Philosophy of Language in the Twentieth Century

Jason Stanley
Rutgers University

In the Twentieth Century, Logic and Philosophy of Language are two of the few areas of philosophy in which philosophers made indisputable progress. For example, even now many of the foremost living ethicists present their theories as somewhat more explicit versions of the ideas of Kant, Mill, or Aristotle. In contrast, it would be patently absurd for a contemporary philosopher of language or logician to think of herself as working in the shadow of any figure who died before the Twentieth Century began. Advances in these disciplines make even the most unaccomplished of its practitioners vastly more sophisticated than Kant. There were previous periods in which the problems of language and logic were studied extensively (e.g. the medieval period). But from the perspective of the progress made in the last 120 years, previous work is at most a source of interesting data or occasional insight. All systematic theorizing about content that meets contemporary standards of rigor has been done subsequently.

The advances Philosophy of Language has made in the Twentieth Century are of course the result of the remarkable progress made in logic. Few other philosophical disciplines gained as much from the developments in logic as the Philosophy of Language. In the course of presenting the first formal system in the Begriffsscrift, Gottlob Frege developed a formal language. Subsequently, logicians provided rigorous semantics for formal languages, in order to define truth in a model, and thereby characterize logical consequence. Such rigor was required in order to enable logicians to carry out semantic proofs about formal systems in a formal system, thereby providing semantics with the same benefits as increased formalization had provided for other branches of mathematics. It was but a short step to treating natural languages as more complex versions of formal languages, and then applying to the study of natural language the techniques developed by logicians interested in proving semantic results about formal theories. Increased formalization has yielded dividends in the Philosophy of Language similar to those in mathematics. It has enabled philosophers to provide better and more fruitful definitions and distinctions.

Progress in Philosophy of Language and logic has positively affected neighboring disciplines such as metaphysics and meta-ethics. Because of this, some philosophers have thought that Philosophy of Language was some kind of "first philosophy", as Descartes viewed what we would now call "epistemology". But the fact that Philosophy of Language has progressed significantly does not mean that it provides us with a first philosophy. One can recognize that a discipline has advanced more than others without thinking that it holds the key to all advancement. The Twentieth Century was the century of "linguistic philosophy", not because all or even most philosophical problems have been resolved or dissolved by appeal to language, but because areas of philosophy that involved meaning and content became immeasurably more sophisticated.

My purpose in this chapter is to explain some of the key developments in the Philosophy of Language. Discussions of content in other fields, such as philosophy of mind or meta-ethics, are reflections of the distinctions drawn and categories developed in thinking about languages, both formal and natural.

I. Frege

It is difficult to write about the development of the Philosophy of Language in the Twentieth Century without reaching back to the latter parts of the Nineteenth, for the story of the revolution in logic and philosophy of language that occurred in the last century begins with the work of Gottlob Frege. Frege's project was not principally directed at language; it was rather primarily epistemological (see "The Birth of Analytic Philosophy"). Frege set out to show that the truths of arithmetic were analytic in nature, by deriving them from the axioms and definitions of logic. In order to carry out this project, Frege needed to show that the theorems of arithmetic could be derived from the theorems of logic without appeal to any synthetic (non-analytic) step. To show that his deductions achieved this goal, Frege devised a formal language for carrying out his proofs. The formal language allowed for the characterization of a set of precise syntactic transformations, each of which was an instance of a purely logical inference rule. Frege's concern with using natural language to carry out his proofs was that natural language was too vague and imprecise to allow the characterization of precise syntactic transformations that expressed instances of purely logical inference rules.

In the *Begriffsschrift*, Frege says remarkably little about how his formal language is to be interpreted. In contrast to the sophistication of the syntax, Frege's few remarks about content are typical of the pre-modern era. Indeed, it is easy to think of Frege's naïve conception of content as being principally about signs, rather than an extra-linguistic reality. First, Frege notoriously takes the identity relation to be a relation between signs (section 8). Secondly, Frege's later ontological distinction between function and argument is presented as a distinction between *expressions* (see Section 9). However, an expression may either be viewed as the function or the argument of a sentence, so that what later is an ontological distinction now reflects merely how we apprehend either the content or the presentation of the content (Ibid.). Frege does speak in rather contorted terms of *Begriffliche Inhalt*, but here too there is much confusion and obscurity. We are never told what an *Inhalt* [content] of any expression is, and Frege only hints at when two sentences have the same *Inhalt* (when they have the same "möglichen Folgerungen" [possible consequences]). Some contemporary philosophers (e.g. Brandom (1994, p. 94)) have tried to read back into Frege's confused remarks about *Begriffliche Inhalt* some controversial modern doctrine about inferential semantics. But Frege did not at this time have sophisticated thoughts about content; indeed, no settled doctrine about content that met Fregean standards of clarity and rigor was to emerge until the early 1890s.

It is instructive to reflect upon what led Frege to an essentially modern way of thinking about content. As Frege started to develop the logicist project, he adopted the Platonist position that arithmetic is about an independently existing domain of abstract objects, namely numbers, and rejected the formalist view that arithmetic is about signs. There are

two parts to the logicist task: deriving the theorems of arithmetic from logical principles, and showing that the concepts of arithmetic are logical concepts. Frege took the fact that numerical terms function as singular terms in arithmetic to be conclusive evidence that numbers are objects. His logicism thereby impels him to identify logical objects that are the numbers. The first part of the logicist project also requires that the syntactic transformations on expressions of the *Begriffsscrift* express inference rules that are indisputably logical. So Frege is led to the project of giving a rigorous interpretation to his formal system for two reasons. First, he is proving facts about numbers, not facts about signs. This position forces Frege to be more specific about the relation between signs and what they are about, since he denies the formalist view that arithmetic is simply about signs. Secondly, he needs to ensure that the syntactic transformations express transitions that are instances of genuinely logical inferences. This in turn forces him to develop a theory of content for his formal language.[1]

Frege's remarkable syntactic achievement in the Begriffsschrift of 1879 was to arrive at a notation that represented reasoning with quantifiers and variables (see "Birth of Analytic Philosophy"). Frege's remarkable semantic achievements occurred later. Frege (1966, Part I) provided a compositional semantics for the *Begriffsschrift* notation. In his seminal paper "On Sense and Reference" (Frege1993a), he also isolated a series of puzzles and topics that provided much of the groundwork for Twentieth century philosophy of language.

Here is not the place to delve into all the details of Frege's mature theory of content. But it is important to sketch it, because Frege laid down the elements upon which all subsequent investigations of content are predicated. Frege's ontology is divided into two kinds of entities, *objects* and *functions*. Though the distinction between object and function is a fundamentally ontological, Frege explains it by appeal to language. For Frege, an object is the kind of thing that is named by a "complete expression", or proper name, such as "Bill Clinton", and a function is the kind of thing that is named by an "incomplete expression", for example a predicate such as "is running" or a one-place functional expression, such as "x^2". Frege took the category of complete expressions or proper names to include sentences, and so he took the reference of sentences to be objects, in particular *truth-values*. Since Frege took one-place predicates to denote a kind of function, and sentences to denote truth-values, he treated the referents of one-place predicates as functions from objects to truth-values. Frege's called functions whose values are truth-values, *concepts*. So a predicate such as "is red" denotes a concept that takes an object to the truth-value True if and only if that object is red; otherwise, it takes that object to the truth-value False. The sentence "That apple is red" denotes the truth-value True if and only if the concept denoted by "is red" takes the object denoted by "that apple" to the truth-value True; otherwise, the sentence denotes the truth value False. Frege also provided an account of the semantics of quantifiers. According to this account (Frege (1966, section 20)), the expression for the universal quantifier denotes a "second-level function", one which takes a first-level function to the True if and only if the first-level function maps every object onto the True. So the occurrence of "everything" in "everything is red" denotes a function from concepts to truth-values. It takes a function, such as that denoted by "is red", to the true if and only the function denoted by "is red"

yields the True as value for every argument. Frege's account of quantifiers as second-level functions has proven to have lasting impact in natural language semantics, as it is a standard way quantifiers are treated in Montague Grammar, the dominant contemporary tradition in natural language semantics.

Frege's ontology gave him the resources to provide a particularly elegant characterization of the conditions under which a sentence of his formal language was true. Indeed, Frege uses it to provide just such a characterization in Part I of the Grundgesetze der Arithmetik. As Frege writes:

> Every name of a truth-value expresses a sense, a thought. Through our stipulations it is determined under what conditions it denotes the True.[2]

So, not only did Frege provide a number of technical suggestions that were to affect the development of semantics, but he also had a clear conception of the semantic project of giving a recursive characterization of the truth-conditions of sentences of a language, via assignment of semantic contents to the basic meaningful parts.

Frege's seminal paper "On Sense and Meaning" ("Über Sinn und Bedeutung") raises topics that are even more germane for the study of natural languages than they are for the study of formal languages. The paper is famous for the modern statement of *the problem of cognitive significance*; how can two expressions that denote the same object in the world (such as "Hesperus" and "Phosphorus", or "Cicero" and "Tully") nevertheless have differing cognitive significance? Why is it that "Hesperus is Phosphorus" and "Cicero is Tully" are cognitively significant, but "Hesperus is Hesperus" and "Cicero is Cicero" are not? Frege's solution involves the introduction of yet another element in his theory of meaning, the notion of *sense* [Sinn]. The sense of a term is (roughly) the way that term presents its referent. So "Hesperus" and "Phosphorus" both refer to the same object, namely Venus, but present this referent in different ways, and therefore have different senses.[3]

Frege uses the notion of sense to give an account of the meaning of *propositional attitude ascriptions*, which are sentences involving propositional attitude verbs such as "believes", "doubts", and "knows". A propositional attitude ascription, such as "John believes that Hesperus is a planet", appears to relate an agent to a thought (or proposition); in the case of this sentence, it appears to relate John to the thought that Hesperus is a planet. According to Frege, while the *referent* of a sentence is a truth-value, the *sense* of a sentence –the way it presents its referent—is a thought. Within the scope of a propositional attitude verb, an expression denotes, not its ordinary referent, but rather its ordinary sense. So propositional attitude verbs (such as "believes") create what are called *opaque contexts*, linguistic contexts in which substitution of co-referring expressions fails.[4] The claim that propositional attitude verbs create opaque contexts accords with our intuition that "John believes that Hesperus is a planet" may be true, whereas "John believes that Phosphorus is a planet" is false, even though "Hesperus" and "Phosphorus" refer to the same object, viz. the planet Venus. "Hesperus" and "Phosphorus" cannot be substituted for one another in the scope of a propositional

attitude verb, despite the fact that they have the same referent. Frege's account of the meaning of propositional attitude constructions explains this, because according to this account, "Hesperus" and "Phosphorus", within the scope of a propositional attitude verb, refer to their ordinary senses, rather than the object Venus. Therefore, within the scope of a propositional attitude verb, "Hesperus" and "Phosphorus" do not after all have the same referent.

Frege's reflections on natural language are not limited to propositional attitude ascriptions. "On Sense and Meaning" contains important and influential discussions of a number of other topics, including the topic that linguists now discuss under the rubric of *presupposition*. Frege's discussions of natural language reflect a great deal of insight about language, especially for someone whose primary interest was mathematics. For example, in addition to his contributions to the study of quantification, propositional attitude constructions, presupposition, Frege had important insights on the topics of plural reference and mass terms (in the course of discussing whether number was a property of objects in Frege (1980b)). Frege's late paper "The Thought" ("Der Gedanke") contains a remarkably lucid discussion of indexicals and demonstratives, expressions such as "I", "now", "today", "this" and "that". An indexical expression changes its referent from context to context; when Bill Clinton uses "I", it denotes a different object than when Hillary Clinton uses "I", despite the fact that the two uses of "I" have the same linguistic meaning. The context-sensitivity of indexicals raises certain difficulties for Frege's notion of sense and for characterizations of linguistic meaning generally, about which Frege showed clear awareness.[5] The sophistication of Frege's reflections about natural language was no doubt due to the fact that, despite his mistrust of its vagueness and context-sensitivity, Frege nevertheless took ordinary linguistic categories to reflect ontological ones.

II. Russell

In Frege's theory of meaning, each expression is associated with at least two semantic values, its ordinary referent [Bedeutung] and its ordinary sense [Sinn].[6] The semantic theory Frege provided for his formal language does not, however, involve the assignment of senses to any expressions. Partly, this is due to the absence of opacity-inducing expressions in his formal language, such as propositional attitude verbs.[7] The compositional semantic theory Frege provides in Part I of Frege (1893) does not explicitly involve assigning thoughts (the senses of sentences) directly to senses. In some sense, Frege seemed to think that giving the truth-conditions of *Begriffsschrift* sentences was enough to represent the thoughts they expressed.[8]

In England, a somewhat different conception of meaning was emerging from the pages of the journal *Mind*, one that was to have an equally large impact on subsequent thought about content. G.E. Moore (1899) argued that in judgment, we are related to contents which he called *propositions*. Moore's conception of propositions was not exactly contemporary; for example, he thought an each object was in fact an existential proposition (for more discussion, see "The Birth of Analytic Philosophy"). But the idea of judgment as expressing a relation between an agent and a complex of existing entities

that formed a distinctive kind of content, a proposition, was taken up by Bertrand Russell. In Russell's seminal (1905), he provided a rather modern characterization of propositions. On Frege's view, the contents of judgments were *Gedanken* [Thoughts], the senses of sentences, which were themselves composed out of ways of thinking of objects and properties. In contrast, Russell's propositions contained actual objects and properties. As Russell wrote to Frege in his famous letter of December 12, 1904 (Frege (1981, p. 98):

> Ich glaube dass der Mont Blanc selbst, trotz aller seiner Schneefelder, Bestandtheil desses ist was eigentlich behauptet wird im Satze "Der Mont Blanc ist mehr also 4000 meter hoch." Man behauptet nicht den Gedanken, der ja psychologische Privatsache ist: man behauptet das Objekt des Gedankens, und dies ist meines Erachtens ein gewisser Complex (ein objektiver Satz, koennte man sagen) worin der Mont Blanc selber ein Bestandtheil ist.
>
> [I believe that Mont Blanc itself, despite all of its snowfields, is a constituent of what would be asserted by the sentence 'Mont Blanc is higher than 4000 meters'. One does not assert the thought, which is a psychologically private matter: one asserts the object of the thought, and this is according to my conception a certain complex (an objective sentence, one might say) of which Mont Blanc itself is a constituent.] [9]

Russell's motivation for developing his theory of propositions was also distinct from Frege's. Russell, like Frege, thought that mathematics posed certain epistemological problems that could be solved by resting it upon a logical foundation. But the epistemological problems Russell thought were posed by mathematics were slightly different than the ones that exercised Frege.

Russell's theory of meaning emerged from his desire to account for our ability (most obvious in the mathematical domain) to think about an infinite class of objects, despite our inability to survey an infinite domain. According to Russell's 1903 theory (Russell (1996, Chapter 5)), what accounts for our ability to grasp propositions that are about an infinite class of objects is the fact that such propositions contain *denoting concepts*. As Russell writes (1996, p. 53):

> A concept *denotes* when, if it occurs in a proposition, the proposition is not *about* the concept, but about a term connected in a certain peculiar way with the concept. If I say 'I met a man', the proposition is not about *a man*; this is a concept which does not walk the streets, but lives in the shadowy limbo of the logic-books. What I met was a thing, not a concept, an actual man with a tailor and a bank-account or a public-house and a drunken wife.

When we grasp a proposition that is about an infinite domain of objects, this is because the proposition contains a denoting concept that is about that infinite domain of objects (Russell (1996, p. 73); see also "The Birth of Analytic Philosophy"). A finite mind can grasp a concept that denotes an infinite class, but not the infinite class itself. Russell

developed his theory of meaning to explain the essentially epistemological problem, so evident in mathematical thought about unsurveyable domains, of how we could grasp the proposition expressed by an occurrence of a sentence, despite the fact that we did not have the ability to grasp all the things the proposition was about.

The story of Russell's dissatisfaction with his 1903 theory of denoting, and his 1905 development of the theory of descriptions is told at some length elsewhere in this volume (see sections 2.3 and 2.4 of "The Birth of Analytic Philosophy"); I direct the reader's attention to those pages. Here, I shall just briefly summarize some of the main differences between Russell's 1905 theory of meaning and its subsequent development and Frege's mature theory of meaning, differences that will be important in our discussion of subsequent developments.

For Russell, sentences express propositions, which are the ultimate objects of truth and falsity. Propositions are non-linguistic entities that contain as constituents objects and properties. Grasping a proposition requires bearing a privileged epistemological relation to each of its constituents (which relation, as we have seen, Russell does not think we can bear to an infinite class).[10] A *logically proper name* is an expression that contributes the object to which it refers to the proposition expressed by a sentence containing it. Thus, if "Jason Stanley" is a logically proper name, grasp of the proposition expressed by "Jason Stanley is a philosopher" would require bearing this epistemologically special relationship to the object Jason Stanley. Following Russell, we shall use the term *acquaintance* for the epistemologically special relationship one must have to the constituents of a proposition in order to grasp that proposition. Russell changed his mind throughout his career about what is required to have acquaintance with an object; soon after 1905 he came to the view that the only objects one could be acquainted with are sense-data and perhaps oneself.[11] Since we clearly do grasp many propositions that are not about objects that Russell thought we had acquaintance with (e.g. objects in the distant past, or people we have never met), Russell did not think that most ordinary proper names were logically proper names. For example, we clearly do grasp the proposition expressed by an occurrence of "Bismarck was a clever man", though we do not (according to Russell) have acquaintance with Bismarck. Therefore, the proposition expressed by this occurrence of "Bismarck was a clever man" does not contain Bismarck as a constituent (else we would not grasp it). Thus, "Bismarck", the ordinary proper name, is not a *logically* proper name.

According to Russell, most ordinary proper names were *disguised definite descriptions*. Since (according to Russell's theory of descriptions, see 2.3 of "Birth of Analytic Philosophy") definite descriptions contribute only universals (i.e. properties) to the propositions expressed by sentences containing, and Russell was fairly liberal about acquaintance with universals, the propositions expressed by sentences containing ordinary proper names are capable of being grasped by ordinary people. For example, the ordinary proper name "Bill Clinton" would be, for Russell, a disguised definite description, perhaps the definite description "the President of the United States of America between 1992 and 2000" (though the proper name "the United States of America" is also presumably a disguised definite description). The proposition expressed

by “Bill Clinton is a Democrat” would then contain as constituents only universals. Hence, it could be grasped by someone with no acquaintance with Bill Clinton.

Frege’s “thoughts”, or propositions, were composed of ways of thinking of objects and properties; the thought that Jason Stanley is a philosopher, on this view, consists of a way of thinking of Jason Stanley and a way of thinking of the property of being a philosopher. Russell’s 1905 theory of meaning differs from Frege’s in that it involves no notion of sense, no “ways of thinking” about things. Russell’s propositions are composed of objects and properties (universals), not ways of thinking of them. Russell also had epistemological motivations for certain of his views that are absent in Frege. Russell thought we were not acquainted with many objects, but nevertheless could grasp propositions that seemed to be about them. So, for epistemological reasons, Russell took ordinary proper names to be disguised definite descriptions, and analyzed definite descriptions away using the apparatus of quantificational logic (see 2.3 of “Birth of Analytic Philosophy”).

Though Russell’s motivation was primarily epistemological, his description theory of ordinary proper names, coupled with his account of the semantics of sentences containing definite descriptions, also allowed him to resolve certain philosophical puzzles. According to Russell, a sentence of the form “The F is G” expressed a proposition whose logical form was more complex than the grammatical form of “The F is G”. In particular, “The F is G” expresses the proposition that there exists an x which is F, there is only one F, and x is G.[12] So a sentence containing a definite description expresses an existentially quantified proposition, together with a *uniqueness* condition, to the effect that the nominal complement of “the F” (which is the instance of ‘F’) is satisfied by one and only one object.[13]

Russell applied his description theory of ordinary proper names, together with his semantic account of sentences containing definite descriptions, to a number of problems. Most famously, Russell applied his views to *the problem of negative existentials*. The problem of negative existentials is raised by sentences such as “Pegasus does not exist”, which clearly express truths, despite the fact that they contain non-referring terms (in this case, “Pegasus”). If “Pegasus does not exist” expresses a true proposition, then “Pegasus” must refer to something that lacks the property of existence. But if “Pegasus” refers to something that does not exist, then there are things that do not exist. This argument in favor of a realm of shadowy non-existents is the problem of negative existentials.

To solve the problem of negative existentials, Russell first applied his description theory of ordinary proper names to conclude that “Pegasus does not exist” expresses the same proposition as (say) “The winged horse of mythology does not exist”. The proposition expressed by “The winged horse of mythology does not exist”, according to Russell’s theory of descriptions, has a true reading, one in which the denoting phrase “the winged horse of mythology” has a *secondary occurrence* (takes narrow-scope) with respect to the negation “not”. According to this reading, the sentence expresses the proposition that it is not the case that there exists an x, x is a winged horse of mythology, and for all y, if y is a winged horse of mythology, then y=x, and x exists. This proposition is clearly true, and

its truth does not commit us to a mysterious ontology of non-existent things. Russell's dissolution of the problem of negative existentials is the paradigm example of using linguistic analysis to resolve metaphysical quandaries.

Russell also repeatedly applies his theory of descriptions to the problem of cognitive significance. The reason that "Scott is the author of Waverly" is cognitively significant, while "Scott is Scott" is not is that "Scott is the author of Waverly" expresses the proposition that there exists an author of Waverly, and only one author of Waverly, and he is Scott, whereas "Scott is Scott" (taking "Scott" to be a logically proper name) expresses a trivial proposition of the form a=a (Russell & Whitehead, 1910, Chapter 3 of Introduction). Using his description theory of ordinary proper names, Russell also can explain why a sentence containing an "is" of identity between two distinct ordinary proper names (such as "Hesperus is Phosphorus" or "Cicero is Tully") is cognitively significant, whereas a sentence containing an "is" of identity between two occurrences of the same name (such as "Hesperus is Hesperus") is not cognitively significant. For distinct ordinary proper names are treated, by Russell, as standing in for distinct definite descriptions. So, "Hesperus is Phosphorus" expresses the same proposition as "The morning star is the evening star", or perhaps "The planet called 'Hesperus' is the planet called 'Phosphorus'" (as in Chapter 16 of Russell (1919)). So, where Frege appealed to distinct senses associated with distinct proper names to resolve the problem of cognitive significance, Russell maintained that distinct ordinary proper names corresponded to distinct definite descriptions.

III. From Frege and Russell to Tarski

Frege, Moore, and Russell had several doctrines in common that have since become widely accepted, yet were not clearly adopted or even understood in previous philosophical work.[14] First, all three philosophers clearly distinguished the *act* of judging from the *object* of the judgment (and similarly the *relation* of believing from the contents particular beliefs). Secondly, all three philosophers thought of the object of judgment as being a complex, mind-independent that was the object of knowledge and belief (though, as we have seen, they differed amongst one another as to the nature of this mind-independent entity). Third, all three philosophers thought of these entities as the primary bearers of truth and falsity.

Frege, Moore, and Russell were not the only philosophers at the time to clearly make these distinctions and adopt these views. For example, Alexius Meinong clearly thought of the objects of judgment as complex, mind independent entities that we are related to in knowledge and belief, and provided sophisticated arguments for this conclusion (see Chapter 3 of Meinong (1910)). Nor were all three of these founding fathers of analytic philosophy terribly consistent in retaining these positions. Moore and Russell abandoned their belief in the existence of propositions soon after they developed them, because of the concern that positing false propositions was ontologically profligate (see e.g. Russell (1994)). Indeed, after 1910, Russell used "proposition" as a way to talk about sentences, and went so far as to abandon the act-object conception of judgment, with the *multiple-relation theory of judgment* (see "The Birth of Analytic Philosophy", section 2.7).

Nevertheless, the clarity and cogency of these views withstood even their abandonment by some of their chief proponents, and survive as presumptions of virtually all contemporary discussions of content.

The twenty years of philosophy that followed Frege and Russell's greatest accomplishments were relatively unimportant to subsequent work in the philosophy of language. Russell spent the years between 1910 and 1920 developing an idiosyncratic version of phenomenalism, according to which ordinary objects were "logical fictions", and names for them were to be treated as "incomplete symbols" to be analyzed away, so that we are left just with reference to sense-data and universals (see section 2.5 of "The Birth of Analytic Philosophy"). Ludwig Wittgenstein's Tractatus Logico-Philosophicus was also devoted to rather large scale metaphysical endeavors, and was not written with the level of clarity that is so characteristic of the writings of Frege and early Russell. However, unlike Frege and Russell, Wittgenstein took seriously the modal notions of *possibility* and *necessity*.[15] For Wittgenstein, a meaningful content divided the space of possibilities. For a proposition to be meaningful, both it and its negation had to be possible; otherwise the proposition did not *divide* the space of possibilities into those in which the proposition is true and those in which it was false.

The Tractatus impeded progress in philosophy, because it led philosophers (in particular the Logical Positivists) to expend their energies in the pursuit of developing and honing a criterion of meaningfulness, and using the criterion to argue that traditional philosophical theses failed to satisfy it, and were hence meaningless (see "Wittgenstein and After"). This project has been moribund for decades. It is nevertheless the quest of a criterion of meaningfulness that can be put to anti-metaphysical use that the lay intellectual public unfortunately most clearly associates with analytic philosophy.

However, the influence of the Tractatus has not been uniformly negative. As we shall see, other philosophers took up some of the metaphysical apparatus developed in the Tractatus and applied it to the study of content. As we shall see, this research program has turned out to be extraordinarily fruitful, not just in subsequent investigations in the philosophy of language, but also in metaphysics. So Wittgenstein's belief that modality and meaningfulness were intimately related has, somewhat ironically, fueled something of a revolution in just the kind of philosophy he wanted to use it to undermine.

IV. Tarski's Theory of Truth

As we have seen, in Part I of Frege (1966), Frege provides a semantic theory for his formal language; a set of "Festsetzungen" [stipulations] that determines under what conditions an arbitrary sentence of the Begriffsschrift is true. Frege also provides semantic proofs about the formal theory of the Grundgesetze, including, rather notoriously, an attempted semantic consistency proof.[16] As Saul Kripke has recently emphasized, semantical proofs also occur in Russell's Principia.[17] Furthermore, Frege proves some model-theoretic results about his formal system of arithmetic within the naïve set-theoretical framework of the Grundgesetze; for example, it is plausible to take the proof of theorem 263 to be a categoricity theorem for his axioms of arithmetic (see

Heck (1993, section 7)). However, Frege and Russell's concern was ultimately to place mathematics on the secure foundations of logic, and they appealed to semantics chiefly in the service of this project. In contrast, the focus of the Polish logician Alfred Tarski (born Alfred Teitelbaum) was on the discipline of semantics itself. Tarski set himself the task of setting *semantics* on the secure foundations of mathematics, by providing mathematical definitions of semantical concepts such as *truth* and *logical consequence*.

Tarski's motivation for setting semantical concepts on secure foundations was distinct from Frege and Russell's motivations for logicism, though related to the reasons for its failure. The logicist program of reducing mathematics to logic was undermined by the fact that the systems that were powerful enough to provide a foundation for mathematics were not plausibly regarded as logical. First, Frege's system of naïve set-theory turned out to be inconsistent, as Russell's paradox demonstrated. Secondly, Russell's system involved axioms that were too controversial to be regarded as logical (see "The Birth of Analytic Philosophy"). But the fate of naïve set-theory was to have repercussions in many areas of mathematics. In particular, it focused attention on the fact that intuitive principles governing a fundamental concept (such as that of an aggregation of objects) could lead to paradox, and that the paradox could be evaded by greater mathematical subtlety.

As in the case of naïve principles governing the aggregation of objects, some of the most obvious principles governing semantical concepts lead quickly to paradoxes. For example, restricting ourselves only to sentences that contain no context-sensitive vocabulary, the following claim seems to be an obvious truism, one that follows from the meaning of the word "true":

(1) 'S' is a true sentence if and only if S.

To illustrate this point (using *quotation-names* as names of sentences) consider the obviousness of the following:

(2) "Snow is white" is a true sentence if and only if snow is white.
(3) "Grass is green" is a true sentence if and only if grass is green.

Few claims are as uncontroversial as (2) and (3). Yet schema (1) seems to lead fairly directly to a contradiction. Consider the following:

(4) (4) is not a true sentence.

Sentence (4) contains no context-sensitive vocabulary. So it should be unproblematic to place it in for "S" in schema (1). But if we plug "(4) is not a true sentence" in for "S" in the right-hand side of (1), together with a name of it on the left-hand side, we obtain:

(5) (4) is a true sentence if and only if (4) is not a true sentence

Since (5) is a contradiction, the intuitive principle about truth that the schema in (1) exemplifies is false.

So much the worse for truth, one might think. After all, the concept of truth is one that seems to belong to metaphysics, which is not the most reputable of disciplines. However, by 1930, it had become clear that the use of semantical concepts such as truth and logical consequence was of genuine mathematical use in describing desirable properties of formal systems. For example, one desirable property of a formal theory of a given subject matter is *completeness*, which is the question of whether that formal system is adequate to proving every sentence that is true in virtue of the subject matter in question. Another desirable property is satisfiability; is it possible for the axioms of the formal system to all express truths, or do some axioms result in contradictions (it is this that Frege was trying to demonstrate in sections 29-32 of the Grundgesetze). There are a number of other semantical properties of formal systems that are defined in terms of semantical notions, and much work in the foundations of mathematics in the 1920s was in the service of proving semantical claims about formal systems. So as the century progressed, it became clear the semantical concepts were not just the philosopher's concern, but the mathematician's as well.[18]

In his landmark paper, "The Concept of Truth in Formalized Languages", Tarski set out to show that, for many languages, one could consistently define a truth-predicate for that language, though this definition must be given in a language that is expressly richer than the original language. Moreover, given the right theoretical resources, one can derive all the instances of (1), for sentences of the original language, from this definition. Thus, Tarski shows how to define the relative concept of Truth-in-L, for a specific language L.

Tarski's method of defining truth for the language L involves inductively defining the notion of *a sequence satisfying an open formula of L*, and truth is defined in terms of thus defined notion of satisfaction (a true sentence is one that is satisfied by all sequences).[19] In his original paper, Tarski focused on defining what he called "the absolute concept of truth" (Tarski, 1983b, p. 199). But the more important notion is the notion of truth for a language L *relative to a model*, of which Tarski's absolute notion is a special case (Ibid.). A model is, intuitively, an interpretation of the language, relative to a domain of objects – the "universe" of that model. The reason to define the more general notion of truth for a language relative to a model (and not just truth for a language) is that the notion of truth in a model is what is required to capture fundamental semantic notions such as *logical validity* and *logical consequence*.[20] A sentence S is a *logical consequence* of sentences $\alpha_1 \ldots \alpha_n$ if and only if S is true in every model M in which $\alpha_1 \ldots \alpha_n$ are true; S is *logically valid* if and only if S is true in every model (logical validity is the limiting case of logical consequence; a logically valid sentence is a logical consequence of the empty set of sentences). The motivation behind these definitions is that a sentence S is logically valid if and only if it is true no matter how one interprets the non-logical vocabulary in S and no matter what objects there are in the domain. Models therefore serve the dual function of providing alternative interpretations of the non-logical vocabulary in the sentence, and varying the objects quantified over by the quantifiers. With the model-theoretic definition of logical consequence, one can give mathematical perspicuity to some of the

fundamental notions in e.g. the completeness theorem for first-order logic. Since the ultimate goal was to give a mathematical characterization of the fundamental semantic notions of validity and consequence, defining truth in a model, rather than truth, should be the desired goal of the meta-mathematician seeking to legitimize the semantical notions most useful for logic.

To illustrate the Tarskian method of defining truth for a language relative to a model, it is instructive to look at a simple example in detail. In what follows, I will define a simple language L, and show how, with the use of Tarski's notion of satisfaction, to give a definition of truth for that language (readers who wish to avoid these details may skip the next few pages without loss).

The Language L

Alphabet of L:

A, B, ...,E	Name Letters (Constants)
F^n, G^n,..., Z^n	n-place Predicate Letters
P, Q, ...,Z	Sentence Letters
a, b, c, ..., w, x, y, z	Variables
~, $\rightarrow$, $\leftrightarrow$, v, &	Sentential Connectives
$\forall$, $\exists$	Quantifiers (Universal, Existential)

Grammar of L:

Termhood of L:

(i) All name letters and variables are terms.
(ii) Nothing else is a term.

Well-formed Formula (wff) of L:

(i) 0-place predicate letters are wffs.
(ii) $\varphi\alpha_1...\alpha_n$ is a wff if φ is an n-place predicate letter, and each of $\alpha_1....\alpha_n$ is a term.
(iii) $\sim\varphi$ is a wff if φ is a wff.
(iv) $(\varphi \rightarrow \psi)$ is a wff if φ is a wff and ψ is a wff.
(v) $(\varphi \leftrightarrow \psi)$ is a wff if φ is a wff and ψ is a wff.
(vi) $(\varphi \vee \psi)$ is a wff if φ is a wff and ψ is a wff.
(vii) $(\varphi \& \psi)$ is a wff if φ is a wff and ψ is a wff.
(viii) $\forall\alpha\varphi$ is a wff if φ is a wff and α is a variable.
(ix) $\exists\alpha\varphi$ is a wff if φ is a wff and α is a variable.
(x) Nothing else is a wff.

So, some well-formed formulae of L are:

R^2xy
$\exists xH^2Ayx$
$\exists x((F^1x \& G^0) \rightarrow \forall zJ^3xzA)$

We shall define *truth relative to a model*, where a model consists of a *domain of discourse* (intuitively, the things spoken of, or quantified over) and an assignment of values to (an interpretation of) the non-logical expressions (the name letters and predicate letters). More formally, a model M for L consists of an ordered pair of sets, <D, $\mathfrak{I}$>. D is a set of objects, called the *universe*, or *domain*, of M, and $\mathfrak{I}$ is a function which (i) assigns to each name letter of L a member of D, (ii) assigns to each 0-place predicate letter a truth-value (either the true or the false), and (iii) assigns to each n-place predicate letter (n>0), a set of n-tuples of members of D. So, $\mathfrak{I}$ 'interprets' the non-logical constants (name letters, sentence letters, and predicate letters) of L.

Notice that we have yet to give a method for interpreting variables. For that, we will need Tarski's notion of *satisfaction*, which enters in below, in the definition of truth-in-a-model. The ultimate goal is to characterize what it is for an arbitrary sentence (a sentence is a well formed formulae with no free variables) to be true in a model. For this, we appeal to the notion of satisfaction. A sentence is true in a model if and only if it is satisfied by all sequences of that model. With the use of standard notation:

$\models_M \varphi$ read: φ is true in (model) M.

$\models_{M,s} \varphi$ read: φ is *satisfied* by (sequence) s in M.

Df.: $\models_M \varphi$ iff for all sequences s of M, $\models_{M,s} \varphi$.

A sequence s of a model M is a function which assigns, to each variable of the language of L, a member of D. In other words, sequences assign values to variables, and the values they assign are members of the domain of discourse of the model M. For every model M, for each member o of the universe of M, and for each variable x of L, there is a sequence of M which assigns o to x (this ensures that we have all the sequences we need).

Since a sentence is true in a model M if and only if it is satisfied by all sequences of that model M, we have now reduced the problem of defining truth in a model to that of defining satisfaction in a model of a sentence by a sequence.[21] We now turn to the inductive definition of satisfaction in a model. For this we will need two additional definitions. First, we will need the concept of *denotation relative to a sequence*:

Df. : Where t is a term, and s a sequence, Den(t,s) = $\mathfrak{I}$(t) if t is a name letter, and Den(t,s) = s(t) if t is a variable.

The denotation function is defined with the use of the interpretation function $\mathfrak{I}$ of the model, and is used to interpret the terms. We will also require the concept of an *s' variant of a sequence s*, which will help us in giving the interpretation clauses for the quantifiers (clauses (viii) and (ix) below):

Df. : $s' \approx_x s$ read: s' is identical to s except at most in assigning something different to the variable 'x' than s does (so, for every

variable $y \neq x$ of L, $s'(y) = s(y)$, and possibly, $s'(x) \neq s(x)$).

With the use of these notions, we may now turn to the inductive definition of satisfaction (for "sats" read "satisfies", and I have suppressed, for convenience's sake, the reference to the model M):

(i) If φ is a 0-place predicate letter, s sats φ iff $\Im(\varphi)$ = the true.
(ii) If φ is an n-place predicate letter, and $\alpha_1 ... \alpha_n$ are terms, s sats $\varphi\alpha_1 ... \alpha_n$ iff $<Den(\alpha_1,s),...,Den(\alpha_n,s)>$ is in $\Im(\varphi)$.
(iii) If φ is of the form '$\sim\psi$', then s sats φ iff s does not sat ψ.
(iv) If φ is of the form '$(\psi \rightarrow \chi)$', then s sats φ iff either s does not sat ψ or s sats χ.
(v) clause for '&'
(vi) clause for 'v' [clauses for '&', 'v', and '$\leftrightarrow$' are left to the reader]
(vii) clause for '$\leftrightarrow$'
(viii) If φ is of the form '$\exists x\psi$', then s sats φ iff for some $s' \approx_x s$, s' sats ψ.
(ix) If φ is of the form '$\forall x\psi$', then s sats φ iff for every $s' \approx_x s$, s' sats ψ.

With the use of these definitions, one may derive theorems that give the conditions under which an arbitrary sentence S of the language L is satisfied by a sequence in a model. For example, one can prove that:

(6) A sequence s of a model M satisfies "$\exists x(F^1x \,\&\, G^1x)$" if and only if for some sequence $s' \approx_x s$, $s'(x)$ is in $\Im(F^1)$ and in $\Im(G^1)$.

The sentence "$\exists x(F^1x \,\&\, G^1x)$" is a sentence of the object-language. The meta-language is the language in which the truth-conditions of this sentence are given. For example, the expressions occurring on the right-hand side of "if and only if" are in the meta-language. The meta-language is a combination of English together with some set-theoretical and logical vocabulary. Intuitively, what (6) says is that "$\exists x(F^1x \,\&\, G^1x)$" is satisfied by a sequence s of a model M if and only if there is something in the domain of discourse of M that falls within the extension of both the predicate F and the predicate G. Since "$\exists x(F^1x \,\&\, G^1x)$" contains no free variables, if one sequence satisfies it, then all sequences will satisfy it. So "$\exists x(F^1x \,\&\, G^1x)$" is true a model M if and only if it is satisfied by at least one sequence of M, and (6) states the conditions under which "$\exists x(F^1x \,\&\, G^1x)$" is satisfied by an arbitrary sequence of M. So, with the use of the inductive definition of satisfaction, together with the definition of truth in a model in terms of satisfaction, we can derive the conditions under which an arbitrary sentence L is true in a model M.

So, a definition of truth for the language principally takes the form of an inductive assignment of satisfaction conditions to sentences of that language, couched in an appropriate meta-theory. Finally, Tarski had a condition of adequacy for a definition of truth for a language. A definition of truth for a language was *materially adequate* if and only if the definition has, as consequences, all instances of the following schema (where

'S' is replaced by structural-descriptive names of sentences of the language, and 'p' is replaced by meta-language translations of the sentence named):

S is true if and only if p

So, a definition of truth is materially adequate if and only if it produces, as theorems, for each sentence S of the language under consideration, a statement in the meta-language of the truth-conditions of S (if the meta-language contains the object-language, this statement can simply be the object-language sentence, as in (1) above). For example, I pointed out above that the definition of truth I provided for the language L produces (6) as a theorem, which adequately gives (on the right hand side of "if and only if") the satisfaction conditions for "$\exists x(F^1x \ \& \ G^1x)$" in the meta-language (which is English plus some set-theory and logic). The above definition of truth for the language L is adequate according to convention T if and only if it produces theorems where the right-hand side of "if and only if" is a genuine translation into the meta-language of the intended interpretation of L. Since Tarski's condition of adequacy contains reference to the notion of translation, it is often said that Tarski defined truth by assuming translation.

So, Tarski showed how to define truth for a language, in an expressively richer meta-language. Tarski also provides a famous negative result. His negative result concerns the impossibility of defining truth for a language that is sufficiently "rich" as to allow for "all concepts and all grammatical forms of the metalanguage" (Tarski (1983b, 254)) to be interpreted in that language. Tarski's particular example of a language that has this character is what he calls the "general theory of classes", which contains variables ranging over entities of any order (e.g. classes, classes of classes, etc.). Tarski proves that, on pain of contradiction, one cannot define a one-place predicate of that language that is true of all and only the true sentences of the general theory of classes.[22] The negative result places a limit on the positive results of the paper.

Tarski shared Frege's suspicion of natural language, but had additional reasons for so doing. Tarski thought that natural language shared a feature with languages of powerful theories, such as the general theory of the calculus of classes. Both languages have a "universal character" that allows for the formulation of the "structural-descriptive concepts" of the meta-language within them. It is this "universal character" that allows for self-reference, and thereby leads to the formulation of the paradox, precluding the possibility of providing a consistent definition of truth. In other words, there is no meta-language for a natural language such as English or German the resources of which cannot be appropriated within English or German (since natural languages are "universal"). Since there is no expressively richer language than such languages, there is no possibility of consistently defining truth for these languages. Since Tarski thought natural language, unlike the language of the general theory of classes, did contain its own truth-predicate, he rather puzzlingly (see Putnam (1975a, p. 73)) declared natural languages to be *inconsistent*.[23]

Tarski's theory of truth has come in for some serious criticism as a contribution to an understanding of the nature of truth (as opposed to a piece of meta-mathematics). First,

Tarski seemed to think he had "reduced" the semantical concept of truth to non-semantic concepts. In the opening remarks of "The Concept of Truth in Formalized Languages", he famously declares "I shall not make use of any semantical concept if I am not able previously to reduce it to other concepts" (Tarski (1983b, p. 153)). In his paper "The Establishment of Scientific Semantics", Tarski writes that in an adequate definition of truth, "the semantical concepts are defined in terms of the usual concepts of the metalanguage and are thus reduced to purely logical concepts, the concepts of the language being investigated, and the specific concepts of the morphology of language." (Tarski, 1983a, p. 406). However, Tarski either failed to recognize, or was ignoring for rhetorical purposes, the fact that a definition of truth appeals to primitive *interpretive* semantic notions. A definition of truth presupposes an assignment of semantic values to primitive expressions of the language. As one can see from clause (ii) of the inductive definition of satisfaction, and the definition of the denotation relation, the interpretation function $\mathfrak{I}$ of the model M is what interprets the predicates and name letters of the language. But the characterization of the interpretation function $\mathfrak{I}$ does not follow from some general account of denotation. It is simply provided as a mapping from expressions to values. This is not a reduction of denotation to non-semantical notions, or indeed an explanation of denotation at all (Field, 1972).

Perhaps Tarski wished to maintain that $\mathfrak{I}$ is not a presupposed list of expressions and their semantic values (and so masks appeal to a primitive notion in need of a theoretical explanation), but simply the product of a stipulative mathematical definition. But then the theorems of Tarskian truth-definitions for languages (or fragments of languages) would be necessary truths, since they would follow from stipulative definitions and logic. However, intuitively an instance of Tarski's schema T such as (7) is not a necessary truth at all:

(7) "Bill Clinton is smart" is true if and only if Bill Clinton is smart.

(7) is not a necessary truth, because the sentence "Bill Clinton is smart" could have meant something other than it does. For example, "is smart" could have expressed the property of being from Mars, in which case (7) would be false. (7) is therefore a contingent truth, rather than a necessary truth. So if Tarski intends $\mathfrak{I}$ to be the result of a stipulative mathematical definition, then what his truth-definition produces will be necessary truths of mathematics, not contingent truths of semantics. In short, if Tarski's purpose was to reduce semantical concepts to non-semantical ones, he certainly did not succeed.

Tarski's purpose was not to use his theory of truth to give an account of the meaning of natural language sentences. It was rather to place semantics, construed as a branch of meta-mathematics, on a scientific grounding by showing that one could give a consistent definition of truth for a language of a theory, and perhaps (unsuccessfully) reduce the semantical concepts to non-semantical ones (either mathematical or physical). Within philosophy, Tarski's work is notable for the lively industry on the liar paradox to which it has given birth (see section 3 of "Philosophical Logic"). But Tarski's work has also had much broader influence. For as a number of philosophers recognized (Davidson, 1967), a Tarskian definition of truth appears to give us a tractable form for a theory of meaning. If

instead of defining truth by appeal to translation, one takes truth as a primitive notion in the system, a recursion on truth provides a statement of the truth-conditions of sentences in the language in question (as we have seen, this much actually seems to have been recognized by Frege).[24] As subsequent decades showed, the idea that the proper form of a theory of meaning for a natural language is a recursive characterization of the conditions under which a sentence is true has been extraordinarily fruitful, perhaps the most fruitful insight in the long history of the study of meaning. In other words, perhaps unintentionally, Tarski discovered the proper *form* of a theory of meaning. Tarski's work, in addition to being a contribution to meta-mathematics, indeed gave birth to a science of semantics, but on a very different understanding of that science than Tarski intended.

V. Necessity and Analyticity in Carnap and Quine

Tarski's work on truth was very much taken in the vein in which it was intended, as rehabilitating the scientific respectability of the semantic notions. Rudolf Carnap, one of the principle members of the Vienna school of logical positivists, was particularly influenced by Tarski in this regard. In his early work, Carnap had shunned semantics (or "semasiology" as Carnap called it), and "intensional logic" in particular. As Carnap (1949, p. 259) writes:

> All questions in the field of logic can be formally expressed and are then resolved into syntactical questions. *A special logic of meaning is superfluous*; 'non-formal logic' is a *contradiction in adjecto*. *Logic is syntax*.

However, Tarski's work convinced him that semantics was worthy of scientific study. In his work Meaning and Necessity (Carnap, 1958, first published 1947), Carnap turned to the project of using semantics in the service of advancing his positivist program.

Carnap sought to show how one could, with the use of semantical rules, set up *linguistic frameworks*. According to Carnap, the linguistic framework one decides to employ is not a factual question; it is simply a question of how to talk. The decision to adopt a linguistic framework is "a practical, not a theoretical question…The acceptance cannot be judged as being either true or false because it is not an assertion. It can only be judged as being more or less expedient, fruitful, conducive to the aim for which the language is intended." (Carnap, 1958a, p. 214). Once one has decided upon a linguistic framework, then a number of factual questions may be formulated with the use of that linguistic framework. Philosophical (and particular metaphysical) disputes arise because people confuse questions about which framework to adopt (what Carnap called *external questions*) with questions that arise within the framework, either of a factual or an analytic nature (what Carnap called *internal questions*). The metaphysical question of whether there are properties or universals is either the external question of whether to accept a linguistic framework that assigns properties as semantic values of predicates, or is the internal question of whether there are properties, which has only a trivial answer. In a linguistic framework in which properties are assigned to predicates, it is analytically true that there are properties. So many philosophical questions are either pseudo-questions (that is, without cognitive content), or have only trivial analytically true or false answers.

Carnap's concern with necessity comes from his desire to make a distinction between two kinds of internal questions: questions whose answers are analytically true (true in virtue of the semantical stipulations of the framework), on the one hand, and questions that are "factual" in nature, on the other (that is, questions whose answer is not determined by the semantical rules of the language). So, Carnap's purpose in Meaning and Necessity is to continue central features of the program of logical positivism. Nevertheless, within this work, Carnap presented the sort of intensional semantic account of an idealized language (one that nevertheless represents a fragment of natural language in crucial respects) that is recognizably contemporary in character. As a result, Carnap's work was to have resounding influence in the decades to come, long after its central philosophical task had been abandoned as hopeless.

As we have seen, Frege and Russell contributed much of importance and interest to our understanding of the semantics of non-extensional contexts, such as the linguistic contexts created by propositional attitude verbs like "believes" and "doubts". But in their formal semantical work, they focused on formal languages lacking expressions that created non-extensional contexts. Furthermore, though Frege and Russell attempted to address the thorny problem of propositional attitude contexts, they did not take the modal notions of necessity and possibility seriously. Carnap, in contrast, provided both a semantic theory for sentences containing modals expressions such as "necessary" and "possible", as well as propositional attitude verbs.

To treat the problem of giving a semantic theory adequate to giving the truth-conditions of sentences containing "necessary" and "possible", Carnap introduced the notion of a *state-description*, which is intended to be a representation of the metaphysical notion of a *possible world*, or *way in which the world could be* (Carnap, 1958, pp. 9-10). A state-description is a set of sentences that is supposed to give a complete description of a possible state of the universe. A sentence S is necessarily true if and only if that sentence is true in every state description, that is, true in every possible world. However, given the philosophical project discussed above, Carnap viewed necessity as *analyticity*, namely truth in virtue of the semantical rules of the linguistic framework (or, more briefly, *truth in virtue of meaning*). Indeed, Carnap laid down as a condition of adequacy of any definition of necessary truth that the necessary truths are all and only those sentences whose truth can be established on the basis of the semantical rules of the language alone (Ibid., p. 10).

In Meaning and Necessity, Carnap made various semantical distinctions that have since become standard. Every term has an *intension* and an *extension*; the intension of an expression is a function from possible worlds (state-descriptions) to its extension at that world. The intension of a term was what Carnap called an *individual concept*, which is a function from possible worlds to objects. The intension of a one-place predicate is a function from possible worlds to classes; the intension of a sentence is a proposition, which is a function from possible worlds to truth-values (the truth-value of that sentence at that world). These identifications have since become part of the basic landscape of the study of content.

Carnap's identification of intensions of terms with individual concepts, functions from possible worlds to objects, allowed him to give a distinct account of the problem of cognitive significance than Frege and Russell. According to Frege, "Scott is the author of Waverly" is cognitively significant, whereas "Scott is Scott" is not, because "the author of Waverly" has a different sense than "Scott". Russell employs his theory of descriptions to explain why "Scott is the author of Waverly" is significant and "Scott is Scott" is not. For Carnap, "Scott is the author of Waverly" is cognitively significant, because it is *factual*; it is neither a necessary truth nor a necessary falsehood. There are some state-descriptions with respect to which the extension of "the author of Waverly" is not the same as the extension of "Scott". In contrast, "Scott is Scott" is necessarily true, and so is not factual. In short, "Scott is the author of Waverly" is cognitively significant because it is *contingent*, whereas "Scott is Scott" is not informative, because it is *necessary*.

We saw in the discussion of Frege that propositional attitude verbs create *opaque linguistic contexts*, that is, contexts in which substitution of co-referential terms may change the truth-value of the sentence containing them. Using the notions of intension and extension, Carnap was also able to provide more rigorous distinctions between types of linguistic contexts in which expressions may occur. Abstracting from extra-linguistic context-sensitivity, which Carnap never considered, we may say that an occurrence of an expression e is within an *extensional context* in a sentence S if and only if one can substitute for that occurrence of e any expression with the same extension as e, without changing the truth-value of S. An occurrence of an expression e is within an *intensional context* in a sentence S if and only if that occurrence of e is not within an extensional context, and one can substitute for that occurrence of e any expression having the same intension, without changing the truth-value of S. If an occurrence of an expression in a sentence is not extensional and not intensional, then Carnap said that the occurrence of that expression in that sentence was *neither extensional nor intensional* (in contemporary vernacular, we would call that occurrence *hyper-intensional*).

As Carnap recognized, an occurrence of an expression within the scope of the modal expressions "necessarily" and "possibly" is within an intensional context. For example, from the fact that the president of the United States of America in 2005 in the youngest son of George H.W. Bush, and the fact that necessarily, if there is a unique president of the United States of America in 2005, the president of the United States of America in 2005 is a president, it does not follow that necessarily, if there is a unique president of the United States of America in 2005, the youngest son of George H.W. Bush is a president. So, substitution of co-extensional terms is not generally permitted within the scope of a modal operator such as "necessarily". But substitution of expressions with the same intension is permitted within the scope of "necessarily".

The fact that modal expressions create intensional contexts raises *the problem of de re modality*. A *de re* modal sentence is a sentence that contains a free variable in the scope of a modal operator, such as "$\exists x \Box Fx$", or in English, a sentence like "There is something such that necessarily it is rational" (a *de dicto* modal sentence is a sentence that contains a modal operator with no free variables in its scope). In *de re* modal sentences, a quantifier

such as “something” or “everything” binds a variable within the scope of a modal operator, such as “necessarily” or “possibly”. Since expressions occurring within the scope of modal operators occur in intensional contexts, this raises the worry that quantification into such positions is somehow illegitimate. After all, the quantifier that occurs outside the scope of the modal operator (to its left) presumably ranges over ordinary objects, extensions of singular terms. Yet, because the expressions occurring within the scope of the modal operator occur in intensional contexts, what is relevant for the truth of the sentence containing them are intensions, rather than extensions. In a series of influential papers, W.V.O. Quine tried to make this worry for the coherence of de re modal attributions more precise (Quine, 1943, 1947, 1953).[25]

Over the years, considerable effort has been expended in laying out these arguments in detail (see in particular David Kaplan’s masterful (1986), Fine (2005a, 2005b), and Neale (2000)). The core of Quine’s objection involves the following kind of contrast:

(1) $\Box$(the number of planets ≥ 7)
(2) $\Box(9 \geq 7)$

As Quine points out, (1) is false, and (2) is true. Yet (3) of course is also true:

(3) the number of planets = 9

Thus, it appears that two co-extensional terms (such as “the number of planets” and “9”) cannot be substituted for one another *salva veritate* (without change in truth-value) under the scope of a modal operator. According to Quine, this shows that objectual quantification into modal operators is not permissible. Somewhat less enthymematically, Quine argues that failure of substitutivity of co-extensional expressions in a linguistic context entails that the context is not “purely designative”, where “[a]n occurrence of the name in which the name simply refers to the object designated” is a purely designative occurrence of that name (Quine, 1943, p. 114). He then proceeds to argue that the coherence of quantification into a position requires that expressions occurring in that position are purely designative (Ibid., pp. 116-118).[26] Behind all of these arguments is the thought that the semantically relevant value of a variable in a non-extensional context is not just the object that is the value of that variable, but also how that object is thought of or named, and that this fact undermines the coherence of objectual quantification into that position.[27]

Carnap’s own response to the problem of de re modality involved his “method of extension and intension”, which involved simultaneously assigning to each expression, including variables, both an intension and an extension (Carnap, 1958, 42-46). When a variable occurs within the scope of a modal operator, the value that is relevant is the intension, rather than the extension. But nevertheless, even when the relevant value of an occurrence of an expression is its intension, that occurrence still has an extension as one of its semantic values. Carnap contrasted his “method of extension and intension” with what he called *the method of the name relation* (Carnap, 1958, Chapter 3). According to the method of the name relation, each occurrence of an expression has only one semantic

value. The method employed by Frege and the influential developer of Frege's method, Alonzo Church (Church, 1951) is a special case of the method of the name relation. Frege took each occurrence of an expression in a sentence to have only one semantic value. If the occurrence is within a non-extensional context, then the occurrence has as its semantic value something other than its ordinary reference; if the occurrence is embedded under only one intensional operator, then it has its ordinary sense as its referent. Carnap objected to Frege's method on the ground that it led to the *problem of the hierarchy of senses*. Frege takes an opaque context to shift the references of the expressions in its scope to the senses of those referents. So it looks as if Frege is fundamentally committed to the thesis that an expression embedded under two propositional attitude verbs (such as the occurrence of "Hesperus" in "John believes that Mary believes that Hesperus is a planet") must have, as its referent, its "indirect" sense, that is, a way of thinking of its ordinary sense. A Fregean semantic theory therefore involves the assignment of an infinite number of semantic values to each expression type.[28] Carnap's method of extension and intension does not suffer from this defect.[29]

From a contemporary perspective both the Frege/Church method and Carnap's method of treating the problem of de re modality are species of the same genus. Both treat variables occurring in modal contexts as special in some way, as not *simply* contributing ordinary referents, relative to an assignment function, to the semantic value of the sentence in which they occur. For both Church and Carnap, the semantically relevant value of an occurrence of a variable occurring in a non-extensional context, relative to an assignment function, is not an extension of a singular term, but something like an intension.[30]

In contrast to modal expressions, propositional attitude verbs, according to Carnap, do not create intensional contexts. Rather, positions within the scope of propositional attitude verbs are *neither intensional nor extensional*. Recall that for Carnap the intension of a sentence is a proposition, and two sentences express the same proposition if and only if those sentences have the same truth-value with respect to every possible world (are "L-equivalent", in Carnap's terminology). Suppose that John believes that 2+2 = 4. But John disbelieves that Peano Arithmetic is incomplete (say he has been misinformed). But "2+2 =4" and "Peano Arithmetic is incomplete" express the same proposition, according to Carnap's criterion of identity for propositions. Both sentences are necessarily true, and according to Carnap's criterion of identity for propositions, there is only one necessarily true proposition. So "believes" does not create an intensional context, because one can substitute expressions with the same intension into its scope, and alter the truth-value of the whole sentence.

According to Carnap's criterion of identity for propositions, "2+2=4" and "Peano Arithmetic is incomplete" express the same proposition, despite having very different structure. Indeed, unlike Frege's thoughts and Russell's propositions, there is no reason to think that Carnap's propositions have structure at all. We could take a Carnapian proposition to be a function from possible worlds to truth-values, or alternatively to be the set of possible worlds in which the proposition is true. But Carnap's account of the semantics of sentences containing propositional attitude verbs does involve the recognition that their truth-values do depend upon structure. Since Carnap's propositions

were not structured, Carnap made the truth-value of sentences containing propositional attitude verbs depend upon the structure of the *sentences* that occur within their scope. Two sentences are "intensionally isomorphic" if and only if "they are built in the same way out of designators such that any two corresponding designators are L-equivalent [have the same intension]" (Carnap, 1958, p. 56). Carnap employed the notion of intensional isomorphism in his account of the truth-conditions of a propositional attitude ascription such as "John believes that D". According to this account, "John believes that D" is true if and only if there is a sentence S in a language understood by John that is intensionally isomorphic to D, and John is disposed to an affirmative answer to S. Carnap's analysis of propositional attitude constructions is a model for subsequent analyses that take the objects of propositional attitudes to be sentences, rather than extra-linguistic entities such as propositions.[31]

There is a similar problem to the problem of de re modality facing any account of the semantics of sentences containing propositional attitude verbs. *A de re attitude ascription* is a sentence containing a free variable in the scope of a propositional attitude verb, such as "∃x(N believes that x is F)", or in English, a sentence like "Some mayor is such that John believes he is not in politics." Since propositional attitude verbs at the very least create intensional contexts, de re attitude ascriptions should be at least as puzzling as de re modal sentences. In both cases, one has a quantifier that seemingly ranges over objects (extensions of terms) binding a variable that occurs within a context in which individual concepts, rather than their extensions, are semantically relevant. Interestingly, however, at least for a time, Quine's belief in the impossibility of regimenting de re modal statements into an acceptable formalism did not extend to de re attitude ascriptions; with regards to the latter Quine (1955, p. 188) notes, "…we are scarcely prepared to sacrifice the relational construction…" In that paper, Quine proposes a way of rescuing the truth of de re attitude ascriptions, by regimenting them into a formalism in which one is not quantifying into a non-extensional context after all (Kaplan (1986) in fact shows that one can use the same mechanism to regiment de re modal claims). The distinction Quine made between the problem of de re modality and the problem of de re propositional attitude ascriptions was no doubt due to his belief that making sense out of de re modality ultimately involved accepted the coherence of dubious metaphysical notions, such as *essentiality*, whereas making sense out of de re attitude ascriptions involved no such metaphysical commitments.

However, Quine (1976, 1977) eventually came around to the same distrust of the possibility of regimenting de re attitude ascriptions as he always had in the case of modality de re. In particular, Quine concluded that de re attitude ascriptions were subject to the same inconstancies as de re modal ascriptions. One and the same de re attitude ascription could be true in one context (with the object thought of one way) and false in another (with the object thought of in another way). Following Hintikka (1962, p. 153), with whose influential work on epistemic logic Quine was engaging, Quine took the de re propositional attitude ascription "∃x(N knows(b = x))" to be synonymous with the claim that *N knows who b is*. And as Quine (1976, p. 863) notes:

> It is very ordinary language indeed to speak of knowing who or what something is. However, ordinarity notwithstanding, I make no sense of the idiom apart from context. It is essentially indexical.

Quine concludes that the inconstancy of de re attitude ascriptions makes them no more susceptible to regimentation than de re modal ascriptions. As Quine writes (Ibid.), "I do not see the makings here of a proper annex to austere scientific language."

Quine's criticism of Carnap's semantical system did not just involve suspicions with quantified modal logic. Recall that Carnap set as an adequacy condition on any definition of necessary truth that the necessary truths are all and only the analytic truths. Within the semantics described by Carnap, the way this condition of adequacy was implemented was via *meaning-postulates*. For Carnap, a possible world is, in the first instance, a maximally consistent set of sentences. Though no such set can contain a logical contradiction, such as "S" and "~S", nothing prevents such a set from containing a sentence such as "Bachelors are married". Given the semantics, it would then come out as possible that bachelors are married, and "Bachelors are unmarried" therefore would not be a necessary truth (see Quine, 1951). To prevent this, each term is associated with an analytic definition (or "meaning postulate"), and state-descriptions are constrained to make all such analytic definitions true (Carnap, 1958b). So, in the case of "bachelor", the meaning postulate is that "bachelor" means the same as "unmarried man", and any possible state-description must be one that contains the sentence "bachelors are unmarried men". Thus, meaning postulates eliminate state-descriptions containing "bachelors are married men" (since this sentence is logically inconsistent with the meaning postulate that "bachelors are unmarried men"), and "bachelors are unmarried men" comes out necessarily true (and hence analytic and non-factual) after all.

The notion of analyticity is at the core of Carnap's semantical system. In Quine's seminal 1951 paper "Two Dogmas of Empiricism" (Quine, 1961a) he launched his influential attack on the coherence of the notion of analyticity. According to Quine, there is no coherent way of making a distinction between synthetic truths and analytic truths; that is, there is no way of forging Carnap's distinction between *factual* and *non-factual* internal questions. The majority of Quine's arguments take the form of showing that explanations of analytic truth always appeal to notions that are equally problematic. For example, Carnap attempts to ground the notion of analyticity by stating that analytic truths are those that are true in virtue of the semantical rules of the language. But as Quine points out, the notion of a semantical rule is no more lucid than that of an analytic truth; nothing demarcates the statements that are semantical rules from the statements that are not semantical rules besides "…appearing on a page under the heading 'Semantical Rules'…" (Quine, 1961a, p. 34). After finding no notion that can explicate the notion of analyticity, Quine then rejects it as ill-founded.[32] If it is ill-founded, then the project of dividing the genuinely empirical claims of science from the non-factual claims of metaphysics, a project that was at the heart of Carnap's semantics, is doomed.

The precise forms of Quine's arguments against analyticity have been a subject of continual debate ever since the publication of "Two Dogmas" and I cannot provide a

lengthy discussion of them here. It suffices to say that though Quine's grounds remain somewhat murky, his rejection of analyticity is widely (though certainly not universally) accepted.[33] But whatever one thinks of Quine's attack on analyticity, Carnap had developed and refined the tools of intensional semantics to such a degree that one could employ them independently of Carnap's intended interpretation of necessity. So as we shall see later, while Carnap's own motivation for his semantic project became bogged down in debates over the coherence of its central notion, and the popularity of his anti-metaphysical philosophical project waned, a number of logicians and philosophers took up the tools developed by Carnap and applied them to some of the traditional questions in the theory of meaning.

VI. Strawson and the Challenge from Ordinary Language Philosophy

All of the philosophers we have been discussing were suspicious of the possibility of using logical tools to investigate natural language. Frege regarded natural language as too vague and context-sensitive to conduct scientific investigation, and Tarski thought that the "universal character" of natural languages rendered them inconsistent. Carnap's semantical systems were intended to be systems of analytic stipulations governing the meanings of the terms in the language; Carnap had no interest in applying his formal tools to the project of empirical semantics.[34] The uneasiness these philosophers had about applying the tools of logic to natural language was mirrored by philosophers who focused mainly on natural language, the so-called "ordinary language" school of philosophy, whose work is best exemplified in the writings of J.L. Austin and Peter Strawson; I will focus on the latter in explaining the doctrines.

The semantical tools developed by Frege, Russell, Tarski, and Carnap involved giving a characterization of the conditions under which sentences of a given language were true (perhaps relative to a model). As we have seen, such a characterization involves assignment of reference to terms and satisfaction conditions to predicates. Central to the ordinary language philosopher's view is the thesis that properties such as truth and reference do not apply to linguistic expressions but are rather properties of what people *do* with linguistic expressions. It is a *use* of a singular term by a person that refers, and it is an assertion of a sentence that has a truth-value; one cannot speak of a *term* having reference, or a *sentence* having a truth-value. In short: words do not refer, people do. If reference and truth are not properties of linguistic expressions, then giving an account of linguistic meaning in terms of reference and truth is fundamentally misguided. Carnap and Tarski were right to focus their attention on formal languages, because the kind of account of meaning they were trying to give (in terms of reference and truth) was inapplicable to natural languages.

Since truth and reference are not properties of linguistic expressions, and linguistic meanings are properties of linguistic expressions, ordinary language philosophers sought an alternative account of linguistic meaning. According to it, the linguistic meaning of an expression is a *rule for its proper use*. As Strawson writes in his classic 1950 paper, "On Referring" (Strawson , 1996, pp. 219-220)):

> To give the meaning of an expression (in the sense in which I am using the word) is to give *general directions* for its use to refer to or mention particular objects or persons; to give the meaning of a sentence is to give general directions for its use in making true or false assertions….The meaning of an expression cannot be identified with the object it is used, on a particular occasion, to refer to. The meaning of a sentence cannot be identified with the assertion it is used, on a particular occasion, to make. For to talk about the meaning of an expression or sentence is not to talk about its use on a particular occasion, but about the rules, habits, conventions governing its correct use, on all occasions, to refer or to assert.

In "On Referring", Strawson does not just state that meaning is use. He shows by a detailed example that two sentences can be used to express the same truth-conditions, yet differ on their use-conditions, and that this difference is a matter of the conventional meaning of the words used. He thus shows, by detailed consideration of a particular case, that giving the truth-conditions of an occurrence of a sentence, or giving the proposition it expresses, is to miss something about the conventional meaning of that expression. Unsurprisingly, the example Strawson uses is the case of *definite descriptions*.

For Russell, sentences containing definite descriptions express existential propositions. A sentence such as "The shortest spy is nice" expresses, for Russell, a proposition whose initial quantifier is existential in force (the proposition that there is a shortest spy, that everything that is a shortest spy is identical to her, and that she is nice). But construing sentences containing definite descriptions as expressing the same proposition as a sentence that contains only existential and universal quantifiers is to miss a crucial distinction in use-conditions between definite and indefinite descriptions. Furthermore, these use-conditions are clearly part of the conventional meanings of definite and indefinite descriptions. As Strawson (1996, p. 228) writes:

> The difference between the use of definite and indefinite articles is, very roughly, as follows. We use "the" either when a previous reference has been made, and when "the" signalizes that the same reference is being made; or when, in the absence of a previous indefinite reference, the context (including the hearer's assumed knowledge) is expected to enable the hearer to tell *what* reference is being made. We use "a" either when these conditions are not fulfilled, or when, although a definite reference *could* be made, we wish to keep dark the identity of the individual to whom, or to which, we are referring.

So even if we grant to Russell the thesis that "The shortest spy is nice" expresses the same proposition as that expressed by "There is a shortest spy, and everything that is a shortest spy is her, and she is nice", it does not follow that the two sentences have the same *meaning*. There are very different conditions of use associated with the two sentences, despite the agreement in truth-condition. Furthermore, these distinctions in use clearly have something to do with conventional properties of definite and indefinite descriptions. As Strawson notes, definite descriptions are typically used to refer entities already introduced (familiar entities), and indefinite descriptions are typically used to

introduce novel entities into the discourse. Strawson's objection here to Russell's theory of descriptions is that, by ignoring non-truth conditional features of use, it ignores crucial differences in conventional meaning between definite and indefinite descriptions.

Strawson's famous concluding sentence (Ibid., p. 230) in "On Referring" is that "[n]either Aristotelian nor Russellian rules give the exact logic of any expression of ordinary language; for ordinary language has no exact logic." This claim is best seen in the light of the discussions in Strawson (1952) of the differences between the truth-tables analyses of the logical connectives of propositional logic, and the ordinary words "and", "or", "if…then", and "if and only if". In each case, Strawson argued that there was a large gulf between the logical connective and its alleged ordinary language counterpart. For example, Strawson (1952, p. 80) objects to the truth-table meaning as a suitable characterization of the meaning of the English "and" that occurs between sentences as follows (I use '&' to denote the connective defined by the truth-tables):

> It might be conceded that 'and' has functions that '&' has not…, and yet claimed that the rules that hold for 'and', where it is used to couple clauses, are the same as the rules that hold for '&'. Even this is not true. [By the truth-table for '&'], 'p & q' is logically equivalent to 'q & p'; but 'They got married and had a child' or 'He set out to work and found a job' are by no means logically equivalent to 'They had a child and got married' or 'He found a job and set out to work'.

Strawson also rejected any kind of meaning equivalence between the material conditional '→' and the 'if…then' of ordinary language. As he says about the latter (Ibid., p. 37), "…in general its employment in linking two clauses indicates that a statement made by the use of the first would be a ground or a reason for a statement made by the use of the second." More explicitly, Strawson writes (Ibid., p. 88):

> …I have spoken of a 'primary or standard' use of 'if…then…', or 'if', of which the main characteristics were: that for each hypothetical statement made by the use of 'if', there could be made just *one* statement which would be the antecedent of the hypothetical and just *one* statement which would be its consequent; that the hypothetical statement is acceptable (true, reasonable) if the antecedent statement, if made or accepted, would be a good ground or reason for accepting the consequent statement; and that the making of the hypothetical statement carries the implication either of uncertainty about, or of disbelief in, the fulfillment of both antecedent and consequent.

Certainly, none of these facts about the primary or standard use of 'if…then' in English are captured by the truth-table for the material conditional.

Similar points apply to the "standard use" of disjunctive statements in English, instances of the schema 'P or Q'. It is reasonable to assert an instance of 'P or Q' only if one is unsure about the truth-value of both disjuncts. For example, it is odd for someone fully aware of the political facts to utter, in 2006, "Either George Bush is president now or Bill Clinton is president now". The fact that it is not reasonable to assert an instance of 'P or

Q' unless one is unsure of the truth-value of both disjuncts is clearly a fact about the standard use of sentences containing "or", and not captured by the truth-table for the logical connective for disjunction. A further difference between the truth-table for disjunction and the English word "or" is "…that in certain verbal contexts, 'either…or…' plainly carries the implication 'and not both…and….', whereas in other contexts it does not." (Strawson, 1952, p. 92). So, there are two distinct uses of the English word "or", and hence apparently two distinct meanings – one corresponding to the truth-table, or inclusive sense of "or", and the other corresponding to the "exclusive" sense of "or" (the "and not both" reading indicated by Strawson). So "or" is ambiguous, whereas the logical connective for disjunction is not (though both distinct meanings are of course characterizable with the use of truth-tables).

The central challenge of ordinary language philosophy is that reference and truth are inappropriate notions to employ in explicating linguistic meaning for a language with pervasive context-sensitivity mediating the relation between word and world. Instead, we need the notion of a rule of proper use. Reference and truth do not help in the analysis of most rules of use; the rules of proper use governing words are not subject to rigorous semantic analysis. For example, truth-tables are clearly hopeless in explaining both the connection between the antecedent and the consequent of a natural language hypothetical statement, and the fact that a statement of this kind is assertible only if the speaker disbelieves the antecedent and consequent; they are equally useless in explaining the similar facts pertaining to natural language disjunctions.

VII. Grice and the Semantics-Pragmatics Distinction

In his extraordinarily influential paper "Logic and Conversation" (Grice, 1989a), Grice set out to defend the truth-table analysis of the meaning of the natural language logical particles from the ordinary language onslaught. Recall that when Strawson spoke of the connection between antecedent and consequent that is part of the "primary use" of an English conditional statement, he spoke of the *acceptability*, *truth*, or *reasonability* of a use of a conditional statement. This suggests that Strawson did not distinguish the *truth* of an utterance from the *acceptability* of that utterance. The key to Grice's defense of the truth-table analysis of the meanings of "and", "or" and "if…not" is that these notions can (and often do) come apart. A given utterance can be true, even though uttering it is not acceptable, because it violates conversational norms. In explaining this distinction, Grice provided the foundations for a theory of conversational norms. The theory Grice gives clearly explains how an utterance may be true, though unacceptable as an assertion due to specific facts about the conversation and its participants. Grice then used the distinction between the truth of a statement and its conversational acceptability in a defense of the thesis that the connectives of propositional logic were correct explications of their natural language counterparts. More specifically, Grice assumed that the natural language logical particles have the truth-table meanings of their logical counterparts, and argued that features of the uses of these expressions that are not explicable by the truth-tables are due to facts about the norms governing conversation, rather than the meanings of the words.

According to Grice, conversation is a cooperative rational activity; each conversation has a purpose. This fact about conversations imposes as a norm what Grice (1989a, p. 26) calls the *Cooperative Principle*, which is "Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged." The Cooperative Principle is the overarching principle guiding conversation. Following it imposes a number of more specific norms on conversational participants. For example, lying involves one kind of uncooperative conversational behavior, being purposely irrelevant involves another, and not being sufficiently informative is a yet a third kind of uncooperative behavior. Following the Cooperative Principle isn't always a matter of saying something true, relevant, and maximally informative; it is also a matter of how one says what one says. According to the maxim of *manner*, one should try to list events in the order in which they occur, and to cite causes before effects.

Using these conversational principles, Grice attempts to explain many of the facts about standard use cited by Strawson without giving up the thesis that the same truth-table analysis for the logical connectives also gives the meanings of their natural language counterparts. Consider Strawson's point that "or" in ordinary language is ambiguous between an *exclusive* use ("but not both") and an *inclusive* use. Assuming that "or" unambiguously means inclusive "or" (the meaning of the logical connective for disjunction), one can explain the fact that "or" is often used exclusively by general conversational principles. Suppose Hannah uttered an instance of "P or Q", but in fact believed that both P and Q were true. Then Hannah would not be maximally informative; she would be violating Grice's conversational maxim of *quantity*. So if someone believes that P and Q, if they wish to follow conversational norms, they should say the more informative P and Q, rather than the less informative P or Q (which is compatible with the truth of only one of P and Q). So, when someone utters an instance of 'P or Q', they convey (without asserting, as part of the linguistically determined content) that they do not know that P and Q. The fact that this is part of what is conveyed by following conversational principles, rather than what is asserted as part of the linguistically determined content, can be ascertained by appeal to Grice's central criterion for distinguished what is part of what is said (the linguistically determined asserted content) from what is merely conversationally conveyed, which is the test of *cancellability*. One can *cancel* the implication conveyed by an utterance of 'P or Q' (which is that one doesn't know both P and Q) by saying "P or Q; in fact, both P and Q are true", as in "John is with Bill or he is with Frank; in fact he is with both". So, consistently with the assumption that "or" unambiguously *means* inclusive "or", one can explain why "or" is often *used* as if it meant exclusive "or".

One can use the very same kind of explanation to dissolve the sense that it is part of the meaning of a disjunctive statement that the speaker is unaware of the truth of either disjunct. If Hannah knows that John was at the party, it would be a violation of the maxim of quantity for her to assert that either John was at the party or he was at home. She would not be being maximally informative by asserting the disjunctive statement, and hence would be violating the maxim of quantity. Furthermore, the implication that the speaker is unaware of the truth of either disjunct can be cancelled, as in Grice's

example (1989b, pp. 44-5) “The prize is either in the garden or in the attic. I know that because I know where I put it, but I’m not going to tell you.” Thus, one can explain the fact that a disjunctive statement is usually only proper if the speaker is unaware of the truth of either disjunct, without making that fact part of the conventional meanings of any words.

Grice also attempted to provide pragmatic explanations (that is, explanations from general principles governing conversation) for the divergences between the truth-table meaning for the conditional and ordinary indicative conditionals (Grice, 1989c).[35] In attempting to account for the connection thesis, the thesis that a conditional is only assertible if the antecedent provides a ground or good reason to accept the consequent, Grice (1989c, pp. 61-2) appealed to the conversational maxims, in particular that of quantity, which directs interlocutors to always assert the strongest claim consistent with their evidence, and that of quality, which directs them to have adequate evidence for their assertions. If the indicative conditional is the material conditional, then it is true if and only if the antecedent is false or the consequent is true. If the speaker knows that the antecedent is false, adherence to the maxim of quantity requires that the speaker simply assert the negation of the antecedent, rather than the whole conditional; *mutatis mutandis* for the truth of the consequent. So a conditional is only assertible if the speaker is unaware of the truth values of the antecedent and consequent. But the maxim of quality requires anyone who asserts a conditional to have evidence for the truth of the material conditional. Since, for the reasons just given, the evidence cannot be truth-functional (that is, the speaker’s grounds cannot be knowledge of the truth-values of the antecedent or consequent), the speaker must have non-truth-functional grounds for her assertion of the material conditional, if she is adhering to the maxims of quantity and quality. So, asserting an indicative conditional, on the supposition that it has the meaning of the material conditional, requires the speaker to have non-truth-functional grounds for her assertion. More specifically, it requires the speaker to know or believe that the antecedent would be a good ground for the consequent.

As we shall see, there are a number of problems with Grice’s defense of the material conditional analysis of indicative conditionals. But Grice’s defense of the thesis that the meaning of “or” is exhausted by the truth-table for inclusive “or” has been widely accepted, as have a number of other Gricean explanations of use-facts. The moral of Grice’s work is that the facts of linguistic use are a product of two factors, meaning and conversational norms. Failure to absorb this fact undermines many of the main theses of ordinary language philosophy.

However, recall that there were two aspects of the ordinary language philosopher’s position. The first involved emphasizing the divergences in use between the logical terms and their ordinary language counterparts. The second involved the fact that natural languages involve context-sensitive words (e.g. “I”, “here”, and “now”), and that many words only have reference relative to a context of use, and many sentences only have truth-values relative to a context of use. Since reference and truth-value are only properties of *uses* of expressions, they are inappropriate notions to use in the analysis of the linguistic meanings of expressions. In general expression types in natural language do

not have references or truth-values, only uses of them do. So employing the apparatus of semantic theory, which crucially avails itself of notions such as reference and truth, is not the right way to give a theory of meaning for natural language; the meaning of expression types is given by rules of use. Grice's response to the ordinary language philosopher only speaks to the first of these aspects of the ordinary language philosopher's position. But a response to the second aspect of the ordinary language philosopher's position was to emerge from the work of those who developed and refined intensional semantics.

VIII. The Development of Intensional Semantics: From Montague to Kaplan

As we saw in section V, Carnap's semantic theory crucially exploits the notion of a possible world in defining semantic values of expressions. Each expression has, as its primary semantic value, an intension, which is a function from a possible world to the extension of that expression at that world. In the case of sentences, the intension of a sentence is a function from possible worlds to truth-values. Carnap's semantic theory has, as its "central notion" (in Michael Dummett's sense), the notion of *truth with respect to a possible world*. The logician Richard Montague, a student of Tarski's, argued that a theory of meaning should take the more general form of *truth with respect to a context of use*, where possible worlds are but one feature of a context of use (Montague, 1974a, p. 96). Montague treated a context of use as an *index*, a collection of semantically relevant aspects of the context of use. If the language in question contained tenses and modal operators, then the indices involved in the semantic interpretation of that language would contain times and worlds. If the language also contained the indexical terms "I" and "here", the indices would also have persons and places as aspects. Montague then generalized Carnap's notion of intension; instead of an intension being a function from possible worlds to extensions, an intension, for Montague, was a function from indices to extensions. For example, the intension of a sentence such as "I am tired" would be a function from indices to truth-values; it would take an index whose aspects were times, worlds, and persons to the true if and only if the person at the index was tired at the time and world of the index.

The interpretation of modal operators in Montague's system was also a generalization from their interpretation in modal semantics. In Carnap's system, the function of modal operators was to *shift* the evaluation of a proposition from one possible world to another; a modal operator took an intension, and evaluated that intension at other possible worlds. On this account, a sentence such as "possibly S" is true relative to a world w if and only if the intension of S is true is some (possibly distinct) world w'. So the function of "possible", for Carnap, is to shift the evaluation of the intension of S from w to w'; "possibly S" is true in w if and only if S is true in w' (and the function of "necessarily" is to shift the evaluation the content of the embedded sentence to all possible worlds). In Montague's system, modal and tense operators evaluate intensions at *indices* rather than just possible worlds. On this account, a sentence such as "possibly S" is true at an index i if and only if the intension of S is true at i', where i' differs from i at most in its world feature. So rather than truth with respect to a possible world being the fundamental notion, truth with respect to an index is Montague's fundamental notion, with worlds being one element of an index.[36] This apparatus allowed Montague to generalize the

apparatus of intensional semantics to treat *context-sensitivity* in natural language, without sacrificing the elegant treatment of modal and tense operators. As we shall see below, this leads to an alternative kind of response to the challenge from ordinary language philosophy than the one developed by Grice.

Montague's contributions to the systematic study of language went well beyond generalizing intensional semantics to capture tense and context-sensitivity. Montague's most influential papers focused on intensional constructions in natural language. In Montague (1974c), he gave an account of a number of intensional constructions other than the classic cases of propositional attitude verbs and modal and temporal contexts. For example, Montague provided a semantic analysis of *intensional transitive verbs*, such as "seek" and "worship". The difference between intensional transitive verbs and extensional transitive verbs (such as e.g. "kick" and "meet") is that, while one cannot meet a unicorn or kick a unicorn (since there are no unicorns), one can nevertheless *seek* a unicorn. So, whereas satisfying an instance of the predicate "meeting N" requires that that there is some existent entity that one meets, satisfying an instance of the predicate "seeking N" does not require that there is some existent entity that one seeks.

Intensional transitive verbs had generally been ignored in the literature beginning with Frege and Russell, largely because of the influence of Russell's theory of descriptions, the standard method of dissolving apparent reference to non-existent entities. Russell's theory involves providing a contextual definition of definite descriptions; meanings are assigned only to sentences containing definite descriptions, rather than the definite descriptions themselves. Russell's theory helps us analyze away apparent reference to non-existent entities in a construction such as "John believes that the fountain of youth is in Peru", since we can apply the theory to the sentence "the fountain of youth is in Peru", and arrive at an object of John's belief, without there being a fountain of youth. In contrast, one cannot use Russell's theory to arrive at an object of seeking for a construction such as "Pizarro sought the fountain of youth", since that theory gives us no way of treating the definite description "the fountain of youth" in isolation. For this reason, Quine (1960, section 32) regimented intensional transitive verbs away in favor of propositional attitude verbs (so the intensional transitive construction "x looks for y" becomes the propositional attitude construction "x endeavors that x finds y").[37] The fact that analytic philosophers had not produced a successful analysis of intensional transitive verbs must be viewed as a bit of an embarrassment. The problem of intensional transitive verbs is one of the original motivations for Twentieth Century discussions of content. For example, it was salient in the minds of Brentano and his students, who sought to render consistent the thesis that the characteristic feature of mental states was that they were about things, with the fact that one could have a mental state the object of which did not exist. Montague's discussion of intensional transitive verbs was thus a watershed moment in the theory of meaning. It has subsequently given rise to a lively literature in semantics and philosophy of language on the topic (e.g. Partee (1974), Zimmerman (1993), Forbes (2000), Richard (2001)).

Montague's semantic theory was not just distinctive for its focus on intensional constructions in natural language. Montague also returned philosophers of language and

semanticists to a tradition that was lost or at the very least obscured in the kind of semantic theory favored by Tarski (and Davidson). Recall that Frege treated the traditional relation between the subject of a sentence and its predicate as that of an *argument* to a *function*. That is, Frege regarded the fundamental relationship between the semantic values of expressions in a sentence to be one of *functional application*. Though quantifiers, for Frege, had the function of binding variables within their scope, they also had determinate semantic values, namely *second-level functions*.[38] For example, as we saw above, "everything" denoted a function from first-level functions to truth-values. The denotation of "everything" is a function that takes any first-level function that takes every object to the true, to the true, and takes every other entity to the false. Similarly, the denotation of "something" takes to the true any first-level function that had the true for at least one value, and everything else to the false. So Frege operated with an ontology that was stratified into *types*; there were objects, then functions from objects to truth-values (first-level functions), then functions from first-level functions to truth-values (second-level functions), and on up. In Tarski's work, by contrast, no use is made of functional application as a relation between semantic values. Quantifiers are not assigned functions of various kinds; an object-language universal quantifier over objects is interpreted via the use of a meta-language quantifier over sequences. Montague's semantics returned philosophers of language to the Fregean tradition of treating semantic values as functions from arguments to values, with functional application as the primary mode of semantic composition. There are lively foundational debates between advocates of Montague's *type-theoretic* approach to semantics and advocates of the more Tarskian approach, such as James Higginbotham.

Montague's marriage of intensional semantics with type-theory was extraordinarily fruitful, and led (with the help of the work of his distinguished student Barbara Partee) to the emergence of *semantics* as a new discipline within linguistic theory. But that is not to say that the generalization of intensional semantics that was at the heart of his program has been universally accepted; in fact, the majority of philosophers of language today regard it as incorrect. The mistake made by Montague was to think that the study of modality was a branch of *pragmatics*, the study of context-sensitivity in natural language. Recall that Montague's generalization of intensional semantics consisted of treating possible worlds as features of the more general notion of a *context of use*. He then generalized the treatment of operators as shifting the evaluation of the truth of a content from one world to the next, to shifting the evaluation of the truth of a content from one context of use (or index) to the next. It is this generalization that is widely (but not universally) regarded as an error.

The first hint that something was amiss in the assimilation of modal and temporal operators to the general study of truth relative to a context of use came from Hans Kamp's work on the temporal indexical "now". Kamp, a student of Montague's, established several theses about temporal logic. The first is that representing the full range of English tensed sentences with just a past tense operator and a future tense operator is not possible. There are certain sentences that cannot be expressed in a tense logic with just past and future tense operators. For example, the sentence in (1) (from Kamp (1971, p. 231)) cannot be expressed in such a language:

(1) A child was born that will become the ruler of the world.

Sentence (1) means something like "In the past, a child was born who, in the future of the present moment, becomes the ruler of the world". In order to express (1) in a language with only temporal operators, one needs to have at one's disposal an operator with the meaning of the English word "now", whose function is to evaluate its embedded content *at the present moment*. Kamp then established that a satisfactory semantics for "now" requires having two times in the Montagovian "index" that is supposed to represent a context of use. One of the times would be shifted by temporal operators such as "it was the case that" and "it will be the case that". The other time would be intuitively the time of the utterance, and would never be shifted by any operators. Its function would be to allow for the interpretation of any occurrences of "now" in the sentence. The present moment feature of the Montagovian index could not be shifted by any operators, because otherwise, in interpreting any embedded occurrences of "now" (that is, embedded inside other temporal operators), one could no longer access the present moment, and thereby successfully interpret "now".

Kamp's insights about "now" carry over directly to the modal indexical "actual". In order to provide a successful interpretation of embedded occurrences of "actual" (that is, occurrences of "actual" inside other modal operators), each Montagovian index must contain two worlds, one that would be shifted by modal operators, and the other that would be the world of utterance. Interpreting an embedded occurrence of "actual" (that is, one that occurs within the scope of other modal operators) requires keeping track of the world of utterance. For the function of the initial modal operator is to shift the evaluation of the content of the embedded sentence to another possible world, and one needs to retain the information about the actual world of utterance, in order to interpret any occurrences of "actual" within that embedded sentence. So Montague's indices each would contain two kinds of features. First, they would contain features (worlds and times) that were shifted by operators. Secondly, to interpret indexical operators such as "now" and "actual", the indices would contain features that were not capable of being shifted by operators, but would always represent features of the actual context of use of the sentence being uttered.

So, Kamp's work suggests that within a single Montagovian index, there are two quite different sorts of features. First, there are features that are shifted by operators, such as "necessarily" and "possible" (and the past and future tense, assuming that they are operators). Secondly, there are features that intuitively represent features of the actual context of use. These include the moment at which the utterance was made, and the world at which the utterance was made, which are required, respectively, to interpret indexical operators such as "now" and "actual". Furthermore, these features are not capable of being shifted by operators, or else one could not interpret embedded occurrences of indexical operators. So, for example, in evaluating the truth of the intension of say "Necessarily S" with respect to an index i, one would evaluate the intension of S at all indices i' that differed from i at most in their world index, and shared with i all the features relevant for interpreting indexical operators; that is, all those features that

represent aspects of the context of use in which "Necessarily S" was uttered. This suggests that Montague's indices are not natural kinds. Each index contained two kinds of information: information relevant for interpreting modal and temporal operators, on the one hand, and information that represented features of the context of use, which are relevant for interpreting indexical expressions such as "now", "actual", "I", and "here".

There were also other reasons to be suspicious of Montague's index-theoretic approach. In 1970, Robert Stalnaker pointed out (Stalnaker (1999a, pp. 36ff.)) that Montague's semantics (or as Stalnaker calls it, Montague's "Semantics-Pragmatics") did not allow for the representation of *propositions*. For Montague, there is only one semantic content of an utterance (or occurrence) of the sentence "I am tired", and that is a function from contexts of use to truth-values. If Hannah utters "I am tired", and John utters "I am tired", the only difference there is between the contents of their utterances is that one may be true and the other may be false (that is, the value of the semantic content of "I am tired" may be different, because it is being evaluated relative to distinct indices). But, as Stalnaker emphasizes, there are additional differences between their utterances. Intuitively, *what Hannah said* when she uttered "I am tired" is distinct from *what John said* when he uttered "I am tired"; they expressed different propositions. But there is no semantic value in Montague's system that represents the different propositions in question. There is just a function from indices to truth-values associated with "I am tired", and this is not the proposition expressed by either of these utterances of "I am tired" (since they express *different* propositions).

Kamp's work clearly shows the need for "double indexing". The first kind of index is required to interpret indexical expressions occurring within a sentence. The second kind of index is required to give the proper semantics for operators on content, such as "necessarily" and "possible" (and the tenses, if they are operators). But it took another student of Montague's, David Kaplan, to draw out the real moral behind the need for double indexing. In his seminal work "Demonstratives" (Kaplan (1989)), Kaplan argues that the two kinds of indices correspond to two kinds of semantic values.[39] The first kind of index represents the dependence of semantic value upon context. The semantic content of a context-dependent sentence such as "I am tired" depends upon features of the context of use. If John is the speaker in the context of use, then "I am tired" expressed the proposition that John is tired; if Hannah is the speaker in the context of use, then it expresses the proposition that Hannah is tired. The second kind of index represents the dependence of truth of a semantic content on a *circumstance of evaluation* (such as a possible world, or a time if tenses are operators on contents), and is required to give a satisfactory semantics for sentence-operators. A proposition may be true at one possible world, but false at another.

Accordingly, expressions are associated with two kinds of semantic values, which Kaplan called *character* and *content* respectively. The character of an expression is a function from a context of use to the content of that expression relative to that context. According to Kaplan, the character of an expression is also the linguistic meaning of that expression. So, the linguistic meaning of the first-person pronoun "I" is a function from contexts of use to persons (intuitively, the speakers of those contexts). Any use of "I" has the same

meaning as any other use, though a possibly distinct semantic content. Kaplan took the semantic contents of singular terms, such as proper names and indexicals such as "I" to be their referents, in Frege's sense, and he took the semantic contents of sentences relative to contexts to be propositions. Sentences operators such as "necessarily" and "possibly" shifted the world feature of the index that represented the circumstance of evaluation. The index that represented the context of use did not contain any features that were shifted by operators in the language (Kaplan, 1989, pp. 510ff.).

By dividing features of indices into contexts of use and circumstances of evaluation, Kaplan's semantic theory represents a clear advance over Montague's. It explains why only certain features are shiftable by operators and, more importantly, it gives a semantic representation of *propositions* (the values of the characters of sentences).[40] As a result, Kaplan's distinctions have been widely adopted in philosophy of language over the last thirty years. In particular, most philosophers have come to accept that context-dependent expressions show that there are two levels of semantic value; first, linguistic meaning, and secondly, the content of an occurrence of an expression on an occasion. Different occurrences of an expression might have different semantic contents, despite sharing a linguistic meaning, as is so clearly the case with the first-person pronoun "I" and other indexicals.

The work done by Montague and then Kaplan allows for another kind of reply to the ordinary language philosopher's skepticism about the possibility of giving a rigorous semantics for a natural language than the one provided by Grice. Recall that the ordinary language philosopher's skepticism arose from the conviction that truth and reference were properties of *uses* of expressions, rather than properties of expressions, and meanings were rules for using those expressions. Kaplan's semantic theory undermines these considerations. It does make perfect sense to speak of singular terms having reference, albeit *relative to a context*, and it makes perfect sense to speak of sentences having truth-values, also relative to a context. So it makes perfect sense to attribute reference and truth to expression types, once contextual relativity is factored into the semantic theory. Whereas the notion of a rule of use is vague and mystical, Kaplan's notion of the character of an expression is not only clear, but set theoretically explicable in terms of fundamental semantic notions; the character of an expression is a function from a context to the reference of that expression in that context. Far from context-sensitivity being an impediment to giving a proper account of linguistic meaning in terms of reference and truth, appeal to these semantic notions allows us to give a considerably more explicit characterization of linguistic meaning than the ordinary language philosophers were capable of providing.

IX. Necessity Regained

Though Montague and his descendents refined and extended the intensional semantic framework developed by Carnap, they did not at all share with him the interpretation of necessity as analyticity. Instead, a growing consensus developed around the idea that *metaphysical* necessity was a legitimate interpretation of modality. However, the

consensus built up slowly, and in large part as a reaction to Quine's influential criticisms of de re modality.

Recall that Carnap's solution to the problem of de re modality involved assigning a dual interpretation to each expression; every occurrence of an expression had both an intension and an extension, including variables. What mattered for the truth of an open sentence embedded inside a modal operator were the value-intensions of the variables occurring within it. The Frege/Church method, by contrast, involved taking expressions occurring inside modal contexts as denoting intensions rather than extensions. Variables occurring within modal contexts, according to this approach, ranged only over intensions. Both of these approaches treated the semantically relevant values of variables occurring within modal contexts as intensions, and in the case of individual-level variables, *individual concepts* (functions from possible worlds to objects).

The individual concept approach to de re modal quantification concedes that quantification into modal contexts is special in some way. The insight that led to the current consensus about modality lies in recognizing that quantification into modal contexts is not special in any way; quantification into modal contexts should be treated just like quantification into extensional contexts. In other words, on this *objectual* conceptual of quantification, the semantically relevant values of variables in modal contexts are just the same as the semantically relevant values of variables in extensional contexts, namely normal objects.

Let us return to Quine's worry, and in particular, the distinction in truth-value between (1) and (2):

(1) □(the number of planets ≥ 7)
(2) □(9 ≥ 7)

Quine's concern is that objectual quantification into the position of the variable "x" in the open modal sentence "□(x ≥ 7)" is incoherent, because whether or not a sequence satisfies "□(x ≥ 7)" will depend not just upon the object that sequence assigns to the variable "x", but also on how we *describe* that object (as "9" or as "the number of planets"). The objectual conception involves rejecting the thought that in quantifying into a modal context one needs to have any description at all of the objects that are the values of the variables. The value of a variable is simply the object it designates, and so variables, no matter where they occur, are *purely designative* in Quine's sense, and hence permitted to be bound by quantifiers even when they occur in the scope of a modal operator.

But recall Quine's argument, presented in section V, for the incoherence of quantifying into modal contexts. Quine inferred from the fact that co-extensive terms such as "the number of planets" and "9" could not be substituted for one another inside the scope of a modal operator without change in truth-value, to the conclusion that the position of the variable "x" in the open modal sentence "□(x ≥ 7)" is not a purely designative position.

So how is it possible to hold, in the face of the facts in (1) and (2), that the variable "x" in "□(x ≥ 7)", is purely designative?

From the perspective of the advocate of objectual quantification, Quine's mistake was to infer from the premise that the two co-extensional terms "9" and "the number of planets" could not be substituted for one another in "□(x ≥ 7)", to the conclusion that the position occupied by the variable "x" is not purely designative. As Kaplan (1986, p. 235) has clearly emphasized, all that follows from the premise is that at least one of the two occurrences of "9" and "the number of planets" in (1) and (2) is not purely designative. Nothing whatever follows about the *position* that these terms occupy. In particular, it may be that "the number of planets" does not have a purely designative occurrence in (1), whereas "9" has a purely designative occurrence in (2), and the variable "x" has a purely designative occurrence in "∃x □(x ≥ 7)".

According to the advocate of objectual quantification, the function of any occurrence of a variable is simply to be purely designative. So the advocate of objectual quantification endorses substitutivity of identity in the form:

(3) $\forall x \forall y(x=y \rightarrow (\Phi(x) \leftrightarrow \Phi(y)))$.

The fact that (1) is false and (2) is true does not in the least threaten the truth of (3). For according to the advocate of objectual quantification, though not all occurrences of terms are purely designative, variables are always purely designative. The fact that (1) and (2) differ in truth-value demonstrates that at least one of the terms "the number of planets" and "9" has a non-purely designative occurrence in "□(x ≥ 7)". But it does not show that (3) is false. Given the objectual interpretation of variables, what (3) expresses is *Leibniz's Law*, which, as Cartwright (1971) clearly shows, is an obviously true metaphysical principle not to be confused with the false principle that any two co-extensional terms (including descriptive terms) can be substituted *salva veritate* in modal contexts.

Another way of thinking of the failure of Quine's argument, emphasized in Stanley (1997a, p. 561), is that the failure of the substitutivity of identity with variables as stated in (3) only follows from the failure of substitutivity with terms (as in (1) and (2)) if we think of the quantifiers *substitutionally*, as allowing for arbitrary substitution of singular terms (including descriptions) for variables. The advocate of objectual quantification rejects this construal of quantification. The reason that Quine construes quantification into modal contexts substitutionally is because his targets interpreted necessity as analyticity (as we have seen with Carnap). Since analyticity is fundamentally a property of sentences, it is natural to construe quantification into an open modal sentence in terms of the analyticity of a sentence with no free variables (Neale, 2000, pp. 302-303). But the advocate of objectual quantification rejects this interpretation of necessity and with it the corresponding non-objectual account of quantification. The natural interpretation of "□" on the objectual construal of quantification is as *metaphysical* necessity; on this interpretation, an object satisfies "□Fx" if and only if that object has F as an essential property. Thus Quine's charge (Quine, 1953) that objectually quantifying into modal contexts involves "Aristotelian essentialism" is partially vindicated; at the very least, the

coherence of essentialist attributions is presupposed by this construal of quantification into a necessity operator interpreted metaphysically.[41]

What, then, of the failure of substitution in pairs such as (1) and (2)? Quite early on, philosophers had started to recognize that Quine's argument seemed to play upon some feature peculiar to definite descriptions as opposed to proper names. In a review of Quine (1947), Arthur Smullyan wrote, with reference to the claim that it is not necessary that the evening star is identical to the morning star:

> We now may ask what sense of the word "constant" is needed in order to justify application of the principle of existential quantification. It is possible that by "constant" is meant what is commonly understood by "proper name". Under this interpretation it appears evident to this reviewer that the principle of existential generalization is true. However, we observe that if "Evening Star" and "Morning Star" proper-name the same individual they are *synonymous* and therefore [the claim is false]. (Smullyan, 1947, p. 140)

So, Smullyan claims that co-extensional terms *used as proper names* are substitutable for one another salva veritate in modal contexts (and in particular, if "Evening Star" and "Morning Star" are used as proper names, it is necessary that evening star is morning star). So it was because Quine was using the terms as descriptions rather than proper names that they were not substitutable for another in modal contexts.[42] In a similar vein, some years later, Ruth Barcan Marcus wrote in her classic 1961 paper "Modalities and Intensional Languages":

> Now, suppose we come upon a statement like
>
> (15) Scott is the author of Waverley.
>
> and we have a decision to make….If we decide that 'the evening star' and 'the morning star' are proper names for the same thing, and that 'Scott' and 'the author of Waverly' are proper names for the same thing, then they must be intersubstitutable in every context. (Barcan Marcus, 1993a, p. 10).

Indeed, Barcan Marcus concludes that:

> What I have been arguing is that to say truly of an identity (in the strongest sense of the word) that it is true, it must be tautologically true or analytically true. (Ibid., p. 12).

So Barcan Marcus maintains that when "a" and "b" are being used as names (presumably logically proper names, in Russell's sense), as opposed to being used as descriptions (Ibid., pp. 10-12), then "a=b" is analytically true, if true at all.[43]

What emerges from the suggestions of Smullyan and Barcan Marcus is that there is a distinction between *using terms as proper names* and *using terms as descriptions*. If one

uses two terms as names, and they refer to the same object, then the two terms are synonymous, and the identity is analytically true. As a result, substitution is permitted even in modal contexts. So, if "the number of planets" is being used as a proper name of 9, then (1) and (2) are both true and one of Quine's premises (the falsity of (1)) is undermined. On the other hand, if (1) is false, then "the number of planets" is being used as a description and not a name, and (1) is not a genuine identity after all.

The problem with Smullyan and Barcan Marcus's suggestions is that it is quite implausible to take a sentence such as Barcan Marcus's (15) to be analytically true, and it is equally implausible to take "The Evening Star" and "The Morning Star" to be synonymous (and hence "The Evening Star is the Morning Star" to be analytically true). By not clearly disassociating the metaphysical notion of necessity from its epistemic cousins such as analyticity and a priority, Smullyan and Barcan Marcus failed to make plausible the thesis that true identities were necessary.

In 1972, Saul Kripke published Naming and Necessity, which transformed what had been certain abstract formal possibilities essentially into common sense and eventual philosophical orthodoxy. First, Kripke clearly distinguished the metaphysical notions of necessity and contingency from the epistemic notions of a priority and a posteriority (Kripke, 1980, pp. 34ff.). As he points out, the notion of a priority is a concept from epistemology, and means roughly that a statement is knowable independently of experience. Though "necessary" can express an epistemic concept (and indeed, as Kripke points out, can sometimes be used to express the property of a priority), it can also be used to express *metaphysical necessity*, which is a concept that has nothing whatever to do with epistemology, but rather is a concept of *metaphysics*. Again very roughly, a truth is metaphysically necessary if and only if the world could not have been different in such a way as to make that proposition false. There is no prima facie reason to think that a priority, the concept from epistemology, coincides with metaphysical necessity, the concept from metaphysics, and indeed Kripke produces examples of metaphysical necessities that are not a priori and metaphysical contingencies that are a priori.

For example, Kripke (1980, pp. 100-3, pp. 108-9) argues that Barcan Marcus and Smullyan were correct to maintain that true identity statements involving names, such as "Hesperus is Phosphorus" and "Cicero is Tully", are necessarily true. However, Kripke (Ibid., pp. 103-4) rejects Barcan Marcus's thesis that the statement that Hesperus is Phosphorus and the statement that Cicero is Tully are analytic, since he rejects that they are a priori, and construes "analytic statement" to entail that a statement is necessary and a priori.[44] So, true identity statements involving ordinary proper names are, for Kripke, instances of statements that are both necessary and (as Frege pointed out) a posteriori.[45] Kripke also gives other examples of statements that are both necessary and a posteriori besides identity sentences such as "Hesperus is Phosphorus" and "Cicero is Tully". For example, Kripke argues that we have intuitions about the essential properties of things and among the essential properties of such things are their *origins*. A person essentially is the product of the sperm and egg that actually produced her; it makes no sense to think of *the same person* being produced by a different sperm or egg (Ibid., p. 113). If a table is made from a hunk of wood, then it is essentially made from that hunk of wood; it could

not be the very same table and be made from (e.g.) metal. Since it is not a priori what a thing's origins are, such necessities as are for instance expressed by sentences such as "Elizabeth originated from this sperm and this egg" or "This table is made out of this hunk of wood" are both necessary and a posteriori. Finally, Kripke also argues that *theoretical identification statements*, such as "Heat is mean molecular motion", "Water is H20", and "Gold is the element with atomic number 79" are necessary a posteriori. However, his arguments here are more controversial, involving the topic of the "rigidity" of general terms (for an excellent recent introduction to the difficulties here, see Soames (2002, Chapter 9)). Kripke also produces examples of statements that are both contingent and a priori. For example, consider the "standard meter" (henceforth *stick S*) in Paris, which is used to fix the reference of the expression "one meter". Stick S could have been slightly longer or slightly shorter (suppose, for example, that heat had been applied to it). However, "for someone who has fixed the metric system by reference to Stick S" (Ibid., p. 56), the statement "Stick S is one meter long" is a priori. So the statement "Stick S is one meter long", for a person who has fixed the metric system by reference to stick S, is both contingent and a priori.[46]

Kripke also provides what are widely accepted as conclusive arguments against Russell's description theory of ordinary proper names. Since Nathan Salmon's discussion in Reference and Essence (Salmon, 1981, pp. 23ff.) it has been standard to distinguish between three sorts of arguments provided by Kripke against the description theory of ordinary proper names; the *modal* argument, the *epistemological* argument, and the *semantic* argument. According to the modal argument against the description theory of proper names, proper names are *rigid designators*, where a designator N of an object o is rigid if and only if N designates o in all possible worlds w in which o exists, and N designates nothing other than o in all possible worlds in which o does not exist. In contrast, the descriptions that plausibly give the meaning of ordinary proper names are not rigid designators. For example, "the last great philosopher of antiquity" plausibly gives the meaning of the name "Aristotle", if any description does. But an utterance of the sentence "Aristotle is the last great philosopher of antiquity" expresses a contingent proposition. The designator "the last great philosopher of antiquity" designates other people than Aristotle relative to some possible worlds in which Aristotle exists; for example, relative to a possible world in which Aristotle did not write any philosophy at all, it designates Plato. So "the last great philosopher of antiquity" is not a rigid designator of Aristotle, while Aristotle is a rigid designator of Aristotle (for a detailed account of the empirical argument for the thesis that names are rigid designators, see Stanley (1997a, pp. 565ff.)).[47] According to the *epistemological argument* against the description theory of names, sentences containing names and the descriptions that supposedly give their meaning are not a priori true, which they should be if the descriptions are synonymous with those names. For example, "Aristotle is the last great philosopher of antiquity" is not an a priori truth, which it should be if "Aristotle" was really the covert definite description "the last great philosopher of antiquity". Finally, according to the *semantic argument* against the description theory of names, someone can still use a name to refer to an object, even if they are completely unaware of the reference-fixing description.[48]

So, according to Kripke, names are not covert definite descriptions. Kripke argues (1980, p. 78) that names are not just rigid designators, but are (in the vocabulary of Salmon (1981)), *obstinate* rigid designators, in the sense that a name refers to the same object relative to every possible world, including worlds in which that object does not exist. If so, then names behave under modal operators *exactly as variables relative to an assignment* according to the objectual interpretation of quantification into modal contexts. On this interpretation of de re modal quantification, relative to a sequence, the occurrence of the variable "x" in the open modal formula "□Fx" has the same value relative to any possible world. Where the sequence s assigns the object o to "x", the value of "x" relative to any possible world is just o. We are to think of "□Fx" being satisfied by an assignment s that assigns the object o to the variable "x" if and only if the object o is F in every possible world. Kripke's point is that names behave in just the same way under modal operators, *whereas definite descriptions do not*. Quine's logical argument against the coherence of de re modal attributions fails, because (like co-extensional variables relative to an assignment) co-extensional names are substitutable in modal contexts. The fact that a definite description such as "the number of planets" cannot be substituted for a co-extensional expression within the scope of a modal operator does not entail that the *position* in which that definite description occurs blocks substitution. It has rather to do with a feature of *definite descriptions*, namely that definite descriptions (unlike names and variables with respect to an assignment) are not rigid designators.

According to Quine, our intuitions about de re modal statements are inconstant, fluctuating depending upon the way we think of an object. If we think of 9 as the number of planets, then it is not necessarily odd, whereas if we think of 9 as the number 9, it is necessarily odd. Kripke argues that our intuitions about de re modal statements are not context-sensitive. A genuine de re modal attribution attributes a property to an object essentially (and even clearer proponent of this position is Plantinga (1974)). Quine's arguments for the inconstancy of de re modal attributions result from not clearly distinguishing de dicto modal statements with definite descriptions from de dicto modal statements with names.[49]

David Lewis presented a very different response to Quine's objections to quantified modal logic. Lewis was a realist about different possible worlds, and believed that no object existed in more than one possible world. Lewis's metaphysical view about possible worlds prevented him from accepting the objectual interpretation of de re modal quantification, because the objectual interpretation requires making sense of an object in the actual world existing and having properties at other possible worlds. Instead, Lewis (1968) proposed what he called a *counterpart theoretic* interpretation of quantified modal logic.[50] According to counterpart theory, a de re modal sentence is true in virtue of *counterparts* of actual objects having properties in other possible worlds. Unlike Kripke, Lewis agrees with Quine's premise that our intuitions about de re modal attributions were inconstant, that is, fluctuated with context.[51] In fact, Lewis thought that this feature of de re modal attributions was the key to solving a number of classical metaphysical problems. But Lewis took the inconstancy of de re modal attributions as evidence that they called for a *context-sensitive semantic theory*. More specifically, Lewis thought that there were distinct *counterpart relations*. An open modal formula, such as "□Fx" is satisfied by an

object o if and only if each counterpart c of o in any world w has the property F. Since there are distinct counterpart relations, "□Fx" may be true of an object relative to one counterpart relation and false of an object relative to another counterpart relation (since different sets of counterparts are determined by the different counterpart relations evoked in different contexts). So, Lewis accepts Quine's claim that de re modal attributions are inconstant. But he does not think that this shows that quantified modal logic is not formally intractable. Instead, Lewis incorporates the contextual relativity into the formal semantics itself.

Lewis's account of quantified modal logic is an instructive example to consider in light of Quine's skepticism about the possibility of regimenting de re modal statements. Quine's skepticism about formalizing de re modal statements (and de re attitude ascriptions) arises in part from his belief that our intuitions about these constructions are context-sensitive, and this context-sensitivity is an impediment to regimentation. However, Lewis's reaction to the apparent context-sensitivity of de re modal sentences is to incorporate the context-sensitivity into the regimentation. What this shows is that the alleged context-sensitivity of a kind of discourse, far from serving as an impediment to regimentation, is simply further fodder for it.[52]

X. Conditionals

Grice's response to the ordinary language philosopher involved defending the material conditional analysis of the indicative conditional. By providing a pragmatic account of the conflicting data, Grice hoped to dispel the challenge to giving a truth-functional analysis of the indicative conditional. But Grice's defense of the material conditional analysis is deeply problematic. For example, Grice's analysis predicts that if one has a high degree of credence in either the negation of the antecedent or the truth of the consequent, the conditional is not assertible. But, as Frank Jackson has emphasized (1987, p. 20), there are many conditionals that are highly assertible even though we have a high degree of credence in the falsity of the antecedent or the truth of the consequent, such as 'If the sun goes out of existence in ten minutes' time, the earth will be plunged into darkness in about eighteen minutes time', or, if we are convinced that Bekele will win the race, 'If Webb runs, Bekele will win the race, and if Webb doesn't run, Bekele will win the race'. In the former case, we are as certain of the falsity of the antecedent as we are of the truth of the conditional, yet it is highly assertible; in the latter case, we are certain of the truth of the consequent, yet the conditional remains assertible. Finally, Grice's theory of the indicative conditional predicts that logically equivalent statements to the material conditional have the same truth-conditions, and are equally assertible. For example, Grice's theory predicts that "If A then B" should have the same truth conditions and be equally assertible as "If not B, then not A". But this prediction is not borne out (Bennett, 2003, p. 32). Grice's defense of the conditional postulates a large gap between standard use of conditionals and their meanings, a gap that he tries to cover with explanations from general conversational principles. Unfortunately, the conversational principles do not succeed in explaining the gap between indicative conditionals, construed as material conditionals, and their standard uses. Another theory is required.

Grice's focus was on indicative conditionals in natural language. In the 1960s, philosophers influenced by theories of meaning for modal languages turned their attention to subjunctive conditionals in natural language. Characterizing precisely the distinction between indicative and subjunctive conditionals is a thorny matter. But the basic contrast between the two classes of conditionals is brought out in the classic pair:

(a) If Oswald didn't shoot Kennedy, someone else did.
(b) If Oswald hadn't shot Kennedy, someone else would have.

The indicative conditional in (a) is true, but the subjunctive conditional in (b) is probably not true. Subjunctive conditionals generally (or always) contain modal terms in their consequents.

The first published account of a modal semantics for subjunctive conditionals was presented in Stalnaker (1968). According to Stalnaker, a subjunctive conditional "If A were the case, then B would be the case" is true at a possible world w if and only if B is true at the closest possible world to w at which A is true. Stalnaker's original analysis is both simple and elegant. But it also has some dramatic consequences. Call an "A-world" a world in which the proposition A is true. According to Stalnaker's theory, whenever a subjunctive conditional "If A were the case, then B would be the case" has a truth-value, there is always a unique closest A-world to the world of evaluation. As a consequence, Stalnaker's theory also validates *conditional excluded middle* (CEM), the principle that "Either if A were the case, then B would be the case, or if A were the case, ~B would be the case". Both of these consequences have been widely held to be problematic (according to David Lewis (1973, p. 79), the validation of CEM is "The principle virtue and the principle vice of Stalnaker's theory"). David Lewis (1973) presented a somewhat more complicated theory of subjunctive conditionals that does not involve the hypothesis of a closest A-world in evaluating the truth of a subjunctive conditional, and does not validate CEM.

Stalnaker's 1968 theory of conditionals, despite being "constructed primarily to account for counterfactual conditionals" (Stalnaker, 1999b, p. 68), was "intended to fit conditional sentences generally, without regard to the attitude taken by the speaker to the antecedent or consequent or his purpose in uttering them, and without regard to the grammatical mood in which the conditional is expressed." (Ibid.) According to this analysis, a conditional "If A, then B" is true if and only if B is true in the most similar A worlds. In the case of indicative conditionals, there is a pragmatic principle governing the context-dependent notion of similarity; worlds that are assumed to be live epistemic possibilities are most similar. Of course, the actual world is always the most similar world to itself, so if the antecedent of a conditional is true, the conditional is true if and only if its consequent is true. But if the antecedent is false, then the truth of the conditional will depend upon the truth of the consequent in the most similar epistemically possible world in which the antecedent is true. Thus, Stalnaker's analysis of conditionals elegantly explains the fact that natural languages employ the same expression to formulate indicative and subjunctive conditionals. Furthermore, Stalnaker's analysis predicts that indicative conditionals are context-sensitive constructions, since they depend

for their truth-value on a parameter that shifts with context, namely the metric of similarity. Stalnaker's theory gives a *semantic* explanation for some of the distinctions that Grice's theory was powerless to explain. For example, Stalnaker's theory predicts that "If A, then B" does not have the same truth-conditions as "If ~B, then ~A". Stalnaker's theory also explains much of the contextual variability in our intuitions about the truth-values of indicative conditionals.

Grice attempted to give a simple semantic analysis of the indicative conditional, and a pragmatic explanation of the divergence between the ordinary uses of conditional sentences and their actual truth-conditions. However, Grice's explanation left an implausibly large gap between our intuitions about the truth-conditions of indicative conditionals, and their semantic content, a gap that developments of similar views have arguably also failed to bridge (see Jackson (1987), and the criticism of Jackson in Bennett (2003, pp. 38ff.)). Stalnaker, in contrast, exploits the tools of intensional logic to provide a context-sensitive semantic theory for indicative conditionals, one that exploits more complex logical mechanisms to bring the semantic content closer to what it intuitively seems to be. The topic of conditionals is another area in which the central dispute is between sophisticated semantic theories that capture intuitive data by incorporating context-sensitivity, and semantic theories that eschew more complex mechanisms and context-sensitivity in favor of attempted pragmatic explanations of the intuitions.

The debate on conditionals has rightfully attracted more attention than other debates in the philosophy of language in the last thirty years, not just because of the centrality of the construction to so many areas of thought, but also because the issues have been complicated by several surprising facts about them. The goal of Stalnaker's semantics for indicative conditionals was to bring the truth-conditions of these constructions closer to what they intuitively seem to be. One arguably intuitive claim about the use of indicative conditionals concerns their *probability*. Generally, people wish to assert propositions that they believe are very likely to be true, which suggests the general thesis that something is assertible for a speaker at a time if and only if it has a high subjective probability for that person. If so, then a conditional is assertible if and only if it has a high subjective probability. It is quite intuitive to take the probability of the indicative conditional 'if A, then C' to be the conditional probability of C given A. This suggests that a semantic theory for indicative conditionals that accords with their use conditions should have as a consequence that the probability of an indicative conditional 'if A, then C' for a person at a time should be equivalent to the conditional probability of C given A. But David Lewis (1976) proved the surprising result that no connective O that links propositions could have the property that the probability of O(A, C) is the conditional probability of C given A, and simultaneously yield a satisfactory (or even close to satisfactory) account of our ordinary intuitions about the probabilities of various conditionals. If having an account of the meaning of indicative conditionals that matches intuitive judgments about their content requires such a connective, then the search is futile. A second fact about indicative conditionals, emphasized in Gibbard (1981), is that they are extremely context-sensitive; one speaker can assert 'if A, then C', and intuitively be correct, and another speaker can assert 'if A, then ~C' to describe the same situation, and also intuitively be correct. Another fact about indicative conditionals (also emphasized in Gibbard (1981)) is

that it is not easy to embed indicative conditionals inside other conditionals. A number of distinguished philosophers have used these considerations, along with others, to motivate the view that indicative conditionals lack truth-conditions altogether.[53] On this view of their meaning, indicative conditionals are an example of one kind of non-normative sentence for which something like the model of meaning endorsed by expressivists about moral discourse is correct.[54]

XI. Conclusion

In the 1960s and 1970s, philosophers started to exploit the resources of semantic theories for formal languages in the analysis of natural language meaning. A formal language differs from a natural language in having a simple, clearly defined syntax. To avoid the complexities of natural language grammar, many of these philosophers gave semantic theories for fragments of natural language regimented in various extensions of the language of first-order predicate logic (such as the language of quantified modal logic, or the language of intensional logic). But of course what we interpret when we understand sentences of natural language are the structures of those sentences, not the sentences of some regimented formal language. So the relevance for the project of giving a theory of meaning for natural languages of semantic theories for various extensions of the language of first-order logic is not completely clear.

However, in the 1960s, work by linguists, in particular the linguist Noam Chomsky, began to show that natural languages, like formal languages, had grammars that could be described formally. Chomsky's work made the project of transferring the tools of the logician to the analysis of meaning considerably more tractable. If natural languages have a systematic syntax, then there is no obstacle to mimicking the formal semantic project directly for natural languages. Using the research of contemporary syntax, one could represent what the objects of natural language interpretation were, using the tools of semantics one could interpret them, and using the norms of discourse described by Grice, one could explain divergences between use and meaning. It took a number of years for philosophers to absorb the lessons of syntax. But since an apparent fact about meaning may be due either to the syntax of a given sentence, its semantics, or general facts about language use, the contemporary philosopher of language must master all three branches of investigation.

The discovery that the notions of reference and truth could be used to give a theory of meaning for natural language, together with the twin developments of syntax and pragmatics, have resolved many of the foundational disputes of mid-century philosophy of language. It is difficult to argue that context-sensitivity undermines the project of giving a systematic theory of reference and truth for natural language when the best models of context-sensitivity appeal to reference in giving the meanings of context-sensitive expressions (as in Kaplan's notion of character). It is difficult to argue that vagueness undermines this project, when sophisticated semantic theories for vague expressions have been developed (e.g. the supervaluational semantics for vagueness developed in Fine (1975)). The very features that are used to cast doubt on the possibility of formalization are always the next challenge for the project of formalization. As a

result, attention has shifted to giving the meaning of particular philosophically interesting constructions in natural-language (of which the conditionals literature is but an especially interesting example). To be sure, in some of these literatures, philosophers frustrated with the intractability of the problems posed by the relevant constructions have tried to draw broader morals, and sometimes within this framework pleas to return to the pessimistic attitudes towards the prospect of a systematic theory of meaning have been advanced (most notably Schiffer's influential (1987)). Nevertheless, cries of frustration with the difficulty of particular constructions have not been met with widespread defeatism, and the program of giving a systematic account of the theory of meaning has continued.

Once many of the foundational issues were settled, a vast amount of work was produced in philosophy of language and (especially) its close relative in linguistics, the field of semantics. It is impossible even to provide a road map to the wealth of work that has been done in the last thirty years on adverbs, anaphora, determiners, mass terms, plurals, adjectives and gradability, modals, tense, aspect, and other topics. In terms of details, since the 1970s much of that sub-part of the investigation of natural language meaning that has been conducted by philosophers been devoted to detailed arguments with respect to various constructions about whether a Gricean response can account for the phenomena that goes beyond a simple semantic analysis, or whether a more complex semantic theory that incorporates context-sensitivity semantically (as in Stalnaker's analysis of conditionals) is plausible. For every construction we have discussed, there are advocates of each view. Predictably, some of these disputes take place on a meta-level, with advocates of a non-semantic account of the phenomena arguing that a Gricean or quasi-Gricean apparatus does much more explaining than is ordinarily recognized, and advocates of semantic accounts arguing for greater attentiveness to the nuances of natural language meaning and form. But there is an overarching agreement even between most disputants at this meta-level – the overarching agreement is meaning and use should never be conflated, and that any adequate account of meaning fundamentally employs the notions of reference and truth.

Bibliography

Adams, Ernest (1975): The Logic of Conditionals (Dordrecht: Reidel).
Azzouni, Jody (2006): Tracking Reason: Proof, Consequence, and Truth (Oxford: Oxford University Press).
Barcan, Ruth (1947): "The Identity of Individuals In A Strict Functional Calculus of Second Order", *The Journal of Symbolic Logic* **12.1**: 12-15.
Barcan Marcus, Ruth (1993): Modalities: Philosophical Essays (Oxford: Oxford University Press).
Barcan Marcus, Ruth (1993a): "Modalities and Intensional Languages", in Barcan Marcus (1993a): 5-23 (first published in 1961).
Bennett, Jonathan (2003): A Philosophical Guide to Conditionals (Oxford: Oxford University Press).
Boghossian, Paul (1997): "Analyticity", in Hale and Wright (1997): 331-368.
Brandom, Robert (1994): Making it Explicit (Boston: Harvard University Press).

Brentano, Franz (1995): Psychology from an Empirical Standpoint (London: Routledge) (First published 1874).
Burge, Tyler (1979) "Frege and the Hierarchy", *Synthese* **40**: 265-81.
Carnap, Rudolf (1949): The Logical Syntax of Language (London: Routledge & Kegan Paul) (first published in English in 1937).
Carnap, Rudolf (1958) Meaning and Necessity: A Study in Semantics and Modal Logic (Chicago: University of Chicago Press). (first published 1947).
Carnap, Rudolf (1958a): "Empiricism, Semantics, and Ontology", in Carnap (1958): 205-221.
Carnap, Rudolf (1958b): "Meaning Postulates", in Carnap (1958): 222-229.
Chomsky, Noam (1969): "Quine's Empirical Assumptions", in Davidson and Hintikka (1969).
Chomsky, Noam (1975): Reflections on Language (New York: Pantheon Books).
Church, Alonzo (1950): "On Carnap's analysis of statements of assertion and belief", *Analysis* **10**: 97-99.
Church, Alonzo (1951): "A Formulation of the Logic of Sense and Denotation", in Structure, Method, and Meaning: Essays in Honor of H.M. Sheffer, ed. By P. Henle, H.M. Kallen, and S.K. Langer (New York: Liberal Arts Press).
Davidson, Donald (1967): "Truth and Meaning", *Synthese* **17**: 304-23. Also reprinted in Davidson (1990).
Davidson, Donald (1990): Inquiries into Truth & Interpretation (Clarendon Press: Oxford 1990).
Davidson, Donald (1990a): "Theories of Meaning and Learnable Languages", in Davidson [1990]: 3-15.
Davidson, Donald and Hintikka, Jakko (1969): Words and Objections: Essays on the Work of W. V. Quine (Dordrecht: Reidel).
Dummett, Michael (1981): Frege: Philosophy of Language (London: Duckworth).
Edgington, Dorothy (1986): "Do Conditionals Have Truth-Conditions?"
Evans, Gareth (1981) "Understanding Demonstratives", in Meaning and Understanding, H. Parret and Jacques Bouveresse (eds.) (Berlin: W. de Gruyter).
Evans, Gareth (1982): The Varieties of Reference (Oxford: Clarendon Press).
Fara, Michael and Williamson, Timothy (2005): "Counterparts and Actuality" *Mind* **114**: pp. 1-30.
Field, Hartry (1972): "Tarski's Theory of Truth", *Journal of Philosophy* **69**: 347-75.
Fine, Kit (1975). "Vagueness, truth and logic", *Synthese* **54**: 235-59.
Fine, Kit (2005): Modality and Tense (Oxford: Oxford University Press).
Fine, Kit (2005a): "The Problem of De Re Modality", in Fine (2005): 40-104.
Fine, Kit (2005b): "Quine on Quantifying In", in Fine (2005): 105-30.
Forbes, Graeme (2000): "Objectual Attitudes", *Linguistics and Philosophy* **23.2**: 141–183.
Frege, Gottlob (1879): Begriffsschrift, in Gottlob Frege (1988), Begriffsschrift und andere Aufsaetze, ed. By Ignacio Angelelli (Hildesheim: Georg Olms Verlag).
Frege, Gottlob (1966): Grundgesetze der Arithmetik (Hildesheim: Georg Olms) [first published 1892).
Frege, Gottlob (1980a): Gottlob Freges Briefwechsel, ed. G. Gabriel, F. Kambartel, and C. Thiel (Felix Meiner Verlag).

Frege, Gottlob (1980b): The Foundations of Arithmetic (Evanston: Northwestern University Press) (First published 1884).
Frege, Gottlob (1993): "On Sense and Meaning", in Geach, Peter and Black, Max Translations from the Philosophical Writings of Gottlob Frege (Oxford: Basil Blackwell): 56-78.
Gibbard, Allan (1981): "Two Recent Theories of Conditionals", in Harper, W., Stalnaker, R., and Pearce, G. (eds.) Ifs (Dordrecht: Reidel): 211-47.
Graff, Delia (2001): "Descriptions as Predicates," *Philosophical Studies* **102**(1): 1–42.
Hale, Bob and Wright, Crispin (1997): A Companion to the Philosophy of Language (Oxford: Blackwell Press).
Grice, Paul (1989): Studies in the Way of Words (Cambridge: Harvard University Press).
Grice, Paul (1989a): "Logic and Conversation", in Grice (1989): 22-40.
Grice, Paul (1989b): "Further Notes on Logic and Conversation", in Grice (1989): 41-57.
Grice, Paul (1989c): "Indicative Conditionals", in Grice (1989): 58-85.
Hazen, Alan (1979): "Counterpart Theoretic Semantics for Modal Logic", *Journal of Philosophy* **76**: 319-338.
Heck, Richard (1993): "The Development of Arithmetic in Frege's Grundgesetze der Arithmetik", Journal of Symbolic Logic 58: 579-602.
Heck, Richard (1998): "*Grundgesetze der Arithmetik* I §§29-32", *Notre Dame Journal of Formal Logic* 38: 437-74.
Heim, Irene and Kratzer, Angelika (1998): Semantics in Generative Grammar (Oxford: Blackwell Press).
Higginbotham, James (1988): "Contexts, Models, and Meanings: A Note on the data of semantics", in Mental Representations: The interface between language and reality (ed. By Ruth Kempson) (Cambridge: Cambridge University Press): 29-48.
Hintikka, Jaakko (1962): Knowledge and Belief: An Introduction to the Logic of the Two Notions (Ithaca: Cornell University Press).
Jackson, Frank (1987): Conditionals (Oxford: Basil Blackwell).
Kaplan, David (1968): "Quantifying In", *Synthese* **19**: 178-214.
Kaplan, David (1986): "Opacity", in The Philosophy of W.V. Quine, edited by Lewis Edwin Hahn and P.A. Schilpp (LaSalle, Open Court): 229-288.
Kamp, Hans (1971): "Formal Properties of 'Now'", *Theoria* 37.3: 227-273.
King, Jeffrey (2003): "Tense, Modality, and Semantic Value", in John Hawthorne and Dean Zimmerman (ed.), *Language and Philosophical Linguistics, Philosophical Perspectives* **17** (Oxford: Blackwell).
Kripke, Saul (1963): "Semantical Considerations on Modal Logic", *Acta Philosophica Fennica*, 16: 83-94.
Kripke, Saul (1980): Naming and Necessity (Cambridge: Harvard University Press).
Larson, Richard, Den Dikken, Marcel, and Ludlow, Peter (forthcoming): "Intensional Transitive Verbs and Abstract Clausal Complementation." *Linguistic Inquiry*.
Lepore, Ernest (1982): "What Model Theoretic Semantics Cannot Do," *Synthese*, 54: 167-187.
Lewis, David (1968): "Counterpart Theory and Quantified Modal Logic", *Journal of Philosophy* **65**: 113-26.
Lewis, David (1971): "Counterparts of Persons and Their Bodies", *Journal of Philosophy* **68**: 203-11.

Lewis, David (1973): Counterfactuals (Cambridge: Harvard University Press).
Lewis, David (1976): "Probabilities of Conditionals and Conditional Probabilities", *Philosophical Review* **85**: 297-315.
Lewis, David (1979): "Scorekeeping in a Language Game", *Journal of Philosophical Logic* **8**: 339-59.
Lewis, David (1981): Index, Context, and Content. In Philosophy and Grammar: Papers on the Occasion of the Quincentennial of Uppsala University. Edited by Stig Kanger and Sven Ohman (D. Reidel, Dordrechet): 79-100.
Meinong, Alexius (1910): Ueber Annahmen.
Montague, Richard (1974): Formal Philosophy (New Haven: Yale University Press).
Montague, Richard (1974a): "Pragmatics", in Montague (1974): 95-118.
Montague, Richard (1974b): "On the Nature of Certain Philosophical Entities", in Montague (1974): 148-187.
Montague, Richard (1974c): "On the Proper Treatment of Quantification in Ordinary English", in Montague (1974): 247-270.
Moore, G.E. (1899): "The Nature of Judgment", *Mind* 5.8: 176-93.
Moore, G.E. (1910): Some Main Problems of Philosophy.
Neale, Stephen (1990): Descriptions (Cambridge: MIT Press).
Neale, Stephen (2000): "On A Milestone of Empiricism", in Alex Orenstein and Petr Kotatko (eds.), Knowledge, Language, and Logic (London: Kluwer): 237-346.
Parsons, Terry (1969): "Essentialism and Quantified Modal Logic", *The Philosophical Review* **78.1**: 35-42.
Parsons, Terry (1981): "Frege's Hierarchies of Indirect Sense and the Paradox of Analysis", *Midwest Studies in Philosophy* 6: 37-57.
Partee, Barbara (1973): "Some Structural Analogies between Tenses and Pronouns in English", *Journal of Philosophy* 70: 601-9.
Partee, B., 1974, "Opacity and Scope", in *Semantics and Philosophy*, M. Munitz and P. Unger (ed.), New York: NYU Press: 81–101.
Perry, John (1977) "Frege on Demonstratives", *The Philosophical Review* 86.4: 474-97.
Plantinga, Alvin (1974): The Nature of Necessity (Oxford: Clarendon Press).
Putnam, Hilary (1975): Mind, Language, and Reality: Philosophical Papers, Volume 2 (Cambridge: Cambridge University Press).
Putnam, Hilary (1975a): "Do True Assertions Correspond to Reality?", in Putnam (1975): 70-84.
Quine, W.V.O. (1943): "Notes on Existence and Necessity", *Journal of Philosophy* 40: 113-27.
Quine, W.V.O. (1947): "The Problem of Interpreting Modal Logic", *Journal of Symbolic Logic* 12: 43-8.
Quine, W.V.O. (1953): "Three Grades of Modal Involvement", in Quine (1976): 158-176.
Quine, W.V.O. (1955): "Quantifiers and Propositional Attitudes", in Quine (1976): 185-96.
Quine, W.V.O. (1960): Word and Object (Cambridge: MIT Press).
Quine, W.V.O. (1961): "From A Logical Point of View" (Cambridge: Harvard University Press).

Quine, W.V.O. (1961a): "Two Dogmas of Empiricism", reprinted in Quine (1961): 20-46. [First published in *The Philosophical Review*, 1951].
Quine, W.V.O. (1976): The Ways of Paradox and other Essays (Cambridge: Harvard University Press).
Quine, W.V.O. (1976): "Worlds Away", *The Journal of Philosophy* 73.22: 859-863.
Quine, W.V.O. (1977): "Intensions Revisited", *Midwest Studies in Philosophy* **2**: 5-11.
Richard, Mark (1987): "Quantification and Leibniz's Law", *Philosophical Review* **96**: 555-78.
Richard, Mark (1990): Propositional Attitudes: An Essay on Thoughts and How we Ascribe Them (Cambridge: Cambridge University Press).
Richard, Mark (2001): "Seeking a Centaur, Adoring Adonis: Intensional Transitives and Empty Terms", in *Figurative Language: Midwest Studies in Philosophy Volume 25*, P. French and H. Wettstein (ed.) (Oxford: Basil Blackwell): 103–127.
Russell, Bertrand (1919): Introduction to Mathematical Philosophy.
Russell, Bertrand (1996): The Principles of Mathematics (New York: Norton & Company) [First published 1903].
Russell, Bertand (1905): "On Denoting"
Russell, Bertrand and Whitehead, Alfred (1910): Principia Mathematica (Cambridge: Cambridge University Press).
Russell, Bertrand (1919): Introduction to Mathematical Philosophy (New York: Simon and Schuster).
Russell, Bertrand (1988): The Problems of Philosophy (Buffalo: Prometheus Books).
Russell, Bertrand (1994): "On the Nature of Truth and Falsehood", in his Philosophical Essays (London: Routledge): 147-59.
Russell, Bertrand (1985): The Philosophy of Logical Atomism (La Salle: Open Court).
Smullyan, Arthus Francis (1947): "Review of W. V. Quine, The Problem of Interpreting Modal Logic", *The Journal of Symbolic Logic* **12**: 139-141.
Salmon, Nathan (1981): Reference and Essence (Princeton: Princeton University Press).
Schiffer, S. (1987) Remnants of Meaning. Cambridge, Massachusetts: MIT Press.
Smullyan, Arthur Francis (1948): "Modality and Description", *The Journal of Symbolic Logic* **13.1**: 31-37.
Soames, Scott (2002): Beyond Rigidity: The Unfinished Semantic Agenda of Naming and Necessity (New York: Oxford University Press).
Sosa, David (2001): "Rigidity in the Scope of Russell's Theory", *Nous* **35**:1-38.
Stalnaker, Robert (1968): "A Theory of Conditionals" in *Studies in Logical Theory* (N. Rescher, ed.), Oxford, 1968, 98-112.
Stalnaker, Robert (1984): Inquiry (Cambridge: MIT Press).
Stalnaker, Robert (1999): Context and Content (Oxford: Oxford University Press).
Stalnaker, Robert (1999a): "Pragmatics", in Stalnaker (1999): 31-46. [First published in *Synthese* 1970].
Stalnaker, Robert (1999b): "Indicative Conditionals", in Stalnaker (1999): 63-77. [First published in *Philosophia* 1975].
Stalnaker, Robert (2003): Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays (Oxford: Oxford University Press).
Stalnaker, Robert (2003a): "Anti-Essentialism", in Stalnaker (2003): 71-85.

Stalnaker, Robert (2003b): "The Interaction of Modality with Quantification and Identity", in Stalnaker (2003): 144-59.
Stanley, Jason (1996): "Truth and Metatheory in Frege", *Pacific Philosophical Quarterly* 77: 45-70.
Stanley, Jason (1997a): "Names and Rigid Designation", in Hale and Wright (1997): 555-585.
Stanley, Jason (1997b): "Rigidity and Content", in Language, Thought, and Logic: Essays in Honor of Michael Dummett, R. Heck, ed. (Oxford: Oxford University Press, 1997): 131-156.
Stanley, Jason (1999): "Understanding, Context-Relativity, and the Description Theory", *Analysis* **59**: 14-18.
Stanley, Jason (2002): "Modality and What is Said", *Philosophical Perspectives 16: Language and Mind,* J. Tomberlin, ed. (Blackwell Press, 2002): 321-44.
Strawson, Peter (1952): Introduction to Logical Theory (London: Metheun).
Strawson, Peter (1996): "On Referring", in The Philosophy of Language, Third Edition, A.P. Martinich, ed. (New York: Oxford University Press): 208-230.
Szabo, Zoltan (2000): 'Descriptions and Uniqueness.' *Philosophical Studies* **101**: 29 – 57.
Tarski, Alfred (1983): Logic, Semantics, Meta-Mathematics (ed. By John Corcoran, translated by J.H. Woodger) (Indianapolis: Hackett Publishing).
Tarski, Alfred (1983a): "The Establishment of Scientific Semantics", in Tarski (1983): 401-8.
Tarski, Alfred (1983b): "The Concept of Truth in Formalized Languages", in Tarski (1983): 152-278.
Zimmerman, T. E. (1993): "On the Proper Treatment of Opacity in Certain Verbs", *Natural Language Semantics*, 1: 149–179.

---

[1] I have discussed this topic at length in my (1996); see especially section IV.
[2] Frege (1966), section 32 (my translation).
[3] The problem of cognitive significance is not specifically a problem about identity sentences, though it is often misleadingly presented as such. Another version of the same problem can be raised by the difference in cognitive significance between "Hesperus is a planet" and "Phosphorus is a planet".
[4] As we shall see below in the discussion of Carnap, there are two kinds of opaque contexts, contexts that are intensional and contexts that are neither intensional nor extensional.
[5] These difficulties are described in some detail in Perry (1977). For an influential reply to Perry on behalf of (a somewhat psychologized version of) Frege, see Evans (1981).
[6] I say "at least two", because of the problem of the "hierarchy of senses", which will be discussed below.
[7] This is presumably not accidental. One way of thinking of Frege's treatment of propositional attitude verbs is that they induce a kind of systematic ambiguity, in which a term has one referent in one context (not under the scope of a propositional attitude verb), and another referent in another context (under the scope of a propositional attitude verb). Since Frege regards ambiguity as a defect of natural languages, there is little wonder that he would want to minimize the appearance of ambiguity-inducing operators into his

formal language (if he did, he would also have to modify the language to include names of senses, so as to avoid ambiguity; see Frege's letter to Russell of December 28, 2002, reprinted in Frege (1980a, pp. 82-5), especially p. 84).

[8] As Dummett (1981, p. 227) writes, in explaining this methodology "…even when Frege is purporting to give the sense of a word or symbol, what he actually states is what its reference is…The sense of an expression is the mode of presentation of the referent: in saying what the referent is, we have to choose a particular way of saying this…In a case in which we are concerned to convey, or stipulate, the sense of the expression, we shall choose that means of stating what the referent is which displays the sense: we might here borrow a famous pair of terms from the Tractatus and say that, for Frege, we *say* what the referent of a word is, and thereby *show* what its sense is."

[9] This is my translation. Frege would of course dispute Russell's contention that thoughts are "psychologically private matters", since Frege takes senses (including thoughts, the senses of sentences) to be objective, mind-independent entities.

[10] A distant relative of this claim survives in contemporary philosophy of language and mind, under the name "Russell's Principle". See Evans (1982) for discussion.

[11] However, Russell remained rather liberal about acquaintance with universals (see Russell (1988, Chapter X)).

[12] In Russell (1905), Russell was not yet fully fluent with the apparatus of quantifiers, and employed the primitive predicate of propositional functions, "is always true". The major contemporary defense of Russell's account of definite descriptions is Neale (1990). One influential recent criticism of Russell (Graff, 2001) involves the topic of predicative uses of definite descriptions, as in "Napoleon was the greatest French general", which pose certain problems for Russell's theory.

[13] There is a large literature challenging the uniqueness clause involved in Russell's theory. The classical attack on Russell's claim that descriptions involve uniqueness is Strawson (1996); Strawson argues that an utterance of "The table is covered with books" can be true even though we are perfectly aware that there is more than one table in the universe, and so Russell's uniqueness clause fails. Influential contemporary challenges to Russell's claim that descriptions involve uniqueness include Lewis (1979, example 3) and Szabo (2000).

[14] In Chapter 7 of Brentano [1995], he rejects the view that judgment differs from presentation in that the former always has complex entities (like propositions) as objects, and the latter can have simple entities as objects. For example, Brentano (Chapter 7, section 5) argues that some judgments have simple entities as objects; in particular, denials and affirmations of existence. Interestingly, as we have seen, the theory of Moore (1899) also has as a consequence that the judgment that A exists only has A as its object, because the object A is identified with the existential proposition that A exists.

[15] Frege and Russell never took the modal concepts of possibility and necessity to be very central; see for example Russell's discussion of "possible" in Chapter 7 of Russell (1985).

[16] This occurs in sections 29-32 of the Grundgesetze (see Heck (1998)). As Frege writes about this attempted proof in his famous letter to Russell of 6/22/02 (Frege (1980a, p. 61)), "Es scheint danach [after the discovery of the paradoxes], dass die Umwandlung der Allgemeinheit einer Gleichheit in eine Werthverlaufsgleichheit…nicht immer erlaubt ist,

dass mein Gesetz V…falsch ist und dass meine Ausfuehrungen im section 31 nicht genuegen, in allen Faellen meinen Zeichenverbindungen eine Bedeutung zu sichern." ["It appears after this, that the transformation of a generality of an identity into an identity of courses-of-value is not always allowed, that my Basic Law V is false, and that my explanations in section 31 are not sufficient, to secure a reference for my expressions in all cases."; my translation]. In other words, Frege's initial reaction to the paradox is to observe that his attempted consistency proof in section 31 fails.

[17] [citation]

[18] As Tarski summarizes the situation(1983a, p. 401):

> Concepts from the domain of semantics have traditionally played a prominent part in the discussions of philosophers, logicians, and philologists. Nevertheless, they have long been regarded with a certain skepticism. From the historical point of view, this skepticism is well-founded; for although the content of the semantical concepts, as they occur in colloquial language, is clear enough, yet all attempts to characterize this content more precisely have failed, and various discussions in which these concepts appeared and which were based on quite plausible and seemingly evident premises, have often led to paradoxes and antinomies.

[19] Tarski transforms the recursive definition into an explicit one, with the use of Frege's ancestral (footnote 1 of Tarski (1983b, p. 193)).

[20] The above mentioned-property of *satisfiability* is also elegantly defined with the use of models; a set of sentences $\alpha_1 \ldots \alpha_n$ is satisfiable if and only there is some model in which they are all true.

[21] It is worth noting that the following two definitions of truth in a model for sentences, that is, well-formed formula with no free variables, are equivalent:

(1) $\models_M \varphi$ iff for all sequences s of M, $\models_{M,s} \varphi$
(2) $\models_M \varphi$ iff for some sequence s of M, $\models_{M,s} \varphi$

It is clear, intuitively, that the equivalence holds. If φ has no free variables, then it won't depend at all for its truth on what s assigns to any variables. However, the proof of the equivalence of these two definitions is somewhat subtle, and I won't attempt it here. Essentially, the trickiness seems to be due to the following. What one might set out to prove is that where s and s' are assignments of M, then, if $FV(\varphi) = \varnothing$, then $\models_{M,s} \varphi$ iff $\models_{M,s'} \varphi$. However, there is a difficulty with proving this directly, since the induction hypothesis will be ineffectual in the interesting case, where φ is a quantified formula. So, one must prove a stronger theorem, to the effect that, if s and s' are assignments of M such that for every variable v that is free in φ, s(v) = s'(v), then s sats φ in M iff s' sats φ in M. The equivalence of (1) and (2) is an immediate consequence of this fact.

[22] Tarski's method of establishing this theorem is very similar to Gödel's method of establishing his first incompleteness theorem, and there are important questions of relative priority (see footnote 1 on p. 247 of Tarski).

[23] Though see Azzouni (2006, pp. 98ff.) for a defense of Tarski's claim.

[24] The relative notion of truth in a model is far less important for this kind of project, that is, for the project of giving a theory of meaning for natural language. In the case of meta-mathematics, truth in a model is the fundamental notion, since it is required in the

definition of notions such as logical consequence. But the semantics of natural language does not concern itself e.g. with completeness theorems; the purpose of the semantics of a natural language is to give a successful account of the meanings of natural language sentences, not to prove desirable semantic properties of formal systems. For another discussion of why the appeal to truth in a model is not central for natural language semantics, see Lepore's classic (1982) and Higginbotham (1988, section 3).

[25] The fact that Carnap interpreted necessity in terms of analyticity is important to bear in mind when assessing Quine's criticisms of the problem of de re modality. Quine is often criticized for treating quantifying into modal contexts as similar to quantifying into quotation marks. But Quine was writing for an audience that shared his assumption that a metaphysical interpretation of necessity was incoherent. Since analyticity is fundamentally a property of sentences, the view that necessity is another way of talking of analyticity makes necessity fundamentally a property of sentences as well. From this perspective, it is clear why Quine thought of the problem of de re modality as akin to quantifying into quotation marks; what "necessarily" appended to, on this view, is a quote-name of a sentence.

[26] It is fair to say that the transition in Quine's argument that has received the most criticism is the transition from the premise of failure of substitution of co-designative expressions in a syntactic position in a sentence to the conclusion that that position is not purely designative. As Kaplan (1986, p. 235) rightly points out, all that follows from the premise is that at least one occurrence of the two expressions is not purely designative; it does *not* follow that the position itself is "opaque" (that is, it does not follow that *every* occurrence of an expression in that position is a not purely designative occurrence of that expression). For a different sort of criticism of the transition, see Fine (2005a, pp. 89-90) and Fine (2005b, pp. 113-115).

[27] If one construes this claim as ruling out the possibility of writing a satisfaction clause for quantifiers where the truth of the embedded formula depends upon more than just the object that is assigned to the embedded variable, then there are apparently clear counterexamples. As Mark Richard (1987) points out, it is simple to write up a clause for a quantifier that depends not only upon what object is quantified over, but also what the embedded variable is. For instance, suppose the variables in the object-language come with numerical subscripts, and consider the following satisfaction clause for the existential quantifier:

(1) If φ is of the form '$\exists x_n \psi$', then s sats φ iff for some s'$\approx_{x\text{-}n}$s, s' sats ψ and n is odd.

According to (1), an existentially quantified formula is satisfied by a sequence only if the variable in the object-language sentence has an odd-numbered numerical subscript. This is formally a perfectly coherent satisfaction clause for quantified sentences (though of course it doesn't correspond to any intuitively natural interpretation of the quantifiers).

[28] As Carnap (1958, p. 131) puts the point:

> The fact that, according to Frege's method, the same name may have different nominata [references] in different contexts has already been mentioned as a disadvantage. But the multiplication of entities goes far beyond Frege's initial distinction between the ordinary and the oblique nominatum of a name. Actually,

these two nominata constitute only the beginning of an infinite sequence of nominata for the same name. If we apply Frege's method to sentences with multiple obliqueness, then we have to distinguish the ordinary nominatum of the name, its first oblique nominatum, its second oblique nominatum, and so forth.

There is now a large literature devoted to the evaluation of this objection to Frege. Classic contributions include Davidson (1990a), Burge (1979), Church (1951), Dummett (1981, pp. 267ff.), and Parsons (1981).

[29] As Jeffrey King has pointed out to me (p.c.), Carnap's method also evades some classic objections to the Frege/Church approach. For example, one classic objection to the Frege/Church version of the "method of the name relation" involves examples in which a single quantifier binds occurrences of variables both inside and outside non-extensional contexts, as in "Every teacher$_i$ John met x_i John believed x_i was a doctor." The first occurrence of "x" is in an extensional context, and the second occurrence is in a non-extensional context. Carnap would have no problem with this sort of example, since both occurrences of "x" have the same semantic values – an intension and an extension. In contrast, it is not clear how the Fregean method would treat this sort of example (though see Kaplan (1968, section 5) for one suggested solution on behalf of the Fregean).

[30] Quine (1947, p. 47) criticizes this account as involving "queer ontological consequences". Essentially, Quine interprets individual concepts as strange sorts of objects, with the justification that "…the ontology of a logic is nothing other than the range of admissible values of the variables of quantification."

[31] The meta-linguistic aspect of Carnap's analysis has also been subject to withering critique (Church, 1950).

[32] Though as Boghossian (1997, pp. 340-41) emphasizes, from a contemporary perspective, it is unclear whether Quine thinks that the predicate "is analytic" fails to have any determinate meaning at all, or whether it has a determinate meaning, but has no instances.

[33] In contrast, Quine's arguments (e.g. Quine, 1960, Chapter 2) for skepticism about meaning facts are generally not accepted. The literature here too is far too extensive to cite. But many philosophers accept Chomsky's famous charge ((1969), (1975, pp. 179ff.)) that the problem with Quine's arguments for meaning skepticism is that the premise of his argument is just a standard instance of under-determination of theory by evidence, and that it raises no issue specific to the case of meaning.

[34] As Carnap (1958b, pp. 223-5) writes "Suppose that the author of a system wishes the predicates 'B' and 'M' to designate the properties Bachelor and Married, respectively. How does he know that these properties are incompatible and that therefore he has to lay down [the relevant meaning postulate]? This is not a matter of knowledge but of decision. His knowledge or belief that the English words 'bachelor' and 'married' are always or usually understood in such a way that they are incompatible may influence his decision if he has the intention to reflect in his system some of the meaning relations of English words."

[35] For the distinction between indicative and subjunctive conditionals, see section X, below.

[36] It is clear that Montague was influenced by Carnap. But, as Jeffrey King has emphasized to me, Montague repeatedly acknowledges a debt to Kripke's classic (1963).

The fact that Kripke provides an elegant semantics for modal logics whose interpretations involve restricted accessibility relations is obviously important for Montague, since such semantics are important for the sort of applications Montague had in mind. But Montague attributes a greater debt than this to Kripke. Montague (1974b, p. 153) criticizes Carnap's treatment of possible worlds as what Montague calls *models* (presumably, he means state-descriptions, i.e. sets of sentences), and attributes to Kripke the discovery that possible worlds are not models, but rather primitive "points of reference" (Montague (1974a, p. 109)).

[37] In recent years, some more empirically minded philosophers and linguists have suggested that Quine's analysis does not need to be taken in the revisionary spirit in which it is intended, because there is some evidence that it is in fact correct as an analysis of natural language (see Larson, Den Dikken, and Ludlow (forthcoming)).

[38] In most recent treatments of quantifiers within the Montague tradition, quantified noun phrases are just treated as denoting second-level functions, though they are taken to introduce lambda abstracts that function to bind variables within their scope. In other words, the two aspects of quantification are formally distinguished (see e.g. Chapter 7 of Heim and Kratzer (1998)).

[39] Kaplan's paper was only published in 1989, but had been widely circulated since the mid-1970s, as "UCLA mimeograph #2".

[40] Though for a dissenting view, see Lewis (1981). Because Kaplan regards tenses and place expressions such as "somewhere" as operators on contents, he is forced to treat sentence contents as true relative to several different kinds of features of circumstances of evaluation. So, the content of an occurrence of "It is raining", for Kaplan, is neutral as to time and place (Kaplan, 1989, p. 504). But classic propositions are not neutral with regard to place or time. For example, the proposition expressed by an utterance of "It's raining" is intuitively about a particular time and place, for example New York City on December 16, 2005. As a result, Kaplan's contents do not "…exactly correspond to the classical conception of a proposition." (Kaplan, Ibid.). Continuing this line of reasoning, Lewis argues that, given the number of sentence operators in the language, Kaplan's sentence contents are going to be no closer to propositions than Montague's index theoretic semantic values. Jeffrey King (King, 2003) has persuasively argued that Kaplan and Lewis are wrong to take basic tenses as sentence operators. Following the tradition of Partee (1973), he argues that they are instead predicates of syntactically represented times. If locational expressions such as "somewhere" are also not operators (but rather, say, quantifiers over location variables in the sentence) then modal sentential operators are the only genuine content operators, and sentence contents are genuinely classical propositions.

[41] Parsons (1969, section VI) maintains that quantified modal logic is not even committed to the *meaningfulness* of essentialism. But what Parsons has in mind here is that one could provide an *alternative* interpretation of the language of quantified modal logic, "some *other* truth-conditions" than the Kripkean ones. Parsons does not supply such an alternative interpretation in his paper (though see Stalnaker (2003a) for a suggestion).

[42] Smullyan (1948) raises the point that substitution of co-extensional descriptions is permitted when the descriptions take wide-scope with respect to a modal operator. Neale (2000, pp. 308ff.) argues convincingly that this merely obfuscates the debate.

[43] In Barcan (1947) she proved that if an identity is true, then it is necessarily true, albeit in a language with only variables as singular terms.
[44] One could raise quibbles about Kripke's definition of "analytic statement", since he defines it in such a way that an analytic statement is both a priori and necessary. Certain truths in virtue of meaning, such as any utterance of "I am here now" seem to be both a priori and contingent. However, Kripke repeatedly emphasizes that his definition of "analytic statement" is intended to be stipulative. Furthermore, this is independent of the genuine issue, which concerns his disagreement with Barcan Marcus. All agree that analytic statements are a priori, so Barcan Marcus is committed to the thesis that true identity statements involving expressions used as names are a priori.
[45] I have here followed Kripke in using the unclear term "statement" to describe the thing that is both necessary and a priori. It is clear that Kripke takes the proposition expressed by an utterance of a sentence to be what is necessary. In contrast, it is not so clear *what* Kripke takes to be a priori or a posteriori. For example, it is not clear that he takes the proposition expressed by an utterance of a sentence to be what is a priori or a posteriori. Perhaps it is the sentence itself or its utterance rather than what it expresses on an occasion that is the bearer of epistemic properties. So perhaps Kripke would deny that there is one thing –a proposition—that is both a posteriori and necessary. I am not certain about the answer to this interpretative question.
[46] However, it must be said the notion of a priority with which Kripke (1972) operates is somewhat non-standard, in that it seems to be relativized to a *person*. Kripke also speaks of *stage*-relative a priori truth in footnote 33 (Kripke, 1980, p. 79), but this is equally unfamiliar.
[47] For a classic response to Kripke's modal argument, see the appendix to Chapter 5 of Dummett (1981). For more recent responses, see Stanley (1997a (section 7), 1997b, 2002)), and Sosa (2001).
[48] For a response to the semantic argument, see Stanley (1999).
[49] An interestingly distinct response to the problems is given in Stalnaker (2003b). Stalnaker does not locate the difficulty with interpreting quantified modal logic in the distinction between names and definite descriptions; according to him, both are singular terms. Instead, Stalnaker thinks that the substitution schema is properly formulated in terms of *predications*, and "one cannot treat sentences generally as predications" (Stalnaker, 2003b, p. 148). To forge the required distinction between sentences and predications, Stalnaker employs a language with complex-predicate forming devices.
[50] Lewis in fact does not provide an interpretation of quantified modal logic, in the sense of a model-theoretic semantics for it, as in Kripke (1963). Rather, Lewis (1968) proposes a *translation* of quantified modal logic into the language of counterpart theory. The success of Lewis's translation schema (in particular, for the language of quantified modal logic augmented with an operator with the meaning of "actually") is challenged in Hazen (1979) and Fara and Williamson (2005).
[51] Lewis's agreement with Quine on this point is obscured by two facts. First, in Lewis's discussion of Quine in his original paper on counterpart theory (Lewis, 1968, section 3), he writes as if there is no inconstancy in de re modal attributions, and indeed in this paper Lewis just talks as if there is one counterpart relation. Secondly, Lewis surely does not agree with Quine's views about the *degree* of inconstancy of de re modal attributions.

Nevertheless, the inconstancy of de re modal attributions, which is explicated by the availability of multiple distinct counterpart relations, is crucial to Lewis's metaphysical applications of counterpart theory (Lewis, 1971).

[52] Very similar moves to Lewis's occurred in the literature on de re propositional ascriptions, in response to Quine's inconstancy worries about such constructions. That is, as we have seen, Quine's later suspicions about regimenting de re propositional ascriptions were due to his belief that our intuitions about them were too context-sensitive. Certain philosophers (in particular (Richard, 1990)) proposed a context-sensitive semantics for propositional attitude sentences, where the contextual relativity of our intuitions is reflected in a formal semantics that incorporates context-sensitivity.

[53] In addition to Gibbard, Adams (1975) and Edgington (1986) have provided arguments for this view of indicative conditionals. I should say that I think a context-sensitive semantics, of the sort given by Stalnaker, provides an elegant explanation of the data presented by Gibbard; see Stalnaker (1984, pp. 108ff.).

[54] Presumably, it is no accident that Gibbard is both a moral expressivist and denies that indicative conditionals have truth-conditions.