Freges konsistente Semantik

Wolfgang Schwarz

19. April 2003

1 Einleitung

Freges Thesen zur Semantik haben, wenn man sie auf sich selbst anwendet, mitunter sonderbare Konsequenzen. So behauptet Frege etwa, man könne auf die Bedeutung des Prädikats "Pferd" nicht mit einem singulären Term Bezug nehmen. Wenn sich diese Behauptung aber auch auf die Sprache bezieht, in der sie selbst formuliert ist, dann kann sich hier der Ausdruck "die Bedeutung des Prädikats 'Pferd'" nicht auf die Bedeutung des Prädikats "Pferd" beziehen. Die Behauptung scheint sich irgendwie selbst den Boden unter den Füßen wegzuziehen.

Mit Schwierigkeiten bei der Anwendung auf die eigene Metasprache ist Freges Semantik aber nicht allein. Eine ganze Reihe von Problemen und Paradoxien stellt sich einem solchen Projekt entgegen, angefangen vom Lügner-Paradox über Russells Paradox bis hin zu Bradleys Regress. Muss am Ende der Versuch einer Semantik für die eigene Sprache als unmöglich aufgegeben werden? Ich glaube nicht. Ich glaube, dass man durchaus konsistente semantische Theorien entwickeln kann, die den Test der Selbstanwendung bestehen. In dieser Arbeit möchte ich zeigen, dass man sogar eine Theorie dieser Art erhält, wenn man Freges Thesen konsequent befolgt.

Im ersten Teil der Arbeit skizziere ich die Grundelemente einer solchen Theorie. Ich stelle zunächst Freges kompositionalen Ansatz vor, diskutiere dann seine strikte Unterscheidung zwischen Funktionen und Gegenständen, und überlege schließlich, wie Freges Semantik dargestellt werden müsste, wenn man mit dieser Unterscheidung ernst macht.

Im zweiten Teil geht es dann um die erwähnten Gefahren und Probleme, denen eine Semantik ausweichen muss, die auf ihre eigene Metasprache anwendbar sein will. Dabei setze ich voraus, dass zumindest Freges Grundidee einer kompositionalen Semantik nicht in Frage steht. Semantische Theorien, die diesen Ansatz nicht teilen, sind von den angeführten Schwierigkeiten in der Regel nicht betroffen. *Wenn* man Freges Grundidee aber akzeptiert, dann erweisen sich seine auf den ersten Blick teils recht merkwürdigen Thesen als hilfreich zur Vermeidung unangenehmer Konsequenzen. Diese Tatsache kann als eingeschränkte Rechtfertigung oder Begründung von Freges Thesen angesehen werden – eingeschränkt, weil sich herausstellen wird, dass man mit deutlich schwächeren Annahmen meist ebenfalls durch kommt.

Eine Bemerkung noch vorab: Obwohl es in dieser Arbeit keineswegs nur um Frege geht, sondern ganz allgemein um mögliche Strategien für eine auf die eigene Sprache anwendbare Semantik, finden sich im Folgenden keine Verweise auf Theorien anderer Autoren. Stattdessen begnüge ich mich in der Regel damit, bestimmte Annahmen, oder bestimmte Kombinationen von Annahmen, als inkonsistent oder gefährlich nachzuweisen. Die Frage, ob diese Annahmen in der einen oder anderen konkreten Theorie (uneingeschränkt) gemacht werden, würde eine sorgfältigere Untersuchung erfordern als hier möglich ist.

2 Eigennamen und Funktionsausdrücke

Eine der Grundideen von Freges Semantik ist, dass die semantischen Eigenschaften eines komplexen Ausdrucks irgendwie systematisch durch die semantischen Eigenschaften seiner Teile bestimmt werden. Dieser kompositionale Ansatz empfiehlt sich nicht nur, weil wir offensichtlich in der Lage sind, Sätze zu verstehen, die uns bislang noch nie begegnet sind, sondern auch, weil in gewöhnlichen Sprachen, künstlichen wie natürlichen, unendlich viele Sätze existieren, ihre Semantik daher nur unter Annahme der Kompositionalität in einer endlichen Theorie beschrieben werden kann.

Kompositionale Semantik geht Hand in Hand mit kompositionaler Syntax. Denn wenn die Semantik komplexer Ausdrücke durch die ihrer Teile bestimmt werden soll, muss erst einmal geklärt werden, was unter den Teilen eines Ausdrucks zu verstehen ist. Die Bedeutung von "Sokrates ist sterblich" ergibt sich zum Beispiel sicher nicht aus den Bedeutungen von "Sokrates i" und "st sterblich". Wir benötigen deshalb Regeln zur Zerlegung komplexer Ausdrücke, bzw. umgekehrt Regeln zum Aufbau komplexer Ausdrücke aus einfacheren. Dazu bietet es sich an, die wohlgeformten Ausdrücke der jeweiligen Sprache in verschiedene Kategorien einzuteilen, so dass man anschließend etwa festlegen kann, dass eine Zeichenkette, die aus einem *einstelligen Prädikat* besteht, dem ein *singulärer Term* folgt, einen Ausdruck vom Typ *Satz* ergibt.

Freges Syntax ist verblüffend einfach. Er teilt alle wohlgeformten Ausdrücke in zwei Kategorien ein: *Eigennamen* und *Funktionsausdrücke*. Wie üblich bei Frege ist die Wahl dieser Bezeichnungen nicht besonders glücklich, denn zu den Eigennamen zählt Frege nicht nur echte Eigennamen, sondern auch definite Kennzeichnungen und Sätze[1]. Eigennamen des Deutschen sind also zum Beispiel die Ausdrücke "Sokrates", "die Hauptstadt von Schweden" und "Cato tötete Cato". Die letzten beiden Beispiele enthalten wiederum Eigennamen als Teile: "Schweden" bzw. "Cato". Funktionsausdrücke sind nun all diejenigen Ausdrücke, die von Eigennamen übrig bleiben, wenn man darin vorkommende Eigennamen entfernt[2]. Entfernt man zum Beispiel aus "die Hauptstadt von Schweden" den Eigennamen "Schweden", so bleibt der Funktionsausdruck "die Hauptstadt von" zurück. Entfernt man aus dem Eigennamen "Cato tötete Cato" das erste Vorkommnis von "Cato", so erhält man den Funktionsausdruck "tötete Cato". Funktionsausdrücke wie "tötete Cato", die bei Ergänzung durch Eigennamen Sätze ergeben, nennt Frege auch *Begriffswörter* (oder *Begriffsnamen*).

[1] vgl. z.B. [FB, 17f.], [SB, 27], [SB, 34], [GgA1, §2]

[2] vgl. z.B. [FB, 17–19], [FB, 27], [WF, 663f.]

Zur Kennzeichnung der Leerstellen, die ergänzt werden müssen, um aus einem Funktionsausdruck einen Eigennamen zu bilden, verwendet Frege die griechischen Buchstaben "ξ" und "ζ". Den Funktionsausdruck "die Hauptstadt von" kann man also auch schreiben als "die Hauptstadt von ξ", "tötete Cato" als "ξ tötete Cato". Diese Schreibweise erlaubt es, anzudeuten, ob an den Leerstellen eines Funktionsausdrucks unterschiedliche Eigennamen eingesetzt werden dürfen oder nicht. Entfernt man zum Beispiel aus "Cato tötete Cato" beide Vorkommnisse von "Cato", so erhält man entweder "ξ tötete ξ" oder "ξ tötete ζ". Durch die zweimalige Verwendung von "ξ" wird mit der ersten Form angedeutet, dass dieser Ausdruck nur durch Einsetzung desselben Eigennamens an beiden Leerstellen zu einem Eigennamen ergänzt werden darf. Aus "ξ tötete ξ", "Brutus" und "Cäsar" kann also nicht der Satz "Brutus tötete Cäsar" aufgebaut werden, sondern nur die Sätze "Brutus tötete Brutus" und "Cäsar tötete Cäsar".

"die Hauptstadt von ξ" und "ξ tötete ξ" sind *einstellige* Funktionsausdrücke, denn sie können nur durch einen einzigen Eigennamen ergänzt werden. "ξ tötete ζ" ist dagegen zweistellig. Wenn wir mit "$F[N]$"[3] den Ausdruck bezeichnen, der aus dem Funktionsausdruck F durch Einsetzung des Eigennamens N an die durch "ξ" gekennzeichneten Leerstellen hervorgeht, und entsprechend mit '$F[N_1, \ldots, N_n]$' die Ergänzung des n-stelligen Ausdrucks F durch $N_1, \ldots, N_n$, so können wir eine einfache Regel für die Zusammensetzung von Ausdrücken aufstellen:

> Wenn F ein n-stelliger Funktionsausdruck und $N_1, \ldots, N_n$ Eigennamen sind, so ist $F[N_1, \ldots, N_n]$ ein Eigenname.

Diese Regel ist für die deutsche Umgangssprache, anders als für Freges Begriffsschrift, nur eingeschränkt anwendbar. Sie erlaubt zum Beispiel Sätze wie "Cato tötete Stockholm ist die Hauptstadt von Schweden" und "Cato oder Schweden", denn diese ergeben sich regelkonform aus den Funktionsausdrücken "Cato tötete ξ" bzw. "ξ oder ζ" und den Eigennamen "Stockholm ist die Hauptstadt von Schweden" bzw. "Cato" und "Schweden". Für eine intuitiv korrekte Grammatik des Deutschen kommt man mit den beiden Kategorien *Eigenname* und *Funktionsausdruck* wahrscheinlich nicht aus. Dennoch ist Freges Syntax eine elegante Näherung, und es ist sicher nicht unvernünftig, zuerst eine idealisierte Syntax und Semantik aufzustellen, und anschließend Zusatzregeln für die unvermeidlichen Eigentümlichkeiten einer natürlichen Sprache einzurichten.

Ein wichtiges Merkmal von Freges Syntax ist die *multiple Zerlegbarkeit* komplexer Ausdrücke. "Cato tötete Cato" zum Beispiel kann nicht nur in den Funktionsausdruck "ξ tötete Cato" und den Eigennamen "Cato" zerlegt werden, sondern auch in "Cato tötete ξ" und "Cato", "ξ tötete ζ" und zweimal "Cato", sowie in "ξ tötete ξ" und "Cato"[4].

Ausgehend von Freges Unterscheidung zwischen Funktionsausdrücken und Eigennamen könnte man im Prinzip noch eine weitere Art der Zerlegung ins Auge fassen, bei

[3] Ich verwende in dieser Arbeit der Einfachheit halber gewöhnliche Anführungszeichen an Stelle von korrekteren, aber komplizierteren Konstruktionen, aus denen ersichtlich wäre, welche Teile einer Zeichenkette durch außerhalb der Kette stehende Quantoren gebunden werden. Dies muss also jeweils aus dem Zusammenhang ermittelt werden.

[4] vgl. z.B. [BS, §9], [GlG2, 372, Fn.4], [NS, 17f.], [NS, 212]

der man statt eines Eigennamens einen Funktionsausdruck entfernt. Entfernt man zum Beispiel aus "Sokrates ist sterblich" den Funktionsausdruck "ξ ist sterblich", so bleibt "Sokrates Φ" übrig. Das Zeichen "Φ" deutet hier an, dass ein Funktionsausdruck entfernt wurde. Umgekehrt gilt daher die Regel: "Sokrates Φ" ergibt einen Eigennamen, wann immer man die Leerstelle durch einen (einstelligen) Funktionsausdruck ergänzt. Funktionsausdrücke, die wie "Sokrates Φ" eine Leerstelle für Funktionsausdrücke enthalten, bezeichnet Frege als *Funktionsausdrücke zweiter Stufe*. "ξ ist sterblich" ist dagegen ein *Funktionsausdruck erster Stufe*. Wenn man der Einfachheit halber Eigennamen als Funktionsausdrücke nullter Stufe definiert, so kann man die obige Regel verallgemeinern:

> Wenn F ein n-stelliger Funktionsausdruck i-ter Stufe und $G_1, \ldots, G_n$ Funktionsausdrücke der Stufe $i-1$ sind, so ist $F[G_1, \ldots, G_n]$ ein Eigenname[5].

Die Zerlegung des einfachen Satzes "Sokrates ist sterblich" in "Sokrates Φ" und "ξ ist sterblich" hielt Frege offenbar nicht für besonders fruchtbar – vielleicht zog er die Möglichkeit auch einfach nie in Erwägung. Funktionsausdrücke zweiter Stufe führt Frege nur ein, wo eine Zerlegung in Eigennamen und Funktionsausdrücke erster Stufe mangels Eigennamen nicht durchführbar ist. In dem Satz "Alles ist sinnlos" taucht zum Beispiel kein Eigenname auf. Enthalten ist hier jedoch der Funktionsausdruck (erster Stufe) "ξ ist sinnlos". Entfernt man diesen, so bleibt der Quantor "Alles Φ" übrig[6]. Dieser kann durch beliebige Funktionsausdrücke erster Stufe zu einem Satz ergänzt werden und gilt deshalb als Funktionsausdruck zweiter Stufe. In der Begriffsschrift gibt es außerdem *Quantoren zweiter Stufe*. Sie können durch Funktionsausdrücke zweiter Stufe ergänzt werden und sind deshalb Funktionsausdrücke dritter Stufe.

3 Gegenstände und Funktionen

Nach diesen Vorbemerkungen zu Freges Syntax nun endlich zur Semantik. Hier sollen, wie gesagt, die semantischen Eigenschaften komplexer Ausdrücke durch die Eigenschaften ihrer Teile bestimmt werden. Ein naheliegender Ansatz besteht deshalb darin, jedem syntaktisch atomaren Ausdruck einen semantischen Wert zuzuordnen, und mit generellen Regeln festzulegen, wie sich die semantischen Werte komplexer Ausdrücke aus den Werten ihrer Teile ergeben.

"Semantischer Wert" verwende ich in einem sehr umfassenden Sinn für jede Art von Entität, die im Rahmen einer Semantik Ausdrücken zugewiesen wird. Das können zum Beispiel Wahrheitswerte sein oder Wahrheitsbedingungen, gewöhnliche Gegenstände, Eigenschaften, Propositionen, inferentielle Rollen, definite Kennzeichnungen, Äquivalenzklassen von Sätzen, Funktionen von Kontexten auf Mengen möglicher Welten, und so weiter.

Frege unterscheidet zwei Arten semantischer Werte: *Sinn* und *Bedeutung*. Wenn wir das Problem nicht-referierender Eigennamen einmal außer Acht lassen, wird in Freges

[5] Die Regel wird noch ein klein wenig komplizierter, wenn man Funktionsausdrücke *gemischter* Stufe berücksichtigt, also z.B. zweistellige Funktionsausdrücke, deren erste Leerstelle durch Eigennamen und deren zweite durch Funktionsausdrücke erster Stufe zu ergänzen ist. Vgl. z.B. [FB, 29].

[6] vgl. z.B. [FB, 26f.], [GlG2, 373]

Semantik jedem Eigennamen und jedem Funktionsausdruck sowohl ein Sinn als auch eine Bedeutung zugewiesen[7]. Semantische Werte (Sinne und Bedeutungen) von Funktionsausdrücken bezeichnet Frege als *Funktionen*, die von Eigennamen als *Gegenstände*[8]. Funktionen, die Bedeutung eines einstelligen Begriffsworts erster Stufe sind, nennt Frege auch *Begriffe*; Gegenstände, die der Sinn eines Satzes sind, *Gedanken*.

Die zentrale Eigenschaft von Funktionen ist, dass sie einen Wert liefern, wenn man sie mit passenden Argumenten ergänzt. Damit können die Regeln zur Bestimmung der semantischen Werte komplexer Ausdrücke sehr einfach formuliert werden. Bezeichnen wir mit "$B(A)$" die Bedeutung des Ausdrucks A, so gilt:

($\mathbf{R}_B$) Wenn F ein n-stelliger Funktionsausdruck i-ter Stufe und $G_1, \ldots, G_n$ Funktionsausdrücke der Stufe $i-1$ sind, so ist $B(F[G_1, \ldots, G_n])$ der Wert von $B(F)$ für die Argumente $B(G_1), \ldots, B(G_n)$.

Völlig analog sieht die Regel zur Bestimmung der *Sinne* komplexer Ausdrücke aus: Bezeichnen wir mit "$S(A)$" den Sinn des Ausdrucks A, so gilt:

($\mathbf{R}_S$) Wenn F ein n-stelliger Funktionsausdruck i-ter Stufe und $G_1, \ldots, G_n$ Funktionsausdrücke der Stufe $i-1$ sind, so ist $S(F[G_1, \ldots, G_n])$ der Wert von $S(F)$ für die Argumente $S(G_1), \ldots, S(G_n)$.

Den Satz "Sokrates ist sterblich" kann man zum Beispiel syntaktisch in den Eigennamen "Sokrates" und den Funktionsausdruck "ξ ist sterblich" zerlegen. Als Bedeutung weist Frege ersterem die Person Sokrates und zweiterem eine Funktion zu, die für alle sterblichen Gegenstände den Wahrheitswert *Wahr* und für alle anderen den Wahrheitswert *Falsch* liefert. Die Bedeutung des ganzen Ausdrucks "Sokrates ist sterblich" ist also der Wert jener Funktion für das Argument Sokrates, das heißt, der Wahrheitswert *Wahr*.

Entsprechend ist die Bedeutung des Allquantors eine Funktion zweiter Stufe, die für jede Funktion erster Stufe F genau dann den Wahrheitswert *Wahr* liefert, wenn F selbst für jeden Gegenstand den Wahrheitswert *Wahr* liefert. Die Bedeutung von "Alles ist sinnlos" ist deshalb genau dann *Wahr*, wenn die Bedeutung von "ξ ist sinnlos" für jeden Gegenstand den Wert *Wahr* liefert[9].

Aus der angegebenen Bestimmung der Ausdrücke "Funktion" und "Gegenstand" folgt unmittelbar:

(**1**) Jeder semantische Wert eines Eigennamens ist ein Gegenstand.

(**2**) Jeder semantische Wert eines Funktionsausdrucks ist eine Funktion.

Darüber hinaus vertritt Frege noch die folgende These:

(**3**) Alles ist entweder ein Gegenstand oder eine Funktion, aber nichts ist beides.

Aus (1), (2) und (3) folgt:

---

[7] vgl. z.B. [NS, 275], [WB, 96], [WB, 127]

[8] vgl. [FB, 18]. Ich komme auf die Frage, ob man dies als Freges offizielle Definition ansehen sollte, gleich zu sprechen.

[9] vgl. z.B. [GlG2, 373f.].

(4) Kein semantischer Wert eines Eigennamens ist eine Funktion.

(5) Kein semantischer Wert eines Funktionsausdrucks ist ein Gegenstand.

Das heißt, dem syntaktischen Unterschied zwischen Funktionsausdrücken und Eigennamen entspricht Frege zufolge ein ontologischer Unterschied zwischen Funktionen und Gegenständen: Semantische Werte von Funktionsausdrücken sind so beschaffen, dass sie prinzipiell nicht als semantische Werte von Eigennamen in Frage kommen. Ebenso schließt die Natur von Gegenständen aus, dass sie semantische Werte von Funktionswörtern sein könnten. (Frege sagt, Funktionen seien wesentlich "ungesättigt", Gegenstände aber "gesättigt".) Ohne die These (3) wäre durchaus zulässig, dass ein Eigenname und ein Funktionsausdruck denselben semantischen Wert haben. Dieser wäre dann sowohl ein Gegenstand als auch eine Funktion.

Erstaunlicherweise gibt Frege in seinen Schriften nirgendwo eine sorgfältige Begründung für die Thesen (3) bis (5). Häufig behandelt er sie, als verstünden sie sich einfach von selbst: "Diesem Unterschiede in den Zeichen [zwischen Funktionsausdrücken und Eigennamen]", lesen wir zum Beispiel in [GlG2, 371], "entspricht natürlich ein solcher im Reiche der Bedeutungen". Wie Funktionsausdrücke, fährt er fort, müssten auch Funktionen eine "Leerstelle" haben, da sie andernfalls nicht an Gegenständen "haften" würden[10]. Das sind, wie Frege zugibt, nur metaphorische Andeutungen. Eine genauere Erklärung sei aber leider nicht möglich, denn der Unterschied zwischen Funktionen und Gegenständen sei eine "logische Urerscheinung", die "nicht auf Einfacheres zurückgeführt werden kann" [GlG2, 371][11].

Erstaunlich ist auch, dass Frege seine semantischen Thesen nicht auf bestimmte Sprachen relativiert. (5) lautet zum Beispiel nicht:

(5*) Kein semantischer Wert eines Funktionsausdrucks *der Begriffsschrift* ist der semantische Wert eines Eigennamens *der Begriffsschrift*.

Das würde erlauben, dass in einer *anderen* Sprache, zum Beispiel im Deutschen, eine Entitität, die der semantische Wert eines besgriffsschriftlichen Funktionsausdruck ist, durch einen Eigennamen herausgegriffen wird.

[10] Ähnliche Bemerkungen finden sich z.B. in [BG, 193], [BG, 205], [WF, 665], [WB, 224], [NS, 192].

[11] vgl. auch hier [BG, 193], [WF, 665]. Ich verwende in dieser Arbeit "Funktion" und "Gegenstand" im Sinne von "etwas, das semantischer Wert eines Funktionsausdrucks sein kann" bzw. "etwas, das semantischer Wert eines Eigennamens sein kann". Frege wäre mit diesen Definitionen wahrscheinlich nicht glücklich, denn er betont ja, dass man "Funktion" und "Gegenstand" gar nicht definieren könne. Diese Auffassung ist allerdings dem Verständnis seiner Position nicht gerade förderlich, denn wie soll man Thesen über Funktionen und Gegenstände beurteilen, wenn gar nicht klar ist, wie diese Ausdrücke eigentlich verwendet werden? Freges eigene Verwendung veranschaulicht das Problem. Manchmal, etwa in [SB, 27], scheint er nicht "Gegenstand" über "Eigenname" bestimmen zu wollen, sondern umgekehrt "Eigenname" zu definieren als "etwas, was einen Gegenstand bezeichnet". Dann wieder, etwa in [BG, 193–195], scheint er die Annahme, eine Funktion könne durch einen Eigennamen bezeichnet werden, nicht nur als falsch, sondern geradezu als widersprüchlich anzusehen. Hier müsste man "Funktion" also definieren als "etwas, was *nur* semantischer Wert eines Funktionsausdrucks sein kann" (entsprechend bei Gegenständen mit Eigennamen). Damit wären die Thesen (4) und (5) analytisch, (1) bis (3) aber substanziell. Für das Folgende erweist sich die von mir gewählte Terminologie, nach der (1) und (2) analytisch sind, zwar nicht als entscheidend, aber als bequem.

Dieser universelle Anspruch von Freges Thesen hat weitreichende Konsequenzen. Er versetzt uns, wie Frege sich ausdrückt, in eine sprachliche "Zwangslage": Wenn F ein Funktionsausdruck ist, dann kann die Bedeutung von F nicht durch einen Eigennamen bezeichnet werden – also auch nicht durch den Eigennamen "die Bedeutung von F".

Der Eigenname "die Bedeutung von F" kann gemäß (1) und (4) nur einen Gegenstand bedeuten. In "Über Begriff und Gegenstand" [BG, 197ff.] erklärt Frege, es handle sich bei diesem Gegenstand um die Extension von F, also die Klasse der Gegenstände, auf die F zutrifft. Diese Klasse ist zwar nicht die Bedeutung von F, denn sie ist ein Gegenstand und keine Funktion, aber sie kann laut Frege diese vertreten, wenn wir das Bedürfnis haben, etwas über die Bedeutung von F auszusagen. Von dieser Art der Vertretung (von Funktionen durch Extensionen, bzw. allgemeiner durch Wertverläufe) macht Frege besonders in den *Grundgesetzen der Arithmetik* ausgiebig Gebrauch. Während er zum Beispiel in den *Grundlagen der Arithmetik* [GlA, §§68–83] die Kardinalzahlen als Klassen von Begriffen definiert, identifiziert er sie in den *Grundgesetzen* [GgA1, §§40f.] mit Klassen von Extensionen. Der Übergang zwischen Aussagen über Funktionen und Aussagen über Extensionen wird dabei jeweils durch Axiom 5 der *Grundgesetze* gerechtfertigt.

Später, als Frege die Inkonsistenz dieses Axioms klar wurde[12], bevorzugt er eine andere Antwort auf die Frage nach der Bedeutung von "die Bedeutung von F": Dieser Ausdruck sei schlicht fehlerhaft und sollte gar nicht verwendet werden: "Eigentlich ist ja der Ausdruck 'die Bedeutung des Begriffsworts [F] zu verwerfen" [NS, 133]. Dasselbe gilt für die Ausdrücke "ξ ist eine Funktion" und "ξ ist ein Begriff". Denn deren Leerstelle kann aus rein syntaktischen Gründen keinen Funktionsausdruck aufnehmen, also kann man mit diesen Ausdrücken keine Funktion als Funktion klassifizieren. "Der Begriff der Funktion muss ja ein Begriff zweiter Stufe sein, während er in der Sprache immer als ein Begriff erster Stufe erscheint" [WB, 218][13].

Frege scheint dies zumindest zum Teil als einen Mangel der deutschen Sprache anzusehen. In der Begriffschrift, schreibt er etwa in [WB, 218], könne man statt "ξ ist eine Funktion" einfach ein Prädikat zweiter Stufe einführen (vgl. auch [NS, 133]). Dummett [5, 216] schlägt den Ausdruck "$\forall x(\Phi[x] \vee \neg\Phi[x])$" vor: Aus diesem ergibt sich bei jeder Ergänzung durch einen Funktionsausdruck ein wahrer Satz[14]. Im Deutschen lässt sich das ungefähr mit "Alles ist entweder Φ oder nicht Φ" übersetzen.

Wenn man Freges Thesen ernst nimmt, muss man die Darstellung seiner Semantik also ziemlich umkrempeln: Von "Bedeutung", "Sinn" und "semantischer Wert" darf nur noch in Bezug auf Eigennamen gesprochen werden. Funktionsausdrücke haben weder Sinn noch Bedeutung. Trotzdem haben sie semantische Eigenschaften, die aber nicht durch Zuweisung semantischer Werte bestimmt werden können[15]. Stattdessen könnte man beispielsweise für "ξ ist sterblich" festlegen:

"ξ ist sterblich" ist genau dann *Wahr* für einen Gegenstand x, wenn x sterb-

[12] mehr dazu unten, im Abschnitt "Russells Paradox"

[13] Vgl. auch [NS, 130f.], [NS, 192], [NS, 210], [NS, 275], [WB, 219], [WB, 224].

[14] Das stimmt nicht ganz. Streng genommen vertritt Dummetts Ausdruck nur "ξ ist eine *einstellige* Funktion *erster Stufe*".

[15] vgl. [WB, 219]: "Man kann von einem Begriffsnamen eigentlich nicht sagen, dass er etwas bedeute; aber man kann sagen, dass er nicht bedeutungslos sei".

lich ist.

Die kompositionale Regel für Sätze würde entsprechend besagen, dass die Bedeutung eines Satzes $F[N]$ genau dann *Wahr* ist, wenn F *Wahr* ist für die Bedeutung von N.

Das lässt sich ohne größere Schwierigkeiten auf andere Funktionsausdrücke übertragen, auch wenn der Ansatz hier ein wenig ungewohnt wirkt:

> Für alle x, y: "die Hauptstadt von ξ" ist genau dann y für x, wenn y die Hauptstadt von x ist.

Die generelle Kompositionsregel lautet also, dass ein Ausdruck $F[N]$ genau dann x bedeutet, wenn F x ist für die Bedeutung von N.

Schwieriger wird es bei höherstufigen Funktionsausdrücken wie Quantoren, denn wenn die Argumentstelle einen Funktionsausdruck aufnimmt, darf von dessen Bedeutung nicht gesprochen werden. Dieser Ansatz ist, soweit ich sehen kann, deshalb nicht mit einer *einheitlichen* semantischen Regel für komplexe Ausdrücke verträglich. Zur Not muss bei höherstufigen Funktionsausdrücken über Ausdrücke quantifiziert werden:

> Für alle Begriffswörter F ist die Bedeutung von "$\forall x F[x]$" genau dann *Wahr*, wenn F für alle Gegenstände *Wahr* ist.

Damit entfällt eine Zusatzregel für die Bedeutung quantifizierter Sätze.

Diese Strategie könnte man auch auf Funktionsausdrücke erster Stufe übertragen:

> Für alle Eigennamen N ist die Bedeutung von "N ist sterblich" genau dann *Wahr*, wenn die Bedeutung von N sterblich ist.

> Für alle Eigennamen N ist die Bedeutung von "die Hauptstadt von N" die Hauptstadt der Bedeutung von N.

Weil hier der Ausdruck "ξ ist *Wahr* für ζ" nicht mehr vorkommt, müsste entsprechend auch die Regel für Quantoren geändert werden, etwa so:

> Für alle Begriffswörter F ist die Bedeutung von "$\forall x F[x]$" genau dann *Wahr*, wenn für alle Eigennamen N die Bedeutung von $F[N]$ wahr ist[16].

Beide Ansätze erscheinen zumindest auf den ersten Blick durchführbar.

Was wird nun aus Freges Thesen (1) bis (5)? Auch sie müssen revidiert werden, wenn man eben diese Thesen ernst nimmt. Denn auch hier ist von "Funktionen" und "semantischen Werten" von Funktionsausdrücken die Rede. Im Satz

> (**3**) Alles ist entweder Gegenstand oder Funktion, aber nichts ist beides

[16] Diese substitutionelle Interpretation setzt voraus, dass alle Gegenstände durch Eigennamen bezeichnet werden. Das lässt sich kompensieren, indem man zum Beispiel über alle *möglichen* Eigennamen quantifiziert (ähnlich wie in der Semantik der prädikatenlogischen Quantoren mit Rekursion über Wahrheit – den *möglichen* Eigennamen entsprechen dort 'neue' Individuenkonstanten gepaart mit Interpretations-Varianten), oder indem man unbenannte Gegenstände generell als ihre eigenen (unaussprechlichen) Eigennamen zulässt.

finden wir zum Beispiel den Quantor "Alles Φ", der durch den Funktionsausdruck erster Stufe "ξ ist entweder ein Gegenstand oder ξ ist eine Funktion" ergänzt wird. An dieser Stelle kann "ξ ist eine Funktion" offenbar nicht nach Dummetts Schema in eine Funktion zweiter Stufe übersetzt werden, weil der Satz sonst grammatisch unkorrekt würde. Es lässt sich aber leicht ein Ausdruck erster Stufe finden, der mit genau denjenigen Eigennamen, die eine Funktion bezeichnen – also gar keinen – einen wahren Satz ergibt, zum Beispiel "$\xi \neq \xi$. Ebenso können wir den Begriff "ξ ist ein Gegenstand" nun definieren als "$\xi = \xi$", denn beide Ausdrücke ergeben genau dann einen wahren Satz, wenn man an die Leerstelle etwas einsetzt, was einen Gegenstand bedeutet.

In einer konsequenten Frege-Semantik müsste man (3) also etwa so paraphrasieren:

> (**3′**) Alles ist entweder mit sich selbst identisch oder nicht mit sich selbst identisch, aber nichts ist sowohl mit sich selbst identisch als auch nicht mit sich selbst identisch[17].

Im Gegensatz zu (3) ist diese These keineswegs merkwürdig, sondern eine ganz uninteressante, logische Wahrheit. Für Frege ist das nicht unbedingt zum Nachteil, denn die Ansicht, dass man Thesen wie (3') nicht weiter begründen kann, sondern als "logische Urerscheinung" akzeptieren muss, klingt durchaus vernünftig. Die Frage nach einer Begründung für Freges Thesen ist damit allerdings nicht völlig vom Tisch. Sie bekommt lediglich eine andere Form: Warum sollten wir Semantik gerade *so* betreiben, und nicht zum Beispiel auch Funktionsausdrücken semantische Werte zuweisen? In den folgenden Abschnitten werde ich einige mögliche Antworten auf diese Frage unter die Lupe nehmen.

Vorher sei noch einmal betont, dass man die hier skizzierte semantische Theorie nur mit Vorbehalten Frege zuschreiben kann. Ganz offensichtlich schwankte Frege selbst zwischen einer solchen konsequenten Anwendung seiner Thesen und einer inkonsequenten Theorie, die für die eigene Metasprache eine Ausnahme macht. In dieser Metasprache gibt es Eigennamen nicht nur für Gegenstände, sondern auch für Funktionen, so dass die Semantik komplexer Ausdrücke *der Objektsprache* stets durch Funktionsausdrücke erster Stufe wie "die Bedeutung von ξ ist der Wert der Funktion ζ für das Argument χ" bestimmt werden kann.

Diese inkonsequente Theorie ähnelt in vieler Hinsicht der mengentheoretischen Standardsemantik für die Prädikatenlogik. Die semantischen Werte der objektsprachlichen *Funktionsausdrücke* werden in dieser Metasprache – typischerweise die Sprache der Zermelo-Fränkel-Mengenlehre ZF – durch *singuläre Terme* bezeichnet. Das Ergebnis ist einigermaßen bizarr, denn einerseits ist die Sprache von ZF selbst nichts anderes als Prädikatenlogik erster Stufe, andererseits kann die Semantik von ZF erwiesenermaßen *nicht* (korrekt) in ZF angegeben werden[18]. Die Standardsemantik der Prädikatenlogik

[17] Mit ganz ähnlichen Überlegungen und etwas Mühe kann man auch die Thesen (1), (2), (4) und (5) in triviale logische Wahrheiten überführen. (1) zum Beispiel wird zu

> (**1′**) Jeder semantische Wert eines Eigennamens ist mit sich selbst identisch.

[18] Sonst müsste man beispielsweise dem ZF-Ausdruck "ξ ist eine Menge" als semantischen Wert die Menge aller Mengen zuweisen. In ZF ist aber andererseits beweisbar, dass es keine Menge aller Mengen gibt.

erster Stufe ist also nicht auf ihre eigene Sprache anwendbar, obwohl diese Sprache selbst Prädikatenlogik erster Stufe ist.

Daran wird deutlich, wie schwierig es ist, eine Semantik für die eigene Sprache zu entwickeln, das heißt eine Semantik, die auch für ihre eigene Metasprache gilt. Im Rest dieser Arbeit werde ich einige Gefahren diskutieren, denen ein solches Projekt ausgesetzt ist, und untersuchen, ob eine konsequente Frege-Semantik – anders als die Standardsemantik der Prädikatenlogik – diesen entgehen kann.

4 Semantische Paradoxien

Die vielleicht bekanntesten Schwierigkeiten für eine Semantik der eigenen Sprache sind Paradoxien, die duch uneingeschränkte Anwendung von Disquotationsschemata entstehen. Solche Schemata sind zum Beispiel:

> (**Tarski**) "p" ist genau dann wahr, wenn p.
>
> (**Grelling**) "F" ist genau dann wahr für x, wenn $F[x]$.

Setzen wir nun den Satz

> (**L**) Der erste Satz in "Freges konsistente Semantik", der mit den Worten "Der erste" beginnt, ist nicht wahr

in das Tarski-Schema ein, so erhalten wir:

> "Der erste Satz in 'Freges konsistente Semantik', der mit den Worten 'Der erste' beginnt, ist nicht wahr" ist genau dann wahr, wenn der erste Satz in "Freges konsistente Semantik", der mit den Worten "Der erste" beginnt, nicht wahr ist.

Nun ist L selbst der erste Satz in "Freges konsistente Semantik", der mit den Worten "Der erste" beginnt. Mit der harmlosen Zusatzannahme, dass wenn ein Satz p identisch ist mit einem Satz q, dann p und q unter genau denselben Bedingungen wahr sind, erhalten wir den Widerspruch:

> "Der erste Satz in 'Freges konsistente Semantik', der mit den Worten 'Der erste' beginnt, ist nicht wahr" ist genau dann wahr, wenn "Der erste Satz in 'Freges konsistente Semantik', der mit den Worten 'Der erste' beginnt, ist nicht wahr" nicht wahr ist.

Entsprechend führt das Grelling-Schema zu einem Widerspruch, wenn man Prädikate wie "ξ ist nicht wahr für ξ" betrachtet. Denn Einsetzung in das Schema ergibt, dass für alle x,

> "ξ ist nicht wahr für ξ" ist genau dann wahr für x, wenn x nicht wahr ist für x.

Also gilt insbesondere für $x =$ "ξ ist nicht wahr für ξ":

> "ξ ist nicht wahr für ξ" ist genau dann wahr für "ξ ist nicht wahr für ξ", wenn "ξ ist nicht wahr für ξ" nicht wahr ist für "ξ ist nicht wahr für ξ".

Zwei mögliche Antworten auf die Paradoxien möchte ich an dieser Stelle ohne Diskussion ausschließen. Die erste ist die Annahme, die paradoxen Einsetzungen existierten gar nicht, weil zum Beispiel der Satz (L) *syntaktisch* nicht wohlgeformt sei. Das scheint mir hoffnungslos: Gegen welche syntaktische Regel verstößt (L)? Für ebenso hoffnungslos halte ich zweitens den Dialethismus, der die paradoxen Instanzen einfach akzeptiert: (L) zum Beispiel ist aus Sicht des Dialethismus sowohl wahr als auch nicht wahr. Diese Position ist inkonsistent und deshalb falsch[19].

Wenn die paradoxen Einsetzungen also erstens existieren und zweitens das Resultat nicht akzeptiert werden kann, dann bleibt als einziger Ausweg die Ablehnung der Schemata. Das heißt, jede Theorie – insbesondere jede semantische Theorie –, der zufolge diese Schemata *uneingeschränkt* gelten, ist nachweislich falsch.

In den Grundregeln der Fregeschen Semantik, so wie ich sie im letzten Abschnitt vorgestellt hatte, kam kein Schema dieser Art vor. Die Paradoxie kann sich also höchstens noch bei der Bestimmung der semantischen Eigenschaften einzelner Ausdrücke einschleichen. Besonders gefährlich sind hier natürlich die Ausdrücke "ξ ist wahr" und "ξ ist wahr für ζ" (bzw. der allgemeinere Ausdruck "ξ ist χ für ζ"). Die Paradoxien zeigen, dass zum Beispiel "ξ ist wahr" nicht definiert werden kann über die Regel

> (**R**$_w$*) Für alle N, die Bedeutung von "N ist wahr" ist genau dann *Wahr*, wenn die Bedeutung von N *Wahr* ist[20].

Diese Definition ist zwar naheliegend, aber nichts in einer Fregeschen Semantik zwingt uns, sie zu akzeptieren.

Die semantischen Paradoxien lehren meiner Ansicht nacht vor allem, dass man wachsam gegenüber Definitionen sein sollte. Oft werden die Paradoxien so dargestellt, als seien die fraglichen Ausdrücke – "wahr", "heterologisch" usw. – zweifelsfrei durch Regeln wie (Tarski), (Grelling) oder (R$_w$*) definiert. Weil diese Regeln inkonsistent sind, ist es dann in der Tat schwierig, einen Widerspruch zu vermeiden. Eben diese Inkonsistenz der Regeln beweist aber in Wirklichkeit nur, dass man mit ihnen nichts definieren kann. Man kann eine solche Definition *versuchen*, aber das würde ebenso scheitern wie der Versuch, "ξ ist qwer" zu definieren über die Regeln

> (**a**) Eine gerade Zahl ist qwer genau dann, wenn sie größer ist als 3.

[19] Das soll keine *Widerlegung* des Dialethismus sein. Konseqente Dialethisten kann man meiner Ansicht nach nicht widerlegen. Graham Priest [7] zum Beispiel gibt zu, dass seine Position inkonsistent ist, lehnt das aber als Widerlegung des Dialethismus ab, weil der Witz des Dialethismus ist ja gerade in der Annahme bestünde, dass widersprüchliche Positionen korrekt sein können.

[20] Setzen wir für N wieder den Satz L ein, und nehmen wir an, dass die Bedeutung von "nicht p" genau dann *Wahr* ist, wenn die Bedeutung von p nicht *Wahr* ist, so erhalten wir wegen $L =$ "L ist nicht wahr" aus (R$_w$*) "die Bedeutung von L ist genau dann nicht *Wahr*, wenn die Bedeutung von L *Wahr* ist.

(**b**) Eine Primzahl ist qwer genau dann, wenn sie nicht größer ist als 3.

Wäre diese Definition zulässig, dann ließe sich leicht beweisen, dass 2 sowohl größer als auch nicht größer ist als 3. Gar so leicht sollte die Verteidigung des Dialethismus aber nicht sein!

5 Bradleys Regress

Sehen wir uns noch einmal Freges ursprüngliche Regel für die Bedeutung komplexer Ausdrücke an:

($\mathbf{R}_B$) Wenn F ein n-stelliger Funktionsausdruck i-ter Stufe und $G_1, \ldots, G_n$ Funktionsausdrücke der Stufe $i-1$ sind, so ist $B(F[G_1, \ldots, G_n])$ der Wert von $B(F)$ für die Argumente $B(G_1), \ldots, B(G_n)$.

Diese Regel widerspricht Freges semantischen Thesen, weil hier mit Eigennamen ("$B(F)$") auf die Bedeutung von Funktionen verwiesen wird. In [BG, 204f.] deutet Frege ein Argument an, warum ein solcher Ansatz seinem eigenen unterlegen ist.

Nehmen wir als einfaches Beispiel wieder den Satz "Sokrates ist sterblich". Gemäß (R_B) gilt:

> Die Bedeutung von "Sokrates ist sterblich" ist der Wert der Bedeutung von "ξ ist sterblich" für die Bedeutung von "Sokrates".

Der Ausdruck, der hier die Bedeutung des Satzes angibt, setzt sich syntaktisch aus den Bestandteilen "die Bedeutung von 'ξ ist sterblich'", "die Bedeutung von 'Sokrates'" und "der Wert von ξ für ζ" zusammen.

Wenn wir wissen wollen, wie die Bedeutung *dieses* Ausdrucks aus den Bedeutungen seiner Teile hervor geht, benötigen wir daher eine weitere Instanz von (R_B):

> Die Bedeutung von "der Wert der Bedeutung von 'ξ ist sterblich' für die Bedeutung von 'Sokrates'" ist der Wert der Bedeutung von "der Wert von ξ für ζ" für die Bedeutungen von "die Bedeutung von 'ξ ist sterblich'" und "die Bedeutung von 'Sokrates'".

Offenbar führt das in einen Regress: Der Ausdruck, der den semantischen Wert bestimmen soll, wird immer komplizierter. Wenn man diesen Regress, nennen wir ihn *Bradleys Regress*, vermeiden will, muss man dafür sorgen, dass der semantische Wert von $F[N]$ mit einem Ausdruck bestimmt werden kann, der nicht komplexer ist als $F[N]$ selbst. Andererseits verlangt der kompositionale Ansatz, dass in diesem Ausdruck "$B(F)$" und "$B(N)$" vorkommen. Freges Forderung, "$B(F)$" dürfe selbst nur als Funktionsausdruck verwendet werden, scheint dieses Problem zu lösen. Denn die Bedeutung von $F[N]$ muss dann durch eine Regel dieser Form bestimmt werden:

$$B(F[N]) = B(F)[B(N)]$$ [21].

[21] Zumindest auf Deutsch lässt sich das nicht ohne Weiteres ausbuchstabieren, denn weil die Bedeutung

Wenn man außerdem annimmt, dass Iterationen von "$B()$" gekürzt werden können, weil etwa die Bedeutung von "die Bedeutung von 'Sokrates'" nichts anderes ist als die Bedeutung von "Sokrates", ist der Regress blockiert.

Besonders überzeugend scheint mir dieses Argument für Freges Ansatz jedoch nicht zu sein. Erstens ist keineswegs klar, dass Iterationen von "$B()$" immer gekürzt werden können. Im Gegenteil, das Disquotationsschema

(**Berry**) Die Bedeutung von "t" ist t

hat wie die im letzten Abschnitt besprochenen Schemata widersprüchliche Instanzen, zum Beispiel für t = "die kleinste Zahl, die nicht die Bedeutung eines deutschen Ausdrucks mit 77 Buchstaben ist".

Zweitens fordert der dargestellte Lösungsansatz lediglich, dass man *in der semantischen Regel für komplexe Ausdrücke* mit Funktionsausdrücken auf die Bedeutung von Funktionsausdrücken verweisen muss. Daraus folgt nicht, dass man auf diese Bedeutungen nicht auch mit Eigennamen verweisen kann. Eine zwingende Begründung für Freges Thesen (3) bis (5) gibt das Argument also nicht her. (Allenfalls könnte man sagen, dass Eigennamen für semantische Werte, die in keiner semantischen Regel vorkommen, nutzlos und überflüssig sind. "Nutzlos und überflüssig" ist aber nicht genau dasselbe wie "unmöglich".)

Drittens müsste zu allererst geklärt werden, ob Bradleys Regress nicht einfach akzeptiert werden kann. Unbefriedigend wäre sicher, wenn man dadurch, dass zur Formulierung einer semantischen Regel zunehmend komplizierte Ausdrücke nötig sind, die weitere Regeln erfordern, am Ende unendlich viele semantische Regeln bräuchte. Dem ist aber nicht so, weil aus einer hinreichend allgemein formulierten Regel wie (R_B) all diese einzelnen Regeln abgeleitet werden können. Unbefriedigend wäre auch, wenn die semantischen Regeln den Anspruch einer philosophischen *Analyse* oder *Reduktion* hätten, also einer Zurückführung auf einfachere, grundlegendere Tatsachen. Ich sehe aber keinen Grund, warum das metaphysische Projekt einer Analyse mit dem semantischen Projekt einer systamtischen Bestimmung semantischer Eigenschaften in jeder Hinsicht übereinstimmen muss.

6 Russells Paradox

Wenn Frege nach 1902 einen Grund für die strikte Trennung zwischen Funktionen und Gegenständen gab, dann verwies er meist auf Russells Paradox. In seiner ursprünglichen Form entsteht diese Paradoxie durch das Begriffswort "ξ ist eine Menge, die nicht Element von sich selbst ist". Frege war bis 1902 davon ausgegangen, dass man jedem Be-

des Satzes $F[N]$ ein Wahrheitswert ist, muss wohl auch "$B(F)$" ein Begriffswort sein und damit $B(F)[B(N)]$ ein Satz. Aus meinen oben vorgeschlagenen Skizzen Fregescher Semantiken erhält man dabei so etwas wie: "Die Bedeutung von 'Sokrates ist sterblich' ist 'ξ ist sterblich' ist *Wahr* für die Bedeutung von 'Sokrates'". (Dem Ausdruck "$B(F)$" entspricht hier "F ist *Wahr* für ξ".) Weil im Deutschen Sätze aber nicht als Eigennamen dienen, muss man das umstellen, etwa zu "Die Bedeutung von 'Sokrates ist sterblich' ist genau dann *Wahr*, wenn 'ξ ist sterblich' *Wahr* ist für die Bedeutung von 'Sokrates'".

griffswort eindeutig die Menge derjenigen Dinge zuordnen kann, auf die es zutrifft. Dem angegebenen Begriffswort wäre also eine Menge zuzuordnen, die genau all jene Mengen enthält, die sich nicht selbst enthalten. Eine solche Menge aber kann es nicht geben, denn sie würde sich genau dann selbst enthalten, wenn sie sich nicht selbst enthält. Frege schließt aus diesem Widerspruch, dass Ausdrücke wie "die Menge der F" oder "die Extension von F" gar nicht erst verwendet werden sollten. So schreibt er an Richard Hönigswald:

> Das Wesentliche dieses in ein Gestrüpp von Widersprüchen führenden Verfahrens ist in folgendem zusammenzufassen. Man sieht die unter F fallenden Gegenstände als ein Ganzes, einen Gegenstand an und bezeichnet es mit dem Namen "Menge der F" [...]. Man verwandelt hiermit ein Begriffswort "F" in einen Gegenstandsnamen (Eigennamen) "Menge der F". Dieses ist unzulässig wegen der wesentlichen Verschiedenheit von Begriff und Gegenstand. [...]
>
> Eine weithin sichtbare Warntafel muss aufgerichtet werden: Niemand lasse sich einfallen, einen Begriff in einen Gegenstand zu verwandeln! [WB, 86f.][22]

Sehen wir uns die Entstehung der Paradoxie etwas genauer an. Die verhängnisvolle Annahme ist nicht, dass jedem Begriffswort F ein Gegenstand $\{x : F[x]\}$ als Extension zugeordnet werden kann. Entscheidend ist vielmehr die zusätzliche Forderung, dass Begriffswörter, die auf verschiedene Dinge zutreffen, stets verschiedene Extensionen haben. Genau dies besagt Axiom 5 von Freges *Grundgesetzen* [GgA1, §20]:

(Axiom 5) $\{x : F[x]\} = \{x : G[x]\} \;\leftrightarrow\; \forall x(F[x] \leftrightarrow G[x])$

Wenn wir nun den Ausdruck "$\xi \in \zeta$" folgendermaßen definieren:

$$x \in y \;\Leftrightarrow_{Df.}\; \exists F(F[x] \;\wedge\; y = \{x : F[x]\})$$

so erhalten wir[23]:

(RP) Zu jedem Begriffswort F gibt es einen Eigennamen "$\{x : F[x]\}$", so dass gilt: $\forall x(F[x] \;\leftrightarrow\; x \in \{x : F[x]\})$.

Also gilt insbesondere für F = "$\xi \notin \xi$":

[22] vgl. auch [NS, 196ff.], [WB, 121].

[23] Sei F ein beliebiges Begriffswort. Wegen $\forall x(F[x] \leftrightarrow F[x])$ gilt aufgrund von Axiom 5:

(i) $\{x : F[x]\} = \{x : F[x]\}$.

Aus der Definition von $\in$ ergibt sich ferner:

(ii) $x \in \{x : F[x]\} \;\leftrightarrow\; \exists G(G[x] \;\wedge\; \{x : F[x]\} = \{x : G[x]\})$.

Außerdem erhalten wir erneut aus Axiom 5:

(iii) $\{x : F[x]\} = \{x : G[x]\} \;\leftrightarrow\; \forall x(F[x] \leftrightarrow G[x])$

und daher

(iv) $\exists G(G[x] \;\wedge\; \{x : F[x]\} = \{x : G[x]\}) \;\leftrightarrow\; F[x]$.

(RP) folgt unmittelbar aus (ii) und (iv). Die Verwendung von Prädikatenlogik zweiter Stufe ist zur Ableitung des Widerspruchs übrigens unverzichtbar, vgl. [4].

$$\forall x(x \notin x \;\leftrightarrow\; x \in \{x : x \notin x\}).$$

Und daraus wiederum folgt für $x = \{x : x \notin x\}$ der Widerspruch

$$\{x : x \notin x\} \notin \{x : x \notin x\} \;\leftrightarrow\; \{x : x \notin x\} \in \{x : x \notin x\}.$$

Für die Ableitung des Widerspruchs aus (RP) werden keine Annahmen über Extensionen und die $\in$-Beziehung mehr gebraucht. Wir können das Ergebnis also verallgemeinern: Jede Theorie, die eine Aussage der folgenden Form enthält, ist inkonsistent:

> **(RPA)** Zu jedem Begriffswort F gibt es einen Eigennamen N_F, so dass gilt:
> $F[x]$ genau dann, wenn $x \ \Psi\ N_F$.

Grellings Paradox erweist sich nun als eine Instanz von (RPA). Dabei ist N_F das Begriffswort F in Anführungszeichen und Ψ der Ausdruck "ζ ist wahr für ξ". Mit dieser Einsetzung ergibt sich aus [RPA] genau das Grelling-Schema

$F[x]$ genau dann, wenn "F" wahr ist für x.

(RPA) ist aber erheblich stärker, denn es umfasst auch viele Thesen, die mit Disquotation überhaupt nichts zu tun haben, wie eben

$F[x]$ genau dann, wenn $x \in \{x : F[x]\}$

oder auch

$F[x]$ genau dann, wenn x die F-Eigenschaft instanziiert.

Russells Paradox beweist, dass keine These dieser Art uneingeschränkt akzeptiert werden darf. Und das zeigt, dass die Annahme, man könne in der Semantik jedem Begriffswort eindeutig einen Gegenstand als semantischen Wert zuordnen, zumindest gefährlich ist. Eine direkte Widerlegung dieser Annahme ergibt sich jedoch soweit ich sehen kann nicht. Denn weil in der Semantik über sprachliche Ausdrücke geredet wird, bekäme (RPA) hier etwa diese Form:

> **(RPA′)** Zu jedem Begriffswort F gibt es einen Eigennamen N_F, so dass gilt:
> "$F[x]$" ist genau dann wahr, wenn $x \ \Psi\ N_F$.

Und das führt nur mit einer Anwendung des Tarski-Schemas zu einem Widerspruch. Ob Russells Paradox einer solchen Theorie zum Verhängnis wird, hängt deshalb zum einen von der genauen Form der These (RPA′) ab, und zum andern von der Anwendbarkeit des Tarski-Schemas auf Aussagen dieser Form. – Auch wenn dieses Schema nicht generell zulässig ist, besteht doch kein Zweifel, dass *fast alle* seiner Einsetzungen korrekt sind. Es wäre zum Beispiel wenig plausibel, eine semantische Theorie dadurch retten zu wollen, dass man zwar den Satz "die durch 'ξ instanziiert sich nicht selbst' ausgedrückte Eigenschaft instanziiert sich selbst" für wahr erklärt, gleichzeitig aber bestreitet, dass die durch "ξ instanziiert sich nicht selbst" ausgedrückte Eigenschaft sich selbst instanziiert.

Alles in allem liefert Russells Paradox also immerhin eine gewisse Unterstützung für Freges Ansatz in der Semantik. Allerdings widerlegt die Paradoxie bestenfalls die Annahme, dass es für *jedes* Begriffswort einen Eigennamen für dessen semantischen Wert geben kann. Daraus folgt noch lange nicht, dass dies für *kein* Begriffswort möglich ist, wie Frege meint.

7 Cantors Theorem

In seiner Korrespondenz mit Frege erwähnt Russell, dass schon allein deshalb nicht alle Funktionen Gegenstände sein können, weil es wegen Cantors Theorem mehr Funktionen als Gegenstände gibt. Cantors Theorem besagt, dass man zwar jedem Element einer beliebigen Menge M eindeutig eine Funktion auf M zuordenen kann, aber nicht umgekehrt jeder Funktion auf M eindeutig ein Element von M. In mengentheoretischer Terminologie heißt das: Die Kardinalität von M ist kleiner als die Kardinalität der Funktionen auf M.

Die Anwendbarkeit von Cantors Theorem in der Semantik ist jedoch fragwürdig. Zunächst einmal steht nicht fest, dass die Gegenstände überhaupt eine Menge bilden, und nur unter dieser Voraussetzung gilt Cantors Theorem. Wenn man aber beispielsweise Mengen selbst als Gegenstände akzeptiert, dann gibt es keine Menge aller Gegenstände, denn es gibt keine Menge, die alle Mengen enthält.

Darüber hinaus deckt sich der mengentheoretische Funktionsbegriff nicht unbedingt mit Freges Verwendung, wo "Funktion" einfach "semantischer Wert eines Funktionsausdrucks" heißt[24]. Das gilt natürlich schon deshalb, weil in der Mengenlehre Funktionen als Mengen (von n-Tupeln) und damit als Gegenstände definiert werden. Diese Definition ist für Cantors Theorem aber unerheblich. Wichtig ist lediglich diese Annahme:

> (**CT**) Für beliebige Dinge existiert eine Funktion, die für genau jene Dinge den Wert *Wahr* liefert[25].

Übertragen auf die Semantik heißt das: Für beliebige Dinge ist es möglich, ein Begriffswort zu definieren, welches auf genau jene Dinge zutrifft. Diese Annahme klingt zwar auf den ersten Blick nicht abwegig, aber soweit ich weiß gibt es keinen zwingenden Grund, sie zu akzeptieren.

Frege selbst ist jedenfalls nicht darauf verpflichtet. Worauf er sich verpflichtet ist lediglich die Annahme, dass *jedem Funktionsausdruck* eine Funktion entspricht[26]. Über

[24] jedenfalls in meiner Rekonstruktion von Freges Verwendung.

[25] Dass es sich um den Wert *Wahr* handelt, spielt keine Rolle. Ein beliebiger anderer Gegenstand würde den Zweck ebenso erfüllen.

Hier der Beweis, dass es keine Bijektion f von M auf die Menge der Funktionen auf F geben kann: Gemäß (CT) gäbe es andernfalls eine Funktion d mit

$$d(x) = \begin{cases} \mathit{Wahr} & \text{wenn der Wert von } f(x) \text{ für } x \textit{ Falsch} \text{ ist} \\ \mathit{Falsch} & \text{andernfalls.} \end{cases}$$

d ist eine Funktion auf M und müsste folglich im Wertebereich von f liegen. D.h. es gibt ein $m \in M$, so dass $f(m) = d$. Was ist nun der Wert von d für m? Gemäß Definition müsste gelten:

$$d(m) = \begin{cases} \mathit{Wahr} & \text{wenn der Wert von } d \text{ für } m \textit{ Falsch} \text{ ist} \\ \mathit{Falsch} & \text{andernfalls} \end{cases}$$

Das ist aber ein Widerspruch.

[26] Im System der *Grundgesetze* ist dieses Komprehensionsprinzip durch die Substitutionsregel [GgA1, §48.9] vertreten. Die stärkere These (CT) gilt nur in den Standardmodellen der Prädikatenlogik zweiter Stufe, nicht aber in Nonstandard-Henkin-Modellen. (Beide Modelltheorien hätte Frege natürlich abgelehnt, weil in ihnen Funktionen als Werte von Eigennamen behandelt werden.)

Funktionen, für die es keinen Ausdruck gibt, ist damit nichts gesagt. Und nicht alle Funktionen, von denen in Cantors Beweis die Rede ist, können durch Funktionsausdrücke bezeichnet werden, da diese selbst Gegenstände sind (vgl. [WB, 218]) und also zahlenmäßig den Cantorschen Funktionen unterlegen.

Cantors Theorem beweist also nur, dass nicht alle Begriffe durch Eigennamen bezeichnet werden können, *wenn* man annimmt, dass es erstens eine Menge aller Gegenstände gibt und zweitens für beliebige Dinge ein Begriff existiert, unter den genau jene Dinge fallen.

8 Russell schlägt zurück: Unnormale Ben Lomond-Gedanken

Bisher spielte die Beschaffenheit der semantischen Werte keine große Rolle. Für die nächsten (und letzten) beiden Probleme ist das anders. Diese zeigen nämlich, dass man semantische Werte von Sätzen nicht allzu feinkörnig individuieren sollte. In Freges Semantik hat, wie bereits erwähnt, jeder Satz zwei semantische Werte: einen Wahrheitswert als Bedeutung und einen *Gedanken* als Sinn. Der einem Satz zugewiesene Gedanke bestimmt den kognitiven Gehalt, den der Satz ausdrückt. Während zum Beispiel "Der Mont Blanc ist in Frankreich" und "Ben Lomond ist in Schottland" dieselbe Bedeutung haben (sie sind beide wahr), drücken sie unterschiedliche Gedanken aus.

Nach der kompositionalen Grundidee ergibt sich der durch einen Satz ausgedrückte Sinn aus den Sinnen seiner syntaktischen Bestandteile. Der durch "Ben Lomond ist in Schottland" ausgedrückte Gedanke wird beispielsweise durch den Sinn von "Ben Lomond" und den von "ξ ist in Schottland" bestimmt. Nennen wir solche Gedanken, in die ein Sinn von Ben Lomond eingeht, *Ben Lomond-Gedanken*. Ein Ben Lomond-Gedanke ist also ein Gedanke *über* Ben Lomond.

In [8] präsentiert Adam Rieger nun das folgende Argument gegen Freges semantische Thesen:

> Each Ben Lomond thought is the thought that Ben Lomond falls under some concept. Let us call this the concept *associated* with that thought. A few Ben Lomond thoughts fall under their associated concepts. [...] For example, the thought expressed by "Ben Lomond is a thought" does so. Most, however, do not: that expressed by "Ben Lomond is an aardvark", for example. Call the latter thoughts the *ordinary* Ben Lomond thoughts. There is thus a first-level concept *ordinary Ben Lomond thought*, and a [...] thought expressed by "Ben Lomond is an ordinary Ben Lomond thought". This thought is false, to be sure [...]; but is it ordinary? By definition, it is an ordinary Ben Lomond thought iff it fails to fall under its associated concept, that is, iff it fails to fall under the concept "ordinary Ben Lomond thought", that is, iff it is *not* an ordinary Ben Lomond thought. [8, 327f.]

Die Paradoxie ist bei näherem Hinsehen eine weitere Instanz der Russell-Paradoxie (RPA), denn sie ergibt sich aus dieser Annahme:

> Zu jedem Begriffswort F gibt es einen Eigennamen "der durch 'Ben Lomond ist F' augedrückte Gedanke", so dass gilt: $F[x]$ genau dann, wenn x unter den Begriff fällt, der mit dem durch "Ben Lomond ist F" augedrückten Gedanken assoziiert ist.

Riegers Paradox demonstriert in der Tat eine Gefahr für semantische Theorien. Aber es kann nur formuliert werden, wenn Eigennamen für Begriffe zur Verfügung stehen – insbesondere Eigennamen der Form "der mit ... assoziierte Begriff". Ob Freges konsequente Semantik von der Paradoxie überhaupt betroffen ist, kann deshalb bezweifelt werden. Dagegen scheinen andere Ansätze, die ganz unverblümt Funktionsausdrücken semantische Werte zuordnen, viel eher betroffen. Nehmen wir zum Beispiel an, in einer Theorie wird jedem Begriffswort eine *Eigenschaft* als Wert zugewiesen, und jedem Satz eine *Proposition*. Dann kann man mit Rieger wie folgt argumentieren:

Für jeden Gegenstand x und jede Eigenschaft F gibt es die Proposition, dass x F instanziiert. Nennen wir F die *assoziierte Eigenschaft* einer solchen Proposition. Manche Propositionen instanziieren ihre assoziierte Eigenschaft, andere nicht. Nennen wir die letzteren *normale Propositionen*. Ist nun die durch "Ben Lomond ist eine normale Proposition" ausgedrückte Proposition normal? Sie ist es genau dann, wenn sie es nicht ist.

Rieger meint, zur Vermeidung des Widerspruchs sollte man Gedanken nicht als Gegenstände einstufen. Dann wäre "ξ ist ein normaler Ben Lomond-Gedanke" ein Funktionsausdruck zweiter Stufe (wir müssten "Φ" statt "ξ" schreiben), und "Ben Lomond ist ein normaler Ben Lomond-Gedanke" kein wohlgeformter Satz. Das ist aber eine ziemlich eigenartige Lösung: Wie hat man sich einen Funktionsausdruck vorzustellen, dessen Bedeutung ein Gedanke sein soll? Durch welche Argumente würde er gesättigt? Mir scheint, Riegers Paradox hat eine ganz andere Voraussetzung, die viel leichter über Bord geworfen werden kann. Diese wird deutlich, wenn wir das Argument auf Fregesche *Bedeutungen* übertragen:

Für jeden Gegenstand x und jeden Begriff F gibt es den *Wahrheitswert*, dass x unter F fällt. Nennen wir F den *assoziierten Begriff* eines solchen Wahrheitswertes. Manche Wahrheitswerte fallen unter ihren assoziierten Begriff, andere nicht. Nennen wir die letzteren *normale Wahrheitswerte*. Ist nun der Wahrheitswert von "Ben Lomond ist ein normaler Wahrheitswert" normal? Er ist es genau dann, wenn er es nicht ist.

Die entscheidende Voraussetzung ist, dass man aus dem semantischen Wert von $F[N]$ den Wert von F eindeutig rekonstruieren kann. Bei Wahrheitswerten ist das sicher nicht so: Aus dem Wahrheitswert von "Ben Lomond ist ein Berg", also dem Wert *Wahr*, kann man den semantischen Wert von "ξ ist ein Berg" nicht mehr rekonstruieren. Rieger verdunkelt diese Möglichkeit, indem er ohne Rechtfertigung von *dem* mit einem Gedanken assoziierten Begriff spricht[27].

[27] Sehen wir uns an, wie ohne diese Voraussetzung die Paradoxie blockiert wird. Wenn man nicht annimmt, dass es zu jedem Ben Lomond-Gedanken genau einen assoziierten Begriff gibt, muss "normaler Ben Lomon-Gedanke" ein wenig anders definieren, etwa so:

> x ist ein *normaler* Ben Lomond-Gedanke genau dann, wenn es einen Begriff F gibt, so dass x der Gedanke ist, dass Ben Lomond unter F fällt, x aber nicht selbst unter F fällt.

Riegers Paradox widerlegt also keineswegs Freges Semantik – nicht einmal die inkonsequente Formulierung seiner Semantik –, sondern vielmehr gewisse Annahmen über die Individuation von Gedanken (allgemein, von semantischen Werten von Sätzen): Zur Vermeidung des Widerspruchs muss es möglich sein, dass ein- und derselbe Gedanke *sowohl* der Gedanke ist, dass Ben Lomond unter N fällt, *als auch* der Gedanke, dass Ben Lomond unter M fällt, obwohl N und M verschiedene, nicht koextensive, Begriffe sind. Besonders gefährdet sind hier offenbar Theorien, die Sätzen hyperintensionale "strukturierte Propositionen" als semantische Werte zuweisen.

Wenn wir die Beziehung zwischen dem semantischen Wert eines Satzes und dem Wert der syntaktischen Teile des Satzes einmal (unpassenderweise) als "Zerlegung" bezeichnen, können wir diese Lehre auch so formulieren: Die These der multiplen Zerlegbarkeit (siehe Abschnitt 2) muss nicht nur für Sätze gelten, sondern auch für deren semantische Werte: Ein einziger Gedanke kann mitunter auf verschiedene Weise in verschiedene semantische Werte von Satzbestandteilen zerlegt werden[28].

Ist nun der durch "Ben Lomond ist ein normaler Ben Lomond-Gedanke" ausgedrückte Gedanke, nennen wir ihn G, normal? Das heißt, gibt es einen Begriff F, so dass G ausdrückt, dass Ben Lomond unter F fällt, obwohl G selbst nicht unter F fällt? Der Widerspruch entsteht, wenn der einzige Begriff F, der diese Bedingung erfüllt, der durch "ξ ist ein normaler Ben Lomond-Gedanke" ausgedrückte Begriff ist. Nennen wir diesen N. Es könnte aber einen *anderen* Begriff $M \neq N$ geben, so dass G auch der Gedanke ist, dass Ben Lomond unter M fällt und G selbst nicht unter M fällt. Aufgrund der Definition von N würde dann G unter N fallen. Da $M \neq N$ ergibt sich daraus jedoch kein Widerspruch.

Gibt es einen solchen Begriff M? Hier ist ein guter Kandidat:

x fällt unter M $\Leftrightarrow$ Es gibt einen Begriff F, so dass x der Gedanke ist, dass x unter F fällt, x aber nicht unter F fällt.

Wir müssen überprüfen, dass a) der Gedanke, Ben Lomond falle unter M, vernünftigerweise mit dem Gedanken, dass Ben Lomond unter N fällt, identifiziert werden kann, obwohl b) N und M nicht auf dieselben Gegenstände zutreffen, weil insbesondere G selbst zwar unter N, nicht aber unter M fällt.

Zu (a): Den Gedanken, dass Ben Lomond unter N fällt, können wir ausdrücken, indem wir "Ben Lomond" an die Stelle von "x" in unsere Definition von "normal" einsetzen:

(*) Es gibt einen Begriff F, so dass Ben Lomond der Gedanke ist, dass Ben Lomond unter F fällt, Ben Lomond aber nicht selbst unter F fällt.

Den Gedanken, dass Ben Lomond unter M fällt, können wir entsprechend ausdrücken, indem wir "Ben Lomond" an Stelle von "x" in die Definition von M einsetzen. Als Resultat erhalten wir genau denselben Satz (*). Wenn aber der Gedanke, dass Ben Lomond unter N fällt durch genau denselben Satz ausgedrückt wird wie der Gedanke, dass Ben Lomond unter M fällt (und dieser Satz keine mehrdeutigen Ausdrücke enthält), liegt die Annahme nahe, dass die beiden Gedanken identisch sind.

Zu (b): Wir überprüfen zunächst, dass G nicht unter M fällt. Nach obiger Definition fällt G genau dann unter M, wenn es einen Begriff F gibt, so dass G der Gedanke ist, dass G unter F fällt, G aber nicht unter F fällt. Diese Bedingung ist wohl kaum erfüllt, denn G ist überhaupt kein Gedanke über G, sondern über Ben Lomond. Bleibt noch zu überprüfen, dass G unter N fällt, also ein normaler Ben Lomond-Gedanke ist. Nach Definition gilt dies genau dann, wenn es einen Begriff F gibt, so dass G der Gedanke ist, dass Ben Lomond unter F fällt, G selbst aber nicht unter F fällt. Tatsächlich ist diese Bedingung für den Begriff M erfüllt, wie wir gerade sahen. Also fällt G unter N.

[28] Die Rede von der Zerlegbarkeit von Gedanken hat sich unter Frege-Schülern eingebürgert, sie ist aber insofern unpassend, als sie den Eindruck vermittelt, Gedanken bestünden *mereologisch* aus den Sinnen von Satzbestandteilen. Das ist eine zusätzliche und unabhängige These, die gesondert behandelt werden sollte. Offensichtlich wird das, wenn man die Forderung nach multipler "Zerlegbarkeit" auf

In den letzten Jahren gab es eine recht lebhafte Diskussion darüber, ob Frege selbst an die multiple Zerlegbarkeit von Gedanken glaubte – oder wenigstens hätte glauben sollen (vgl. z.B. [1], [2], [6], [3]). Die Diskussion wird dadurch belebt, dass Frege in seinen Schriften permanent zwischen zwei schwer vereinbaren Positionen wechselt:

(**A**) Gedanken haben eine Struktur, die in etwa der Struktur der Sätze entspricht, die sie ausdrücken[29].

(**B**) Strukturell verschiedene Sätze können denselben Gedanken ausdrücken[30].

Einige Autoren, darunter David Bell, argumentieren, Frege sollte (A) zugunsten von (B) aufgeben. Michael Dummett will dagegen an (A) festhalten und stattdessen (B) verwerfen. Riegers Paradox entscheidet diese Auseinandersetzung nicht direkt, denn auch (A) ist noch mit multipler Zerlegbarkeit (im oben bestimmten Sinn) verträglich: Wenn der durch "Cato tötete Cato" ausgedrückte Gedanke strukturell isomorph ist zu diesem Satz, so könnte er immer noch wie jener auf verschiedene Weise in Sinne zerlegt werden, zum Beispiel in den Sinn von "Cato" und den von "ξ tötete Cato", oder in den von "Cato" und den von "ξ tötete ξ". Riegers Paradox widerlegt aber eine starke Variante von (A), der zumindest Dummett zugeneigt ist (siehe [6, 302]): dass nämlich die Struktur eines Gedankens eine *eindeutige* Zerlegung in Sinne ermöglicht.

9 Cantor schlägt zurück: Zu viele Gedanken

Rieger hat noch einen zweiten Einwand gegen Freges Ontologie:

> Suppose there are κ objects. Then there are 2^κ concepts. [...] Pick an arbitrary object, say Ben Lomond. For each concept, there is a thought (a Ben Lomond thought) that Ben Lomond falls under that concept. Clearly, distinct concepts give rise to distinct Ben Lomond thoughts. Hence there are at least as many thoughts, and so at least as meny objects, as concepts – contradiction. [8, 327]

Dieser Widerspruch beruht auf den folgenden Annahmen:

(**1**) Für beliebige Dinge gibt es einen Begriff, unter den genau diese Dinge fallen.

Bedeutungen überträgt.

[29] So schreibt Frege z.B. in [NS, 275]: "Der Satz kann als Abbildung des Gedankens betrachtet werden in der Weise, dass dem Verhältnisse bom Teil zum Ganzen bei den Gedanken im Großen und Ganzen dasselbe Verhältnis bei den Sätzen und Satzteilen entspricht." Ähnliche Aussagen finden sich in [BP, 369], [Ver, 155–157], [Ggf, 36f.], [WB, 127], [WB, 231], [NS, 209], [NS, 243], [NS, 262].

[30] So schreibt Frege z.B. in [NS, 213f.] und [WB, 105f.], ein nicht-analytischer Gedanke A sei genau dann identisch mit einem Gedanken B, wenn die Annahme, dass A wahr und B falsch sei, zu einem Widerspruch führt. Entsprechend benutzt er häufig als Test für die Identität von Gedanken, ob es möglich ist, den einen für wahr und den anderen für falsch zu halten, vgl. etwa [FB, 14], [SB, 47], [WB, 128], [WB, 236], [WB, 240], [NS, 152f.]. Weitere Unterstützung für (B) findet sich in [GlA, §§64-67], [BG, 199], [WB, 164], [Ggf, 48–51], [NS, 209], [NS, 203], [NS, 218], [NS, 204].

(**2**) Für jeden Begriff gibt es den Gedanken, dass Ben Lomond unter ihn fällt.

(**3**) All diese Gedanken sind verschieden.

(**4**) Es gibt eine Menge aller Gegenstände.

(**5**) Alle Gedanken sind Gegenstände.

Aus (1)–(4) folgt wegen Cantors Theorem, dass es mehr Gedanken gibt als Gegenstände, im Widerspruch zu (5).

In Abschnitt 7 habe ich argumentiert, dass Frege selbst nicht auf die Annahmen (1) und (4) verpflichtet ist. Wieder trifft dieses Argument also weniger Frege, als mögliche Alternativen, die an (1) und (4) festhalten. Dies kann man tun, wenn man, wie in Abschnitt 7 dargestellt, dafür die Annahme aufgibt, dass alle Begriffe Gegenstände sind, also durch Eigennamen bezeichnet werden können. Riegers zweites Paradox zeigt, dass damit die Gefahr noch nicht gebannt ist.

Wenn meine Überlegungen im vorangegangenen Abschnitt korrekt waren, muss wegen Riegers erstem Paradox die Annahme (3) ohnehin verworfen werden. Ich fürchte aber, allein damit lässt sich die vorliegende Paradoxie noch nicht vermeiden.

An dieser Stelle ist es wichtig, zwischen dem Sinn und der Bedeutung eines Begriffswortes zu unterscheiden. Der Gedanke, dass Ben Lomond unter einen Begriff F fällt, ist streng genommen nicht in F und den Sinn von "Ben Lomond" zerlegbar, sondern in den *Sinn* von "F" – bzw. eine *Art des Gegebenseins* von F[31]– und den Sinn von "Ben Lomond". Die Annahme (2) kann also in zwei Annahmen unterteilt werden:

(**2a**) Für jeden Begriff gibt es mindestens eine Art des Gegebenseins, einen Sinn.

(**2b**) Für jeden dieser Sinne gibt es einen Gedanken, der in jenen Sinn und den Sinn von "Ben Lomond" zerlegt werden kann.

Für Riegers zweites Paradox kann Prämisse (3) damit erheblich abgeschwächt werden: Es ist nicht nötig, dass *all* die Ben Lomon-Gedanken verschieden sind. Es reicht, wenn es für unterschiedliche Begriffe *jeweils wenigstens eine* Art des Gegebenseins gibt, so dass die damit zusammengesetzten Ben Lomond-Gedanken alle verschieden sind.

Ein Beispiel: "Ben Lomond tötete Ben Lomond" drückt sowohl den Gedanken aus, dass Ben Lomond unter die Bedeutung von "ξ tötete ξ" fällt, als auch den Gedanken, dass Ben Lomond unter die Bedeutung von "ξ tötete Ben Lomond" fällt, obwohl diese Begriffe verschieden sind, das heißt, auf verschiedene Gegenstände zutreffen. Die Bedeutung von "ξ tötete Ben Lomond" trifft nun aber auf genau dieselben Gegenstände zu wie die von "ξ ist ein rundes Quadrat", nämlich auf gar keine. Der *Sinn* von "ξ ist ein rundes Quadrat", also in etwa der kognitive Gehalt dieses Ausdrucks, ist aber wohl ein anderer als der von "ξ tötete Ben Lomond". Entsprechend sind auch die mit diesen Sinnen kon-

[31] Wenn x die *Bedeutung* eines Ausdrucks A ist und y der *Sinn* von A, dann bezeichnet Frege y als *Art des Gegebenseins* von x.

struierten Gedanken verschieden. Daraus, dass ein und derselbe Ben Lomond-Gedanke auf mehrfache Weise zerlegt werden kann, folgt also nicht, dass man nicht jedem Begriff eindeutig einen Ben Lomond-Gedanken zuordnen kann.

Wenn man an den Annahmen (1) und (4) festhalten will, muss man also wahrscheinlich (2a) oder (2b) aufgeben. Mir erscheint dabei eine Ablehung von (2a) plausibler, denn die Ablehnung von (2b) würde bedeuten, dass manche Funktionen, die Art des Gegebenseins eines Begriffs sind, für den Sinn von "Ben Lomond" einfach keinen Wert liefern, und das klingt sonderbar. Es verstößt auf jeden Fall gegen Freges Ansicht, semantische Funktionen müssten stets total definiert sein. Eine Ablehung von (2a) dagegen wäre weniger problematisch. Wegen Cantors Theorem steht ohnehin fest, dass nicht alle Begriffe durch Begriffswörter bedeutet werden können. Die Ablehnung von (2a) würde also nicht implizieren, dass einige Begriffswörter zwar eine Bedeutug, aber keinen Sinn haben, sondern allenfalls, dass es für einige der Begriffe, die nicht durch irgend etwas bedeutet werden, auch keine Art des Gegebenseins gibt.

10 Fazit

Es ist möglich, eine konsistente kompositionale Semantik für die eigene Sprache zu entwerfen. Man muss sich dabei jedoch an gewisse Grundregeln halten, um nicht in Widersprüche zu geraten. So darf man die gängigen Disquotationssschemata für semantische Ausdrücke nicht uneingeschränkt akzeptieren; Man darf Sätzen keine allzu feinkörnig individuierten semantischen Werte zuordnen; Und zumindest einigen Prädikaten sollte jeglicher semantische Wert abgesprochen werden.

Setzt man Freges Thesen zur Semantik, insbesondere seine strenge Trennung zwischen den semantischen Werten von Eigennamen und Funktionsausdrücken, konsequent um, so erhält man eine Theorie, die all diesen Anforderungen genügt. Eine konsequente Frege-Semantik ist also, anders als etwa die mengentheoretische Semantik der Prädikatenlogik, auf ihre eigene Metasprache anwendbar.

Andererseits liefert diese Tatsache nur eine eingeschränkte Rechtfertigung von Freges Ansatz. So folgt aus der Tatsache, dass man nicht allen Prädikaten einen semantischen Wert zuordnen sollte, noch lange nicht Freges radikale These, dass *kein* Prädikat einen solchen Wert hat. Anwendbarkeit auf die eigene Metasprache schränkt also lediglich den Kreis möglicher Alternativen zu Freges Ansatz ein. Die Frage nach den Vorzügen und Nachteilen der verbleibenden Theorien muss hier ebenso offen bleiben wie die Frage, ob Selbstanwendung überhaupt ein erstrebenswertes Ziel einer Semantik ist.

Literatur

[BS] Frege, G. *Begriffsschrift*, Halle (Saale), 1879

[GlA] Frege, G. *Die Grundlagen der Arithmetik*, Beslau 1884,

[FB] Frege, G. "Funktion und Begriff", Vortrag gehalten in der Sitzung vom 9.1.1891 der Jenaischen Gesellschaft für Medicin und Naturwissenschaft, Jena: Hermann Pohle 1891

[BG] Frege, G. "Über Begriff und Gegenstand", *Vierteljahreszeitschrift für wissenschaftliche Philosophie* 16 (1892): 192–205

[SB] Frege, G. "Über Sinn und Bedeutung", *Zeitschrift für Philosophie und philosophische Kritik* 100 (1892): 25–50

[GgA1] Frege, G. *Grundgesetze der Arithmetik*, Band I, Jena: Verlag Hermann Pohle, 1893

[BP] Frege, G. "Über die Begriffsschrift des Herrn Peano und meine eigene", *Berichte über die Verhandlungen der königlich-sächsischen Gesellschaft der Wissenschaften zu Leipzig; Mathematisch-Physische Klasse*, 48 (1896): 361-378

[GlG2] Frege, G. "Über die Grundlagen der Geometrie II", *Jahresbericht der deutschen Mathematiker-Vereinigung*, 12 (1903): 368–375

[WF] Frege, G. "Was ist eine Funktion?", *Festschrift für L. Boltzmann*, 1904, 656–666

[Ver] Frege, G. "Die Verneinung: Eine logische Untersuchung", *Beiträge zur Philosophie des deutschen Idealismus* 1 (1919): 143–157

[Ggf] Frege, G. 1923. "Logische Untersuchungen, Dritter Teil: Gedankengefüge", *Beiträge zur Philosophie des deutschen Idealismus* 3: 36–51

[NS] Frege, G. *Nachgelassene Schriften*, Zweite Auflage, Hamburg: Meiner, 1983

[WB] Frege, G. *Wissenschaftlicher Briefwechsel*, Hamburg: Meiner, 1976

[1] Bell, D. "Thoughts", *Notre Dame Journal of Formal Logic* 28 (1987): 36–50

[2] Bell, D. "The Formation of Concepts and the Structure of Thoughts", *Philosophy and Phenomenological Research* 56 (1996): 583–596

[3] Bermudez, J.L. "Frege on Thoughts and Their Structure", *Philosophiegeschichte und Logische Analyse*, 4 (2001): 87-105.

[4] Burgess, J.P. "On a Consistent Subsystem of Frege's *Grundgesetze*", *Notre Dame Journal of Formal Logic* 39 (1998): 274–278

[5] Dummett, M. *Frege: Philosophy of Language*, 2. Auflage, Cambridge (Mass.): Harvard University Press 1981

[6] Dummett, M. "More about Thoughts", *Notre Dame Journal of Formal Logic*, 30 (1989): 1–19, und in Dummett, M. *Frege and Other Philosophers*, Oxford: Clarendon Press, 1991, 289–314

[7] Priest, G. "Contradiction, Belief and Rationality", *Proceedings of the Aristotelean Society* 86 (1986/87), 99–116

[8] Rieger, A. "Paradox without Basic Law V: A Problem with Frege's Ontology", *Analysis* 62 (2002): 327–30