What is Frege's Theory of Descriptions?

Francis Jeffry Pelletier and Bernard Linsky

1 Introduction[1]

When prompted to consider Frege's views about definite descriptions, many philosophers think about the meaning of proper names, and some of them can cite the following quotation taken from a footnote Frege's 1892 article "Über Sinn und Bedeutung."[2]

> In the case of an actual proper name such as 'Aristotle' opinions as to the *Sinn* may differ. It might, for instance, be taken to be the following: the pupil of Plato and teacher of Alexander the Great. Anybody who does this will attach another *Sinn* to the sentence 'Aristotle was born in Stagira' than will a man who takes as the *Sinn* of the name: the teacher of Alexander the Great who was born in Stagira. So long as the *Bedeutung* remains the same, such variations of *Sinn* may be tolerated, although they are to be avoided

[1]We are grateful to Harry Deutsch, Mike Harnish, Theo Janssen, Gregory Landini, Jim Levine, Leonard Linsky, Nathan Salmon, Gurpreet Rattan, Peter Simons, Peter Stanbridge, Kai Wehmeier, an anonymous referee, and audiences at Simon Fraser University, the Society for Exact Philosophy, the Western Canadian Philosophical Association, and the Russell vs. Meinong Conference for helpful comments on various aspects of this paper.

[2]In our quotations, we leave *Bedeutung, Sinn*, and *Bezeichnung* (and cognates) untranslated in order to avoid the confusion that would be brought on by using 'nominatum', 'reference', and 'meaning' for *Bedeutung*, and 'meaning' or 'sense' for *Sinn*. We got this idea from Russell's practice in his reading notes (see Linsky 2004). Otherwise we generally follow Black's translation of "Über Sinn und Bedeutung" (in the 3rd Edition), Furth's translation of *Grundgesetze*, and Austin's tranlsation of *Grundlagen*.

> in the theoretical structure of a demonstrative science and ought not to occur in a perfect language. (p.58)

Many readers, following Kripke (1980), have taken it to be definitive of Frege's views on the meaning of proper names that they can be expressed by a description or are equivalent to a description in the manner indicated by this footnote. And many of these readers have thought that Kripke's arguments against that view have thoroughly discredited Frege. Perhaps so. But our target is different. We wish to discern Frege's views about descriptions themselves, and their logical properties; we do not wish to defend or even discuss whether or not they have any relation to the meaning of proper names. And so we shall not even enter into a discussion of whether Kripke has given us an accurate account of Frege's position on proper names; and if accurate, whether his considerations are telling against the view.

We will also not concern ourselves with the issue of "scope" in what follows. Some scholars point to the seeming ambiguity of (1) and the seeming lack of a similar ambiguity in (2) as further evidence that names and descriptions are radically different.

> (1) George believes the inventor of the bifocals was very clever.

> (2) George believes Benjamin Franklin was very clever.

Although the analysis of these sentences is rather murky, some have held that the presence of a description with its alleged implicit quantifier and consequent capacity to participate in scope ambiguities could yield this difference:

> (1a) George believes that there was a unique inventor of the bifocals and he was very clever

> (1b) There is a unique inventor of the bifocals and George believes him to be very clever[3]

But, it is further claimed, there is no similar ambiguity to be represented as:

[3]Theorists might find further scopal ambiguities here, such as:
(i) There is a unique person such that George believes he is an inventor of the bifocals and is very clever.
(ii) There is a person such that George believes he is the unique inventor of the bifocals and is very clever.

(2a) George believes (there was a) Benjamin Franklin (and he) is very clever

(2b) Benjamin Franklin is such that George believes he is very clever

Not only has this consideration been used against the identification of the "meaning" of proper names with definite descriptions, but it has also been used as a reason to consider them to be of different semantic types: definite descriptions are complex quantified terms (which are syntactically singular, but semantically not, and their semantic representation contains a quantifier), while proper names are simple, unanalyzable singular terms (both syntactically and semantically). This would allow for a scope ambiguity in (1) and explain why there is no scope ambiguity in (2).

Although this is a topic we shall talk only obliquely about in what follows (and that only because Frege, unlike Russell, took definite descriptions to be semantically singular terms), we note that a person could maintain that (2) *does* exhibit an ambiguity, and that it is represented (more or less) as in (2a) and (2b). Of course, such a move would require some sort of strategy to account for the "quantifying in" force of (2b). But there are options available, such as that pursued by Kaplan (1968) under the rubric of "vivid names." And, if this is so, a Fregean could treat definite descriptions as semantically singular, and still be able to account for the ambiguity in (1). We do not suggest that Frege is committed to this specific ploy; only that there are options open to the Fregean in this realm. And so we feel excused from dwelling on the issue of scope in what follows.

2 Background

Frege discussed definite descriptions in two main places, his 1892 article "Über Sinn und Bedeutung" and in Volume I of his *Grundgesetze der Arithmetik* (1893). We think it may still be fruitful to discuss the doctrine(s) of those works since some readers may disagree as to their main points. It is also helpful to put Frege in the context of Russell's view, since in many ways Frege might usefully be seen as a foil to Russell.

Theories of descriptions concern the analysis of sentences containing definite descriptions. For example 'The present queen of England is married', 'The positive square root of four is even', and 'The heat loss

of humans at -20^0C in calm wind is 1800 watts/m^2' all contain definite descriptions. Such sentences arise naturally in ordinary discourse, and just as naturally in semi-formalized theories such as mathematics and science. Theories of descriptions can therefore be seen as trying to account for our ordinary usage and for the usage in semi-formalized situations. In giving such a theory, one will feel tugs from different directions: on one side is the grammatical tug, which encourages the theoretician to mirror the syntactic features of these natural sentences when giving a theoretical account of descriptions. Another side tugs from "rationality", which would have the theoretician give a formal account that matches the intuitive judgments about validity of natural sentences when used in reasoning. And yet a third side tugs from considerations of scientific simplicity, according to which the resulting theoretical account should be complete in its coverage of all the cases but yet not postulate a plethora of disjointed subtheories. It should instead favor one overarching type of theoretical analysis that encompasses all natural occurrences with one sort of entity whenever possible so that the resulting system exhibits some favored sort of simplicity.

The strength of these different tugs has been felt differently by various theorists who wish to give an account of descriptions. If Kaplan (1972) is right, and we think he is, despite the views of certain revisionists, Russell decided that the grammatical tug was not as strong as the others, and he decreed that the apparent grammatical form of definite descriptions (that they are singular terms on a par with proper names) was illusory. In his account it was not merely that the truth value of a sentence with a definite description (e.g., 'The present queen of England is married') has the same truth value as some other one (e.g., 'There is a unique present queen of England and she is married'), but that the correct, underlying character of the former sentence actually contains no singular term that corresponds to the informal definite description. As Russell might say, we must not confuse the true (logical) form of a sentence with its merely apparent (grammatical) form.

As we know, Russell had his reasons for this view, and they resulted from the tugs exerted by the other forces involved in constructing his theory: especially the considerations of logic. Russell seems to have thought that one could not consistently treat definite descriptions that employed certain predicates, such as 'existing golden mountain' or 'nonexisting golden mountain', as singular terms. Carnap (1956: p.33) feels the tug of the third sort most fully. He says, of various choices for improper descriptions, that they

> are not to be understood as different opinions, so that at least one of them is wrong, but rather as different proposals. The different interpretations of descriptions are not meant as assertions about the meaning of phrases of the form 'the so-and-so' in English, but as proposals for an interpretation and, consequently, for deductive rules, concerning descriptions in symbolic systems. Therefore, there is no theoretical issue of right or wrong between the various conceptions, but only the practical question of the comparative convenience of different methods.

In a similar vein, Quine (1940: p.149) calls improper descriptions "uninteresting" and "waste cases", which merely call for some convenient treatment, of whatever sort. We do not share this attitude and will discuss in our concluding section where we think further evidence for a treatment of improper descriptions might come from.

Our plan is to consider the various things that Frege offered about the interpretation of definite descriptions, categorize them into support for different sorts of theories, and finally to describe somewhat more formally what the details of these theories are. After doing this we will consider some Russellian thoughts about definite descriptions in the light of Fregean theories. Our view is that there are four distinct sorts of claims that Frege has made in these central works, but that it is not clear which language Frege intends to apply them to (natural language or an ideal language [*Begriffsschrift*]), nor which of the four theories he suggests should be viewed as his final word. One of these theories, we argue, was given up by Frege, to be replaced by another theory, although it is not so clear as to which theory is the replacement. We will give various alternative accounts of Frege's use for these other three. We will note that the four theories are all opposed to one another in various ways, so that it is difficult to see how Frege might have thought that they all had a legitimate claim in one or another realm. And of course, it is then difficult to see how more than one could have a legitimate claim to being "*Frege's* theory of definite descriptions". We will further argue that the one he explicitly proffers for one of the realms is incomplete, or perhaps inconsistent, and it is not clear how to emend it while satisfying all the desiderata which Frege himself proposes for an adequate theory.

In all the theories suggested by Frege's words, he sought to make definite descriptions be *terms*, that is, be name-like in character. By

this we mean that not only are they *syntactically* singular in nature, like proper names, but also that (as much as possible) they behave *semantically* like paradigm proper names in that they designate some item of reality, i.e., some object in the domain of discourse. Indeed, Frege claims that definite descriptions are proper names: "The *Bezeichnung* <indication> of a single object can also consist of several words or other signs. For brevity, let every such *Bezeichnung* be called a proper name" (1892, p.57). And although this formulation does not explicitly include definite descriptions (as opposed, perhaps, to compound proper names like 'Great Britain' or 'North America'), the examples he feels free to use (e.g., 'the least rapidly convergent series', 'the negative square root of 4') make it clear that he does indeed intend that definite descriptions are to be included among the proper names. In discussing the 'the negative square root of 4', Frege says "We have here the case of a compound proper name constructed from the expression for a concept with the help of the singular definite article" (1892, p.71).

In this desire to maintain the name-like character of definite descriptions, Frege is at odds with Russell's theory, which as we indicated, claims that these sentences contain no singular terms in their "true" logical form. While maintaining this name-like character may be seen as a point in Frege's favor when it comes to a theory of natural language, we should look at the semantic and logical properties of the resulting theories before we decide that Frege is to be preferred to Russell in this regard, a topic to which we shall return.

The next section, Section 3, consists of four parts, each devoted to a possible theory of definite descriptions suggested by some of Frege's words. In these four parts we will marshal the textual evidence relevant to these theories and make some comments about some of their informal properties. In Section 4 we look at some considerations that might be relevant to determining Frege's attitude about the domain of application of the different theories. In Section 5 will consider the more formal properties of the four theories, and discuss whether Frege would be happy with the formal properties of any of the theories. In the sixth section, we will consider some of Russell's arguments. In Section 7 we return to a discussion of reasons to prefer one type of theory over another, and to some different sorts of evidence that might be relevant.

3 Fregean Theories of Definite Descriptions

As we mentioned, we find Frege saying things that might be seen as endorsing four different types of theories. We do not think they all enjoy the same level of legitimacy as being "Frege's Theory of Definite Descriptions". Nonetheless, the theories are of interest in their own rights, and each of them has at some time been seen as "Fregean" (although we will not attempt to name names in this regard). We shall evaluate the extent to which they each can be seen as Fregean.

3a A Frege-Hilbert Theory

In "Über Sinn und Bedeutung" Frege famously remarks:

> A logically perfect language [*Begriffsschrift*] should satisfy the conditions, that every expression grammatically well constructed as a proper name out of signs already introduced shall in fact *bezeichne* an object, and that no new sign shall be introduced as a proper name without being secured a *Bedeutung*. (1892, p.70)

And in discussing the 'the negative square root of 4', he says:

> We have here the case of a compound proper name constructed from the expression for a concept with the help of the singular definite article. This is at any rate permissible if the concept applies to one and only one single object. (1892, p.71)

One could take these statements as requiring that to-be-introduced proper names must first be shown to have a *Bedeutung* before they can be admitted as real proper names. Before 'the negative square root of 4' is admitted to the language as a name, it must be proved that it is proper. (This is the sort of procedure pursued by Hilbert & Bernays 1934, according to Carnap 1956, p.35. And for this reason we call it the 'Frege-Hilbert' theory.) A difficulty with this method is that it makes well-formedness be a consequence of provability or of some factual truth. Before we know whether a sentence employing the (apparent) name 'the negative square root of 4' is well-formed, we need to prove the propriety of that name. And in order to know whether the sentence 'The planet most distant from the Sun is cold' is grammatical (much less true), we would have to know that there is a unique planet most distant from the

Sun. We see with these examples (and others) that the issue of meaninglessness of apparently well-formed names and sentences would arise in mathematics, astronomy and physics, just as much as in ordinary language, according to the Frege-Hilbert theory. Once one allows "contingent" expressions to be used in forming singular terms, one is liable to find sentences that seem to be grammatically impeccable suddenly becoming meaningless, and then not meaningless as the world changes.

In his 1884 *Grundlagen der Arithmetik*, §74 fn. 1, Frege says:

> The definition of an object in terms of a concept under which it falls is a very different matter. For example, the expression "the largest proper fraction" has no content, since the definite article claims to refer to a definite object. On the other hand, the concept "fraction smaller than 1 and such that no fraction smaller than one exceeds it in magnitude" is quite unexceptionable: in order, indeed, to prove that there exists no such fraction, we must make use of just this concept, despite its containing a contradiction. If, however, we wished to use this concept for defining an object falling under it, it would, of course, be necessary first to show two distinct things:
> 1. that some object falls under this concept;
> 2. that only one object falls under it.
>
> Now since the first of these propositions, not to mention the second, is false, it follows that the expression "the largest proper fraction" is *sinnlos* (senseless). (1884, pp. 87–88)

This quotation seems pretty clearly to be in favor of a Frege-Hilbert theory, especially with its use of *sinnlos* to describe definite descriptions that do not have a unique referent. (Although it must be borne in mind that this was from the time before Frege made his *Sinn-Bedeutung* distinction, and so it is not completely clear what sense of *sinnlos* is intended).

This approach is taken to set names in some presentations of set theory. Various axioms have the consequences that there are sets of such and such a sort, something that is usually proved by finding a large enough set and then producing what is wanted by using the axiom of separation. The axiom of extensionality then yields the result that there is exactly one such set; the so-called "existence and uniqueness" results. When it is been shown that there is exactly one set of things that are φ, then one introduces the expression $\{x : \varphi x\}$, which is henceforth treated

as a singular term.[4] The 'Frege-Hilbert' proposal is to treat definite descriptions in the same manner: One proves or otherwise concludes that there is exactly one φ thing, and then '$\iota x F x$' is introduced as a singular term on a par with other proper names.

We will bring forth evidence in the next three subsections that Frege did not adopt this theory of definite descriptions in his later writings, and that it was thus a feature only of his earlier works, such as the 1884 *Grundlagen*.

3b A Frege-Strawson Theory

Frege also considered a theory in which names without *Bedeutung* might nonetheless be used so as to give a *Sinn* to sentences employing them. He remarks,

> It may perhaps be granted that every grammatically well-formed expression figuring as a proper name always has a *Sinn*. But this is not to say that to the *Sinn* there also corresponds a *Bedeutung*. The words 'the celestial body most distant from the Earth' have a *Sinn*, but it is very doubtful if there is also have a *Bedeutung*. In grasping a *Sinn*, one is certainly not assured of a *Bedeutung*. (1892, p.58)

> Is it possible that a sentence as a whole has only a *Sinn*, but no *Bedeutung*? At any rate, one might expect that such sentences occur, just as there are parts of sentences having *Sinn* but no *Bedeutung*. And sentences which contain proper names without *Bedeutung* will be of this kind. The sentence 'Odysseus was set ashore at Ithaca while sound asleep' obviously has a *Sinn*. But since it is doubtful whether the name 'Odysseus', occurring therein, has a *Bedeutung*, it is also doubtful whether the whole sentence does. Yet it is certain, nevertheless, that anyone who seriously took the sentence to be true or false would ascribe to the name 'Odysseus' a *Bedeutung*, not merely a *Sinn*; for it is of the *Bedeutung* of the name that the predicate is affirmed or denied. Whoever does not admit a *Bedeutung* can neither apply nor withhold the predicate. (1892, p.62)

[4] See, for example, Shoenfield (1967) pp. 241-242.

> The thought loses value for us as soon as we recognize that the *Bedeutung* of one of its parts is missing ... But now why do we want every proper name to have not only a *Sinn*, but also a *Bedeutung*? Why is the thought not enough for us? Because, and to the extent that, we are concerned with its truth-value. This is not always the case. In hearing an epic poem, for instance, apart from the euphony of the language we are interested only in the *Sinn* of the sentences and the images and feelings thereby aroused ... Hence it is a matter of no concern to us whether the name 'Odysseus', for instance, has a *Bedeutung*, so long as we accept the poem as a work of art. It is the striving for truth that drives us always to advance from the *Sinn* to *Bedeutung*. (1892, p.63)

It seems pretty clear that Frege here is not really endorsing a theory of language where there might be "empty names", at least not for use in any "scientific situation" where we are inquiring after truth; nonetheless, it could be argued that this is his view of "ordinary language as it is" — there are meaningful singular terms (both atomic singular terms like 'Odysseus' and compound ones like 'the author of *Principia Mathematica*') which do not *bedeuten* an individual. And we can imagine what sort of theory of language is suggested in these remarks: a Frege-Strawson theory[5] in which these empty names are treated as having meaning (having *Sinn*) but designating nothing (having no *Bedeutung*), and sentences containing them are treated as themselves meaningful (have *Sinn*) but having no truth value (no *Bedeutung*) — the sentence is neither true nor false. As Kaplan (1972) remarks, if one already had such a theory for empty proper names, it would be natural to extend it to definite descriptions and make improper definite descriptions also be meaningful (have *Sinn*) and sentences containing them be treated as themselves meaningful (have *Sinn*) but as having no truth value (no *Bedeutung*).

Theories of this sort can be seen as falling into two camps: the "logics of sense and denotation", initiated by Church (1951) and described by Anderson (1984), allow that expressions (presumably including definite descriptions, were there any of them in the language) could lack a denotation but nonetheless have a sense. A somewhat different direction is taken by "free logics", which in general allow singular terms not to

[5]The name 'Frege-Strawson' for this theory is due to Kaplan, 1972, thinking of Strawson (1950, 1952).

denote anything in the domain, thereby making (some) sentences containing these be truth-valueless. (See Lambert & van Fraassen (1967), Lehmann (1994), Moscher & Simons (2001). In these latter theories, there is a restriction on the rules of inference that govern (especially) the quantifiers and the identity sign, so as to make them accord with this semantic intuition. Even though Frege does not put forward the Frege-Strawson theory in his formalized work on the foundations of mathematics, it has its own interesting formal features to which we will return in §5b. And some theorists think of this theory as accurately describing Frege's attitude toward natural language.

3c A Frege-Carnap Theory

In "Über Sinn und Bedeutung", Frege gave the outlines of the Frege-Carnap theory[6] of definite descriptions, which along with the Frege-Strawson theories, are the ones that are the most formally developed of the theories associated with Frege in this realm. In initiating this discussion Frege gives his famous complaint (1892, p.69): "Now, languages have the fault of containing expressions which fail to *bezeichen* an object (although their grammatical form seems to qualify them for that purpose) because the truth of some sentence is a prerequisite", giving the example 'Whoever discovered the elliptic form of the planetary orbits died in misery', where he is treating 'whoever discovered the elliptic form of planetary orbits' as a proper name that depends on the truth of 'there was someone who discovered the elliptic form of the planetary orbits'. He continues:[7]

> This arises from an imperfection of language, from which even the symbolic language of mathematical analysis is not altogether free; even there combinations of symbols can occur that seem to *bedeuten* something but (at least so far) are without *Bedeutung* <*bedeutungslos*>, e.g., divergent infinite series. This can be avoided, e.g., by means of the special stipulation that divergent infinite series shall *bedeuten* the number 0. A logically perfect language (*Begriffsschrift*) should

[6]Once again, the name is due to Kaplan (1972), referring to Carnap (1956), especially pp. 32-38.

[7]Note that the second half of this quote was employed by the Frege-Hilbert theorists as justification for attributing that theory to Frege. But we see here, from placing it in the context of the preceding sentences, that Frege in fact does not hold that theory; instead, he is pointing to the Frege-Carnap theory.

> satisfy the conditions, that every expression grammatically well constructed as a proper name out of signs already introduced shall in fact *bezeichne* an object, and that no new sign shall be introduced as a proper name without being secured a *Bedeutung*. (1892, p.70)

In discussing the 'the negative square root of 4', Frege says (as we quoted above)

> We have here the case of a compound proper name constructed from the expression for a concept with the help of the singular definite article. This is at any rate permissible if one and only one single object falls under the concept. (1892, p.71)

But this does not really support the Frege-Hilbert theory, as can be seen from the continuation of this statement with the footnote:

> In accordance with what was said above, an expression of the kind in question must actually always be assured of a *Bedeutung*, by means of a special stipulation, e.g., by the convention that its *Bedeutung* shall count as 0 when the concept applies to no object or to more than one.[8]

Frege is also at pains to claim that it is not part of the "asserted meaning" of these sorts of proper names that there is a *Bedeutung*; for, if it were, then negating such a sentence would not mean what we ordinarily take it to mean. Consider again the example 'Whoever discovered the elliptic form of the planetary orbits died in misery' and the claim that 'whoever discovered the elliptic form of planetary orbits' in this sentence depends on the truth of 'there was a unique person who discovered the elliptic form of the planetary orbits'. If the sense of 'whoever discovered the elliptic form of planetary orbits' included this

[8] As we said when discussing the Frege-Hilbert theory, Frege's requirements [(i) that every expression grammatically well constructed as a proper name out of signs already introduced shall in fact designate an object, and (ii) that no new sign shall be introduced as a proper name without being secured a *Bedeutung*] do not by themselves make it necessary to supply a special *Bedeutung* for improper names. One might instead withhold the status of "proper name", and that is the option that is pursued by the Frege-Hilbert theory. But this footnote proves that Frege did not wish to withhold the status of being a proper name in such a circumstance, thus denying the Frege-Hilbert interpretation and lending some possible support to the Frege-Carnap theory's being Frege's preferred view.

thought, then the negation of the sentence would be 'Either whoever discovered the elliptic form of the planetary orbits did not die in misery or there was no unique person who discovered the elliptic form of the planetary orbits'. And he takes it as obvious (1892, p.70) that the negation is not formed in this way.[9]

Securing a *Bedeutung* for all proper names is an important requirement, not just in the case of abstract formal languages, but even in ordinary discourse. In one of the very few politically charged statements he makes anywhere in his published writings, he says that failure to adhere to it can lead to immeasurable harm.

> The logic books contain warnings against logical mistakes arising from the ambiguity of expressions. I regard as no less pertinent a warning against proper names without any *Bedeutung*. The history of mathematics supplies errors which have arisen in this way. This lends itself to demagogic abuse as easily as ambiguity does – perhaps more easily. 'The will of the people' can serve as an example; for it is easy to establish that there is at any rate no generally accepted *Bedeutung* for this expression. It is therefore by no means unimportant to eliminate the source of these mistakes, at least in science, once and for all.[10]

[9]Contrary, perhaps, to Russell's opinion as to what is and isn't obvious.

[10]The way in which Frege thinks that 'the will of the people' serves as an example is not quite clear from this or other published works. It may just be a hackneyed example in common use at the time. Or it might relate to some of his views about politics. In either case it presumably revolves around the idea that there are just too many things that are "wills of the people", and hence it is an improper description for this reason. In an unpublished work he makes various assertions that could be relevant to understanding his reasoning. "Vorschläge für ein Wahlgesetz" (Proposal for an Electoral Law), which is not merely a proposed law but also a discussion of its justification, was found in the archives of the politician Clemens von Delbrück (1856-1921), and dates from probably 1918 (see Gabriel & Dathe, 2001, pp. 185-296). In this document Frege remarks that, while it was all right for the Americans and English to follow the heresy of the French *égalité* in allowing women the right to vote, "we Germans think differently." "For us Germans" the basis of society is the family and not the individual, and therefore only families should vote, and "the representation of the family belongs to the husband." His actual proposal contained the following text: "I would support that the right to vote can only be obtained if the citizen (1) is beyond reproach, (2) has fulfilled his military service, (3) is married or was married." These differing conceptions of "the will of the people"—as judged by the French vs. the Germans—might be traced to such underlying thoughts that Frege had. Or, as we said above, this may just be a hackneyed example used by all writers on popular politics. (Thanks to Theo Janssen for bringing this unpublished

These are the places that Frege puts forward the Frege-Carnap theory. It will be noted that there is no formal development of these ideas (or any other ones) in "Über Sinn und Bedeutung". In Section 5 below we consider the sort of formal system that these statements suggest, particularly the K-M theory (which follows the lead of Carnap, 1956).[11]

3d Frege-*Grundgesetze* Theory

In the 1893 *Grundgesetze*, where Frege develops his formal system, he also finds room for definite descriptions — although his discussion is disappointingly short. The relevant part of the *Grundgesetze* is divided into two subparts: a rather informal description that explains how all the various pieces of the language are to be understood, and a more formal statement that includes axioms and rules of inference for these linguistic entities.

Frege retains the central point of the Frege-Carnap theory that he put forward in "Über Sinn und Bedeutung" (about formal languages[12]) by proclaiming (§28) "the following leading principle: *Correctly-formed names must always* bedeuten *something*", and (§33) "every name correctly formed from the defined names must have a *Bedeutung*". These claims of Frege's show pretty conclusively that Frege did not adopt the Frege-Hilbert theory in the *Grundgesetze*, for here he is maintaining that syntactic well-formedness is all that is required for a term to have a *Bedeutung*, rather than requiring a *Bedeutung* in order to be well-formed. In contrast to the Frege-Hilbert theory, we do not need to prove that a description is proper before we can employ the name in a sentence, nor do we need to determine a descriptions propriety by any empirical methods before we can use it.[13]

In the *Grundgesetze*, Frege uses the symbols $\grave{\epsilon}\ F\epsilon$ to indicate the "course of values" (*Werthverlauf*) of a concept F, that is, the set of

work to our attention).

[11] 'K-M' is our name for the overall theory given in Montague & Kalish (1957), Kalish & Montague (1964), and Kalish, Montague, Mar (1980). Our references to these works will henceforth be MK, KM, and KMM, respectively. When we wish to refer to the overarching thoughts or general theory in these works, as we do here, we will call it 'the K-M view/theory/system/etc'.

[12] It is less clear whether Frege also intended this to be a prescription for ordinary language.

[13] Frege does, notoriously, prove that every expression in his system has a *Bedeutung*, in §31, but this is not seen as a preliminary to introducing descriptions, but is proved for other reasons. Whether the argument is in fact an attempted consistency proof, or can even be seen as such, is not relevant to our issue here.

things that are F.[14] In §11 of the *Grundgesetze*, Frege introduces the symbol '\', which is called the "substitute for the definite article".[15] He distinguishes two cases:

> 1. If to the argument there corresponds an object Δ such that the argument is $\dot{\epsilon}\,(\Delta = \epsilon)$, then let the value of the function $\backslash\, x$ be Δ itself.
>
> 2. If to the argument there does not correspond an object Δ such that the argument is $\dot{\epsilon}\,(\Delta = \epsilon)$, then let the value of the function be the argument itself.

And he follows this up with the exposition:

> Accordingly $\backslash\dot{\epsilon}\,(\Delta = \epsilon) = \Delta$ is the True, and "$\backslash\,\dot{\epsilon}\,\Phi(\epsilon)$" *bedeutet* the object falling under the concept $\Phi(\xi)$, if $\Phi(\xi)$ is a concept under which falls one and only one object; in all other cases "$\backslash\dot{\epsilon}\,\Phi(\epsilon)$" *bedeutet* the same as "$\dot{\epsilon}\,\Phi(\epsilon)$".

He then gives as examples (a) "the item when increased by 3 equals 5" designates 2, because 2 is the one and only object that falls under the concept being equal to 5 when increased by 3; (b) the concept *being a square root of 1* has more than one object falling under it, so "the square root of 1" designates $\dot{\epsilon}\,(\epsilon^2 = 1)$[16]; (c) the concept *not identical with itself* has no object falling under it, so it designates $\dot{\epsilon}\,(\epsilon \neq \epsilon)$[17]; and (d) "the x plus 3" designates $\dot{\epsilon}\,(\epsilon + 3)$[18] because *x plus 3* is not a concept at all (it is a function with values other than the True and the False)

In the concluding paragraph of this section, Frege says his proposal has the following advantage:

---

[14] Much as we would use $\{x{:}Fx\}$ or $\hat{x}Fx$ to designate the set of Fs.

[15] Morscher & Simons (2001: 20) take this turn of phrase to show that Frege did not believe that he was giving an analysis of natural language, but of a substitute language. To us, however, the matter does not seem so clear: How else would Frege have put the point if in fact he were trying to give a logical analysis of the natural language definite article?

[16] That is, it *bedeutet* the course of values of "is a square root of 1", i.e., the set $\{-1, 1\}$.

[17] The course of values of "is non-self-identical", i.e., the empty set.

[18] The course of values of the function of adding 3, that is, the set of things to which three has been added.

> There is a logical danger. For, if we wanted to form from the words "square root of 2" the proper name "the square root of 2" we should commit a logical error, because this proper name, in the absence of further stipulation, would be ambiguous, hence even without *Bedeutung* <*bedeutungslos*>. If there were no irrational numbers – as has indeed been maintained – then even the proper name 'the positive square root of 2' would be, at least by the straightforward sense of the words, without a denotation, without special stipulation. And if we were to give this proper name a *Bedeutung* expressly, this would have no connection with the formation of the name, and we should not be entitled to infer that it was a positive square root of 2, while yet we should be only too inclined to conclude just that. This danger about the definite article is here completely circumvented, since "$\grave{\backslash\epsilon}\ \Phi(\epsilon)$" always has a *Bedeutung*, whether the function $\Phi(\xi)$ be not a concept, or a concept under which falls no object or more than one, or a concept under which falls exactly one object. (pp. 50-51)

There seem to be two main points being made here. First, there is a criticism of the Frege-Carnap theory on the grounds that in such a theory the arbitrarily stipulated entity assigned to "ambiguous" definite descriptions "would have no connection to the formation of the name." This would pretty clearly suggest that Frege's opinion in *Grundgesetze* was against the Frege-Carnap view of definite descriptions. And second there is the apparent claim that in his theory, the square root of 2 is a square root of 2, or more generally that the denotation of improper descriptions, at least in those cases where the description is improper due to there being more than one object that satisfies the predicate, manifests the property mentioned in the description.

At this point there is a mismatch between Frege's theory and his explanation of the theory. On this theory, in fact the square root of 2 is *not* a square root of 2 — it is a course of values, that is to say, a set. So it looks like we cannot "infer that the square root of 2 is a square root of 2" even though "we should be only too inclined to conclude just that." [On Frege's behalf, however, we could point out that everything in (= which is a member of) that course of values will be a square root of 2; so there is some connection between the object that the definite description refers to and the property used in the description. But the course of

values itself will not be a square root of 2. Thus, the connection won't be as close as saying that the *Bedeutung* of 'the F' is an F. Indeed, we will see in §5d that this latter suggestion is in fact impossible for the Frege-*Grundgesetze* theory to maintain.]

Is that which we are "only too inclined to conclude" something that we in fact shouldn't? But if so, why is this an objection to the proposal to just stipulate some arbitrary object to be the *Bedeutung*? Perhaps Frege doesn't feel that his *Grundgesetze* theory is a case of giving descriptions a *Bedeutung* "expressly". Yet his own theory seems to be an arbitrary choice from among other alternative possibilities. So we don't know what to make of Frege's reason to reject the Frege-Carnap account in this passage, since his apparent reason is equally a reason to reject the account being recommended.

We call this theory the Frege-*Grundgesetze* theory of definite descriptions. Regardless of one's attitude concerning the applicability of the various theories to natural language, it is clear at least that Frege put forward the Frege-*Grundgesetze* theory in his most fully considered work on the features of a *Begriffsschrift* for mathematics, and that none of the other theories is envisioned at this late date in his writings as being appropriate for this task. But as we will see, both in this subsection and more fully in §5d, not all is well with this theory, even apart from the issue of Basic Law V.

In the *Grundgesetze*, definite descriptions are dealt with by means of

Basic Law (VI): $a = \grave{\epsilon}\,(a = \epsilon)$

(See §18). This Law is used only to derive two further formulas (in §52, stated here using some more modern notation). Frege first cites an instance of Va, one direction of Basic Law V:

$$[(\alpha)(f(\alpha) \equiv (a = \alpha)) \supset \grave{\epsilon}\, f(\epsilon) = \grave{\epsilon}\,(a = \epsilon)]$$

From that he derives a lemma:

$$[(\alpha)(f(\alpha) \equiv (a = \alpha)) \supset (a = \grave{\epsilon}\,(a = \epsilon) \supset a = \grave{\epsilon}\, f(\epsilon))]$$

and then a corollary to the Basic Law, Theorem (VIa):

$$[(\alpha)(f(\alpha) \equiv (a = \alpha)) \supset a = \grave{\epsilon}\, f(\epsilon)]$$

That is, "If a is the unique thing which is f, then a is identical with the f." These are the only theorems about the description operator which

are proved in the introductory section. The description operator is used later in *Grundgesetze*, but this last-mentioned corollary is all that is needed for those uses. It is interesting to note that the notorious Basic Law V is used crucially in this proof.[19] We say that the description operator is used later in *Grundgesetze*, but in fact it is only used in one definition, and then in the proof of only one theorem. The definition is of the notation $a \cap u$, Frege's expression for 'a is an element of u'. Definition A on page 53 is:

$$\backepsilon\dot{\alpha}\ [\neg\forall g(u = \dot{\epsilon}\ g(\epsilon) \supset\ \neg g(a) = \alpha)] \equiv a \cap u$$

or, in other words, 'a is an element of u' has the same truth value as 'there is some g such that u is the course of values of g, and $g(a)$'. (Speaking more closely to the actual formula, it says that the unique element of the course of values of the concept "truth value α such that $[\neg\forall g(u =\dot{\epsilon}\ g(\epsilon) \supset \neg g(a) = \alpha)]$" is identical with the truth value of $a \cap u$). The theorem which makes use of this definition is Theorem 1, on page 75:

$$\text{Theorem 1: } f(a) \equiv a \cap \dot{\epsilon}\ f(\epsilon)$$

This is in effect an abstraction principle: a is f if and only if a is in the course of values of f; and it is used as a lemma for later theorems, but the definition of '$\cap$' is not used again. This abstraction principle of course leads to Russell's paradox directly if one substitutes '$\neg(\xi \cap \xi)$' for '$f(\xi)$', and '$\dot{\epsilon}\ [\neg(\epsilon \cap \epsilon)]$' (the "set of all sets that are not members of themselves") for 'a' and so making $\dot{\epsilon}\ f(\epsilon)$ on the right hand side become $\dot{\epsilon}\ [\neg(\epsilon \cap \epsilon)]$. While the possibility of deriving the inconsistency of *Grundgesetze* can be traced to Basic Law V, it is with Frege's Theorem 1 that it is fully in the open. While not responsible for the contradiction itself, Frege's theory of descriptions does keep bad company.[20]

Frege says nothing else in the *Grundgesetze* about definite descriptions and formulas derived from Basic Law (VI). But it seems clear to us

[19] Basic Law V amounts to an unrestricted comprehension principle, and so is responsible for the inconsistency of *Grundgesetze*. It is safe to conjecture that Basic Law VI with its consequence of Theorem VIa, if added as an axiom, would not lead to an inconsistency alone. What is more, the half of Basic Law V which is used (Va) is not itself responsible for the contradiction. So, while Frege's theory of descriptions does not involve him in the contradiction, it does crucially use the notion of "course of values", which does lead him into trouble.

[20] Morscher & Simons (2001: 21) say "the definite article was implicated by association with the assumptions leading to the paradox Russell discovered in Frege's system."

that there is something missing from this formal development, however: there is no discussion of how improper descriptions are to be logically treated. Basic Law (VI) just does not say anything about this case, being instead relevant only to the case of proper descriptions. This is very puzzling indeed, given the space Frege had committed to detailing just why there needs to be a treatment of improper descriptions. Many commentators seem to have simply passed over this point.[21]

Michael Dummett does notice it, however, in his (1981: p.405) but says only, in the midst of a discussion of Frege's stipulation of interpretations for other expressions, that Frege:

> ...stipulates, for his decription function $\backslash\xi$, both that its value for a unit class as argument shall be the sole member of that unit class, and that its value for any argument not a unit class shall be that argument itself; but, when he formulates the axiom of the system governing the description operator, Axiom VI, it embodies only the first of these two stipulations. For Frege, it is essential to guarantee a determinate interpretation for the system, and, for this purpose, to include, in the informal exposition, enough to determine the referent of every term; but it is unnecessary to embody in the formal axioms more of these stipulations than will actually be required to prove the substantial theorems.

This suggests that "stipulations" about the interpretation of improper descriptions are limited to the informal introduction to a system, and not part of the logical truths that the axioms are intended to capture. In his 1991, Dummett again briefly discusses Axiom VI, saying this time:

> This stipulation is not needed for proving anything in the formal theory that Frege needed to prove; if it had been, it

[21] Beaney (1996, p.248), for example, despite having a discussion of the formal theory of descriptions, does not mention this point. Morscher & Simons (2001: p.21) do in fact mention the fact that it is not derivable, but do not suggest the sort of addition that would be necessary. C.A. Anderson gives an axiom (11^{β}) for definite descriptions which handles the case of improper descriptions in his presentation of the Logic of Sense and Denotation in (Anderson,1984, p.373). Axiom (11^{β}) is:

$(f).(x_\beta)[fx_\beta \rightarrow (\exists y_\beta)[fy_\beta \,.\, y_\beta \neq x_\beta]] \rightarrow .(\iota f) = (\iota x_\beta)F_0$.

In this logic, improper descriptions will not denote anything, but they will have a sense, (ιf), which must be given a value, a "wastebasket" value, a function that takes senses into the preferred sense of The False. Alonzo Church's own formulation of "Alternative 0" had followed Frege in not having an axiom for improper descriptions. But as we said, this does not allow for completeness of the system of descriptions.

> would have been incorporated into the axiom, as it could easily have been. (1991: p.158)

Dummett's view in 1991 now appears to be that the missing axiom could have been included in what is presumably is a system of logical truths, and so is a logical truth. It is left out simply because it is not needed for further theorems.

It is very peculiar that anyone should take this view, since it is so very easy to construct improper descriptions in the language of mathematics. As Frege has explicitly said, one can form such expressions as "the square root of 4" in mathematics, and therefore the underlying logical system needs to be able to deal with these types of terms. And perhaps there are no "atomic truths" to be proved about *the square root of 4*, but nonetheless there are truths that could be proved, such as that it is not green maybe, or that it is self-identical. And of course, we need to be able to use these phrases in reductio proofs, in conditional proofs, and more generally, in all non-assertoric positions of a proposition. It is hard to believe that Frege decided in the end, after pointing out how such expressions are a part of the science that we are formalizing, that we need not find any *Bedeutung* for them because they are not "needed" for further theorems. Regardless of whether they think Frege might have held the view, it is even stranger for modern commentators to cite the view with approval, since as we now all know, one cannot have a complete theory without some account of all the terms in the language.

Tichy (1988, p.121) also explicitly discusses the issue and he decides (as do we) that Frege cannot derive the required VI*:[22]

(VI*) $[\neg(\exists\alpha)(a =\dot{\epsilon}\ (\epsilon = \alpha)] \supset \backslash a = a$ [23]

despite Frege's informal claim that this is the appropriate improper

[22] Tichy tries to disarm the point by saying "The only possible explanation for the lacuna is that in axiomatizing his system Frege did not aim at logical completeness in an absolute sense, but only at a completeness relative to the specific task he set himself in *Grundgesetze*, namely that of deriving the basic truths of arithmetic." Tichy (1988, p.181). Klement (2002: p. 55) also suggests (VI*) and this addendum, explicitly following Tichy.

[23] It is also unlikely that adding VI* as an axiom for improper descriptions would by itself produce an inconsistency (even though it does explicitly introduce courses of values) without more of the force of V than is used here. Morscher & Simons (2001) agree, saying that "the fault [of having a contradiction] lies squarely elsewhere" than with Frege's Basic Law VI and definite descriptions generally.

description rule.[24] This VI* would be a good candidate for a seventh Basic Law.

It appears that the only theory of descriptions which can be definitively attributed to Frege has problematic formal features, as we will outline in §5d, and at the very least, is incomplete in the sense of not allowing for the proof of all semantically valid truths.

4 To What do Frege's Theories of Descriptions Apply?

In this section we will make some inconclusive and partial suggestions about Frege's views concerning the formalization of theories of definite descriptions. Our primary question is to discern what Frege thought he was describing when he gave his various theories of definite descriptions. Or perhaps better put, did Frege think he was telling us how natural language worked? Or how it should work? Or was he engaged instead in telling us how a formal theory suited for mathematics should work? Do the different theories represent changes of mind on his part? Or perhaps they are intended to apply to the different realms, natural language vs. mathematics?

Let us start with the Frege-Hilbert theory. We have provided evidence that seems to show pretty conclusively that this theory was *not* advocated by Frege in either "Über Sinn und Bedeutung" or the *Grundgesetze*. We have shown that the evidence that is sometimes adduced for this view in fact supports different theories: either the Frege-Strawson or the Frege-Carnap theories. The place where the Frege-Hilbert theory is most prominent, we think, is in the *Grundlagen*; and as evidenced by the quotation we cited from it, it seems that Frege there is concerned with a language for mathematics and with the properties that one would need to prove in order to introduce a definite description into his formal language. It does not seem that Frege is making any claims here about how definite descriptions do or should work in natural language.

But this is a view that he gave up when he came to write "Über Sinn und Bedeutung" and the *Grundgesetze*, where the other three views are

[24] It is also unlikely that adding VI* as an axiom for improper descriptions would by itself produce an inconsistency (even though it does explicitly introduce courses of values) without more of the force of V than is used here. (Morscher & Simons (2001) agree, saying that "the fault [of having a contradiction] lies squarely elsewhere" than with Frege's Basic Law VI and definite descriptions generally.)

put forward.

It is not clear to us whether Frege intended the Frege-Carnap and Frege-*Grundgesetze* theories to apply to different realms: the Frege-Carnap view that is put forth in "Über Sinn und Bedeutung" theory applies perhaps to a formalized version of natural language while the Frege-*Grundgesetze* theory to a formal account of mathematics. Frege himself never gives an explicit indication of this sort of distinction between realms of applicability, although it is very easy to see him as engaging simultaneously in two different activities: constructing a suitable framework for the foundations of mathematics, and then a more leisurely reflection on how these same considerations might play out in natural language.

Various attitudes are possible here; for example, one who held that the Frege-Strawson theory represented Frege's attitude to natural language semantics would want to say that both the Frege-Carnap and the Frege-*Grundgesetze* theories were relevant only to the formal representation of arithmetic. This then raises the issue of how such an attitude would explain why Frege gave both the Carnap and the *Grundgesetze* theories for arithmetic. Possibly, this attitude might maintain, the *Grundgesetze* theory was Frege's "real" account for arithmetic, but in "Über Sinn und Bedeutung" he felt it inappropriate to bring up such a complex theory (with its courses of values and the like) in those places where he was concerned to discuss formal languages – as opposed to those places where he was discussing natural language (and where he put forward the Frege-Strawson account). So instead he merely mentioned a "simplified version" of his theory. In this sort of picture, not only is the Frege-Hilbert theory an inappropriate account of Frege's views, but so too is the Frege-Carnap theory, since it is a mere simplified account meant only to give non-formal readers something to fasten on while he was discussing an opposition between natural languages and *Begriffsschriften*. According to this attitude, the real theories are Strawson for natural language and *Grundgesetze* for arithmetic.

Another attitude has Frege being a language reformer, one who wants to replace the bad natural language features of definite descriptions with a more logically tractable one. In this attitude, Frege *never* held the Strawson view of natural language. His talk about Odysseus was just to convince the reader that natural language was in need of reformation. And he then proposed the Frege-Carnap view as preferable in this reformed language. According to one variant of this view, Frege thought that the Carnap view was appropriate for the reformed

natural language while the *Grundgesetze* account was appropriate for mathematics. Another variant would have Frege offer the Carnap view in "Über Sinn und Bedeutung" but replace it with a view he discovered later while writing the *Grundgesetze*. As evidence for this latter variant, we note that Frege did seem to reject the Carnap view when writing the *Grundgesetze*, as we discussed above. However, a consideration against this latter variant is that Frege would most likely have written the relevant portion of the *Grundgesetze* before writing "Über Sinn und Bedeutung". And a consideration against the view as a whole in both of its variants is that Frege never seems to suggest that he is in the business of reforming natural language.[25]

Michael Beaney describes Frege's attitude towards improper descriptions as follows:

> ...descriptions can readily be formed that lack a referent, or that fail to uniquely determine a single referent. Ordinary language is deficient in this respect, according to Frege, whereas in a logical language a referent must be determined for every legitimately constructed proper name. (1996: p.287)

And Morscher & Simon say:

> [Frege] thought sentences containing empty terms would lack reference themselves, and since for him the reference of a sentence was a truth-value this would mean having truth-value gaps in the midst of serious science. So in his own terms Frege's solution is reasonable since he was not attempting anything like a linguistic analysis of actual usage, rather a scientifically better substitute. (2001: 21)

This suggests that the Frege-Strawson view is an account of descriptions as they occur in actual ordinary languages, but that for a "logical language" some referent must be found. Although this neutral statement leaves open the question of whether Frege should be seen as a "reformer" who thought that natural language should be changed so as to obey this requirement that is necessary for a logical language, or whether he was content to leave natural language "as it is", both

[25] Although consider the remarks in fn. 9 above, which can be seen as a recommendation that natural language assign a *Bedeutung* to such natural language phrases as 'The will of the people'.

Beaney and Morscher & Simon, at least in the quoted material here, seem to suggest that Frege is a reformer. (Beaney's "Ordinary language is deficient ..." and Morscher & Simons "better substitute" suggest this). Others[26] quite strongly take the opposing view that Frege was concerned with only a description of natural language, not a reformation, and that this description amounts to the Frege-Strawson account as a background logic.

One might assume that "Über Sinn und Bedeutung", because it was published in 1892, would be an exploratory essay, and that the hints there of the Frege-Carnap view were superseded by the final, official *Grundgesetze* view. Yet clearly the *Grundgesetze* was the fruit of many years work, and it is hard to imagine that by 1892 Frege had not even proved his Theorem 1, in which the description operator figures.[27] But even if it is Frege's considered opinion, not all is easy with the *Grundgesetze* account, as we will soon see.

5 Formalizing Fregean Theories of Descriptions

In this section we mention some of the semantic consequences of the different theories, particularly we look at some of the semantically valid truths guaranteed in the different theories, as well as some valid rules of inference. One tug in the construction of theories for definite descriptions comes from reflection on these topics, so one way to choose which of the theories should be adopted is to study their semantic consequences. Hence we now turn to these features.

We start by listing a series of formulas and argument forms to consider because of their differing interactions with the different theories. The formulas and answers given by our four different Fregean theories are summarized in a TABLE, along with the answer in Russell's theory. Although the justifications for the answers are brought out in the next four subsections, we present the table here at the beginning in order to be able to refer to the formulas easily.

We rely mostly on informal considerations of what the sentences assert in a theory that embraces the principles mentioned in the last

[26] E.g., Mike Harnish in conversation.

[27] On the other hand, it might be noted that in the Introduction to the *Grundgesetze* (p.6) Frege remarks that "a sign meant to do the work of the definite article in everyday language" is a new primitive sign in the present work. And it is of course well known that Frege says that he had to "discard an almost-completed manuscript" of the *Grundgesetze* because of internal changes brought about by the discovery of the *Bedeutung-Sinn* distinction.

section for the different views on definite descriptions. With regards to Russell's approach, it is well-known what this theory is: classical first-order logic plus some method of eliminating descriptions (that we will discuss shortly). (Indeed, working out this theory is the goal of *14 of *Principia Mathematica*.) The Frege-Carnap theory is developed in Chapter 7 of Kalish & Montague (1964)[28], but we needn't know all the details in order to semantically evaluate our formulas. All we need to do is focus on the sort of interpretations presumed by the theory: namely, those where every improper description designates the same one thing in the domain and this thing might also be designated in more ordinary ways. The Frege-*Grundgesetze* theory similarly can be conceived semantically as containing both objects and courses-of-values of predicates (sets of objects that satisfy the predicate) in the domain. And we can informally evaluate the formulas simply by reflecting on these types of interpretations: improper descriptions designate the set of things that the formula is true of – which will be the empty set in the case of descriptions true of nothing, and will be the set of all instances in those cases where the descriptions are true of more than one item in the domain.[29]

There might be many ways to develop a Frege-Strawson theory, but we concentrate on the idea that improper definite descriptions do not designate anything in the domain and that this makes sentences containing such descriptions be neither true nor false. This is the idea developed by (certain kinds of) free logics: atomic sentences containing improper descriptions are neither true nor false because the item designated by the description does not belong to the domain. (It might, for example, designate the domain itself, as in Simon & Morscher 2001, and Lehman 1994[30]). In a Frege-like development of this idea, we want

[28] And also as Chapter 6 in Kalish, Montague, & Mar (1980).

[29] There will be difficulties in giving a complete and faithful account of the Frege-*Grundgesetze* theory, since its formal development by Frege is contradictory. Even trying to set aside problems with Basic Law V, there will be difficulties in giving an informal account of improper descriptions, because they are supposed to denote a set. And so this set must be in the domain. But we would then want to have principles in place to determine just what sets must be in a domain, given that some other sets are already in the domain. None of this is given by Frege, other than by his contradictory Basic Law V. Some of our evaluations of particular sentences will run afoul of this problem; but we will try to stick with the informal principles that Frege enunciates for this theory, and give these "intuitive" answers.

[30] Kalish, Montague, & Mar (1980), Chapter 8, have what they call a "Russellian" theory that is formally similar to this in that it takes "improper" terms to designate something outside the domain. But in this theory, all claims involving such terms

the lack of a truth-value of a part to be inherited by larger units. Frege wants the *Bedeutung* of a unit to be a function of the *Bedeutungen* of its parts, and if a part has no *Bedeutung*, then the whole will not have one either. In the case of sentences, the *Bedeutung* of a sentence is its truth value, and so in a complex sentence, if a subsentence lacks a truth value, then so will the complex. In other words, the computation of the truth value of a complex sentence follows Kleene's (1952: 334) "weak 3-valued logic", where being neither true nor false is inherited by any sentence that has a subpart that is neither true nor false[31]

An *interpretation* of a language is an assignment of semantic values to the syntactic items of the language. For example, an interpretation could assign a set of things to each (one-place) predicate, with the intuitive meaning that according to this interpretation the predicate is true of each item in the set. And it might assign an individual thing as the interpretation of a name, for example. Different underlying theories might require different sorts of semantic values for the same syntactic item, or one theory may only allow a proper subset of the interpretations allowed by some other underlying theory. Although an interpretation assigns some semantic item to every symbol in the language, it is normally the case that the assignments to syntactically complex items are computed on the basis of the assignments to the syntactically simple items. Also, while an interpretation assigns something to every symbol in the language, in fact when we consider the assignment made to some specific syntactic item by an interpretation, we need not consider what the interpretation does to items not mentioned in the specific syntactic item. In order, for example, to discover what a particular interpretation assigns to the syntactic item $\forall x(Fx \supset Gx)$, we need not consider what the interpretation has to say about predicates other than F and G.

As one can see, there are many, many different interpretations for a language even for just one underlying theory. But sometimes all interpretations yield the same result. For example, in the special case of sentences, whose interpretation is a truth value, it may turn out that every interpretation (which is legitimatized by the underlying theory)

are taken to be false, rather than "neither true nor false". (It seems wrong to call this a "Russellian" theory, since singular terms are not eliminated. It might be more accurate to say that it is a theory that issues forth with sentences that have singular-term definite descriptions that have the same truth value as the Russellian sentences do when descriptions are eliminated.)

[31] There are certainly other 3-valued logics, but Frege's requirement that the *Bedeutung* of a whole be a function of the *Bedeutungen* of the parts requires the Kleene "weak" interpretation.

assigns the same value. In these cases we say that the sentence is *logically true* or *logically false* according to the underlying theory, depending on whether all interpretations say it is true or they all say it is false.

Since we are usually looking at the cases where the descriptions are improper, most interpretations we consider will be called *i-interpretations* (for "improper description interpretations"). In an i-interpretation for a particular formula, all definite descriptions mentioned in the formula are improper. If it should turn out that the formula under consideration is false in every i-interpretation (of the sort relevant to the theory under consideration), then we will call it *i-false*, i.e., false in every interpretation for the theory where the descriptions mentioned in the formula are improper. Similarly, we call some formulas *i-true* if they are true in every i-interpretation that is relevant to the theory. Of course, if a formula is true (or is false) in *every* interpretation (not restricted to i-interpretations) for the theory, then it will also be i-true (or i-false); in these cases we say that the formula is *logically true* or *logically false* in the theory. If a formula is true in some i-interpretations and false in other i-interpretations, then it is called *i-contingent*. Of course, an i-contingent formula is also simply contingent (without the restriction to i-interpretations). Similar considerations hold about the notion of *i-validity* and *i-invalidity*. An argument form is i-valid if and only if all i-interpretations where the premises are true also make the conclusion true. If an argument form is valid (no restriction to i-interpretations), then of course it is also i-valid.

In the case of the Frege-Strawson theory, sentences containing an improper description are neither true nor false in an i-interpretation. We therefore call these *i-neither*. When we say that an argument form is *i-invalid** (with the *), we mean that in an i-interpretation the premise can be true while the conclusion is neither true nor false (hence, not true).

Things are more complex in Russell's theory. For one thing, the formulas with definite descriptions have to be considered "informal abbreviations" of some primitive sentence of the underlying formal theory. And there can be more than one way to generate this primitive sentence from the given "informal abbreviation", depending on how the scope of the description is generated. If the scope is "widest", so that the existential quantifier corresponding to the description becomes the main connective of the sentence, then generally speaking[32], formulas with

[32]But not always; see formula #4 in our TABLE.

improper descriptions will be i-false. But often they will be contingent (without the i) because there will be non-i-interpretations in which there is such an item and others in which there is not. Sometimes the description itself is contradictory and therefore the sentence is logically false (because the wide-scope elimination would assert the existence of an entity with the contradictory property), and hence also i-false.

Furthermore, there might be definite descriptions that are true of a unique object as a matter of logic, such as 'the object identical to a'; and in these latter cases, if the remainder of the sentence is "tautologous" then the sentence could be logically true... for example, "Either the object identical with Adam is a dog or the object identical with Adam is not a dog", whose wide scope representation would be (approximately) 'There exists a unique object identical with Adam which either is a dog or is not a dog.'[33]

We will therefore take all descriptions in Russell's theory to have narrow scope, and so our claims in the TABLE about i-truth, i-falsity, i-contingency, i-validity, and i-invalidity in Russell should be seen as discussing the disambiguation of the "informal abbreviation" with narrowest scope for all the descriptions involved, and then assuming that there is no unique object that satisfies the property mentioned in the description.

One final remark should be made about the interpretation of the TABLE. It was our intent that the various Fs and Gs that occur in the formulas should be taken as variables or schema, so that any sentence of the form specified would receive the same judgment. But this won't work for some of our theories, since predicate substitution does not preserve logical truth in them. For example, in Frege-Strawson, if we substitute a complex predicate containing a non-denoting definite description for the predicate F in a logical truth, a logical falsehood, or a contingent formula, then the result becomes neither true nor false. So the Frege-Strawson theory does not preserve semantic properties under predicate substitution. In Russell's theory, substitution of arbitrary predicates for the Fs and Gs can introduce complexity that interacts with our decision to eliminate all descriptions using the narrowest scope.

[33] *Principia Mathematica* had no individual constants, so this description could not be formed. It is not clear to us whether there is any formula that can express the claim that it is logically true that exactly one individual satisfies a formula, if there are no constants. Since Russell elsewhere thinks that proper names of natural languages are disguised descriptions, it is also not clear what Russell's views about forming these 'logically singular' descriptions in English might be.

For example, #3, when F stands for 'is a round square', generates Russell's paradigm instance of a false sentence: 'The round square is a round square'. This judgment of falsity is generated when we eliminate the definite description in Russell's way, and is what we have in our TABLE. But were we to uniformly substitute $\neg F$ for F in formula #3, we generate $\neg F\imath x \neg Fx$; and eliminating the description in this formula by narrowest scope we have

$$\neg\exists x(\neg Fx \wedge \forall y(\neg Fy \supset x = y) \wedge Fx)$$

It can be seen that the formula inside the main parentheses is logically false (regardless of whether the description is or isn't proper), and so there can be no such x. And therefore the negation of this is logically true. Yet, we followed Russell's rule in decreeing that the original formula #3 is false when the description is eliminated with narrowest scope. This example shows that Russell's theory allows one to pass from logical falsehood to logical truth by uniform predicate substitution. Were we to start with $\neg F\imath x \neg Fx$ (which, as we have just seen, is logically true regardless of whether the description is or isn't proper) and do a predicate substitution of $\neg F$ for F, we would get $\neg\neg F\imath x \neg\neg Fx$, that is, $F\imath xFx$. But we have just seen that this is i-false. So Russell's theory does not preserve logical truth under predicate substitution, unless one is allowed to alter the scope of description-elimination.[34] To avoid all these difficulties, therefore, we are going to restrict our attention to the case where the Fs and Gs are atomic predicates in the theories under consideration here. And so we will not be able to substitute $\neg F$ for F in #3, with this restriction.

[34] Gödel (1944: 126) expresses concern about whether Russell's theorems and definitions hold up under substitutions, and connects this with the issue of scope distinctions.

TABLE: Five Theories of Descriptions' View of Some Formulas.

	Formula or Rule	F-H	F-S	F-C	F-Gz	Russell
1	$i)\forall xFx \vDash F\imath xGx$	ill-form.	i-invalid*	valid	valid	i-invalid
	$ii)F\imath xGx \vDash \exists xFx$	valid	valid	valid	valid	valid
	$iii)\forall xFx \supset F\imath xGx$	ill-form.	i-neither	log.true	log.true	i-false
	$iv)F\imath xGx \supset \exists xFx$	ill-form.	i-neither	log.true	log.true	log.true
2	$\exists y\ y = \imath xFx$	ill-form.	i-neither	log.true	log.true	i-false
3	$F\imath xFx$	ill-form.	i-neither	i-contin.	i-false	i-false
4	$(P \vee \neg P) \vee G\imath xFx$	ill-form.	i-neither	log.true	log.true	log.true
5	$F\imath xFx \vee \neg F\imath x\neg Fx$	ill-form.	i-neither	log.true	i-false	log.true
6	$(\exists x\forall x(Fx \equiv x = y)) \supset F\imath xFx$	ill-form.	i-neither	log.true	log.true	log.true
7	$G\imath xFx \vee \neg G\imath xFx$	ill-form.	i-neither	log.true	log.true	log.true
8	$\imath xFx = \imath xFx$	ill-form.	i-neither	log.true	log.true	i-false
9	$\imath x\ x\neq x = \imath x\ x\neq x$	ill-form.	i-neither	log.true	log.true	log.false
10	$\imath x\ x=x = \imath x\ x\neq x$	ill-form.	i-neither	log.true	i-false	log.false
11	$(\imath xFx = \imath x\ x\neq x) \vee (\imath x\neg Fx = \imath x\ x\neq x)$	ill-form.	i-neither	log.true	i-contin.	log.false
12	$(G\imath xFx \wedge G\imath x\neg Fx) \supset G\imath xGx$	ill-form.	i-neither	log.true	i-contin.	log.true
13	$(\imath xFx = \imath xGx) \supset G\imath xFx$	ill-form.	i-neither	i-contin.	i-false	log.true
14	$\forall x(Fx \equiv Gx) \vDash \imath xFx = \imath xGx$	ill-form.	i-invalid	valid	valid	i-invalid
15	$(G\imath xFx \wedge F\imath xGx) \supset \imath xFx = \imath xGx$	ill-form.	i-neither	log.true	log.true	log.true
16	$i)\ \forall x(Sxa \equiv x = b)$ $ii)\ b = \imath xSxa$	interder. not equiv.	interder. not equiv.	not inter-derivable	not inter-derivable	log.equiv. interder.

5a Frege-Hilbert

As we said above, the Frege-Hilbert treatment requires that descriptions be proper before they can even be used in forming a sentence. That is, the propriety is a precondition of well-formedness. This will obviously lead to problems in giving an account of what the well-formed formulas of the language are, although Carnap remarks on how this may not be such an issue in the context of formalizing mathematics. In this context, he suggested, before using any description, a mathematician will first prove it to be proper. And only then will it appear in formulas. We shall return to this alleged amelioration shortly, after discussing some

logical features of such a system.

No formula of the Frege-Hilbert system that contains a definite description can be guaranteed to be well-formed unless there are constants that can be used to describe "logically singular" predicates, as discussed above. And some of them can even be guaranteed to be ill-formed, when they contain "impossible" definite descriptions like $\iota x\, x{\neq}x$. These are all marked as "ill-formed" in the TABLE, even though of course some instances of the formula (namely, when the descriptions are proper) will be well-formed (and true). Although none of these formulas must be well-formed, we can nonetheless have valid arguments employing them, because if an argument has an "empirical" (i.e., not logically impossible) definite description in its premise, then since a valid argument is one where, if the premise is true so is the conclusion, we are given that the premise is true and therefore its description is proper. Hence, for example, #1(ii) in TABLE, $F\iota xGx \vDash \exists xFx$ is valid, since whenever the premise is true so is the conclusion. When a description appears in the conclusion, however, matters are different. The arguments #1(i) and #14 in TABLE are marked as "ill-formed" because the description mentioned in the conclusion can be improper while the premise is true.

#16 is interesting: if 16(i) is a premise, then there is a unique thing that bears the S relation to a (so the description is proper), and that thing is b (and hence the conclusion, 16(ii) is true). If 16(ii) is true, then the description is proper, and b is the unique thing which bears S to a; and thus 16(i) must be true. But although 16(i) and 16(ii) are thus interderivable, they are not equivalent, since the description might be improper and hence their biconditional could be ill-formed. This shows a peculiarity in the Frege-Hilbert method. If one can prove independently that there is a unique F, then one can use that conclusion to introduce the definite description ιxFx. But one cannot assert as a theorem that these are equivalent facts unless one has an independent premise that there is a unique F!

It should be noted that some of the expressions in TABLE cannot be well-formed in F-H: the definite descriptions in formulas #9-11, where we form descriptions from self-identity and non-self-identity, yield these "non-empirical" definite descriptions. Certainly $\iota x\, x{\neq}x$ must be improper. Hence all of #9-11 are ill-formed on the Frege-Hilbert theory, despite some of them looking like instances of $(P \vee \neg P)$ and others looking like instances of $a = a$. But if, on the other hand, we were to consider just instances of the formulas in TABLE where the descriptions are proper (thus excluding #9-11), so that the formulas are well-formed,

then there are no real surprises in the Frege-Hilbert theory; for, these sorts of definite descriptions act exactly like ordinary proper names.[35] All such descriptions are just "ordinary names" that happen to have a descriptive component. Being "ordinary names", they designate an object, and therefore raise no issues over and above the issues raised by proper nouns generally (such as, objectual vs. substitutional quantification, substitution into intensional contexts, etc.). Where we given that the descriptions are proper,

$$\#2\ \exists y\ y = \imath x Fx.$$
$$\#3\ F\imath x Fx$$
$$\#6\ (\exists x \forall x(Fx \equiv x = y)) \supset F\imath x Fx$$
$$\#8\ \imath x Fx = \imath x Fx$$
$$\#13\ (\imath x Fx = \imath x Gx) \supset G\imath x Fx$$
$$\#14\ \forall x(Fx \equiv Gx) \vDash \imath x Fx = \imath x Gx$$

will be logically true (or valid, our #14). (Of course, a premise of the form in #14 can be true without the descriptions being proper, and so the conclusion may be ill-formed. But we are not considering that case. If the conclusion is well-formed, it follows that there is exactly one F and exactly one G; and the premise then guarantees that they are one and the same object.)

This is perhaps not the only way of visualizing the Frege-Hilbert method. We mentioned above the suggestion about mathematical usage, and we might more charitably interpret Frege-Hilbert as introducing definite descriptions as abbreviations. They come about by first proving the existence and uniqueness of the description, and then it is allowed to be used just as any name is. In this view, the "logical form" of $G\imath x Fx$, then, would become $\exists x(\forall y(Fy \equiv x = y)\ \&\ Gx)$, rather as Russell has it. But here, since all descriptions are proper, they all take widest scope, rather than being ambiguous as in Russell's theory. It seems to us, however, that under this interpretation, we no longer have a Fregean view, but rather (an alternative version of) a Russellian view.

[35] Because they are treated as "ordinary names", descriptions cannot be used in the rules of Existential Instantiation and Universal Generalization . . . any more than ordinary names can. The rules prohibit "non-arbitrary" names, and so we cannot use a definite description in place of α in the rules

$\exists x Fx \vDash F\alpha$

$F\alpha \vDash \forall x Fx$.

Another way to ameliorate the difficulty of having meaningless sentences because of improper descriptions is to explicitly build in the possibility of failure of reference when introducing the descriptions. For example, we might define "The set of F's" as[36]:

$$\{x : Fx\} = y \text{ iff}$$
$$\forall w(w \in y \equiv ((y \text{ is a set } \& Fw) \vee (y = \emptyset \ \& \neg \exists b \forall z(z \in b \equiv Fz))))$$

Note here that if there is no set all of whose members (and only them) are F, then 'the set of F's' is said to designate the empty set. But now we no longer have a Frege-Hilbert theory, and rather have a Frege-Carnap theory. What this shows is that there is always an easy transition from a Hilbert theory to a Carnap theory when one uses definitions to establish propriety of descriptions. For, this latter style of definition explicitly builds in the "failure of denotation" into the last disjunct and thereby provides a *Bedeutung* for the description even in the case of apparent denotation failure, and thus avoids meaninglessness. But despite the fact that a Hilbert theory can be turned into a Carnap theory by this formal trick, the two types of theory are very different: in one theory we have meaninglessness while in the other we have truth and falsity.

Alonzo Church's "Logic of Sense and Denotation" (Church 1951) may be seen, from one point of view, as formalizing the Frege-Hilbert account.[37] Though it does not include constants as Church originally presented the theory, the logic can be supplemented with an expression naming the sense of the definite description 'the f', where nothing, or more than one thing, is f. It will not, however, allow an expression denoting an individual which is f; in particular, it will not contain a symbolization of 'the f'. There can be a name for its sense, but there can be no description of an individual that the sense denotes! Thus descriptions for individuals (or senses) can only be introduced when guaranteed a denotation, even though senses that don't denote objects can nevertheless be named. (See Anderson 1984, p.375.)

5b Frege-Strawson

A Frege-Strawson approach to definite descriptions is one where improper descriptions have no designation (at least, not in the universe

[36] See Suppes, 1960, §2.5, pp.33 ff.

[37] As we remaked earlier, it can also be seen as a development of the Frege-Strawson theory.

of objects), and sentences containing such descriptions have no truth value. As stated, this principle would decree that, when '$\imath xFx$' is improper, such sentences as:

#2 $\exists y\ y = \imath xFx$
#4 $(P \vee \neg P) \vee G\imath xFx$
#6 $(\exists x \forall x(Fx \equiv x = y)) \supset F\imath xFx$
#7 $G\imath xFx \vee \neg G\imath xFx$
#7a $G\imath xFx \supset G\imath xFx$
#8 $\imath xFx = \imath xFx$

have no truth value, however implausible this might sound. A way to semantically describe such a logic[38] is to maintain the classical notion of a model (as containing a nonempty domain D, an interpretation function defined on names, including descriptive names, and an interpretation function that interprets n-ary predicates as subsets of D^n), but to allow (some) names to designate D rather than an element of D and to modify the compositional interpretation rules. Variants of this approach have been in vogue for free logics, where some names lack a denotation in the domain, and can be equally well applied to the case of definite descriptions. (See Lambert & van Fraassen 1967; see also Lehmann 1994 for a variant.)

In a Frege-Strawson semantics, the interpretation in a model of all simple names is either some element of D or else D itself. We will use $[\![\phi]\!]$ to mean the semantic value of ϕ in the interpretation under consideration. The interpretation of descriptive names in an interpretation is similar to that of simple names, but subject to this proviso:

> If a is the unique element of D which is F,
> then $[\![\imath xFx]\!] = a$, otherwise $[\![\imath xFx]\!] = D$

Truth-in-(interpretation)-I for atomic formulas is defined as:

- $F^n(a_1, \ldots a_n)$ is true-in-I iff $<[\![a_1]\!], \ldots [\![a_n]\!]> \in [\![F^n]\!]$
- $F^n(a_1, \ldots a_n)$ is false-in-I iff $<[\![a_1]\!], \ldots [\![a_n]\!]> \notin [\![F^n]\!]$ and $\forall i\ [\![a_i]\!] \in D$
- $a_1 = a_2$ is true-in-I iff $[\![a_1]\!] = [\![a_2]\!]$, $[\![a_1]\!] \in D$ and $[\![a_2]\!] \in D$

[38] There are other ways, but they don't seem so natural to us. (See Morscher & Simons 2001 for a survey).

- $a_1 = a_2$ is false-in-I iff $[\![a_1]\!] \neq [\![a_2]\!]$, $[\![a_1]\!] \in D$ and $[\![a_2]\!] \in D$

True-in-I and false-in-I for the propositional connectives $\neg$ and $\vee$ (which can serve as exemplars for the others) are:

$\neg\Phi$ is true-in-I iff Φ is false-in-I

$\neg\Phi$ is false-in-I iff Φ is true-in-I

$(\Phi \vee \Psi)$ is true-in-I iff either: Φ is true-in-I and Ψ is true-in-I
or Φ is true-in-I and Ψ is false-in-I
or Φ is false-in-I and Ψ is true-in-I

$(\Phi \vee \Psi)$ is false-in-I iff Φ is false-in-I and Ψ is false-in-I

(We note that, for instance, if the atomic formula Fa is neither true-in-I nor false-in-I because $[\![a]\!] = D$, then both $\neg Fa$ and $(Fa \vee P)$ will likewise be neither true-in-I nor false-in-I). Quantification in this logic is over elements in the domain only, and therefore needs no special treatment different from classical quantification theory. (Legitimate values of assignment functions are always in the domain).

Let us consider some of the semantic consequences of this conception. None of:

#2 $\exists y = \iota x Fx$

#3 $F\iota x Fx$

#7 $G\iota x Fx \vee \neg G\iota x Fx$

#8 $(\iota x Fx = \iota x Fx)$

#8a $(\iota x Fx = \iota x Fx) \vee (\iota x Fx \neq \iota x Fx)$

#14 $\forall x(Fx \equiv Gx) \vDash \iota x Fx = \iota x Gx$

are logically true, since '$\iota x Fx$' might be improper and hence not designate anything in D. So as a consequence of denying that definite descriptions always designate something in the domain (#2), in Frege-Strawson we are not guaranteed that instances of tautologies are also tautologies (#4, #8a), nor that self-identity is a law (#8), nor the identity of co-extensionals (#14). If a description does not designate anything in D, then the atomic sentence in which it occurs is neither true-in-I nor false-in-I in such an interpretation. And therefore any more complex formula in which it occurs will be neither true-in-I nor false-in-I. Similarly,

#1i $\forall x Gx \vDash G\iota x Fx$

is not a valid rule of inference, but as in the Frege-Hilbert theory,

$$\#1\text{ii } G\iota xFx \models \exists yGy$$

is a valid rule. (Given that the premise is true, it follows that ιxFx is proper, and hence the conclusion would be true). However, the corresponding conditional

$$\#1\text{iv } G\iota xFx \supset \exists yGy$$

is not logically true, since the antecedent might lack a truth value, thereby making the whole formula truth-valueless. A similar remark can be made about #13 and #13a, where the conditional of #13 is replaced by a $\models$:

$$\#13\ (\iota xFx = \iota xGx) \supset G\iota xFx$$
$$\#13\text{a}\ (\iota xFx = \iota xGx) \models G\iota xFx$$

#13a is a valid argument form in Frege-Strawson, since in order for the premise to be true, the descriptions must be proper, and in such a case we would have $G\iota xGx$ but also the premise that $\iota xFx = \iota xGx$ and hence $G\iota xFx$. But #13 is i-neither true nor false.

One can imagine, following Lambert & van Fraassen (1967), modifications of the Universal Instantiation rule that would be valid for this interpretation of definite descriptions, for example:

$$\forall xGx, \exists yy = \iota xFx \models G\iota xFx$$

And similarly, restrictions could be employed to single out the instances of the above list that are semantically valid; for example:

$$\exists y\ y = \iota xFx \models \iota xFx = \iota xFx$$
$$\exists y\ y = \iota xFx \models F\iota xFx$$

and so on. It can once again be seen that the deduction theorem does not hold here, for, although the former are valid inferences ("if the premise is true then so is the conclusion"), the following are not logically true because they have no truth value if ιxFx is improper:[39]

$$(\exists y\ y = \iota xFx) \supset (\iota xFx = \iota xFx)$$
$$(\exists y\ y = \iota xFx) \supset F\iota xFx.$$

[39]It can be seen that the Frege-Strawson system given here is a sort of "gap theory", where atomic sentences containing improper descriptions are truth-valueless, and the evaluation rules for the connectives follow Kleene's weak logic (1952: 334).

An inspection of the TABLE will reveal that there is no logical difference between the Frege-Hilbert and the Frege-Strawson theories, other than a choice of words: whether to call the formulas and arguments with improper descriptions "ill-formed" or "neither true nor false". But whatever one calls them, the difference is only "fluff", since this alleged semantic difference never makes a logical difference.

5c Frege-Carnap

According to the Frege-Carnap theory of descriptions, each interpretation provides a referent for improper definite descriptions, indeed, in each interpretation it is the same referent that is provided for all improper descriptions. This referent is otherwise just one of the "ordinary" items of the domain, and it has whatever properties the interpretation dictates that this "ordinary" item might happen to have. If this object is the number 0, then sentences like 'The square root of 4 is less than 1' will be true in that interpretation. If the object is the null set then sentences like 'The prime number between 47 and 53 is a subset of all sets' will be true in that interpretation. A formula is logically true if it is true in every interpretation, and for a formula with a description this means that it is true regardless of which item in the domain is chosen as the referent for all improper descriptions and no matter what properties this object has. When Carnap (1956: 36ff) developed the theory, he used 'a*' as the designation of all improper descriptions. It presumably is because Carnap used this special name that Montague and Kalish have said that the method has the feature of being "applicable only to languages which contain at least one individual constant" (MK, p.64). But this is not true, for there is at least one description which is improper in every interpretation: $\iota x\ x{\neq}x$; and so we can use this descriptive name as a way to designate the referent of every improper description. And in doing so, we will not require any non-descriptive name at all, because in each model every other improper description will denote the same as $\iota x\ x{\neq}x$.[40] We will explain what we mean by this by describing how the theory plays out in K-M.

A crucial feature of the theory is that there is always a referent for $\iota x\ x{\neq}x$, and so this means that rules of universal instantiation and existential generalization can be stated in full generality:

[40] It is also true in the Frege-Carnap theory that all definite descriptions can be eliminated, except for occurrences of $\iota x\ x{\neq}x$, and the result will be logically equivalent. We will not prove this, but it follows from Thm 426 of KMM (p. 406).

#1i $\forall x Gx \models G\imath xFx$

#1ii $G\imath xFx \models \exists y Gy$

and the deduction theorem holds, or at least, #1iii and #1iv do not form counterexamples to it:

#1iii $\forall x Fx \supset F\imath xGx$

#1iv $F\imath xGx \supset \exists x Fx$

It also means that self-identities can be stated in full generality, since

#8 $\imath xFx = \imath xFx$

#9 $\imath x\ x{\neq}x = \imath x\ x{\neq}x$

and it means that

#2 $\exists y\ y = \imath xFx$

is logically true. Because every description has a referent, we also have

#7 $G\imath xFx \vee \neg G\imath xFx$

#7a $G\imath xFx \supset G\imath xFx$

#14 $\forall x(Fx \equiv Gx) \models \imath xFx = \imath xGx$.[41]

K-M say that the "essence of Frege" ('Frege' being their name for the Frege-Carnap theory) is:

$G\imath xFx \equiv (\exists y[\forall x(Fx \equiv x = y)\&Gy] \vee [\neg\exists y\forall x(Fx \equiv x = y)\&G(\imath x\ x{\neq}x)])$

that is, 'The F is G' is true just in case either there is exactly one F and it is G, or there isn't exactly one F but the denotation of $\imath x\ x{\neq}x$ is G. Alternatively, and slightly more weakly, we might say:

$F\imath xFx \equiv (\exists y\forall x(Fx \equiv x = y) \vee F(\imath x\ x{\neq}x))$

That is, 'The F is F' is true just in case either there is exactly one F or else the denotation of improper descriptions is F. K-M develop the Frege-Carnap theory by using two very simple rules of inference, Proper Description and Improper Description:

[PD] $\exists x\forall y(Fx \equiv x = y) \models F\imath xFx$

[ID] $\neg\exists x\forall y(Fx \equiv x = y) \models \imath xFx = \imath x\ x{\neq}x$

[41] Morscher & Simons (2001: p.21) call this "the identity of coextensionals" and say it is an "obvious truth" that should be honored by any theory of descriptions.

Given how well this theory comports with intuition on the above logical features, one might be tempted to adopt it despite the slight unnaturalness involved in saying that all improper descriptions designate the same object, which is some "ordinary" object that has "ordinary" properties and can also be designated in some more "ordinary" manner. However, there are some logical features of the theory that may give one pause. We do not have the (natural-sounding)

#3 $F\iota xFx$

#13 $\iota xFx = \iota xGx \supset G\iota xFx$

for, if ιxFx is improper, then on this theory the object denoted is not necessarily an F. For example, although 'the golden mountain' denotes something, what it denotes is not necessarily golden (nor a mountain). And this denoted object might be identical with the object denoted by ιxGx, but this is no guarantee that it will have the property G, for that depends in part on whether there is a unique G or whether the chosen object in the domain has property G.

Furthermore, since all improper descriptions designate the same object we have the somewhat peculiar logical truth:

#10 $\iota x\ x{=}x = \iota x\ x{\neq}x$.

(In domains where there is exactly one object, then this object is the unique self-identical object, and as well must serve as the designation for all improper descriptions, such as $\iota x\ x{\neq}x$; in any larger domain, both of the descriptions are improper and hence designate the same object of the domain, whatever it may be). And we have the decidedly peculiar logical truths in Frege-Carnap:

#5 $F\iota xFx \vee \neg F\iota x\neg Fx$

#11 $(\iota xFx = \iota x\ x{\neq}x) \vee (\iota x\neg Fx = \iota x\ x{\neq}x)$

(For formula #5, in any model where there is a unique thing that is F or a unique thing that is $\neg F$, then the formula will be true. But if otherwise, then both ιxFx and $\iota x\neg Fx$ are improper, and hence both denote that object in the domain which is chosen for all improper descriptions. But that object is either F or $\neg F$. Thus #5 is logically true. For formula #11, if either ιxFx or $\iota x\neg Fx$ is improper the formula will be true. But if otherwise, then both are proper; and this can happen only in a two-element domain where one element is F and the other is $\neg F$. But in such a domain, one or the other of these must be the denotation

chosen for improper descriptions, and hence one or the other disjunct will be true.

Further seemingly implausible candidates for logical truth are these formulas, which turn out to be logical truths in Frege-Carnap:

$$\#12\ (G\iota xFx\ \&\ G\iota x\neg Fx) \supset G\iota xGx$$
$$\#15\ (G\iota xFx\ \&\ F\iota xGx) \supset \iota xFx = \iota xGx\ .$$

#12 cannot be falsified, because to make the consequent false, we would need both for ιxGx to be an improper description and for G not to be true of $\iota x\ x{\neq}x$. But consider the antecedent. If either ιxFx or $\iota x\neg Fx$ is improper, then G is true of $\iota x\ x{\neq}x$. On the other hand, the only way for both ιxFx and $\iota x\neg Fx$ to be proper is in a two element domain, where one of these elements is F and the other is $\neg F$. In that case, though, one or the other of these would have to be the denotation of $\iota x\ x{\neq}x$, and so even in this case an object would have to be G, so #12 is logically true. With respect to #15, if both ιxFx and ιxGx are improper, then the consequent (and hence the whole formula) is true. If at least one of them is proper, the following happens (let's assume it is ιxFx that is proper). Since ιxFx is proper, there is exactly one F, and the antecedent says that ιxGx has this property. But again, since ιxFx is proper, we have $F\iota xFx$; and therefore ιxFx and ιxGx must be the same.

There are also difficulties of representing natural language in the Frege-Carnap theory. Consider #16i and ii, under the interpretation "Betty is Alfred's only spouse" and "Betty is the spouse of Alfred", represented as

$$\#16\text{i}\ \ \forall x(Sxa \equiv x = b)$$
$$\#16\text{ii}\ b = \iota xSxa$$

While the two English sentences seem interderivable (if you know that one was true, you could derive the other), the symbolized sentences are not, in Frege-Carnap: consider Alfred unmarried and Betty being the designated object. Then the first sentence is false but the second is true, so they cannot be interderivable. Since they are not interderivable, they cannot be equivalent.

Another (arguable) mismatch between the Frege-Carnap theory and natural language is that there is no notion of "primary vs. secondary scope of negation" in this theory. The two apparent readings of a sentence like 'The present king of France is not bald' turn out to be equivalent in Frege-Carnap (see KMM p. 405). This seems like a bad result

for the theory, and even KMM (who otherwise favor the theory) admit that "the Russellian treatment is perhaps closer to ordinary usage" in this regard.[42] These are but small pieces of the general problem with the Frege-Carnap theory as an account of definite descriptions in English. However, they should be enough to show that the often-made claim that improper descriptions are "waste cases" or "uninteresting" or "it doesn't matter which decision we take about them" will not stand up to scrutiny. It makes a difference what we say about improper descriptions. It wont do simply to pick any old arbitrary thing to which they will refer, without considering the logical features of such a choice. This is a topic to which we return later.

5d Frege-*Grundgesetze*

On a very intuitive level, many of the desirable logical features of the Frege-Carnap theory hold also in the Frege-Grundgesetze theory, since there is always a *Bedeutung* for every definite description:

$$\#2\ \exists y\ y = \text{ı}xFx.$$

Hence

$$\#5\ G\text{ı}xFx \vee \neg G\text{ı}xFx\ ,$$
$$\#8\ \text{ı}xFx = \text{ı}xFx\ ,$$
$$\#11\ \text{ı}x\ x{\neq}x = \text{ı}x\ x{\neq}x$$

will always be true, whether there are no Fs, just one, or more than one. And the identity of co-extensionals, #14, will be valid because the *Bedeutung* of '$\text{ı}xFx$' is a function of what F is true of:

$$\#14\ \forall x(Fx \equiv Gx)\ \vDash \text{ı}xFx = \text{ı}xGx\ .$$

The various rules fall out as follows:

$$\#1\text{i}\ \forall xGx\ \vDash\ G\text{ı}xFx$$
$$\#1\text{ii}\ G\text{ı}xFx\ \vDash\ \exists yGy$$

are valid rules of inference, because the *Bedeutung* of '$\text{ı}xFx$' is in the domain of the quantifers. As a result the corresponding conditionals:

[42] It is not clear from the context whether KMM mean the informal notion of a Russell-like theory when they say "Russellian treatment" or whether they mean their own 'Russellian theory', which, as we remarked, is not really Russellian.

#1iii $G\iota xFx \supset \exists yGy$

#1iv $\forall xGx \supset G\iota xFx$

are logically true.

However, when it comes to more intricate formulas, discussion of the Frege-*Grundgesetze* theory is hampered by Frege's apparent lack of understanding of sets. In one way or another, this is tied up with his Basic Law V and its seeming appeal to naïve set theory. Frege apparently needs to have sets in his domains, even in those cases where we are not considering the development of set theory as a mathematical object. For improper descriptions designate *Werthverlaufen* (courses of values, i.e., sets) even in the most mundane settings. But then the quantifiers will range over these sets, and we will need principles that tell use which sets are in a domain, given that other sets are already in it. Frege's answer, to judge from Basic Law V, is that *all* sets are in it, and that *every* open formula can designate a set. In such a conception, if ιxFx is improper, then the set $\{x{:}Fx\}$ will be in the domain; and then so will the set $\{x{:}\neg Fx\}$. But we know that such an unrestricted principle will yield a contradiction. Indeed, one can wonder whether Frege would be aware that if both F and $\neg F$ were true of more than one object, then at least one of the apparent improper descriptions ιxFx and $\iota x\neg Fx$ would have to be impossible. For if ιxFx is improper, then $\{x{:}Fx\}$ would be a set in the domain; but then $\{x{:}\neg Fx\}$ could not be a set (it would have to be a "proper class", as we now call them), and so $\iota x\neg Fx$ could have no semantic value.

In Frege's intuitive view (insofar as it can be made out from the introductory remarks in the *Grundgesetze*), it would seem that if ιxFx is improper and thereby designates $\{x{:}Fx\}$, then this set is an entity of which many things could be asserted. We might say it has property G or H; indeed, we might want to say that it even has property F. But this would lead to a contradiction, with its violation of the (later) axiom of foundation: If F is true of $\{x{:}Fx\}$, then $\{x{:}Fx\}$ would have to contain itself.

In a similar but somewhat more complicated vein, let us consider formula

#15 $(G\iota xFx \;\&\; F\iota xGx) \supset \iota xFx = \iota xGx$

Let us first look at the case where both descriptions are proper. Since ιxFx is proper, it follows that $F\iota xFx$. Since there is exactly one F, if the antecedent of #15 is true, then ιxGx must be identical to ιxFx, so the

conclusion of #15 is true. (And also in this case, both F and G hold of both ιxFx and ιxGx). Now consider the case where both descriptions are improper. In this case, $\iota xFx = \{x{:}Fx\}$ and $\iota xGx = \{x{:}Gx\}$. But if the antecedent of #15 is true, then $\{x{:}Fx\} \in \{x{:}Gx\}$ and also $\{x{:}Gx\} \in \{x{:}Fx\}$. But this is impossible, according to the axiom of foundation. So in this case the antecedent of #15 is false, and hence #15 is true. Now consider the remaining case where just one of the descriptions is proper, say, ιxFx is proper. Since it is proper, we have $F\iota xFx$, and since there is just one F and $F\iota xGx$, it follows that $\iota xFx = \iota xGx$, and so the consequent of #15 is true. However, it also follows that $\iota xFx = \{x{:}Gx\}$, and since we are given in the antecedent of #15 that $G\iota xFx$, it would follow that $G\iota xGx$, which we already know to be impossible. Thus it is not possible for even one of the descriptions to be improper, if the antecedent of #15 is to be true.

There are some morals here concerning the difficulties encountered when the domain of an interpretation contains both "ordinary" objects and also sets of these objects (and sets of sets, etc.). It is not at all clear that a truly coherent theory of definite descriptions can be constructed from the *Grundgesetze* theory. So we continue with our account, based on intuitive principles that seem to be accepted by Frege but always with a worry that there is an underlying incoherency involved (even apart from Basic Law V).

An important principle in the Frege-*Grundgesetze* theory is that

#3 $F\iota xFx$

is i-false. for if F is not true of exactly one thing, 'ιxFx' will have a course of values as its *Bedeutung*, and as we saw above, F cannot be true of this set, under pain of contradiction. Of course, if 'ιxFx' is proper, then '$F\iota xFx$' will be true, so #3 is not logically false, just i-false. This striking feature, that it is contradictory to assume $F\iota xFx$ if there is more than one F, is evidence against Frege's informal claim that an improper description like 'the square root of 2' must be a square root of 2. In fact, it is logically impossible for an improper description to have this property, in the Frege-*Grundgesetze* theory.

It is of some interest to note just where the Frege-Carnap and the Frege-*Grundgesetze* theories differ. Examination of TABLE reveals, besides #3, which is i-contingent in Frege-Carnap but i-false in Frege-*Grundgesetze*, as we just discussed, the following

#5 $F\iota xFx \vee \neg F\iota x\neg Fx$

#10 $\iota x\ x{=}x = \iota x\ x{\neq}x$.

#11 $(\iota xFx = \iota x\ x{\neq}x) \vee (\iota x{\neg}Fx = \iota x\ x{\neq}x)$

#12 $(G\iota xFx\ \wedge\ G\iota x{\neg}Fx) \supset G\iota xGx$

#13 $\iota xFx = \iota xGx \supset G\iota xFx$

as places where they differ.

Formula #13 is i-contingent in Frege-Carnap, but is i-false in Frege-*Grundgesetze*. For, if both descriptions are improper, then in particular we could not have $G\iota xGx$. But the antecedent says that $\iota xFx = \iota xGx$; so we also cannot have $G\iota xFx$. Although it is i-false, it is not logically false because it is true when the descriptions are proper.

#5, #10, #11, and #12 are logically true in the Frege-Carnap theory, as explained in the preceding section. But matters are different in Frege-*Grundgesetze*. With regards to #5, if both ιxFx and $\iota x{\neg}Fx$ are improper, then they designate the sets $\{x{:}Fx\}$ and $\{x{:}\neg Fx\}$, respectively. But the former set cannot be F nor can the latter one be $\neg F$, under pain of contradiction.[43] So it is i-false. But it is not logically false, for if one of the descriptions is proper, then #5 is true. With regards to #10, in the Frege-*Grundgesetze* theory, $\iota x\ x{\neq}x$ always designates the empty set, while $\iota x\ x{=}x$ will designate the set of all entities in the domain. This latter description is improper in any domain with more than one element, and the set thereby designated will be different from the empty set. So it is i-false. But it is not logically false, for in a one-element domain $\iota x\ x{=}x$ is proper and denotes that element. And if the element is the empty set, then #10 is true.[44] For #11, since $\iota x\ x{\neq}x$ always designates the empty set on the Frege-*Grundgesetze* theory, one of the disjuncts of #11 is true if either there are no F's or there are no $\neg F$'s. But in any other i-interpretation, #11 would be false because neither the set of F's nor the set of $\neg F$'s would be identical to the empty set.[45] Hence #11 is i-contingent. If all three descriptions in #12 are improper, and if we allow both ιxFx and $\iota x{\neg}Fx$ to designate sets in the domain, then it is possible that there be a predicate G that is true of both these sets (such as "does not contain exactly one member"). In this case the antecedent

[43] Of course, Frege may not have appreciated that this led to a contradiction. And as we remarked before, if F is allowed to determine a set, then $\neg F$ cannot do so unrestrictedly. But again, Frege seems not aware of this.

[44] Once again, it is not clear whether Frege's unstated background theory will allow a domain that consists of only the empty set.

[45] Again, it is not clear that Frege acknowledges that if F has a set as its *Wertverlauf* then $\neg F$ cannot have one.

of #12 would be true; but we have already seen that the consequent cannot be true if $\backprime x Gx$ is improper. On the other hand, it is also possible for the antecedent of #12 to be false, as whenever we use a predicate G that is not true of the two sets. In this case the false antecedent makes #12 be true. Thus, #12 is i-contingent.

As we have been at pains to remark, there are conceptual problems lurking when one adds courses of values (sets) to the domain of individuals, but Frege seems unperturbed by them. There will be a raft of sentences about improper descriptions that will be i-true, and inferences that one can make from a premise that a description is improper. But of course what sentences are true in all interpretations depends on what counts as an interpretation, and this will be complicated by the addition of sets to the domain. The language of *Grundgesetze* includes terms for courses of values, as in Basic Law VI: $a = \dot{\epsilon}\ (a = \ \epsilon)$. What semantic values are allowed for '$\dot{\epsilon}\ (a = \epsilon)$' ? If this must be the singleton set containing the semantic value of 'a', then there will be a range of sentences about courses of values that will be true on every interpretation, including many about the members, identity, and memberships of various courses of values. If it is an interesting question of what would make for a consistent system including these axioms, it is much more difficult to understand what it would even mean to have a complete system for descriptions that employed the *Grundgesetze* framework.

6 Russellian Considerations

Russell criticizes Frege as follows (where Russell says 'denotation' understand 'Bedeutung'; where he says 'meaning' understand 'Sinn'):

> If we say, 'the King of England is bald', that is, it would seem, not a statement about the complex meaning of 'the King of England', but about the actual item denoted by the meaning. But now consider 'the King of France is bald'. By parity of form, this also ought to be about the denotation of the phrase 'the King of France'. But this phrase, though it has a meaning, provided 'the King of England' has a meaning, has certainly no denotation, at least in any obvious sense. Hence one would suppose that 'the king of France is bald' ought to be nonsense; but it is not nonsense, since it is plainly false. (1905a, p.165)

This sort of criticism misses the mark against the most plausibly-Fregean theories, holding only against the Frege-Hilbert theory. (In the Frege-Hilbert theory, the phrase 'the present King of France' does *not* have a meaning (*Sinn*), and is in this way different from 'the present King of England.' It is not part of the Frege-*Grundgesetze* Theory (nor of the Frege-Carnap theory) that 'the King of France is bald' is nonsense. It is, of course, a feature of the Frege-Strawson account that it lacks a truth value, which is still some way from nonsense; for, although it lacks a *Bedeutung* it still has a *Sinn*. A further criticism of the Frege-Strawson view is contained in the sentences just following the above quote:

> Or again consider such a proposition as the following: 'If u is a class which has only one member, then that one member is a member of u', or, as we may state it, 'If u is a unit class, the u is a u'. This proposition ought to be always true, since the conclusion is true whenever the hypothesis is true. ... Now if u is not a unit class, 'the u' seems to denote nothing; hence our proposition would seem to become nonsense as soon a u is not a unit class.
>
> Now it is plain that such propositions do not become nonsense merely because their hypotheses are false. The King in The Tempest might say, 'If Ferdinand is not drowned, Ferdinand is my only son'. Now 'my only son' is a denoting phrase, which, on the face of it, has a denotation when, and only when, I have exactly one son. But the above statement would nevertheless have remained true if Ferdinand had been in fact drowned. Thus we must either provide a denotation in cases which it is at first absent, or we must abandon the view that the denotation is what is concerned in propositions which contain denoting phrases. (1905a, p. 419)

Russell here is arguing against the Frege-Strawson view on which sentences with non-denoting descriptions come out neither true nor false (Russell's "meaningless"?), because if the antecedent of a conditional hypothesizes that it is proper then the sentence should be true. (Our #6 captures this). But as even Russell says, one needn't abandon all singular term analyses in order to obey this intuition. So it is strange that he should think he has successfully argued against Frege, unless it is the Frege-Strawson view that Russell is here attributing to Frege.

And yet Russell had read the relevant passages in "Über Sinn und Bedeutung" as well as *Grundgesetze* in 1902, making notes on them for his "Appendix A on 'The Logical Doctrines of Frege' " to be published in his *The Principles of Mathematics*. Indeed elsewhere in "On Denoting" he does in fact attribute the *Grundgesetze* theory to Frege:

> Another way of taking the same course <an alternative to Meinong's way of giving the description a denotation> (so far as our present alternative is concerned) is adopted by Frege, who provides by definition some purely conventional denotation for the cases in which otherwise there would be none. Thus 'the King of France', is to denote the null-class; 'the only son of Mr. So-and-so' (who has a fine family of ten), is to denote the class of all his sons; and so on. But this procedure, though it may not lead to actual logical error, is plainly artificial, and does not give an exact analysis of the matter. (1905a, p.165)

Even granting Russell's right to call it "plainly artificial", he does not here find any logical fault with Frege's *Grundgesetze* theory. In any case, he joins the other commentators in not remarking on the different theories of descriptions Frege presented in different texts. Indeed he states two of them without remarking on their obvious difference.

On the other hand Russell's presentation of one of his "puzzles" for a theory of descriptions does touch on Frege (cf. our #5):

> (2) By the law of excluded middle, either 'A is B' or 'A is not B' must be true. Hence either 'the present King of France is bald', or 'the present King of France is not bald' must be true. Yet if we enumerated the things that are bald, and then the things that are not bald, we should not find the present King of France in either list. Hegelians, who love a synthesis, will probably conclude that he wears a wig. (1905a, p. 166)

Presumably the empty set, the Frege-*Grundgesetze* theory's *Bedeutung* for 'the present King of France', will be in the enumeration of things that are not bald. With the Frege-Carnap view we just don't know which enumeration it will be in (but it will be in one of them), and with the Frege-Hilbert view we can't use the expression 'the present King

of France' in the first place. Perhaps Russell is attributing the Frege-Strawson theory to the Hegelians. (It is neither true nor false that the present King of France is bald, so he must be wearing a wig!)

Russell's discussion is unfair to Frege's various accounts. Russell's main arguments are directed against Meinong, and since both Meinong and Frege take definite descriptions to be designating singular terms, Russell tries to paint Frege's theory with the same brush as he uses on Meinong's theory. Although there is dispute on just how to count and individuate the number of different Russell arguments against Meinong, we see basically five objections raised in Russell's 1905 works: (a) Suppose there is not a unique F. Still, the sentence 'If there were a unique F, then $F\iota xFx$ ' should be true (cf. our #6). (b) The round square is round, and the round square is square. But nothing is both round and square. Hence 'the round square' cannot denote anything. (c) If 'the golden mountain' is a name, then it follows by logic that there is an x identical with the golden mountain, contrary to empirical fact. (d) The existent golden mountain would exist, so one proves existence too easily. (e) The non-existing golden mountain would exist according to consideration (d) but also not exist according to consideration (b).

But however much these considerations hold against Meinong, Ameseder, and Mally (who are the people that Russell cites), they do not hold with full force against Frege.[46] The first consideration holds perhaps against the Frege-Hilbert and Frege-Strawson views, but we should note that #6 is logically true in both Frege-Carnap and Frege-*Grundgesetze*, just as it is in Russell's theory. Against the second consideration, Frege has simply denied that $F\iota xFx$ is i-true (and it might be noted that Russell's method has this effect also, as can be seen in the TABLE #3), and that is required to make the consideration have any force. Against the third consideration, Frege could have said that there was nothing wrong with the golden mountain existing, so long as you don't believe it to be golden or a mountain. Certainly, whatever the phrase designates does exist, by definition in the various theories of Frege. And against the fourth consideration, Frege always disbelieved that existence was a predicate, so he would not even countenance the case. Nor would the similar case of (e) give Frege any pause.

So, Russell's considerations do not really provide a conclusive argument against all singular term accounts of definite descriptions. And it is somewhat strange that Russell should write as if they did. For as

[46] Perhaps the arguments do not hold against Meinong, Ameseder, and Mally either. But that is a different topic.

we mentioned earlier, in 1902 he had read both "Über Sinn und Bedeutung" and *Grundgesetze*, making notes on Frege's theories.[47] Yet he betrays no trace here of his familiarity with them, saying that they do not provide an "exact analysis of the matter", but never saying how they fall short. In fact, a glance at TABLE reveals that there is one commonality among *all* the theories of descriptions we have discussed: they never treat

$$\#3\ F\iota xFx$$

as logically true, unlike Meinong and his followers. (When the description is improper, the F-C theory treats it as i-contingent – sometimes true, sometimes false; the Frege-*Grundgesetze* theory and Russell's theory treat it as i-false. The Frege-Hilbert and Frege-Strawson theories also treat it as not always true.) It is this feature of all these theories that allows them to avoid the undesirable consequences of a Meinongian view, and it is rather unforthcoming of Russell to suggest that there are really any other features of his own theory that are necessary in this avoidance. For, each of Frege's theories also has this feature.

We wish to compare our various test sentences with the formal account of definite descriptions in *Principia Mathematica* $*14$, so that one can see just what differences there are in the truth values of sentences employing definite descriptions between Russell's theory and the various Frege theories, as summarized in TABLE. Let us first see the ways where Russell's theory differs from the Frege-Carnap and Frege-*Grundgesetze* theories. As can be seen from TABLE, there are many such places. The places where both of these Frege theories agree with one another and disagree with Russell are:

	Formula	Frege	Russell
$\#1i$	$\forall xFx \models G\iota xGx$	valid	i-invalid
$\#1iii$	$\forall xFx \supset F\iota xGx$	log.true	i-false
$\#2$	$\exists y\ y = \iota xFx$	log.true	i-false
$\#8$	$\iota xFx = \iota xFx$	log.true	i-false
$\#9$	$\iota x\ x{\neq}x = \iota x\ x{\neq}x$	log.true	log.false
$\#14$	$\forall x(Fx \equiv Gx) \models \iota xFx = \iota xGx$	valid	i-invalid
$\#16i$	$\forall x(Sxa \equiv x = b)$	not-	
$\#16ii$	$b = \iota xSxa$	interder.	log.equiv.

Most of these differences are due to the fundamental #2. Given that different choice, it is clear that #1i and #1iii must differ as they

[47] See Linsky (2004) for the notes.

do. And Russell's interpretation of a definite description as asserting that there exists a unique satisfier of the description (and his related "contextual definition") will account for all the cases where the formula (or argument) is i-false (or i-invalid) in Russell. However, since it is logically impossible that there be a non-self-identical item, #9 must be logically false, and not just i-false, in the Russell framework.

Besides the already-discussed #3, there are only two of our formulas that are treated differently by each of Frege-Carnap, Frege-*Grundgesetze*, and Russell:

$$\#11\ (\iota xFx = \iota x\ x{\neq}x) \vee (\iota x{\neg}Fx = \iota x\ x{\neq}x)$$

$$\#13\ \iota xFx = \iota xGx \supset G\iota xFx$$

We already described why #11 is logically true in Frege-Carnap: if one or the other of ιxFx and $\iota x{\neg}Fx$ is improper, then #11 is true. But if they are both proper, this requires a two-element domain where one element is F and the other $\neg F$. But in that case, one or the other of the two elements has to be the denotation of $\iota x\ x{\neq}x$, and so even then #11 is true. It is i-contingent in Frege-*Grundgesetze* (assuming we allow both descriptions[48]), because $\iota x\ x{\neq}x$ always designates the empty set. Sometimes one of F or $\neg F$ might have a null extension and then the resulting definite description will also designate the empty set, and #11 will be true. And for other some other cases neither F nor $\neg F$ will have a null extension and then neither of the resulting descriptions will designate the empty set, and #11 will then be false. In Russell there cannot be any object that is not self identical, and thus each disjunct will be false in any interpretation. Hence #11 is logically false.

#13 is logically true in Russell because if the descriptions are improper then the antecedent is false (and hence #13 is true), but if the antecedent is true then both descriptions must be proper, and hence the consequent must be true. But as we discussed before, in Frege-Carnap if the descriptions are both improper, then they designate the same entity and so the antecedent is true. But the consequent might or might not be true in an interpretation depending on whether the object chosen to be the denotation of all improper descriptions happens to have the property G or not. So it is i-contingent. And in Frege-*Grundgesetze*, #13 must be i-false because if the two improper descriptions are identical

[48]Recall that at least one of F and $\neg F$ will have a proper class as its course of values, and that this will raise problems of interpretation for at least one of the two disjuncts.

then $G\iota xFx$ would be equivalent to $G\iota xGx$. And we know that this latter is impossible. Yet #13 is not logically false because if the descriptions are proper then #13 is true.

Most simple statements about descriptions will only hold in Russell if the description is proper, which Whitehead and Russell (1910) indicate with $E!\iota xFx$. There is a theorem with that as an antecedent and $\iota xFx = \iota xFx$ as consequent ($*14{\cdot}28$).

$$\iota xFx = \iota xFx \vee \iota xFx \neq \iota xFx$$
$$G\iota xFx \vee \neg G\iota xFx$$

are true unconditionally, provided that the scope of the descriptions is read narrowly, so that this is an instance of $(P \vee \neg P)$. The following are all theorems, given the assumption that the description is proper:

$$\iota xFx = \iota yFy$$
$$F\iota xFx \ \ (*14{\cdot}22)$$
$$\exists y\ y = \iota xFx\ (*14{\cdot}204)$$
$$\forall x(Fx \equiv Gx) \supset \iota xFx = \iota xGx\ (*14{\cdot}27).$$

The rules are similarly affected:

$$\forall x\ Gx \vDash G\iota xFx$$

requires the additional premise $E!\iota xFx$ ($*14{\cdot}18$), though

$$G\iota xFx \ \vDash \exists y\ Gy$$

does not ($*14{\cdot}21$), as the propriety of the description follows from the truth of the premise. The corresponding conditional:

$$G\iota xFx \supset \exists y Gy$$

is thus logically true.

7 Concluding Remarks

The fundamental divide in theories of descriptions now, as well as in Russell's time, is whether definite descriptions are "really" singular terms, or "really" not singular terms (in some philosophical "logical form" sense of 'really'). If they are "really" not singular terms then this might be accommodated in two rather different ways. One such way is Russell's: there is no grammatically identifiable unit of any sentence in

logical form that corresponds to the natural language definite description. Instead there is a grab-bag of chunks of the logical form which somehow coalesce into the illusory definite descriptions. A different way is more modern and stems from theories of generalized quantifiers in which quantified terms, such as 'all men', are represented as a single unit in logical form and this unit can be semantically evaluated in its own right–this one perhaps as the set of all those properties possessed by every man. In combining this generalized quantifier interpretation of quantified noun phrases into the evaluation of entire sentences, such as 'All men are mortal', the final, overall logical form for the entire sentence becomes essentially that of classical logic. So, although quantified noun phrases are given an interpretable status on their own in this second version, neither does their resulting use in a sentence yield an identifiable portion of the sentence that corresponds to them nor does the interpretation of the quantified noun phrase itself designate an "object" in the way that a singular term does (when it is proper). It instead denotes some set-theoretic construct.

If we treat definite descriptions as a type of generalized quantifier, and thereby take this second way of denying that definite descriptions are "really" singular terms, the logical form of a sentence containing a definite description that results after evaluating the various set-theoretic constructions will (or could, if we made Russellian assumptions) be that which is generated in the purely intuitive manner of Russell's method. So these two ways to deny that definite descriptions are singular terms really amount to the same thing. The only reason the two theories might be thought different is due to the algorithms by which they generate the final logical form in which definite descriptions "really" are not singular terms, not by whether the one has an independent unit that corresponds to the definite description. In this they both stand in sharp contrast to Fregean theories.

These latter disagreements are pretty much orthogonal to those of the earlier generation. The contemporary accounts, which have definite descriptions as being "nearly" a classical quantifier phrase, agree with the Russellian truth conditions for sentences involving them. Although these truth conditions might be suggested or generated in different ways by the different methods (the classical or the generalized quantifier methods) of representing the logical form of sentences with descriptions, this is not required. For one could use either the Russellian or Frege-Strawson truth conditions with any contemporary account. It is clear, however, that we must first settle on an account of improper

descriptions.

We remarked already on the various tugs that a theorist may feel when trying to construct a theory of definite descriptions, and the various considerations that might move a theorist in one direction or another. We would like to point to one further consideration that has not, we think, received sufficient consideration.[49] It seems to us that whatever semantic treatment is advocated for "regular" proper names, that treatment should apply to all simple proper names, regardless of whether they are denoting or non-denoting. There is simply no intuitive, syntactic way to distinguish non-denoting from denoting names in natural language: 'Pegasus' should therefore be given the same semantic treatment as 'Benjamin Franklin', in the sense that the same semantic evaluation rule for proper names should be applied to them both. It also seems to us that improper descriptions have much in common with non-denoting names like 'Pegasus', and should be treated similarly. Just as there is no intuitive way to distinguish non-denoting from denoting proper names in natural language, so too is there no intuitive way to distinguish (empirically) non-proper vs. proper descriptions. So, all these singular terms should be dealt with in the same way. If definite descriptions are to be analyzed away à la Russell, then the same procedure should be followed for 'Pegasus' and its kin. And if for 'Pegasus', then for 'Benjamin Franklin' and its kin. If, on the other hand, 'Benjamin Franklin' is taken to be a singular term that is evaluated semantically as designating an entity, then so too should proper names like 'Pegasus'. And whatever account is given for non-denoting names like 'Pegasus' should also be given for improper descriptions: if non-denoting names are banned from the language, then we should adopt the Frege-Hilbert theory of improper descriptions. If such names have a sense but no denotation in the theory, then we should adopt the Frege-Strawson theory of improper descriptions. If we think we can make meaningful and true statements about Pegasus and its cohort, then we should adopt either the Frege-Carnap or the Frege-*Grundgesetze* theory of improper descriptions.

In any case we should care about the present king of France.

[49] Except from certain of the free logicians, who take the view that sentences which contain non-denoting names are neither true nor false, and this ought to be carried over to non-denoting definite descriptions as well.

References

Anderson, C.A. (1984) "General Intensional Logic", *Handbook of Philosophical Logic*, Vol. II, Reidel: Dordrecht, 355-385.

Beaney, M. (1996) *Frege: Making Sense*. London: Duckworth.

Carnap, R. (1956) *Meaning and Necessity*. 2nd Edition. Chicago: University of Chicago Press.

Church, A. (1951) "A Formulation of the Logic of Sense and Denotation", in P. Henle, H. Kallen, and S. Langer (eds.) *Structure, Method, and Meaning: Essays in Honor of Henry M. Sheffer*. New York: Liberal Arts Press, 3-24.

Dummett, M.A.E. (1981) *The Interpretation of Frege's Philosophy*. Cambridge, Mass: Harvard University Press.

Dummett, M.A.E. (1991) *Frege: Philosophy of Mathematics*. Cambridge Mass: Harvard University Press.

Frege, G. (1884) *Die Grundlagen der Arithmetik: eine logisch-mathematische Untersuchung über den Begriff der Zahl*. W. Koebner: Breslau. Translated by J. Austin (1953, 2nd Ed.) as *The Foundations of Arithmetic. A logic-mathematical enquiry into the concept of number* Oxford: Blackwell.

Frege, G. (1892) "Über Sinn und Bedeutung" *Zeitschrift für Philosophie und philosophische Kritik* 100, 25-50. Translated by M. Black (1952) *Selections from the Philosophical Writings of Gottlob Frege* , 3rd ed., (1980), (Oxford: Blackwell) as "On Sense and Meaning", 56-78. Quotations and page references are to this translation.

Frege, G. (1893) *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet*. Vol. 1. Verlag Hermann Pohle: Jena. Translated M. Furth (1964) as *The Basic Laws of Arithmetic: Exposition of the System*. Los Angeles: University of California Press. Quotations and page references are to this translation.

Gabriel, G. & Dathe, U. (2001) *Gottlob Frege: Werk und Wirkung*. Padeborn: Mentis. Frege's "Vorschläge für ein Wahlgesetz", 297-313.

Gödel, K., (1944) "Russell's Mathematical Logic", in *The Philosophy of Bertrand Russell*, P.A.Schilpp ed., Open Court: LaSalle, 125-153.

Hailperin, T. (1954) "Remarks on Identity and Description in First-Order Axiom Systems". *Journal of Symbolic Logic* 19: 14-20.

Hilbert, D. & P. Bernays (1934) *Grundlagen der Mathematik*, Vol. I. Berlin: Springer.

[KM] Kalish, D. & R. Montague (1964) *Logic: Techniques of Formal Reasoning*. New York: Harcourt Brace Jovanovich.

[KMM] Kalish, D., R. Montague, G. Mar (1980) *Logic: Techniques of Formal Reasoning*. 2nd Edition. New York: Harcourt Brace Jovanovich.

Kaplan, D. (1968) "Quantifying In". *Synthese* 19, 178-214.

Kaplan, D. (1972) "What is Russell's Theory of Descriptions?" in D. Pears *Bertrand Russell: A Collection of Critical Essays*. Garden City, New York: Anchor Books, 227-244.

Kleene, S. (1952) *Introduction to Metamathematics*. New York: Van Nostrand.

Klement, K. (2002) *Frege and the Logic of Sense and Reference*. New York: Routledge.

Kripke, S. (1980) *Naming and Necessity*. Cambridge, Mass: Harvard University Press.

Lambert, K. & van Fraassen, B. (1967) "On Free Description Theory" *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 13, 225-240.

Lehmann, S. (1994) "Strict Fregean Free Logic". *Journal of Philosophical Logic* 23, 307-336.

Linsky, B. (2004) "Russell's Notes on Frege for Appendix A of *The Principles of Mathematics*". *Russell: The Journal of Bertrand Russell Studies*, n.s. 24, 133-72.

[MK] Montague, R. & D. Kalish (1957) "Remarks on Descriptions and Natural Deduction (Parts 1 and 2)". *Archiv für mathematische Logik und Grundlagenforschung* 8, 50-64, 65-73.

Morscher, E. & P. Simons (2001) "Free Logic: A Fifty-Year Past and Open Future" in E. Morscher & A. Hieke (eds.) *New Essays in Free Logic: In Honour of Karel Lambert* Dordrecht: Kluwer, 1-34.

Neale, S. (1990) *Descriptions*. Cambridge Mass: MIT Press.

Quine, W.V. (1950)*Mathematical Logic*. (Revised edition). New York: Harper Torchbooks.

Russell,B. (1905a) "On Denoting", *Mind* **14**: 479–493. Reprinted in *The Collected Papers of Bertrand Russell, Vol.IV: Foundations of Logic 1903-05* (Routledge: London 1994) A.Urquhart & A. Lewis (eds.), 414-427. Page references to this reprint.

Russell, B. (1905b) "Review of Meinong and Others, *Untersuchungen zur Gegenstandstheorie und Psychologie*", *Mind* 14, 530-538.

Scott, D. (1967) "Existence & Description in Formal Logic", in Schoenman, R. (ed.) *Bertrand Russell: Philosopher of the Century*. London: Allen and Unwin, 121-200.

Schoenfield, J. (1967) *Mathematical Logic*. Reading, MA: Addison-Wesley.

Strawson, P. (1950) "On Referring". *Mind* 59, 320-344.

Strawson, P. (1952) *Introduction to Logical Theory*. London: Methuen.

Suppes, P. (1960) *Axiomatic Set Theory*. New York: Van Nostrand.

Tichy, P. (1988) *The Foundations of Frege's Logic*. Berlin: Walter de Gruyter.

Whitehead, A. & B. Russell (1910)*Principia Mathematica*. Cambridge University Press: Cambridge. Second Edition 1925-1927. Reprinted in paperback as *Principia Mathematica to* $*56$. Cambridge: Cambridge University Press, 1962.