandes ist. Am deutlichsten ist dieses Argument von Kripke für Freges Beispiel des Abendsterns ausformuliert:

„When the mythical agent first saw Hesperus, he may well have fixed his reference by saying, ‚I shall use ‚Hesperus‘ as a name of the heavenly body appearing in yonder position in the sky.‘ He then has fixed the reference of ‚Hesperus‘ by its apparent celestial position. Does it follow that it is part of the meaning of the name that Hesperus has such and such position at the time in question? Surely not: if Hesperus had been hit earlier by a comet, it might have been visible at a different position at that time. In such a counterfactual situation we would say that Hesperus would not have occupied that position, but not that Hesperus would not have been Hesperus. The reason is that ‚Hesperus‘ rigidly designates a certain heavenly body and ‚the body in yonder position‘ does not – a different body, or no body might have been in that position, but no other body might have been Hesperus (though another body, not Hesperus, might have been called ‚Hesperus‘. Indeed, as I have said, I will hold that names are always rigid designators.“ [[Kri72] p. 276]

Kann mit den nun zur Verfügung stehenden Mitteln die Frage beantwortet werden, wie es möglich ist, daß ein Satz wie „Der Morgenstern ist identisch mit dem Abendstern“ für uns informativ sein kann? Nur zum Teil, und nicht so, wie Frege es wollte und wie es für eine rekursive Semantik wichtig wäre.

Zunächst fällt uns bei der Untersuchung auf, daß eine Aussage wie „Hesperus = Phosphorus“ nicht a priori wahr ist. Daraus folgt, wie wir jetzt wissen, noch nicht, daß der Satz nur kontingent wahr sei. Wenn wir wissen wollen, ob der Morgenstern notwendig mit dem

---

<sup>2</sup>Ein Designator ist – im Gegensatz zu einem Prädikat – ein Ausdruck, der ein Ding bezeichnet, also das, was bei uns die ganze Zeit schon Gegenstandsbezeichnung heißt. Als Designatoren gelten Eigennamen, nichtabzählbare Namen (zum Beispiel natural kind terms, die ich im Gegensatz zum normalen Sprachgebrauch „natürliche Prädikate“ nenne), Kennzeichnungen und Pronomen im Singular. Allerdings sind natürlich nicht alle diese Designatoren starr, Kennzeichnungen zum Beispiel offensichtlich nicht. Zum Begriff des Designators vgl. z.B. [Ho91].

Abendstern identisch ist, dann müssen wir fragen, ob es möglich gewesen wäre, daß der Morgenstern nicht der Abendstern ist.

Sicher wäre es möglich, daß der mythische Akteur zwei verschiedene Sterne „Morgenstern“ (Phosphorus) und „Abendstern“ (Hesperus) genannt hat. Dann hätten wir eine Situation, in der wir morgens einen anderen Himmelskörper sehen als abends. Allerdings ist dies keine Situation, in der der Morgenstern nicht der Abendstern *gewesen* wäre, denn:

„We use ‘Hesperus’ as the name of those bodies in all possible worlds. If, in fact, they are the same body, then in any other possible world we have to use it as a name of that object. And so in any other possible world it will be true that Hesperus is Phosphorus.“  
[[Kri72] p. 308]

Wenn „Abendstern“ (Hesperus) und „Morgenstern“ (Phosphorus) starre Bezeichner sind, dann bezeichnen sie in jeder Welt denselben Gegenstand, unabhängig davon, welche Kennzeichnungen auf diesen Gegenstand jeweils zutreffen. Und wenn das die Venus ist, dann werden wir finden, daß der Morgenstern notwendig der Abendstern ist, weil die Venus notwendig die Venus ist. Daß wir, ohne es zu wissen, dieselbe Identität doppelt bezeichnen, kann sich aber als informativ herausstellen.

Um diesen Fall zu verdeutlichen stelle man sich die Begriffe „Morgenstern“ und „Abendstern“ durch entsprechende Kennzeichnungen ersetzt vor. Der Satz „Der Stern, der abends an der Position  $(x_1, x_2, x_3)$  zu sehen ist, ist identisch mit dem Stern, der morgens an der Position  $(y_1, y_2, y_3)$  zu sehen ist“ drückt eine lediglich kontingente Wahrheit aus.

Kripkes Theorie steht, wie ich finde, unseren Intuitionen näher, als es zunächst den Anschein hat: Der Name eines Gegenstandes drückt die Selbstidentität dieses Gegenstandes aus, die in jedem Fall eine notwendige Eigenschaft ist. Namen sind daher starre Bezeichner, sie bezeichnen in jeder Welt dieselbe Entität. Freges Antwortstrategie, wonach sich die Namen, die in einen Identitätssatz eingehen, in ihrer kognitiven Signifikanz unterscheiden, ist damit vom Tisch. Wir stehen ohne Antwort auf unser Problem da, obwohl uns jetzt die Beziehung von Kontingenz und Notwendigkeit in Freges Beispiel klarer ist.

Kripke hat trotzdem noch nicht erklärt, wie die Dinge zu ihren Namen kommen – mit dieser Frage werde ich das dritte Hauptmotiv seiner Vorlesung aufnehmen.

## II.4 Name und Notwendigkeit

Man kann die Kennzeichnungstheorien von Frege und Russell ebenso wie die Bündeltheorie in zwei Weisen auffassen, zum einen als eine Theorie der Bedeutung, wie sie von Frege und Russell wohl intendiert war: in dieser Version der Theorie würde die Bedeutung eines Ausdrucks tatsächlich durch die Kennzeichnung beziehungsweise das Bündel von Kennzeichnungen festgelegt. Daß dies nicht funktionieren kann, haben wir anhand von Kripkes bisheriger Argumentation gesehen. Doch man kann die Bündeltheorie ebenso wie die Kennzeichnungstheorien von Frege und Russell in schwächerer Lesart auch als Theorien der Referenz auffassen. Dies würde bedeuten, daß das Kennzeichnungsbündel, das wir mit einem Begriff verbinden, zwar nicht die Bedeutung eines Namens ausmacht, aber dessen Referenz festlegt. Das hätte den Vorteil, daß bei einer Unsicherheit über die Referenz des



Namens auf die Kennzeichnung zurückgegriffen werden könnte. Doch die meisten Kennzeichnungen im Sinne der Kennzeichnungstheoretiker sind, wie wir sahen, so beschaffen, daß sie kontingente Tatsachen bezeichnen. Daher setzt sich auch eine Referenztheorie der Kennzeichnung dem Problem aus, sagen zu müssen, daß der Begründer der Theorie der natürlichen Zahlen Peano ist, egal, wer sie entdeckt hat. Also referiert „Peano“ auf Dedekind. (Die Bedeutungstheorie der Kennzeichnung hätte womöglich dazu verleitet, zu behaupten, Dedekind sei Peano, obwohl sie natürlich so nicht gemeint war.)

Die Bündeltheorie gibt in jeder Lesart ein bestimmtes Bild davon, wie Referenz zustande kommt: ich will einen bestimmten Gegenstand benennen. Also ersinne ich eine Weise, in der ich ihn so beschreiben kann, daß er als einziger herausgegriffen wird und lege fest, daß ich mit dem Namen „Einstein“ den Menschen bezeichne, der die Relativitätstheorie aufgestellt hat und für die Entwicklung der Atombombe eingetreten ist. Immer jedoch gibt es vor der eigentlichen Benennung eine Bedingung, eine Eigenschaft, an die diese Benennung geknüpft ist. Die gesamte Sprachgemeinschaft, der ich angehöre, ebenso wie die Beschaffenheit der Welt wird davon ausgeschlossen. Die Taufe eines Gegenstandes auf einen bestimmten Namen geschieht diesem Bild zufolge in einer Art mentaler Zeremonie in meinem Kopf. Danach ist dann der Name durch das Kennzeichnungsbündel mit dem Gegenstand verbunden.

Kripkes Argument gegen dieses Bild und besonders gegen die These (6) führt in eine Konzeption, die die Philosophie der Bedeutung, besonders die der Namen, die wir bisher kennengelernt haben, grundlegend verändert hat, weil die Idee, daß ein Name als Wort so etwas wie einen eigenen Sinn, eine deskriptive Bedeutung habe, endgültig fallengelassen wird<sup>3</sup>. Allerdings ist die These, die Kripke dem Bild entgegensetzt, das die Bündeltheorie gibt, nur sehr skizzenhaft ausgeführt.

Kripke sagt:

„Obviously if the only descriptive senses of names we can think of are of the form ‘the man called such and such’, ‘the man called ‘Walter Scott’, ‘the man called ‘Socrates’’, then whatever this relation of calling is is really what determines the reference and not any description like ‘the man called ‘Socrates’’. [[Kri72] p. 284]

Es ist also die Relation des *Nennens*, die die Referenz des Namens bestimmt, nicht eine oder mehrere Kennzeichnungen. Nach der „Taufe“ eines Gegenstandes auf einen bestimmten Namen ist dieser Name ein Name dieses Gegenstandes; das ist es, was die Verwendung des Namens „Einstein“ zu einem Namen von Einstein macht. Die Referenz des Namens wird dadurch erhalten, daß die Sprachgemeinschaft den aufwachsenden Sprechern diesen Namen beibringt. Der Name wird von Sprecher zu Sprecher weitergegeben, wobei diese Weitergabe nicht nur ahistorisch, sondern auch über Generationen hinweg erfolgen kann. Das Bild dieser Kausalkette zeichnet Kripke so:

„Someone, let’s say, a baby, is born; his parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by a chain. A speaker who is on the far

<sup>3</sup>Mit diesem Gedanken geht Kripke zu Mill zurück, dessen Auffassung, wonach Eigennamen keine deskriptive Bedeutung haben, wir in I.2 kurz besprochen haben.

end of this chain, who has heard about, say Richard Feynman, in the market place or elsewhere, may be referring to Richard Feynman even though he can't remember from whom he first heard of Feynman or from whom he ever heard of Feynman. He knows that Feynman was a famous physicist. A certain passage of communication reaching ultimately to the man himself does reach the speaker. He then is referring to Feynman even though he can't identify him uniquely." [[Kri72] p. 299]

Das hat nun allerdings einen Nachteil: wenn ich mir über die Referenz eines Namens unsicher bin, müßte ich, um Klarheit zu erlangen, theoretisch die ganze kausale Referenzkette bis zur Taufe zurückverfolgen. In der Praxis ist es wohl eher so, daß Enzyklopädien die Aufgabe übernehmen, die Referenz eines Namens aufzuzeigen, doch auch auf die Handhabung und das Entstehen von Wörterbüchern könnte man die kausale Theorie der Referenz, wie dieses Bild genannt wird, ausdehnen. Natürlich muß die Absicht gegeben sein, die Referenz des Namens auch korrekt weiterzugeben.

Damit also ein Name  $N$  einen Gegenstand  $g$  bezeichnet, müssen zwei Bedingungen erfüllt sein:

- a) Es muß eine Kausalkette existieren, die, von einer ursprünglichen Taufsituation herführend, in der  $a$  den Namen  $N$  erhalten hat, bis zum jetzigen Gebrauch von  $N$  reicht. Diese Kette besteht aus Verwendungen des Namens durch verschiedene Sprecher derart, daß der Name und seine Verwendung von Sprecher zu Sprecher weitergegeben werden.
- b) Jeder Sprecher muß bei einer Verwendung von  $N$  die *Intention* haben, den Namen  $N$  so zu verwenden, wie der Sprecher, von dem er die Verwendung von  $N$  gelernt hat. Das heißt, wenn er weiß,  $N$  bezeichnet genau  $a$ , wird er gegenüber anderen Sprechern  $N$  nie anders gebrauchen als derart, daß  $N$   $a$  bezeichnet.

Wenn a) und b) gegeben sind, liegt darüberhinaus eindeutig eine starre Designation vor.  $N$  wird in allen möglichen Welten, in denen  $a$  existiert,  $a$  bezeichnen, gleich welche Eigenschaften  $a$  dort hat, eben weil  $N$  *per definitionem*  $a$  bezeichnet und  $a$  in jeder möglichen Situation notwendig  $a$  ist. Dies ist die Verbindung, die zwischen Name und Notwendigkeit besteht.

Mehr ist an Substanz aus den Ausführungen Kripkes zur kausalen Theorie der Referenz nicht zu gewinnen. Zwar räumt er ein, daß es Taufszenarien wie das der Bündeltheorie gebe, doch es sei kein verallgemeinerbares Modell. Die Kausaltheorie, wie ich sie ab sofort verkürzend nennen werde, hat jedoch den Anspruch, auf alle Namen anwendbar zu sein. Das kann man sich leicht vorstellen, denn in a) und b) ist nichts darüber ausgesagt, *auf welche Weise*  $N$  mit  $a$  verbunden wird. Die Sprecher können alle möglichen Arten von Kriterien oder Motivationen haben,  $N$  für  $a$  zu wählen und es kann durchaus sein, daß dieser Wahl bestimmte Kennzeichnungen zugrundeliegen, die  $a$  erfüllt – nur kann man davon nicht ausgehen.

Die Kausaltheorie ist jedoch eine Konstruktion, die wesentlich genauer ausgearbeitet werden müßte<sup>4</sup>, um zu echter theoretischer Schlagkraft zu kommen. Kripke selbst sagt, es sei eigentlich keine Theorie oder solle wenigstens keine sein, sondern lediglich ein vorstellbares Bild dessen liefern, was geschieht. Doch dieses Bild ist noch lückenhaft.

<sup>4</sup>Vergleiche hierzu beispielsweise [Ev73] und besonders [Dev81].

So hat Kripke die Referenz zwar kausal erklärt, jedoch nicht auf semantischer Ebene, was die Bündeltheorie versucht hat. Insofern hat Kripke auch die Frage des „eigentlichen“ Gegenstandsbezugs von Namen seit Carnap nicht weiterentwickelt. In den Begriffen der Kausaltheorie, wie sie hier entwickelt wurde, läßt sich nicht sagen, *wie* ein Name sich auf ein Ding bezieht, dafür benötigen wir eine semantische Theorie der Sprachbeschreibung, wie Carnap sie geliefert hat. Mit Kripke ist nur etwas darüber gesagt, wie man einen Bezug herstellen kann und wie Namen gelernt werden, nicht aber, wie dieser Bezug sich abbilden läßt oder wie er beschaffen ist.

## II.5 Eine Bestandsaufnahme

Erstens: Die Fregesche Lösungsstrategie zum Problem der Informativität von Identitätssätzen lautete im Prinzip, daß ein Satz, der die Identität zweier Dinge behauptete, deshalb eine erhellende Information enthalten kann, weil die Intensionen von Eigennamen so beschaffen sind, daß Sätze, in denen sie vorkommen, nicht in allen kontrafaktischen Situationen wahr sind – die Intensionen der beiden Ausdrücke, zwischen denen Identität behauptet wird, unterscheiden sich schließlich. Informativität aber ist das Gegenteil von Apriorität: auch Informativität – oder, wie man aus Sicht des Subjekts sagen könnte, Erkenntnis a posteriori – ist eine epistemische Kategorie. Kripke aber sagt, daß wir Namen damit mißverstehen: sie sind starre Designatoren. Deshalb kommen den aus ihnen gebildeten Identitätssätzen ihre Wahrheitswerte notwendig zu. Man könnte, Kripke dehnend, auch sagen: die Intensionen von Namen sind starr. Das heißt aber, daß man durch sie keine Informativität erklären kann, denn man kann nicht aufgrund der Informativität eines Satzes auf die Kontingenz seines Wahrheitswertes schließen. Damit ist die gesamte Denkrichtung, in die die Tradition sich von Frege über Russell, Carnap, Strawson und Searle in diesem Punkt bewegt hat, versperrt.

Vieles, was erst Kripke in aller Deutlichkeit herausgearbeitet hat, und vieles, was sich bei ihm erstmalig andeutet, wird uns wiederbegegnen: seine Arbeit legt den Grundstein für eine Entwicklung, die erst in den zwei Texten, die ich in Kapitel IV bespreche, zu ihrer vollen Blüte kommt. So wird uns die Lehre von der starren Designation wieder begegnen, von Putnam in unerwarteter Richtung erweitert. Wie Kripke, so wird auch Putnam von der These ausgehen, daß starre Bezeichnung in jeder möglichen Welt festlegt, „was was ist“ – und daher auch, das werden wir sehen, was was *nicht* ist. Nur wird er daraus noch weitreichendere Schlußfolgerungen ziehen.

Doch das weitaus Interessanteste ist eine Auffassung, die bei Kripke zwar nur in spärlichem Raum figuriert, doch bei Putnam und Burge zentrale Bedeutung gewinnt: die Rolle der Sprachgemeinschaft für die Referenz und die Bedeutung eines sprachlichen Ausdrucks. Im Rahmen des Themas dieser Arbeit markiert Kripke in dieser Hinsicht den eigentlichen Wendepunkt, nicht nur durch die Zertrümmerung der Fregeschen Strategie und der folgenreichen Erwägungen zur Trennung von Apriorität und Notwendigkeit. Kripke hat als erster gesehen, wie zentral es für die Begriffe von Bedeutung und Referenz ist, daß Namen *gelernt* werden, daß sie auch von *anderen* Sprechern verwendet und von einer Sprachgemeinschaft geformt werden. Kripke kritisiert in diesem Zusammenhang die Auffassung der Bündeltheo-

rie:

„The picture which leads to the cluster-of descriptions theory is something like this: One is isolated in a room; the entire community of other speakers, everything else, could disappear; and one determines the reference for himself by saying – ‘By ‘Gödel’ I shall mean the man, whoever he is, who proved the incompleteness of arithmetic’. Now you can do this if you want to. [. . .] If that’s what you do, then if Schmidt discovered the incompleteness of arithmetic you do refer to him when you say ‘Gödel did such and such’.“ [[Kri72] p. 298]

Es ist leicht ersichtlich, inwiefern in der Kausaltheorie die anderen Sprecher einer Sprachgemeinschaft an Bedeutung für die Referenz gewinnen. Explizit heißt es:

„In general our reference depends not just on what we think ourselves, but on other people in the community, the history of how the name reached one, and things like that. It is by following such a history that one gets to the reference.“ [[Kri72] p. 301]

In diesen zwei Textstellen „schlummert“ ein Gedanke, der auf sehr grundlegender Ebene nicht nur von Frege und Russell wegführt, sondern, geeignet entwickelt, jene internalistische Tradition in Frage stellen *könnte*, der die Philosophie seit ihren Anfängen folgt: die Auffassung, daß es allein von uns selbst und unserem Inneren abhängt, was wir denken.

Wenn man Kripkes Argumente akzeptiert, stößt man also im Rahmen der Kausaltheorie auf die Frage, inwiefern Referenz einerseits tatsächlich durch internalistische Elemente – mentale Konstellationen des Sprechers – festgelegt ist, inwiefern dieser andererseits den Bezug und die Bedeutung von Ausdrücken durch die Prägung übernimmt, die er von seiner Sprachgemeinschaft erfährt, und inwiefern diese Prägung für die Beschaffenheit des Bedeutungsbegriffs wichtig ist. Mit anderen Worten: wir können, von Kripke ausgehend, eine zweite Art von Merkmalen identifizieren, die für die Beschaffenheit von Bedeutung und Referenz entscheidend sein könnten und die bisher niemand berücksichtigt hat: Elemente, die *nicht* im Kopf des Sprechers sind, sondern der Sprachgemeinschaft kollektiv zukommen. Auch der Ruf nach subjektivitätskompetenten Bedeutungsbegriffen, die subjektive wie objektive Dimensionen abbilden können, wird damit lauter. Bisher kam dieser Ruf von der semantischen Seite der Bedeutungstheorie, weil wir extensionale Wahrheitsbedingungen für Überzeugungszuschreibungen haben wollten. Jetzt aber deutet sich an, daß wir auch von der kognitiven Seite der Bedeutungstheorie her mit der Forderung nach erweiterten Bedeutungsbegriffen konfrontiert werden, weil es nämlich bedeutungsbestimmende Faktoren abseits des Sprecherbewußtseins gibt – und, wie wir noch sehen werden, weil Referenz seitens des Sprechers intentional ist. Doch in diesem Stadium sind das noch Spekulationen. Allerdings deutet sich an, daß Kripkes Thesen den Internalismus, der im Rahmen von Kapitel I nie in Frage gestellt wurde, in einen unangenehmen Erklärungs- und Rechtfertigungsdruck bringen können.

Diese spannenden Gedanken werde ich in Kapitel IV wieder aufnehmen. Doch vorher werden wir noch einen Begriff erarbeiten, der von zentralem Gewicht für spätere Überlegungen in der Externalismusdebatte werden wird und der ermöglicht hat, Freges Lösungsstrategie für das Informativitätsproblem trotz Kripkes Gedankenwende noch zu verwirklichen: den Begriff des Charakters, der beschrieben und entwickelt wurde von David Kaplan.

# III Eine Theorie des Indexikalischen: Kaplan

## III.1 Einige wichtige Unterscheidungen

David Kaplan verfaßte 1977 ein folgenreiches Manuskript „Demonstratives“, das erst 1989 als [Kap89a] veröffentlicht wurde. Bereits während dieser zwölf Jahre (und auch danach) entfaltete die Theorie, die in diesem Text dargestellt wird, eine ungeahnte Wirkung, die man als Aufkommen einer neuen Orthodoxie in der Theorie der Bedeutung bezeichnen muß.

Kaplan zeigt in diesem Text, wie die Lösungsstrategie, die Frege für das Informativitätsproblem von Identitätssätzen anbot, trotz der Kripkeschen Einwände partiell verwirklicht werden kann – beschränkt auf Sätze mit bestimmten indexikalischen Wörtern, also solcher Wörter, deren Extension vom Äußerungskontext abhängt, wie zum Beispiel „er“, „der dort“, „hier“, „jetzt“ oder „ich“. Außerdem bietet Kaplan eine überaus einleuchtende Theorie der Bedeutung indexikalischer Wörter an, womit wir uns nach Kripkes Korrektur des philosophischen Rahmens vorübergehend wieder der rekursiven Semantik und der rein objektiven Bedeutungsbegriffe zuwenden.

Im Rahmen von Kaplans Theorie wird – das ist das Interessanteste – erstmals der Begriff des „Charakter“ verwandt: der Charakter eines Wortes bezeichnet eine Funktion, mit deren Hilfe die kognitive Signifikanz des betreffenden Wortes semantisch repräsentiert werden kann. Insofern kann Kaplans Theorie als ein weiterer wichtiger Baustein zu einer allgemeinen Theorie der kognitiven Signifikanz von sprachlichen Ausdrücken (und Gedanken) interpretiert werden, obwohl er – ähnlich wie Carnap – sich auf die ahistorische Zustandsbeschreibung einer bestehenden Sprache beschränkt.

Kaplans Theorie bezieht sich ursprünglich auf Gegenstandsbezeichnungen, das heißt Designatoren. Hinsichtlich der Designatoren kann man zwei hauptsächliche Arten unterscheiden, die absoluten Designatoren, deren Referenz nicht vom Kontext abhängt (was nicht mit Starrheit verwechselt werden darf) und die indexikalischen Designatoren oder kurz Indexikale<sup>1</sup>, deren Referenz von Kontext zu Kontext variiert (auch viele solcher Designatoren können in einem etwas erweiterten Sinne von Starrheit als starr gelten). Bezüglich der indexikalischen Designatoren unterscheidet Kaplan zwei Arten: die reinen Indexikale einerseits, deren Referenz allein durch semantische Regeln hinreichend festgelegt wird, und die Demonstrativa, bei denen eine Kontextinformation, im deutlichsten Fall eine Zeigegeste hinzutreten muß, um die Referenz eindeutig festzulegen. Diese Kontextinformation nennt Kaplan eine „Demonstration“. Beispiele für absolute Designatoren sind zum Beispiel Namen, reine Indexikale sind Wörter wie „ich“ oder „gestern“ und Demonstrativa sind beispielsweise „der da“, „das dort“, „sie“ undsofort.

---

<sup>1</sup>„Indexikal“ ist in der Tat die Übersetzung von „Indexical“. Dabei soll nicht an ein substantiviertes Adjektiv gedacht werden wie zum Beispiel bei „die Himmlischen“ (=ein „Indexikalisches“), sondern an ein eigenständiges Substantiv, ähnlich wie bei „Differential“ oder „Integral“.

Die Unterscheidung zwischen reinen Indexikalen und Demonstrativa kann man auch so formulieren: Die Festlegung des Referenten eines reinen Indexikals hängt vom Kontext ab und die Festlegung des Referenten eines Demonstrativums von der assoziierten Demonstration.

Das kann man leicht nachvollziehen: das Wort „hier“ zum Beispiel kann man in dem Satz „Ich bin jetzt hier“ als reines Indexikal gebrauchen. In dem Satz „Nächste Woche bin ich hier“ gebrauche ich, wenn ich bei der Äußerung von „hier“ auf einen Punkt auf der Landkarte deute, dagegen dieses Wort als ein Demonstrativum.

Bisher wurde viel über mögliche Welten und Kontexte von Äußerungen geredet. Kaplan trifft hier eine fundamentale Unterscheidung, die sich letztlich als der entscheidende Rettungsanker für Freges Strategie zur Erklärung der Informativität und als zentral für die weitere Sprachphilosophie erweisen wird. Kaplan unterscheidet zwischen solchen möglichen Welten, in denen ein Satz *geäußert* wird und solchen, in denen er *ausgewertet* wird. Ich werde in diesem Zusammenhang zur Bezeichnung der ersteren Art möglicher Welten von Äußerungskontexten [*context of utterance*] oder kurz Kontexten sprechen, zur Bezeichnung der letzteren Art von Auswertungssituationen [*circumstances of evaluation*] oder kurz Situationen sprechen<sup>2</sup>.

Mit diesen Differenzierungen lassen sich nun die Erweiterungen herausarbeiten, die die Fregesche Theorie braucht, um überleben zu können.

### III.2 Direkte Referenz, Äußerungskontexte und Auswertungssituationen

Wichtig für die korrekte Darstellung von Kaplans Semantik der Indexikale ist der Begriff der direkten Referenz.

Nach Kaplan beging die Frege und Carnap verpflichtete Tradition der Bedeutungstheorie neben der Gleichsetzung von Namen mit Kennzeichnungen, die Kripke widerlegt hat, noch einen weiteren, großen Fehler, indem sie für alle sprachlichen Ausdrücke dieselbe Art von Referenz annahm. Diesen Referenztyp nennt Kaplan deskriptiv.

Fregesche deskriptive Referenz kennen wir noch aus Kapitel I: Die Konventionen und Regeln einer natürlichen Sprache bestimmen die Intension (Frege: den Sinn) eines sprachlichen Ausdrucks. Diese Intension bestimmt in Abhängigkeit von der Beschaffenheit der Welt, die als Argument übergeben wird, die aktuelle Extension des Ausdrucks. Die Festlegung der Extension eines Ausdrucks  $\alpha$  erfolgt also durch die logische „Hintereinanderschaltung“ dieser beiden Bestimmungsfaktoren; ohne sie ist keine Extension zugeordnet.

Nach Kaplan gibt es jedoch noch eine andere Art der Referenz, die sogenannte direkte Referenz. Bei Ausdrücken mit direkter Referenz bestimmen die sprachlichen Regeln und Konventionen für jeden gegebenen Kontext unmittelbar die entsprechende Extension – eben

---

<sup>2</sup>Kaplan will sich nicht auf die Mögliche-Welten-Semantik einlassen und ich will versuchen, dies in meinem Sprachgebrauch zu berücksichtigen. Auch Sätze und Propositionen faßt er nicht als Angaben über mögliche Welten auf, sondern lediglich als eine Form strukturierter Entitäten. Deshalb rede ich wie er von Kontexten und Situationen, nicht von Welten. Wir werden bald sehen, daß Kontexte auch nicht ganz dasselbe sind wie Welten.

*direkt*, ohne Umweg über eine vermittelnde Intension. Dabei steht nicht, wie bei Kripke, der Gedanke im Vordergrund, daß sich *herausstellen* könnte, daß ein Ausdruck denselben Referenten in allen möglichen Situationen bezeichnet, sondern daß die semantischen Regeln für einen gegebenen Ausdruck *direkt* dazu führen, daß in allen möglichen Situationen der Referent derselbe ist (vgl. [Kap89a] p. 493). Die Intension wird dann erst in einem zweiten Schritt festgelegt – und zwar durch die Extension. Dies liegt daran, daß es der durch den Kontext und die sprachlichen Konventionen festgelegte Referent selbst ist, dessen Verhalten in kontrafaktischen Umständen untersucht wird.

Ein gutes Beispiel für direkte Referenz ist das Wort „ich“. Daß das reine Indexikal „ich“ immer denjenigen bezeichnet, der dieses Wort äußert, ist eine semantische Regel und sie bestimmt *direkt* den Referenten. Ein anderes gutes Beispiel für direkt referentielle Ausdrücke sind zum Beispiel Eigennamen.

Wenn ich zum Beispiel sage: „Ich bin heute sehr müde“, so wird dieser Satz analysiert, indem kontrafaktische Umstände betrachtet werden, in denen *ich* mich befinde – nicht ein beliebiger müder Sprecher. Die Referenz von „ich“ bin nämlich ich – und damit haben wir auch eine Intension für „ich“. Das funktioniert natürlich nur, wenn diese Extension als konstant angenommen wird, wenn man also annimmt, daß die Intension auch tatsächlich durch eine bestimmte Extension eindeutig festgelegt wird. Das aber heißt, direkt referentielle Ausdrücke sind in Kripkes Sinne starr.

Wenn das Gesagte ausgewertet wird, so wird nämlich nur *ein bestimmtes Objekt* für *alle* möglichen Auswertungssituationen relevant sein. Der Kontext der Äußerung kann zwar eines aus vielen Objekten herausgreifen, doch dieses Objekt ist dann für alle möglichen Auswertungssituationen das richtige (nicht aber für alle Kontexte). Damit ist bereits ein wichtiger Grund für die Unterscheidung zwischen Kontexten und Auswertungssituationen aufgezeigt; darüber hinaus ist dies eine einleuchtende Erklärung für die Starrheit solcher Gegenstandsbezeichnungen, die bei Kripke weitgehend unerklärt bleibt.

Man könnte auch versöhnlicher sagen: direkt referentielle Ausdrücke sind starre Designatoren im *starken Sinne*, das heißt sie bezeichnen in jeder Welt dasselbe Objekt, egal, ob es dieses Objekt gibt oder nicht. Allerdings sind nicht alle starren Designatoren direkt referentiell. Kaplan widerspricht Kripke auch in der Auffassung, daß starre Designatoren keine deskriptive Bedeutung (oder auch: Sinn) hätten – denn reine Indexikale zum Beispiel *sind* starre Designatoren und sie *haben* in dem Sinne eine deskriptive Bedeutung, daß zum Beispiel „ich“ in jeder möglichen Welt den Sprecher bezeichnet. Das ist eine Form deskriptiver Bedeutung. (Eigennamen allerdings, da stimmt Kaplan zu, haben keine deskriptive Bedeutung. Starrheit wird uns in diesem Kapitel nicht weiter interessieren: bei der Besprechung Putnams wird sie wieder von zentraler Bedeutung sein.)

Nach Kaplan sind alle Indexikale direkt referentiell, sowohl reine Indexikale wie auch Demonstrativa. Für beide Arten von Indexikalen läßt sich diese Behauptung auch durch starke Argumente stützen, die ich hier jedoch nicht wiedergeben will (vgl. [Kap89a] ch. V (p. 498 ff.) und IX (p. 512 ff.)).

Reine Indexikale haben eine deskriptive Bedeutung, die in der Regel ziemlich einfach anzugeben ist, wie im Fall von „ich“. Aber diese Bedeutung ist nur relevant für die Festlegung des Referenten in einem *Gebrauchskontext*, nicht dagegen in einer *Auswertungssituation*. Die deskriptive Bedeutung ist für die Bestimmung des Referenten in einer gegebenen Auswer-

tungssituation entweder unanwendbar oder irrelevant. Auch dafür gibt es Beweise, die ich nicht explizieren mag, doch man muß sich, um das zu verstehen, das folgende Beispiel ansehen:

Wenn ich zu einem Zeitpunkt  $t_0$  sage „Alles, was jetzt schön ist, wird vergehen“ und beabsichtige, diesen Satz zu einem späteren Zeitpunkt  $t_1$  auszuwerten, welche Zeit ist dann für die Bestimmung des Wahrheitswerts des Satzes relevant,  $t_0$  oder  $t_1$ ? Offensichtlich der Zeitpunkt des Äußerungskontextes, also  $t_0$ , – denn ich meine mit „jetzt“  $t_0$ . Die Auswertungssituation zum Zeitpunkt  $t_1$  ist für die Feststellung des Wahrheitswertes dieses Satzes dagegen einfach nicht anwendbar.

Die Unterscheidung von Gebrauchskontexten und Auswertungssituationen ist also für die Beschreibung direkter Referenz von zentralem Gewicht, während man diese Unterscheidung für die Darstellung aller hier besprochenen Theorien rekursiver Semantik vor Donnellan nicht benötigt. Welche theoretische Erklärungskraft diese Unterscheidung entfaltet – und warum – wird gleich klar werden.

### III.3 Die doppelte Abhängigkeit der Indexikale

Aus dem, was wir bisher erarbeitet haben, läßt sich nur folgendes schließen:

- a) Bei Indexikalen fallen intuitiver Bedeutungsbegriff und Intension nicht zusammen. Wäre das so, dann würde die Angabe einer Intension ausreichen, um die „Bedeutung“ von Wörtern wie „ich“ anzugeben. Das ist aber nicht der Fall, weil die Intensionsfunktion nicht zwischen Äußerungskontexten und Auswertungssituationen entscheidet. Was uns zu Punkt
- b) führt: offensichtlich unterliegt die Extension eines indexikalischen Ausdrucks einer doppelten Abhängigkeit. Einerseits ist für jeden möglichen Äußerungskontext eine Intension durch die direkt ermittelbare Extension festgelegt. Wen „ich“ nämlich bedeutet, hängt von diesem Äußerungskontext ab, der an die Intensionsfunktion angeschlossen ist. Andererseits aber bestimmt die Intension auch für jede mögliche Auswertungssituation eine Extension.

Das verdient weitere Erklärung: sehen wir uns die Regel an, wonach die Extension von „ich“ in jedem Gebrauchskontext derjenige ist, der dieses Wort äußert. Diese Regel beschreibt offensichtlich *nicht* die Intension von „ich“. Zwar formuliert sie eine Abhängigkeit der Extension von der Beschaffenheit einer möglichen Welt – doch das ist nicht die Abhängigkeit, um die es bei Intensionen geht. Um die Intension eines Ausdrucks angeben zu können, ist es nötig, zu wissen, welche Extensionen dieser unter kontrafaktischen Umständen hat. Nun ist es aber so, daß die angegebene Regel sagt, daß ich mit dem Wort „ich“ immer mich meine – nicht einen kontrafaktischen Sprecher, selbst wenn er eindeutig bestimmt wäre. Wenn ich sage: „Alles wäre anders gewesen, hätte ich doch das Stoppschild gesehen“, so rede ich über eine mögliche Welt, in der alles anders ist als hier, weil *ich* das Stoppschild gesehen habe, nicht der Fahrer des Wagens oder sonstwer. Das heißt, „ich“ bezeichnet den



Sprecher auch in den möglichen Welten, von denen er redet. Anders gesagt: Um die Bedeutung von „ich“ zu beschreiben – das drückt der obige Punkt a) aus – *genügt* es eben nicht, nur eine Intension anzugeben, denn es wäre nur *eine*. Zwar kann ich sagen, daß „ich“ immer mich bezeichnet, doch die Intension von „ich“ ordnet nur denjenigen Gebrauchskontexten den Wahrheitswert „wahr“ zu, in denen ich „ich“ sage. Doch auch ein anderer kann „ich“ sagen und einen wahren Satz damit formen. Was die Regel, daß „ich“ immer den Sprecher bezeichnet, ausdrückt, ist also keine Intension.

Jetzt sehen wir zwei Dinge: erstens ist „ich“ ein starrer Designator – dieser Ausdruck bezeichnet überall den Sprecher, der den Ausdruck äußert. Das ist nach den Ausführungen von Abschnitt 2 aber keine Überraschung mehr. Zweitens ist offensichtlich nicht nur die Referenz von „ich“ kontextabhängig, sondern auch die Intension. Wir müssen, um die Bedeutung von „ich“ anzugeben, nicht nur die Intension aufführen, sondern auch angeben, inwiefern diese vom Äußerungskontext abhängt. Um diese doppelte Abhängigkeit zu beschreiben, benötigen wir einen neuen Begriff.

### III.4 Die zwei Arten der Bedeutung: Intension und Charakter

Wenn Heike sagt: „Ich war gestern am See spazieren“, sagt sie etwas völlig anderes, als ich, wenn ich denselben Satz äußere. Das kann man leicht sehen, indem man sich vor Augen führt, daß sie, wenn sie den Satz heute äußert, einen wahren Satz formt, während ich einen falschen Satz äußern würde. Selbst wenn auch ich aber spazieren gewesen wäre, ist es zumindest *möglich*, daß die Wahrheitswerte unserer Sätze sich in anderen Kontexten unterscheiden würden.

Insofern unterscheiden sich unsere Sätze in etwas, das Kaplan *Inhalt* nennt. Inhalt kommt jedem Bedeutung tragenden Ausdruck innerhalb eines Kontexts zu, nicht nur Sätzen. Der Inhalt eines Satzes ist zum Beispiel eine Proposition, das Resultat der Auswertung wird ein Wahrheitswert sein. Generell gilt, daß das Resultat einer Auswertung des Inhalts eines wohlgeformten Ausdrucks  $\alpha$  innerhalb einer Auswertungssituation eine geeignete Extension für  $\alpha$  ist (vgl. [Kap89a] p. 501). Das kommt uns zu Recht bekannt vor: wenn man den Inhalt eines Ausdrucks als eine Funktion von Auswertungssituationen in Extensionen repräsentieren kann, dann läßt sich der Inhalt in diesem Sinne als Intension darstellen, allerdings unter Berücksichtigung der Unterscheidung zwischen Kontexten und Auswertungssituationen. Nach Kaplan sind Inhalt und Intension *nicht* dasselbe, der Inhalt läßt sich lediglich als Intension *darstellen*. Auf diese subtile Unterscheidung möchte ich zwar hinweisen, doch macht sie für die weitere Darstellung keinen Unterschied<sup>3</sup>.

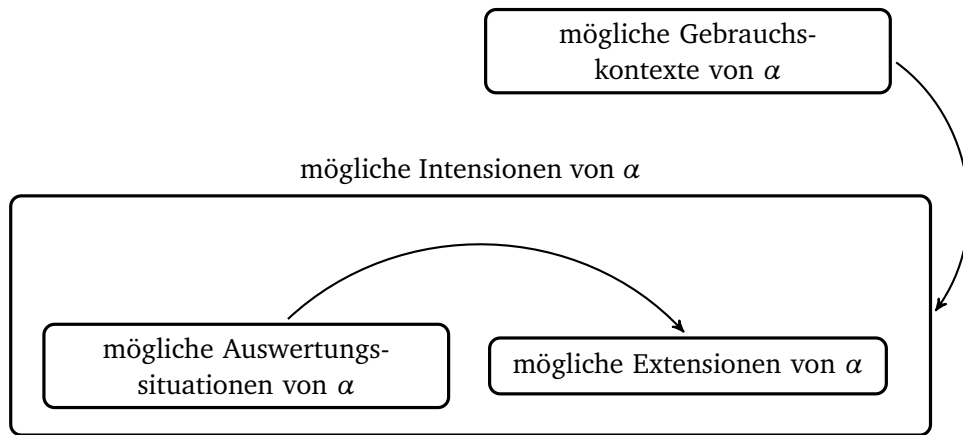
Kommen wir nun zu Kaplans zweitem und wesentlich interessanteren Begriff von Bedeutung. Die Bedeutung eines Ausdrucks, so Kaplan, besteht in seinem *Charakter*. Der Charakter eines Ausdrucks ist eine Funktion von Gebrauchskontexten in Intensionen. Das bedeutet, der Charakter des Ausdrucks legt fest, in welchen Kontexten der Ausdruck welche Intension

---

<sup>3</sup>Wie man in Kapitel V sehen wird, wurde der Begriff des Inhalts (in anderem Sinne) ziemlich ins Dunkle gezerzt, weswegen ich aus Gründen der Deutlichkeit den Inhalt im Kaplanschen Sinne im folgenden vereinfachend als ‚Intension‘ bezeichnen werde.

hat. Ein Charakter mit einem Kontext liefert eine Intension, eine Intension und eine Auswertungssituation liefern eine Extension:

### Charakter von $\alpha$



(Die Pfeile sollen darstellen, daß es sich um eine Funktion handelt – daß also eine Zuordnung bzw. Abbildung von Elementen der einen Menge in Elemente der anderen Menge stattfindet.)

Der Charakter eines Ausdrucks ordnet also jeder Situation, in der wir diesen Ausdruck verwenden, die Bedeutung zu, die dieser Ausdruck unter den gegebenen Umständen hat.

Das heißt aber, daß wir jetzt ein gutes Mittel gefunden haben, die *kognitive Signifikanz* mindestens indexikalischer Wörter tatsächlich abzubilden – oder anders gesagt: die Gestalt, die solche Wörter „in unserem Kopf“ haben, die Rolle, die sie beim Verstehen spielen. Der Charakter eines Wortes ordnet jedem Kontext, in dem es verwendet wird, eine Intension zu, die uns sagt, wie wir dieses Wort auswerten müssen, um zur korrekten Extension zu kommen. Auch das hervorstechendste Merkmal der Indexikale können wir jetzt beschreiben:

„Indexicals have a context-sensitive character. It is characteristic of an indexical that its content varies with context. Nonindexicals have a fixed character. The same content is invoked in all contexts. This content will typically be sensitive to circumstances, that is, the nonindexicals are typically not rigid designators but will vary in extension from circumstance to circumstance.“ [[Kap89a] p. 506]<sup>4</sup>

Dazu ein Beispiel: Wie sähe eine Analyse des Satzes „Ich bin jetzt hier“ mithilfe des Charakterbegriffs aus?

Es sei für jeden Ausdruck  $\alpha$  einer gegebenen Sprache  $L$  der Charakter  $\|\alpha\|$  dieses Ausdrucks eine Funktion, die jedem Äußerungskontext  $c$  aus der Menge  $\mathcal{C}$  der möglichen Kontexte die Intension  $\|\alpha\|(c)$  von  $\alpha$  in  $c$  zuordnet. Die Intension  $\|\alpha\|(c)$  von  $\alpha$  in  $c$  ist eine Funktion, die jeder möglichen Auswertungssituation  $w$  aus der Menge  $\mathcal{W}$  der möglichen

<sup>4</sup>„Starrer Designator“ heißt hier: starr in dem von Kripke intendierten Sinn, daß die Extension in jeder Welt dieselbe ist. In einem modifizierten Sinne, das hatten wir gesagt, sind auch Indexikale starr.

Auswertungssituationen die Extension  $\|\alpha\|(c)(w)$  von  $\alpha$  in  $c$  und  $w$  zuordnet. Die Extension  $\|\alpha\|(c)(w)$  von  $\alpha$  in  $c$  und  $w$  ist eine Entität, die – wie bei Carnap – der syntaktischen Natur von  $\alpha$  angemessen ist: ein Gegenstand, falls  $\alpha$  ein Designator ist, eine Menge von Gegenständen, falls  $\alpha$  ein einstelliges Prädikat ist, ein Wahrheitswert, falls  $\alpha$  ein Satz ist, undsofort. Dann gilt:

$\|\text{ich}\|(c)(w) = \text{der Sprecher von } c \text{ (sofern er in } w \text{ existiert, sonst undefiniert)}$

$\|\text{jetzt}\|(c)(w) = \text{die Zeit von } c$

$\|\text{hier}\|(c)(w) = \text{der Ort von } c$

und demnach:

$\|\text{„Ich bin jetzt hier“}\|(c)(w) = \text{wahr}$

$\Leftrightarrow \|\text{„Ich befinde mich jetzt hier“}\|(c)(w) = \text{wahr}$

$\Leftrightarrow \langle \|\text{ich}\|(c)(w), \|\text{jetzt}\|(c)(w), \|\text{hier}\|(c)(w) \rangle \in \|\text{sich befinden}\|(c)(w)$

$\Leftrightarrow \text{der Sprecher von } c \text{ befindet sich zur Zeit von } c \text{ am Ort von } c.$

Hierbei bleiben Fragen offen, die in der gerafften Darstellung nicht geklärt werden können, zum Beispiel, was Kontexte genau sind. Kontexte sind normalerweise geordnete Folgen aus Elementen, die Äußerungsumstände sprachlicher Ausdrücke festlegen, also zum Beispiel Zeit, Ort, Sprecher undsofort. (Man nennt eine solche geordnete Folge auch „Index“.) Doch welche Kontextaspekte wirklich relevant für die Auswertungssituation werden, ist eine delikate Frage: tatsächlich kann die ganze Kontextwelt relevant sein, weshalb Kontexte extrem schwer formal darstellbar sind. Außerdem müßte man auch genauer erörtern, daß Charakter partiell definierte Funktionen sind, wie durch den Klammerzusatz beim oben dargestellten Charakter von „ich“ deutlich werden soll. Dieser Vorbehalt dient der Genauigkeit und ist gerade im Hinblick auf Kripkes Überlegungen zur starren Designation wichtig. Außerdem sieht man, daß Charakter rekursiv aufgebaut sind: der Charakter eines komplexen Ausdrucks setzt sich aus den Charakteren seiner Teilausdrücke zusammen. Auch das verdiente genauere Beleuchtung, die ich hier nicht leisten kann. Jedenfalls ist sichtbar, daß der Charakter ein zentraler Baustein einer rekursiven Semantik ist.

Wörter wie „ich“ jedenfalls unterscheiden sich, wie wir gesehen haben, in einem sehr wichtigen Sinne von Wörtern wie „Holz“<sup>5</sup> oder „Cäsar“: sie haben einen *variablen* Charakter, der verschiedenen Gebrauchskontexten verschiedene Intensionen zuordnet, während nicht-indexikalischen Wörtern in jedem Gebrauchskontext dieselbe Intension zugeordnet wird. Nichtindexikalische Wörter haben daher einen konstanten Charakter. Der Charakter eines Wortes fällt mit seiner objektiven Bedeutung zusammen und gibt dessen kognitive Signifikanz an: ein Synonym für ein Wort angeben heißt, ein Wort mit dem gleichen Charakter angeben. Ein Wort mit der gleichen Intension angeben, reicht nicht aus, denn „ich“, von

<sup>5</sup>Wir werden im nächsten Kapitel sehen, daß auch natürliche Prädikate wie „Holz“ einer Spielart von Indexikalität unterliegen. Auch diese Art von Indexikalität könnte man mit dem Charakterbegriff explizieren, doch das ist nicht Gegenstand meiner Betrachtung, obwohl ich aus Gründen der gedanklichen Vollständigkeit darauf hinweisen will.

mir geäußert, *bedeutet* in diesem Sinne dasselbe, wie „du“ von dir zu mir gesagt. Das heißt nicht, daß „ich“ und „du“ Synonyme sind.

Über die deskriptive Bedeutung von Indexikalen sagt Kaplan:

„Insofar as meaning is given by the rules of a language and is what is known by competent speakers, I would be more inclined to say in the case of direct referential words and phrases that their reference is no part of their meaning. The meaning of the word ‚I‘ does not change when different persons use it. The meaning of ‚I‘ is given by the rules (D1), (D2), and (D3) above.“ [[Kap89a] p. 520]<sup>6</sup>

„The meaning (character) is directly associated, by convention, with the word. The meaning determines the referent; and the referent determines the content.“ [ibid. fn. 44]

Ich möchte noch ein Beispiel einfügen, das illustrieren soll, weswegen der Charakter eines Wortes dessen kognitive Signifikanz abbildet: Frege spricht in „Der Gedanke“ über das Problem der Mitteilbarkeit von ichbezogenen Gedanken. Nehmen wir Gedanken in Freges Sinn als Propositionen (was sie faktisch sind), so hieße das, daß Propositionen, die das Wort „Ich“ enthalten, nie restlos mitteilbar sind, weil jedes Individuum sich selbst (durch sein Bewußtsein) in besonderer, sozusagen privilegierter Weise gegeben ist. Wenn also Dr. Lauben denkt „Ich bin verwundet worden“, so kann er diesen Gedanken nicht restlos mitteilen. Für dieses Problem gibt es eine raffinierte semantische Lösung (p. 535, fn. 59), doch hier soll uns nur seine Umformulierung interessieren: Kaplan sagt, daß diese (semantisch nur scheinbare) Privilegierung darauf zurückgeht, daß Dr. Lauben der einzige ist, der sich selbst als Extension unter dem Charakter von „Ich“ gegeben ist (p. 533). Diese Tatsache ist, finde ich, ein schönes Beispiel, inwiefern der Charakter die kognitive Signifikanz eines Wortes abbildet.

Das heißt: es gibt eine Möglichkeit, abzubilden, welche kognitive Rolle ein korrekt gebrauchtes Wort innerhalb einer Sprache spielt, die Angabe seines Charakters. Die Anschlußfähigkeit an Kripkes Überlegungen zur starren Designation wurden im Text aufgezeigt. Der Charakter als Begriff der kognitiven Signifikanz ist jedoch auch zu Kripkes kausaler Theorie der Referenz kompatibel, denn Kaplan stellt keine Theorie darüber auf, wo der Charakter „herkommt“ oder wie er entsteht. Seine These betrifft ausschließlich semantische Merkmale einer als gegeben vorgestellten Sprache. Zugleich ist dies, wie in Carnaps Fall, eine gewisse Einschränkung: in Kapitel IV werden wir auf Probleme stoßen, die wir nicht alle mit dem Charakterbegriff lösen können. Wir wollen jedoch noch sehen, inwieweit wir Kripkes Begrifflichkeiten mithilfe des Charakterbegriffs abbilden können, um zu zeigen, wie fruchtbar und hilfreich dieser Begriff ist<sup>7</sup>.

---

<sup>6</sup>Die drei Bedingungen sind semantische Regeln: (D1): „Ich“ ist ein Indexikal, verschiedene Verwendungen werden verschiedene Inhalte (Intensionen) haben. (D3): „Ich“ ist in jeder Äußerung direkt referentiell. (D2): In jeder Äußerung wird „ich“ denjenigen bezeichnen, der diese Äußerung tut.

<sup>7</sup>Es sollte klar sein, daß wir auch Donnellans über Carnap hinausgehende Begriffe der attributiven und der referentiellen Lesart bei Kennzeichnungen mithilfe des Charakterbegriffs darstellen können. Diese Lesarten unterscheiden sich nur durch die Art der Bezugnahme des Charakters auf die jeweils angegebene Kontextwelt. In attributiver Lesart ist der Charakter

### III.5 Kripke und Kaplan

Kripke hat die Begriffe der Apriorität und der Notwendigkeit entscheidend revidiert. Danach ist, wir erinnern uns, Apriorität eine epistemische, Notwendigkeit hingegen eine metaphysische Modalität. Lassen Kripkes Begriffe dieser Modalitäten sich mithilfe von Kaplans Charakterbegriff darstellen?

Sehen wir uns den oben analysierten Beispielsatz „Ich bin jetzt hier“ an. Der Charakter von „Ich bin jetzt hier“ könnte für drei verschiedene Kontexte  $c_i$  und drei verschiedene Auswertungssituationen  $w_i$  folgendermaßen aussehen:

„ich bin jetzt hier“	$w_1$	$w_2$	$w_3$
$c_1$	wahr	falsch	wahr
$c_2$	wahr	wahr	falsch
$c_3$	falsch	falsch	wahr

Was können wir über die Notwendigkeit des Satzes „Ich bin jetzt hier“ mithilfe dieser Tabelle aussagen? Offensichtlich hat der Satz für jeden Kontext  $c$  zumindest in einigen Auswertungssituationen  $w$  den Wahrheitswert „falsch“. Damit wissen wir bereits, daß der Satz in keinem Kontext in Kripkes Sinne notwendigerweise wahr sein kann. Ein Satz  $S$  ist dann in einem bestimmten Kontext  $c$  notwendigerweise wahr, wenn es für diesen Kontext keine Auswertungssituation  $w$  gibt, derart, daß  $S$ , in  $c$  geäußert, in  $w$  den Wahrheitswert „falsch“ annimmt; anders gesagt: ein Satz  $S$  ist metaphysisch notwendig genau dann, wenn für alle Auswertungssituationen  $w$   $\|S\|(c)(w) = \text{wahr}$  ist, also in der Zeile oder *Horizontale*, die zu  $c$  gehört, nie etwas anderes als „wahr“ steht.

Wie sieht es nun hinsichtlich der Apriorität von „Ich bin jetzt hier“ aus? Dieser Satz kann nie falsch geäußert werden, also ist er in jedem Äußerungskontext wahr, denn in jedem Kontext, in dem der Sprecher diesen Satz äußert, befindet er sich zur Kontextzeit am Kontextort. Doch wir müssen den Bezug zu einer Auswertungswelt herstellen, um etwas über die Wahrheit des Satzes aussagen zu können. Eine Äußerung findet immer in einer Kontextwelt statt. Wahrheit eines Satzes  $S$  in einem Kontext  $c'$  heißt also, daß  $S$  wahr ist, sofern  $w$  die Kontextwelt von  $c$  ist, Kontextwelt und Auswertungssituation also zusammenfallen. Anders ausgedrückt, wenn  $w_c$  die Welt sei, in der der Kontext  $c$  situiert ist, dann gilt  $\|S\|(c)(w) = \text{wahr}$ . In der Tabelle kann man dies dadurch umsetzen, daß man  $w_1$  als Kontextwelt für  $c_1$  annimmt,  $w_2$  für  $c_2$  undsofort. Dann ist für  $n \in 1, 2, 3 \dots$   $\| \text{ich bin jetzt hier} \|(c_n)(w_n) = \text{wahr}$ , das heißt, die Diagonale der obigen Tabelle enthält durchweg den Wahrheitswert „wahr“. Wenn ein Satz also in jedem Äußerungskontext wahr ist, so wissen wir auch ohne genaue

des Ausdrucks „der Erbauer des Eiffelturms“ folgender:  $\| \text{der Erbauer des Eiffelturms} \|(c)(w) =$  das Individuum, das in  $w$  den Eiffelturm erbaut hat, sofern es genau ein solches Individuum gibt.

In referentieller Lesart, in der wir uns auf ein bestimmtes Individuum beziehen, hieße das:  $\| \text{der Erbauer des Eiffelturms} \|(c)(w) =$  das Individuum, das in  $w_c$  den Eiffelturm erbaut hat, sofern es in  $w_c$  genau ein solches Individuum gibt und es in  $w$  existiert. Eigentlich müßte man jedoch den Charakter für die referentielle Lesart so weit ausdehnen, daß die Referenz auch glückt, wenn niemand den Eiffelturm erbaut hat, denn das beabsichtigt Donnellan. Es sollte jedenfalls klar sein, daß dies kein ernstes Problem mehr ist.

Kenntnis der Beschaffenheit des jeweiligen Kontexts, daß dieser Satz wahr sein wird und in genau diesem Sinne ist er a priori wahr. (Kaplan hat damit den Aprioritätsbegriff auf fruchtbare Weise präzisiert, die jedoch in der Literatur zum Aprioritätsbegriff leider weitgehend ignoriert wird.)

Als analytisch wahr gilt ein Satz, wenn er aufgrund der Bedeutung seiner Bestandteile und deren Zusammensetzung wahr ist. Ein solcher Satz kann niemals falsch werden. Ein Satz ist demnach analytisch wahr genau dann, wenn  $\|S\|(c)(w) = \text{wahr}$  für alle  $c$  und  $w$ , wenn also der Satz in allen Kontexten und Auswertungswelten wahr ist, das heißt, in jedem Kontext notwendig wahr oder aber a priori wahr und in jedem Kontext starr (von konstanter Intension). Der Analytizitätsbegriff Kripkes, der einen Satz als analytisch einstuft, wenn er notwendig und a priori wahr ist, wird dadurch im Großen und Ganzen gestützt.

### III.6 Frege und Kaplan

Kaplan bietet für Freges Problem der Informativität von Identitätssätzen eine partielle Lösung an, die sich auf Sätze mit Demonstrativa bezieht. Wie Kaplan sich Demonstrativa vorstellt, haben wir bisher nicht gesagt.

Demonstrativa sind Wörter wie „dieser da“, „dieser  $F$  dort“, „jenes  $x$ “. Sie unterscheiden sich von den reinen Indexikalen, wie wir schon gesagt haben, dadurch, daß ihre Extension nicht ausschließlich durch semantische Regeln bestimmt wird, sondern daß irgendeine Art von spezifizierender Maßnahme (im deutlichsten Fall: eine Zeigegeste) den Gebrauch des Demonstrativums begleiten muß, um einen Referenten zu bestimmen. Kaplan nennt dies eine *Demonstration*. Diese Demonstration ist eine Eigenschaft, die die semantische Funktionsweise der Demonstrativa der referentiell verwendeten Kennzeichnungen ähnlich macht: man kann sich den Gebrauch eines Demonstrativums mit flankierender Demonstration auch wie eine Kennzeichnung vorstellen. Das heißt: Mit jeder Verwendung eines Demonstrativums  $\alpha_d$  in einem Kontext  $c$  ist eine Eigenschaft  $F_{\alpha_d,c}$  assoziiert, die den Referenten eindeutig bestimmt. Diese Demonstration ist allerdings nicht vollständig im Demonstrativum spezifiziert, weswegen die Vervollständigung der Demonstration Regeln folgt, die diffus und ziemlich schwer zu präzisieren sind. Die Referenz auf den demonstrierten Gegenstand ist jedoch, wie wir wissen, direkt, so daß die Intension, die sich ergibt, konstant bleibt:

$\|\alpha_d\|(c)(w) = \text{das Individuum, das in } w_c \text{ die Eigenschaft } F_{\alpha_d,c} \text{ besitzt}$   
 (sofern es in  $w_c$  genau ein solches Individuum gibt und dieses in  $w$  existiert)

Auch Demonstrativa sind starre Designatoren, wie man daran sehen kann, daß die Extension von  $\alpha_d$  nicht von  $w$  abhängt; die Indexikalität von  $\alpha_d$  zeigt sich darin, daß in verschiedenen Kontexten verschiedene  $F_{\alpha_d,c}$  mit  $\alpha_d$  assoziiert sind und daran, daß Kontexten, in denen dieselbe Eigenschaft  $F_{\alpha_d,c}$  mit  $\alpha_d$  assoziiert wird, jeweils ein anderer Gegenstand diese Eigenschaft erfüllen kann.

Dann aber findet Freges Problem zu einer recht einfachen Lösung, sofern Demonstrativa beteiligt sind.

Der mythische Akteur zeigt abends auf den Abendstern und sagt: „Dieser Stern (der Hesperus heißt) ist. . .“ und am Morgen auf den Morgenstern und sagt „. . . mit diesem Stern

identisch (den ich Phosphorus nenne)“ – der Zeitgenosse müßte nur langsam genug sprechen. Nach Kaplans Theorie der direkten Referentialität sind beide Demonstrativa starre Designatoren, weswegen der Zeuge, sofern er recht hat, etwas notwendig Wahres aussagt. Daß seine Aussage informativ ist, wissen wir. Gleichzeitig aber wissen wir seit Kripke, daß Informativität und Notwendigkeit nichts miteinander zu tun haben. Mit „dieser Stern“ wird im Vordersatz eine Eigenschaft  $F$  assoziiert (die Demonstration) und im Nachsatz mit „diesem Stern“ eine andere Eigenschaft  $G$  (die zweite Demonstration). Der Satz heißt demnach:

S14) Der in  $w_c$   $F$  ist, ist in  $w_c$   $Q$ .

Diese Aussage ist nicht in allen Kontexten wahr, denn der Akteur hätte auf verschiedene Sterne zeigen können. Also wird in der Diagonale des Charakters dieses Satzes auch der Wahrheitswert „falsch“ auftauchen. Die Diagonale wird den Satz auf diese Weise als informativ ausweisen. Man beachte dabei, daß Freges Lösungsstrategie, eine Gegenstandsbezeichnung mit einer beschreibenden Kennzeichnung zu assoziieren dem Prinzip nach aufrecht erhalten und mit Kripkes Einwänden harmonisiert wird. Nur faßt Kaplan diese Kennzeichnung als auf den Kontext und nicht als auf die Auswertungssituation bezogen auf (man kann auch sagen: er liest sie referentiell, nicht attributiv).

Der Begriff des Charakters ermöglicht die formale Darstellung aller bisher besprochenen Bedeutungstheorien. Außerdem bettet Kaplan durch diesen Begriff auch die formale Konkretisierung der kognitiven Signifikanz der Gegenstandsbezeichnungen in die Erklärung mit ein, die gleichzeitig Raum für die Unterbringung metaphysischer und epistemischer Modalitäten bietet. Das ist – auch wenn die Lösung eine nur partielle ist – eine philosophische Leistung.

Der Charakter eines Ausdrucks wird bei Kaplan durch die Konventionen und Regeln der Sprache festgelegt. Damit läßt Kaplan offen, wie diese Signifikanz konkretisiert ist. Auf welche Weise der Charakter eines Ausdrucks etwas abbildet, das der Sprecher im Kopf hat, bleibt unerklärt. Insofern ist Kaplans Beitrag zu dieser Diskussion – ähnlich wie zum Beispiel der von Carnap – auf die semantische Seite der Bedeutungstheorie beschränkt. Der Charakterbegriff und das Konzept der direkten Referenz lassen nichtsdestoweniger Raum für die Andeutungen Kripkes, was die Rolle der Sprachgemeinschaft für die Festlegung der Charakter angeht als auch für die kausale Theorie der Referenz. Insofern nimmt Kaplan zu der Frage, *wie* Bedeutungen konkretisiert sind, nicht wirklich Stellung (dagegen sagt er mit dem Charakterbegriff *wodurch*).

Das nächste Kapitel wird ebenfalls an Kripke anschließen und dessen Überlegungen in eine andere, unerwartete Richtung weiterentwickeln. Wenn wir herausfinden, „wo“ Bedeutungen sind, dann können wir auch eine kausale Erklärung für die kognitive Signifikanz von Ausdrücken angeben. Und bisher gingen wir davon aus, daß Bedeutungen „im Kopf“ sind – könnten sie etwa woanders sein?

## IV Der semantische Externalismus: Putnam und Burge

### IV.1 Das internalistische Konzept der Bedeutung und seine Begrifflichkeiten

Die historischen Positionen, die wir im Rahmen dessen, was ich als Kennzeichnungstheorie bezeichnet habe, bisher betrachtet haben, hatten ein grundlegendes Strukturmoment gemeinsam: alle gingen implizit davon aus, daß Bedeutung etwas sei, das nur in Ansehung des jeweiligen Subjektes beschrieben werden könne und das ausschließlich von dessen inneren Dispositionen abhängt, die zu dessen intrinsischen Eigenschaften zählen. (Kripke kritisiert das, entwickelt jedoch mit der Kausaltheorie ein eher unspezifisches Gegenbild und Kaplan umgeht die Relevanz dieses Einwands durch seine rein formale Orientierung.) Dieser Standpunkt war mehr oder minder unbewußt zu der generellen These verallgemeinert worden, daß die Kenntnis der Bedeutung eines jeweiligen Ausdrucks mit einem bestimmten mentalen Zustand, einer psychischen Disposition, gleichgesetzt werden könne. (Wir erinnern uns an Carnap: eine Intension wäre demnach das semantische Kondensat eines psychischen Zustandes, der uns dazu bringt, einem Ausdruck in jeder Situation eine bestimmte Extension zuzuordnen. In dieser Weise leuchtet ein, inwiefern Bedeutungen im Kopf sein sollen.) Man versteht dabei die Kenntnis der Bedeutung eines Ausdrucks ebenso wie daraus gebildete Überzeugungen als sogenannte *propositionale Einstellungen* [*propositional attitudes*]. Ein Beispiel für eine solche Einstellung ist: „*x* weiß, daß *p*“. Doch um das transparent zu machen, müssen wir ein wenig ausholen: Als propositional bezeichnet man diese Einstellungen, weil man ihnen sogenannten propositionalen Inhalt zuspricht. Dieser propositionale Inhalt ist das, was wir an der Einstellung in Form des daß-Satzes wiedergeben können. Wenn wir einem Individuum *x* eine bestimmte propositionale Einstellung zuschreiben [*ascribe/attribute*], dann sagen wir damit, daß *x* in einer bestimmten Beziehung zu einer Proposition *p* steht, also zum Beispiel „*x* glaubt, daß *p*“ oder „*x* weiß, daß *p*“. Propositionale Einstellungen werden durch solche Sätze Individuen zugeschrieben, daher bezeichne ich diese Sätze als Einstellungszuschreibungen.

Es gibt zwei als basal angesehene Prototypen von propositionalen Einstellungen: erstens solche, die als eine Art von individueller Disposition bezüglich gewisser Informationen gelten können. Zum Beispiel bezeichnen die beiden letztgenannten Zuschreibungsformen Einstellungen dieser Art. Solche propositionalen Einstellungen bezeichnet man als Überzeugungen [*beliefs*], ihre Zuschreibungen demzufolge als Überzeugungszuschreibungen. Die zweite Art von propositionalen Einstellungen bestehen in einer Tendenz eines Individuums, wie zum Beispiel „*x* wünscht sich, daß *p*“, „*x* will vermeiden, daß *p*“ oder „*x* fürchtet, daß



$p$ “. Diese Art von Einstellung subsumiert man gewöhnlich dem (mißverständlichen) Oberbegriff „Wunsch“ [*desire*]<sup>1</sup>. In der Regel werden alle propositionalen Einstellungen einer dieser zwei basalen Kategorien zugeordnet. (Ich werde manchmal abkürzend von Einstellung statt von propositionaler Einstellung sprechen, ebenso verwende ich in Zukunft die Wörter „Überzeugung“ und „Wunsch“ im hier vorgestellten Gebrauch.)

Auf die Bedeutungstheorie kann man den Begriff der propositionalen Einstellung nun anwenden, indem man den psychischen Zustand, der (in der internalistischen Tradition) mit der Kenntnis der Bedeutung  $i$  eines Ausdrucks  $\alpha$  korreliert, als eine propositionale Einstellung wie folgt auffaßt: „ $x$  glaubt, daß  $\alpha$  die Bedeutung  $i$  hat.“ Nimmt man die Extension hinzu, so kann man sich für eine Entität  $e$  und einen Prädikatausdruck  $F$  leicht Einstellungen dieser Art vorstellen: „ $x$  glaubt, daß  $Fp$ “. Ein Beispiel hierfür ist: „Jakob glaubt, daß die Substanz in dem Glas vor ihm Wasser ist.“

Diese propositionalen Einstellungen fallen, so die herkömmliche Vorstellung, insofern mit mentalen Zuständen zusammen, als daß sich mentale Zustände durch Einstellungszuschreibungen angeben lassen<sup>2</sup>. Mentale Zustände korrelieren dabei ihrerseits mit physikalischen Gehirnzuständen. Der frühe Hilary Putnam hat eine philosophische Position beschrieben, die üblicherweise als Funktionalismus bezeichnet wird und die er selbst seit einigen Jahren mit großer Energie bekämpft. Für eine gewisse Zeit jedoch war der Funktionalismus eine Orthodoxie, gegen die innerhalb der Philosophie des Geistes praktisch niemand mehr opponierte. Der Funktionalismus identifiziert mentale Zustände (und in einigen seiner Versionen auch die Gehirnzustände) anhand ihrer jeweiligen funktionalen, das heißt kausalen Eigenschaften. Die Gesamtheit der funktionalen Eigenschaften eines mentalen Zustands wird auch als seine funktionale Rolle bezeichnet. Die funktionale Rolle eines mentalen Zustands  $C$  angeben heißt, angeben:

- 1.) mit welchen Sinneswahrnehmungen,
- 2.) welchen motorischen Aktionen und
- 3.) welchen anderen mentalen Zuständen  $C$  in welcher Weise verknüpft ist, das heißt, von was  $C$  hervorgerufen wird und was  $C$  selbst hervorruft.

Aus funktionalistischer Sicht werden mentale Zustände mit ununterscheidbaren funktionalen Eigenschaften als (typen)identisch betrachtet.

Nimmt man all dies zusammen, so heißt das letztlich, daß jemand, der die Bedeutung eines bestimmten Begriffs kennt, sich in einem mentalen Zustand mit bestimmten funktionalen Eigenschaften befindet, anhand derer wir diesen Zustand identifizieren können. Dieser psychische Zustand soll natürlich nicht als eine Art permanenter physikalischer Zustand

---

<sup>1</sup>Konsequenterweise müßte ich die Zuschreibung einer solchen Einstellung unüblich als „Wunschzuschreibung“ bezeichnen. Da solche Wunschzuschreibungen jedoch in der weiteren Besprechung nicht vorkommen, muß der Leser sich mit solchen Wortmonstern nicht belasten. Der Ausdruck „Überzeugungszuschreibung“ wird auch weiterhin im hier vorgestellten Sinne verwandt.

<sup>2</sup>Diese These ist in den späten Achtziger Jahren von einigen Philosophen kritisiert worden. Der Inhalt der Einstellungszuschreibung, der eingebettete Komplementsatz, gebe, so das Argument, gar nicht wirklich wieder, was im Kopf ist. Eine solche Position werden wir in V.2 kennenlernen. Generell ist jedoch die Auffassung, daß Überzeugungszuschreibungen mentale Zustände denotieren, recht breit akzeptiert.

interpretiert werden. Er kann logisch paraphrasiert werden als ein zweistelliges Prädikat, dessen Argumente ein Individuum und ein Zeitpunkt sind.

Putnam identifiziert zwei Lesarten des Begriffs des psychischen Zustands, eine enge und eine weite.

Psychische Zustände im weiteren Sinne sind schlicht alle psychischen Zustände, wie zum Beispiel „x ist eifersüchtig auf y“ oder „x hat Schmerzen“. Doch es geht Putnam bei seiner unten vorgestellten Theorie um psychische Zustände im engeren Sinne: dies sind solche, die der auf Descartes zurückgehenden Prämisse des methodologischen Solipsismus genügen, das heißt Zustände, die nichts voraussetzen außer der Existenz des Subjekts, das sich in diesem Zustand befindet. In diesem engen Sinne zählt „x hat Schmerzen“ weiterhin als psychischer Zustand, doch „eifersüchtig sein“, ein Ausdruck der logisch als (mindestens) zweistelliges Prädikat paraphrasiert werden muß, zählt in diesem engen Sinne von „Zustand“ nicht als psychischer Zustand. Bei der Darstellung von Putnams Theorie wird es in diesem Kapitel im folgenden immer um psychische Zustände im engeren Sinne gehen.

Die Prämisse, nur psychische Zustände im engeren Sinne als Gegenstand der psychologischen Erklärung herzunehmen, nennt Putnam „methodologischen Solipsismus“. Natürlich ist der methodologische Solipsismus eine präzisierte Form dessen, was ich im Rahmen dieser Arbeit als Internalismus bezeichne. In [Bu79] findet sich in diesem Zusammenhang der Begriff „Individualismus“ und in [Sti83] wird vom „Prinzip der Autonomie“ gesprochen. Fodor unterscheidet in Kapitel 2 von [Fo87] Individualismus und methodologischen Solipsismus, doch es wird nicht klar, ob er unter Individualismus dasselbe versteht wie Burge. Mindestens zwei der drei Bezeichnungen werden von einigen Autoren unterschieden, doch die generelle Idee, die dem Ganzen zugrundeliegt, ist Internalismus in folgendem Sinne: daß psychische Zustände und Eigenschaften, die in wissenschaftlich-psychologischer Erklärung beschrieben und bezeichnet werden, intrinsische Zustände und Eigenschaften des Organismus sind, dessen Verhalten erklärt werden soll.

Frege verdrängte zwar den traditionellen Gedanken, daß Bedeutungen geistige Entitäten seien zugunsten eines Ansatzes, der sie als abstrakte Entitäten ausweist. Die Begründung dafür lautete, daß Bedeutungen<sup>3</sup> „öffentlich“ seien; eine Theorie der Bedeutung müsse der Tatsache Rechnung tragen, daß alle Sprecher einer gegebenen Sprachgemeinschaft auf die Bedeutung ihrer Ausdrücke Zugriff haben müßten. Dennoch blieb aber das Erfassen dieser abstrakten Entitäten durch das Subjekt nach wie vor ein individueller psychischer Akt und das wurde in keiner der Theorien in Kapitel I und auch durch Kaplan nicht angegriffen, der eine Theorie der Beschaffenheit von Bedeutung anbietet, nicht von deren Bezug zum Sprecher. Im Prinzip blieb die Gleichsetzung der Bedeutung mit einem psychischen Zustand damit erhalten und dies bedeutete für die Sprachphilosophie vor wie nach Frege natürlich auch, sich auf die internalistische Annahme einzulassen, daß Bedeutungen in Form der kognitiven Signifikanz von Ausdrücken etwas seien, das jeder Sprecher in seinem Kopf mit sich herumtrage. An dieser Auffassung hat sich bisher nicht wirklich etwas geändert, obwohl wir in Kapitel II gesehen haben, daß die Route zu einer Kritik des Internalismus von Kripke

---

<sup>3</sup> „Bedeutung“ wird hier und auch im folgenden natürlich nicht mehr im Fregeschen Sinne gebraucht, sondern im Sinne des englischen *meaning*. Freges Gebrauch von „Bedeutung“ ist, wie bereits erwähnt, nicht deckungsgleich mit dem intuitiven Gebrauch und hat sich auch spätestens seit Carnap nicht mehr durchsetzen können.

bereits angegeben, wenn auch nicht eingeschlagen wurde.

Die erste philosophische Position, die die unbewußte These des Internalismus wirksam angegriffen hat, ist jene, die von Hilary Putnam, Tyler Burge und Stephen Stich in den Siebziger Jahren entwickelt wurde und die man gemeinhin als semantischen Externalismus bezeichnet. (vgl. dazu [Put75], [Bu79] und [Sti78]. Stich wird hier nicht besprochen, da sein Zugang zur externalistischen Argumentation sich stärker auf die psychologische als die semantische Seite des Problems bezieht. Nichtsdestoweniger muß er im gleichen Atemzug mit Putnam und Burge genannt werden.)

Der locus classicus des semantischen Externalismus ist Putnams populärer Aufsatz „The Meaning of ‚Meaning““ (= [Put75]). Für Putnam ist dieser Text ein wichtiger Baustein in seiner Auseinandersetzung mit dem logischen Empirismus, dessen eminent einflußreiche Position in der Erkenntnistheorie er auf diese Weise als unverdient zu entlarven versucht<sup>4</sup>. Putnams Ausführungen sind jedoch ebenso gegen die oben besprochene Position des Funktionalismus in all ihren Spielarten gerichtet.

## IV.2 Die zwei Grundlagen des internalistischen Bedeutungsbegriffs

Ging es in der Bedeutungsdiskussion von Frege bis zu Donnellan im Prinzip um ein semantisches Modell innerhalb einer als unproblematisch akzeptierten philosophischen Theorie, so formuliert Putnam, indem er an Kripke anschließt, nun auf sehr fundamentaler Ebene eine These gegen die bisherige Bedeutungstheorie *in toto*. Die Widerlegung des internalistischen Bedeutungsbegriffs geschieht dabei auf einem sehr eingängigen und beeindruckend eleganten Weg. Es lohnt sich, ihn hier von Anfang an zu verfolgen.

Putnam identifiziert die erkenntnistheoretische Position des logischen Empirismus vor allem mit zwei Prämissen:

**P1)** „That knowing the meaning of a term is just a matter of being in a certain psychological state (in the sense of ‘psychological state’, in which states of memory and psychological dispositions are ‘psychological states’; no one thought that knowing the meaning of a word was a continuous state of consciousness, of course).“ [[Put75] p. 219]

P1 ist eine kondensierte Darstellung der in Abschnitt 1 dieses Kapitels dargelegten internalistischen Tradition der Bedeutungstheorie. Sie stellt eine weitgehend eigenständige Weiterentwicklung der Fregeschen Annahmen aus I.2-4 dar und ist uns – in weniger scharfer Form – bereits in I.4 begegnet, auch wenn Frege selbst diese Theorie natürlich nicht so formuliert.

**P2)** „That the meaning of a term (in the sense of ‘intension’) determines its extension (in the sense that sameness of intension entails sameness of extension).“ [[Put75] p. 219]

---

<sup>4</sup>cf. [Put75] pp. 216-219. Im Abschnitt „Meaning and Extension“ findet sich ein Überblick über das, was Putnam am Empirismus zu beseitigen gedenkt.

Die Bedeutung eines Ausdrucks, hier im Sinne einer Carnapschen Intension zu verstehen, bestimmt seine Extension. Das heißt, aus Gleichheit der Intension folgt Gleichheit der Extension. Das haben wir schon in Kapitel I bei der Besprechung der Positionen Freges und Carnaps gesehen.

Die Pointe ist nun, daß keine Theorie formuliert werden kann (oder muß) in der diese beiden Prämissen wahr sind.

### IV.3 Wie die internalistische Bedeutungstheorie entwickelt wird

Es seien  $\alpha$  und  $\beta$  zwei Ausdrücke mit verschiedener Extension<sup>5</sup>.

Gemäß P1 sind dann „die Bedeutung von  $\alpha$  kennen“ und „die Bedeutung von  $\alpha$  kennen“ psychische Zustände (im engen Sinne). Gemäß P2 folgt, daß aufgrund der verschiedenen Extension von  $\alpha$  und  $\beta$  auch deren Intension (also: ihre „Bedeutung“) verschieden sein muß (vgl. p. 221).

Die psychischen Zustände, als die wir die Kenntnis der Bedeutungen (Intensionen) von  $\alpha$  und  $\beta$  beschreiben, müssen aber, sofern sie im Kopf des Sprechers diesen Intensionen tatsächlich entsprechen, genauso wie die Intensionen die Extensionen von  $\alpha$  und  $\beta$  bestimmen<sup>6</sup>. Das heißt, aus der Gleichheit der psychischen Zustände folgt die Gleichheit der Intensionen und kraft P2 die Gleichheit der Extensionen. Mit anderen Worten: „die Bedeutung von  $\alpha$  kennen“ kann nicht derselbe Zustand sein wie „die Bedeutung von  $\beta$  kennen“. (Das wird schnell klar, wenn man sich verdeutlicht, daß  $\alpha$  und  $\beta$  mit Donnellan attributiv gelesen werden können: die Bedeutung von  $\alpha$  kennen, heißt natürlich auch, zu wissen, daß es die Bedeutung von  $\alpha$  ist. Wenn  $\alpha$  und  $\beta$  also verschiedene Ausdrücke sind, dann ist „die Bedeutung von  $\alpha$  kennen“ selbst dann etwas anderes als „die Bedeutung von  $\beta$  kennen“, wenn  $\alpha$  und  $\beta$  dieselbe Extension haben. Es ist – in attributiver Lesart – einfach deshalb etwas anderes, weil  $\alpha$  und  $\beta$  verschiedene Ausdrücke sind, und zu wissen, daß  $I$  die Intension von  $\alpha$  ist, ist etwas anderes als zu wissen, daß  $I$  die Intension von  $\beta$  ist.)

Es kann aufgrund dieser Folgerung also keine zwei verschiedenen, logisch möglichen Welten geben derart, daß ein gegebener Sprecher in  $W_1$  den Ausdruck  $\alpha$  mit der Intension  $I_1$  und in  $W_2$  mit der Intension  $I_2$  versteht und sich dennoch in  $W_1$  und in  $W_2$  im selben psychischen Zustand befindet.

Vereinfacht ausgedrückt: zwei Sprecher können sich hinsichtlich eines Ausdrucks  $\alpha$  nicht im selben psychischen Zustand befinden, sofern die Extensionen von  $\alpha$  nicht in den Idiolekten beider Sprecher identisch sind. Und genau das ist nach Putnam falsch.

---

<sup>5</sup>Putnam verwendet die Begriffe „Intension“ und „Extension“ nicht in der scharfen Kontur, in der sie in Kapitel 1 herausgearbeitet wurden, sondern in einem recht intuitiven Sinn; vgl. hierzu seine mehr als laxer Darstellung auf loc.cit. p. 217. Seine Theorie läßt jedoch auch den präzisierten Gebrauch der Begriffe zu, den wir uns in Kapitel 1 angeeignet haben; daher werde ich von diesem Gebrauch nicht absehen.

<sup>6</sup>Man beachte: eine Intension ist eine Funktion, ein logisches Konstrukt. Die Kenntnis der Intension dagegen ist ein psychischer Zustand eines Individuums. Diese beiden Begriffe sind miteinander korreliert, doch nicht identisch. Kraft dieser Korrelation beeinflusst der psychische Zustand die Extension genau dann, wenn die Intension die Extension beeinflusst.

## IV.4 Erster Schritt der Widerlegung: Bedeutungen sind nicht im Kopf

Zwei Sprecher können sich, so behauptet Putnam, hinsichtlich eines Ausdrucks  $\alpha$  eben doch im selben psychischen Zustand befinden, obwohl sich die Extension von  $\alpha$  im Idiolekt des einen von der Extension von  $\alpha$  im Idiolekt des anderen unterscheidet. Die unmittelbare Folgerung aus dieser Annahme wäre, daß entgegen der klassisch-internalistischen Annahme die Extension eines Ausdrucks nicht vom psychischen Zustand festgelegt wird, der mit dessen Intension korreliert. Bedeutungen wären demnach nicht „im Kopf“.

Einmal angenommen, man akzeptierte das, so gäbe es zwei Möglichkeiten:

- 1.) Der psychische Zustand des Sprechers bestimmt nicht die Bedeutung (was heißt, daß P1 zu verwerfen sei); oder
- 2.) die Intension bestimmt nicht die Extension (was heißt, daß P2 zu verwerfen sei).

Gelänge also der Beweis, daß Bedeutungen nicht im Kopf sind, wäre damit gezeigt, daß unmöglich beide der Annahmen, mit denen Putnam die internalistische Bedeutungstheorie identifiziert, wahr sein können. Mindestens eine müßte verworfen werden. Putnam entscheidet sich für P1.

Den Beweis dafür, daß psychische Zustände nicht die Extensionen von Ausdrücken bestimmen und Bedeutungen also nicht im Kopf sind, führt Putnam in einem sehr eigenwilligen Gedankenexperiment: Man stelle sich eine mögliche Welt vor, die der Erde in jeder nur denkbaren Hinsicht gleicht, sie hat physikalisch die gleiche Entwicklungsgeschichte hinter sich wie die Erde, wird von den gleichen Individuen bewohnt und ist mit ihr bis auf die molekulare Ebene hinab vollständig identisch. Putnam nennt diese Welt die Zwillingserde [*Twin Earth*]. Zwischen der Erde und der Zwillingserde gibt es nur einen einzigen, wichtigen Unterschied: auf der Zwillingserde existiert die Substanz  $H_2O$  nicht (die wir bekanntlich „Wasser“ nennen). Allerdings gibt es dort eine Substanz, die von den Bewohnern der Zwillingserde „Wasser“ genannt wird; sie sieht genauso aus, sie schmeckt und riecht genauso und hat exakt die gleichen Eigenschaften wie  $H_2O$ , obwohl sie nicht  $H_2O$  ist. Diese Substanz hat vielmehr eine so komplizierte chemische Zusammensetzung, daß wir sie hier der Einfachheit halber als XYZ abkürzen. Im folgenden verwende ich „Wasser“<sub>E</sub>‘, um das deutsche Wort „Wasser“ zu bezeichnen, „Wasser“<sub>TE</sub>‘ für das zwillingsdeutsche Homonym, „Wasser“<sub>E</sub>‘ um  $H_2O$  und „Wasser“<sub>TE</sub>‘ um XYZ zu bezeichnen.

Überall, wo also auf der Erde  $H_2O$  vorkommt, kommt auf der Zwillingserde XYZ vor, es füllt Seen, fällt vom Himmel, kommt aus dem Wasserhahn (der auch im Zwillingsdeutschen „Wasserhahn“ heißt) und man kocht damit Tee und Kaffee. Auf der Erde hat der Ausdruck „Wasser“ im Deutschen also die Extension  $H_2O$ , auf der Zwillingserde, die ich manchmal verkürzend „Zwerde“ nennen werde, hat der Ausdruck „Wasser“ (im Zwillingsdeutschen) dagegen die Extension XYZ.

Stellen wir uns nun vor, daß jemand von der Erde zur Zwillingserde reiste. Er wäre von der Ähnlichkeit der beiden Welten zweifelsohne sehr überrascht, doch noch mehr würde ihn die Existenz einer Substanz überraschen, die unserem Wasser<sub>E</sub> in praktisch allem gleicht, nur ganz anders aufgebaut ist.

Wie also verhält es sich angesichts dieses spektakulären Falls interstellarer Kommunikation mit den Bedeutungen des Wortes „Wasser“? Fürs erste ist der Fall unproblematisch: was die Bewohner der Zwerde „Wasser“ nennen, ist kein Wasser<sub>E</sub>. Im Deutschen ist die Extension des Wortes „Wasser“ H<sub>2</sub>O (=Wasser<sub>E</sub>), im Zwillingsdeutschen ist die Extension von „Wasser“ dagegen XYZ (=Wasser<sub>TE</sub>).

Um das Problem nun sichtbar zu machen, stelle man sich in einer zweiten Version des Szenarios vor, daß im Jahre 1750, noch bevor die molekulare Struktur von Wasser<sub>E</sub> bekannt war, ein Bewohner der Erde zur Zwillingserde reist. Er hat zwar gewisse Methoden zur Identifizierung von Wasser<sub>E</sub> anhand von Aussehen, Beschaffenheit, Geruch, Geschmack undsofort, die man toleranterweise als eine Art operationaler „Definition“ von Wasser<sub>E</sub> ansehen könnte, doch er kann jetzt, wo es darauf ankommt, nicht wirklich *beweisen*, ob das, was er auf der Zwillingserde zu sehen bekommt, Wasser<sub>E</sub> ist oder nicht. Er hat nicht einmal einen vernünftigen Grund, daran zu zweifeln.

Jetzt ist deutlich, worum es in dem Beispiel geht: Der Erdling und der Zwerdling können in der zweiten Version des Experiments nicht bemerken, daß die Extension von „Wasser“ im Deutschen und die Extension von „Wasser“ im Zwillingsdeutschen sich signifikant unterscheiden. Wie also sollte man behaupten können, daß die psychischen Zustände der beiden, in denen die jeweiligen Bedeutungen von „Wasser“ verkörpert sind, sich unterscheiden sollten? Der Erdling und der Zwerdling haben denselben „Begriff“ von Wasser und daß sie von unterschiedlichen Dingen reden, können sie erst feststellen, wenn die Wissenschaft Methoden zur sicheren Identifizierung von Wasser<sub>E</sub> beziehungsweise Wasser<sub>TE</sub> gefunden hat. Also stellt die Situation in dem Gedankenexperiment eine eklatante Verletzung der Schlußfolgerungen aus P1 und P2 dar: Offensichtlich können sich zwei Individuen (beispielsweise ich und mein zwirdischer Doppelgänger) in ein und demselben psychischen Zustand befinden und doch zwei Ausdrücke benutzen, deren Extensionen nicht einmal eine gemeinsame, nichtleere Schnittmenge haben.

Der Internalist, der sowohl P1 als auch P2 aufrecht erhalten möchte, kann *nicht erklären*, wie so etwas möglich ist.

Auch in [Bu79] findet sich ein ähnliches, doch weniger spektakuläres Gedankenexperiment. Es ist in zwei Stufen gegliedert:

- 1) Ein Patient *P* hat eine große Zahl von Überzeugungen bezüglich der Bedeutung des Wortes „Arthritis“ und die meisten Inhalte dieser Überzeugungen werden von den Sprechern seiner Sprachgemeinschaft geteilt: *P* denkt (korrekt), daß er seit Jahren Arthritis hat, daß die Arthritis in seinen Fingern schmerzhafter ist als die in seinen Fußknöcheln, daß Arthritis immer noch besser ist als Krebs, daß Gelenkschmerzen ein Symptom von Arthritis sind undsofort. Diese propositionalen Einstellungen sind auch bei anderen verbreitet und nicht überraschend. Zusätzlich denkt *P* jedoch irrtümlich, daß die Arthritis auch seinen Oberschenkel befallen habe. *P* besitzt souveräne Sprachkompetenz und ist normal intelligent. Er geht zu seinem Arzt und berichtet ihm seine Vermutung. Der Arzt sagt ihm, daß seine Vermutung falsch ist, weil Arthritis eine Krankheit sei, die ausschließlich Entzündungen der Gelenke hervorruft. *P* ist überrascht und fragt sich, was denn dann mit seinem Oberschenkel los sein könnte.
- 2) Der zweite Schritt besteht in einer kontrafaktischen Annahme: Wir stellen uns eine Si-

tuation vor, in der der Patient bis zu dem Zeitpunkt, da er die Praxis betritt und seinen Verdacht äußert, exakt die gleiche Biographie hat und in allen raumzeitlichen und nichtintentionalen<sup>7</sup> (!) mentalen Gegebenheiten identisch mit dem Patienten in der ersten Version ist. Demnach ist seine Überzeugung „Ich habe Arthritis im Oberschenkel“ das Resultat der gleichen Abläufe und Ereignisse. Generell unterscheidet sich die kontrafaktische Situation von der Situation in 1) nur durch ein einziges Merkmal: „Arthritis“ ist in der Sprachgemeinschaft, die wir uns nun vorstellen, ein sehr weitgefaßter Ausdruck, er bezeichnet alle möglichen Arten rheumatoider Erkrankungen, nicht ausschließlich Gelenkentzündungen. Auch Ärzte und Fachleute gebrauchen dieses Wort so. Demzufolge schließt dieser kontrafaktische Gebrauch von „Arthritis“ die Erkrankung im Oberschenkel des Patienten ein und sein Gebrauch von „Arthritis“ ist nicht abweichend. Ansonsten besteht zwischen den Sprachgemeinschaften in 1) und 2) kein Unterschied.

Was zeigt das? In 2) hat der Patient keine der Überzeugungen bezüglich Arthritis, die er in 1) hat. Das kommt daher, daß das, was er in 2) als „Arthritis“ bezeichnet, nicht Arthritis ist, denn „Arthritis“ ist ein starrer Designator. Zum Beispiel wird der Patient *nicht* glauben, daß er Arthritis im Oberschenkel habe, daß er seit vielen Jahren Arthritis hat, daß Gelenkschmerzen Symptome von Arthritis sind und sofort, denn alle diese Überzeugungen beziehen sich auf eine Krankheit, die nicht Arthritis ist, obwohl sie in 2) so *genannt* wird. Daher unterscheiden sich die propositionalen Einstellungen von *P* in 2) von denen, die *P* in 1) hat. In keinem Fall kann man in 2) das Vorkommen des Wortes „Arthritis“ in einem Überzeugungssatz von *P* durch einen koreferentiellen Ausdruck ersetzen<sup>8</sup>:

„However we describe the patient's attitudes in the counterfactual situation, it will not be with a term or phrase extensionally equivalent with ‚arthritis‘. So the patient's counterfactual attitude contents differ from his actual ones.“ [[Bu79] p. 79]

Genaugenommen können wir *keine* der Einstellungen, die *P* in 2) zugeschrieben werden, korrekt wiedergeben, indem wir das Wort „Arthritis“ oder ein koreferentielles Wort gebrauchen, denn woran *P* in 2) denkt, ist nicht Arthritis, weil diese Krankheit nicht in die Extension von „Arthritis“ fällt. (Es ist erst seit Kripke möglich, so zu argumentieren, denn natürlich ist „Arthritis“ starr.)

Das Resultat des Arthritis-Experiments ist daher folgendes:

„The upshot of these reflections is that the patient's mental contents differ while his entire physical and non-intentional mental histories, considered in isolation from their

<sup>7</sup>Das ist wichtig, denn seitens des Patienten besteht die Intention, sich mit „Arthritis“ auf eine bestimmte Krankheitserscheinung zu beziehen. Wenn wir also bei einer Lesart bleiben, die Intentionalität berücksichtigt, dann müssen wir intentionale mentale Zustände von der Gleichheitsklausel ausklammern.

<sup>8</sup>Burge nennt die Erscheinung eines Ausdrucks  $\alpha$  in einem Satz „oblique occurrence“, wenn  $\alpha$  nicht *salva veritate* gegen einen koreferentiellen Ausdruck ausgetauscht werden kann: „When an expression like ‚water‘ functions in a content clause so that it is not freely exchangeable with all extensionally equivalent expressions, we shall say that it has oblique occurrence.“ (loc.cit. p.76) Das entspricht Freges „ungeradem Gebrauch“ und der Satz dem, was wir als intensionalen Kontext bezeichnet haben. In 2) können wir *P* keine propositionale Einstellung korrekt zuschreiben, die eine solche „ungerade“ Erscheinung von „Arthritis“ enthält, in 1) schon.

social context, remain the same.“ [[Bu79] p. 79]<sup>9</sup>

Wir können, was die gedanklichen Inhalte des Patienten angeht, auf jeden Fall feststellen, daß sie sich hinsichtlich des Wortes „Arthritis“ von 1) zu 2) beträchtlich unterscheiden, obwohl das Individuum in 2) ansonsten mit dem in 1) identisch in Bezug auf alle nichtintentionalen mentalen Eigenschaften ist. Offensichtlich bestimmt auch in diesem Fall nicht allein der psychische Zustand des Individuums die Extension des Ausdrucks.

„The thought experiment indicates that certain „linguistic truths“ that have often been held to be indubitable can be thought yet doubted. And it shows that a person's thought content is not fixed by what goes on in him, or by what is accessible to him simply by careful reflection. The reason for this last point about „accessibility“ need not be that the content lies too deep in the unconscious recesses of the subject's psyche. Contents are sometimes „inaccessible“ to introspection simply because much mentalistic attribution does not presuppose that the subject has fully mastered the content of his thought.“ [[Bu79] p. 105]

Die Überzeugungsinhalte des Patienten in 1) unterscheiden sich nämlich von denen in 2) nur aufgrund der unterschiedlichen Sprachpraxis innerhalb der Sprachgemeinschaft. Die Extension des Ausdrucks „Arthritis“ wird von den mentalen Konfigurationen von *P* jedoch nicht beeinflusst. (Burge folgert außerdem, wie im vorangegangenen Zitat angedeutet, daß Sprecher mit Überzeugungsinhalten auch dann mit einer gewissen Kompetenz umgehen können, wenn sie diese nicht vollkommen verstehen. Dadurch wird das ganze kartesiansche Programm der Introspektion über den Haufen geworfen, weil ein Sprecher einen nicht richtig verstandenen Ausdruck auch durch noch so sorgfältige Introspektion nicht verstehen kann. Er müßte einfach jemanden *fragen*.)

Fassen wir also Schritt 1 des externalistischen Programms zusammen:

Zwei Sprecher können sich hinsichtlich eines Ausdrucks  $\alpha$  im selben psychischen Zustand (im engen Sinne) befinden, obwohl sich die Extension von  $\alpha$  im Idiolekt des einen von der Extension von  $\alpha$  im Idiolekt des anderen unterscheidet. Die unmittelbare Folgerung aus dieser Annahme ist, daß die Extension eines Ausdrucks nicht von einem psychischen Zustand festgelegt wird, der mit dessen Bedeutung korreliert.

Wie schon angedeutet, heißt das, daß P1 und P2 nicht beide wahr sein können. Entweder also bestimmt der psychische Zustand des Sprechers nicht die Bedeutung oder die Bedeutung bestimmt nicht die Extension.

---

<sup>9</sup>Der Ausdruck „content“, der uns bald häufiger begegnen wird, heißt hier einfach „Inhalt“. Gemeint sind die Inhalte sogenannter „content clauses“, die Burge wie folgt definiert: „In an ordinary sense, the noun phrases that embed sentential expressions in mentalistic idioms provide the content of the mental state or event. We shall call that-clauses and their grammatical variants „content-clauses“.“ [[Bu79] p. 74] Burge meint offensichtlich jene Sätze, die propositionale Einstellungen verbalisieren.



## IV.5 Zweiter Schritt: Bedeutung, Realismus und versteckte Indexikalität

Das Twin-Earth-Experiment funktioniert mit dem natürlichen Prädikat „ist Wasser“ und es ließe sich leicht für beliebige andere natürliche Prädikate umgestalten. Welche Auswirkungen aber haben nun die überraschenden Ergebnisse des Experiments bezüglich der Bedeutung solcher Prädikate?

Putnam erläutert diese Auswirkungen anhand des Beispiels des Versuchs einer Ostensivdefinition<sup>10</sup>: nehmen wir an, ich auf der Erde und mein Zwilling auf der Zwillingserde zeigen auf ein Glas ‚Wasser‘ und sagen „Die Flüssigkeit in diesem Glas ist Wasser.“ Natürlich ist die Flüssigkeit in dem irdischen Glas H<sub>2</sub>O und die in dem zwirdischen Glas XYZ. Es gibt dann zwei Möglichkeiten, die Bedeutung des Ausdrucks „Wasser“ zu beurteilen:

- a) Das Wort „Wasser“ hat eine konstante relative Bedeutung, das heißt es bezeichnet konstant überall dasselbe Phänomen, seine Extension ist jedoch relativ zum Kontext bestimmt. Der Ausdruck „Wasser“ bedeutet daher in beiden Welten dasselbe, nur hat Wasser auf der Erde die molekulare Struktur H<sub>2</sub>O und auf der Zwillingserde die Struktur XYZ.
- b) Wasser ist in jeder Welt H<sub>2</sub>O und „Wasser“ hat daher in beiden Welten verschiedene Bedeutung. In jeder anderen möglichen Welt ist eine Flüssigkeit nur dann Wasser, wenn sie mit dem, was wir in unserer Welt „Wasser“ nennen (=Wasser<sub>E</sub>), molekular identisch ist.

Nach Putnam ist eindeutig b) die richtige Sichtweise. Sie kann als das zentrale Element der Auffassung bezeichnet werden, die er selbst „Realismus“ nennt: die Bedeutungen der Ausdrücke hängen nicht von anderen möglichen Welten, sondern von unserer Welt ab. Wie aber kommt er dazu?

An dieser Stelle nun zeigt sich der Zusammenhang zwischen Putnams Überlegungen und denen Kripkes. Erstens übernimmt und ergänzt Putnam Kripkes Auffassung von der Starrheit von Designatoren. Die Starrheit des Wortes „Wasser“<sub>E</sub> – und das ist das zentrale Argument für b) – folgt aus der Tatsache daß ich als Sprecher, wenn ich auf ein Glas zeige und sage „diese Flüssigkeit“, eben nicht eine konstante relative Bedeutung ausdrücke wie in a), sondern eben gerade *dieses* Glas Wasser<sub>E</sub> vor mir in *dieser* meiner Welt meine, also der Logik von b) und nicht der von a) folge. Mit Kripke schließt Putnam, daß es einfach keinen Fall geben kann, in dem Wasser<sub>E</sub> nicht Wasser<sub>E</sub> ist, weshalb es auch keine mögliche Welt geben kann, in der Wasser nicht H<sub>2</sub>O ist.

Zweitens schließt Putnam seine Überlegungen an Kripkes Thesen zur Differenzierung von Notwendigkeit und Apriorität an. Eine Aussage kann metaphysisch notwendig und dennoch epistemisch kontingent, sprich: a posteriori erkennbar sein. Genauso verhält es sich nämlich, wenn auf der Erde und der Zwerde gegen Ende des 18. Jahrhunderts plötzlich entdeckt wird, daß „Wasser“ auf der Erde und der Zwerde unterschiedliche Bedeutung hat. Da

<sup>10</sup>Wie Putnam an anderer Stelle ausführt, sind Ostensionen als Bedeutungserklärungen unzulänglich. Für das folgende Beispiel ist auch nur die Geste des Zeigens von besonderer Wichtigkeit.

„Wasser“<sub>E</sub> ein starrer Designator ist, ist XYZ notwendig kein Wasser<sub>E</sub>. Erkennen konnten wir das aber nicht a priori.

Drittens übernimmt Putnam von Kripke implizit die Auffassung, daß mögliche Welten nur als Konzept eine Berechtigung haben und nicht als qualitativ oder gar als tatsächlich existierende, mit der wirklichen Welt vergleichbare Entitäten aufgefaßt werden sollten, wie zum Beispiel Lewis in [Lew88] behauptet. Würden wir mögliche Welten qualitativ auffassen, so wäre unklar, wieso wir b) der Annahme a) vorzuziehen hätten.

Putnam ergänzt außerdem das Denken über das Verhältnis zwischen Indexikalität und Starrheit um die These von der *versteckten Indexikalität*.

Zwar ist die offenkundige Indexikalität bei Wörtern wie „ich“, „hier“, „gestern“ und sogar „Bundeskanzler“ oder „Miss World“ seit längerem bekannt. Doch es ist niemand auf die Idee gekommen, den Geltungsbereich von P2, Intension bestimme Extension, auf solche Wörter auszudehnen<sup>11</sup>. Beispielsweise legt mein Zwilling auf der Zwillingserde wenn er „ich“ sagt, denselben Begriff von sich zugrunde wie ich von mir, wenn ich „ich“ sage. Jedenfalls ist nicht ersichtlich, weshalb diese Begriffe sich bei ihm und mir unterscheiden sollten. Dagegen unterscheiden sich die Extensionen dieser Begriffe beträchtlich, denn die Extension von „ich“ in seinem Gebrauch ist er und die Extension von „ich“ in meinem Gebrauch bin ich. Für offenkundig indexikalische Wörter stimmt Annahme P2 also nicht. (Das ist natürlich noch kein Grund, sie aufzugeben: wie Kaplan gezeigt hat, kann man entgegen Putnams strenger Bemerkung P2 auch in einer Logik der Indexikalen relativ problemlos unterbringen, wenn man in P2 „Intension“ durch „Bedeutung“ ersetzt und diese im Satzzusammenhang von P2 als geordnetes Paar aus Charakter und Intension auffaßt.)

Für natürliche Prädikate wie „ist Wasser“ nun läßt sich folgendes feststellen: nämlich daß, wie oben in Anschluß an Kripke bereits gezeigt, etwas, um Wasser<sub>E</sub> zu sein, in einer bestimmten Relation (Identität der molekularen Struktur) zu unserem Wasser<sub>E</sub> hier in der wirklichen Welt stehen muß, denn alles andere ist eben kein Wasser<sub>E</sub>. In dieser Überlegung spielt der Begriff der *Querweltrelation* eine tragende Rolle: die Beziehung zwischen Wasserkandidaten in etwaigen Welten  $W_x$  und unserer Welt kann, folgt man Putnam in diesem Punkt, nur als Querweltrelation dargestellt werden, das heißt, als eine Relation, deren Argumente Teil verschiedener möglicher Welten sind. Ob etwas zur Extension von „Wasser“<sub>E</sub> gehört, bestimmen wir, indem wir den Wahrheitswert des Querweltsatzes für Wasser<sub>E</sub> ablesen: „Ist  $x$  in  $W_x$  mit Wasser<sub>E</sub> in unserer Welt molekular identisch?“ Falls es so ist, gehört  $x$  in die Extension von „Wasser“<sub>E</sub>, falls nicht, gehört  $x$ , völlig egal wer wo und wann welchen Begriff von  $x$  hat, nicht zur Extension von „Wasser“<sub>E</sub>. Insofern sind auch natürliche Prädikate *versteckt indexikalisch*.

Folgendes wird dabei deutlich: wenn wir sagen, daß Wasser<sub>TE</sub>, das zwar Wasser<sub>E</sub> ober-

---

<sup>11</sup>So, wie ich die Theoriegeschichte hier darstelle, ist natürlich längst jemand auf diesen Gedanken gekommen, nämlich Kaplan. Die Erklärung: Kaplans Text kursierte seit 1977 als Manuskript und war in der Fachwelt weithin bekannt. Veröffentlicht wurde er jedoch erst in [Kap89a], zu einem Zeitpunkt, als der Charakterbegriff bereits in weitreichende Orthodoxie gemündet hatte. Als Putnam dieses Argument entwickelte, konnte er jedoch noch nicht auf Kaplan zurückgreifen. Ich habe Kaplan vor Putnam besprochen, weil Kaplan im Rahmen dieser Arbeit für die an Putnam anschließende Diskussion um den engen Inhalt und die Weiterentwicklung Freges wichtig ist und ich Kaplan nicht zwischen Putnam und das Kapitel über engen Inhalt klemmen wollte, was der philosophischen Bedeutung Kaplans nicht gerecht würde.

flächlich ähnlich ist, kein Wasser<sub>E</sub> ist, obwohl der Begriff, den die Zwerdlinge von Wasser<sub>TE</sub> haben, sich nicht von dem Begriff, den wir von Wasser<sub>E</sub> haben, unterscheidet, dann unterstellen wir natürlich, daß ein Extensionsunterschied einen Bedeutungsunterschied ausmacht – genau wie bei Kaplanschen Indexikalen<sup>12</sup>. Wenn die Bedeutungen des deutschen „Wasser“ und des zwillingsdeutschen „Wasser“ sich aber infolgedessen unterscheiden und die geistigen Begriffe, die man jeweils davon hat, nicht – dann können Bedeutungen keine geistigen Begriffe sein. Basta.

Putnam räumt allerdings selbst ein, daß eine andere Lesart möglich wäre, nämlich jene, zu sagen, „Wasser“ habe zwar in beiden Welten dieselbe Bedeutung, doch eben unterschiedliche Extension (Lesart a). In dieser Art würde man nämlich offenkundig indexikalische Begriffe wie „ich“ behandeln. Dann müßte man allerdings die Behauptung aufgeben, daß die Intension die Extension bestimme und das will Putnam nachgerade nicht, jedenfalls nicht endgültig und universell. (Bei Ausdrücken mit direkter Referenz gilt P2 ohnehin nicht, wie wir in Kapitel III gesehen haben. Das heißt nicht, daß P2 völlig verfehlt gewesen wäre.)

## IV.6 Dritter Schritt: die Sprachgemeinschaft und die Welt bestimmen die Extension

Nach all diesen Überlegungen sehen wir, daß P1 auf jeden Fall zu verwerfen ist, P2 jedoch mindestens einen Teil ihrer Gültigkeit behält.

Doch wie wird nun die Extension tatsächlich bestimmt, wenn nicht allein durch eine Intensionsfunktion? Für Putnam sind zur Bestimmung der Extension vor allem zwei Faktoren wichtig – gerade jene, die die bisherige Bedeutungstheorie bis auf Kripke nie zur Kenntnis genommen hat.

Einen Faktor haben wir in diesem Kapitel bereits kennengelernt: die wirkliche Welt. Was zur Extension eines Ausdrucks gehört und was nicht, das hat mit der Beschaffenheit der Welt zu tun, wie das Querweltbeispiel zeigt: etwas, das wir mit dem Namen *N* benennen wollen, muß der Entität, die wir starr mit *N* bezeichnen, hinreichend ähnlich sein, um in die Extension von *N* fallen zu können. Beispielsweise haben wir für alles, was „Wasser“<sub>E</sub> genannt werden soll, gefunden, daß es mit unserem Wasser<sub>E</sub> molekular identisch sein muß, um Wasser<sub>E</sub> zu sein. Das jedoch hängt von den Merkmalen ab, die wir bezüglich des Wassers herausbekommen haben. Dinge, die so sind, bezeichnen wir starr als „Wasser“<sub>E</sub> und mit Begriffen im Kopf hat das nichts zu tun. Die eigentliche Bezeichnungsfunktion ist starr; nicht nur sie, sondern auch die Beschaffenheit der wirklichen Welt ist für die Zugehörigkeit zu einer Extension entscheidend. Die Kenntnis der Beschaffenheit der Welt hängt dabei jedoch vom Fortschritt der Wissenschaft ab: wir können durchaus entdecken, daß das, was wir lange Zeit als „Wasser“<sub>E</sub> bezeichnet haben, unter Umständen in weitere natürliche Sorten zerfällt. Die Entdeckung solcher Ähnlichkeiten und Unterschiede, die es ermöglicht, die Querweltsätze in dieser Form zu formulieren und zu präzisieren, ist die Aufgabe der Wissenschaft.

---

<sup>12</sup>Wir müssen daran denken, daß Kaplan sein Manuskript zwei Jahre nach Putnams Aufsatz veröffentlicht hat und die Überlegung, daß in manchen Fällen die Extension die Intension bestimmt, in Kapitel III nicht als Gedankenwende inszeniert wurde, weil Putnam hier der erste war.

Der andere Faktor sind die anderen Sprecher. Die Experimente aus Abschnitt 4 zeigen nämlich, daß der psychische Zustand, oder, wie Burge ihn nennt, der gedankliche Inhalt, nicht ausschließlich von den mentalen Phänomenen des Sprechers geprägt wird, sondern von der Sprachgemeinschaft. Burge betont an mehreren Stellen:

„Even apart from reversals of the thought experiment, it is plausible [. . .] that our well-understood propositional attitudes depend partly for their content on social factors independent of the individual, asocially and non-intentionally construed.“ [[Bu79] p. 85]

„The claim that is crucial is [. . .], rather, that the social differences between the actual and counterfactual situations affect the content of the subject's attitudes.“ [[Bu79] p. 86/87]

Das unvollständige Verstehen gewisser Wörter ist für viele Sprecher einer Sprachgemeinschaft, so Burge, nicht nur keine abweichende oder ungenügende Kompetenz, sondern vollkommen normal. Wörter wie „Kontrakt“, „Sonate“, „Rezession“, „Vorverstärker“, „gotisch“ oder „Fermentation“ müssen nicht von allen Sprechern einer Sprachgemeinschaft vollkommen korrekt benutzt werden können. Innerhalb einer Sprachgemeinschaft ist es also zulässig, daß gewisse Wörter nur von bestimmten Experten in allen Fällen korrekt benutzt werden, der normale Sprecher muß deren Wissen nicht haben.

Die Verwendung solcher Wörter (und es sind nicht wenige) unterliegt, so Putnams These, innerhalb eines Sprachkollektivs dem Phänomen der „sprachlichen Arbeitsteilung“. In jeder Sprachgemeinschaft gibt es Wörter, deren Extension nur von bestimmten Experten sicher bestimmt werden kann<sup>13</sup>, die aber nichtsdestoweniger im Sprachcode des als normal angesehenen Wortschatzes enthalten sind und demzufolge auch von der Mehrzahl der Sprecher verwendet werden, die keine Experten sind. Der kollektive Gebrauch solcher Wörter hängt von der gesellschaftlichen Kooperation der expertischen und der nichtexpertischen Sprecher ab. Ein einleuchtendes Beispiel hierfür ist zum Beispiel das Wort „Gold“. Wir alle wissen, was Gold „ist“, aber verifizieren, ob ein gegebenes Stück Material wirklich in die Extension von „Gold“ fällt, können allenfalls Leute, die zumindest über ein gewisses „Expertenwissen“ verfügen (und selbst dann kann sich die zugrundeliegende Definition noch als unzureichend herausstellen). Ähnliche Beispiele lassen sich im Bereich der natürlichen Prädikate in Hülle und Fülle finden (vgl. „Gorilla“, „Kohlendioxid“, „Diamant“ und sofort, aber auch „Jupiter“). Natürlich gibt es auch Wörter, bei denen keine sprachliche Arbeitsteilung besteht („Stuhl“), doch das Phänomen dieser Arbeitsteilung enthebt den normalen Sprecher der Pflicht, in allem Experte zu sein. Bei Begriffen, die der sprachlichen Arbeitsteilung unterliegen, genügt es, wenn das sprachliche Kollektiv als Ganzes sie korrekt verwendet, der einzelne Sprecher muß dies nicht können. Statt eines differenzierten Charakters lernt er ein Stereotyp, das er

<sup>13</sup>1.) Hierbei spielen natürlich Thesen über die Wissenschaft als kulminativen Prozeß eine Rolle, die hier nicht besprochen werden können: Wasserexperten gibt es natürlich erst, wenn die Wissenschaft soweit entwickelt ist, daß sie Expertenwissen hervorbringt. Vor 1750 waren in diesem Sinne wohl fast alle Leute Wasserexperten. 2.) Auch Experten können in die Situation kommen, festzustellen, daß ihre vermeintlich genaue Feststellung immer noch mehrere Entitäten bezeichnet, die man zu unterscheiden lernen kann, sobald die wissenschaftlichen Mittel dafür zur Verfügung stehen. Auch hier greift das Beispiel H<sub>2</sub>O vs. XYZ.

dann verwendet und bei Bedarf auch an andere Sprecher weitergibt. (Ein „Tiger“-Stereotyp zum Beispiel wäre: ‚große, schwarz-gelb gestreifte Raubkatze‘.) Von der Rolle der Stereotypen wird noch die Rede sein – zunächst ist folgendes festzuhalten: aufgrund der sprachlichen Arbeitsteilung in einer Sprachgemeinschaft lernt der einzelne Sprecher, wenn er einen Ausdruck zu gebrauchen lernt, der der sprachlichen Arbeitsteilung unterliegt, nichts, was dessen Extension festlegt, denn ein Stereotyp allein reicht dazu natürlich nicht aus (sonst wäre das ganze Twin-Earth-Experiment hinfällig). Infolgedessen aber kann man natürlich auch nicht länger behaupten, daß der psychische Zustand des Sprechers die Extension festlege. Denn auf welcher Basis sollte das geschehen, wenn keine „Bedeutung“ im eigentlichen Sinne gelernt wird, da diese der Arbeitsteilung unterliegt und nur den Experten zugänglich ist?

Dies zeigt aber, daß die soziale Umgebung des Sprechers die Extension eines Ausdrucks offensichtlich ebenso festlegt, wie die Beschaffenheit der Welt selbst dies tut. Burge bemerkt:

„The intentional or semantic role of mental states and events is not a function merely of their functionally specified roles in the individual. The failure of these accounts of intentional mental states and events derives from an underestimation of socially dependent features of cognitive phenomena.“ [[Bu79] p. 107/108]

Wieder einmal sehen wir uns also der Notwendigkeit gegenüber, einzugestehen, daß Bedeutungen nicht im Kopf sind. Doch wenn es nun nicht die Bedeutungen sind, die im Kopf sind, was ist dann im Kopf?

#### **IV.7 Vierter Schritt: die untergeordnete Rolle der individuellen Sprachkompetenz**

Auch wenn die Extension eines Ausdrucks nicht wirklich vom psychischen Zustand des Sprechers festgelegt wird, so muß dennoch die individuelle Sprachkompetenz irgendeine Rolle spielen, wenn es um die Verwendung von Begriffen geht. Jemandem, der einen Begriff wie „Tiger“ in einigen wenigen Fällen korrekt verwenden kann („Tiger haben Masse“), ohne aber etwas über Tiger zu wissen (zum Beispiel daß sie gestreifte Raubkatzen sind), würden wir absprechen, die Bedeutung des Wortes „Tiger“ zu kennen. Irgendetwas muß also im Kopf sein.

Laut Putnam kennt jemand die Bedeutung des Wortes „Tiger“, wenn das, was er unter einem „Tiger“ versteht, wie folgt korrekt festgelegt ist, also wenn

- 1) der Gebrauch, den der Sprecher von „Tiger“ macht, vom Kollektiv anerkannt wird, und
- 2) wenn seine Stellung in der Welt und seine Sprachgemeinschaft derart beschaffen sind, daß die Extension des Wortes „Tiger“ in seinem Sprachgebrauch die Menge aller Tiger ist. (Beispielsweise haben die Zwerdlinge das deutsche Wort „Wasser“ nicht erworben.)

Ob von einem Sprecher gesagt werden kann, daß er ein bestimmtes Wort korrekt erworben habe, hängt aufgrund von 1) also auch davon ab, ob er ein Mindestmaß an Kompetenz

im Umgang mit diesem Wort erkennen läßt. Dieses Mindestmaß an Kompetenz ist sozial bestimmt und kann als das jeweils zu diesem Wort gehörende Stereotyp beschrieben werden. Ein Stereotyp eines *X*

„[...] is a conventional (frequently malicious) idea (which may be wildly inaccurate) of what an *X* looks like or acts like or is.“ [[Put75] p. 249]

Nach dieser Definition wird also gemäß Bedingung 1) von jemandem zwar nicht verlangt zu wissen, wie die wissenschaftliche Definition eines Zoologen für Tiger aussehen würde, aber es wird verlangt zu wissen, daß stereotypische Tiger gestreift sind. Das heißt, es gibt ein Tiger-Stereotyp, dessen Kenntnis die Sprachgemeinschaft als Kompetenzminimum voraussetzt und ohne dessen Kenntnis man nicht als jemand gilt, der das Wort „Tiger“ erworben hat.

Dabei muß das Stereotyp eines Begriffs paradoxerweise keineswegs repräsentativ oder neutral, nicht einmal zutreffend sein. Zum Beispiel heißt es immer noch, die Erde „kreise“ um die Sonne: seit Kepler wissen wir, daß dies eine inkorrekte Wiedergabe der geometrischen Verhältnisse ist, die das relative Bewegungsverhältnis von Erde und Sonne beschreibt. Leicht lassen sich auch Stereotypen denken, die schlicht Unsinn enthalten oder sogar böseartig sind; man denke sich Stereotypen über Personen anderer Ethnie, die aus Uninformiertheit, Angst oder Desinteresse einer gegebenen Sprachgemeinschaft entstehen können. Dennoch werden gewisse Stereotypen als Kompetenzminimum vom Kollektiv vorausgesetzt, das als Ganzes im Normalfall selbstverständlich in der Lage ist, ein Wort viel exakter zu bestimmen, als das entsprechende Stereotyp im Kopf eines Einzelnen es tut. In der Regel kann jedoch davon ausgegangen werden, daß Stereotypen wenigstens einige paradigmatische Eigenschaften des entsprechenden Wortes einfangen wie im Beispiel von „Tiger“.

Putnam behauptet also, daß das basale Element der individuellen Sprachkompetenz in der Beherrschung mindestens der Stereotypen liegt. Doch hier täte sich, folgt man dem bedeutungstheoretischen Ansatz Putnams, ein weites Feld unerledigter Arbeit auf: man müßte eine Definition für Stereotypen finden und ein geeignetes formales System, in dem sie repräsentiert werden können. Putnam hat lediglich angedeutet, was man sich unter einem Stereotyp ungefähr vorzustellen hat und welche Rolle es für den Bedeutungsbegriff ungefähr spielt.

## IV.8 Fünfter Schritt: der Normalformtyp des Bedeutungsvektors

Wie soll ein Bedeutungsbegriff aussehen, der den vorangegangenen Überlegungen Rechnung trägt? Er müßte zweifelsohne eine Extension enthalten, denn diese wirkt, wie wir bereits gesehen haben, auf die Bedeutung zurück. Außerdem müßte ein neuer Begriff auch den Erkenntnissen über die soziale Arbeitsteilung Rechnung tragen. Also gehört auch die Angabe eines Stereotyps dazu. Damit wären die Bereiche „physikalische Welt“ und „Sprachgemeinschaft“ als Determinanten der Bedeutung abgedeckt.

Darüberhinaus will Putnam sogenannte syntaktische und semantische Marker in den Bedeutungsbegriff aufnehmen. Ein *syntaktischer Marker* ist einfach eine Variable, die Informationen über den syntaktischen Charakter des Wortes enthält. Syntaktische Marker für „Tiger“

wären zum Beispiel „Substantiv“ oder „konkret“. Ein *semantischer Marker* ist eine Variable, die etwas über die Einordnung des Begriffs in bestimmte semantische Kategorien unserer Sprache aussagt. Semantische Marker für „Tiger“ wären zum Beispiel „natürliche Art“, „Lebewesen“ oder „Raubtier“. Das Stereotyp enthält Angaben wie „schwarz-gelb gestreift“ oder vielleicht auch „menschenfressend“ und in ein paar Jahren sicherlich „vom Aussterben bedroht“. Das Stereotyp wird von der Sprachgemeinschaft geprägt. Wie das genau vor sich geht, können wir nicht im Rahmen einer rein logischen Explikation erklären, hier wäre die Soziolinguistik und die Sprachsoziologie gefragt. Das Stereotyp ist infolgedessen das am schwersten formal faßbare Element des Bedeutungsbegriffs. Die Extension des Wortes „Tiger“ schließlich ist, wie gehabt, die Menge aller Tiger.

Die Bedeutung des Wortes „Wasser“ sähe demnach so aus:

syntaktische Marker:	Kontinuativum, konkret
semantische Marker:	Flüssigkeit, natürliche Art
Stereotyp:	farblos, geschmacksneutral, durchsichtig, durstlöschend etc.
Extension:	H <sub>2</sub> O (mit und ohne Beimengungen)

Nach diesem revidierten Bedeutungsbegriff, den Putnam als „Bedeutungsvektor“ bezeichnet, kann man nun sowohl dem Erdling als auch dem Zwerdling dieselbe sprachliche Kompetenz zuschreiben, denn bis auf das vierte Element des Vektors würde sich die Bedeutung des Wortes „Wasser“ im Deutschen nicht von der zwillingsdeutschen Bedeutung unterscheiden. Im Bedeutungsvektor von „Wasser“<sub>TE</sub> stünde unter „Extension“ statt H<sub>2</sub>O eben XYZ.

## IV.9 Und nun?

Wenn wir nun auf das fertige Gedankengebäude zurückschauen, wird uns auffallen, daß Bedeutung nach wie vor Extension bestimmt (P2) – „per constructionem“ wie Putnam sagt. Hingegen ist die Vorstellung verworfen worden, daß es sich bei Bedeutungen um Entitäten handelte, die in den Köpfen der Menschen steckten (P1).

Natürlich wirft dies das weitaus Meiste dessen, was in Kapitel I über Bedeutungsbegriffe gesagt wurde, über den Haufen. Stattdessen ist ein Vorwurf aufgetaucht, der die gesamte individualistische Sprach- und Geistesphilosophie in ihrer Substanz angreift: der Gedanke, daß Bedeutungen nichts Privates sind und daß wir andere Dinge als nur den Sprecher ansehen müssen, um hinter ihre Natur zu kommen. Mit anderen Worten, daß Gedanken – bezogen auf das Subjekt – keine intrinsischen Eigenschaften sind, sondern einen Bezug zur Außenwelt beinhalten. Anklänge an diese These haben sich bereits bei Kripke gefunden. Auch der Gedanke, daß Kontexte die Bedeutung gewisser Äußerungen beeinflussen können, spielt seit Strawson eine Rolle, doch auch für Strawson war diese Bedeutung deshalb noch nicht irgendwo anders als im Kopf des Sprechers.

Das heißt aber, daß unsere Frage nach der kognitiven Signifikanz von Ausdrücken in einer Krise steckt: zwar könnte der Bedeutungsvektor als Wiedergabe der kognitive Signifikanz eines Ausdrucks funktionieren, doch enthält er keine Information, inwiefern die individuelle Sprachkompetenz zur Bedeutung beiträgt. Das Stereotyp legt offensichtlich keine Bedeutungen fest. Putnam erklärt nicht, wie jene Elemente, die tatsächlich im Kopf sind,

an der Festlegung der Bedeutung beteiligt sind. Also bleibt unsere Frage nach der kognitiven Signifikanz auch durch Putnams Ausführung eigentlich bei Kaplans Position stehen. Die Kaplansche Semantik aber findet keine Lösung auf die Twin-Earth-Probleme, also ist eine Weiterentwicklung oder Umarbeitung vonnöten – vorausgesetzt natürlich, wir lassen uns auf Putnams Argument ein und werfen den Internalismus über Bord.

Das haben aber nicht sehr viele Philosophen getan. Bevor wir in Kapitel VI eine Versöhnungsmöglichkeit internalistischer und externalistischer Argumente kennenlernen, wollen wir sehen, was passiert, wenn man stattdessen Putnams Externalismusargument zurückweist und den Internalismus in seiner bisherigen Form zu verteidigen versucht.

Eine der am breitesten diskutierten Verteidigungsstrategien wurde von Jerry Fodor aufgebracht, einem Freund und ehemaligen Schüler Putnams. Fodor hat die Position des semantischen Externalismus scharf kritisiert und den (zur Zeit) bekennenden Realisten Putnam in einen längeren Dialog über diese fundamentale Annahme des Internalismus verwickelt – wobei man einräumen muß, daß die Position Fodors hinsichtlich des Externalismus zu keinem festen Profil gefunden hat. Fodor trat eine Diskussion los, die eine der dunkelsten und unstrukturiertesten der jüngeren Sprachphilosophie ist: die Diskussion um den Begriff des „engen Inhalts“, der seiner Meinung nach das sein soll, was im Kopf ist, und daher das, was bei den Zwillingen in Putnams Beispielen gleich ist, wobei deren psychische Zustände sich in Wirklichkeit unterscheiden könnten. Wir werden uns im nächsten Kapitel an diesen Begriff und seine Rolle in der internalistischen Verteidigungsstrategie anhand einer kleinen Auswahl von Texten herantasten, ohne das Thema „enger Inhalt“ allzusehr zu vertiefen.



# V Der „Enge Inhalt“ als Verteidigung gegen den Externalismus

## V.1 Wie man Putnam begegnen könnte

Wir sind der historischen Wendung, die die Bedeutungsdiskussion genommen hat, gefolgt. Die Frage nach der kognitiven Signifikanz von Überzeugungszuständen ist in den Vordergrund der Diskussion getreten, während die drängendsten semantischen Fragen seit Kaplan erst einmal gelöst zu sein scheinen.

Das Twin-Earth-Experiment ist für den Internalismus deshalb so gefährlich, weil es die extensionale Identität der Gedanken von deren „inhaltlicher“ Identität (im intuitiven Sinne) insofern trennt, als daß das eine kein Beweis mehr für das andere ist. Dies liegt an der Art, wie in Putnams Experiment propositionale Einstellungen identifiziert werden. In Abschnitt 4 dieses Kapitels werden wir ein starkes Argument gegen diesen Trennungsversuch besprechen.

Generell geht es in diesem Kapitel um den Inhalt von Gedanken. Nach wie vor wollen wir wissen, wie wir die Beschreibung der kognitiven Signifikanz von Wörtern mit einem Grundriß rekursiver Semantik verbinden können. Putnams Argumente gegen P1 stellen hierfür jedoch ein Problem dar: wenn Bedeutungen nicht im Kopf sind, dann sieht die Rolle, die die Verbindung von Kognition und Semantik spielt, ganz anders aus, als wir bisher dachten und auch die Verbindung selbst muß neu gedacht werden. Wenn man Putnams Folgerungen nun ablehnt und P1 retten will, so gibt es die Möglichkeit zu behaupten, daß die psychischen Zustände der Zwillinge gleich sind, doch daß nicht die psychischen Zustände, sondern der Inhalt der Gedanken den extensionalen Unterschied in den Überzeugungssätzen der Zwillinge auf mentaler Ebene abbildet und die Determination von Extension durch das, was im Kopf ist, deshalb nicht widerlegt ist. Diesen Weg geht Fodor, dessen Position wir in Abschnitt 4 und 5 dieses Kapitels besprechen. Er faßt den engen Inhalt als einen Begriff auf, der ähnlich wie Kaplans Charakterbegriff aussehen soll.

Die andere Möglichkeit, besteht darin, abzulehnen, daß der Inhalt der Gedanken von Bedeutungen wirklich bestimmt wird. Das würde heißen, der Komplementsatz in einer Überzeugungszuschreibung (Julia glaubt, Wasser sei XYZ) legt den eigentlichen Inhalt der Gedanken gar nicht fest. Dieser Strategie folgt Loar in [Lo88]. Allerdings bemüht Loar im Gegensatz zu Fodor die überaus starke These, daß der Inhalt von Gedanken – respektive von propositionalen Einstellungen, gar nicht wirklich propositional festgelegt ist: vielmehr gebe es für jeden Gedanken Elemente, die zu ihm gehören und nicht Teil des daß-Satzes innerhalb des Zuschreibungssatzes sind. Loar psychologisiert dadurch das Argument und faßt engen Inhalt nicht wie White und Fodor als einen charakterartigen Begriff auf.

Der Begriff des engen Inhalts wurde in den Achtzigern und Anfang der Neunziger Jahre

breit diskutiert. Es geht mir nicht darum, hier eine Darstellung dieser Diskussion oder auch nur einen repräsentativen Querschnitt zu geben. Ich möchte nur die beiden Antwortmöglichkeiten, die ich angesprochen habe, anhand der exemplarischen Positionen von Loar und Fodor aufzeigen. Da Fodors Begriff des engen Inhalts formal undurchdacht und eigenartig begründet ist, will ich ihm außerdem die viel durchdachtere und theoretisch sauberere Konzeption gegenüberstellen, die White in [Whi82] vorgelegt hat. Dann werde ich Fodors Inhaltsbegriff der Kritik durch Schiffer unterziehen, der behauptet, daß Fodors Position keinen solchen Begriff erfordert. Außerdem wird die Position von Ned Block besprochen werden, wonach ein Inhaltsbegriff, wie Fodor ihn beabsichtigt, unmöglich ist. Ich hoffe, daß durch diese Darstellung deutlich wird, welche Route das internalistische Hauptargument gegen Putnam eingeschlagen hat und woran diese Argumentationsrichtung krankt. Wird dies deutlich, so hat dieses Kapitel seinen Zweck schon erfüllt.

## V.2 Der Inhalt einer Einstellung ist nicht propositional festgelegt:

### Loar

In [Lo88] findet sich ein Grundriß zu einer Vorgehensweise gegen den semantischen Externalismus, die sich in erster Linie gegen [Bu79] richtet. Historisch gesehen ist Loar natürlich nicht der erste, der einen solchen Grundriß liefert, doch für die systematische Darstellungsweise, die ich für diese Kapitel gewählt habe, bietet es sich an, seine Position der eigentlichen Diskussion um engen Inhalt voranzustellen, weil sie auf noch fundamentalerer Ebene als die späteren Positionen gegen den Externalismus ansetzt.

Wenn es gelänge, zu zeigen, daß neben dem normalen propositionalen Inhalt eines Überzeugungssatzes eine Art „psychologischer Inhalt“ beziehungsweise „enger Inhalt“ bestehen kann, ohne daß dieser psychologische Inhalt zwangsläufig von den im daß-Satz angegebenen Wahrheitsbedingungen abhängt und ohne daß er demzufolge durch diesen daß-Satz vollständig bestimmt oder wiedergegeben werden kann, dann wäre es möglich zu behaupten, daß Burge und Putnam zwar recht haben, wenn sie sagen, daß in bestimmten Fällen von Überzeugungssätzen, die einen Gebrauch von Namen oder anderen Gegenstandsbezeichnungen (denn auf solche beschränken wir uns hier) in intensionalem Kontext [*oblique occurrence*] beinhalten, die herkömmliche Art der Zuschreibung einer Überzeugung [*ascription of belief*] versagt. Andererseits jedoch wäre das nicht mehr schlimm, weil man sich nicht mehr auf die Folgerung einlassen müßte, daß der eigentliche Inhalt der Einstellung von irgendwelchen externen Elementen – zum Beispiel durch Besonderheiten der sprachlichen Sozialisation – signifikant mitbestimmt werde. Man könnte nämlich darauf verweisen, daß es keine Wahrheitsbedingungen gebe, durch die sich dieser Inhalt angeben lasse.

Bei Putnam und Burge war das gängige Argument folgendes.

Stelle ich zwei Personen, einen Erdling und seine zwillingsirdische Version vor ein Glas Wasser<sub>E</sub>, würden zwar beide sagen: „Das ist Wasser.“ Doch in diesen Aussagen ist das Wort „Wasser“ – da der eine Deutsch, der andere Zwillingsdeutsch spricht – aufgrund seiner unterschiedlichen Bedeutungen nicht *salva veritate* austauschbar. Auf deutsch kann man Wasser als H<sub>2</sub>O beschreiben, doch der zwillingsdeutsche Ausdruck „Wasser“ muß sowohl auf Deutsch wie auf Zwillingsdeutsch anders umschrieben werden. Abgesehen davon, daß Put-

nam daraus schließt, daß Sozialisation und physikalische Konfiguration der Welt die Bedeutung des Wortes „Wasser“ beeinflussen, muß er auch auf folgendes schließen:

Der eine ist überzeugt, daß „dies hier“ Wasser<sub>E</sub> ist, der andere ist überzeugt, daß „dies hier“ Wasser<sub>TE</sub> ist. Also unterscheiden sich die beiden Überzeugungen ganz erheblich, was ihren nichtintentionalen *Inhalt* im Sinne dessen, was gesagt wird, angeht, denn der eine ist von einer Aussage über H<sub>2</sub>O überzeugt, der andere von einer Aussage über XYZ. Dies entspricht völlig Putnams Auffassung, wonach Überzeugungen über Gegenstände, die nicht völlig identisch sind, verschiedene Überzeugungen sind.

Es scheint, folgt man Putnam, so zu sein, als ob sich diese psychologischen Inhalte der Überzeugungssätze der beiden Probanden unterscheiden, und zwar deshalb, weil die Extensionen der Homonyme „Wasser“ sich unterscheiden. Hinter diesem Argument steht folgende Auffassung:

**T1)** Unterschiede bezüglich der Zuschreibung von Überzeugungen, die nichtextensionale Vorkommnisse allgemeiner Termini enthalten, implizieren Unterschiede im psychologischen Inhalt der Überzeugung<sup>1</sup>.

Man kann T1 als Konversion des folgenden, etwas genauer gefaßten Satzes auffassen:

**T2)** Die Identität eines nichtextensionalen Vorkommens eines allgemeinen Ausdrucks in zwei Überzeugungszuschreibungen impliziert, sofern alles andere gleich bleibt (*ceteris paribus*), die Identität der psychologischen Inhalte der Überzeugungen, die jeweils zugeschrieben werden.

Dies ist nach Loar der Punkt, an dem der Externalist sich in Widersprüche verstricken muß, denn man kann beweisen, daß T2 falsch ist, womit auch die Falschheit von T1 gezeigt wäre. Der Beweis besteht in einer leichten Abwandlung eines Gedankenexperiments, das sich bereits in [Kri79] findet und als Gegenstrategie genau das Problem aus I.6 aufzeigt, das uns bewog, innerhalb der Carnapschen Semantik Einstellungskontexte als hyperintensional zu behandeln: das Problem der Subjektivitätskompetenz.

Beweis: Man stelle sich einen Sprecher – Pierre – vor, der monolingual in Frankreich aufgewachsen ist. Er hat von einer schönen Stadt gehört, die „Londres“ genannt wird, und er sagt

S15) „Londres est jolie“.

Er zieht schließlich nach London, ohne zu wissen, daß es „Londres“ ist und landet unglücklicherweise in einem recht trüben Stadtteil. Traurig sagt er

---

<sup>1</sup>Hier ist zu beachten, daß „psychologischer Inhalt“ [*psychological content*] bei Loar und „mentaler Inhalt“ [*mental content*] bei Burge etwas anderes ist als das, was Putnam als „mental state“ bezeichnet: Burge und Loar meinen den Inhalt einer bestimmten Überzeugung während Putnams Ausdruck lediglich klarmachen soll, daß innerhalb der traditionellen Theorie die individuenbezogene Entsprechung eines Bedeutungsausdrucks ein irgendwie mentaler „Zustand“ sei. Was ein solcher „Inhalt“ sein könnte, wird in Abschnitt 3 erläutert, wenn es um Fodors Vorschlag geht.

S16) „London ist nicht schön“.

(Kripke nimmt dies als Beweis dafür, daß unsere herkömmlichen Prinzipien der Zuschreibung von Überzeugungen uns anleiten, daß Pierre nun sowohl glaubt, daß London schön ist, als auch, daß London nicht schön ist.)

Pierre könnte aber auch mehr Glück gehabt haben und stattdessen in einem anderen Stadtteil eine Wohnung bekommen haben, so daß er statt S16) denkt

S17) „London ist schön“.

Nehmen wir also folgenden Satz:

S18) Pierre glaubt, London sei schön.

S18 ist wahr aufgrund der früheren Fakten über Pierres sprachliche Sozialisation (vgl. S15) und aufgrund der späteren Fakten (Pierre in London). Diese doppelte Wahrheit von S18 hängt nicht, wie es vielleicht scheint, von Übersetzungsproblemen ab, sondern davon, daß die koreferentiellen Ausdrücke „London“ und „Londres“ nichtextensional vorkommen, also nicht *salva veritate* durch andere, koextensionale Ausdrücke ersetzt werden können: das Gleiche würde passieren, wenn jemand sich beispielsweise über die Bedeutung eines Wortes irrt und es als Bezeichner zweier verschiedener Dinge annimmt, wobei er nur eines davon korrekt bezeichnet. Doch nun taucht die Frage auf, wieviele Überzeugungen Pierre mit S15 und S17 denn tatsächlich hat.

Der common sense<sup>2</sup> kennzeichnet S15 und S17 als zwei distinkte Überzeugungen, ebenso wie „Paris ist schön“ und „Rio ist schön“. S15 und S17 haben auch unterschiedliche funktionale Rollen, sie würden mit anderen Überzeugungen und Annahmen unterschiedlich interagieren. (Pierre könnte zum Beispiel aufgrund von S15 denken, daß er, wenn er einmal nach Londres kommen sollte, die Stadt besucht, in der Oscar Wilde gelebt hat. Doch wenn er dann in London wohnt, muß ihm das nicht auffallen.)

Was noch entscheidender ist: die beiden Überzeugungen unterscheiden sich in ihrem Inhalt:

„These beliefs not only are individuated by commonsense psychology as distinct in their psychological roles; it also seems quite appropriate to regard them as distinct in content. The differences in their interactive properties flow from differences in how Pierre conceives things, in how he takes the world to be, in what he regards the facts

---

<sup>2</sup>Der Ausdruck „common sense“ ist hier eine Abkürzung für den sperrigen Ausdruck „commonsense belief/desire psychology“, der normalerweise eine Psychologie folgender Form bezeichnet: wenn  $p$   $x$  will und  $y$  tun muß, um  $x$  zu bekommen, dann wird  $p$   $y$  tun. Bei der Behandlung von Fodors Position wird die Inbrunst auffallen, mit der Philosophen über diese „CBDP“ sprechen. Auch im folgenden werde ich die Abkürzung „common sense“ verwenden, weil mir der Ausdruck „common-sense-Psychologie“ oder „Volkspsychologie“ konstruiert und mißverständlich erscheint. Wir wollen ja gerade zwischen dem common sense und dem, was die Psychologie sagt, unterscheiden! Für eine genauere Besprechung des Begriffs des common sense vgl. V.4.

as being – that is, differences in some semantic or intentional dimension.“ [[Lo88] p. 103]

Damit ist T2 bereits widerlegt, denn offensichtlich kann es Sätze geben, für die gilt, daß sie, obwohl sie beide jeweils nichtextensionale Vorkommnisse von koreferentiellen Ausdrücken enthalten, unterschiedliche korrespondierende psychologische Inhalte verkörpern.

Wie der aufmerksame Leser bemerkt hat: Loar behauptet damit, daß das Substitutionsprinzip für Intensionen in diesen Fällen nicht gilt. Das Problem dabei ist, daß wir es mit einem Kontext zu tun haben, den Carnap nicht als intensional vorgesehen hatte – laut Carnap sind Einstellungskontexte nicht intensional, deshalb kann keine Gültigkeit dieses Prinzips postuliert werden. Das ist keine Überraschung. Dennoch könnte, wie wir bald sehen werden, diese Erkenntnis ein Problem für den Externalisten darstellen. Der Witz ist nämlich, daß Loar aus dem Pierre-Beispiel schließt, daß es eine Art von psychologischem Inhalt gibt, der nicht durch die daß-Sätze innerhalb der Überzeugungszuschreibungen ausgedrückt werden kann – sonst müßte er derselbe sein. Es ist also, wie schon in I.6 bemerkt, leicht zu sehen, daß die Verortung kognitiver Signifikanz von Überzeugungssätzen das Projekt der rekursiven Semantik direkt beeinflusst, weswegen, wie wir in I.4 ausführlich besprochen haben, Bedeutungsbegriffe nötig sind, die subjektive und objektive Dimensionen abbilden.

Mit Loars Schlußfolgerung aus dem Pierre-Beispiel rückt der Begriff des Inhalts in den Mittelpunkt der Überlegungen. Loar meint hier einen „psychologischen Inhalt“, den man nicht mit dem Inhalt in Kaplans Sinne und erst recht nicht mit dem propositionalen oder „weiten“ Inhalt verwechseln darf, der in dieser Lesart mit dem daß-Satz innerhalb des Überzeugungssatzes zusammenfällt. Was sollen wir uns unter dem psychologischen Inhalt einer Überzeugung vorstellen? Loar sagt dazu:

„By psychological content I shall mean whatever individuates beliefs and other propositional attitudes in commonsense psychological explanation, so that they explanatorily interact with each other and with other factors such as perception in familiar ways.“  
[[Lo88] p. 99]

Das ist ziemlich vage. Wir haben es mit einer in erster Linie psychologischen Größe zu tun, die offensichtlich eine gewisse Erklärungskraft besitzt, ohne aber im herkömmlichen Sinne semantisch auswertbar zu sein. Man könnte an diesem Punkt bereits einhaken und Loar antworten, daß ein Begriff, der nicht präzise semantisch ausgewertet werden kann, für eine semantische Theorie – und eine solche müßte auf die semantische Theorie von Putnam und Burge antworten – wohl keinen großen Wert besitzt.

Loars psychologischer Inhalt darf auch nicht mit Freges „Gedanken“ verwechselt werden, denn der Gedanke fällt mit der Proposition zusammen und die ist durch den Komplementsatz, also den weiten Inhalt, ausgedrückt.

Doch bevor wir Einwände erheben, soll Loars Argument entwickelt werden. Deshalb werden wir fürs erste seine Definition des psychologischen Inhalts verwenden.

Wenn wir Loars Strategie näher untersuchen, werden wir feststellen, daß sie nicht schwer zu verstehen ist:

Wenn es nicht die daß-Sätze sind, die den psychologischen Inhalt einer Überzeugung beschreiben, dann bekommt Putnams Angriff auf P1 einen ganz anderen Charakter. Denn dann

muß es „zwischen“ dem Wort und dem psychischen Zustand, der dessen Bedeutung mittelbar festlegt, so etwas wie einen psychologischen – oder „engen“ – Inhalt (der Überzeugung, was das Wort bedeute) geben, der mit der Bedeutung zu tun hat und die Tatsache erklärt, daß der psychische Zustand derselbe sein kann, wenn die Bedeutungen des Wortes differieren. Der Begriff dieses Inhalts würde es zum Beispiel ermöglichen, zu argumentieren, daß nicht wirklich der psychische Zustand (im engen Sinne) des Erdlings und des Zwerdlings gleich sein muß, sondern es etwas anderes gibt, das gleich geblieben ist, und wir P1 gegen die scharfsinnigen Folgerungen aus dem Twin-Earth-Experiment verteidigen können. Mit anderen Worten: die internalistische Tradition kann zwar zugeben, daß es etwas gibt, worin ich und mein Zwilling uns gleichen, wenn wir an Wasser denken oder darüber sprechen, aber es ist letzten Endes nicht der psychische Zustand, sondern etwas anderes und faktisch können unsere psychischen Zustände sich anderweitig unterscheiden. Oder noch einfacher: Wenn Putnam P1 widerlegen will, dann soll er erst einmal beweisen, daß es tatsächlich die psychischen Zustände (im engen Sinne) des Erdlings und des Zwillingserdlings sind, die gleich sind und nicht etwas anderes, das wir finden und erklären können.

Doch es gibt ein Problem: der psychologische Inhalt darf nicht die Bedeutung der Ausdrücke und Sätze in dem Sinne bestimmen, in dem der propositionale, weite Inhalt sie bestimmt, denn sonst liefert man sich der Twin-Earth-Argumentation aufs Neue aus. Er muß aber mit der Bedeutung zusammenhängen und eine Spielart psychologischer Bedeutung irgendwie repräsentieren, denn sonst liefert er keine Grundlage für eine Verteidigung. Hier zeichnet sich bereits die gravierendste Schwierigkeit mit diesem Begriff ab, die bald noch deutlicher werden wird.

Wenn man nun der Intuition folgt und davon ausgeht, daß meine Sätze über Wasser und die Sätze meines Zwillings über das, was er „Wasser“ nennt, insofern denselben Inhalt haben, als dieser von der Psychologie identifiziert wird, beziehungsweise denselben „engen Inhalt“ [*narrow content*] – dann könnte man, so Loar, noch immer zwei Dinge einwenden:

- 1) Enger Inhalt kann keine intentionalen Eigenschaften beschreiben. Das kann man daran sehen, daß die jeweiligen Counterpart-Überzeugungen im Arthritis-Experiment und im Twin-Earth-Experiment keine Wahrheitsbedingungen gemeinsam haben. Intentionalität beinhaltet eine gewisse Richtung der Gedanken auf die Konstellation von Entitäten, also auf Bedingungen, unter denen sie wahr sind. Ich intenziere schließlich die Angabe bestimmter Wahrheitsbedingungen mit meiner Aussage.
- 2) Es ist nicht zu sehen, wie man denn den gemeinsamen engen Inhalt solcher Überzeugungspaare beschreiben sollte. Ein Begriff, der nichts bezeichnet, das man beschreiben kann, ist nutzlos.

Einwand 1) bewirkt, das räumt Loar ein, daß der enge Inhalt nicht die Wahrheitsbedingungen von Überzeugungen festlegen kann. Indem Loar zwischen Wahrheits- und Realisierungsbedingungen unterscheidet, schafft er jedoch die Möglichkeit, einen Begriff des Inhalts zu entwickeln, der sich gar nicht auf Wahrheitsbedingungen, wie wir sie kennengelernt haben<sup>3</sup>, bezieht. Die Forderung einer internalistischen (also: engen oder auch: individualisierten) Erklärung von Intentionalität ist dadurch zu beantworten. Zwar weiß ich beim

<sup>3</sup>Wahrheitsbedingungen müssen in Satzform angebar sein, das bezweifelt auch Loar nicht. Wir können uns

Lesen oder Hören von Sätzen nie, ob dem Gebrauch der Wörter eine Fehlkonzeption zugrundeliegt, doch ich weiß, wie die Welt, die durch diese Sätze beschrieben wird, aussehen würde, wenn der Gebrauch keine Fehlkonzeptionen enthielte. Auf diese Weise bestimmen Gedanken eine Klasse möglicher Welten, in denen sie wahr sind, wenn sie aus korrekt verwendeten Begriffen zusammengesetzt sind. Diese Klasse soll die Realisierungsbedingungen des jeweiligen Gedankens angeben, die nicht mit den Wahrheitsbedingungen identisch sind, die ihrerseits durch den daß-Satz gegeben sind, denn *dieser* Satz ist, wie Burge gezeigt hat, nicht von rein internalistischen Elementen abhängig. Die Wahrheitsbedingungen eines Gedankens sind einfach die möglichen Welten, in denen der Satz, durch den ich den Gedanken repräsentiere, wahr ist. Wahrheits- und Realisierungsbedingungen fallen deshalb nicht zusammen, weil Realisierungsbedingungen Intentionalität berücksichtigen sollen, Wahrheitsbedingungen nicht.

Finde ich zum Beispiel ein Tagebuch, in dem steht: „Heute war es heiß und sonnig; war mit Betty am Strand“, so weiß ich auch ohne Datumsangabe, was ich mir vorzustellen habe, das heißt, ich weiß, wie eine Welt aussehen würde, in der dieser Satz wahr ist. Was ich mir unter „heiß und sonnig“ vorstelle, ist eine Realisierungsbedingung des Satzes „Heute war es heiß und sonnig“, ganz im Gegensatz zur Wahrheitsbedingung, daß es an dem Tag, als der Schreiber den Satz schrieb, heiß und sonnig gewesen sein muß. Die Realisierungsbedingung ist im Gegensatz zur Wahrheitsbedingung in gewissem Sinne kontextunabhängig. Um also die psychologische Erklärung, warum ich am Strand verabredet war, verstehen zu können, muß man nur die Realisierungsbedingungen des Satzes verstehen, während seine Wahrheitsbedingungen nicht entscheidend sind. Wo psychologische Erklärung auf Intentionalität angewiesen ist, wird diese Intentionalität also durch kontextunabhängige Realisierungsbedingungen repräsentiert, nicht durch die Wahrheitsbedingungen, die ihrerseits mit dem weiten, das heißt propositionalen Inhalt des Gedankens zusammenfallen.

Was 2) angeht: zwar haben wir keine Spezifikationsmethode für engen Inhalt, so wie wir sie für weiten Inhalt haben, der ganz einfach mit dem daß-Satz, respektive den Wahrheitsbedingungen oder auch dem, was Frege einen Gedanken nennt, zusammenfällt und den Loar „sozialen Inhalt“ nennt (weil er mitteilbar ist: hier zeigt sich die Nähe zu Freges Begrifflichkeiten). Aber wir können psychologische Erklärungen liefern, die von solchen Spezifikationen unabhängig sind. Wir können Überzeugungen identifizieren, feststellen ob sie typenidentisch sind und haben perfekte Intuitionen, wann wir die Überzeugungen verschiedener Leute zu unterscheiden haben. Und enge Inhalte sind nicht insofern diffus, als daß wir sie nicht approximieren könnten:

- a) Wir können daß-Sätze benutzen und dennoch Kontextfakten berücksichtigen, beispielsweise in dem Fall, wo wir mit Pierre über London sprechen. Wir könnten die Bedeutungsdifferenz durch Erfragen von Kontextfaktoren herausbekommen und korrigieren.
- b) Wir haben praktisch nie einzelne Sätze vor uns, sondern können immer andere Äußerungen des Sprechers hinzuziehen, die uns helfen, den engen Inhalt, der durch einen bestimmten Satz ausgedrückt ist, zu spezifizieren.

---

hier an Frege erinnern, der Gedanken als Wahrheitsbedingungen von Sätzen annahm. Die Gedanken in Freges Sinn fallen hier mit dem propositionalen Gehalt von Einstellungen zusammen. Es ist beeindruckend, wie weit Frege hier gesehen hat.

- c) Wir schreiben engen Inhalt zu, indem wir Wörter benutzen, die für uns denselben engen Inhalt haben: Londres kann ich auch als schön bezeichnen, indem ich sage „London is beautiful“.

Die Approximation von engen Inhalten ist so gesehen eine der fundamentalen Techniken der wissenschaftlich-psychologischen Erklärung. Daß-Sätze, so Loars These, sind zwar bedeutend, wenn es um die linguistische Repräsentation von Gedanken geht, doch sie sind nur sehr lose mit dem verknüpft, was er engen Inhalt nennt – was Antwort a) wieder relativiert, denn dann ist nicht zu sehen, woher die Approximation durch Sätze ihre Legitimität bezieht. Was Burge dagegen mit seinen Beispielen aufzeigt, ist nach Loar eine Art weiter Inhalt, der angibt, wie man sich korrekt ausdrücken würde, wenn die Sprachgemeinschaft genauso sprechen würde, wie man selbst.

Mit Loar haben wir uns nun auf den schmalen Grat zwischen semantischen und psychologischen Problemen begeben, den wir nun kurz wieder verlassen werden, um die formale Konzeption Whites zu besprechen, um mit Fodor bald auf dieses dünne Seil zurückzukehren.

Was ist mit Loar erreicht? Die obige Strategie führt uns dazu, zu postulieren, daß es eine Art psychologischen Unterschied der Putnam-Zwillinge gibt, nämlich einen Unterschied im psychologischen, das heißt engen Inhalt von deren Überzeugungen. Das Problem ist, daß wir den engen Inhalt nicht in Sprache darstellen können (außerdem bedarf er eines Trägers und gehört zur Innenwelt des Sprechers, weshalb ich finde, daß er dem ähnelt, was Frege eine Vorstellung nannte). Außerdem muß er von Wahrheitsbedingungen, die der propositionale Inhalt bereits angibt, unabhängig sein – doch dann müssen wir zwischen Wahrheits- und Realisierungsbedingungen unterschieden und das Auffinden beliebiger Beispiele, in denen diese auseinanderfallen, heißt noch nicht, daß es eine Unterscheidung von theoretischer Tragfähigkeit ist. Außerdem bedeutet dieser „entpropositionalisierte“ – eben enge – Charakter, daß der enge Inhalt intentionale Eigenschaften einschließen muß, mit anderen Worten, daß wir enge Inhalte nicht so unterscheiden dürfen, wie Putnam es tut, wenn er sagt, die Überzeugung, daß  $i$  die Intension von  $\alpha$  sei, sei etwas anderes als die Überzeugung, daß  $i$  die Intension von  $\beta$  sei, selbst wenn  $i(\alpha)$  und  $i(\beta)$  identisch sind. Das macht den Begriff des engen Inhalts immer diffuser. Denn schließlich muß der Inhalt irgendwie mit der Extension der Ausdrücke verknüpft sein, weil er sonst nicht als Verteidigung taugt. Andererseits aber darf er diese Extension, wie wir gesehen haben, auch nicht bestimmen. Irgendwie muß er aber auch Teil einer rekursiven Semantik sein, denn sonst kann diese mit den Twin-Earth-Problemen nicht umgehen.

Besondere Probleme bereitet uns vorderhand die sprachliche Approximation des engen Inhalts. Die drei Möglichkeiten, die Loar hier einräumt, stehen auf tönernen Füßen. b) scheint noch am vielversprechendsten zu sein, denn a) ist bereits relativiert, wenn wir sagen, daß-Sätze haben nichts mit engem Inhalt zu tun, und c) krankt an dem gleichen Problem. Und b) bezieht sich auf pragmatische Hilfsmittel, die auch fehlen können. Wenn wir also behaupten, daß die Inhalte von Überzeugungen nicht oder wenigstens nicht ausschließlich propositional festgelegt sind, dann räumen wir ein, daß enger Inhalt in Fällen, wo Kontextinformationen fehlen, sich als irreduzibel privat und nicht mitteilbar herausstellen könnte. Das ist eine unattraktive Möglichkeit.



### V.3 Inhalt in formalem Gewand: White

Die These, daß Bedeutung wesentlich extra-individualistisch bestimmt sei, richtete die allgemeine Aufmerksamkeit schnell auf die Erforderlichkeit einer genauen Unterscheidung zwischen internalistischen und nichtinternalistischen Elementen der Bedeutung. Diese Unterscheidung ist durch Putnam in Form der Stereotypenlehre allenfalls rudimentär getroffen worden. Abgesehen davon rief Putnams impliziter Angriff auf den von ihm selbst aufgebrauchten Funktionalismus – die Annahme, daß man Überzeugungen in Form funktionaler Zustände des Geistes beziehungsweise des Gehirns darstellen könne – Stimmen auf den Plan, die diese Lehre zu verteidigen versuchten, auch ohne daß primäre Interessen an der Weiterentwicklung der bisherigen Bedeutungstheorie bestanden. So nahmen auch die Antworten auf externalistische Argumente verschiedene Formen an, je nachdem, ob man Putnam, Burge und Stich als Semantiker oder als Psychologen verstand.

Eine der ersten (und durchdachtsten) Entgegnungen auf die Kernthese des Externalismus findet sich in [Whi82]. White antwortet jedoch nicht nur auf die bedeutungstheoretischen Überlegungen Putnams, sondern auch auf den Angriff auf den Funktionalismus in der Philosophie des Geistes, der besonders deutlich von Stich geführt wurde. White will den Funktionalismus mithilfe einer Differenzierung zwischen internen und externen Komponenten der Bedeutung verteidigen, wobei er den überwältigenden Erfolg von Kaplans Charakterbegriff durch geschickte Definition auf den Begriff des engen Inhalts (den er noch gar nicht verwendet) zu übertragen versucht. Obwohl die Frage nach der Korrelierbarkeit von Überzeugungen mit funktionalen Gehirnzuständen eng mit unseren Überlegungen verwandt ist, sollen uns hier hauptsächlich die bedeutungstheoretisch relevanten Teile der Diskussion interessieren.

White kritisiert an Putnam in erster Linie vier Aspekte:

- 1) Die von Putnam skizzierte Stereotypenlehre repräsentiert nicht alles, was individualistisch verstanden werden muß. Jeder Sprecher einer Sprachgemeinschaft muß, um verstanden werden zu können, die Sprache in derselben Weise gebrauchen wie die Gemeinschaft; er muß ein Interesse am korrekten Gebrauch der Ausdrücke haben. Diese Verbindlichkeit wird durch die Intentionen repräsentiert, mit denen der Sprecher sich äußert und die Ausdrücke der Sprache gebraucht. Diese Art von Intentionalität *ist* im Kopf und bestimmt die Extension auf jeden Fall mit. Damit muß ein Bedeutungsbegriff, der aufzeigt, was im Kopf ist, Intentionalität abbilden können.
- 2) Stereotypisch schwer unterscheidbare Wörter wie „Ulme“ und „Buche“ sind eher die Ausnahme als die Regel. Der Wortschatz eines noch als kompetent gelten könnenden Sprechers darf nicht zu viele Wörter enthalten, die er nur mit diesen „dünnen“ Stereotypen verknüpft. Wäre das nämlich der Fall, so läge eine uninterpretierte Sprache vor. Wir können Sprecher als kompetent akzeptieren, die Ulmen und Buchen nicht voneinander unterscheiden können, doch Leute, die nicht einmal wenigstens die eine oder andere korrekte Ostension liefern können, wenn man sie fragt, was denn ein Stuhl sei, die akzeptieren wir nicht als kompetente Sprecher. Also kann es nicht sein, daß an Stereotypen derart niedrige Anforderungen gestellt werden, wie Putnam behauptet, denn

jeder Sprecher muß wenigstens einige „reiche“ Stereotypen [*thick stereotypes*] haben, um überhaupt korrekte Referenz herstellen zu können.

- 3) Putnam nimmt keinen Bezug auf die rekursive Semantik, das heißt, er erklärt nicht, wie die Wahrheitsbedingungen von komplexen Sätzen durch Analyse der Wahrheitsbedingungen ihrer einfacheren Bestandteile festgelegt werden. Zum Beispiel kann ein bilingualer Sprecher das gleiche Stereotyp mit „Buche“ wie mit „beech“ verbinden. Dennoch kann er bei einem bestimmten Baum zwar glauben, daß es eine Buche sei, doch in der Frage, ob es eine „beech“ sei, agnostisch sein. Wenn die Wahrheitswerte einiger Glaubenssätze sich in verschiedenen Repräsentationen derselben Proposition unterscheiden, dann muß erklärt werden, wie so etwas geht. Wie Putnam schon einräumt, zeigt diese Repräsentationsabhängigkeit, daß die Extension von mehr als dem Stereotyp abhängt, doch zusätzlich muß auch gezeigt werden, wie dieses darüber Hinausgehende Teil einer rekursiven Semantik sein kann. (Schließlich geht es uns in dieser Arbeit um die Verbindung von epistemischen und formal-semantischen Elementen und dazu schweigt Putnam sich aus, wie wir gesehen haben.)
- 4) Es gibt noch einen weiteren Aspekt, der unbestreitbar im Kopf ist und die Extensionen unserer Begriffe mitbestimmt, nämlich die jeweilige Klasse von subjektbezogenen Dispositionen, die das individuelle Urteil über Ähnlichkeit und Übertragbarkeit von Überzeugungen von bekannten Situationen auf unbekannte Situationen regelt. Zwar hängen hier die Begriffe, die der Einzelne „besitzt“, nicht vollständig davon ab, was in *seinem* Kopf ist, doch ebensowenig wird das von Experten geregelt. Prädikate wie „ist rot“ oder „ist lustig“ sind zum Teil durch die extra-linguistische Tradition geregelt, das eigene Urteil ist dabei dem der Sprachgemeinschaft anzupassen.

Was White abgesehen von diesen speziellen Aspekten an Putnam stört, ist etwas sehr Generelles:

„What is being criticized is the lack of an adequate theory of the way in which what is in the head contributes to the determination of meaning and reference.“ [[Whi82] p. 353]

Zwar erklärt Putnam zumindest ungefähr, inwiefern extraindividualistische Elemente die Bedeutung der Ausdrücke beeinflussen, doch er läßt völlig offen, wie das, was nun tatsächlich im Kopf des Einzelnen ist, dafür eine Rolle spielt. Die Stereotypenlehre kann die von White genannten Punkte nicht plausibel erklären und Putnam verschweigt generell, inwiefern das, was im Kopf ist, zur Festlegung von Bedeutung und Referenz beiträgt. Zwar trennt Putnam individuelle und externe Elemente von Bedeutung, doch er erklärt nicht den Mechanismus der Verbindung zwischen internen Komponenten eines linguistischen Ausdrucks und der Extension. (Das wird im Twin-Earth-Experiment deutlich: wir sehen nicht, wie die Stereotypen oder das, was sonst im Kopf ist, die Extension beeinflusst.)

Ergo braucht man eine Theorie, die nicht nur eine Trennung von internen und externen Komponenten der Bedeutung möglich macht, ohne die von White beschriebenen Defekte aufzuweisen, sondern auch eine, die etwas über das Verhältnis von Bedeutung und dem,

was im Kopf ist, sagt. Und eine solche Theorie will White mithilfe des Begriffs des „partiellen Charakters“ liefern. Dieser Begriff ist ein eng verwandter historischer Vorläufer dessen, was Fodor als „engen Inhalt“ bezeichnet – nur ist White sehr viel präziser in der Ausführung dieses Begriffs und auch etwas scharfsinniger in seiner Kritik Putnams: die genannten vier Kritikpunkte rekurren ausschließlich auf semantische Aspekte – dadurch vermeidet White den Psychologismus Loars und öffnet seine Theorie der Kritik und Weiterentwicklung auf formaler Ebene.

White bietet eine Theorie an, die für jedes Wort, beziehungsweise für jeden Begriff zwischen individuellen und externen Komponenten unterscheidet. Den Schwerpunkt des begrifflichen Instrumentariums für diese Theorie entlehnt White aus [Kap89a]. Er übernimmt die Unterscheidung zwischen Äußerungskontexten und Auswertungssituationen sowie die Begriffe des Charakters und der Intension.

White faßt Putnams Experiment als Beweis der These von der versteckten Indexikalität auf, und zwar insofern, als daß auch bisher als nichtindexikalisch geltende Wörter wie „Wasser“ in dem Sinn als indexikalisch aufgefasst werden müssen, daß ihre Intension und Extension mit dem Kontext ihres *Erwerbs* variiert. (Wir erinnern uns: ich und mein Zwilling sind zu jedem Zeitpunkt unserer Existenz molekular identisch. Trotzdem meinen wir mit „Wasser“ völlig verschiedene Dinge, weil unsere Erwerbskontexte dieses Wortes – die jeweilige physikalische Beschaffenheit der Welt – sich unterscheiden. Ich habe gelernt, Wasser als „Wasser“ zu bezeichnen, er hat gelernt, XYZ als Wasser zu bezeichnen.)

White will daher eine formale Verbindung zwischen dem Erwerbskontext (um dessen Begriff er das Instrumentarium Kaplans erweitert) und dem Charakter eines Wortes, die anhand individualistischer Merkmale erklärt, wie diese versteckte Indexikalität möglich ist. Dazu jedoch ist es nötig, den Erwerbskontext genauer zu definieren (das definiendum ist im folgenden jeweils kursiv dargestellt):

**Definition (Möglicher Erwerbskontext):** Ein möglicher Erwerbskontext  $KE$  ist ein geordnetes Paar  $\langle w, S \rangle$  aus einer möglichen Welt  $w$  und einem funktionalen Zustand  $S$  (genauer: einer einzelnen Instanz eines funktionalen Zustands, das heißt, eines psychischen Zustands mit bekannter funktionaler Rolle), der in dieser Welt existiert und dessen kausale Geschichte durch die Beschaffenheit dieser Welt fixiert ist.

**Definition (Funktionale Äquivalenz von Erwerbskontexten):** Zwei mögliche Erwerbskontexte  $KE_1 = \langle w_1, S_1 \rangle$  und  $KE_2 = \langle w_2, S_2 \rangle$  sind *funktional äquivalent*, wenn die jeweils in ihnen enthaltenen funktionalen Zustandsinstanzen  $S_1$  und  $S_2$  typenidentisch sind.

**Definition (Erwerbskontext/Möglicher Erwerbskontext eines Worttyps):** Ein möglicher Erwerbskontext  $KE = \langle w_0, S \rangle$  (mit  $w_0 =$  unsere Welt) ist genau dann ein *Erwerbskontext des Worttyps  $W$* , wenn der in  $KE$  vorkommende funktionale Zustand  $S$  dem Gebrauch von  $W$  zugrunde liegt. Für  $KE = \langle$  mögliche Welt  $w_x, S \rangle$  ist  $KE$  ein *möglicher Erwerbskontext des Worttyps  $W$*  genau dann, wenn  $KE$  funktional äquivalent zu einem Erwerbskontext des Worttyps  $W$  ist.

**Definition (Interpretation des Worttyps  $W$ ):** Funktional äquivalente Erwerbskontexte des Worttyps  $W$  bilden Äquivalenzklassen. Eine *Interpretation  $I$  eines Worttyps  $W$*  ist ein Element der Menge dieser Äquivalenzklassen.

**Definition (Fixierung):** Ein möglicher Erwerbskontext  $KE = \langle \text{mögliche Welt } w_x, S \rangle$  eines Worttyps  $W$  fixiert den Charakter von  $W$  als  $C$  genau dann, wenn eine korrekte kausale Theorie der Referenz auf der Basis der Kausalbeziehungen zwischen  $S$  und seiner sozialen und physikalischen Umgebung in  $w_x$  festlegt, daß  $C$  der Charakter von  $W$  ist.

**Definition (Partieller Charakter):** Der *partielle Charakter*  $C_W$  eines Worttyps  $W$  relativ zu einer Interpretation  $I_W$  von  $W$  ist eine Funktion  $f : I \rightarrow \mathcal{C}$  von  $I$  in die Menge der Charakter  $\mathcal{C}$  derart, daß für  $KE \in I$  die Funktion  $f(KE) = C$  ist, derart, daß  $KE$  den Charakter von  $W$  als  $C$  fixiert.

Der partielle Charakter eines Worttyps  $W$  ist insofern partiell, als daß er auf eine einzelne Äquivalenzklasse  $I$  von Erwerbskontexten beschränkt bleibt und sich mit dieser verändert.

Ein wenig vereinfacht kann man sagen: Der partielle Charakter eines Worttyps  $W$  ist eine Funktion, die jedem einzelnen Erwerbskontext eines Worttyps  $W$  (aus einer gegebenen Klasse funktional äquivalenter Erwerbskontexte dieses Worttyps natürlich) genau den Charakter von  $W$  zuordnet, der durch diesen Erwerbskontext fixiert ist, das heißt, der durch eine kausale Theorie der Referenz (vgl. Kripke!) dem Worttyp  $W$  zugeordnet wird, und zwar nicht von ungefähr, sondern aufgrund der Kausalbeziehungen zwischen der möglichen Welt in  $KE$  und dem funktionalen (psychischen) Zustand im engen Sinne in  $KE$ . Das heißt klarerweise, daß der partielle Charakter in Relation zu einem gegebenen Erwerbskontext abbildet, wie die Beschaffenheit dieses Erwerbskontextes bestimmt, welcher Charakter dem in diesem Kontext erworbenen Wort  $W$  zugeordnet wird. Insofern hätten wir erstmalig im Verlauf dieser Abhandlung einen Bedeutungsbegriff, der semantische Merkmale mit kognitiven Merkmalen verbindet, das heißt der zugleich formal und subjektivitätskompetent ist. Whites Begriff des partiellen Charakters läßt es nämlich zu, anzugeben, *woher* Charakter kommen – genau das Merkmal, das wir bei Kaplan vermißt haben.

Die Unterscheidung zwischen einem Charakter eines Wortes und seinem partiellen Charakter kommt erst im Fall von Eigennamen und natürlichen Prädikaten wirklich deutlich zum Tragen: wenn ich zum Beispiel glaube, in diesem Glas sei Wasser so glaubt mein Zwilling auf der Zwillingserde ebenfalls, daß in dem Glas vor ihm Wasser sei – was immer Wasser ist.

Die Überzeugungssätze der beiden Hohenadels unterscheiden sich zunächst hinsichtlich ihrer Extension, denn es geht um unterschiedliche Gläser und Substanzen.

Sie unterscheiden sich auch in ihrer Intension, weil „Wasser“ nicht in beiden Welten dasselbe bedeutet; das heißt, unsere irdische Intension von „Wasser“ würde in einem Auswertungskontext, in dem innerhalb eines wahrheitswertfähigen Satzes „Wasser“ etwas anderes als  $H_2O$  bezeichnet, diesem Satz nicht den Wahrheitswert „wahr“ zuordnen.

Außerdem unterscheiden sich die beiden Sätze in ihrem Charakter, da sie sich bereits in ihrer Intension unterscheiden, also die Charaktere der beiden Sätze demselben Verwendungskontext unterschiedliche Intensionen zuordnen.

Worin sich jedoch beide Sätze gleichen, ist der partielle Charakter des Wortes „Wasser“:

Der Erwerbskontext meines Zwillings ist mit meinem funktional äquivalent, da der Zwerd-ling sich, wie Putnam bereits gezeigt hat, in demselben funktionalen Zustand  $S$  befindet wie

ich. Also gehören unsere Erwerbskontexte der gleichen Äquivalenzklasse an. Und der partielle Charakter des Wortes „Wasser“ ordnet – relativ zu der Äquivalenzklasse von Erwerbskontexten, der unsere beiden Interpretationen von „Wasser“ angehören – unseren Interpretationen von „Wasser“ entsprechend verschiedene Charaktere zu. Doch es ist deshalb nicht plausibel, anzunehmen, daß es zwei verschiedene partielle Charakter des Wortes „Wasser“ gebe. Vielmehr ist es dieselbe Funktion, die den funktional äquivalenten Erwerbskontexten verschiedene Funktionswerte in Form verschiedener Charakter zuordnet. Das heißt aber, daß in einem sehr starken Sinne das Teilen einer Überzeugung das Teilen des partiellen Charakters bedeutet. Hier zeigt sich die kognitive Qualität von Whites Begriff.

Selbst wenn nämlich der Zwerdling nun hierher käme und das Glas vor mir sähe, würde er etwas sagen wie „Ach, ich habe genauso eins und trinke auch immer Wasser daraus“. Doch es bliebe ein Satz über sein Glas in seiner Welt, dessen Inhalt für ihn mit dem Wort „Wasser“ bezeichnet wird und dem dort ein anderer Charakter zugeordnet ist als hier, während ich nur etwas über meine Welt und mein Glas Wasser glaube, und keine Ahnung habe, was für verrücktes Zeug er so trinkt. Das Ergebnis dieser Zuordnung geht dann aber auf dieselbe Funktion zurück.

Um nun funktionalistische Orientierungen verteidigen zu können, muß White zeigen, daß das Aufzeigen eines gemeinsamen partiellen Charakters mit dem Aufzeigen einer gemeinsamen Überzeugung in einem weiten Sinne bedeutungsgleich ist.

In gewisser Weise ist die von White entwickelte Darstellungsweise der Gemeinsamkeit semantischer Eigenschaften funktionaler Duplikate natürlich eine Gemeinsamkeit *per constructionem*. Die Legitimität solcher Konstruktionen bei funktionalen Duplikaten steht und fällt mit der Legitimität der Zuschreibung semantischer Ähnlichkeiten, die sich nicht auf funktionale Ähnlichkeiten reduzieren lassen. (Sonst müßten wir befürchten, die Ähnlichkeiten bestünden nur bei funktionalen Zwillingen, wie Block später zeigen wird.) Welche Motivation könnten wir haben, jenseits funktionaler Ähnlichkeiten Begriffe für solche semantischen Ähnlichkeiten einzuführen?

Ein gutes Beispiel sind die leeren Designatoren. Beispielsweise stehen wir noch immer vor dem aus Kapitel I ererbten Problem, zugeben zu müssen, daß Einhörner eigentlich dasselbe wie Greifen sind, weil beide Wörter dieselbe Extension haben: nichts. Auch eine Übertragung in die Mögliche-Welten-Semantik ändert daran nichts: Zwar können wir uns das übliche Duplikat unserer Welt vorstellen, in der unserer Vorfahren an Greifen und Einhörner glaubten und es sie auch tatsächlich *gab* – doch „Einhorn“ und „Greif“ sind starre Designatoren und das heißt, sie bezeichnen überall dasselbe: nichts. Also haben die beiden Wörter dieselbe Intension und die Mögliche-Welten-Semantik wird sie nicht unterscheiden können, denn es gibt keine mögliche Welt mit Einhörnern und Greifen darin, so einfach ist das.

Nehmen wir nun noch einmal die übliche Zwillingserde an, in der es Kreaturen gibt, die die Beschreibung erfüllen, die wir normalerweise von Einhörnern und Greifen geben würden. Nennen wir diese Tiere *Es* und *Gs*. Ein Einwohner dieser Welt wird also mit unserem Ausdruck „Einhorn“ eine andere Intension verbinden als mit dem Ausdruck „Greif“, sofern wir ihm diese Wörter erklärten. Das heißt aber *per definitionem*, daß wir auch *hier* in der wirklichen Welt einen Unterschied im partiellen Charakter dieser beiden Wörter ausmachen können, denn die mit diesen Wörtern jeweils assoziierten funktionalen Zustände  $S_1$  und  $S_2$  können sich unterscheiden: wenn der Sprecher diese Wörter in jener möglichen Welt

erworben hätte, dann würde  $S_1$  mit einer Intension korrelieren, die mögliche Welten der Klasse von *Es* zuordnet und  $S_2$  mit einer Intension, die mögliche Welten der Klasse von *Gs* zuordnet. Und das ergibt definitionsgemäß einen Unterschied im partiellen Charakter.

Was White damit gleistet hat, ist, etwas zu definieren, das es uns möglich macht, Unterscheidungen in unserer Welt zu postulieren, wenn es logisch mögliche Welten gibt, in denen die zugrundeliegenden Unterschiede existieren, eine Art „doppelter“ Modalität.

Außerdem hat White – und das ist interessant – mit dem partiellen Charakter ein extensionales Kriterium für die Übereinstimmung von Überzeugungszuständen geliefert. Damit entspricht seine Theorie in einem hohem Maße der in I.6 aufgestellten Forderung der Subjektivitätskompetenz der rekursiven Semantik, ein großer Vorteil, den kein anderer Inhaltstheoretiker (sofern man White als solchen interpretieren möchte) geliefert hat.

White ist historisch der erste, der Kaplans Konstruktion des Charakters in die Diskussion um eine Antwort auf Putnams Vorstoß bringt. White ist ebenfalls der erste, der die Notwendigkeit sieht, Putnam dadurch zu begegnen, daß inhaltliche Aspekte von Überzeugungssätzen als abhängig vom Kontext des Erwerbs der darin vorkommenden Wörter angesehen werden – einer Überlegung, die ursprünglich durch Putnam selbst inspiriert ist. White hat damit nicht nur zwei wichtige Aspekte der internalistischen Verteidigungsstrategie vorgezeichnet, die sich nicht mehr ändern werden, sondern die Sauberkeit der Argumentation auch auf ein Niveau gehoben, das von den wesentlich breiter diskutierten Thesen Fodors wieder unterschritten wird<sup>4</sup>. Seine Antwort auf Putnam zeichnet sich desweiteren dadurch aus, daß er das Twin-Earth-Experiment nicht zurückweist, sondern seine Schwächen gezielt zu beheben versucht, indem zwischen internen und externen Elementen sauber unterschieden wird. Außerdem leistet White als erster die Konstruktion eines gleichermaßen subjektivitätskompetenten und formal präzisen Bedeutungsbegriffs.

Whites Begriff des partiellen Charakters entspricht den Anforderungen, die Loar an den engen Inhalt stellt, weit besser, als Loars eigener Begriff: der partielle Charakter kann von verschiedenen Personen geteilt werden, er bestimmt die intuitive Bedeutung (den Charakter) der Wörter, ohne die Extension einfach festzulegen (das vermeidet White durch die raffinierte Kontextabhängigkeit) und er kann sprachlich approximiert werden, ohne auf pragmatische Komponenten expliziten Bezug nehmen zu müssen – und, was sehr wichtig ist, er bildet Intentionalität ab. Dafür erbt Whites Charakterbegriff die Schwierigkeiten von Kaplans Begriff, Dinge wie einen „Kontext“ formal definieren zu müssen. Verglichen mit Loars Problemen sind das allerdings geringe Schwierigkeiten<sup>5</sup>.

---

<sup>4</sup>Ärgerlich ist, daß sehr viele Unklarheiten in der folgenden Diskussion problemlos hätten vermieden werden können, wenn White aufmerksamer gelesen worden wäre. Die Vernachlässigung seiner präzisen Begrifflichkeiten und die Ignoranz, die diesem durchdachten Ansatz entgegengebracht wurde, fordert die Ungehaltenheit des interpretierenden Philosophen besonders heraus, wenn man auf die vielen begrifflichen Schlamperien stößt, durch die die Diskussion um engen Inhalt sich sonst auszeichnet. Der Begriff des partiellen Charakters birgt viele Möglichkeiten, auf die hier gar nicht eingegangen werden kann.

<sup>5</sup>Historisch wurde White von nahezu der gesamten Inhaltsdiskussion übersehen, die vermutlich ganz anders – nämlich wesentlich kürzer und transparenter – verlaufen wäre, wenn Fodor, Block, Dennett und Loar den *Pacific Philosophical Quarterly* aufmerksamer gelesen hätten.

## V.4 Wie man Einstellungen doch individualistisch verstehen könnte: Fodor

Die Position des quasi „klassisch“ physikalistisch orientierten Internalisten nimmt Jerry Fodor in [Fo87] ein. Dieses Buch ist eine Art großangelegter Apologie des Internalismus, ohne allerdings mit wirklich stichhaltigen Ansätzen zu der bekannten Problematik aufzuwarten, die sich mit Putnams Behandlung von P1 verbindet. Fodor will P1 retten, indem er erst die Unanwendbarkeit des Twin-Earth-Szenarios auf die wissenschaftliche Psychologie aufzeigt: er versteht Putnams Experiment als Beweis, daß die Standards, nach denen propositionale Einstellungen vom common sense identifiziert werden, für die wissenschaftliche Psychologie nicht hinreichend sind. Also könne die Psychologie es sich leisten, sofern sie eigene Standards der Identifizierung von Einstellungen habe, das Twin-Earth-Problem zu ignorieren. Dennoch geht Fodor dann zu einer Enger-Inhalt-Strategie über, auf die ich auch zu sprechen kommen werde und die der Loars ähnelt, die sie mit einer formalen Konstruktion des engen Inhalts zu verknüpfen versucht, die ihrerseits der von White entfernt ähnelt.

Fodors Argumenten wird trotz ihres problematischen Charakters deshalb eine so breite Darstellung gewidmet, weil sie lange Zeit die einzige dauerhaft vorgetragene und breit diskutierte Antwort auf Putnam waren. (Erst in Kapitel VI, bei der Besprechung von Stalnakers Analyse, kommt es zu einer echten „Aufarbeitung“ dieser Argumente.)

Fodor läutet seine Argumentation mit einer breit angelegten Verteidigung des psychologischen common sense ein, die hier wiederzugeben ohne Witz wäre. Da die Psychologie des common sense jedoch eine wichtige Rolle in der Internalismus/Externalismus-Debatte spielt und schon im vorangegangenen Kapitel in den Rang einer Fußnote abgedrängt wurde, will ich sie hier ein wenig erläutern. Unter „Commonsense-Belief/Desire-Psychologie“ versteht die analytische Philosophie des Geistes etwas, das man auch als Alltagspsychologie oder Alltagsrationalität bezeichnen könnte. Ich habe sie im letzten Absatz verkürzend als „common sense“ bezeichnet und bei dieser Bezeichnung werde ich auch bleiben. Der common sense ist jene Art Psychologie, die wir im Alltag anwenden, wenn es darum geht, auf Einstellungen oder Handlungen von Personen zu schließen oder deren Verhalten zu verstehen, und die beschreibt, wie wir normalerweise mit Dingen, Situationen und mit anderen Menschen umgehen, bei denen wir die gleiche Art common sense-Rationalität voraussetzen. Das Vokabular des common sense umfaßt die bisher verwendete Rede von propositionalen Einstellungen, Überzeugungen und Wünschen. Ein Beispiel:

Betty denkt, daß es ihr Spaß macht, an den Strand zu gehen, wenn es heiß und sonnig ist. (Unnötig zu sagen: Betty ist froh, wenn sie Spaß hat und Betty möchte froh sein.) Also wünscht Betty sich, daß es morgen heiß und sonnig sein möge.

Das Muster aus diesem ganz einfachen Beispiel, in dem ein Wunsch direkt aus Überzeugungen resultiert, ist im Prinzip schon alles, was man braucht, um den common sense erfolgreich anwenden zu können. Ein anderes Schlußmuster ist zum Beispiel: Wenn  $x$  will, daß  $P$ , und  $x$  glaubt, daß  $\neg P$  bis  $Q$ , und  $x$  glaubt, daß  $Q$  herbeiführbar ist, dann wird  $x$  versuchen,  $Q$  herbeizuführen.

Die Philosophie des Geistes geht, wie wir gesehen haben, davon aus, daß propositionale Einstellungen durch psychische Zustände (hier jetzt im weiten Sinne) festgelegt sind. Die-

se psychischen Zustände „supervenieren“ dabei auf physischen Gehirnzuständen, das heißt, daß bei jeder Änderung von psychischen Zuständen sich auch eine Änderung von Gehirnzuständen zeigen muß (cf. [Fo87] p. 30). Aufgrund dieser letzten These ist die Konzeption der common-sense-Rationalität auch von Putnams Angriff auf Annahme P1 akut betroffen. Fodors gesamtes Konzept der „Psychosemantik“ ist jedoch von einem Begriff des common sense abhängig, der die Geist-Gehirn-Supervenienz einschließt. Er muß also beides gegen Putnam verteidigen.

Fodor weigert sich daher, aus den Experimenten von Putnam und Burge die Konsequenz zu ziehen, daß Bedeutungen, also auch propositionale Einstellungen, *nicht* im Sinne einer Geist-Gehirn-Supervenienz auf Gehirnzuständen beruhen könnten. Putnam hat gezeigt, daß zwei Gehirne, die auf molekularer Ebene identisch sind, dennoch unterschiedliche propositionale Einstellungen entwickeln können, ohne diese Identität zu verlieren. Fodor gibt dies zwar zu, doch er wendet ein, daß Putnam die propositionalen Einstellungen seiner „Probandenpaare“ (wie „Wasser ist naß“, „Wasser löscht Durst“ und sofort) nicht auf eine Weise identifiziert, die für die Anforderungen wissenschaftlicher Psychologie ausreichend ist, sondern lediglich auf dem common sense beruhen. Daher zeige das Experiment auch nur, daß die Standards, die der common sense zur Identifizierung [*individuation*] propositionaler Einstellungen heranzieht, nicht für eine wissenschaftlich relevante Behandlung ausreichen. Ergo kann keine Aussage über Gehirnzustände in Putnams Schlußfolgerung involviert werden und ergo ist P1 nicht wirklich widerlegt. Soweit die Route.

Das Hauptproblem des Twin-Earth-Experiments besteht nach Fodor darin, daß in diesem Gedankenexperiment die propositionalen Einstellungen der Probanden *relational* identifiziert werden, was das Vorgehen des common sense kennzeichnet: Putnam sagt, mein zwirrdischer Zwilling glaube *deshalb* nicht wirklich, „daß Wasser naß ist“, weil das, von dem er glaubt, daß es naß sei, kein Wasser (=Wasser<sub>E</sub>) ist. Das heißt, seine Einstellung wird in Abhängigkeit von dem Gegenstand Wasser<sub>E</sub> identifiziert. Wir verwenden diesen Gegenstand, um die Identitätskriterien propositionaler Einstellungen wie „Wasser ist naß“ anzugeben. Insofern ist diese Art der Identifizierung von Einstellungen relational, denn sie ist von einer Relation dieses daß-Satzes, also ihres weiten Inhaltes zur wirklichen Welt abhängig. Außerdem ist sie nichtinternalistisch, denn Wasser hat nichts mit meinem Gehirn oder meinen intrinsischen Eigenschaften zu tun. Eine internalistische Identifizierung meiner Einstellung läge dann vor, wenn die Identifizierungsmethode meine Einstellung nur aufgrund meiner inneren Konfigurationen identifiziert. Wie würde so etwas aussehen?

Eine nichtrelationale, internalistische<sup>6</sup>

Identifikation der propositionalen Einstellung „Wasser ist naß“ sieht den Inhalt dieser Überzeugung meines Zwillings und den Inhalt meiner eigenen Überzeugung, „daß Wasser naß sei“, als identisch (genauer: als typenidentisch) an, weil wir beide über das, was wir „Wasser“ nennen, *dasselbe glauben*. Da wir Wasser<sub>E</sub> und Wasser<sub>TE</sub> nur durch wissenschaftlichen Aufwand unterscheiden können, kann ohne unser Wissen um diesen Unterschied, unser beider Begriff von „Wasser“ sich nicht unterscheiden, weswegen auch die kausalen

---

<sup>6</sup>Die Ausdrücke „nichtrelational“ und „individualistisch“ bzw. „internalistisch“ sind nicht synonym, auch wenn sie de facto oft zusammenfallen. Erwägungen, in denen dieser Unterschied eine Rolle spielt, kommen bei Fodor zwar vor, werden hier aber nicht erörtert. Ich verwende die Ausdrücke daher austauschbar.



Eigenschaften unserer Wasser-Überzeugungen (das Verhalten, das sie auslösen, und die Art der Verknüpfung mit anderen Einstellungen und Sinnesdaten) sich nicht unterscheiden können. Da die Psychologie propositionale Einstellungen anhand ihrer kausalen Eigenschaften identifiziert, kann sie die Einstellungen von mir und meinem Zwilling als typenidentisch betrachten und dann liegt kein Verstoß mehr gegen P1 vor: wir haben typenidentische psychische Zustände (im weiten Sinne), die mit Einstellungen korrelieren, die die Psychologie ebenfalls als typenidentisch betrachtet.

Die Folgerung daraus ist, daß die Psychologie nicht die relationalen Standards der Identifikation propositionaler Einstellungen übernehmen darf, die der common sense verwendet, vielmehr muß sie eigene, neue Identitätskriterien für propositionale Einstellungen formulieren. Das aber ist kein prinzipielles Problem, denn die Neuformulierung solcher Kriterien berührt nicht die Grundlage, auf der die Gemeinsamkeit zwischen wissenschaftlicher Psychologie und common sense sich gründet, nämlich darauf, daß psychischen Zuständen (im weiten Sinne) sowohl kausale Verursachungskraft als auch semantischer Gehalt zugeschrieben werden kann. Außerdem sind die erforderlichen Identifizierungsstandards nach Fodor durchaus formulierbar. Darauf werden wir zurückkommen.

Doch wieso sind die kausalen Eigenschaften der propositionalen Einstellungen und nicht deren relationale Eigenschaften für die Angabe ihrer Identitätskriterien wichtig?

Weil die Psychologie, wie jede Wissenschaft, ihre Gegenstände anhand von deren *Kausaleigenschaften* klassifiziert. Das heißt, Gehirnzustände ebenso wie psychische und selbst molekulare Zustände müssen durch kausale Eigenschaften identifiziert werden. Die kausalen Eigenschaften einer propositionalen Einstellung bestehen in der Art, wie sie durch welche Sinnesdaten beeinflusst wird, wie sie auf andere propositionale Einstellungen wirkt und von diesen beeinflusst wird und wie sie sich auf das Verhalten auswirkt. Die Gesamtheit dieser drei Arten von Kausaleigenschaften habe ich als die funktionale Rolle einer propositionalen Einstellung bezeichnet. Es ist leicht zu sehen, daß in der funktionalen Rolle einer propositionalen Einstellung keine relationalen Eigenschaften vorkommen.

Warum? Weil Eigenschaften wie die, ein psychischer Zustand (im weiten Sinne) einer Person zu sein, die in einer Welt lebt, wo es statt  $H_2O$  nur XYZ gibt (und alle ähnlichen relationalen Eigenschaften psychischer, physiologischer und molekularer Zustände), irrelevant für die funktionale Rolle dieses Zustandes sind und ergo irrelevant für jede wissenschaftliche Erklärung oder Taxonomie. Das heißt, die funktionalen Rollen der propositionalen Einstellungen von uns Zwillingen unterscheiden sich nicht, also muß die Psychologie auch unsere propositionalen Einstellungen nicht unterscheiden, obwohl sie unterschiedliche Bedeutungsvektoren enthalten.

Doch sind die funktionalen Rollen unserer Einstellungen wirklich gleich? Es gibt hier zwei Einwände.

**Einwand 1:** Wenn ich um Wasser bitte, dann gibt es, je nachdem, wo ich bin, höchst unterschiedliche Effekte: hier bekomme ich  $H_2O$ , dort bekomme ich XYZ. Das sind verschiedene Ereignisse, ergo verschiedene Verhaltensweisen.

Aber: die Identität funktionaler Rollen darf nicht *zwischen* verschiedenen Kontexten überprüft werden, sondern muß *in* verschiedenen Kontexten gelten. Relevant in diesem Zusam-

menhang ist nicht, daß dieselbe Äußerung (oder derselbe Gedanke) in verschiedenen Kontexten Verschiedenes bewirkt: für die Identität der funktionalen Rollen unserer Einstellungen ist entscheidend, daß unser beider Äußerung sowohl in meinem wie in seinem Kontext geäußert, jeweils identische Effekte haben. Würde mein Zwilling nämlich in *meinem* Äußerungskontext hier auf der Erde sagen, daß er Wasser will, dann würde er Wasser<sub>E</sub> bekommen. Die Effekte beider Äußerungen können in verschiedenen Kontexten verschieden ausfallen, sie dürfen sich nur nicht innerhalb *desselben* Kontexts voneinander unterscheiden, sofern wir davon ausgehen wollen, daß sie identische funktionale Rollen haben. Doch innerhalb desselben Kontexts unterscheiden sie sich ja auch nicht: Auch wenn ich zu meinem zwirdischen Besucher sage, er möge „Wasser“ bringen, wird er mir bringen, was aus dem Hahn kommt. Und das ist H<sub>2</sub>O.

**Einwand 2:** Wenn man Einstellungen intentional beschreibt und mit Putnam und Kripke davon ausgeht, daß XYZ nicht Wasser<sub>E</sub> ist, dann kann der Zwerdling – genau wie der kontrafaktische Sprecher in Stufe 2 des Burge-Experiments – *gar keine* Wasser<sub>E</sub>-Einstellungen *haben* (höchstens „Wasser“-Einstellungen). Ich jedoch habe eine Menge davon: Ich denke, daß Wasser<sub>E</sub> den Durst löscht, daß man darin schwimmen kann, daß eine kalte Dusche morgens erfrischend sein kann undsofort. Doch mein Zwilling denkt nichts davon. Er benutzt denselben Ausdruck, doch seine Überzeugungen beziehen sich nicht auf H<sub>2</sub>O, ergo nicht auf Wasser<sub>E</sub>. Also könnte man sagen, daß unsere propositionalen Einstellungen, das heißt unsere psychischen Zustände (im weiten Sinne) sich unterscheiden.

Allerdings: wenn ich zugebe, daß die psychischen Zustände der Zwillinge sich unterscheiden, wird es schwierig, zu behaupten, daß ihre Gehirnzustände dieselben seien. Beide Gehirnzustände bringen dieselben verbalen Äußerungen hervor und wenn diese Äußerungen aus den eben genannten Gründen als distinkt gelten sollen, dann *unterscheiden* sich unsere Verhaltensweisen, weswegen unsere Gehirnzustände unterschiedliche Kausaleffekte haben und nicht typenidentisch sein können. Dann kann ich sie aber auch als Wissenschaftler guten Gewissens unterscheiden, ohne gegen P1 zu verstoßen.

Wir können also davon ausgehen, daß die funktionalen Rollen unserer propositionalen Einstellungen gleich sind, obwohl die darin vorkommenden Bedeutungsvektoren von „Wasser“ sich durch ihre Extension unterscheiden.

Zusammenfassend läßt sich sagen: der Fehler Putnams ist es nach Fodor gewesen, relational definierte psychische Zustände so zu behandeln, als sei ihre Unterscheidung für die Psychologie relevant, die psychische Zustände nur anhand nichtrelationaler, kausaler Eigenschaften unterscheidet. Putnam unterscheidet die Bedeutungsvektoren der beiden „Wasser“-Begriffe, doch diese Vektoren sind Teil von propositionalen Einstellungen, die identische funktionale Rollen haben, also von der Psychologie nicht unterschieden werden. Also hat Putnam nichts gezeigt, was für die Psychologie relevant ist. Sie kann weiterhin von einem internalistischen Programm zur Beschreibung dessen, was im Kopf ist, ausgehen und kann die Bedeutungsvektoren ignorieren. Das Twin-Earth-Experiment unterbricht also die (kausale) Verbindung zwischen extensionaler Identität und inhaltlicher Identität der Gedanken nicht wirklich: sie relativiert sie nur in Hinsicht auf den entsprechenden Kontext.

Im Grunde ist dies schon die ganze Substanz von Fodors Hauptargument gegen Putnam. Es gibt noch ein Nebenargument:

Man kann zwar die relationalen Eigenschaften einer propositionalen Einstellung (und auch anderer Dinge) durch alles mögliche beeinflussen, auch *per definitionem*. Um von einem Ding jedoch sagen zu können, daß es die *kausalen* Eigenschaften eines anderen beeinflusst, muß man irgendein Gesetz oder einen Mechanismus entdecken können, der für diese Beeinflussung verantwortlich ist. Wenn ich also (Fodors Hauptargument ignorierend) sagen würde, daß die beiden Einstellungen unterschiedliche funktionale Rollen haben, sich also signifikant unterscheiden, dann wohl nur aufgrund der unterschiedlichen Bedeutungsvektoren. Diese werden aber laut Putnam von der Sprachgemeinschaft und der physikalischen Beschaffenheit der Welt festgelegt. Das heißt, die Sprachgemeinschaft und die physikalische Beschaffenheit der Welt verursachen mittelbar einen Unterschied in den Kausaleigenschaften meiner „Wasser“-Einstellung und der „Wasser“-Einstellung meines Zwilling, also unserer psychischen Zustände (im engen Sinne). Es ist jedoch plausibel, daß die einzigen Faktoren, die eine Beeinflussung psychischer Zustände durch Umwelteffekte möglich machen, neurologischer Natur sind. Im Twin-Earth-Experiment ist jedoch nicht klar, *wie* die kausalen Eigenschaften der psychischen Zustände eines Sprechers von der Welt oder der Sprachgemeinschaft, die beide als identisch vorausgesetzt werden, beeinflusst werden könnten, *ohne* dabei die Physis des Sprechers zu involvieren. Würde Putnam eine Beeinflussung durch die Physis (also neurologisch) zulassen, dann läge kein Verstoß gegen P1 vor, zumindest aber kein Verstoß gegen die Geist/Gehirn-Supervenienz.

Es ist also nicht gezeigt, wie die Kausaleigenschaften von Wasser<sub>E</sub>-Einstellungen und Wasser<sub>TE</sub>-Einstellungen sich unterscheiden können, wo doch das, wodurch sie beeinflusst sein sollen, nämlich die Umgebung des Sprechers, die entsprechenden Kausaleigenschaften anders als durch dessen Physis beeinflussen soll. Die Art und Weise dieser Beeinflussung von Bedeutungen durch die Umwelt bleibt bei Putnam also unklar.

Auch diese Aporie kann man vermeiden, indem man die psychischen Zustände der Sprecher innerhalb eines *wissenschaftlichen* Ansatzes als gleich ansieht.

## V.5 Fodors Zuordnungstheorie des engen Inhalts und ihre Herleitung

Fodor hat nun die Unterschiede zwischen den Identifizierungsstandards für propositionale Einstellungen aufgezeigt, die zwischen dem common sense und der wissenschaftlichen Psychologie bestehen. Die Standards des common sense sind, wie wir gesehen haben, für die Psychologie unbrauchbar.

Nun ist es aber laut Fodor andererseits auch so, daß der common sense eine gute Grundlage für die psychologische Bewertung von Verhalten liefert. Es wäre sehr problematisch zu behaupten, daß eine wissenschaftliche Psychologie Identitätskriterien für propositionale Einstellungen verwendet, die dem common sense widersprechen. Vielmehr soll die Psychologie den common sense wenigstens in den meisten Fällen bestätigen (einzelne Abweichungen mag es freilich geben). Die Psychologie bestätigt den common sense am wirksamsten dann, wenn sie den psychischen Zuständen (im weiten Sinne) Eigenschaften zuspricht, die der common sense auch den propositionalen Einstellungen in seinem Sinne zuspricht.

Dies ist der Punkt, an dem Fodors Ausführungen problematisch werden und die logische Harmonie des letzten Abschnitts verlassen. Im nächsten Abschnitt werden wir sehen, was genau an dem Punkt merkwürdig ist, an dem wir uns jetzt befinden. Wir müssen uns merken, daß Fodor die Forderung nach einer Unterstützung des common sense durch die Psychologie aufstellt. In diesem Abschnitt geht es darum, zu zeigen, wie diese Forderung bei ihm in den Begriff des engen Inhalts mündet.

Wie nun kann man diese Forderung einlösen? Das Problem ist bekanntlich, daß der common sense die Einstellungen der Zwillinge in Putnams Experiment als verschieden betrachtet, denn meine Einstellung bezieht sich auf Wasser<sub>E</sub>, die meines Zwillings auf Wasser<sub>TE</sub>. Wenn der common sense den propositionalen Inhalt zur semantischen Auswertung heranzieht, wird er immer Relationalität in seine Identifikationsmethoden einschleppen und die ist, wie wir im letzten Abschnitt gesehen haben, in der psychologischen Identifikation von Einstellungszuständen verboten. Der Psychologe betrachtet daher die Einstellungszustände als identisch, weil sie identische funktionale Rollen haben. Folglich darf die Psychologie psychische Zustände (im weiten Sinne) nicht wie der common sense in Begriffen des propositionalen Inhalts identifizieren, sie benötigt hierfür Standards, die die Geist/Gehirn-Supervenienz berücksichtigen und internalistisch sind. Eigenschaften wie „eine Überzeugung sein, daß Wasser naß ist“, ist für Kausalerklärungen taxonomisch irrelevant.

Doch – und hier setzt nun Fodors neue Argumentation ein – die Psychologie *muß* psychische Zustände (im weiten Sinne) auch durch ihren *Inhalt* kennzeichnen, denn psychische Zustände sind essentiell durch ihre Inhalte bestimmt, so daß die Identität der ersteren die Identität der letzteren garantiert. Wir sind damit aus anderer Richtung wieder bei dem Problem angelangt, das Loar schon hatte: wie kann der Inhalt eines psychischen Zustandes (beziehungsweise: Gedankens) diesen essentiell identifizieren, wenn er nicht auf dem propositionalen Gehalt des Gedankens supervenieren darf?

Fodors Antwort lautet: Der erforderliche internalistische Begriff des Inhalts ist nicht der des weiten Inhalts, der mit dem propositionalen Inhalt beziehungsweise den Wahrheitsbedingungen des Gedankens zusammenfällt, sondern ein „enger Inhalt“, dessen Konzeption (auch) Fodor an Kaplans Begriff des Charakters anlehnt. Ein enger Inhalt eines Gedankens ist etwas, das einen weiten Inhalt in Relation zu einem Kontext festlegt, also eine Funktion von Erwerbskontexten<sup>7</sup> in weite Inhalte/Wahrheitsbedingungen (das haben wir ähnlich bei White schon ausgeführt gesehen, nur weniger tendenziös und technisch sauberer). Diese Fixierung von Inhalten auf Kontexte ist es auch, was das Twin-Earth-Experiment ausdrückt:

„For even if all the intuitions about Twin-Earth are right, and even if they have all the implications that they are said to have, extensional identity still constrains intentional identity because contents still determine extensions relative to a context.“ [[Fo87] p. 47]

---

<sup>7</sup>Fodor ist in seinem Sprachgebrauch hier sehr unpräzise: er erklärt nicht, ob er Äußerungs- oder Erwerbskontexte meint. (Gehen wir davon aus, daß Auswertungssituationen ohnehin nicht in Frage kommen.) Die Art, wie er den Begriff „Kontext“ verwendet, legt nahe, daß es sich nicht um Äußerungskontexte handeln kann, sondern daß Erwerbskontexte wie bei White gemeint sind. Durch diese fehlende Festlegung jedoch werden viele Schärpen, die seine Argumentation sonst hätte, ausgewaschen. Bei der Lektüre von [Fo87] scheint es fast, als habe er [Whi82] gar nicht gelesen, obwohl Elemente seiner Auffassung White stark verpflichtet sind, cf. Abschnitt 1.

Was Fodor damit formuliert hat, ist ein Kriterium der Identität, mit dessen Hilfe die Psychologie psychische Zustände (im weiten Sinne) beziehungsweise Gedanken so identifizieren können soll, daß sie die Zustände der Zwillinge in Übereinstimmung mit dem common sense als inhaltlich identisch ansehen kann, ohne einen Verstoß gegen die Geist/Gehirn-Supervenienz hinnehmen zu müssen:

„Two thought contents are identical only if they effect the same mapping of thoughts and contexts onto truth conditions. Specifically, your thought is content-identical to mine only if in every context in which your thought has truth condition *T*, mine has truth condition *T* and vice versa.“ [[Fo87] p. 48]

Damit versucht Fodor die oben gestellte Forderung einzulösen. Enger Inhalt kann Intentionalität abbilden: die Wasser-Einstellungen meines Zwillinges und meine sind intentional identisch, denn sie haben denselben Inhalt, obwohl die Kontexte sich hinsichtlich der Wahrheitsbedingungen unserer Gedanken unterscheiden, wie auch bei Loar. Auf diese Art läßt sich ein extensionales Kriterium für den engen Inhalt formulieren. Der weite Inhalt dagegen ist das, was man in Form eines Satzes angeben kann: er besteht aus einem spezifizierten engen Inhalt und einem fixierten Kontext. Dann können wir sagen, daß ich und mein Zwilling in Überzeugungszuständen mit demselben engen Inhalt sind, doch da unsere Kontexte sich unterscheiden, unterscheiden sich unsere weiten Inhalte.

In Whites Terminologie würde das heißen, daß die Zwillinge sich denselben partiellen Charakter des Wortes „Wasser“ teilen, doch daß ihre Überzeugungen sich hinsichtlich der Kaplanschen Charakter unterscheiden. Whites extensionales Kriterium ist formal viel transparenter ausgeführt als das von Fodor und berücksichtigt Intentionalität ebenso. (Außerdem ist der Begriff des partiellen Charakters ein „echter“ Bedeutungsbegriff, weil seine Wertemenge aus „echten“ Bedeutungsbegriffen, nämlich Kaplanschen Charakteren besteht. Fodors Begriff dagegen ist in diesem Sinne keine „echte“ Bedeutung, solange es nur um Wahrheitsbedingungen geht. Für den engen Inhalt von Sätzen könnte Fodors Begriff funktionieren, doch der enge Inhalt von Wörtern bedarf einer viel genaueren Explikation.)

Fodor hat nun den eingangs geforderten Standard der Identifikation für propositionale Einstellungen, der sowohl internalistisch ist als auch Inhalt hat. Mit dem engen Inhalt als Standard soll es wissenschaftlicher Psychologie gelingen, den common sense in der Identifizierung propositionaler Einstellungen zu stützen statt ihn zu widerlegen.

Warum nun nenne ich das eine Zuordnungstheorie des engen Inhalts? Fodors Theorie fußt – im Gegensatz zu Loar – auf einem Begriff des engen Inhalts als einer Funktion, mithin auf der *Zuordnung* von weiten Inhalten zu Kontexten. (In diesem Sinne war auch White ein Zuordnungstheoretiker. Es gibt, wie wir gesehen haben, andere Theorien des engen Inhalts, wie zum Beispiel die von Loar, von denen Inhalt nicht als Funktion aufgefaßt wird.)

Wir müssen also, folgen wir Fodor, folgendes ablehnen, was Putnams Ausführungen vorschlagen:

- 1) Die Extension eines Ausdrucks (Begriffs, Gedankens) ist keine unabhängige Komponente des Bedeutungsvektors.
- 2) Innerhalb der Psychologie kann der Begriff des Inhalts nicht durch den Begriff des Stereotyps ersetzt werden.

Forderung 1) kann deshalb nicht erfüllt werden, weil Fodor gezeigt hat, daß enge Inhalte (die man jetzt sozusagen psychologisch als „Bedeutung“ ansehen könnte) bei gleichbleibendem Kontext sehr wohl eine Extension festlegen. Und daraus folgt auch gleich die Ablehnung von 2), denn Stereotypen legen gar nichts fest und können darum Inhalt nicht ersetzen.

Doch da stimmt etwas nicht: im Zusammenhang mit Loar haben wir gesehen, daß der enge Inhalt eines Ausdrucks dessen Bedeutung nicht bestimmen darf. Bei Fodor geht das plötzlich. Warum?

Weil Fodor die Festlegung der Inhalte durch die Extensionen nur relativ zu einem gegebenen Kontext gelten läßt. Putnam dagegen, der von Kripkes Lehre der starren Designation ausgeht, legt die Bedeutung eines Ausdrucks durch seine Bedeutung in *dieser* Welt für alle möglichen Welten fest. Er schafft daher Vergleichbarkeit zwischen Kontexten, die Fodor verneint. Bei gleichem Kontext gilt nach Fodor, daß die Inhalte sich unterscheiden, sofern sich die Extensionen unterscheiden. Das heißt aber, daß das Twin-Earth-Experiment in dieser Hinsicht ungefährlich ist. Wir stehen nämlich nicht – wie Loar – unter dem Zwang, einerseits behaupten zu müssen, daß die Identität des Inhalts in jedem Fall Identität der Extension bedeute, andererseits aber auf der Identität der gedanklichen Inhalte der Putnam-Zwillinge bestehen zu müssen, denn deren Kontexte unterscheiden sich. Dies ist ein wichtiger Unterschied zwischen Fodor und Loar, der sich erklärt, wenn man bedenkt, daß Loar strikter als Fodor die propositionale Festlegung von Einstellungen generell in Zweifel zieht. Fodors ganze Argumentation greift also nur, wenn wir uns nicht schon im Vorfeld auf Loar einlassen.

Nun steht man mit der Einführung des engen Inhalts aber dennoch vor einem Problem, das auch Loar hatte: wie sieht das, was an den Gedanken der Zwillinge gleich ist, denn nun wirklich aus? Kann man in einem Satz sagen, welchen Inhalt sie teilen? Natürlich nicht, denn was ein Satz ausdrückt, ist niemals enger Inhalt, weil es keine „enge“ Sprache gibt. Sätze drücken immer zuvorderst weiten Inhalt aus. Das größte Problem mit dem engen Inhalt ist hier dasselbe wie bei Loar: er ist nicht semantisch auswertbar und kann es auch nicht sein, weswegen er eigentlich auch gar kein Inhalt im üblichen Sinne ist, obwohl das durch seine Definition als Funktionsbegriff zunächst so aussieht:

„Indeed, if you mean by content what can be semantically evaluated, then what my water-thoughts share with Twin ‘water’-thoughts isn’t content. Narrow content is radically inexpressible, because it’s only content potentially; it’s what gets to be content when – and only when – it gets to be anchored. We can’t [...] say what Twin thoughts have in common. This is because what can be said is ipso facto semantically evaluable; and what Twin-thoughts have in common is ipso facto not. [...] You can’t, in absolute strictness, express narrow content; but [...] there are ways of sneaking up on it.“ [[Fo87] p. 50]

Was ich im weiteren Sinne mit meinem Zwilling teile, ist ein psychischer Zustand (im engen Sinne) und psychische Zustände sind in Relation zu einem gegebenen Kontext sehr wohl semantisch auswertbar. Wenn wir von gewissen hypothetischen semantischen Eigenschaften der engen Inhalte ausgehen, dann können wir – wie Loar – sprachliche Ausdrücke benutzen, um diese engen Inhalte zu approximieren. Und insofern kann man sagen, daß der Ausdruck „der Gedanke, daß Wasser naß ist“ etwas beschreibt, das mein Zwilling und ich gemeinsam haben.

Natürlich ist dies fast ein „inhaltsloser“ Begriff des engen Inhalts. Das einzige, was Fodor wirklich über den engen Inhalt sagen kann, ist, daß er eine Funktion von Kontexten in Wahrheitsbedingungen, also in weite Inhalte sein soll, so daß die engen Inhalte sich im Hinblick auf die Kontexte unterscheiden lassen. Das allein ist jedoch für eine schlüssige Erklärung des engen Inhalts zuwenig. Präzisierende Begrifflichkeiten, die das ganze formal unterfüttern, wie bei White, liefert Fodor nicht.

In diesem Abschnitt haben wir gesehen, daß Fodor mit dem Begriff des engen Inhalts zwar seine Forderung einlöst, die Psychologie solle den common sense unterstützen, indem sie Einstellungen inhaltlich identifiziert – im nächsten Abschnitt werden wir dafür sehen, daß er diese Forderung gar nicht hätte stellen müssen.

## V.6 Warum enger Inhalt überflüssig ist: Fodor in der Kritik von Schiffer

In [Sch90] findet sich eine Argumentation, die die Überflüssigkeit des engen Inhalts, wie Fodor ihn sich vorstellt, behauptet.

Fodor argumentiert für die mentale Gleichheit von Überzeugungszuständen wie folgt:

Meine Überzeugung „Wasser ist naß“ ist insofern typenidentisch mit der Überzeugung meines Zwillings „Wasser ist naß“, als daß sie denselben „mentalen Satz“ instanziiert, nämlich „Wasser ist naß“. In meinem inneren, mentalen Code bedeutet dieser Satz, daß Wasser<sub>E</sub> naß ist, in dem formal identischen mentalen Code meines Zwillings bedeutet er, daß Wasser<sub>TE</sub> naß ist.

Außerdem besitzen diese beiden Instanzen des neuralen Satzes in beiden Köpfen dieselbe funktionale Rolle. Diese funktionale Rolle des Satzes soll von einer Form wissenschaftlicher Psychologie auswertbar sein. Die kognitive Psychologie, die Fodor anstrebt, ist insofern „syntaktisch“, als daß sie ohne direkten Bezug zur eigentlichen Bedeutung mentaler Ausdrücke arbeitet, das heißt, sie ordnet mentalen Ausdrücken funktionale Rollen zu, ohne die semantischen Eigenschaften dieser Ausdrücke im Kopf der Sprecher in diese funktionale Identifizierung einzubeziehen. Semantik ist also nur vermittels Syntax im Kopf.

Die beiden Überzeugungen unterscheiden sich allerdings durch ihren propositionalen Inhalt. Fodor behauptet, daß der common sense propositionale Einstellungen relational identifiziert, also anhand ihres weiten Inhalts. Das heißt aber, daß diejenigen Eigenschaften von Überzeugungszuständen, die signifikant in common-sense-Erklärungen und common-sense-Generalisierungen vorkommen, von folgender Form sind:

(F1) „ist eine Überzeugung, daß  $p$ “.

Die Psychologie dagegen kann mit diesen propositionalen Inhalten nichts anfangen, sondern benötigt Identifizierungsstandards, die auf der Grundlage von Gehirnzuständen entwickelt werden müssen.

Und schon wird es eigenartig. Wenn man von einer „syntaktischen“ Psychologie nach Fodor ausgeht, heißt das doch, daß die relevanten Eigenschaften einer Überzeugung die Form

haben:

(F2) „einen mentalen Satz glauben, der die funktionale Rolle  $f$  hat“.

Die Crux: Eine solche Psychologie braucht weder weiten noch engen Inhalt, denn in F2 kommt keine Inhaltsvariable oder Inhaltsfunktion vor, nur eine Variable für die Gesamtheit der Kausaleigenschaften, nämlich die funktionale Rolle  $f$ .

Fodor sagt, die Weise, in der die Psychologie Überzeugungszustände identifiziere, „garantiere“, daß diese durch engen Inhalt gekennzeichnet werden könnten [[Fo87] p. 45]. Schiffer bemerkt, daß es zwei Arten gebe, dies zu lesen, und beide unabhängig voneinander inkonsistent seien.

**I) Die starke Lesart:** Die kognitive Psychologie soll expliziten Bezug zu Überzeugungseigenschaften herstellen, die die Form haben:

(F3) „einen mentalen Satz mit dem engen Inhalt  $s$  glauben“

Für die Identifikation von Überzeugungszuständen werden enge Inhalte gebraucht, weil von der psychologischen Theorie auf sie Bezug genommen wird. (Wenn diese Theorie niedergeschrieben wird, wird sie den Ausdruck „enger Inhalt“ enthalten.) Oben haben wir gesehen, daß sich aus Fodors Ausführungen der Schluß ziehen läßt, daß kein enger Inhalt nötig ist. Aus Fodors oben dargestellter Äußerung (von p. 45 in [Fo87]) in der starken Lesart muß man den Schluß ziehen, daß ein charakterähnlicher Begriff des engen Inhalts absolut nötig ist. Das ist ein Widerspruch.

Außerdem ist nicht klar, weshalb die kognitive Psychologie den common sense in der Identifikation von psychischen Zuständen (im weiten Sinne) ausgerechnet dadurch unterstützen sollte, daß sie deren propositionalen Inhalten Rechnung trägt. Es ist zwar richtig, daß jede ernstzunehmende Psychologie mit der Tatsache konsistent sein muß, daß meine Überzeugung „Wasser ist naß“ sich von der Überzeugung meines Zwillinges, „Wasser ist naß“ durch die relationale Eigenschaft unterscheidet, daß ich mich auf  $H_2O$  beziehe und er sich auf XYZ. Daraus folgt aber nicht, daß wissenschaftliche Psychologie diesen Unterschied im propositionalen Inhalt unserer Überzeugungen auch zur Kenntnis nehmen oder gar thematisieren müßte! Wenn jemand den Physikalismus Fodors teilt, also auch die Rückführung intentionaler beziehungsweise repräsentierender Eigenschaften auf physikalische Eigenschaften zu betreiben versucht, dann wird er auch versuchen, propositionale Einstellungen in empirischen Begriffen zu beschreiben. Das ist es, was Fodor in Kapitel 4 von [Fo87] tut: eine Theorie entwickeln, die den engen Inhalt schon aufgrund ihrer Begrifflichkeiten nicht braucht und ihn dennoch entwickelt.

**II) Die schwache Lesart:** Die Überzeugungszustände müssen von der kognitiven Psychologie zwar nicht unter explizitem Gebrauch des engen Inhalts identifiziert werden, aber die wissenschaftlichen Standards müssen sich in einer Typologie darstellen lassen, die mit der Konzeption des engen Inhalts konform ist. Das Problem dabei ist, daß nicht ersichtlich ist, wie oder warum psychologische Prädikate der Form F2 mit common-sense-Prädikaten der



Form F3 koextensiv sein sollten, denn das würde Fodor damit behaupten. Anders formuliert: warum sollte die Tatsache, daß die Psychologie Einstellungen internalistisch identifiziert, bedeuten, daß sie einen engen Begriff des Inhalts annimmt, geschweige denn benötigt?

Schlußfolgerung: Fodor könnte, so Schiffer, den engen Inhalt ignorieren und statt seiner in V.5 dargestellten Strategie wie folgt argumentieren:

„Die Eigenschaften von Überzeugungszustandsinstanzen sind im Gebrauch des common sense von der Form F1. Aber diese Eigenschaften können von der kognitiven Psychologie nicht benutzt werden, denn sie sind aufgrund ihrer Inkohärenz mit der These der Geist/Gehirn-Supervenienz für die Klassifizierung funktionaler Rollen der Eigenschaften irrelevant. Das heißt, kognitive Psychologie muß internalistisch (solipsistisch) sein und benötigt nur Eigenschaften der Form F2.

Andererseits sind Eigenschaften der Form F1, wie der common sense sie den Überzeugungszuständen zuschreibt, obwohl kausal irrelevant, offensichtlich Eigenschaften, die tatsächlich vorliegen und die man braucht, um Vorgänge in der Außenwelt beschreiben zu können. Jede ernstzunehmende Psychologie muß also mit der Tatsache, daß Überzeugungszustände propositionalen Gehalt haben, kompatibel sein. Sie muß diese Tatsache jedoch in der Formulierung ihrer eigenen Gesetzmäßigkeiten nicht gesondert berücksichtigen. Eine Unterstützung des common sense kann man durch philosophisch-kausale Theorien erreichen.“

Mit dieser Argumentation könnte Fodor vollständig auf den Begriff des engen Inhalts verzichten.

Gibt es den engen Inhalt, wie Fodor ihn sich vorstellt, überhaupt? Schiffer bejaht dies und gibt eine Herleitung an ([Sch90], p. 92-94), die im Rahmen dieser Arbeit nicht auch noch behandelt werden kann. Diese Herleitung mündet in einer Funktion, die wertlos ist, weil sie keine Erklärungskraft bezüglich irgendetwas hat. Schiffer argumentiert daher, der enge Inhalt sei etwas, über das weiterhin nachzudenken sich nicht lohne.

Doch wenn es diesen engen Inhalt in Fodors Sinne gibt (und nach Schiffer gibt es diesbezüglich keine Gegenanzeigen), dann wollen wir wenigstens noch in Erfahrung zu bringen versuchen, ob er denn wenigstens so funktionieren könnte, wie Fodor es sich gedacht hat. Außerdem ist es lehrreich zu sehen, warum dieser Begriff so problematisch ist.

## **V.7 Enger Inhalt zwischen Syntax und Holismus: Fodor in der Kritik von Block**

Der enge Inhalt soll die Kennzeichnung propositionaler Einstellungen in einem Sinne möglich machen, der für die Erklärungsversuche wissenschaftlicher Psychologie relevant ist: mithilfe des engen Inhalts einer Einstellung oder Äußerung soll auf Begründungsstrukturen und Motivationen geschlossen werden können. Daraus und aus dem, was wir bereits von Fodor über den engen Inhalt gehört haben, folgt:

- a) Der enge Inhalt muß Relevanz für Erklärungsversuche der wissenschaftlichen Psychologie besitzen. Das heißt, er muß etwas sein, das Personen gemeinsam haben können, denn ohne eine gewisse Generalität dieses Begriffs kann die Psychologie damit nichts anfangen.
- b) Der enge Inhalt soll tatsächlich „eng“ sein: er soll auf dem neurophysiologischen Zustand des jeweiligen Sprecher supervenieren, also auf physikalischen, nichtrelationalen Eigenschaften. Aus a) und b) folgt, daß alle mentalen Zwillinge sich denselben engen Inhalt teilen.
- c) Enge Begriffe sollen hinreichend differenziert sein: der enge Begriff<sup>8</sup> des Wortes „Silber“ soll sich vom engen Begriff des Wortes „Aluminium“ unterscheiden, der enge Begriff des Wortes „Pandabär“ vom engen Begriff des Wortes „Grizzlybär“ und sofort. Unterscheidungen bei engen Begriffen sind nötig, weil sonst die Erklärungskraft des engen Inhalts verloren geht. (Der Psychologe will schließlich erklären, warum ich meine Kamera zücke, wenn ich höre, ein Pandabär sei hinter der Ecke, doch weglaufe, wenn ich herausfinde, daß es in Wirklichkeit ein Grizzly ist. Wenn die engen Begriffe dieser Wörter sich aber nicht unterscheiden, was ist dann gewonnen?)

Fodor will den engen Inhalt auch aus einem Grund in der psychologischen Bedeutungstheorie installieren, den wir bisher nicht besprochen haben: um die psychologische Bedeutungstheorie vor einem Holismus der Bedeutung retten zu können.

Diese Gefahr muß trotz der gedrängten Darstellung näher erklärt werden: Bedeutungsholismus, wie Fodor ihn versteht, besteht in der These, daß die Identität einer propositionalen Einstellung erst durch die Gesamtheit aller Sätze (beziehungsweise Einstellungen) festgelegt wird, deren Wahrheitswerte der Sprecher für die semantische Auswertung des Inhalts dieser Einstellung als relevant erachtet. Man muß also, um eine Einstellung identifizieren zu können, eine Fülle anderer Einstellungen kennen.

Ein Beispiel: Ich habe eine Biographie von Skrjabin gelesen und finde, daß er vermutlich ziemlich verrückt war. Außerdem habe ich Vladimir Horowitz in einem Interview sagen hören, er habe als kleiner Junge Skrjabin erlebt und auch er sei der Auffassung, Skrjabin sei absolut verrückt gewesen. Ich denke, daß Horowitz als sensibler Künstler das sicher beurteilen kann und stimme ihm zu. Meine Bekannte Betty glaubt allerdings, daß auch Horowitz verrückt sei, und spricht ihm daher die Fähigkeit ab, beurteilen zu können, ob Skrjabin verrückt gewesen sei. Ihrerseits glaubt sie vielmehr, Skrjabin sei (vielleicht auf seine Art) keineswegs verrückt (allenfalls „religiös ungefestigt“), sondern ein genialer Komponist und großer ästhetischer Visionär gewesen.

Wir sehen also: ob der Satz „x glaubt, daß Skrjabin verrückt gewesen sei“ von mir oder von Betty wahrheitsgemäß ausgesagt werden kann, hängt von den Wahrheitswerten sehr unterschiedlicher anderer Sätze ab. Für mich ist Horowitz' Urteil über Skrjabin für die semantische Auswertung des Satzes „Skrjabin war verrückt“ wichtig, für Betty nicht. (Ich werde im folgenden so verfahren zu sagen, Horowitz' Urteil spiele für mich die Rolle eines

---

<sup>8</sup>Ein enger Begriff ist der enge Inhalt eines einzelnen Wortes, während herkömmlicher enger Inhalt sich auf Sätze bezieht. Mein Sprachgebrauch hier lehnt sich an [Blo91] an.

*epistemischen Indikators*<sup>9</sup> für den Satz „Skrjabin war verrückt.“) Wenn der Bedeutungsholismus wahr ist, folgt daraus, daß die Identität der Überzeugung „Skrjabin war verrückt“ von der Gesamtheit aller anderen, für die Auswertung relevanten Einstellungen oder Urteile abhängt. Diese Gesamtheit variiert offensichtlich von Sprecher zu Sprecher. Ergo: „Skrjabin war verrückt“ ist eine rein individuelle Überzeugung, die keine zwei Sprecher gemeinsam haben können. Jede Psychologie muß an dieser Hürde scheitern.

Offensichtlich ist die Schlußfolgerung aus dem Bedeutungsholismus inakzeptabel, denn selbstverständlich könnte auch Betty das Buch lesen und ebenfalls die Auffassung entwickeln, daß Skrjabin verrückt war, ohne jedoch ihre Meinung über Horowitz deshalb ändern zu müssen.

Das Ur-Argument für den Bedeutungsholismus lautet folgendermaßen:

1.) Zeige, daß einige der epistemischen Indikatoren einer bestimmten Einstellung ihren intentionalen Inhalt bestimmen.

2.) Zeige, daß epistemische Indikatoren vom Sprecher abhängig sind, wir also in Bezug auf eine einzelne Einstellung/einen einzelnen Satz nicht wissen können, welche epistemischen Indikatoren den intentionalen Inhalt bestimmen.

Aus 1.) und 2.) folgt per Modus tollens:

3.) Alle denkbaren epistemischen Indikatoren beeinflussen den intentionalen Inhalt jeder Einstellung.

Fodor diskutiert sehr umfangreich die „drei Wege“, die ausgehend von diesem Urargument in den Bedeutungsholismus führen und für jede dieser drei argumentativen Routen bietet er Gegenargumente an. Obwohl eine Einfügung von Fodors Holismus-Diskussion überaus interessant wäre, muß sie im Hinblick auf den begrenzten Umfang dieser Arbeit an dieser Stelle ausgelassen werden. Für unsere Zwecke ist es ausreichend, zu sehen, daß Bedeutungsholismus für Fodors Begriff des engen Inhalts ein Problem ist.

Fodor sprach davon, daß Putnams Experiment neue Identitätskriterien für die Beschreibung psychischer Zustände (im weiten Sinne) nötig mache. Doch wie sollen wir solche Kriterien bekommen, wenn es keine geben kann?

Der Bedeutungsholismus, so Fodor, ist eine Gefahr für alle Bedeutungsbegriffe, die den Anspruch haben, in den kausalen Erklärungen wissenschaftlicher Psychologie Aussagekraft zu haben. Es ist unschwer zu erkennen, daß die These des Bedeutungsholismus mindestens gegen die Forderungen a) und c) verstößt, die der Begriff des engen Inhalts erfüllen soll. Gegen a) verstößt die Folgerung aus dem Bedeutungsholismus deshalb, weil keine zwei Personen mehr dieselbe Einstellung haben könnten (vgl. Skrjabin-Beispiel) und gegen c), weil die Differenzierung zu feinkörnig würde, um subjektunabhängige enge Begriffsprofile wie „Panda“ und „Grizzly“ eindeutig bestimmen und unterscheiden zu können. Stattdessen hätte jeder Sprecher seine eigenen engen Begriffe. Dies mündet in genau dieselbe Problematik, die wir mit Frege bereits in I.4 besprochen haben: kein Gedanke wäre mehr mitteilbar, weil jeder Sprecher nur über seine eigenen Begriffe spricht. Wissenschaft in Form kumulativen

---

<sup>9</sup>Fodor nennt es „epistemic liaison“. Es ist ein informeller Ausdruck, der keine große Rolle spielt. Der epistemische Indikator ist kein semantisches, sondern ein psychologisches Merkmal: es muß keine objektivierbare Beziehung zwischen ihm und der jeweiligen propositionalen Einstellung geben. Ob etwas ein epistemischer Indikator für eine bestimmte Einstellung ist, hängt ganz von der jeweiligen Einschätzung des Sprechers ab.

Wissens wäre undenkbar. Aus diesem Grund können wir die Schlußfolgerungen aus dem Holismusargument nicht akzeptieren.

Wenn enger Inhalt als theoretische Konstruktion lebensfähig sein soll, muß demnach der Bedeutungsholismus widerlegt werden, weil sonst kein Inhalt sich sowohl als begrifflich zusammengesetzt, also generalisierbar (a), *und* hinreichend differenziert (c) verstehen lassen könnte, um wirkliche Erklärungskraft zu entfalten.

Wir werden nun aber feststellen, daß auch Fodors eigene Konzeption des engen Inhalts nicht in seinem Sinne funktionieren kann. Sie bringt schwerwiegende Probleme mit sich, weil sie den engen Inhalt in eine Position zwischen Skylla und Charybdis drängt: jeder Begriff des engen Inhalts, der auf der Zuordnungstheorie beruht, die Fodor aufgestellt hat, also durch eine Funktion von Kontexten in Wahrheitsbedingungen dargestellt wird, ist dazu verurteilt, entweder in reine Syntax zu implodieren oder in Holismus zu explodieren. Und sofern sich zwischen diesen Polen enge Inhalte formulieren lassen, sind sie für psychologische Erklärungen unbrauchbar.

Diese Bedrohung für den Inhaltsbegriff ist am deutlichsten in [Blo91] beschrieben. Laut Fodor besteht der enge Inhalt einer Einstellung oder eines Satzes in einer partiellen Funktion, die verschiedenen Erwerbskontexten bestimmte Wahrheitsbedingungen (weite Inhalte) zuordnet. Nach Block gibt es mehrere Wege, diesen zuordnungstheoretischen Begriff des engen Inhalts zu entwickeln:

I. Der enge Inhalt der Äußerung „Schlamm ist ein gutes Shampoo“ ordnet jedem Element einer Menge von nichtaktuellen Äußerungskontexten die Wahrheitsbedingungen zu, die dieser Satz in dem jeweiligen Kontext hat. Das heißt, der Definitionsbereich der Funktion sind all die Kontexte, in denen ich jeweils aufgewachsen bin und zu sprechen gelernt habe. Der Wertebereich der Funktion sind die jeweiligen Wahrheitsbedingungen meiner Äußerung dieser Worte im jeweiligen Kontext als Sprecher, der in diesem Kontext erwachsen geworden ist.

In dieser Version hat der enge Inhalt praktisch nichts mit der Beschaffenheit meines Geistes zu tun: vielmehr hängt, was ich mit dem Satz sagen will, in den meisten Erwerbskontexten von meinem Aufwachsen in der physikalischen und sozialen Umgebung ab. Wenn ein Psychologe erklären wollte, warum ich mein Haar mit Schlamm wasche, kann er auf nichts zurückgreifen, was nicht physikalisch oder sozial erklärbar wäre, jedenfalls nichts, was eindeutig in meinem Kopf wäre. Wir müssen uns vergegenwärtigen, daß ich eine andere Sprache sprechen könnte: der Satz „Schlamm ist ein gutes Shampoo“ könnte tatsächlich alles mögliche bedeuten, weil jeder hörbare Laut, den ich produziere, für Sprachzeichen mit vollkommen beliebiger Bedeutung stehen könnte. Ich könnte, je nach sprachlicher Sozialisation, mit dem Satz „Schlamm ist ein gutes Shampoo“ auch meinen, daß Eisenhower eine Ente ist. Alles was ich mit meinen Zwillingen teile, ist daher die Tatsache, daß ich in jedem Kontext die gleichen Laute produziere. Das heißt aber, daß der *einzigste konstante Unterschied* zwischen den Äußerungen „Schlamm ist ein gutes Shampoo“ und „Eisenhower ist eine Ente“, die beide jeweils alles bedeuten können, nicht semantischer oder mentaler, sondern syntaktischer Natur ist. Der Begriff, der diesen Unterschied einfängt, gehört nicht in die Semantik, sondern zur Syntax der Sprache, die ich spreche. Also ist der enge Inhalt in Syntax kollabiert. Wieso ist das eine Gegenanzeige zu unserem Vorhaben? Weil syntaktisch identische Objekte sehr unterschiedliche funktionale Rollen innehaben können. Wir sollten

die engen Inhalte allerdings anhand ihrer funktionalen Rollen unterscheiden können; jede Theorie, die das nicht zuläßt, ist zu undifferenziert, um den Zwecken der Psychologie zu genügen. Wenn ich ein Lebewesen im Wasser sehe, hängt es von meinem engen Inhalt des Begriffs für dieses Lebewesen ab, ob ich davonschwimme, oder es neugierig betrachte, denn ich habe unterschiedliche enge Inhalte von „Hai“ und „Frosch“.

II. Nehmen wir nun in einer zweiten Version der Theorie an, der Definitionsbereich sei die Menge aller Kontexte, in denen ich einfach als Sprecher der Worte „Schlamm ist ein gutes Shampoo“ sozusagen „erscheine“, ohne jeweils aufgewachsen sein zu müssen. Der Wertebereich sind die Wahrheitsbedingungen, die in diesen Kontexten jeweils für diesen Satz bestehen.

Für jede Welt, in der also mein plötzliches Erscheinen physikalisch oder sonstwie unmöglich ist, ist die partielle Inhaltsfunktion einfach nicht definiert. Wenn ich aber erscheine, dann müßte die Funktion meiner Äußerung die Wahrheitsbedingungen zuordnen, die in dem jeweiligen Kontext für diesen Satz bestehen. Abgesehen davon, daß völlig unklar ist, wie diese aussehen könnten, können wir annehmen, es seien dieselben wie in unserer Welt. Wir haben nun etwas Nutzloses definiert, nämlich eine Funktion, die immer dieselben Wahrheitsbedingungen zuordnet. Dieser enge Inhalt erfüllt seinen Zweck nicht, denn er würde von mir und meinem Zwilling nicht geteilt, wenn wir daran denken, daß Schlamm, um Schlamm zu sein, Dreck und Wasser (!) enthalten muß.

Das Problem des engen Inhalts hier ist, daß wir ihn gerne konstant halten würden: wir brauchen eine Funktion, die jedem Kontext denselben engen Inhalt zuordnet, der in unserer Welt mit der Äußerung im jeweiligen Erwerbkontext verknüpft wird. Das können wir jedoch nicht einfach *per definitionem* festlegen, denn dann würden wir eine leere Theorie schaffen. Was eigentlich gebraucht wird, ist eine Bedingung, die bewirkt, daß der enge Inhalt in jedem Kontext derselbe ist wie jeweils in unserer Welt. Dafür bräuchte man eine Beschränkung des Definitionsbereichs, ein Kriterium, das diese Ausfilterung nutzloser Kontexte sicherstellt, ohne selbst den Begriff des engen Inhalts vorauszusetzen (cf. [Blo91], p. 39).

Das können wir versuchen: Nehmen wir an, der enge Inhalt ordnet jedem Kontext genau die Wahrheitsbedingungen zu, die eine typenidentische Äußerung in jeder Welt haben würde, in der der Sprecher aufwächst und zu sprechen lernt und zum Zeitpunkt der Äußerung ein physikalisches und/oder funktionales Duplikat von mir ist. Für alle Kontexte, in denen der Sprecher kein Duplikat von mir ist, ist die Funktion nicht definiert. Diese dritte Version verstößt offensichtlich und sehr eklatant gegen Forderung a), denn im Normalfall psychologischer Verhaltensklärung geht es nunmal nicht um funktionale Duplikate. Enge Inhalte, die nur von solchen Duplikaten geteilt werden, sind nicht ausreichend allgemein.

III. Doch auch eine andere Herangehensweise, zum Beispiel über den engen Inhalt von Wörtern, führt zu Problemen:

Nehmen wir an, der enge Inhalt meiner Äußerung von „Wasser“ besteht in der Funktion, die jedem Kontext in dem ich sie äußere, die jeweils farb- und geruchlose Flüssigkeit zuordnet, die vom Himmel fällt und Seen und Flüsse füllt. Das kann schon deshalb nicht sein, weil „Wasser“ auch in Kontexten innerhalb des Definitionsbereichs alles mögliche heißen kann, zum Beispiel „Aua“, „später“, „wenn“ oder „blau“.

Also nehmen wir an, der enge Inhalt von „Wasser“ ordne eben jedem Kontext, in dem ich sie äußere, *um* die jeweils farb- und geruchlose Flüssigkeit, die vom Himmel fällt und Seen und Flüsse füllt, herauszugreifen, eben genau jene Flüssigkeit zu, auf die diese Beschreibung in der jeweiligen Welt zutrifft. Kontexte, in denen Wasser etwas anderes bedeutet, sind außerhalb des Definitionsbereiches. Doch es kann Kontexte geben, in denen die Sprecher kein Konzept von Farblosigkeit, Geruchlosigkeit oder von Flüssen oder Seen haben. (Das kann zum Beispiel an veränderter Wahrnehmung oder hochentwickelter Wissenschaft liegen. Ausschließen kann man es nicht.) Das würde aber heißen, daß diese Sprecher Wasser bezeichnen könnten, ohne etwas zu besitzen, das unserem engen Inhalt vergleichbar wäre und dann können wir uns auch diesen Versuch sparen. Wir wollen aber, wie oben beschrieben, daß jeder Sprecher in jedem Kontext immer denselben engen Inhalt hat, wie wir in unserer Welt und das ist noch immer nicht geschafft.

Man könnte noch versuchen, nur solche Kontexte zuzulassen, in denen das Wort „Wasser“ mit demselben weiten Inhalt gebraucht wird, wie wir in unserer Welt es tun, wobei weiter Inhalt für ein einzelnes Wort hier bedeuten soll, daß wir mit diesem Wort immer dieselben Überzeugungssätze verknüpfen. Dann wäre der enge Inhalt von „Wasser“ die Funktion, die jedem Kontext, in dem das Wort „Wasser“ mit dem Begriff der jeweils anwesenden farb- und geruchlosen Flüssigkeit verbunden ist, die vom Himmel fällt und Seen und Flüsse füllt, genau diese Flüssigkeit zuordnet.

Das Problem ist nun, daß wir aus dem engen Inhalt etwas gemacht haben, das nicht mehr eng ist: es könnte jetzt sehr wohl ein funktionales Duplikat von mir geben, das einen anderen engen Inhalt von „Wasser“ hat als ich.

Nach all diesen Versuchen zeichnet sich bereits ab, daß nicht klar ist, wie wir es schaffen sollen, einen zuordnungstheoretischen Begriff von engem Inhalt zu entwickeln, der alle gestellten Anforderungen erfüllt.

Wir sind mit drei Anläufen in drei verschiedene Sackgassen gerannt:

- I.) Wir haben es nicht geschafft, engen Inhalt zu entwickeln, weil er in Syntax kollabiert ist. Syntaktische enge Inhalte sind zu undifferenziert, um kausale Unterschiede syntaktisch identischer Objekte erklären zu können (Verstoß gegen c).
- II.) Wir haben es geschafft, engen Inhalt zu entwickeln, aber er wird nur von funktionalen Doppelgängern geteilt und ist daher für die Psychologie wertlos (Verstoß gegen a).
- III.) Wir haben es geschafft, Inhalt zu entwickeln, aber er ist nicht eng und kann hinreichend durch den weiten Inhalt ausgedrückt werden. Außerdem muß er von funktionalen Doppelgängern nicht geteilt werden (Verstoß gegen b und den Begriff des „Engseins“).

Es sieht schlecht für Fodors Projekt aus. Wir wissen jetzt, daß es keinen Begriff des engen Inhalts geben kann, der in der von Fodor beschriebenen Weise konstruiert ist und in der von ihm intendierten Weise funktioniert. Der nächste Abschnitt wird zeigen, welchen Begriff von engem Inhalt wir laut Block haben *können*.

## V.8 Blocks Versuch des holistischen engen Inhalts

Gibt es nun eine Zuordnungstheorie des engen Inhalts, die die genannten Schwierigkeiten vermeiden kann? Blocks Antwort lautet: ja, aber sie mündet in Holismus.

Er geht von einem Ansatz aus, den er die „weiter-Inhalt“-Zuordnung nennt. Nehmen wir an, der enge Inhalt meines Wortes „Wasser“ sei eine Funktion von Kontexten in Wahrheitsbedingungen wie gehabt. Der Input dieser Funktion sind alle und nur die Kontexte, in denen der Sprecher alle meine Überzeugungen bezüglich Wasser teilt. Der Output sind die Erfüllungsbedingungen für „Wasser“ in diesen Kontexten, das heißt Farb- und Geruchlosigkeit, Flüssigkeit und sofort. (Überzeugungen bezüglich Wasser heißen in diesem Zusammenhang: alle Überzeugungen bezüglich Wasser mit Ausnahme der Überzeugung, daß Wasser als „Wasser“ bezeichnet wird.)

Wenn mein zwirtdischer Doppelgänger und ich darin übereinstimmen, daß Wasser durstlöschend ist, daß es Seen und Flüsse füllt, daß es farb-, geruch- und geschmacklos ist, dann haben wir jetzt denselben engen Inhalt von Wasser.

Dem Kollaps in Syntax können wir vorbeugen, indem wir versuchen, schon vor der Einführung des engen Inhalts zwischen allen Begriffen semantische Unterscheidungsmerkmale einzuführen. Block schlägt dafür folgende kausale Theorie der Referenz vor.

Jedes Wort, dessen Referenz in der aktuellen Welt von der physikalischen und sozialen Umgebung des Sprechers festgelegt ist, *könnte theoretisch* auch durch einen kennzeichnenden Ausdruck so eingeführt werden, daß seine begriffliche Konzeption genau diejenigen Objekte bezeichnet, die dem Wort von der Sprachgemeinschaft zugeordnet wurden, daß also die Referenz des Wortes durch eine Kennzeichnung fixiert wird, ähnlich wie die frühen Semantiker es sich dachten.

Zum Beispiel könnte Sherlock Holmes, wenn er rätselhafte Morde im Dschungel untersucht, zu dem Schluß kommen, daß es eine Spezies gibt, die noch unbekannt ist, und orange-schwarz gestreift, ein Fleischfresser und eine Raubkatze sein muß. Diese Spezies nennt er „Tiger“ und legt damit die Referenz des Ausdrucks „Tiger“ auf alles fest, was diese Beschreibung erfüllt. In diesem Fall spielt die Konzeption des Ausdrucks eine entscheidende Rolle für die Referenz. Die engen Inhalte von „Tiger“ und „Panda“ werden sich, wenn Holmes auch noch Pandas entdeckt, auf jeden Fall signifikant unterscheiden.

Freilich werden längst nicht alle Taufen so ablaufen, es gibt vielmehr auch Kontexte, in denen die Sprachgemeinschaft einem Gegenstand auch einfach einen Namen gibt, ohne daß irgendetwas außer der Relation des Nennens den Referenten mit seinem Namen verbindet.

Doch wir *könnten* uns für fast alle Wörter auch nachträglich Unterscheidungen ausdenken, die auf Kennzeichnungen basieren, also semantischer Art sind, so daß dem Kollaps in Syntax vorgebeugt würde. Diese Unterscheidungen muß die Zuordnungstheorie des engen Inhalts berücksichtigen.

Wie kommt es nun zum Holismus?

Es gibt für jedes Wort natürlich mögliche Erwerbskontexte, in denen die Referenz eines Wortes von mehr als einer Person festgelegt wird. Und in vielen dieser Kontexte werden verschiedene Täufer verschiedene Konzeptionen zugrundelegen. Zum Beispiel kann man sich uneins sein, ob Tiger schwarze Streifen auf orangem Körper haben oder orangene Streifen auf schwarzem Körper. Oder man kann sich streiten, wie lang der Schwanz eines Tigers

sein soll. Für jeden solchen Unterschied gibt es mögliche Erwerbskontexte, in denen dieser Unterschied bewirkt, daß der entsprechende Ausdruck nur eine von mehreren tatsächlich vorhandenen, distinkten Spezies bezeichnet. Zum Beispiel könnte es zwei Arten von Tigern geben, eine orange-schwarz gestreifte und eine schwarz-orange gestreifte, eine mit langem Schweif und eine mit Stummelschwanz. Wir können für jede denkbare Vielfalt begrifflicher Konzepte eines Wortes einen Kontext konstruieren, in dem diese Vielfalt auf tatsächlichen Unterschieden basiert, und in denen folglich der Ausdruck keine eindeutige Referenz hat.

Wenn nun jeder der beiden Täufer die Absicht hat, die Referenz des Ausdrucks „Tiger“ so festzulegen, daß nur die Elemente, die jeweils *seine* Konzeption erfüllen, Tiger sind, dann folgt, daß es zwei verschiedene Arten von Tigern geben wird, jede entsprechend der zugrundeliegenden Konzeption. Schließlich ist alles, was wir brauchen, um zwischen zwei Begriffen  $p_1$  und  $p_2$  einen Unterschied zu verursachen, ein einziger Erwerbskontext, in dem sich die Extensionen von  $p_1$  und  $p_2$  unterscheiden. Also: jede Unterscheidung der Konzeptionen der Täufer bewirkt einen Unterschied – und sei er noch so klein – in den jeweiligen zuordnungstheoretischen engen Inhalten dieses Wortes. Und das ist sprecherbezogener Holismus auf einer sehr grundsätzlichen Ebene.

Die Frage ist nun, ob das den Theoretiker des engen Inhalts stören muß. Zwar ist es laut Block, wenn man den Holismus auf dieser Ebene akzeptiert, etwas leichter, eine funktionalistische Strategie gegen Putnam zu entwickeln, doch wenn enge Inhalte tatsächlich holistisch sind, dann können psychologische Gesetze, die sich auf diese Inhalte berufen, nicht die Generalität haben, die Fodor sich vorstellt und die in a) ausgedrückt ist. Außerdem wären enge Inhalte nicht viel mehr als die Inhalte, die wir uns vorstellen, wenn wir im Jargon des common sense bleiben.

Für den Begriff des engen Inhalts, wie Fodor ihn sich vorstellt, ist Holismus unduldbar. Daraus folgt endgültig, daß es den engen Inhalt, den Fodor will, nicht geben kann.

## V.9 Epilog zum engen Inhalt

Fassen wir zusammen: es kann keinen zuordnungstheoretischen Begriff des engen Inhalts geben, wie Fodor sich ihn gewünscht hätte. Ein solcher Begriff würde sich entweder in Syntax auflösen oder für die Psychologie unbrauchbar sein. Jeder brauchbare zuordnungstheoretische Begriff des engen Inhalts scheint aber, folgt man [Blo91], holistisch zu sein. Läßt man dies gelten, so müßte jede künftige Verteidigung des engen Inhalts sich mit der Frage auseinandersetzen, inwiefern ein holistischer Inhaltsbegriff in Kauf genommen werden kann. Und dem steht Schiffers Argument gegenüber, das wir bereits besprochen haben, wonach enger Inhalt (als Funktionsbegriff) ohnehin überflüssig ist.

Doch ich will auch auf folgendes hinweisen:

- 1.) Die starken Argumente, die Block ins Felde führt, beziehen sich nur auf die Zuordnungstheorie des engen Inhalts. Werden enge Inhalte anders als durch Funktionen entwickelt, so muß man auch andere Gegenargumente finden. Eine andere, prominente Enge-Inhalt-Strategie, die völlig anderen Typs ist als die hier vorgestellten, findet sich in [De82]. Schlagende Argumente gegen diese Strategie werden von Stalnaker in [Sta89] expliziert.



- 2.) Der hier besprochene Begriff des engen Inhalts, der [Fo87] entnommen ist, ist nicht der einzige Begriff des engen Inhalts, der von Fodor entwickelt wurde. In [Fo83] wird beispielsweise eine Version entwickelt, die gegen einige der aufgeführten Argumente gefeit ist.
- 3.) Die hier besprochenen Gegenargumente sind natürlich nicht alle, die gegen Fodor, seine in [Fo87] dargestellte Repräsentationstheorie des Geistes und seine Zuordnungstheorie des engen Inhalts aufgeführt wurden. Deutlichst sei hingewiesen auf [Put88], wo die ganze Problematik der Argumente gegen den Funktionalismus, der Repräsentationstheorie des Geistes, des engen Inhalts und des semantischen Externalismus von Putnam noch einmal in eingängiger Weise diskutiert wird. Putnam geht in diesem Zusammenhang auch auf die Argumente von Fodor und Block ein.

Im Rahmen dieses Kapitels ging es mir daher vor allem darum, die Position des *advocatus diaboli* für den Begriff des engen Inhalts, nämlich die Fodors, genauer darzustellen, weil sie eigentlich die einzige wirklich breit diskutierte Antwort auf Putnams externalistischen Vorstoß darstellt und eine Flut von Antwortliteratur auslöste, obwohl Fodors widersprüchlicher Ansatz keine derart breite Rezeption erfordert hätte. Vielmehr hätte es sich gelohnt, die Position Whites weiterzuentwickeln, die alle Merkmale aufweist, um nach einer entsprechenden Ausarbeitung als ebenso gute Lösung zum Problem des subjektivitätskompetenten, formalen Bedeutungsbegriffs gelten zu können wie die im nächsten Kapitel besprochene Theorie von Stalnaker.

Der Versuch, einen engen Inhalt zu definieren, markiert den Versuch, Putnam dadurch zu begegnen, daß man zeigt, daß das Twin-Earth-Experiment den Determinationszusammenhang zwischen extensionaler und „inhaltlicher“ (hier im intuitiven Sinne) Identität von Gedanken nicht, wie Putnam behauptete, wirklich unterbricht (und so zeigt, daß Bedeutungen nicht im Kopf sind). Stattdessen will man an der These festhalten, daß der Extensionsunterschied in dem von beiden Zwillingen geäußerten Satz „Wasser ist naß“ aufgrund der in diesem Satz steckenden Intentionalität der Referenz irgendeinen Niederschlag in deren mentalen Gegebenheiten findet, den man durch einen neuen Begriff beschreiben kann. Das ist keine schlechte Strategie, nur wird der Begriff des engen Inhalts ihr nicht gerecht.

Auch hier zeigt sich also die Notwendigkeit, kognitive Dimensionen von Bedeutungsbegriffen mit formal-semantischen Aspekten zu verknüpfen, wie White es getan hat. Wie schon das in I.6 und V.2 gezeigte Problem der Einstellungskontexte, so zeigt auch das Problem der Abbildung von Intentionalität, daß ein tragfähiger Bedeutungsbegriff subjektivitätskompetent sein muß: dies ist die Verbindung zwischen subjektiven und objektiven Bedeutungsdimensionen, kognitiven und formalen Aspekten eines Begriffs. Einen solchen Begriff hat White konstruiert. Der Begriff des partiellen Charakters berücksichtigt Intentionalität der Referenz und ist daher subjektivitätskompetent.

Ein wichtiger Punkt, den alle drei hier besprochenen Theoretiker – White, Loar, Fodor – gesehen haben, ist, daß in Putnams Experiment keine ausreichende Unterscheidung zwischen intentionalen und nichtintentionalen psychischen Konstellationen getroffen wurde, weil Putnam propositionale Einstellungen relational identifiziert. In gewisser Weise könnte man ihm vorwerfen, in seinem Beweis des Externalismus also schon „vorne“ hineingelegt

zu haben, was „hinten“ herauskommen soll. Meines Erachtens hat White diesen Punkt am konsequentesten (wenn auch unauffälligsten) in seine Argumentation eingebaut.

Wäre der enge Inhalt als theoretische Konstruktion geglückt, so hätte er endlich etwas angegeben, das wir als intrinsische kognitive Signifikanz eines Gedankens hätten bezeichnen können. Der partielle Charakter gibt zwar so gut wie kein anderer bisher behandelter Begriff die kognitive Signifikanz eines Gedankens an – doch er ist nicht wirklich intrinsisch und richtet sich nicht wirklich gegen Putnams Argumente, sondern versucht diese lediglich in die rekursive Semantik einzufügen. Der enge Inhalt hätte darüberhinaus auch dem intuitiven Begriff dessen entsprochen, was ein Gedanke bedeutet. Man kann aus der Tatsache, daß es auch dem Begriff des engen Inhalts nicht gelungen ist, kognitive Signifikanz zu präzisieren, nur den Schluß ziehen, daß wir, wenn wir diesen Weg wählen, nicht weiter sind, als Frege war.

Der Begriff des engen Inhalts wurde seit seinem Aufkommen gegen Mitte der Achtziger Jahre sehr kontrovers diskutiert. Es gilt heute als sicher, daß er nur unter großen Schwierigkeiten konsistent entwickelt werden kann und – egal in welcher Konzeption – sehr wahrscheinlich nicht als wirksame Verteidigung gegen das Twin-Earth-Argument dienen kann. Bei der Lösung der Probleme, die Putnam aufgezeigt hat, war der enge Inhalt bisher weitgehend ohne Wirkung. In den letzten Jahren kann man ein Abebben der Diskussion um diesen Begriff feststellen. Entscheidende Durchbrüche sind nicht mehr zu erwarten.

Im nächsten Kapitel nun werden wir erleben, wie externalistische und internalistische Argumente Eingang in einen neuen Begriff zur Beschreibung von Gedanken gewinnen.

# VI Zur Semantik der Überzeugungen: Stalnaker

## VI.1 Relationalität und Intrinsicität

Wie bringen wir nun die Ansprüche aus I.6 und V.9 mit den Einwänden aus Kapitel II und den Präzisierungen aus Kapitel III einerseits mit den Argumenten aus Kapitel IV zusammen ohne einen problematischen Inhaltsbegriff zu brauchen? Die Lösung, die wir gerne sähen, wäre die, eine Semantik zu konstruieren, in denen sich Putnams Argumente darstellen lassen, das heißt, eine Semantik, die den Satz „Wasser ist naß“, den ich sage, von dem Satz „Wasser ist naß“, den mein zworldischer Kollege sagt, zwar plausibel unterscheidet, aber dennoch auch angibt, was die Zwillinge gemeinsam haben. Was nämlich die Sätze unterscheidet, ist der formal-semantische Aspekt, der die Wörter „Wasser“<sub>E</sub> und „Wasser“<sub>TE</sub> aufgrund ihres Extensionsunterschiedes differenziert. Ein subjektivitätskompetenter Bedeutungsbegriff jedoch muß in der Lage sein, außerdem auch abzubilden, was die Zwillinge gemeinsam haben – nämlich den intentionalen Aspekt, sich auf die Erscheinung zu beziehen, die jeweils „Wasser“ heißt.

Dies heißt nichts anderes, als daß wir für die Zwillingssätze eine rekursive Bedeutungszuweisung realisieren wollen, die die beiden Einstellungen plausibel unterscheidet. Wenn wir das schaffen, so können wir auch dem Problem, daß Gedanken nicht im Kopf seien, einen fundierten Entwurf entgegensetzen – allerdings müßte die geforderte Semantik sich dann nicht nur auf Äußerungen von Sätzen, sondern auch auf Überzeugungszustände beziehen und das ist ein gewaltiges Hindernis.

Die Mittel zur Meisterung dieses schwierigen Projekts liegen bereit: sie wurden von Robert Stalnaker in einer Reihe von Aufsätzen konkretisiert, hauptsächlich in [Sta78] und [Sta87]. Nimmt man außerdem [Sta88] und [Sta89] hinzu, so wird auch der Bezug auf das Twin-Earth-Problem explizit. Dieses Kapitel hat zwei Ziele: erstens müssen wir irgendwie auf Fodors Argument antworten, daß relationale Identifikation von Einstellungszuständen einfach nicht erlaubt sei, um überhaupt den Weg für eine Auseinandersetzung mit Putnams Argument freizumachen. In einem zweiten Schritt soll dann eine Semantik für Überzeugungszustände konstruiert werden, mit deren Hilfe sich die Gedanken der Zwillinge unterscheiden lassen und die die offenen Enden der in den Kapiteln I-IV geführten Diskussion unter Berücksichtigung aller vorgetragenen Argumente zusammenbringt.

Zunächst zum ersten unserer Ziele. Wir erinnern uns, daß die internalistische Sprachphilosophie implizit von der These ausging, daß Bedeutungen – und demnach Überzeugungszustände – intrinsische Eigenschaften des jeweiligen Individuums seien. Der Einwand von Putnam und Burge lautete, daß Überzeugungszustände – demzufolge Gedanken und Bedeutungen – keine intrinsischen Zustände seien, sondern daß sie, konsequent relational interpretiert, nicht „im Kopf“ seien. Fodors Argument dagegen war, daß in wissenschaftlicher Erklärung funktionale Rollen wichtig sind, und daß die Zwillinge hinsichtlich der funktio-

nalen Rollen ihrer semantisch distinkten „Wasser“-Überzeugungen sich *nicht* unterscheiden. Also würde ihr Verhalten von einem wissenschaftlichen Ansatz aus als identisch betrachtet und die diesem Ansatz zugrundeliegende Theorie wäre nicht gezwungen, eine Unterscheidung zwischen ihren Bewußtseinszuständen zu treffen. Fodor nahm außerdem an, daß es möglich sein müßte, relational identifizierte Einstellungen auch intrinsisch zu identifizieren, und diese Identifikation in der Psychologie zu verwenden. Das sollte der enge Inhalt leisten.

Das Twin-Earth-Experiment, so Stalnakers Kritik, führt in gewisser Hinsicht zur Verdunklung: Die Beziehung zwischen Bewußtsein und Umwelt scheint bei Putnam wesentlich mysteriöser zu sein, als sie tatsächlich ist; es ist verwunderlich, daß die Tatsache, daß intentionale und semantische Eigenschaften auch relationale Eigenschaften sind, die Philosophie so sehr überrascht. Natürlich hängt, was jemand denkt, insofern nicht *nur* von dem ab, was in seinem Kopf ist, als daß die Art des Bezugs der Gedanken auf die Welt von der Beschaffenheit der Welt und der linguistischen Sozialisation durch die Sprachgemeinschaft abhängt. Das ist nicht das Ergebnis einer neuen und revolutionären Theorie, sondern eine offensichtliche Binsenweisheit (die sich auch bei Frege und Carnap finden läßt, die darüber nur nicht reflektierten). Kurios sind dagegen die konsequenten Schlußfolgerungen:

- C1) Daß wir, wenn Gedanken nicht im Kopf sind, keine Kontrolle mehr darüber zu haben scheinen, was wir meinen oder denken,
- C2) daß die Rolle, die intentionalen psychischen Zuständen in den Kausalerklärungen der Psychologie zugeordnet ist, mit deren relationalem Charakter in Konflikt zu geraten scheint, und
- C3) daß nicht klar zu sein scheint, wie relationale Zustände Kausalkraft entwickeln können, wenn sie doch von Gegenständen außerhalb des Agenten abhängen.

Stalnaker weigert sich, nur aufgrund dieser Überlegungen den von Fodor geforderten Ausschluß relationaler Eigenschaften zu vollziehen: relationale Eigenschaften, so Stalnakers These, können eine sehr wichtige Rolle in wissenschaftlicher Erklärung spielen. Die folgenden Ausführungen kann man daher als Antwort auf C2 und C3 lesen.

Man stelle sich folgendes vor: es gibt einen Zeitgenossen – Herrn Schlich – der jeden Morgen hundert Liegestützen ausführt (er leidet unter wirklich schlimmem Fitneßwahn). Schlich hat mit dieser Betätigung erst in seinem fünfzigsten Lebensjahr angefangen (vielleicht als Reaktion auf eine verspätete Midlifecrisis?). Da er diese Gewohnheit sehr regelmäßig fortsetzt und kaum je von seinen 100 morgendlichen Liegestützen abgehalten wird, hat er in den dreißig Jahren bis zu seinem achtzigsten Geburtstag 1.000.000 Liegestütze absolviert. Schlich hat zwei Zwillingbrüder (im Ganzen sind es also Drillinge), die, durch seine Sportlichkeit angespornt, über die Jahre hinweg auch versucht haben, dieses Training durchzuhalten. Leider haben alle drei Brüder einen angeborenen Herzfehler, der in dieser Familie von Generation zu Generation weitergegeben wird und daher werden die beiden Zwillingbrüder kurz nach der 700.000. Liegestütze von einem Infarkt dahingerafft. Schlich selbst ist der Einzige, der es geschafft hat, seine 1.000.000. Liegestütze zu absolvieren. Recherchen in der Familienchronik ergeben, daß in der Generation davor ebenfalls sämtliche

Geschwister kurz nachdem sie ihre 700.000. Liegestütze absolviert haben, einem Herzinfarkt erlegen sind. Und die Geschwister in der Generation davor auch. Und sofort.

Schlichs Hausarzt ist die Sache ein Rätsel und er fragt sich, wieso Schlich immer noch lebt, denn natürlich hat auch er den typischen Herzfehler. Also ist die wissenschaftliche Frage: „Wie ist es wissenschaftlich erklärbar, daß angesichts der bekannten Tatsachen über Schlichs Herz, die Eigenschaft, seine einmillionste Liegestütze absolviert zu haben, auf Schlich zutreffen kann?“

Die Eigenschaft, seine einmillionste Liegestütze absolviert zu haben, zerfällt in zwei Eigenschaften, eine relationale, externe und eine nichtrelationale, intrinsische: da ist zunächst die relationale Eigenschaft, 999.999 Liegestützen absolviert zu haben, die eine rein historische Beschreibung Schlichs beinhaltet. Andererseits aber ist die Eigenschaft, eine Liegestütze absolviert zu haben, eine interne, nichtrelationale Eigenschaft<sup>1</sup>. Also ist die Eigenschaft, seine einmillionste Liegestütze zu absolvieren, eine – wie Stephen Stich sagen würde – hybride Beschreibung, die aus externen und intrinsischen Eigenschaften zusammengesetzt ist.

Wir generalisieren über Herzfehler, indem wir alle Schlichs, die, sagen wir, 650.000 Liegestütze absolviert haben, als „gefährdet“ einstufen. Angesichts der traurigen Ereignisse in dieser Familie scheint das ein überaus vernünftiges Vorgehen zu sein. Wenn nun ein blutsverwandter Schlich nach achtzig Lenzen und einer Million Liegestützen noch immer weitermacht, so ist die Frage berechtigt, *wie das sein kann*.

Offensichtlich zeigt das Liegestützbeispiel, daß Generalisierungen auf der Basis relationaler Eigenschaften widerlegt werden können: schließlich konnte Schlich entgegen der Generalisierung 1.000.000 Liegestütze absolvieren. Das Beispiel zeigt hingegen nicht, daß solche Generalisierungen unzulässig oder ohne Erklärungskraft wären, was die Annahmen C2 und C3 widerlegt. Auch zeigt es nicht, daß es nicht Fälle geben kann, in denen relationale Eigenschaften *tatsächlich* ohne Belang sind. Es zeigt nur, daß es Fälle gibt, in denen die Wissenschaft sinnvollerweise eine Erklärung relationaler Zustände zu entwickeln versucht. (Und zwar *ohne* daß man die Frage nach diesen relationalen Zuständen ohne erheblichen theoretischen Aufwand in eine Frage nach intrinsischen Zuständen umformulieren könnte.)

Dagegen könnte Fodor antworten: Wenn eine Theorie auch über Eigenschaften, die durch ihre Kausalgeschichte oder in anderer Weise relational identifiziert werden, generalisieren kann, so wäre es doch trotzdem methodologisch einfach *besser*, sie würde stattdessen in eine Theorie umgewandelt, die diese Eigenschaften *enger* identifiziert. Wären solche Generalisierungen nicht tiefer oder genauer? Generalisierungen über Herzschwäche und sportliche Aktivität oder andere relationale Eigenschaften werden immer Ausnahmen haben. Um sie dennoch als legitim bezeichnen zu können, müssen sie in irgendeiner Weise von intrinsischen Eigenschaften der Beobachtungsgegenstände abhängig gemacht werden und wir werden nicht wissen, weshalb die Generalisierungen funktionieren, bis wir nicht die relevanten intrinsischen Eigenschaften benennen können. Dann aber ist es wahrscheinlich, daß wir wissen, warum Schlich eine Million Liegestütze machen konnte und immer noch lebt.

Es ist natürlich, so Stalnaker, völlig zutreffend, daß das kausale Erklärungsmodell, das

---

<sup>1</sup>Man kann argumentieren, daß eine Liegestütze sich immer noch irgendwie auf die Welt bezieht, da sie Schwerkraft und einen Boden voraussetzt. Doch es wäre nicht schwer, ein Zwillingserde-Beispiel zu konstruieren, in dem das, was Zwillingss-Schlich macht, im Zwillingssdeutschen nicht als Liegestützen zählt.

Fodor darstellt, eine Spielart der wissenschaftlichen Erklärung ist. Doch es gibt drei Gründe, relationale, beziehungsweise kausalgeschichtliche Eigenschaften dennoch als Basis von Generalisierungen anzunehmen.

- 1) Wir brauchen diese Generalisierungen, um Erklärungen über interne Mechanismen überhaupt anstreben zu können. Das Tier beispielsweise hat bestimmte Kapazitäten, Feinde zu erkennen und sich vor ihnen zu schützen. Die Beschreibung des neurophysiologischen Apparates, der das Tier dazu bringt, auf die sinnliche Repräsentation der-und-der Art in der-und-der Weise zu reagieren ist Teil einer Erklärung im Sinne des Physikalisten Fodor. Doch damit ist nicht erklärt, warum *das Tier vor einem Feind flüchtet*. Der Grund, warum wir an der physikalistischen Erklärungskomponente interessiert sind, ist der, daß wir diese Komponente mit einer verhaltensinterpretierenden Komponente verbinden wollen, die sich auf relationale Eigenschaften bezieht: wir wollen wissen, *wie* das Tier in der Lage sein kann, bei der Repräsentation einer bedrohlichen Situation auf diese zu reagieren.
- 2) Wir wissen in der Regel nicht genug, um über interne Eigenschaften generalisieren zu können. In diesem Zusammenhang sind Black-Box-Experimente interessant. Wir wissen, wenn Zustand A der Black Box übergeben wird, erhalten wir Zustand B. Wir wissen vielleicht schon einiges über das Innere der Black Box, doch eben nicht genug, um unsere Fragen oder die vielleicht greifbaren partiellen Antworten in rein intrinsischen Begriffen formulieren zu können. Wir müssen daher auf den Fortschritt der Wissenschaft warten und uns einstweilen mit relationalen Begrifflichkeiten bescheiden.
- 3) Der wichtigste Grund: es gibt Generalisierungen, die nur in nichtinternalistischen Zuständen und Eigenschaften beschrieben werden *können*. Bei verschiedenen Lebewesen wird die physikalistische Erklärung des Mechanismus', der die Verhaltensänderung in Gegenwart von Feinden auslöst, verschieden ausfallen – man braucht nur Löwen mit Amseln und Igel zu vergleichen. Doch das Gemeinsame an diesen Erklärungen können wir gar nicht anders generalisieren, als über die *Eigenschaft*, auf die Gegenwart von Feinden zu reagieren. Der strenge Physikalist ist nicht in der Lage, in dem Verhalten von Amseln, Löwen und Igel, die mit einem Tyrannosaurus Rex konfrontiert werden, ein Merkmal zu erkennen, daß diese drei Verhaltensinstanzen zu Instanzen derselben Verhaltensklasse macht. Die Idee des Funktionalismus besteht darin, Entitäten anhand ihrer funktionalen Rollen zu identifizieren und solche Entitäten, die funktional identisch sind, als identisch anzusehen. Doch ich kann – auch in dem Black-Box-Experiment – immer zwei *Arten* von Fragen stellen: solche, die generell danach fragen, wie es sein kann, daß alle Black Boxes als Output B-Zustände produzieren, wenn man ihnen A-Zustände übergibt. Oder ich kann fragen, *wie* sie die A-Zustände bearbeiten, um B-Zustände zu produzieren. Zur Verdeutlichung: Wenn ich frage, *wie* das Chamäleon die Farbe des Hintergrundes annimmt, so würde der Physikalist korrekt antworten: weil in seiner Haut gewisse chemische Prozesse stattfinden. Wenn ich nun frage, *warum* das Chamäleon die Farbe des Hintergrundes annimmt, so würde der Physikalist dasselbe antworten. Was ich aber hören will, ist eine andere Antwort, die Relationalität und vor allem *Intentionalität* ausdrückt: *weil* die dadurch hergestellte Tarnung dem Chamäleon beim Überleben hilft.

Der Physikalist, der von intrinsischen Zuständen ausgeht, hat für diese Antwort gar kein Vokabular<sup>2</sup>.

Das führt uns zu generellen Überlegungen, mit deren Hilfe auch C1 sich als Scheinargument herausstellt: Jede psychologische Theorie, ob common sense oder Wissenschaft, versteht, sofern sie psychische Zustände relational identifiziert, Personen als Empfänger und Benutzer von Informationen. Nach Stalnaker ist es also die Pointe des Twin-Earth-Experiments, daß es zeigt, daß das Empfangen und Benutzen von Informationen nicht nur eine Tatsache über uns ist, sondern auch über die Art unseres Bezogenseins auf die Umwelt; und insofern ist Bedeutung und psychische Konstellation (im engen Sinne) externalistisch. Denn in verschiedenen Umgebungen können ununterscheidbare psychische Konstellationen verschiedene Informationen beinhalten und verarbeiten. Daraus läßt sich aber nicht einfach (wie in C1) schließen, daß wir keine Kontrolle über unsere Gedanken hätten. Selbstverständlich spricht nichts dagegen, daß der Sprecher die Art des Bezugs und den Umgang mit Informationen mitbestimmen kann. Zwar ist das dadurch noch nicht bewiesen, doch die Schlußfolgerungen aus dem Twin-Earth-Experiment führen ebensowenig deduktiv zu C1.

Es ist, wie Stalnaker hinzusetzt, trotz dieser recht einleuchtenden Ausführungen nicht so einfach, rein interne Elemente einer Überzeugung anzugeben. Der enge Inhalt taugt dazu aufgrund der Komplikationen, die seine Definition hervorruft, nicht. Allerdings ist das kein so furchtbar drängendes Problem mehr, denn wir haben gesehen, daß die Merkmale von semantischem Inhalt und funktionaler Rolle, die wir haben, wohl doch hinreichen, um die wichtigen Phänomene, die das Twin-Earth-Experiment zeigt, beschreiben zu können.

Wenn nun Schlich glaubt, sein Keller stehe unter Wasser, ist dann diese Überzeugung in seinem Kopf oder nicht? Stalnakers Antwort lautet: Natürlich ist die Überzeugung in seinem Kopf, genauso, wie der Mückenstich auf seiner Nase tatsächlich auf seiner Nase oder sein Fußabdruck im allmählich durchnässten Wohnzimmer Teppich wirklich in diesem Teppich ist. Wir können mithilfe dieser Überzeugung erklären, weshalb Schlich den Mop holt und die Feuerwehr ruft und warum er seine Nase kratzt. Wir identifizieren Überzeugungen und Eigenschaften in Begriffen, die ihre Interaktion mit der Umgebung aufzeigen und benützen diese Konstruktionen, um zu erklären, warum Agenten sich so verhalten, wie sie es tun. Es gibt keinen Grund, damit aufzuhören oder anzunehmen, wir könnten das nicht tun.

Damit ist im Gegensatz zu Fodor eine echte Aufarbeitung der Frage geleistet, inwiefern Gedanken im Kopf sind, weil Stalnakers Erklärung sowohl mit unseren Intuitionen und unserer wissenschaftlichen Praxis übereinstimmt, als auch eine Antwort auf C1-3 findet. Putnams Argumente haben damit ihren Platz in unserem Projekt der Frage nach dem Ort der Gedanken gefunden. Wir müssen, auf Stalnaker aufbauend, auch Freges Erklärung des „dritten Reiches“ nicht umstoßen, wenn wir postulieren, daß das Reich der Gedanken eine kognitive Qualität hat, andererseits jedoch genau wie das Bewußtsein im Kopf des Sprechers ist. Das würde dem Diktum der Öffentlichkeit von Gedanken nicht widersprechen, denn es wurde nie behauptet, daß etwas nicht im Kopf ausgewertet und dennoch öffentlich sein kann.

---

<sup>2</sup>Folgendes Beispiel macht die Sache noch klarer. Heike wirft mir vor: „Du hast keine Ahnung, wie es in mir aussieht!“ Als einfühlsamer Physikalist kaufe ich mir daraufhin einen Anatomieatlas. Es dürfte klar sein, daß alles, was ich aus dem Atlas erfahre, mich, sofern ich nicht allwissend bin, nicht in die Lage versetzt, den Vorwurf, den sie meint, auszuräumen.

Und den Satz des Pythagoras könnten wir als Ausdruck dieses Gedankens ansehen, als eine konkretisierte Information, die in Form einer Repräsentation in die Außenwelt eingegangen ist und die Überzeugung, daß im rechtwinkligen Dreieck die Summe der Kathetenquadrate gleich dem Hypothenusenquadrat ist, als Information transportiert, die von späteren Sprechern ausgewertet werden kann. Das paßt zusammen.

Es scheint, als hätten wir einen konsistenten Weg gefunden, die Relationalität von Identifizierungsstandards für propositionale Einstellungen zuzulassen. Das ermöglicht es uns, gegen Putnam zu behaupten, daß Gedanken im Kopf sind, ohne aber der bisher als einzige Alternative vorgestellten Ansicht anheimzufallen, daß Überzeugungen tatsächlich rein intrinsische Eigenschaften des Einzelnen sind. Die gleichzeitige Beibehaltung dieser beiden Prämissen erinnert stark an Freges Auffassung, wie sie in Kapitel I entwickelt wurde. Stalnakers Argumentation bietet einen Weg, den Internalismus um die Umgehensweise mit relationalen Eigenschaften zu erweitern, ohne seine Auffassung, Gedanken seien im Kopf, generell aufzugeben. Insofern kann man sagen, Stalnaker entnimmt der philosophischen Tradition des Internalismus die strenge Doktrin des methodologischen Solipsismus. Man kann jetzt natürlich streiten, ob er nicht doch Externalist ist. . .

## VI.2 Behauptungen und Propositionalkonzepte

Zwar ist dies eine erfreuliche Harmonisierung der bisherigen Widersprüchlichkeiten, doch fehlt noch die Semantik, die Stalnakers Auffassungen auch formal tragfähig macht. Wenn wir davon ausgehen, daß Gedanken im Kopf sind, dann ist klar, daß wir eine Semantik brauchen, die sich nicht auf Behauptungssätze, also geäußerte Überzeugungen beschränkt, sondern die nicht geäußerte Überzeugungen abbilden kann, sonst ist die Argumentation ohne Pointe.

Dafür brauchen wir aber ein neues Instrumentarium. Wenn Kaplans Semantik auch bisher eine gute Grundlage war, so muß sie hier aus einem einleuchtenden Grund versagen: der Begriff des Charakters ist konstruiert, um bestehende Bedeutungen zu beschreiben, nicht um über das Zustandekommen dieser Bedeutung etwas zu sagen oder zu sagen, was jemand sagen würde, wenn er andere Überzeugungen hätte oder eine andere Sprache spräche. Der Charakterbegriff orientiert sich an der aktuellen Beschaffenheit einer gegebenen Sprache, nicht aber daran, wie es auch hätte sein können. Überzeugungen jedoch geben meistens an, wie es auch hätte sein können, nämlich wenn sie falsch sind. Wenn Julia sagt (oder glaubt), Hesperus sei Mars, dann behauptet sie, daß eine mögliche Welt, in der Hesperus Mars ist<sup>3</sup>, der gegebenen Welt entspricht. Es ist nicht Kaplans Anspruch, in diesem Fall den abweichenden, nur von Julia verwandten Charakter der Wörter „Hesperus“ und „Mars“ zu untersuchen; selbst in diesem einfachen Beispiel würde das nicht funktionieren, weil wir nicht wissen, was für Fehlannahmen Julia sonst noch verwendet. Der Charakterbegriff trägt dem intentionalen Aspekt von Julias Behauptung keine Rechnung und hat deshalb

---

<sup>3</sup>Mit Kripke haben wir gesehen, daß so etwas nichts als mögliche Welt zählt. Stalnaker umgeht dieses Verbot auf elegante Weise, wie wir noch sehen werden. Fürs erste lassen wir diese kontrafaktische Annahme ausnahmsweise einmal zu. Wie wir nachher sehen werden, wird Kripkes Auffassung durch Stalnaker nicht revidiert.



trotz seiner Fähigkeit, einen bekannten Unterschied abzubilden, keinen zufriedenstellenden Erklärungswert.

Wenn ich also sage „Wasser ist naß“ und mein Zwilling sagt „Wasser ist naß“, dann hat zwar das Wort „Wasser“ in diesen beiden Sätzen unterschiedlichen Charakter. Doch das erklärt sich schlicht daraus, daß ich deutsch spreche und mein Zwilling zwillingsdeutsch. Daß zwei verschiedene Sprachen Homonyme mit unterschiedlichem Charakter hervorbringen, ist für sich genommen nichts Überraschendes.

In [Sta78] entwickelt Stalnaker einen dem Charakterbegriff formal ähnlichen Begriff, dessen funktionale Konzeption jedoch von diesem grundverschieden ist. Dies ist der Begriff des Propositionalkonzepts, den ich in diesem Abschnitt erklären möchte. Der Begriff des Propositionalkonzepts soll uns helfen, die oben beschriebene Semantik zu entwickeln. Stalnaker entwickelt diesen Begriff zunächst mit Bezug auf Behauptungsaussäuerungen, doch Propositionalkonzepte lassen sich – mit leichten Schwierigkeiten – auf Überzeugungszustände übertragen. Zunächst sehen wir uns an, wie Stalnaker mit Behauptungsaussäuerungen verfährt. (Ich beschränke mich bei der Darstellung im folgenden auf die semantischen Aspekte dieses Begriffs und stelle die von Stalnaker eingebundenen pragmatischen Aspekte zurück.)

Für eine Behauptungsaussäuerung gilt, daß sie einen Inhalt hat und in einem Kontext steht und der Inhalt, das wissen wir, in vielen Fällen vom Kontext abhängt. Um nun zu zeigen, wie Inhalt und Kontext innerhalb einer Theorie der gesprochenen Sprache interagieren, wählt Stalnaker den Begriff des Propositionalkonzepts, den er wie folgt herleitet.

Wir erinnern uns, daß Sätze Propositionen repräsentieren. Propositionen können in Form ihrer Intension dargestellt werden: als Funktionen von möglichen Welten in Wahrheitswerte (welche bekanntlich die Extensionen von Sätzen sind).

Der Wahrheitswert der Proposition  $P$ , die durch den Satz „Schlamm ist ein gutes Shampoo“ ausgedrückt wird, könnte demnach für die möglichen Welten  $w_i, w_j, w_k$  folgendermaßen aussehen:

	$w_i$	$w_j$	$w_k$
$P$	falsch	wahr	falsch

Nun ist es zwar so, daß eine Proposition (oder ein Gedanke, wir können ruhig dabei bleiben) eine Menge möglicher Welten festlegt, in denen sie wahr ist, ganz einfach indem die Fakten in dieser Welt die Proposition erfüllen. Doch es gibt noch eine andere, zweite Art, in der Propositionen mit möglichen Welten verknüpft sind.

Wir haben gesehen, wie es sich mit dem Satz „Das da ist naß“ verhält. Ich äußere, auf Wasser<sub>E</sub> deutend, mit diesem Satz die Behauptung, H<sub>2</sub>O habe die Eigenschaft, naß zu sein, mein Zwilling dagegen äußert die Behauptung, XYZ habe die Eigenschaft, naß zu sein. Das sind verschiedene Propositionen. Also kann eine Äußerung, die Indexikale enthält, verschiedene Propositionen instanziiieren.

Doch wir müssen differenzieren: Neben der offenkundigen Indexikalität, die Kaplan aufgezeigt hat, steht Putnams Entdeckung, daß es Wörter und Ausdrücke gibt, die eine Art versteckte Indexikalität besitzen, nämlich die, daß ihre Extension in der gegebenen Welt der Extension, die der Ausdruck in unserer Welt hat, hinreichend ähnlich sein muß, damit eine

korrekte Verwendung des Ausdrucks vorliegt, daß also eine gewisse Kontextabhängigkeit in Bezug auf die Korrektheit der Verwendung solcher Ausdrücke besteht. Putnams Beispiel waren natürliche Prädikate wie zum Beispiel „ist Wasser“.

Aus dieser versteckten Indexikalität folgt natürlich, daß auch eine Äußerung ohne Kaplansche Indexikale verschiedene Propositionen instanziiieren kann. Allerdings insofern, als daß die Proposition, die durch die Äußerung ausgedrückt wird – wie bei den Kaplanschen Indexikalen – selbst etwas ist, das von Welt zu Welt variiert. Wenn ich sage „Wasser ist naß“ und mein Zwilling sagt „Wasser ist naß“, dann instanziiieren unsere Äußerungen nicht denselben Gedanken.

Diesen zweiten Aspekt kann man sich auch klar machen, indem man sich erinnert, daß jeder Adressat eine Äußerung auch falsch verstehen kann. Wenn zum Beispiel jemand „Schlamm“ für eine Shampoosmarke hält, dann versteht er den Inhalt meiner Äußerung „Schlamm ist ein gutes Shampoo“ in der Weise, daß er denkt, sie drücke eine Proposition aus, die besagt, daß das Shampoo der Marke *Schlamm* ein gutes Shampoo sei. Ich meinte diese Äußerung aber in dem Sinne, daß Matsch oder Lehm, wie er draußen vorkommt, ein gutes Shampoo sei. Das sind höchst verschiedene Propositionen.

Diesen zweiten Aspekt des Zusammenhangs zwischen Propositionen und möglichen Welten nun kann man mithilfe des Propositionalkonzepts der entsprechenden Äußerung abbilden. Das Propositionalkonzept einer Behauptungsäußerung ist eine Funktion von möglichen Welten in Propositionen. Der Definitionsbereich dieser Funktion sind mögliche Welten, die als Äußerungskontexte für diese Behauptung in Frage kommen, der Wertebereich ist die Menge der Propositionen, die die Äußerung faktisch und kontrafaktisch ausdrücken kann.

Um zu verdeutlichen, wie ein Propositionalkonzept funktioniert, hier folgendes Beispiel, das [Sta78] entnommen ist:

Ich sage zu O’Leary: „Du bist ein Trottel.“ Obwohl (oder: weil?) er ein Vollidiot ist, hält O’Leary sich natürlich nicht für einen Trottel. Daniels, der die Äußerung zufällig mitbekam, denkt, ich hätte ihn, Daniels, damit gemeint. Daniels ist allerdings kein Trottel und das weiß er auch. Allerdings hält Daniels O’Leary korrekterweise für einen Trottel. (O’Leary dagegen hält Daniels fälschlicherweise für einen ziemlichen Dummkopf.) Es sei nun  $w_i$  die gegebene Welt,  $w_j$  die Welt, wie O’Leary sie sieht und  $w_k$  die Welt, wie Daniels sie sieht. Dann legt die Äußerung „Du bist ein Trottel“ folgendes Propositionalkonzept fest:

	$w_i$	$w_j$	$w_k$
„Du bist ein Trottel.“	$P_i$ wahr	falsch	wahr
	$P_j$ wahr	falsch	wahr
	$P_k$ falsch	wahr	falsch

Die Horizontalen dieser Tabelle, die das Propositionalkonzept des Satzes „Du bist ein Trottel“ wiedergibt, entsprechen dabei dem jeweiligen Inhalt, den die Äußerung in jeder Welt hat, also der Proposition, die sie ausdrückt. Zum Beispiel läßt  $P_i$  sich durch „O’Leary ist ein Trottel“ ausdrücken,  $P_k$  dagegen durch „Daniels ist ein Trottel“.

Die Vertikalen dieser Tabelle geben dagegen den Äußerungskontext der Äußerung wieder, die Spalte  $w_i$  ist die gegebene Welt als Äußerungskontext,  $w_j$  ist die Welt, wie O’Leary sie

sieht, als Äußerungskontext und  $w_k$  ist der Äußerungskontext der Welt, wie Daniels sie sieht.

Also sehen wir, daß der Satz, den ich meine (entspricht  $P_i$ ) in  $w_i$  tatsächlich wahr ist, weil O'Leary dämlich ist. Dagegen versteht Daniels den Satz so, als sei er selbst damit gemeint (entspricht  $P_k$ ), so daß, da er ein schlauer Kerl ist, „sein“ Satz sowohl in  $w_i$  als auch in  $w_k$  falsch ist, nicht jedoch in  $w_j$ , weil der Idiot O'Leary ja irrtümlich glaubt, Daniels sei blöd.

Hinsichtlich des (jeweiligen!) *Inhalts* meines Satzes stimmen O'Leary und ich überein, daher sind die Horizontalen von  $w_i$  und  $w_j$  identisch. Hinsichtlich des Wahrheitswertes der von mir intendierten Proposition ( $=P_i$ ) stimmen Daniels und ich überein, deshalb sind die Vertikalen von  $w_i$  und  $w_k$  identisch. Wir sehen also, daß das Propositionalkonzept – anders als der Charakterbegriff – Intentionalität berücksichtigen kann. Dies ist die wichtigste Voraussetzung, um als Instrument für eine Semantik der Überzeugungen funktionieren zu können. Wir haben endlich ein subjektivitätskompetentes Instrument zur Abbildung von semantischen Beziehungen gefunden!

Jede konkrete Äußerung kann mit dem Propositionalkonzept, das sie festlegt, assoziiert werden. Allerdings kann man ein solches Propositionalkonzept durchaus auch anders darstellen, zum Beispiel als Funktion eines geordneten Paares möglicher Welten in Wahrheitswerte. Die hier gewählte Form ist dieselbe, die ich in Kapitel III für den Charakterbegriff gewählt habe (der sich ebenfalls anders darstellen ließe), und die das Verständnis der Tabellen, die ich im folgenden Propositionaltabellen nennen will, erleichtert.

Stalnaker berücksichtigt, wie ich schon andeutete, pragmatische Komponenten. Auch im Propositionalkonzept sind pragmatische Komponenten insofern berücksichtigt, als daß die vertikal aufgetragenen Äußerungskontexte so beschaffen sind, daß sie den jeweiligen *Präsuppositionen* der Sprecher oder Agenten entsprechen. Zum Beispiel präsupponiert O'Leary, daß Daniels ein Trottel sei, er selbst hingegen nicht, während diese Auffassung weder Teil meiner noch Daniels' Sprecher-Präsupposition sind.

Diese Auffassung von Präsupposition unterscheidet sich ein wenig von der in Kapitel I verwandten Begrifflichkeit: hier sind wirklich Annahmen des jeweiligen Sprechers beziehungsweise Agenten gemeint, nicht Sätze. Außerdem geht es dabei nicht wie bei Strawson um Wahrheitswertfähigkeit, sondern schlicht um Information und die Festlegung möglicher Welten: eben um das im Äußerungskontext vom Sprecher für wahr gehaltene. Diesem Unterschied werde ich Rechnung tragen, indem ich Stalnakers Konzeption im folgenden als Sprecherpräsupposition bezeichne.

Die Sprecherpräsupposition ist die Menge aller Sätze, deren Wahrheitswert der Sprecher vor dem Hintergrund der Äußerung als bekannt voraussetzt. Sprecherpräsuppositionen sind also gewissermaßen die jeweils angenommene Gesprächsgrundlage, obwohl natürlich nicht alle Sprecherpräsuppositionen faktisch den vermuteten Wahrheitswert haben müssen. Noch dazu tritt in der normalen Sprache wohl nie der Fall ein, daß eine Gruppe von Teilnehmern einer Konversation genau dieselben Sprecherpräsuppositionen hat.

Sprecherpräsuppositionen kann man in Form möglicher Welten repräsentieren, die mit den jeweils präsupponierten Propositionen kompatibel sind. Die Menge dieser Welten ist das sogenannte *Kontextset* des Sprechers. Wenn die Kontextsets aller Teilnehmer identisch sind, liegt ein nichtdefekter Kontext vor, andere Kontexte sind defekt und haben eine Tendenz, einen Zustand höherer Stabilität einzunehmen, das heißt, in den nichtdefekten Zustand zu tendieren. Alle normalen Kontexte sind demzufolge defekt, doch ein Kontext ist hinreichend

nichtdefekt, wenn die Divergenzen in den Kontextsets der Sprecher den Gegenstand der Konversation nicht berühren. Das ist eine pragmatische Konvention, die der getroffenen Unterscheidung Tragkraft ermöglichen soll.

Das heißt nun, daß jede Welt, die zum Kontextset des Sprechers gehört, als Äußerungskontext in Frage kommt, was bedeutet, daß  $w_i$ ,  $w_j$ , und  $w_k$  riesige Klassen möglicher Welten repräsentieren. Innerhalb dieses Kontextsets will der Sprecher mit der Behauptung, die er macht, differenzieren, denn jede Behauptung reduziert das Kontextset um genau die Welten, die mit der Wahrheit dieser Behauptung nicht kompatibel sind.

Doch ich „spreche“ in der Klasse  $w_i$  und die besteht aus allen Welten, in denen O’Leary existiert und ein Trottel ist. Das heißt, daß eine Behauptungsäußerung eines Sprechers eine Proposition instanziiert, aber nicht festlegt. Relativ zu allen mit den Sprecherpräsuppositionen kompatiblen Welten legt diese Behauptung zwar ein Propositionalkonzept fest, aber keine Proposition.

Wie kann man nun mithilfe dieser Unterscheidungen die von Stalnaker versprochene Lösung umsetzen? Wir werden, nachdem wir im nächsten Abschnitt den Begriff des Propositionalkonzepts noch etwas gegen Kaplans Charakterbegriff abheben, im folgenden sehen, wie Stalnaker die Voraussetzungen einer solchen Lösung bestimmt und dann seine Lösung in zwei Stufen verwirklicht; Stufe 1 wendet das Instrument des Propositionalkonzepts auf Behauptungsäußerungen an und Stufe 2 überträgt die Ergebnisse auf Überzeugungszustände.

### VI.3 Kaplan und Stalnaker

Es ist wichtig, den Begriff des Propositionalkonzepts vom Charakterbegriff zu unterscheiden. Die einzigen Ähnlichkeiten zwischen diesen Begriffen bestehen darin, daß beide Begriffe, auf Sätze angewandt, Funktionen von möglichen Welten in Wahrheitsbedingungen sind<sup>4</sup>. Außerdem repräsentieren beide Begriffe eine zweidimensionale Konzeption, wie man an der Darstellungsweise, die ich für Propositionalkonzepte gewählt habe, gut sehen kann.

Doch der Charakter gibt die kognitive Signifikanz eines Wortes an, indem er syntaktischen Ausdruckstypen Referenzobjekte zuordnet; insofern ist er eine wichtige Variante dessen, was wir intuitiv als „Bedeutung“ bezeichnen. Sätze haben in gegebenem Kontext den Inhalt, den sie haben, *weil* sie den Charakter haben, der diesem Kontext diesen Inhalt zuordnet.

Wie wir gesehen haben, ist der Charakter ein objektiver Bedeutungs begriff. Ein Propositionalkonzept dagegen gibt keine objektive Bedeutung wieder. Ein Propositionalkonzept gibt eine Verknüpfung zwischen Äußerungsinstanzen und Kontexten an, wobei das Propositionalkonzept seinerseits durch die Äußerung festgelegt wird. Derselbe Satztyp mit demselben Charakter kann, in verschiedenen Kontexten, verschiedene Propositionalkonzepte festlegen. Äußerungen haben den Inhalt, den sie haben, nicht, *weil* sie ein bestimmtes Propositionalkonzept festlegen. Insofern ist das Propositionalkonzept ein subjektiver und relativer Bedeutungs begriff, der die kognitive Signifikanz einer Äußerung abbilden kann und nicht auf

---

<sup>4</sup>Das ist der enge Inhalt auch und der partielle Charakter sieht, auf Sätze angewandt, auch so aus. Diese Tatsache allein ist keine Erfolgsgarantie und kein Güte Merkmal. Auch sagt sie nichts über theoretische Ähnlichkeit zwischen verschiedenen Funktionsbegriffen aus.

Sätze ohne Kenntnis des Kontexts anwendbar ist<sup>5</sup>.

Ein Propositionalkonzept ist nicht nur keine objektive Bedeutung, es ist nicht einmal eine Funktion von objektiver Bedeutung. Um ein solches Konzept formulieren zu können reicht die Kenntnis der faktischen Bedeutung nicht aus, man muß auch Kenntnis gewisser kontrafaktischer Bedeutungen haben.

Dennoch kann ein Propositionalkonzept für das Projekt einer rekursiven Semantik wertvoll und nützlich sein. Das werden wir in den Abschnitten 5 und 6 sehen.

## VI.4 Voraussetzungen für eine Theorie der Überzeugung

Eine Überzeugungszuschreibung „ $x$  glaubt, daß  $P$ “ setzt ein Individuum  $x$  in eine bestimmte Relation zu einem Objekt der Überzeugung, das durch „daß  $P$ “ ausgedrückt wird, egal wie man es definiert. Der Einfachheit halber kann man es als Proposition approximieren. Interessant ist nun aber die Beschaffenheit der Überzeugungsrelation. Für die Grundlegung einer Semantik der Überzeugungen benötigen wir hierfür eine Theorie.

Diese Theorie muß einerseits die semantischen Fakten über Überzeugungszuschreibungen berücksichtigen, wonach für jeden Satz  $P$  ein äquivalenter Satz  $Q$  gefunden werden kann, derart, daß es ein Individuum  $x$  gibt, für das gilt, daß es von  $P$ , doch nicht von  $Q$  überzeugt ist (erinnern wir uns an das Problem aus I.6 und die Beispiele in V.2). Dies ist ein starkes Argument für die Auffassung, daß Überzeugungsobjekte Sätze oder satzähnliche Dinge wie zum Beispiel Propositionen sind.

Andererseits aber müssen wir unsere Folgerungen aus Abschnitt 1 berücksichtigen: Intentionalität läßt sich nicht wegerklären. Sie muß in einer Semantik der Überzeugung ihren Platz finden. Daher muß die Theorie eine philosophische Erklärung der Intentionalität ebenso wie eine empirische Erklärung kognitiver Zustände und Prozesse ermöglichen, die miteinander in Einklang stehen. Dies allerdings legt nahe, daß die Glaubensrelation zwischen einem Individuum („Glauber“ [*believer*]) und einem Inhalt besteht, der nicht mit einem Satz zusammenfällt, ähnlich, wie Loar vermutet hat.

Hier besteht eine Spannung: man muß, um einen subjektivitätskompetenten Bedeutungsbegriff zu erhalten, ein Konzept des Inhalts entwickeln, das Sätze semantisch unterscheidet, die auch von der Überzeugungszuschreibung kognitiv unterschieden werden. Das ist keine leichte Aufgabe.

Sehen wir uns ein kondensiertes Beispiel an:

- 1) Notwendig äquivalente Überzeugungsobjekte sind identisch.
- 2)  $P$  und  $Q$  sind äquivalent.
- 3) Es ist möglich, von  $P$  überzeugt zu sein, ohne von  $Q$  überzeugt zu sein.

Bereits Kripke konstatierte diesen Sachverhalt in [Kri79]. Loar zog, wie wir gesehen haben, daraus den Schluß, daß man einen entsprechend nichtpropositionalen Begriff von In-

---

<sup>5</sup>Stalnaker selbst faßt den Bedeutungsbegriff noch enger und meint, das Propositionalkonzept sei im Grunde gar keine Spielart intuitiver Bedeutung.

halt bräuchte, auf den 1) und 2) einfach nicht mehr anwendbar sind. Er wandte sich insofern gegen den in 3) ausgedrückten Begriff des Überzeugtseins.

Stalnaker jedoch attackiert These 2) und behauptet, daß Sätze, die notwendig äquivalent sind, nicht in allen Kontexten notwendig äquivalent und also nicht in allen Kontexten äquivalent sein müssen. Wie könnte man das zeigen?

Man muß einen Kontext für eine gegebene Überzeugungszuschreibung definieren und dann eine kontingente Proposition finden, die in diesem Kontext definiert ist und *intuitiv* das ausdrückt, wovon die Überzeugungszuschreibung sagt, daß der Sprecher davon überzeugt sei. Gelingt dies, kann geschlossen werden, daß der Komplementsatz der Überzeugungszuschreibung diese Proposition denotiert.

Danach muß man eine systematische Erklärung dafür finden, daß die Proposition mit dem Komplementsatz verknüpft ist. Das werden wir versuchen und wir werden sehen, daß wir das leidige Zwillingserde-Problem damit endlich in den Griff bekommen.

## VI.5 1. Stufe: die Semantik der Behauptungen

Nehmen wir also an, wir haben die Twin-Earth-Zwillinge beide hier auf der Erde vor uns. Der Erdling sagt: „Wasser ist  $H_2O$ “ und der Zwerdling sagt: „Wasser ist XYZ“.

Gehen wir zu der eben skizzierten Strategie über:

Frage 1: Welche kontingente Proposition repräsentiert die Informationen, die durch die Aussagen jeweils gegeben werden?

Frage 2: Wie kann der Komplementsatz diese Information transportieren?

Zu Frage 1: Wie würde die Welt sein, wenn der Zwerdling recht hätte? Ganz einfach: überall auf unserem Planeten Erde würde statt  $H_2O$  ein wasserartiger Stoff vorkommen, der nicht die molekulare Struktur  $H_2O$ , sondern XYZ hat und dieser Stoff müßte naß sein. Nennen wir eine beliebige mögliche Welt aus der Klasse der möglichen Welten, die mit dieser Beschreibung kompatibel sind,  $w_x$ . Die gegebene, wirkliche Welt sei  $w_0$ . Der Zwerdling sagt dann mit seiner Behauptung, daß  $w_0 = w_x$  sei. Dann dürfen wir die These aufstellen, daß der Inhalt der Behauptung des Zwerdlings eben jene kontingente Proposition ist, die in  $w_x$  wahr und in  $w_0$  falsch ist und die wir durch den Satz „Wasser ist XYZ“ ausdrücken.

Doch jetzt greift Kripkes Auffassung ein: wir müssen annehmen, daß „Wasser“ und „ $H_2O$ “ starre Designatoren sind und wollen behaupten, daß „ist“ in dem Satz des Zwerdlings ein Symbol für eine Identitätsbehauptung ist. Dann stecken wir in denselben Schwierigkeiten, in denen Freges Lösung des Informativitätsproblems nach Kripke steckt. Die Proposition, die der Zwerdling als Inhalt seines Satzes ausdrückt, ist, wie wir in Kapitel II gesehen haben, *notwendig* falsch, auch in  $w_x$ . Es ist keine Welt möglich, in der Wasser nicht  $H_2O$  ist, obwohl wir uns intuitiv *vorstellen* können, daß es so sei. Dies genau ist der in Abschnitt 4 beschriebene Konflikt!

Hier setzt nun Stalnakers Argumentation ein: Stalnaker fragt durch das Propositionalkonzept nicht nur, wie Kripke, was der Satz *sagt*, sondern auch, was er sagen *würde*, wenn er

in verschiedenen möglichen Welten geäußert *würde*. Nehmen wir zum Beispiel an, der zwil-  
lingsdeutsche Satz „Wasser ist naß“ würde in einer möglichen Welt  $w_x$  geäußert, in der XYZ  
statt Wasser<sub>E</sub> Flüsse und Seen füllt und vom Himmel fällt. Dann sähe das Propositionalkon-  
zept dieser Äußerung so aus:

	$w_x$	$w_0$
„Wasser ist naß“	$P[w_x]$	wahr
	$P[w_0]$	falsch

Stalnaker wendet hier einen sehr scharfsinnigen Trick an: in  $w_x$  nämlich bezeichnen „Was-  
ser“ und „XYZ“ starr dieselbe Sache, so daß der zwillingsirdische Satz im Kontext  $w_x$  eine  
notwendige Wahrheit ausdrückt. Das heißt, die starre Designation ist kein Hindernis mehr,  
eine solche Welt als mögliche zuzulassen.

Doch wir suchen die kontingente Proposition, die der Intuition nach durch den zwil-  
lingsdeutschen Satz „Wasser ist naß“ ausgedrückt wird. Diese Proposition ist offensichtlich in  
keiner der Horizontalen zu finden, denn  $P[w_x]$  (=„Wasser<sub>TE</sub> ist naß“) ist notwendig wahr  
und  $P[w_0]$  (=„Wasser<sub>E</sub> ist naß“) ist notwendig falsch. Wir suchen aber eine Proposition,  
deren Wahrheitswerte wechseln. Diese kontingente Proposition, so Stalnakers Argument, ist  
offensichtlich die Proposition, die genau dann wahr ist, wenn der horizontale Eintrag der  
Proposition für eine gegebene Welt *in* dieser Welt wahr ist, das heißt, wenn die Proposition,  
die in einem Kontext  $k$  steht, in diesem Kontext wahr ist.

Innerhalb des Propositionalkonzepts wird diese Proposition durch die *Diagonale* repräsen-  
tiert. Die Diagonalproposition des Propositionalkonzepts für „Wasser ist naß“ drückt genau  
die Proposition aus, die dann wahr ist, wenn der Satz „Wasser ist naß“ im betreffenden Kon-  
text wahr ist. Auf diese Art läßt sich, wie leicht zu sehen ist, auch die versteckte Indexikalität  
von „Wasser“ repräsentieren.

Unser Lösungsangebot für Frage 2 und das Dreisatzdilemma aus Abschnitt 4 muß also –  
bezogen auf Äußerungen – lauten:

[H1] *Unter bestimmten Bedingungen ist der Inhalt einer Aussage nicht die Proposition, die  
durch normale semantische Regeln festgelegt wird, sondern die Diagonale des festgelegten Pro-  
positionalkonzepts.*

Mit anderen Worten: für Propositionen, die durch Diagonalen repräsentiert werden, kann  
man systematisch Twin-Earth-Beispiele erzeugen, womit ganz nebenbei auch noch ein ex-  
tensionales Kriterium zur Darstellung versteckter Indexikalität gefunden ist. Das Prinzip  
der Diagonalisierung trägt immer dann, wenn durch Fehlverständnis oder Indexikalität Be-  
deutungsunterschiede ermöglicht werden. Darüber hinaus steht uns jedoch frei, es immer  
anzuwenden, wenn es geboten erscheint.

Insofern haben wir jetzt auch erklärt, was die Zwillinge tatsächlich gemeinsam haben:  
wenn wir den Satz „Wasser ist naß“ ohne Angabe in welcher Sprache er steht, in ein Pro-  
positionalkonzept übertragen, so ist der Satz in allen und genau den Kontexten wahr, wo  
das, was „Wasser“ heißt, naß ist. Diese Proposition würde durch die Diagonale des entspre-  
chenden Propositionalkonzepts ausgedrückt. Also haben wir endlich das erhoffte Kriterium

gefunden, mit dessen Hilfe sich repräsentieren läßt, was es in semantischer Hinsicht ist, das die Zwillinge gemeinsam haben: die Diagonale des Satzes „Wasser ist naß“.

## VI.6 2. Stufe: die Semantik der Überzeugungen

Wir müssen unser überaus zufriedenstellendes Ergebnis jetzt von *Überzeugungsäußerungen* auf *Überzeugungszustände* übertragen.

S19) Schlich sagt: „Wasser ist naß“.

S20) Schlich glaubt, daß Wasser naß ist.

Das Hauptproblem dabei besteht darin, daß wir die Verbindung zwischen dem Gedanken (der Proposition) und dem mitgeteilten Satz nicht kennen, so daß wir den Lösungsansatz aus VI.5 nicht einfach so übertragen können.

H1 erklärt, wie der Komplementsatz in S19 den Inhalt haben kann, den er hat. Wenn ich glaube, daß „das da“ (=XYZ) Wasser sei, so könnte aber eine Hypothese H2, die H1 ähnlich ist, erklären, wie dieses mentale Ereignis diesen Inhalt haben kann, wie also der Komplementsatz von S20 zu seinem Inhalt kommt.

Allerdings gibt es einen wichtigen Unterschied zwischen einer Theorie für Äußerungen und einer für Gedanken: in Abschnitt 5 haben wir es mit Äußerungen zu tun, das heißt, der Sprecher ist derselbe wie der, dessen Einstellungen die Worte ausdrücken sollten – mit-hin mußten wir nur *eine* Perspektive berücksichtigen. Verallgemeinern wir das jedoch auf Überzeugungen, so kann sich die Sichtweise des Sprechers von der des Adressaten unterscheiden, ebenso wie das Kontextset von Sprecher und Adressat sich unterscheiden kann. Wenn ich sage: „Schlich glaubt, Wasser sei naß“, und Schlich also diese Überzeugung zuschreibe, dann wird diese Überzeugungszuschreibung bei meinem Zwilling als Aussage über XYZ ankommen. Auch hier zeigt sich eine Parallelität zu Kaplan: der Hintergrund, vor dem die Aussage ausgewertet wird, kann sich von dem Hintergrund unterscheiden, vor dem sie gemacht wurde, was parallel zu einer Unterscheidung zwischen Kontext und Auswertungssituation ist.

Das macht es aber schwierig, Diagonalisierung als Lösungsmethode anzuwenden, weil die Kriterien für die Formulierung der Propositionalkonzepte schwerer sind:

Ich weiß, daß Wasser<sub>E</sub> H<sub>2</sub>O ist. Daniels weiß auch, daß Wasser<sub>E</sub> H<sub>2</sub>O ist. Nur O’Leary hat etwas durcheinander gebracht und glaubt, Wasser<sub>E</sub> sei XYZ. Daniels sagt nun zu mir: O’Leary glaubt, daß Wasser XYZ ist“.

Das Problem besteht nun darin, daß Daniels und ich nichts über O’Learys Kontextset wissen oder darüber, wie es in ihm aussieht. Daher können wir nicht sagen, welche Proposition der von Daniels geäußerte Komplementsatz in den Welten ausdrückt, die mit O’Learys Überzeugungen kompatibel sind. In allen Welten, über die wir nichts sagen können, existiert diese Äußerung daher nicht.

Aber wir können das Propositionalkonzept erweitern:



Wenn Daniels in einer O'Leary-kompatiblen Welt die gleiche Äußerung täte, wie wäre deren Inhalt? Wenn die physikalische Beschaffenheit unseres Planeten entsprechend aussähe, würden wir (drei!) „Wasser“ sagen, um XYZ zu bezeichnen. In dieser Welt drückt Daniels' Komplementsatz – laut Kripke – eine notwendige Wahrheit aus.

Wenn wir das Propositionalkonzept in dieser Weise erweitern und es für alle Situationen definieren, die O'Leary-kompatibel sind (das heißt: es sein müßten, nach dem, was ich und Daniels wissen), dann entspricht die Diagonale intuitiv der Proposition, den O'Leary Daniels zufolge glaubt.

Das Propositionalkonzept entspricht dabei nicht dem, das O'Leary selbst benutzen würde, sondern dem, das Sprecher (Daniels) und Adressat (Hohenadel) benutzen und verstehen würden, wenn sie in einer möglichen Welt wären, in der es definiert ist.

Dabei ist es auch egal, welche Sprache O'Leary spricht: es geht um Propositionen, nicht um Sätze. Uns interessiert die ausgedrückte Proposition, egal ob Daniels oder O'Leary deutsch oder eine andere Sprache sprechen. O'Learys Überzeugung bezieht sich auf Wasser, nicht auf Sprache.

Eine wichtige Eigenschaft der so skizzierten Semantik ist es, daß sie auch für linguistische Überzeugungen funktioniert. Wenn Schlich zum Beispiel englisch lernt und irrtümlich annimmt, daß „a fortnight“ für eine Periode von 10 Tagen stehe, ist dies eine linguistische Überzeugung über die Bedeutung von Wörtern. Wenn diese Überzeugung wahr wäre, dann ist eine Welt, die mit Schlichs Überzeugung kompatibel ist, eine solche, in der „a fortnight“ starr 10 Tage bedeutet. In solchen Welten ist dieser Satz wahr, ansonsten falsch, und er ist die Diagonale des entsprechenden Propositionalkonzepts.

In Fällen, wo wir nicht diagonalisieren, können wir subjektunabhängiges Vokabular benutzen. Nehmen wir an, ich und Daniels haben eine Privatsprache, in der das Wort „Gluck“ ein starrer Bezeichner für Wasser<sub>E</sub> ist. Dann kann Daniels eine korrekte Überzeugungszuschreibung vornehmen, wenn er sagt: O'Leary glaubt, daß Gluck XYZ ist“, obwohl O'Leary natürlich den Designator „Gluck“ gar nicht in seinem sprachlichen Repertoire hat.

## VI.7 Frege und Stalnaker

Die Folgerung aus all dem ist, daß Überzeugungszuschreibungen in höchstem Maße kontextabhängig sind. Insofern stärkt Stalnaker Putnams semantische These von der versteckten Indexikalität. Ebenso läßt er Raum für Putnams Auffassung über die Relevanz von Sprachgemeinschaft und Welt für die Formung von Bedeutungsbegriffen und bejaht auch die von Putnam unkommentiert vollzogene relationale Identifizierung von Überzeugungszuständen. Insofern ist Stalnakers Konzeption externalistisch.

Nur zieht Stalnaker aus alledem andere Schlußfolgerungen als Putnam und auch andere als die Theoretiker des engen Inhalts. Die Semantik, die als Folge dieser Schlußfolgerungen entsteht, hebt sich von der Kaplans trotz formaler Ähnlichkeiten durch tiefgreifende theoretische Unterschiede ab, ist jedoch mit Kripkes Thesen zur starren Designation konform. Letztlich sehen wir, daß Freges Lösungsversuch für die Verortung von Vorstellungen und Gedanken so falsch nicht war. Zwar stützt Stalnaker Freges Auffassung über die Sonderrolle von Gedanken nicht explizit, doch läßt er für Freges Auffassung Raum. Stalnakers Semantik

kann als semantische Beschreibung des „dritten Reiches“ gelten, das Frege annahm.

Man sieht also, wie die Inhalte aller bisher besprochenen Kapitel von Stalnakers theoretischem Konzept berücksichtigt werden. Frege, Kripke, Kaplan und der Externalismus finden darin ihren Platz, ebenso werden die Argumente der enger-Inhalt-Theoretiker zur Kenntnis genommen. Es scheint, daß sich letztlich doch noch alles zu einem halbwegs harmonischen Bild zusammengefügt hat. Mit Stalnakers Ansatz zur Realisierung der subjektivitätskompetenten Semantik ist daher unsere Annäherung an die internalistische Bedeutungstheorie gewissermaßen zu einem natürlichen Punkt des Abschlusses gelangt.

Wir haben gesehen, daß es keine rein internalistische Bedeutungstheorie geben kann, weil jede Semantik auch Überzeugungszuschreibungen systematisch Bedeutung zuordnen muß – und diese Überzeugungszuschreibungen denotieren Überzeugungen *über* die Welt, die relationale Eigenschaften des gegebenen Individuums sind.

Andererseits haben wir gesehen, daß die Annahme, Gedanken seien nicht im Kopf, in dieser plakativen Form zu stark zu sein scheint. Wir können kohärent annehmen, daß Gedanken trotz ihrer relationalen Eigenschaften genauso im Kopf sind, wie eine Mückenstich auf der Nase sein kann.

Diese hybride Gestalt der Gedanken war es, die Frege zu der Annahme veranlaßt hat, Gedanken befänden sich in einem „dritten Reich“, eben in einem Bereich, der nicht mit den rein intrinsischen Zuständen des Individuums – dem Bewußtsein – zusammenfällt, doch der auch nicht in dem Sinne zur Außenwelt gehört, wie Tische, Stühle und Steine zur Außenwelt gehören. Stalnaker hat diese Annahme in seiner Semantik der Überzeugungen erheblich präzisiert und er hat einen – noch nicht genug systematisierten – Bedeutungsbegriff zur Verfügung gestellt, der sowohl formal-semantische Merkmale (wie bei Kaplan) als auch kognitiv-epistemische Merkmale (wie durch Putnam aufgezeigt) abbilden kann. Einen solchen Begriff bleibt Frege schuldig.

Frege hat die richtigen Fragen gestellt, nur waren seine Antworten nicht immer ausreichend. Seine Bedeutung in der Philosophiegeschichte ist noch nicht absehbar, weil all die Fragen und Probleme, die er aufgezeigt, und all die Steine, die er ins Rollen gebracht hat, noch nicht zu einer neuen Konstellation gefunden haben. Mit Stalnaker jedoch wird deutlich, wohin man kommt, wenn man Freges Problemstellungen konsequent folgt: zum Ansatz einer semantischen Beschreibung von Denkvorgängen. Die Verachtung, die Frege von seinen Zeitgenossen erfuhr, wird angesichts der Reichweite und Relevanz der Fregeschen Problemstellungen zum Armutszeugnis für Freges frühe Kritiker.

Dieses Projekt sprengt die Grenzen der Bedeutungstheorie und daher endet diese Arbeit hier. Eine größere Arbeit würde sich nun den durch Stalnaker aufgebrachten Möglichkeiten und Thesen widmen und in eine Abhandlung münden, die die Philosophie des Geistes ebenso wie die philosophische Semantik und die Erkenntnistheorie signifikant berühren.

## Nachwort

Natürlich ist dies keine vollständige Geschichte der Bedeutungstheorie – doch das sollte es auch nicht sein. Ich will aus Gründen der Vollständigkeit (nicht der Entschuldigung) noch auf einige Positionen und Texte hinweisen, die hier nicht besprochen werden konnten.

Zum Beispiel wurde auf eine Besprechung der Semantik Donald Davidsons verzichtet, die in jede Geschichte der philosophischen Semantik hinein gehört, auch wenn sie durch Putnams Argumente ebenso in Zweifel gezogen wurde, wie die Kennzeichnungstheorie. Außerdem wurde die Philosophie der Namen, ein ganz eigenes Gebiet innerhalb der philosophischen Semantik, nur gestreift und an vielen Stellen nicht genügend präzisiert.

In diesem Zusammenhang möchte ich zum Beispiel auf [Bu73] hinweisen. Die Kernthese dieses Textes ist, daß Namen wie verkappte Prädikate funktionieren, die auf alle Gegenstände zutreffen, die diesen Namen tragen. Natürlich steht diese These der gesamten Behandlungstradition der hier besprochenen Namensphilosophie entgegen – was aber, wenn man die bekannten Thesen von Burge aus Kapitel IV bedenkt, keine Überraschung ist. Im Zusammenhang mit diesem Text sollte [Kri79] gesehen werden, auf den verschiedentlich Bezug genommen wurde. Beide Texte beschäftigen sich aus unterschiedlicher Perspektive mit Adäquanzbedingungen subjektiver Bedeutungsbegriffe und beinhalten wichtige Differenzierungen, die hier nicht zur Sprache kommen konnten.

Wir haben gesehen, daß die Realisierung einer rekursiven Semantik letztlich auf mehrere Probleme stößt:

- I) die Wahrheitswerte und der kognitive Status negativer Existenzsätze
- II) das Problem einer rekursiven Semantik leerer Gegenstandsbezeichnungen
- III) das Informativitätsproblem bei Identitätssätzen
- IV) Twin-Earth-Probleme, was die Beschreibung von Inhalten von Äußerungen angeht
- V) Konsistenzprobleme bei Überzeugungszuschreibungen wie in I.6 und V.2, aufgrund derer Bedeutungsbegriffe subjektivitätskompetent sein müssen.

Diese Arbeit war hauptsächlich an den Problemen III-V orientiert, doch man kann den Akzent der systematischen Besprechung auch anders setzen. Da diese Arbeit am Zusammenhang von kognitiver Beschreibung und rekursiver Semantik aufgehängt ist, nehmen hier die Probleme, die kognitive Dimensionen beinhalten – also eben III-V – natürlicherweise größeren Raum ein. Mit Stalnakers Semantik läßt sich mindestens noch Problem I lösen, wie in [Sta78] ausgeführt wird.

In diesem Zusammenhang möchte ich insbesondere noch auf [Hs95] hinweisen, wo ebenfalls die die Probleme III-V thematisiert und mithilfe einer Weiterentwicklung von Kaplans

Begrifflichkeiten systematisch gelöst werden. Dieser Text ist ein Kondensat von Passagen aus [Hs94], wo das Problem der versteckten Indexikalität bei natürlichen Prädikaten im Zusammenhang mit subjektivitätskompetenten Bedeutungsbegriffen erörtert wird.

Im übrigen möchte ich noch darauf verweisen, daß die Auffassung existiert, daß die hier vorgestellte Verbindung zwischen der Theorie kognitiver Signifikanz und der Theorie der rekursiven Semantik wenigstens für eine Semantik im Sinne einer Theorie über Wahrheitsbedingungen nur scheinbar ist und diese beiden Komplexe miteinander nichts zu schaffen haben. Diese Sicht wird zum Beispiel in [Wet89] vertreten.

Ich denke jedoch, daß im Zusammenhang mit den Konsistenzproblemen, wie sie in I.6 und V.2 entwickelt werden, eine nachvollziehbare Verbindung zwischen diesen Projekten besteht und eine wechselseitige Einwirkung vorliegt. In dieser Arbeit ging es darum, Weise und Entwicklung dieser Beeinflussung grundrißartig deutlich zu machen.

## Literaturverzeichnis

- [Bil87] Bilgrami, Akeel (1987), „An Externalist Account of Psychological Content“, *Philosophical Topics*, Vol. XV, No. 1
- [Bl77] Blau, U. (1977), *Die dreiwertige Logik der Sprache*, Berlin.
- [Blo80] Block, Ned (1980), *Imagery*, MIT Press/Bradford Books.
- [Blo91] Block, Ned (1991), „What Narrow Content is Not“, in: [Loe91], pp. 33-64
- [Bo91] Boghossian, Paul A. (1991), „Naturalizing Content“, in: [Loe91], pp. 65-87
- [Bu73] Burge, Tyler (1973), „Reference and Proper Names“, in: *Journal of Philosophy* 70, pp. 425-439; übersetzt als: Burge, Tyler (1985), „Referenz und Eigennamen“, in: [Wo85], pp. 252-273
- [Bu77] Burge, Tyler (1977), „Belief De Re“, *Journal of Philosophy* 74 (6), pp. 338-362
- [Bu78] Burge, Tyler (1978), „Belief and Synonymy“, in: *Journal of Philosophy*, Vol. 75, pp. 119-138
- [Bu79] Burge, Tyler (1979), „Individualism and the Mental“, in: *Midwest Studies in Philosophy* 4, pp. 73-121
- [Bu82] Burge, Tyler (1982), „Two thought experiments reviewed“, *Notre Dame Journal of Formal Logic* 23, pp. 284-293
- [Bu86] Burge, Tyler (1986), „Individualism and psychology“, *Philosophical Review* 95 (1), pp. 3-46
- [Ca47] Carnap, Rudolf (1947), *Meaning and Necessity*, Chicago, erw. Ausg. 1956; übersetzt als: Carnap, Rudolf (1972), *Bedeutung und Notwendigkeit*, Wien
- [Cho86] Chomsky, N. (1986), *Knowledge of Language. Its Nature, Origin and Use*, New York.
- [Ch54] Church, Alonzo (1954), „Intensional Isomorphism and Identity of Belief“, *Philosophical Studies* 5, pp. 65-73
- [Cri89] Crimmins, Mark (1989), „Having Ideas and Having the Concept“, in: *Mind & Language* 4, pp. 280-294
- [Dav70] Davidson, D. (1970) „Mental Events“ in: [Dav80], pp. 207-228
- [Dav80] Davidson, D. (1980) *Essays on Actions and Events*, Clarendon Press.

- [De82] Dennett, Daniel C. (1982), „Beyond Belief“, in: Woodfield, Andrew (ed.), *Thought and Object: Essays on Intentionality*, Oxford: Clarendon Press, pp. 1-95
- [Dev81] Devitt, M. (1981), *Designation*, New York
- [Don66] Donnellan, Keith (1966), „Reference and Definite Descriptions“ *Philosophical Review* **75**, pp. 281-304 übersetzt als: Donnellan, Keith (1985), „Referenz und Kennzeichnungen“, in: [Wo85], pp. 179-207
- [Don71] Donnellan, Keith (1971), „Proper Names and Identifying Descriptions“, in: Davidson, D., Harman, G. (eds.), (1972), *Semantics of Natural Language*, Dordrecht.
- [Don74] Donnellan, Keith (1974), „Speaking of Nothing“, *Philosophical Review* **83**, pp. 3-31
- [Ev73] Evans, Gareth (1973), „The Causal Theory of Names“, *Proceedings of the Aristotelian Society*, Suppl. Vol. 47, pp. 187-208 übersetzt als: Evans, Gareth: „Die Kausale Theorie der Namen“, in: [Wo85], pp. 309-336
- [Fi78] Field, Hartry (1978), „Mental representation“, in: Block(ed.) *Readings in Philosophy of Psychology*, Vol. 2, HUP, pp. 78-114
- [Fo83] Fodor, Jerry A. (1983), *The Modularity of the Mind: An Essay on Faculty Psychology*, MIT Press/Bradford Books
- [Fo87] Fodor, Jerry A. (1987), *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*, Cambridge, Mass.
- [Fo95] Fodor, Jerry A. (1995), „Concepts: A Potboiler“, in: Villanueva, E.(ed.): *Philosophical Issues* **6**, pp. 1-24
- [Fo98] Fodor, Jerry A. (1998), *Concepts*. Oxford.
- [Fra68] van Fraassen, Bas (1968), „Presuppositions, Implication, and Self-Reference“, *Journal of Philosophy* **63**, pp. 481-495
- [Fr93] Frege, Gottlob (<sup>4</sup>1993), „Der Gedanke“, in: Frege, Gottlob, *Logische Untersuchungen*, Göttingen, pp. 30-53, ursprünglich veröffentlicht als: Frege, Gottlob (1918), „Der Gedanke“, in: *Beiträge zur Philosophie des deutschen Idealismus*, Heft 2, 1.Jg., 1918-19, pp. 58-77
- [Fr94] Frege, Gottlob (<sup>7</sup>1994), „Über Sinn und Bedeutung“, in: Frege, Gottlob, *Funktion, Begriff und Bedeutung*, Göttingen, pp. 40-65 ursprünglich veröffentlicht als: Frege, Gottlob (1892), „Über Sinn und Bedeutung“, in: *Zeitschrift für Philosophie und philosophische Kritik*, NF 100, pp. 25-50
- [Gr57] Grice, H.P. (1957), „Meaning“, *Philosophical Review* **66**, pp. 377-388
- [Hs94] Haas-Spohn, Ulrike (1994), *Versteckte Indexikalität und subjektive Bedeutung*, Tübingen [Dissertation]

- [Hs95] Haas-Spohn, Ulrike (1995), „The Cognitive Significance of Proper Names“, in: Hamm,F,Kolb,J., von Stechow, A.(eds.): *The Blaubeuren Papers*, Vol. 1, Tübingen, Sfs-Report 08-95, pp. 127-147
- [Ho91] Hodges, Wilfrid (1991), *Logic. An Introduction to Elementary Logic* Penguin Books, London, NY
- [Is87] Israel, David (1987), „The role of propositional objects of belief in action“, CSLI Report, Stanford University.
- [Jac89] Jackendoff, R. (1989), „What is a concept that a person may grasp it?“, *Mind and Language*, Vol. 4, 1 and 2, pp. 68-102
- [Ka71] Kamp, Hans (1971), „Formal Properties of 'Now'“ *Theoria* 37, pp. 227-273
- [Ka81] Kamp, Hans (1981), „A Theory of Truth and Semantic Representation“, in: J. Groenendijk, T. Janssen, M. Stokhof (eds.), *Formal Methods in the Study of Language*, Mathematical Centre Tract 135, Amsterdam, pp.277-322
- [Kap68] Kaplan, David (1968), „Quantifying In“, *Synthese*, Vol. 19, pp. 178-214
- [Kap73] Kaplan, David (1973), „Bob and Carol and Ted and Alice“, in: Hintikka, J. (et al., eds.): *Approaches to Natural Language*, Dordrecht.
- [Kap75] Kaplan, David (1975), „How to Russell a Frege-Church“, *The Journal of Philosophy* 72, pp. 716-729
- [Kap78] Kaplan, David (1978), „Dthat“ in: Cole,P: *Syntax and Semantics*, Vol. 9, New York, wieder abgedruckt in: Martinich, A.PP. (ed.) (1985): *The Philosophy of Language*, Oxford.
- [Kap90] Kaplan, David (1990/91), „Words“, *Proceedings of the Aristotelian Society*, pp. 93-119
- [Kap89a] Kaplan, David (1989), „Demonstratives“, ursprüngl. 1977, jetzt in: Almog,J., Perry,J., Wettstein,H. (eds.), (1989), *Themes from Kaplan*, Oxford, pp. 481-563
- [Kap89b] Kaplan, David (1989), „Afterthoughts“, in: Almog,J. Perry,J., Wettstein,H. (eds.), *Themes from Kaplan*, Oxford, pp. 565-614
- [Kri72] Kripke, Saul (1972), „Naming and Necessity“, in: Davidson,D., Harman,G. (eds.), (1972), *Semantics of Natural Language*, Dordrecht, pp. 253-355 übersetzt als: Kripke, Saul (1981), *Name und Notwendigkeit*, Frankfurt a.M.
- [Kri79] Kripke, Saul (1979), „A Puzzle About Belief“, in: Margalit, A.(ed.) (1979), *Meaning and Use*. Dordrecht, pp. 239-283
- [Lew73] Lewis, David (1973), *Counterfactuals*, Oxford.

- [Lew80] Lewis, David (1980), „Index, Context and Content“, in: Kanger,S., Öhmann,S.(eds.), (1980), *Philosophy and Grammar*, Dordrecht pp. 79-100
- [Lew88] Lewis, David (1988), *On the Plurality of Worlds*, Oxford
- [Lo88] Loar, Brian (1988), „Social Content and Psychological Content“, in: Grimm,R., Merrill,D.(eds.), *Contents of Thought*. Tucson, pp. 99-110
- [Lo81] Loar, Brian (1981), *Mind and Meaning*, Cambridge, CUP
- [Lo91] Loar, Brian (1991), „Can We Explain Intentionality?“, in: [Loe91], pp. 119-137
- [Loe91] Loewer, B., Rey,G.(eds.) (1991): *Meaning in Mind. Fodor and His Critics*, Oxford
- [Mi43] Mill, John Stuart (1843), *A System of Logic*, London
- [Pea92] Peacocke, C. (1992), *A Study of Concepts*. Cambridge, Mass.
- [Per77] Perry, John (1977), „Frege on Demonstratives“, in: *The Philosophical Review*, Vol. 86, pp. 474-497
- [Per79] Perry, John (1979), „The problem of the essential indexical“, *Nous* **13**, pp. 3-12
- [Per97] Perry, John (1997), „Reflexivity, Indexicality and Names“, in: Künne,W., Newen,A., Anduschus,M.(eds.): *Direct Reference, Indexicality and Propositional Attitudes*. Stanford, pp. 3-19
- [Put62] Putnam, Hilary (1962), „It ain't necessarily so“, *Journal of Philosophy* **59** (22), pp. 658-671
- [Put75] Putnam, Hilary (1975), „The Meaning of ‚Meaning‘“, in: ders., *Mind, Language and Reality*. Philosophical Papers Vol. 2. Cambridge, pp. 215-271 übersetzt als: Putnam, Hilary (1990), *Die Bedeutung von ‚Bedeutung‘*, Frankfurt a.M. 2. Auflage
- [Put88] Putnam, Hilary (1988), *Representation and Reality*. Cambridge, Mass. übersetzt als: Putnam, Hilary (1991), *Repräsentation und Realität*. Frankfurt a.M.
- [Qui53] Quine, W.V.O. (1953), *From a Logical Point of View and Other Essays*, Harper & Row
- [Qui60] Quine, W.V.O. (1960), *Word and Object*, MIT Press.
- [Qui74] Quine, W.V.O. (1974), *The Roots of Reference*, La Salle, Ill. übersetzt als: Quine, W.V.O. (1976), *Die Wurzeln der Referenz*, Frankfurt.
- [Rus05] Russell, Bertrand (1905), „On Denoting“ in: ders., (1956), *Logic and Knowledge*, London, pp. 39-57 übersetzt als: Russell, Bertrand (1978), „Kennzeichen“, in: Stegmüller,Wolfgang(ed.), *Das Universalienproblem*. Darmstadt, pp. 21-40
- [Sal86] Salmon, N. (1986), *Frege's Puzzle*, Cambridge.



- [Sch90] Schiffer, Stephen (1990), „Fodor’s Character“, in: Villanueva,E.(ed.), *Information, Semantics and Epistemology*. Oxford, pp. 77-101
- [Sch91] Schiffer, Stephen (1991), „Does Mentalese Have a Compositional Semantics?“, in: [Loe91] pp. 181-201
- [Sea58] Searle, John (1958), „Proper Names“, in: *Mind* **67** pp. 966-973
- [Seg73] Segerberg, Krister (1973), „Two-dimensional Modal Logic“, *Journal of Philosophical Logic* **2**, pp. 77-96
- [Spo93] Spohn, Kleemeier, Kurz (1993), *Namen*, Bielefeld (unveröffentlichtes Manuskript)
- [Sta73] Stalnaker, Robert C. (1973), „Presuppositions“, *Journal of Philosophical Logic* **2**, pp. 447-457
- [Sta76] Stalnaker, Robert C. (1976), „Possible Worlds“, *Nous* **10**, pp. 65-75
- [Sta78] Stalnaker, Robert C. (1978), „Assertion“, in: Cole,P (ed.): *Syntax and Semantics*, Vol. 8, New York, pp. 315-332
- [Sta81] Stalnaker, Robert C. (1981), „Indexical Belief“, *Synthese* **49**, pp. 129-151
- [Sta87] Stalnaker, Robert C. (1987), „Semantics for Belief“, in: *Philosophical Topics* **1**, pp. 177-190
- [Sta88] Stalnaker, Robert C. (1988), „Belief Attribution and Context“, in: Grimm,R., Merrill,D.(eds.): *Contents of Thought*. Tucson, pp. 140-156
- [Sta89] Stalnaker, Robert C. (1989), „On What’s in the Head“, in: *Philosophical Perspectives* **3**, pp. 287-316
- [Sta90] Stalnaker, Robert C. (1990), *Narrow Content in propositional Attitudes: The Role of Content in Logic, Language, and Mind*. Stanford
- [Sta91] Stalnaker, Robert C. (1991), „How to Do Semantics for the Language of Thought“ in: [Loe91] pp. 229-238
- [Sta93] Stalnaker, Robert C. (1993), „Twin Earth Revisited“, *Proceedings of the Aristotelian Society*, Vol. 93, pp. 291-311
- [Sti78] Stich, Stephen P (1978), „Autonomous Psychology and the Belief-Desire Thesis“, in: *Monist* Vol. 61, pp. 573-591
- [Sti79] Stich, Stephen P (1979), „Do Animals Have Beliefs?“, *Australasian Journal of Philosophy*, Vol. 57, pp. 15-28
- [Sti83] Stich, Stephen P (1983), *From Folk Psychology to Cognitive Science: The Case Against Belief*. Cambridge, Mass.

- [Sti91] Stich, Stephen P (1991), „Narrow Content Meets Fat Syntax“ in: [Loe91] pp. 239-255
- [Str50] Strawson, Peter F (1950), „On Referring“, in: *Mind* **59**, pp. 320-344 übersetzt als: Strawson, Peter F. (1985), „Über Referenz“, in: [Wo85] pp. 94-126
- [Ta35] Tarski, Alfred (1935), „Der Wahrheitsbegriff in den formalisierten Sprachen“, *Studia Philosophica* **1**
- [Wet89] Wettstein, Howard (1989), „Cognitive Significance Without Cognitive Content“, in: Almog, J., Perry, J., Wettstein, H. (eds.), *Themes from Kaplan*, Oxford, pp. 421-454
- [Whi82] White, Stephen L. (1982), „Partial Character and the Language of Thought“, in: *Pacific Philosophy Quarterly* **63**, pp. 347-365
- [Wit95] Wittgenstein, Ludwig (1995), *Tractatus logico-philosophicus/Philosophische Untersuchungen*, Frankfurt, stw 501 ursprünglich: Wittgenstein, Ludwig (21958), *Philosophical Investigations*, Oxford.
- [Wo85] Wolf, Ursula (ed.) (1985), *Eigennamen*. Frankfurt/M.