

Demonstratives

An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals

David Kaplan¹

¹This paper was prepared for and read (with omissions) at a symposium on Demonstratives at the March 1977 meetings of the Pacific Division of the American Philosophical Association. The commentators were Paul Benacerraf and Charles Chastain. Much of the material, including the formal system of section XVIII, was originally presented in a series of lectures at the fabled 1971 Summer Institute in the Philosophy of Language held at the University of California, Irvine. © 1977 by David Kaplan.

Table of Contents

	Preface	483
I.	Introduction	489
II.	Demonstratives, Indexicals, and Pure Indexicals	489
III.	Two Obvious Principles	492
IV.	Remarks on Rigid Designators	492
V.	Argument for Principle 2: Pure Indexicals	498
VI.	Terminological Remarks	500
VI. (i)	Content and Circumstance	500
VI. (ii)	Character	505
VII.	Earlier Attempts: Index Theory	507
VIII.	Monsters Begat by Elegance	510
IX.	Argument for Principle 2: True Demonstratives	512
IX. (i)	The Arguments	513
IX. (ii)	The Fregean Theory of Demonstrations	514
IX. (iii)	The Fregean Theory of Demonstratives	516
IX. (iv)	Argument Against the Fregean Theory of Demonstratives	516
X.	Fixing the Reference vs. Supplying a Synonym	518
X. (i)	Reichenbach on Token Reflexives	519
XI.	The Meaning of Indexicals	520
XII.	Dthat	521
XIII.	Contexts, Truth, and Logical Truth	522
XIV.	Summary of Findings (so far): Pure Indexicals	523
XV.	Further Details: Demonstratives and Demonstrations	524
XVI.	Alternative Treatments of Demonstrations	527
XVII.	Epistemological Remarks	529
XVIII.	The Formal System	541
XIX.	Remarks on the Formal System	546
XX.	Adding 'Says'	553
XXI.	Russell on Egocentric Particulars and Their Dispensability	557
XXII.	On Proper Names	558

Preface

In about 1966 I wrote a paper about quantification into epistemological contexts. There are very difficult metaphysical, logical, and epistemological problems involved in providing a treatment of such idioms which does not distort our intuitions about their proper use and which is up to contemporary logical standards. I did not then, and do not now, regard the treatment I provided as fully adequate. And I became more and more intrigued with problems centering on what I would like to call the *semantics of direct reference*. By this I mean theories of meaning according to which certain singular terms refer directly without the mediation of a Fregean *Sinn* as meaning. If there are such terms, then the proposition expressed by a sentence containing such a term would involve individuals directly rather than by way of the "individual concepts" or "manners of presentation" I had been taught to expect. Let us call such putative singular terms (if there are any) *directly referential terms* and such putative propositions (if there are any) *singular propositions*. Even if English contained no singular terms whose proper semantics was one of direct reference, could we determine to introduce such terms? And even if we had no directly referential terms and introduced none, is there a need or use for singular propositions?

The feverish development of quantified modal logics, more generally, of quantified intensional logics, of the 1960s gave rise to a metaphysical and epistemological malaise regarding the problem of identifying individuals across worlds—what, in 1967, I called the problem of "Trans-World Heir Lines." This problem was really just the problem of singular propositions: those which involve individuals directly, rearing its irrepressible head in the possible-world semantics that were then (and are now) so popular.

It was not that according to those semantical theories any sentences of the languages being studied were themselves taken to express singular propositions, it was just that singular propositions seemed to be needed in the analysis of the nonsingular propositions expressed by these sentences. For example, consider

$$(0) \quad \exists x(Fx \wedge \sim \Box Fx).$$

This sentence would not be taken by anyone to express a singular proposition. But in order to evaluate the truth-value of the component

$$\Box Fx$$

(under some assignment of an individual to the variable 'x'), we must first determine whether the *proposition* expressed by its component

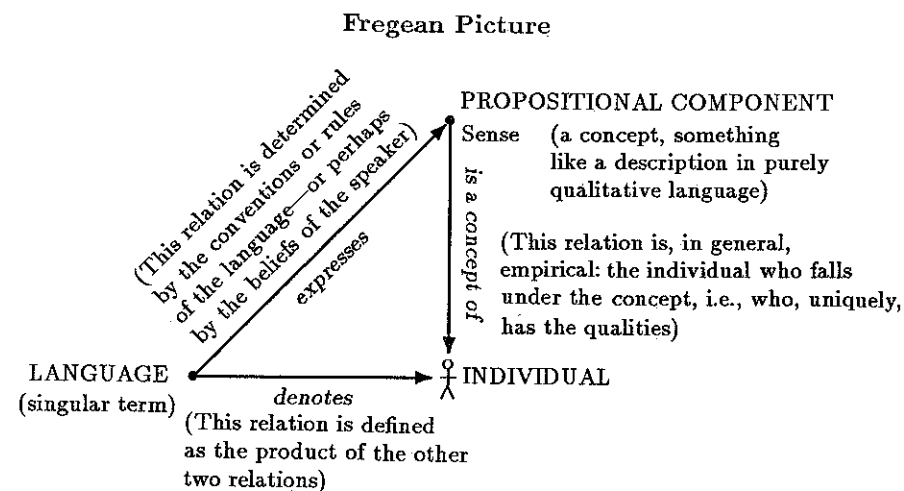
Fx

(under an assignment of an individual to the variable 'x') is a necessary proposition. So in the course of analyzing (0), we are required to determine the proposition associated with a formula containing a *free* variable. Now free variables under an assignment of values are paradigms of what I have been calling *directly referential* terms. In determining a semantical value for a formula containing a free variable we may be given a *value* for the variable—that is, an individual drawn from the universe over which the variable is taken to range—but nothing more. A variable's first and only meaning is its value. Therefore, if we are to associate a *proposition* (not merely a truth-value) with a formula containing a free variable (with respect to an assignment of a value to the variable), that proposition seems bound to be singular (even if valiant attempts are made to disguise this fact by using constant functions to imitate individual concepts). The point is, that if the component of the proposition (or the step in the construction of the proposition) which corresponds to the singular term is determined by the individual and the individual is directly determined by the singular term—rather than the individual being determined by the component of the proposition, which is directly determined by the singular term—then we have what I call a singular proposition. [Russell's semantics was like the semantical theories for quantified intensional logics that I have described in that although no (closed) sentence of *Principia Mathematica* was taken to stand for a singular proposition, singular propositions are the essential building blocks of all propositions.]

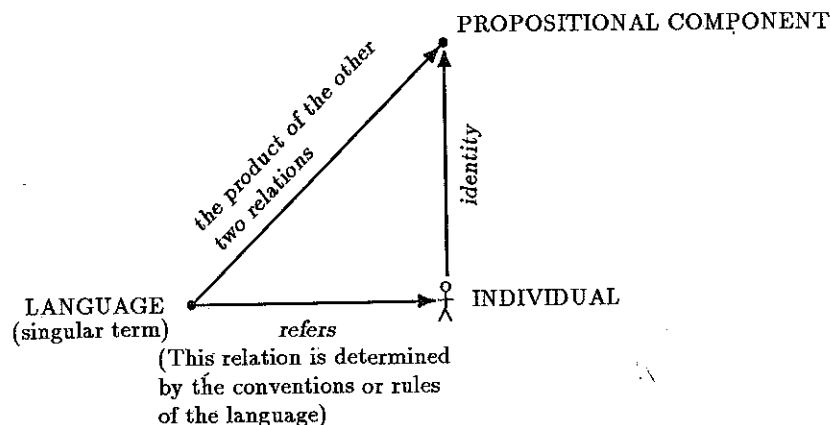
The most important hold-out against semantical theories that required singular propositions is Alonzo Church, the great modern champion of Frege's semantical theories. Church also advocates a version of quantified intensional logic, but with a subtle difference that finesses the need for singular propositions. (In Church's logic, given a sentential formula containing free variables and given an assignment of values to the variables, no proposition is yet determined. An additional assignment of "senses" to the free variables must be made before a proposition can be associated with the formula.) It is no accident that Church rejects *direct reference* semantical theories. For if there were singular terms which referred directly, it seems likely that Frege's problem: how can ' $\alpha = \beta$ ', if true, differ in meaning from ' $\alpha = \alpha$ ', could be reinstated,

while Frege's solution: that α and β , though referring to the same thing, do so by way of different senses, would be blocked. Also: because of the fact that the component of the proposition is being determined by the individual rather than vice versa, we have something like a violation of the famous Fregean dictum that *there is no road back* from denotation to sense [propositional component]. (Recently, I have come to think that if we countenance singular propositions, a collapse of Frege's intensional ontology into Russell's takes place.)

I can draw some little pictures to give you an idea of the two kinds of semantical theories I want to contrast.



Direct Reference Picture



(These pictures are not entirely accurate for several reasons, among them, that the contrasting pictures are meant to account for more than just singular terms and that the relation marked 'refers' may already involve a kind of Fregean sense used to fix the referent.)

I won't go into the pros and cons of these two views at this time. Suffice it to say that I had been raised on Fregean semantics and was sufficiently devout to wonder whether the kind of quantification into modal and epistemic contexts that seemed to require singular propositions really made sense. (My paper "Quantifying In" can be regarded as an attempt to *explain away* such idioms for epistemic contexts.)²

But there were pressures from quarters other than quantified intensional logic in favor of a semantics of direct reference. First of all there was Donnellan's fascinating paper "Reference and Definite Descriptions."³ Then there were discussions I had had with Putnam in 1968 in which he argued with respect to certain natural kind terms like 'tiger' and 'gold', that if their Fregean senses were the kind of thing that one grasped when one understood the terms, then such senses could

²David Kaplan, "Quantifying In," *Synthese* 19 (1968): 178–214; reprinted in *The Philosophy of Language*, ed. A. P. Martinich (Oxford: Oxford University Press, 1985).

³Keith Donnellan, "Reference and Definite Descriptions," *Philosophical Review* 75 (1966): 281–304; reprinted in Martinich, *op. cit.*

not determine the extension of the terms. And finally Kripke's Princeton lectures of spring 1970, later published as *Naming and Necessity*,⁴ were just beginning to leak out along with their strong attack on the Fregean theory of proper names and their support of a theory of direct reference.

As I said earlier, I was intrigued by the semantics of direct reference, so when I had a sabbatical leave for the year 1970–71, I decided to work in the area in which such a theory seemed most plausible: demonstratives. In fall 1970, I wrote, for a conference at Stanford, a paper "Dthat."⁵ Using Donnellan's ideas as a starting point, I tried to develop the contrast between Fregean semantics and the semantics of direct reference, and to argue that demonstratives—although they *could* be treated on a Fregean model—were more interestingly treated on a direct reference model. Ultimately I came to the conclusion that something analogous to Donnellan's referential use of a definite description could be developed using my new demonstrative, "dthat." In the course of this paper I groped my way to a formal semantics for demonstratives rather different in conception from those that had been offered before.

In spring 1971, I gave a series of lectures at Princeton on the semantics of direct reference. By this time I had seen a transcript of *Naming and Necessity* and I tried to relate some of my ideas to Kripke's.⁶ I also had written out the formal semantics for my Logic of Demonstratives. That summer at the Irvine Philosophy of Language Institute I lectured again on the semantics of direct reference and repeated some of these lectures at various institutions in fall 1971. And there the matter has stood except for a bit of updating of the 1971 Logic of Demonstratives notes in 1973.

I now think that demonstratives can be treated correctly only on a direct reference model, but that my earlier lectures at Princeton and Irvine on direct reference semantics were too broad in scope, and that the most important and certainly the most convincing part of my theory is just the logic of demonstratives itself. It is based on just a few quite

⁴Saul Kripke, "Naming and Necessity," in *Semantics of Natural Language*, ed. G. Harman and D. Davidson (Dordrecht: Reidel, 1972); revised edition published as a separate monograph, *Naming and Necessity* (Oxford: Basil Blackwell, 1980). References are to the revised edition.

⁵David Kaplan, "Dthat," in *Syntax and Semantics*, vol. 9, ed. P. Cole (New York: Academic Press, 1978); reprinted in Martinich, *op. cit.*

⁶Although the central ideas of my theory had been worked out before I became familiar with *Naming and Necessity*, I have enthusiastically adopted the 'analytical apparatus' and some of the terminology of that brilliant work.

simple ideas, but the conceptual apparatus turns out to be surprisingly rich and interesting. At least I hope that you will find it so.

In this work I have concentrated on pedagogy. Philosophically, there is little here that goes beyond the Summer Institute Lectures, but I have tried, by limiting the scope, to present the ideas in a more compelling way. Some new material appears in the two speculative sections: XVII (Epistemological Remarks) and XX (Adding 'Says'). It is my hope that a theory of demonstratives will give us the tools to go on in a more sure-footed way to explore the *de re* propositional attitudes as well as other semantical issues.

I. Introduction

I believe my theory of demonstratives to be uncontroversial and largely uncontroversial. This is not a tribute to the power of my theory but a concession of its obviousness. In the past, no one seems to have followed these obvious facts out to their obvious consequences. I do that. What is original with me is some terminology to help fix ideas when things get complicated. It has been fascinating to see how interesting the obvious consequences of obvious principles can be.⁷

II. Demonstratives, Indexicals, and Pure Indexicals

I tend to describe my theory as 'a theory of demonstratives', but that is poor usage. It stems from the fact that I began my investigations by asking what is said when a speaker points at someone and says, "He is suspicious."⁸ The word 'he', so used, is a demonstrative, and the accompanying pointing is the requisite associated demonstration. I hypothesized a certain semantical theory for such demonstratives, and then I invented a new demonstrative, 'dthat', and stipulated that its semantics be in accord with my theory. I was so delighted with this methodological sleight of hand for my demonstrative 'dthat', that when I generalized the theory to apply to words like 'I', 'now', 'here', etc.—words which do *not* require an associated demonstration—I continued to call my theory a 'theory of demonstratives' and I referred to these words as 'demonstratives'.

That terminological practice conflicts with what I preach, and I will try to correct it. (But I tend to backslide.)

The group of words for which I propose a semantical theory includes the pronouns 'I', 'my', 'you', 'he', 'his', 'she', 'it', the demonstrative pronouns 'that', 'this', the adverbs 'here', 'now', 'tomorrow', 'yesterday', the adjectives 'actual', 'present', and others. These words have uses other than those in which I am interested (or, perhaps, depending on how you individuate words, we should say that they have homonyms in which I am not interested). For example, the pronouns 'he' and 'his' are used not as demonstratives but as bound variables in

⁷Not everything I assert is part of my theory. At places I make judgments about the correct use of certain words and I propose detailed analyses of certain notions. I recognize that these matters may be controversial. I do not regard them as part of the basic, obvious, theory.

⁸See "Dthat," p. 320 in Martinich.

For what is a man profited, if he shall gain the whole world, and lose his own soul?

What is common to the words or usages in which I am interested is that the referent is dependent on the context of use and that the meaning of the word provides a rule which determines the referent in terms of certain aspects of the context. The term I now favor for these words is 'indexical'. Other authors have used other terms; Russell used 'egocentric particular' and Reichenbach used 'token reflexive'. I prefer 'indexical' (which, I believe, is due to Pierce) because it seems less theory laden than the others, and because I regard Russell's and Reichenbach's theories as defective.

Some of the indexicals require, in order to determine their referents, an associated demonstration: typically, though not invariably, a (visual) presentation of a local object discriminated by a pointing.⁹ These indexicals are the true demonstratives, and 'that' is their paradigm. The demonstrative (an expression) refers to that which the demonstration demonstrates. I call that which is demonstrated the 'demonstratum'.

A demonstrative without an associated demonstration is incomplete. The linguistic rules which govern the use of the true demonstratives 'that', 'he', etc., are not sufficient to determine their referent in all contexts of use. Something else—an associated demonstration—must be provided. The linguistic rules assume that such a demonstration accompanies each (demonstrative) use of a demonstrative. An incomplete demonstrative is not *vacuous* like an improper definite description. A demonstrative *can* be vacuous in various cases. For example, when its associated demonstration has no demonstratum (a hallucination)—or the wrong kind of demonstratum (pointing to a flower and saying 'he' in the belief that one is pointing to a man disguised as a flower¹⁰)—or too many demonstrata (pointing to two intertwined vines and saying

'that vine'). But it is clear that one can distinguish a demonstrative with a vacuous demonstration: no referent; from a demonstrative with no associated demonstration: incomplete.

All this is by way of contrasting true demonstratives with pure indexicals. For the latter, *no associated demonstration is required, and any demonstration supplied is either for emphasis or is irrelevant*.¹¹ Among the pure indexicals are 'I', 'now', 'here' (in one sense), 'tomorrow', and others. The linguistic rules which govern *their* use fully determine the referent for each context.¹² No supplementary actions or intentions are needed. The speaker refers to himself when he uses 'I', and no pointing to another or believing that he is another or intending to refer to another can defeat this reference.¹³

Michael Bennett has noted that some indexicals have both a pure and a demonstrative use. 'Here' is a pure indexical in

I am in here

and is a demonstrative in

In two weeks, I will be here [pointing at a city on a map].

treatment we would assign as referent for 'he' whatever the speaker *intended* to demonstrate. I intended the example to exemplify a failed demonstration, thus, a case in which the speaker, falsely believing the flower to be some man or other in disguise, but having no particular man in mind, and certainly not intending to refer to anything other than that man, says, pointing at the flower, "He has been following me around all day."

¹¹ I have in mind such cases as pointing at oneself while saying 'I' (emphasis) or pointing at someone else while saying 'I' (irrelevance or madness or what?).

¹² There are certain uses of pure indexicals that might be called 'messages recorded for later broadcast', which exhibit a special uncertainty as to the referent of 'here' and 'now'. If the message: "I am not here now" is recorded on a telephone answering device, it is to be assumed that the time referred to by 'now' is the time of playback rather than the time of recording. Donnellan has suggested that if there were typically a significant lag between our production of speech and its audition (for example, if sound traveled very very slowly), our language might contain two forms of 'now': one for the time of production, another for the time of audition. The indexicals 'here' and 'now' also suffer from vagueness regarding the size of the spatial and temporal neighborhoods to which they refer. These facts do not seem to me to slur the difference between demonstratives and pure indexicals.

¹³ Of course it is certain intentions on the part of the speaker that make a particular vocable the first person singular pronoun rather a nickname for Irving. My semantical theory is a theory of word meaning, not speaker's meaning. It is based on linguistic rules known, explicitly or implicitly, by all competent users of the language.

⁹ However, a demonstration may also be opportune and require no special action on the speaker's part, as when someone shouts "Stop that man" while only one man is rushing toward the door. My notion of a demonstration is a theoretical concept. I do not, in the present work, undertake a detailed 'operational' analysis of this notion although there are scattered remarks relevant to the issue. I do consider, in XVI below, some alternative theoretical treatments of demonstrations.

¹⁰ I am aware (1) that in some languages the so-called masculine gender pronoun may be appropriate for flowers, but it is not so in English; (2) that a background story can be provided that will make pointing at the flower a contextually appropriate, though deviant, way of referring to a man; for example, if we are talking of great hybridizers; and (3) that it is possible to treat the example as a *referential use* of the demonstrative 'he' on the model of Donnellan's referential use of a definite description (see "Reference and Definite Descriptions"). Under the referential use

III. Two Obvious Principles

So much for preliminaries. My theory is based on two obvious principles. The first has been noted in every discussion of the subject.

Principle 1 *The referent of a pure indexical depends on the context, and the referent of a demonstrative depends on the associated demonstration.*

If you and I both say 'I' we refer to different persons. The demonstratives 'that' and 'he' can be correctly used to refer to any one of a wide variety of objects simply by adjusting the accompanying demonstration.

The second obvious principle has less often been formulated explicitly.

Principle 2 *Indexicals, pure and demonstrative alike, are directly referential.*

IV. Remarks on Rigid Designators

In an earlier draft I adopted the terminology of Kripke, called indexicals 'rigid designators', and tried to explain that my usage differed from his. I am now shying away from that terminology. But because it is so well known, I will make some comments on the notion or notions involved.

The term 'rigid designator' was coined by Saul Kripke to characterize those expressions which designate the same thing in every possible world in which that thing exists and which designate nothing elsewhere. He uses it in connection with his controversial, though, I believe, correct claim that proper names, as well as many common nouns, are rigid designators. There is an unfortunate confusion in the idea that a proper name would designate nothing if the bearer of the name were not to exist.¹⁴ Kripke himself adopts positions which seem inconsistent with this feature of rigid designators. In arguing that the object designated by a rigid designator need not exist in every possible world, he seems to assert that under certain circumstances what is expressed by 'Hitler does not exist' would have been true, and not because 'Hitler' would have designated nothing (in *that* case we might have given the sentence *no* truth-value) but because what 'Hitler' would have designated—namely

¹⁴I have discussed this and related issues in "Bob and Carol and Ted and Alice," in *Approaches to Natural Language*, ed. J. Hintikka et al. (Dordrecht: Reidel, 1973), especially appendix X.

Hitler—would not have existed.¹⁵ Furthermore, it is a striking and important feature of the possible world semantics for quantified intensional logics, which Kripke did so much to create and popularize, that variables, those paradigms of rigid designation, designate the same individual in *all* possible worlds whether the individual "exists" or not.¹⁶

Whatever Kripke's intentions (did he, as I suspect, misdescribe his own concept?) and whatever associations or even meaning the phrase 'rigid designator' may have, I intend to use '*directly referential*' for an expression whose referent, once determined, is taken as fixed for all possible circumstances, i.e., is taken as *being* the propositional component.

For me, the intuitive idea is not that of an expression which *turns out* to designate the same object in all possible circumstances, but an expression whose semantical rules provide *directly* that the referent in all possible circumstances is fixed to be the actual referent. In typical cases the semantical rules will do this only implicitly, by providing a way of determining the *actual* referent and no way of determining any other propositional component.¹⁷

We should beware of a certain confusion in interpreting the phrase 'designates the same object in all circumstances'. We do not mean that the expression *could not have been used* to designate a different object.

¹⁵Kripke, *Naming and Necessity*, p. 78.

¹⁶The matter is even more complicated. There are two 'definitions' of 'rigid designator' in *Naming and Necessity*, pp. 48–49. The first conforms to what seems to me to have been the intended concept—same designation in *all* possible worlds—the second, scarcely a page later, conforms to the more widely held view that a rigid designator need not designate the object, or any object, at worlds in which the object does not exist. According to this conception a designator cannot, at a given world, designate something which does not exist in that world. The introduction of the notion of a *strongly* rigid designator—a rigid designator whose designatum exists in all possible worlds—suggests that the latter idea was uppermost in Kripke's mind. (The second definition is given, unequivocally, on page 146 of "Identity and Necessity," in *Identity and Individuation*, ed. M. K. Munitz (New York: New York University Press, 1971).) In spite of the textual evidence, systematic considerations, including the fact that variables cannot be accounted for otherwise, leave me with the conviction that the former notion was intended.

¹⁷Here, and in the preceding paragraph, in attempting to convey my notion of a directly referential singular term, I slide back and forth between two metaphysical pictures: that of possible worlds and that of structured propositions. It seems to me that a truly semantical idea should presuppose neither picture, and be expressible in terms of either. Kripke's discussion of rigid designators is, I believe, distorted by an excessive dependence on the possible worlds picture and the associated semantical style. For more on the relationship between the two pictures, see pages 724–25 of my "How to Russell a Frege-Church," *The Journal of Philosophy* 72 (1975): 716–29.

We mean rather that given a *use* of the expression, we may ask of *what has been said* whether it would have been true or false in various counterfactual circumstances, and in such counterfactual circumstances, which are the individuals relevant to determining truth-value. Thus we must distinguish possible occasions of *use*—which I call *contexts*—from possible circumstances of *evaluation* of what was said on a given occasion of use. Possible circumstances of evaluation I call circumstances or, sometimes, just *counterfactual situations*. A directly referential term *may* designate different objects when used in different *contexts*. But when evaluating what was said in a given context, only a single object will be relevant to the evaluation in all circumstances. This sharp distinction between *contexts of use* and *circumstances of evaluation* must be kept in mind if we are to avoid a seeming conflict between Principles 1 and 2.¹⁸ To look at the matter from another point of view, once we recognize the obviousness of both principles (I have not yet argued for Principle 2) the distinction between contexts of use and circumstances of evaluation is forced upon us.

If I may wax metaphysical in order to fix an image, let us think of the vehicles of evaluation—the what-is-said in a given context—as propositions. Don't think of propositions as sets of possible worlds, but rather as structured entities looking something like the sentences which express them. For each occurrence of a singular term in a sentence there will be a corresponding constituent in the proposition expressed. The constituent of the proposition determines, for each circumstance of evaluation, the object relevant to evaluating the proposition in that circumstance. In general, the constituent of the proposition will be some sort of complex, constructed from various attributes by logical composition. But in the case of a singular term which is directly referential, the constituent of the proposition is just the object itself. Thus it is that it does not just *turn out* that the constituent determines the same object in every circumstance, the constituent (corresponding to a rigid designator) *just is* the object. *There is no determining to do at all*. On this picture—and this is *really* a picture and not a theory—the definite description

- (1) The $n[(\text{Snow is slight} \wedge n^2=9) \vee (\sim\text{Snow is slight} \wedge 2^2=n+1)]$ ¹⁹

¹⁸I think it likely that it was just the failure to notice this distinction that led to a failure to recognize Principle 2. Some of the history and consequences of the conflation of Context and Circumstance is discussed in section VII.

¹⁹I would have used 'snow is white', but I wanted a contingent clause, and so many

would yield a constituent which is complex although it would determine the same object in all circumstances. Thus, (1), though a rigid designator, is not directly referential from this (metaphysical) point of view. Note, however, that every proposition which contains the complex expressed by (1) is *equivalent* to some singular proposition which contains just the number three itself as constituent.²⁰

The semantical feature that I wish to highlight in calling an expression *directly referential* is not the *fact* that it designates the same object in every circumstance, but the *way* in which it designates an object in any circumstance. Such an expression is a *device of direct reference*. This does not imply that it has no conventionally fixed semantical rules which determine its referent in each context of use; quite the opposite. There are semantical rules which determine the referent in each context of use—but that is all. *The rules do not provide a complex which together with a circumstance of evaluation yields an object. They just provide an object.*

If we keep in mind our sharp distinction between contexts of use and circumstances of evaluation, we will not be tempted to confuse a rule which assigns an object to each *context* with a 'complex' which assigns an object to each *circumstance*. For example, each context has an *agent* (loosely, a speaker). Thus an appropriate designation rule for a directly referential term would be:

- (2) In each possible context of use the given term refers to the agent of the context.

But this rule could not be used to assign a relevant object to each circumstance of evaluation. Circumstances of evaluation do not, in general, have agents. Suppose I say,

- (3) I do not exist.

Under what circumstances would *what I said* be true? It would be true in circumstances in which I did not exist. Among such circumstances are those in which no one, and thus, no speakers, no agents exist. To search a circumstance of evaluation for a speaker in order to (mis)apply rule (2) would be to go off on an irrelevant chase.

people (possibly including me) nowadays seem to have views which allow that 'snow is white' may be necessary.

²⁰I am ignoring propositions expressed by sentences containing epistemic operators or others for which equivalence is not a sufficient condition for interchange of operand.

Three paragraphs ago I sketched a metaphysical picture of the structure of a proposition. The picture is taken from the semantical parts of Russell's *Principles of Mathematics*.²¹ Two years later, in "On Denoting,"²² even Russell rejected that picture. But I still like it. It is not a part of my theory, but it well conveys my conception of a directly referential expression and of the semantics of direct reference. (The picture needs *some* modification in order to avoid difficulties which Russell later noted—though he attributed them to Frege's theory rather than his own earlier theory.)²³

If we adopt a possible worlds semantics, all directly referential terms

²¹Bertrand Russell, *The Principles of Mathematics* (London: Allen & Unwin, 1903).

²²Bertrand Russell, "On Denoting," *Mind* 14 (1905): 479–93.

²³Here is a difficulty in Russell's 1903 picture that has some historical interest. Consider the proposition expressed by the sentence, 'The centre of mass of the Solar System is a point'. Call the proposition, '*P*'. *P* has in its subject place a certain complex, expressed by the definite description. Call the complex, 'Plexy'. We can describe Plexy as "the complex expressed by 'the center of mass of the solar system'." Can we produce a directly referential term which designates Plexy? Leaving aside for the moment the controversial question of whether 'Plexy' is such a term, let us imagine, as Russell believed, that we can directly refer to Plexy by affixing a kind of *meaning marks* (on the analogy of quotation marks) to the description itself. Now consider the sentence "'the center of mass of the solar system'" is a point'. Because the subject of this sentence is directly referential and refers to Plexy, the proposition the sentence expresses will have as its subject constituent Plexy itself. A moment's reflection will reveal that this proposition is simply *P* again. But this is absurd since the two sentences speak about radically different objects.

(I believe the foregoing argument lies behind some of the largely incomprehensible arguments mounted by Russell against Frege in "On Denoting," though there are certainly other difficulties in that argument. It is not surprising that Russell there confused Frege's theory with his own of *Principle of Mathematics*. The first footnote of "On Denoting" asserts that the two theories are "very nearly the same.")

The solution to the difficulty is simple. Regard the 'object' places of a singular proposition as marked by some operation which cannot mark a complex. (There always will be some such operation.) For example, suppose that no complex is (represented by) a set containing a single member. Then we need only add {...} to mark the places in a singular proposition which correspond to directly referential terms. We no longer need worry about confusing a complex with a propositional constituent corresponding to a directly referring term because no complex will have the form {*x*}. In particular, Plexy \neq {Plexy}. This technique can also be used to resolve another confusion in Russell. He argued that a sentence containing a nondenoting directly referential term (he would have called it a nondenoting 'logically proper name') would be meaningless, presumably because the purported singular proposition would be incomplete. But the braces themselves can fill out the singular proposition, and if they contain nothing, no more anomalies need result than what the development of Free Logic has already inured us to.

will be regarded as rigid designators in the *modified* sense of an expression which designates the same thing in *all* possible worlds (irrespective of whether the thing exists in the possible world or not).²⁴ However, as already noted, I do not regard all rigid designators—not even all strongly rigid designators (those that designate something that exists in all possible worlds) or all rigid designators in the modified sense—as directly referential. I believe that proper names, like variables, are directly referential. They are not, in general, strongly rigid designators nor are they rigid designators in the original sense.²⁵ What is characteristic of directly referential terms is that the designatum (referent) determines the propositional component rather than the propositional component, along with a circumstance, determining the designatum. It is for this reason that a directly referential term that designates a contingently existing object will still be a rigid designator in the modified sense. The propositional component need not choose its designatum from those offered by a passing circumstance; it has already secured its designatum before the encounter with the circumstance.

When we think in terms of possible world semantics this fundamental distinction becomes subliminal. This is because the style of the semantical rules obscures the distinction and makes it appear that directly referential terms differ from ordinary definite descriptions only in that the propositional component in the former case must be a *constant* function of circumstances. In actual fact, the referent, in a circumstance, of a directly referential term is simply *independent* of the circumstance and is no more a function (constant or otherwise) of circumstance, than my action is a function of your desires when I decide to do it whether you like it or not. The distinction that is obscured by the style of possible world semantics is dramatized by the structured propositions picture. That is part of the reason why I like it.

Some directly referential terms, like proper names, may have no semantically relevant descriptive meaning, or at least none that is specific: that distinguishes one such term from another. Others, like the indexicals, may have a limited kind of specific descriptive meaning relevant to the features of a context of use. Still others, like 'dthat' terms (see below), may be associated with full-blown Fregean senses used to fix the referent. But in any case, the descriptive meaning of a directly referential term is no part of the propositional content.

²⁴This is the *first sense* of footnote 16.

²⁵This is the *second sense* of footnote 16.

V. Argument for Principle 2: Pure Indexicals

As stated earlier, I believe this principle is uncontroversial. But I had best distinguish it from similar principles which are false. I am *not* claiming, as has been claimed for proper names, that indexicals lack anything that might be called 'descriptive meaning'. Indexicals, in general, have a rather easily statable descriptive meaning. But it is clear that this meaning is relevant only to determining a referent in a context of use and *not* to determining a relevant individual in a circumstance of evaluation. Let us return to the example in connection with the sentence (3) and the indexical 'I'. The bizarre result of taking the descriptive meaning of the indexical to be the propositional constituent is that what I said in uttering (3) would be true in a circumstance of evaluation if and only if the speaker (assuming there is one) of the circumstance does not exist in the circumstance. Nonsense! It *that* were the correct analysis, what I said could not be true. From which it follows that

It is impossible that I do not exist.

Here is another example to show that the descriptive meaning of an indexical may be entirely *inapplicable* in the circumstance of evaluation. When I say,

I wish I were not speaking now.

The circumstances desired do not involve contexts of *use* and *agents* who are not speaking. The *actual* context of use is used to determine the relevant individual: *me*—and time: *now*—and then we query the various circumstances of evaluation with respect to *that* individual and *that* time.

Here is another example, not of the inapplicability of the descriptive meaning to circumstances but of its irrelevance. Suppose I say at t_0 , "It will soon be the case that all that is now beautiful is faded." Consider what was said in the subsentence,

All that is now beautiful is faded.

I wish to evaluate that content at some near future time t_1 . What is the relevant time associated with the indexical 'now'? Is it the future time t_1 ? No, it is t_0 , of course: the time of the context of use.

See how rigidly the indexicals cling to the referent determined in the context of use:

- (4) It is possible that in Pakistan, in five years, only those who are actually here now are envied.

The point of (4) is that the circumstance, place, and time referred to by the indexicals 'actually', 'here', and 'now' are the circumstance, place, and time of the *context*, not a circumstance, place, and time determined by the modal, locational, and temporal operators within whose scope the indexicals lie.

It may be objected that this only shows that indexicals always take *primary* scope (in the sense of Russell's scope of a definite description). This objection attempts to relegate all direct reference to implicit use of the paradigm of the semantics of direct reference, the variable. Thus (4) is transformed into,

The actual circumstances, here, and now are such that it is possible that in Pakistan in five years only those who, in the first, are located at the second, during the third, are envied.

Although this may not be the most felicitous form of expression, its meaning and, in particular, its symbolization should be clear to those familiar with quantified intensional logics. The pronouns, 'the first', 'the second', and 'the third' are to be represented by distinct variables bound to existential quantifiers at the beginning and identified with 'the actual circumstance', 'here', and 'now' respectively.

- (5) $(\exists w)(\exists p)(\exists t)[w = \text{the actual circumstance} \wedge p = \text{here} \wedge t = \text{now}$
 $\wedge \Diamond \text{In Pakistan In five years } \forall x(x \text{ is envied} \rightarrow x \text{ is located}$
 $\text{at } p \text{ during } t \text{ in } w)]$

But such transformations, when thought of as representing the claim that indexicals take primary scope, do not provide an *alternative* to Principle 2, since we may still ask of an utterance of (5) in a context c , when evaluating it with respect to an arbitrary circumstance, to what do the indexicals 'actual', 'here', and 'now' refer. The answer, as always, is: the relevant features of the context c . [In fact, although (4) is equivalent to (5), neither indexicals nor quantification across intensional operators is dispensable in favor of the other.]

Perhaps enough has been said to establish the following.

- (T1) *The descriptive meaning of a pure indexical determines the referent of the indexical with respect to a context of use but is either inapplicable or irrelevant to determining a referent with respect to a circumstance of evaluation.*

I hope that your intuition will agree with mine that it is for this reason that:

- (T2) *When what was said in using a pure indexical in a context *c* is to be evaluated with respect to an arbitrary circumstance, the relevant object is always the referent of the indexical with respect to the context *c*.*

This is just a slightly elaborated version of Principle 2.

Before turning to true demonstratives, we will adopt some terminology.

VI. Terminological Remarks

Principle 1 and Principle 2 taken together imply that sentences containing pure indexicals have two kinds of meaning.

VI. (i) Content and Circumstance

What is said in using a given indexical in different contexts may be different. Thus if I say, today,

I was insulted yesterday

and you utter the same words tomorrow, what is said is different. If what we say differs in truth-value, that is enough to show that we say different things. But even if the truth-values were the same, it is clear that there are possible circumstances in which what I said would be true but what you said would be false. Thus we say different things.

Let us call this first kind of meaning—what is said—*content*. The content of a sentence in a given context is what has traditionally been called a proposition. Strawson, in noting that the sentence

The present king of France is bald

could be used on different occasions to make different statements, used 'statement' in a way similar to our use of *content of a sentence*. If we

wish to express the same content in different contexts, we may have to change indexicals. Frege, here using 'thought' for content of a sentence, expresses the point well.

If someone wants to say the same today as he expressed yesterday using the word 'today', he must replace this word with 'yesterday'. Although the thought is the same its verbal expression must be different so that the sense, which would otherwise be affected by the differing times of utterance, is readjusted.²⁶

I take *content* as a notion applying not only to sentences taken in a context but to any meaningful part of speech taken in a context. Thus we can speak of the content of a definite description, an indexical, a predicate, etc. It is *contents* that are evaluated in circumstances of evaluation. If the content is a proposition (i.e., the content of a sentence taken in some context), the result of the evaluation will be a truth-value. The result of evaluating the content of a singular term at a circumstance will be an object (what I earlier called 'the relevant object'). In general, the result of evaluating the content of a well-formed expression α at a circumstance will be an appropriate extension for α (i.e., for a sentence, a truth-value; for a term, an individual; for an n -place predicate, a set of n -tuples of individuals, etc.). This suggests that we can represent a

²⁶From "The Thought: A Logical Inquiry," *Mind* 65 (1956): 289–311. If Frege had only supplemented these comments with the observation that indexicals are devices of direct reference, the whole theory of indexicals would have been his. But his theory of meaning blinded him to this obvious point. Frege, I believe, mixed together the two kinds of meaning in what he called *Sinn*. A *thought* is, for him, the *Sinn* of a sentence, or perhaps we should say a *complete* sentence. *Sinn* is to contain both "the manner and context of presentation [of the denotation]," according to "Über Sinn und Bedeutung" (*Zeitschrift für Philosophie und philosophische Kritik* 100 (1892); trans. as "On Sense and Nominatum," in *Contemporary Readings in Logical Theory*, ed. Copi and Gould (Macmillan, 1967); mistrans. as "On Sense and Meaning," in Martinich, op. cit.). *Sinn* is first introduced to represent the cognitive significance of a sign, and thus to solve Frege's problem: how can ' $\alpha = \beta$ ' if true differ in cognitive significance from ' $\alpha = \alpha$ '. However, it also is taken to represent the truth-conditions or *content* (in our sense). Frege felt the pull of the two notions, which he reflects in some tortured passages about 'I' in "The Thought" (quoted below in XVII). If one says "Today is beautiful" on Tuesday and "Yesterday was beautiful" on Wednesday, one expresses the same thought according to the passage quoted. Yet one can clearly lose track of the days and not realize one is expressing the same thought. It seems then that thoughts are not appropriate bearers of cognitive significance. I return to this topic in XVII. A detailed examination of Frege on demonstratives is contained in John Perry's "Frege on Demonstratives," *Philosophical Review* 86 (1977): 474–97.

content by a function from circumstances of evaluation to an appropriate extension. Carnap called such functions *intensions*.

The representation is a handy one and I will often speak of contents in terms of it, but one should note that contents which are distinct but equivalent (i.e., share a value in all circumstances) are represented by the same intension. Among other things, this results in the loss of my distinction between terms which are devices of direct reference and descriptions which *turn out* to be rigid designators. (Recall the meta-physical paragraph of section IV.) I wanted the content of an indexical to be just the referent itself, but the intension of such a content will be a constant function. Use of representing intensions does not mean I am abandoning that idea—just ignoring it temporarily.

A *fixed content* is one represented by a constant function. All directly referential expressions (as well as all rigid designators) have a fixed content. [What I elsewhere call a *stable content*.]

Let us settle on *circumstances* for possible circumstances of evaluation. By this I mean both actual and counterfactual situations with respect to which it is appropriate to ask for the extensions of a given well-formed expression. A circumstance will usually include a possible state or history of the world, a time, and perhaps other features as well. The amount of information we require from a circumstance is linked to the degree of specificity of contents, and thus to the kinds of operators in the language.

Operators of the familiar kind treated in intensional logic (modal, temporal, etc.) operate on contents. (Since we represent contents by intensions, it is not surprising that intensional operators operate on contents.) Thus an appropriate extension for an intensional operator is a function from intensions to extensions.²⁷ A modal operator when applied to an intension will look at the behavior of the intension with respect to the possible state of the world feature of the circumstances of evaluation. A temporal operator will, similarly, be concerned with

the time of the circumstance. If we built the time of evaluation into the contents (thus removing time from the circumstances leaving only, say, a possible world history, and making contents *specific* as to time), it would make no sense to have temporal operators. To put the point another way, if *what is said* is thought of as incorporating reference to a specific time, or state of the world, or whatever, it is otiose to ask whether what is said would have been true at another time, in another state of the world, or whatever. Temporal operators applied to eternal sentences (those whose contents incorporate a specific time of evaluation) are redundant. Any intensional operators applied to *perfect* sentences (those whose contents incorporate specific values for all features of circumstances) are redundant.²⁸

²⁸The notion of redundancy involved could be made precise. When I speak of building the time of evaluation into contents, or making contents specific as to time, or taking what is said to incorporate reference to a specific time, what I have in mind is this. Given a sentence *S*: 'I am writing', in the present context *c*, which of the following should we take as the content: (i) the proposition that David Kaplan is writing at 10 A.M. on 3/26/77, or (ii) the 'proposition' that David Kaplan is writing? The proposition (i) is specific as to time, the 'proposition' (ii) [the scare quotes reflect my feeling that this is not the traditional notion of a proposition] is neutral with respect to time. If we take the content of *S* in *c* to be (ii), we can ask whether it would be true at times other than the time of *c*. Thus we think of the temporally neutral 'proposition' as changing its truth-value over time. Note that it is not just the noneternal sentence *S* that changes its truth-value over time, but the 'proposition' itself. Since the sentence *S* contains an indexical 'I', it will express different 'propositions' in different contexts. But since *S* contains no *temporal* indexical, the time of the context will not influence the 'proposition' expressed. An alternative [and more traditional] view is to say that the verb tense in *S* involves an implicit temporal indexical, so that *S* is understood as synonymous with *S'*: 'I am writing now'. If we take this point of view we will take the content of *S* in *c* to be (i). In this case *what is said* is eternal; it does not change its truth-value over time, although *S* will express different propositions at different times.

There are both technical and philosophical issues involved in choosing between (i) and (ii). Philosophically, we may ask why the temporal indexical should be taken to be implicit (making the proposition eternal) when no modal indexical is taken to be implicit. After all, we *could* understand *S* as synonymous with *S''*: 'I am actually writing now'. The content of *S''* in *c* is not only eternal, it is perfect. Its truth changes neither through time nor possibility. Is there some good philosophical reason for preferring contents which are neutral with respect to possibility but draw fixed values from the context for all other features of a possible circumstance whether or not the sentence contains an explicit indexical? (It may be that the traditional view was abetted by one of the delightful anomalies of the logic of indexicals, namely that *S*, *S'*, and *S''* are all logically equivalent! See Remark 3, p. 547.) Technically, we must note that intensional operators must, if they are not to be vacuous, operate on contents which are neutral with respect

²⁷As we shall see, indexical operators such as "It is now the case that," "It is actually the case that," and "dthat" (the last takes a term rather than a sentence as argument) are also intensional operators. They differ from the familiar operators in only two ways: first, their extension (the function from intensions to extensions) depends on context, and second, they are directly referential (thus they have a fixed content). I shall argue below (in section VII: Monsters) that all operators that can be given an English reading are 'at most' intensional. Note that when discussing issues in terms of the formal representations of the model-theoretic semantics, I tend to speak in terms of intensions and intensional operators rather than contents and content operators.

What sorts of intensional operators to admit seems to me largely a matter of language engineering. It is a question of which features of what we intuitively think of as possible circumstances can be sufficiently well defined and isolated. If we wish to isolate location and regard it as a feature of possible circumstances we can introduce locational operators: 'Two miles north it is the case that', etc. Such operators can be iterated and can be mixed with modal and temporal operators. However, to make such operators interesting we must have contents which are locationally neutral. That is, it must be appropriate to ask if *what is said* would be true in Pakistan. (For example, 'It is raining' seems to be locationally as well as temporally and modally neutral.)

This functional notion of the content of a sentence in a context may not, because of the neutrality of content with respect to time and place, say, exactly correspond to the classical conception of a proposition. But the classical conception can be introduced by adding the demonstratives 'now' and 'here' to the sentence and taking the content of the result. I will continue to refer to the content of a sentence as a proposition, ignoring the classical use.

Before leaving the subject of circumstances of evaluation I should, perhaps, note that the mere attempt to show that an expression is directly referential requires that it be meaningful to ask of an individual in one circumstance whether and with what properties it exists in another circumstance. If such questions cannot be raised because they are regarded as metaphysically meaningless, the question of whether a particular expression is directly referential (or even, a rigid designator) cannot be raised. I have elsewhere referred to the view that such questions are meaningful as *haecceitism*, and I have described other metaphysical manifestations of this view.²⁹ I advocate this position, although I am

to the feature of circumstance the operator is interested in. Thus, for example, if we take the content of *S* to be (i), the application of a temporal operator to such a content would have no effect; the operator would be vacuous. Furthermore, if we do not wish the iteration of such operators to be vacuous, the content of the compound sentence containing the operator must again be neutral with respect to the relevant feature of circumstance. This is not to say that no such operator can have the effect of *fixing* the relevant feature and thus, in effect, rendering subsequent operations vacuous; indexical operators do just this. It is just that this must not be the general situation. A content must be the *kind* of entity that is subject to modification in the feature relevant to the operator. [The textual material to which this note is appended is too cryptic and should be rewritten.]

²⁹ "How to Russell a Frege-Church." The pronunciation is: "Hex-ee-i-tis-m." The epithet was suggested by Robert Adams. It is not an accident that it is derived from a demonstrative.

uncomfortable with some of its seeming consequences (for example, that the world might be in a state qualitatively exactly as it is, but with a permutation of individuals).

It is hard to see how one could think about the semantics of indexicals and modality without adopting such a view.

VI. (ii) Character

The second kind of meaning, most prominent in the case of indexicals, is that which determines the content in varying contexts. The rule,

'I' refers to the speaker or writer

is a meaning rule of the second kind. The phrase 'the speaker or writer' is not supposed to be a complete description, nor it is supposed to refer to the speaker or writer of the word 'I'. (There are many such.) It refers to the speaker or writer of the relevant *occurrence* of the word 'I', that is, the agent of the context.

Unfortunately, as usually stated, these meaning rules are incomplete in that they do not explicitly specify that the indexical is directly referential, and thus do not completely determine the content in each context. I will return to this later.

Let us call the second kind of meaning, *character*. The character of an expression is set by linguistic conventions and, in turn, determines the content of the expression in every context.³⁰ Because character is what is set by linguistic conventions, it is natural to think of it as *meaning* in the sense of what is known by the competent language user.

Just as it was convenient to represent contents by functions from possible circumstances to extensions (Carnap's intentions), so it is convenient to represent characters by functions from possible contexts to contents. (As before we have the drawback that equivalent characters are identified.³¹) This gives us the following picture:

³⁰ This does not imply that if you know the character and are in first one and then another context, you can *decide* whether the contents are the same. I may twice use 'here' on separate occasions and not recognize that the place is the same, or twice hear 'I' and not know if the content is the same. What I do know is this: if it was the same person speaking, then the content was the same. [More on this epistemological stuff later.]

³¹ I am, at this stage, deliberately ignoring Kripke's theory of proper names in order to see whether the revisions in Fregean semantic theory, which seem plainly required to accommodate indexicals (this is the 'obviousness' of my theory), can throw any light on it. Here we assume that aside from indexicals, Frege's theory

Character: Contexts \Rightarrow Contents

Content: Circumstances \Rightarrow Extensions

or, in more familiar language,

Meaning + Context \Rightarrow Intension

Intension + Possible World \Rightarrow Extension

Indexicals have a *context-sensitive* character. It is characteristic of an indexical that its content varies with context. Nonindexicals have a *fixed* character. The same content is invoked in all contexts. This content will typically be sensitive to circumstances, that is, the non-indexicals are typically not rigid designators but will vary in extension from circumstance to circumstance. Eternal sentences are generally good examples of expressions with a fixed character.

All persons alive in 1977 will have died by 2077

expresses the same proposition no matter when said, by whom, or under what circumstances. The truth-value of that proposition may, of course, vary with possible circumstances, but the character is fixed. Sentences with fixed character are very useful to those wishing to leave historical records.

Now that we have two kinds of meaning in addition to extension, Frege's principle of intensional interchange³² becomes two principles:

is correct, roughly, that words and phrases have a kind of descriptive meaning or sense which at one and the same time constitutes their cognitive significance and their conditions of applicability.

Kripke says repeatedly in *Naming and Necessity* that he is only providing a picture of how proper names refer and that he does not have an exact theory. His picture yields some startling results. In the case of indexicals we do have a rather precise theory, which avoids the difficulty of specifying a chain of communication and which yields many analogous results. In facing the vastly more difficult problems associated with a theory of reference for proper names, the theory of indexicals may prove useful; if only to show—as I believe—that proper names are not indexicals and have no meaning in the sense in which indexicals have meaning (namely a 'cognitive content' which fixes the references in all contexts). [The issues that arise, involving token reflexives, homonymous words with distinct character, and homonymous token reflexives with the same character are best saved for later—much later.]

³²See §28 of Rudolf Carnap's *Meaning and Necessity* (Chicago: University of Chicago Press, 1947).

(F1) The character of the whole is a function of the character of the parts. That is, if two compound well-formed expressions differ only with respect to components which have the same Character, then the Character of the compounds is the same.

(F2) The Content of the whole is a function of the Content of the parts. That is, if two compound well-formed expressions, each set in (possibly different) contexts differ only with respect to components which *when taken in their respective contexts* have the same content, then the content of the two compounds *each taken in its own context* is the same.

It is the second principle that accounts for the often noted fact that speakers in different contexts can say the same thing by switching indexicals. (And indeed they often *must* switch indexicals to do so.) Frege illustrated this point with respect to 'today' and 'yesterday' in "The Thought." (But note that his treatment of 'I' suggests that he does not believe that utterances of 'I' and 'you' could be similarly related!)

Earlier, in my metaphysical phase, I suggested that we should think of the content of an indexical as being just the referent itself, and I resented the fact that the representation of contents as intensions forced us to regard such contents as constant functions. A similar remark applies here. If we are not overly concerned with standardized representations (which certainly have their value for model-theoretic investigations) we might be inclined to say that the character of an indexical-free word or phrase just is its (constant) content.

VII. Earlier Attempts: Index Theory

The following picture seems to emerge. The meaning (character) of an indexical is a function from contexts to extensions (substituting for fixed contents). The meaning (content, substituting for fixed characters) of a nonindexical is a function from circumstances to extensions. From this point of view it may appear that the addition of indexicals requires no new *logic*, no sharp distinction between contexts and circumstances, just the addition of some special new *features* ('contextual' features) to the circumstances of evaluation. (For example, an *agent* to provide an interpretation for 'I'.) Thus an enlarged view of intension is derived. The intension of an expression is a function from certain factors to the extension of the expression (with respect to those factors). Originally such factors were simply possible states of the world, but as it was noticed

that the so-called tense operators exhibited a structure highly analogous to that of the modal operators the factors with respect to which an extension was to be determined were enlarged to include moments of time. When it was noticed that contextual factors were required to determine the extension of sentences containing indexicals, a still more general notion was developed and called an "index." The extension of an expression was to be determined with respect to an index. The intension of an expression was that function which assigned to every index, the extension at that index.

The above example supplies us with a statement whose truth-value is not constant but varies as a function of $i \in I$. This situation is easily appreciated in the context of time-dependent statements; that is, in the case where I represents the instant of time. Obviously the same statement can be true at one moment and false at another. For more general situations one must not think of the $i \in I$ as anything as simple as instants of time or even possible worlds. In general we will have

$$i = (w, t, p, a, \dots)$$

where the index i has many *coordinates*: for example, w is a *world*, t is a *time*, $p = (x, y, z)$ is a (3-dimensional) *position* in the world, a is an *agent*, etc. All these coordinates can be varied, possibly independently, and thus affect the truth-values of statements which have indirect references to these coordinates. [From the *Advice* of a prominent logician.]

A sentence ϕ was taken to be logically true if true at every index (in every 'structure'), and $\Box\phi$ was taken to be true at a given index (in a given structure) just in case ϕ was true at every index (in that structure). Thus the familiar principle of modal generalization: if $\models \phi$, then $\models \Box\phi$, is validated.

This view, in its treatment of indexicals, was technically wrong and, more importantly, conceptually misguided.

Consider the sentence

(6) I am here now.

It is obvious that for many choices of index—i.e., for many quadruples $\langle w, x, p, t \rangle$ where w is a possible world history, x is a person, p is a place,

and t is a time—(6) will be false. In fact, (6) is true only with respect to those indices $\langle w, x, p, t \rangle$ which are such that in the world history w , x is located at p at the time t . Thus (6) fares about on a par with

(7) David Kaplan is in Portland on 26 March 1977.

(7) is empirical, and so is (6).

But here we have missed something essential to our understanding of indexicals. Intuitively, (6) is deeply, and in some sense, which we will shortly make precise, universally, true. One need only understand the meaning of (6) to know that it cannot be uttered falsely. No such guarantees apply to (7). A *Logic of Indexicals* which does not reflect this intuitive difference between (6) and (7) has bypassed something essential to the logic of indexicals.

What has gone wrong? We have ignored the special relationship between 'I', 'here', and 'now'. Here is a proposed correction. Let the class of indices be narrowed to include only the *proper* ones—namely, those $\langle w, x, p, t \rangle$ such that in the world w , x is located at p at the time t . Such a move may have been intended originally since improper indices are like impossible worlds; no such contexts *could* exist and thus there is no interest in evaluating the extensions of expressions with respect to them. Our reform has the consequence that (6) comes out, correctly, to be logically true. Now consider

(8) \Box I am here now.

Since the contained sentence (namely (6)) is true at every proper index, (8) also is true at every proper index and thus also is logically true. (As would be expected by the aforementioned principle of modal generalization.)

But (8) should not be *logically* true, since it is false. It is certainly *not* necessary that I be here now. But for several contingencies, I would be working in my garden now, or even delivering this paper in a location outside of Portland.

The difficulty, here, is the attempt to assimilate the role of a *context* to that of a *circumstance*. The indices $\langle w, x, p, t \rangle$ that represent contexts must be proper in order that (6) be a truth of the logic of indexicals, but the indices that represent circumstances must include improper ones in order that (8) *not* be a logical truth.

If one wishes to stay with this sort of index theory and blur the conceptual difference between context and circumstance, the minimal requirement is a system of *double* indexing, one index for context and

another for circumstance. It is surprising, looking back, that we (for I was among the early index theorists) did not immediately see that double indexing was required, for in 1967, at UCLA, Hans Kamp had reported his work on 'now'³³ in which he had shown that double indexing was required to properly accommodate temporal indexicals along with the usual temporal operators. But it was *four* years before it was realized that this was a general requirement for (and, in a sense, the key to) a logic of indexicals.

However, mere double indexing, without a clear conceptual understanding of what each index stands for, is still not enough to avoid all pitfalls.

VIII. Monsters Begat by Elegance

My liberality with respect to operators on content, i.e., intensional operators (any feature of the circumstances of evaluation that can be well defined and isolated) does not extend to operators which attempt to operate on character. Are there such operators as 'In some contexts it is true that', which when prefixed to a sentence yields a truth if and only if in some context the contained *sentence* (not the content expressed by it) expresses a content that is true in the circumstances of that context? Let us try it:

(9) In some contexts it is true that I am not tired now.

For (9) to be true in the present context it suffices that some agent of some context not be tired at the time of that context. (9), so interpreted, has nothing to do with me or the present moment. But this violates Principle 2! Principle 2 can also be expressed in more theory laden way by saying that indexicals always take primary scope. If this is true—and it is—then no operator can control the character of the indexicals within its scope, because they will simply leap out of its scope to the front of the operator. I am not saying we could not construct a language with such operators, just that English is not one.³⁴ And such operators *could not be added to it*.

There *is* a way to control an indexical, to keep it from taking primary scope, and even to refer it to another context (this amounts to changing its character). Use quotation marks. If we *mention* the indexical rather

than *use* it, we can, of course, operate directly on it. Carnap once pointed out to me how important the difference between direct and indirect quotation is in

Otto said "I am a fool."

Otto said that I am a fool.

Operators like 'In some contexts it is true that', which attempt to meddle with character, I call *monsters*. I claim that none can be expressed in English (without sneaking in a quotation device). If they stay in the metalanguage and confine their attention to sentences as in

In some contexts "I am not tired now" is true

they are rendered harmless and can even do socially useful work (as does, 'is valid' [see below]).

I have gone on at perhaps excessive length about monsters because they have recently been begat by elegance. In a specific application of the theory of indexicals there will be just certain salient features of a circumstance of evaluation. So we may represent circumstances by indexed sets of features. This is typical of the model-theoretic way. As already indicated, all the features of a circumstance will generally be required as aspects of a context, and the aspects of a context may all be features of a circumstance. If not, a little ingenuity may make it so.³⁵

³⁵ Recall that in a particular formal theory the features of a circumstance must include all elements with respect to which there are content operators, and the aspects of a context must include all elements with respect to which there are indexicals. Thus, a language with both the usual modal operators ' \Diamond ', ' \Box ', and an indexical modal operator 'It is actually the case that' will contain a possible world history feature in its circumstances as well as an analogous aspect in its contexts. If a circumstance is an aspect of a context, as seems necessary for the definition of truth, then we only need worry about aspects of contexts that are not features of circumstances. The most prominent of these is the *agent* of the context, required to interpret the indexical 'I'. In order to supply a corresponding nonvacuous feature to circumstances we must treat contents in such a way that we can ask whether they are true for various agents. (Not *characters* mind you, but contents.) This can be done by representing the agent by a *neutral*—a term which plays the syntactical role of 'I' but gets an interpretation only with respect to a circumstance. Let a be a special variable that is not subject to quantification and let b be a variable not in the language. Our variable a is the neutral. We wish to introduce content operators which affect the agent place and which can be iterated. Let R be a relation between individuals, for example ' aRb ' for ' b is an uncle of a '. Then we may interpret the operator $O^R\phi$ as $(\exists b)[aRb \wedge (\exists a)(b = a \wedge \phi)]$. If ϕ is ' a walks', $O^R\phi$ comes to 'an uncle of a walks'. The indexical 'I' can be represented by an operator O^I for which ' aRb ' is just ' $I=b$ '. The result should be that $O^I\phi$ is equivalent to replacing the neutral a by the indexical 'I'.

³³ Published in 1971 as "Formal Properties of 'Now'," *Theoria*.

³⁴ Thomason alleges a counterinstance: 'Never put off until tomorrow what you can do today'. What should one say about this?

We could then represent contexts by the same indexed sets we use to represent circumstances, and instead of having a *logic of contexts and circumstances* we have simply a *two-dimensional logic of indexed sets*. This is algebraically very neat and it permits a very simple and elegant description of certain important classes of characters (for example, those which are true at every pair $\langle i, i \rangle$, though the special significance of the set is somehow diminished in the abstract formulation).³⁶ But it also permits a simple and elegant introduction of many operators which are monsters. In abstracting from the distinct conceptual roles played by contexts of use and circumstances of evaluation the special logic of indexicals has been obscured. Of course restrictions can be put on the two-dimensional logic to exorcise the monsters, but to do so would be to give up the mathematical advantages of that formulation.³⁷

IX. Argument for Principle 2: True Demonstratives

I return now to the argument that all indexicals are directly referential. Suppose I point at Paul and say,

He now lives in Princeton, New Jersey.

Call *what I said*—i.e., the content of my utterance, the proposition expressed—‘Pat’. Is Pat true or false? True! Suppose that unbeknownst to me, Paul had moved to Santa Monica last week. Would Pat have then been true or false? False! Now, the tricky case: Suppose that Paul and Charles had each disguised themselves as the other and had switched places. If that had happened, *and* I had uttered as I did, then the proposition I *would have* expressed would have been false. But in that possible context the proposition I *would have* expressed is not Pat. That is easy to see because the proposition I *would have* expressed, had I pointed to Charles instead of Paul—call this proposition ‘Mike’—not only *would have* been false but actually is false. Pat, I would claim, would still be true in the circumstances of the envisaged possible con-

³⁶See, for example, Krister Segerberg, “Two-dimensional Modal Logic,” *Journal of Philosophical Logic* 2 (1973): 77–96. Segerberg does metamathematical work in his article and makes no special philosophical claims about its significance. That has been done by others.

³⁷There is one other difficulty in identifying the class of contexts with the class of circumstances. The special relationship between the indexicals ‘I’, ‘here’, ‘now’ seems to require that the agent of a context be at the location of the context during the time of the context. But this restriction is not plausible for arbitrary circumstances. It appears that this approach will have difficulty in avoiding the problems of (6) and (8) (section VII).

text provided that Paul—in whatever costume he appeared—were still residing in Princeton.

IX. (i) The Arguments

I am arguing that in order to determine what the truth-value of a proposition expressed by a sentence containing a demonstrative *would be* under other possible circumstances, the relevant individual is not the individual that *would have* been demonstrated had those circumstances obtained and the demonstration been set in a context of those circumstances, but rather the individual demonstrated in the context which *did* generate the proposition being evaluated. As I have already noted, it is characteristic of sentences containing demonstratives—or, for that matter, any indexical—that they may express different propositions in different contexts. We must be wary of confusing the proposition that would have been expressed by a similar utterance in a slightly different context—say, one in which the demonstratum is changed—with the proposition that was actually expressed. If we keep this distinction in mind—i.e., we distinguish Pat and Mike—we are less likely to confuse what the truth-value of the proposition *actually* expressed would have been under some possible circumstances with what the truth-value of the proposition that *would have been* expressed would have been under those circumstances.

When we consider the vast array of possible circumstances with respect to which we might inquire into the truth of a proposition expressed in some context *c* by an utterance *u*, it quickly becomes apparent that only a small fraction of these circumstances will involve an utterance of the same sentence in a similar context, and that there must be a way of evaluating the truth-value of propositions expressed using demonstratives in counterfactual circumstances in which no demonstrations are taking place and no individual has the exact characteristics exploited in the demonstration. Surely, it is irrelevant to determining whether what I said would be true or not in some counterfactual circumstance, whether Paul, or anyone for that matter, *looked* as he does now. All that would be relevant is *where he lives*. Therefore,

(T3) the relevant features of the demonstratum *qua* demonstratum (compare, the relevant features of the x Fx *qua* the x Fx)—namely, that the speaker is pointing at it, that it has a certain appearance, is presented in a certain way—cannot be the essential characteristics used to identify the relevant individual in counterfactual situations.

These two arguments: the distinction between Pat and Mike, and consideration of counterfactual situations in which no demonstration occurs, are offered to support the view that demonstratives are devices of direct reference (rigid designators, if you will) and, by contrast, to reject a Fregean theory of demonstratives.

IX. (ii) The Fregean Theory of Demonstrations

In order to develop the latter theory, in contrast to my own, we turn first to a portion of the Fregean theory which I accept: the Fregean theory of demonstrations.

As you know, for a Fregean the paradigm of a meaningful expression is the definite description, which picks out or denotes an individual, a unique individual, satisfying a condition *s*. The individual is called the *denotation* of the definite description and the condition *s* we may identify with the *sense* of the definite description. Since a given individual may uniquely satisfy several distinct conditions, definite descriptions with distinct senses may have the same denotation. And since some conditions may be uniquely satisfied by no individual, a definite description may have a sense but no denotation. The condition by means of which a definite description picks out its denotation is the *manner of presentation* of the denotation by the definite description.

The Fregean theory of demonstratives claims, correctly I believe, that the analogy between descriptions (short for 'definite descriptions') and demonstrations is close enough to provide a sense and denotation analysis of the 'meaning' of a demonstration. The denotation is the demonstratum (that which is demonstrated), and it seems quite natural to regard each demonstration as presenting its demonstratum in a particular manner, which we may regard as the sense of the demonstration. The same individual could be demonstrated by demonstrations so different in manner of presentation that it would be informative to a competent auditor-observer to be told that the demonstrata were one. For example, it might be informative to you for me to tell you that

That [pointing to Venus in the morning sky] is identical with that [pointing to Venus in the evening sky].

(I would, of course, have to speak very slowly.) The two demonstrations—call the first one 'Phos' and the second one 'Hes'—which accompanied the two occurrences of the demonstrative expression 'that' have

the same demonstratum but distinct manners of presentation. It is this difference between the sense of Hes and the sense of Phos that accounts, the Fregean claims, for the informativeness of the assertion.

It is possible, to pursue the analogy, for a demonstration to have no demonstratum. This can arise in several ways: through hallucination, through carelessness (not noticing, in the darkened room, that the subject had jumped off the demonstration platform a few moments before the lecture began), through a sortal conflict (using the demonstrative phrase 'that *F*', where *F* is a common noun phrase, while demonstrating something which is not an *F*), and in other ways.

Even Donnellans's important distinction between referential and attributive uses of definite descriptions seems to fit, equally comfortably, the case of demonstrations.³⁸

The Fregean hypostatizes demonstrations in such a way that it is appropriate to ask of a given demonstration, say Phos, what *would* it have demonstrated under various counterfactual circumstances. Phos and Hes might have demonstrated distinct individuals.³⁹

We should not allow our enthusiasm for analogy to overwhelm judgment in this case. There are some relevant respects in which descriptions and demonstrations are disanalogous. First, as David Lewis has pointed out, demonstrations do not have a syntax, a fixed formal structure in terms of whose elements we might try to define, either directly or recursively, the notion of sense.⁴⁰ Second, to different audiences (for example, the speaker, those sitting in front of the demonstration platform, and those sitting behind the demonstration platform) the same demonstration may have different senses. Or perhaps we should say that a single performance may involve distinct demonstrations from the perspective of distinct audiences. ("Exactly like proper names!" says the Fregean, "as long as the demonstratum remains the same, these fluctuations in sense are tolerable. But they should be avoided in the system

³⁸I have written elsewhere, in appendices VII and VIII of "Bob and Carol and Ted and Alice," of these matters and won't pursue the topic now.

³⁹It could then be proposed that demonstrations be individuated by the principle: $d_1 = d_2$ if and only if, for all appropriate circumstances *c*, the demonstratum of d_1 in *c* = the demonstratum of d_2 in *c*. An alternative principle of individuation is that the same demonstration is being performed in two different contexts if the standard audience can't determine, from the demonstration alone, whether the contexts are distinct or identical. This makes the individuation of demonstrations more epistemological than the metaphysical proposal above.

⁴⁰Although recent work on computer perception has attempted to identify a syntax of pictures. See P. Suppes and Rottmayer, "Automata," in *Handbook of Perception*, vol. 1 (New York: Academic Press, 1974).

of a demonstrative science and should not appear in a perfect vehicle of communication.”)

IX. (iii) The Fregean Theory of Demonstratives

Let us accept, tentatively and cautiously, the Fregean theory of demonstrations, and turn now to the Fregean theory of demonstratives.⁴¹

According to the Fregean theory of demonstratives, an occurrence of a demonstrative expression functions rather like a place-holder for the associated demonstration. The sense of a sentence containing demonstratives is to be the result of replacing each demonstrative by a constant whose sense is given as the sense of the associated demonstration. An important aim of the Fregean theory is, of course, to solve Frege's problem. And it does that quite neatly. You recall that the Fregean accounted for the informativeness of

That [Hes] = that [Phos]

in terms of the distinct senses of Hes and Phos. Now we see that the senses of the two occurrences of 'that' are identified with these two distinct senses so that the ultimate solution is exactly like that given by Frege originally. The sense of the left 'that' differs from the sense of the right 'that'.

IX. (iv) Argument Against the Fregean Theory of Demonstratives

Let us return now to our original example:

He [Delta] now lives in Princeton, New Jersey

where 'Delta' is the name of the relevant demonstration. I assume that in the possible circumstances described earlier, Paul and Charles having disguised themselves as each other, Delta would have demonstrated Charles. Therefore, according to the Fregean theory, the proposition I just expressed, Pat, would have been false under the counterfactual circumstances of the switch. But this, as argued earlier, is wrong. Therefore, the Fregean theory of demonstratives though it nicely solves Frege's problem, is simply incorrect in associating propositions with utterances.

Let me recapitulate. We compared two theories as to the proposition expressed by a sentence containing a demonstrative along with an asso-

ciated demonstration. Both theories allow that the demonstration can be regarded as having both a sense and a demonstratum. My theory, the direct reference theory, claims that in assessing the proposition in counterfactual circumstances it is the actual demonstratum—in the example, Paul—that is the relevant individual. The Fregean theory claims that the proposition is to be construed as if the sense of the demonstration were the sense of the demonstrative. Thus, in counterfactual situations it is the individual that *would* have been demonstrated that is the relevant individual. According to the direct reference theory, demonstratives are rigid designators. According to the Fregean theory, their denotation varies in different counterfactual circumstances as the demonstrata of the associated demonstration would vary in those circumstances.

The earlier distinction between Pat and Mike, and the discussion of counterfactual circumstances in which, as we would now put it, the demonstration would have demonstrated nothing, argue that with respect to the problem of associating propositions with utterances the direct reference theory is correct and the Fregean theory is wrong.

I have carefully avoided arguing for the direct reference theory by using modal or subjunctive sentences for fear the Fregean would claim that the peculiarity of demonstratives is not that they are rigid designators but that they always take primary scope. If I had argued only on the basis of our intuitions as to the truth-value of

If Charles and Paul had changed chairs, then he (Delta)
would not now be living in Princeton

such a scope interpretation could be claimed. But I didn't.

The perceptive Fregeans among you will have noted that I have said nothing about how Frege's problem fares under a direct reference theory of demonstratives. And indeed, if 'that' accompanied by a demonstration is a rigid designator for the demonstratum, then

that (Hes) = that (Phos)

looks like two rigid designators designating the same thing. Uh Oh! I will return to this in my Epistemological Remarks (section XVII).

⁴¹The Fregean theory of demonstrations is not a part of my obvious and uncontroversial theory of indexicals. On the contrary, it has the fascination of the speculative.

X. Fixing the Reference vs. Supplying a Synonym⁴²

The Fregean is to be forgiven. He has made a most natural mistake. Perhaps he thought as follows: If I point at someone and say 'he', that occurrence of 'he' must refer to the male at whom I am now pointing. It does! So far, so good. Therefore, the Fregean reasons, since 'he' (in its demonstrative sense) means the same as 'the male at whom I am now pointing' and since the denotation of the latter varies with circumstances the denotation of the former must also. But this is wrong. Simply because it is a rule of the language that 'he' *refers* to the male at whom I am now pointing (or, whom I am now demonstrating, to be more general), it does not follow that any synonymy is thereby established. In fact, this is one of those cases in which—to use Kripke's excellent idiom—the rule simply tells us how to *fix the reference* but does not supply a synonym.

Consider the proposition I express with the utterance

He [Delta] is the male at whom I am now pointing.

Call that proposition 'Sean'. Now Sean is certainly true. We know from the rules of the language that any utterance of that form must express a true proposition. In fact we would be justified in calling the *sentence*

He is the male at whom I am now pointing.

almost analytic. ('Almost' because of the hypothesis that the demonstrative is *proper*—that I am pointing at a unique male—is needed.)

But is Sean necessary? Certainly not, I might have pointed at someone else.

This kind of mistake—to confuse a semantical rule which tells how to fix the reference to a directly referential term with a rule which supplies a synonym—is easy to make. Since semantics must supply a meaning, in the sense of content (as I call it), for expressions, one thinks naturally that whatever way the referent of an expression is given by the semantical rules, that *way* must stand for the content of the expression. (Church [or was it Carnap?] says as much, explicitly.) This hypothesis

⁴²I use Kripke's terminology to expound the important distinction he introduces in *Naming and Necessity* for descriptive meaning that may be associated with a proper name. As in several other cases of such parallels between proper names and indexicals, the distinction, and its associated argument, seems more obvious when applied to indexicals.

seems especially plausible, when, as is typical of indexicals,

the semantical rule which fixes the reference seems to exhaust our knowledge of the meaning of the expression.

X. (i) Reichenbach on Token Reflexives

It was from such a perspective, I believe, that Reichenbach built his ingenious theory of indexicals. Reichenbach called such expressions 'token-reflexive words' in accordance with his theory. He writes as follows:

We saw that most individual-descriptions are constructed by reference to other individuals. Among these there is a class of descriptions in which the individual referred to is the act of speaking. We have special words to indicate this reference; such words are 'I', 'you', 'here', 'now', 'this'. Of the same sort are the tenses of verbs, since they determine time by reference to the time when the words are uttered. To understand the function of these words we have to make use of the distinction between *token* and *symbol*, 'token' meaning the individual sign, and 'symbol' meaning the class of similar tokens (cf. §2). Words and sentences are symbols. The words under consideration are words which refer to the corresponding token used in an individual act of speech, or writing; they may therefore be called *token-reflexive* words.

It is easily seen that all these words can be defined in terms of the phrase 'this token'. The word 'I', for instance, means the same as 'the person who utters this token'; 'now' means the same as 'the time at which this token was uttered'; 'this table' means the same as 'the table pointed to by a gesture accompanying this token'. We therefore need inquire only into the meaning of the phrase 'this token'.⁴³

But is it true, for example, that

(10) 'I' means the same as 'the person who utters this token' ?

It is certainly true that

I am the person who utters this token.

⁴³H. Reichenbach, *Elements of Symbolic Logic* (New York: Macmillan, 1947), p. 284.

But if (10) correctly asserted a synonymy, then it would be true that

(11) If no one were to utter this token, I would not exist.

Beliefs such as (11) could make one a compulsive talker.

XI. The Meaning of Indexicals

In order to correctly and more explicitly state the semantical rule which the dictionary attempts to capture by the entry

I: the person who is speaking or writing

we would have to develop our semantical theory—the semantics of direct reference—and then state that

(D1) 'I' is an indexical, different utterances of which may have different contents

(D3) 'I' is, in each of its utterances, directly referential

(D2) In each of its utterances, 'I' refers to the person who utters it.

We have seen errors in the Fregean analysis of demonstratives and in Reichenbach's analysis of indexicals, all of which stemmed from failure to realize that these words are directly referential. When we say that a word is directly referential are we saying that its meaning is its reference (its only meaning is its reference, its meaning is nothing more than its reference)? Certainly not.⁴⁴ Insofar as meaning is given by the rules of a language and is what is known by competent speakers, I would be more inclined to say in the case of directly referential words and phrases that their reference is *no* part of their meaning. The meaning of the word 'I' does not change when different persons use it. The meaning of 'I' is given by the rules (D1), (D2), and (D3) above.

⁴⁴We see here a drawback to the terminology 'direct reference'. It suggests falsely that the reference is not mediated by a meaning, which it is. The meaning (character) is directly associated, by convention, with the word. The meaning determines the referent; and the referent determines the content. It is this to which I alluded in the parenthetical remark following the picture on page 486. Note, however, that the kind of descriptive meaning involved in giving the character of indexicals like 'I', 'now', etc., is, because of the focus on context rather than circumstance, unlike that traditionally thought of as Fregean sense. It is the idea that the referent determines the content—that, contra Frege, there is a road back—that I wish to capture. This is the importance of Principle 2.

Meanings tell us how the content of a word or phrase is determined by the context of use. Thus the meaning of a word or phrase is what I have called its *character*. (Words and phrases with no indexical element express the same content in every context; they have a fixed character.) To supply a synonym for a word or phrase is to find another with the same *character*; finding another with the same *content* in a particular context certainly won't do. The content of 'I' used by me may be identical with the content of 'you' used by you. This doesn't make 'I' and 'you' synonyms. Frege noticed that if one wishes to say again what one said yesterday using 'today', today one must use 'yesterday'. (Incidentally the relevant passage, quoted on page 501, propounds what I take to be a direct reference theory of the indexicals 'today' and 'yesterday'.) But 'today' and 'yesterday' are not synonyms. For two words or phrases to be synonyms, they must have the same content in every context. In general, for indexicals, it is not possible to find synonyms. This is because indexicals are directly referential, and the compound phrases which can be used to give their reference ('the person who is speaking', 'the individual being demonstrated', etc.) are not.

XII. Dthat⁴⁵

It would be useful to have a way of converting an arbitrary singular term into one which is directly referential.

Recall that we earlier regarded demonstrations, which are required to 'complete' demonstratives, as a kind of description. The demonstrative was then treated as a directly referential term whose referent was the demonstratum of the associated demonstration.

Now why not regard descriptions as a kind of demonstration, and introduce a special demonstrative which requires completion by a description and which is treated as a directly referential term whose referent is the denotation of the associated description? Why not? Why not indeed! I have done so, and I write it thus:

dthat[α]

where α is any description, or, more generally, any singular term. 'Dthat' is simply the demonstrative 'that' with the following singular term func-

⁴⁵Pronunciation note on 'dthat'. The word is not pronounced dee-that or duh-that. It has only one syllable. Although articulated differently from 'that' (the tongue begins behind the teeth), the sounds are virtually indistinguishable to all but native speakers.

tioning as its demonstration. (Unless you hold a Fregean theory of demonstratives, in which case its meaning is as stipulated above.)

Now we can come much closer to providing genuine synonyms.

'I' means the same as 'dthat [the person who utters this token]'.

(The fact that this alleged synonymy is cast in the theory of utterances rather than occurrences introduces some subtle complications, which have been discussed by Reichenbach.)

XIII. Contexts, Truth, and Logical Truth

I wish, in this section, to contrast an *occurrence* of a well-formed expression (my *technical* term for the combination of an expression and a context) with an *utterance* of an expression.

There are several arguments for my notion, but the main one is from Remark 1 on the Logic of Demonstratives (section XIX below): I have sometimes said that the content of a sentence in a context is, roughly, the proposition the sentence would express if uttered in that context. This description is not quite accurate on two counts. First, it is important to distinguish an *utterance* from a *sentence-in-a-context*. The former notion is from the theory of speech acts, the latter from semantics. Utterances take time, and utterances of distinct sentences cannot be simultaneous (i.e., in the same context). But in order to develop a logic of demonstratives we must be able to evaluate several premises and a conclusion all in the same context. We do not want arguments involving indexicals to become valid simply because there is no possible context in which all the premises are uttered, and thus no possible context in which all are uttered truthfully.

Since the content of an occurrence of a sentence containing indexicals depends on the context, the notion of *truth* must be relativized to a context.

If *c* is a context, then an occurrence of ϕ in *c* is true iff the content expressed by ϕ in this context is true when evaluated with respect to the circumstance of the context.

We see from the notion of truth that among other aspects of a context must be a possible circumstance. Every context occurs in a particular circumstance, and there are demonstratives such as 'actual' which refer to that circumstance.

If you try out the notion of truth on a few examples, you will see that it is correct. If I now utter a sentence, I will have uttered a truth just in case *what I said*, the content, is true in *these* circumstances.

As is now common for intensional logics, we provide for the notion of a *structure*, comprising a family of circumstances. Each such structure will determine a set of possible contexts. Truth in a structure, is truth in every possible context of the structure. Logical truth is truth in every structure.

XIV. Summary of Findings (so far): Pure Indexicals

Let me try now to summarize my findings regarding the semantics of demonstratives and other indexicals. First, let us consider the non-demonstrative indexicals such as 'I', 'here' (in its nondemonstrative sense), 'now', 'today', 'yesterday', etc. In the case of these words, the linguistic conventions which constitute *meaning* consist of rules specifying the referent of a given *occurrence* of the word (we might say, a given token, or even utterance, of the word, if we are willing to be somewhat less abstract) in terms of various features of the context of the occurrence. Although these rules fix the referent and, in a very special sense, might be said to define the indexical, the way in which the rules are given does not provide a synonym for the indexical. The rules tell us for any possible occurrence of the indexical what the referent would be, but they do *not* constitute the content of such an occurrence. Indexicals are directly referential. The rules tell us what it is that is referred to. Thus, they *determine* the content (the propositional constituent) for a particular occurrence of an indexical. But they are not a *part* of the content (they constitute no part of the propositional constituent). In order to keep clear on a topic where ambiguities constantly threaten, I have introduced two technical terms: *content* and *character* for the two kinds of meaning (in addition to extension) I associate with indexicals. Distinct occurrences of an indexical (in distinct contexts) may not only have distinct referents, they may have distinct meanings in the sense of *content*. If I say "I am tired today" today and Montgomery Furth says "I am tired today" tomorrow, our utterances have different contents in that the factors which are relevant to determining the truth-value of what Furth said in both actual and counterfactual circumstances are quite different from the factors which are relevant to determining the truth-value of what I said. Our two utterances are as different in content as are the sentences "David Kaplan is tired on 26 March 1977" and

"Montgomery Furth is tired on 27 March 1977." But there is another sense of meaning in which, absent lexical or syntactical ambiguities, two occurrences of the *same* word or phrase *must* mean the same. (Otherwise how could we learn and communicate with language?) This sense of meaning—which I call *character*—is what determines the content of an occurrence of a word or phrase in a given context. For indexicals, the rules of language constitute the meaning in the sense of *character*. As normally expressed, in dictionaries and the like, these rules are incomplete in that, by omitting to mention that indexicals are directly referential, they fail to specify the full content of an occurrence of an indexical.

Three important features to keep in mind about these two kinds of meaning are:

1. Character applies only to words and phrases as types, content to occurrences of words and phrases in contexts.
2. Occurrences of two phrases can agree in content although the phrases differ in character, and two phrases can agree in character but differ in content in distinct contexts.
3. The relationship of character to content is something like that traditionally regarded as the relationship of sense to denotation, character is a way of presenting content.

XV. Further Details: Demonstratives and Demonstrations

Let me turn now to the demonstratives proper, those expressions which must be associated with a demonstration in order to determine a referent. In addition to the pure demonstratives 'that' and 'this' there are a variety of demonstratives which contain built-in sortals: 'he' for 'that male', 'she' for 'that female',⁴⁶ etc., and there are demonstrative phrases built from a pure demonstrative and a common noun phrase: 'that man drinking a martini', etc. Words and phrases which have demonstrative use may have other uses as well, for example, as bound variable or pronouns of laziness (anaphoric use).

I accept, tentatively and cautiously, the Fregean theory of demonstrations according to which:

- (1) A demonstration is a way of presenting an individual.
- (2) A given demonstration in certain counterfactual circumstances would have demonstrated (i.e., presented) an individual other than the individual actually demonstrated.
- (3) A demonstration which fails to demonstrate any individual might have demonstrated one, and a demonstration which demonstrates an individual might have demonstrated no individual at all.

So far we have asserted that it is not an essential property of a given demonstration (according to the Fregean theory) that it demonstrate a given individual, or indeed, that it demonstrate any individual at all. It is this feature of demonstrations: that demonstrations which in fact demonstrate the same individual might have demonstrated distinct individuals, which provides a solution to the demonstrative version of Frege's problem (why is an utterance of 'that [Hes] = that [Phos]' informative?) analogous to Frege's own solution to the definite description version. There is some theoretical latitude as to how we should regard such other features of a demonstration as its place, time, and agent. Just to fix ideas, let us regard all these features as accidental. (It may be helpful to think of demonstrations as *types* and particular performances of them as their *tokens*). Then,

- (4) A given demonstration might have been mounted by someone other than its actual agent, and might be repeated in the same or a different place.

Although we are not now regarding the actual place and time of a demonstration as essential to it, it does seem to me to be essential to a demonstration that it present its demonstrata from some perspective, that is, as the individual that looks thusly *from here now*. On the other hand, it does not seem to me to be essential to a demonstration that it be mounted by any agent at all.⁴⁷

⁴⁷If the current speculations are accepted, then in the original discussion of Pat and Mike the emphasis on the counterfactual situation in which the same agent was doing the pointing was misguided and that feature of counterfactual situations is irrelevant. It is the agent of course who focuses your attention on the relevant local individual. But that needn't be done *by* anyone; we might have a convention that whoever is appearing on the demonstration platform is the demonstratum, or the speaker might take advantage of a natural demonstration of opportunity: an explosion or a shooting star.

⁴⁶'Male' and 'female' are here used in the grammatical sense of gender, not the biological sense.

We now have a kind of standard form for demonstrations:

The individual that has appearance *A* from here now

where an appearance is something like a picture with a little arrow pointing to the relevant subject. Trying to put it into words, a particular demonstration might come out like:

The brightest heavenly body now visible from here.

In this example we see the importance of perspective. The same demonstration, differently located, may present a different demonstratum (a twin, for example).

If we set a demonstration, δ , in a context, *c*, we determine the relevant perspective (i.e., the values of 'here' and 'now'). We also determine the demonstratum, if there is one—if, that is, in the circumstances of the context there is an individual that appears that way from the place and time of the context.⁴⁸ In setting δ and *c* we determine more than just the demonstratum in the possible world of the context. By fixing the perspective, we determine for each possible circumstance what, if anything, would appear like that from that perspective. This is to say, we determine a *content*. This content will not, in general, be fixed (like that determined by a rigid designator). Although it was Venus that appeared a certain way from a certain location in ancient Greece, it might have been Mars. Under certain counterfactual conditions, it *would* have been Mars that appeared just that way from just that location. Set in a different context, δ , may determine a quite different content or no content at all. When I look at myself in the mirror each morning I know that I didn't look like that ten years ago—and I suspect that nobody did.

The preceding excursion into a more detailed Fregean theory of demonstrations was simply in order to establish the following structural features of demonstrations:

1. A demonstration, when set in a context (i.e., an *occurrence* of a demonstration), determines a content.

2. It is not required that an occurrence of a demonstration have a fixed content.

In view of these features, we can associate with each demonstration a *character* which represents the 'meaning' or manner of presentation of the demonstration. We have now brought the semantics of demonstrations and descriptions into isomorphism.⁴⁹ Thus, I regard my 'dthat' operator as representing the general case of a demonstrative. Demonstratives are incomplete expressions which must be completed by a demonstration (type). A complete sentence (type) will include an associated demonstration (type) for each of its demonstratives. Thus each demonstrative, *d*, will be accompanied by a demonstration, δ , thus:

$d[\delta]$

The character of a *complete* demonstrative is given by the semantical rule:

In any context *c*, $d[\delta]$ is a directly referential term that designates the demonstratum, if any, of δ in *c*, and that otherwise designates nothing.

Obvious adjustments are to be made to take into account any common noun phrase which accompanies or is built-in to the demonstrative.

Since no immediately relevant structural differences have appeared between demonstrations and descriptions, I regard the treatment of the 'dthat' operator in the formal logic LD as accounting for the general case. It would be a simple matter to add to the syntax a category of 'nonlogical demonstration constants'. (Note that the indexicals of LD are all logical signs in the sense that their meaning [character] is not given by the structure but by the evaluation rules.)

XVI. Alternative Treatments of Demonstrations

The foregoing development of the Fregean theory of demonstrations is not inevitable. Michael Bennett has proposed that only places be demonstrata and that we require an explicit or implicit common noun phrase to accompany the demonstrative, so that:

⁴⁸ Since, as remarked earlier, the speaker and different members of the audience generally have different perspectives on the demonstration, it may appear slightly different to each of them. Thus each may take a slightly different demonstration to have been performed. Insofar as the agent and audience of a given context can differ in location, the location of a context is the location of the agent. Therefore the demonstratum of a given demonstration set in a given context will be the individual, if any, thereby demonstrated from the speaker's point of view.

⁴⁹ We should not, of course, forget the many disanalogies noted earlier nor fail to note that though a description is associated with a particular character by linguistic convention, a demonstration is associated with its character by nature.

that [pointing at a person]

becomes

dthat [the person who is there [pointing at a place]].

My findings do not include the claim that the—or better, a—Fregean theory of demonstrations is correct. I can provide an alternative account for those who regard demonstrations as nonrepeatable nonseparable features of contexts. The conception now under consideration is that in certain contexts the agent is demonstrating something, or more than one thing, and in others not. Thus just as we can speak of agent, time, place, and possible world history as features of a context, we may also speak of first demonstratum, second demonstratum, ... (some of which may be null) as features of a context. We then attach subscripts to our demonstratives and regard the n -th demonstrative, when set in a context, as rigid designator of the n -th demonstratum of the context. Such a rule associates a character with each demonstrative. In providing no role for demonstrations as separable ‘manners of presentation’ this theory eliminates the interesting distinction between demonstratives and other indexicals. We might call it the *Indexical theory of demonstratives*. (Of course every reasonable theory of demonstratives treats them as indexicals of some kind. I regard my own theory of indexicals in general, and the nondemonstrative indexicals in particular, as essentially uncontroversial. Therefore I reserve *Indexical theory of demonstratives* for the controversial alternative to the Fregean theory of demonstrations—the Fregean theory of demonstratives having been refuted.)

Let us call my theory as based on the Fregean theory of demonstrations the *Corrected Fregean theory of demonstratives*. The Fregean theory of demonstrations may be extravagant, but compared with its riches, the indexical theory is a mean thing. From a logical point of view, the riches of the Corrected Fregean theory of demonstratives are already available in connection with the demonstrative ‘dthat’ and its descriptive pseudodemonstrations, so a decision to enlarge the language of LD with additional demonstratives whose semantics are in accord with the Indexical theory need not be too greatly lamented.

If we consider Frege’s problem, we have the two formulations:

that [Hes] = that [Phos]

and

that₁ = that₂

Both provide their sentence with an informative character. But the Fregean idea that that very demonstration might have picked out a different demonstratum seems to me to capture more of the epistemological situation than the Indexicalist’s idea that in some contexts the first and second demonstrata differ.

The Corrected Fregean theory, by incorporating demonstration types in its sentence types, accounts for more differences in informativeness as differences in meaning (character). It thereby provides a nice Frege-type solution to many Frege-type problems. But it can only forestall the resort to directly epistemological issues, it cannot hold them in abeyance indefinitely. Therefore I turn to epistemological remarks.

XVII. Epistemological Remarks⁵⁰

How do content and character serve as objects of thought? Let us state, once again, Frege’s problem

(FP) How can (an occurrence of) ‘ $\alpha = \beta$ ’ (in a given context), if true, differ in cognitive significance from (an occurrence of) ‘ $\alpha = \alpha$ ’ (in the same context)?

In (FP) α, β are arbitrary singular terms. (In future formulations, I will omit the parentheticals as understood.) When α and β are demonstrative free, Frege explained the difference in terms of his notion of sense. A notion which, his writings generally suggest, should be identified with our *content*. But it is clear that Frege’s problem can be reinstituted in a form in which resort to contents will not explain differences in ‘cognitive significance’. We need only ask,

(FPD) How can ‘dthat[α] = dthat[β]’ if true, differ in cognitive significance from ‘dthat[α] = dthat[α]’?

Since, as we shall show, for any term γ ,

‘ $\gamma = \text{dthat}[\gamma]$ ’ is analytic

the sentence pair in (FP) will differ in cognitive significance if and only if the sentence pair in (FPD) differ similarly. [There are a few assumptions built in here, but they are O.K.] Note, however, that the *content* of ‘dthat[α]’ and the *content* of ‘dthat[β]’ are the same whenever ‘ $\alpha = \beta$ ’

⁵⁰This section has benefited from the opportunity to read, and discuss with him, John Perry’s paper “Frege on Demonstratives.”

is true. Thus the difference in cognitive significance between the sentence pair in (FPD) cannot be accounted for in terms of content.

If Frege's solution to (FP) was correct, then α and β have different contents. From this it follows that $\text{'dthat}[\alpha]$ and $\text{'dthat}[\beta]$ have different characters. [It doesn't really, because of the identification of contents with intensions, but let it pass.] Is character, then, the object of thought?

If you and I both say to ourselves,

(B) "I am getting bored"

have we thought the same thing? We could not have, because what you thought was true while what I thought was false.

What we must do is disentangle two epistemological notions: *the objects of thought* (what Frege called "Thoughts") and the *cognitive significance of an object of thought*. As has been noted above, a character may be likened to a manner of presentation of a content. This suggests that we identify objects of thought with contents and the cognitive significance of such objects with characters.

E. Principle 1 *Objects of thought (Thoughts) = Contents*

E. Principle 2 *Cognitive significance of a Thought = Character*

According to this view, the thoughts associated with $\text{'dthat}[\alpha]$ = $\text{'dthat}[\beta]$ and $\text{'dthat}[\alpha]$ = $\text{'dthat}[\alpha]$ are the same, but the thought (not the denotation, mind you, but the *thought*) is *presented* differently.

It is important to see that we have not *simply* generalized Frege's theory, providing a higher order Fregean sense for each name of a regular Fregean sense.⁵¹ In Frege's theory, a given manner of presentation presents the same object to all mankind.⁵² But for us, a given manner of presentation—a character—what we both said to ourselves when we both said (B)—will, in general, present different objects (of thought) to different persons (and even different Thoughts to the same person at different times).

⁵¹ According to Church, such higher order Fregean senses are already called for by Frege's theory.

⁵² See his remarks in "On Sense and Nominatum" regarding the "common treasure of thoughts which is transmitted from generation to generation" and remarks there and in "The Thought" in connection with tensed sentences, that "Only a sentence supplemented by a time-indication and complete in every respect expresses a thought."

How then can we claim that we have captured the idea of cognitive significance? To break the link between cognitive significance and universal Fregean senses and at the same time forge the link between cognitive significance and character we must come to see the *context-sensitivity* (dare I call it ego-orientation?) of cognitive states.

Let us try a Putnam-like experiment. We raise two identical twins, Castor and Pollux, under qualitatively identical conditions, qualitatively identical stimuli, etc. If necessary, we may monitor their brain states and make small corrections in their brain structures if they begin drifting apart. They respond to all cognitive stimuli in identical fashion.⁵³ Have we not been successful in achieving the same cognitive (i.e., psychological) state? Of course we have, what more could one ask! But wait, they believe different things. Each sincerely says,

My brother was born before I was

and the beliefs they thereby express conflict. In this, Castor speaks the truth, while Pollux speaks falsely. This does not reflect on the identity of their cognitive states, for, as Putnam has emphasized, circumstances alone do not determine extension (here, the truth-value) from cognitive state. Insofar as distinct persons can be in the same cognitive state, Castor and Pollux are.

E. Corollary 1 *It is an almost inevitable consequence of the fact that two persons are in the same cognitive state, that they will disagree in their attitudes toward some object of thought.*

The corollary applies equally well to the same person at different times, and to the same person at the same time in different circumstances.⁵⁴ In general, the corollary applies to any individuals x, y in different contexts.

My aim was to argue that the cognitive significance of a word or phrase was to be identified with its character, the way the content is presented to us. In discussing the twins, I tried to show that persons

⁵³ Perhaps it should be mentioned here, to forestall an objection, that neither uses a proper name for the other or for himself—only 'my brother' and 'I'—and that raising them required a lot of environmental work to maintain the necessary symmetries, or, alternatively, a lot of work with the brain state machine. If proper names are present, and each uses a different name for himself (or, for the other), they will never achieve the same *total* cognitive state since one will sincerely say, "I am Castor" and the other will not. They may still achieve the same cognitive state in its relevant part.

⁵⁴ The corollary would also apply to the same person at the same time in the same circumstances but in different places, if such could be.

could be in the same total cognitive state and still, as we would say, believe different things. This doesn't prove that the cognitive content of, say, a single sentence or even a word is to be identified with its character, but it strongly suggests it.

Let me try a different line of argument. We agree that a given content may be presented under various characters and that consequently we may hold a propositional attitude toward a given content under one character but not under another. (For example, on March 27 of this year, having lost track of the date, I may continue to hope to be finished by this March 26, without hoping to be finished by yesterday.) Now instead of arguing that character is what we would ordinarily call cognitive significance, let me just ask why we should be interested in the character under which we hold our various attitudes. Why should we be interested in that special kind of significance that is sensitive to the use of indexicals; 'I', 'here', 'now', 'that', and the like? John Perry, in his stimulating and insightful paper "Frege on Demonstratives" asks and answers this question. [Perry uses 'thought' where I would use 'object of thought' or 'content', he uses 'apprehend' for 'believe' but *note that other psychological verbs would yield analogous cases*. I have taken a few liberties in substituting my own terminology for Perry's and have added the emphasis.]

Why should we care under what character someone apprehends a thought, so long as he does? I can only sketch the barest suggestion of an answer here. *We use the manner of presentation, the character, to individuate psychological states, in explaining and predicting action*. It is the manner of presentation, the character and not the thought apprehended, that is tied to human action. When you and I have beliefs under the common character of 'A bear is about to attack me', we behave similarly. We both roll up in a ball and try to be as still as possible. Different thoughts apprehended, same character, same behavior. When you and I both apprehend that I am about to be attacked by a bear, we behave differently. I roll up in a ball, you run to get help. Same thought apprehended, different characters, different behaviors.⁵⁵

Perry's examples can be easily multiplied. My hope to be finished by a certain time is sensitive to how the content corresponding to the

time is presented, as 'yesterday' or as 'this March 26'. If I see, reflected in a window, the image of a man whose pants appear to be on fire, my behavior is sensitive to whether I think, 'His pants are on fire' or 'My pants are on fire', though the object of thought may be the same.

So long as Frege confined his attention to indexical free expressions, and given his theory of proper names, it is not surprising that he did not distinguish objects of thought (content) from cognitive significance (character), for that is the realm of *fixed* character and thus, as already remarked, there is a natural identification of character with content. Frege does, however, discuss indexicals in two places. The first passage, in which he discusses 'yesterday' and 'today' I have already discussed. Everything he says there is essentially correct. (He does not go far enough.) The second passage has provoked few endorsements and much skepticism. It too, I believe, is susceptible of an interpretation which makes it essentially correct. I quote it in full.

Now everyone is presented to himself in a particular and primitive way, in which he is presented to no one else. So, when Dr. Lauben thinks that he has been wounded, he will probably take as a basis this primitive way in which he is presented to himself. And only Dr. Lauben himself can grasp thoughts determined in this way. But now he may want to communicate with others. He cannot communicate a thought which he alone can grasp. Therefore, if he now says 'I have been wounded', he must use the 'I' in a sense that can be grasped by others, perhaps in the sense of 'he who is speaking to you at this moment', by doing which he makes the associated conditions of his utterance serve for the expression of his thought.⁵⁶

What is the particular and primitive way in which Dr. Lauben is presented to himself? What cognitive content presents Dr. Lauben to himself, but presents him to nobody else? Thoughts determined this way can be grasped by Dr. Lauben, but no one else can grasp *that* thought determined in *that* way. The answer, I believe, is, simply, that Dr. Lauben is presented to himself under the character of 'I'.

A sloppy thinker might succumb to the temptation to slide from an acknowledgement of the privileged *perspective* we each have on ourselves—only I can refer to me as 'I'—to the conclusions: first, that

⁵⁵ John Perry, "Frege on Demonstratives," p. 494.

⁵⁶ Gottlob Frege, "The Thought: A Logical Inquiry," p. 298.

this perspective necessarily yields a privileged *picture* of what is seen (referred to), and second, that this picture is what is intended when one makes use of the privileged perspective (by saying 'I'). These conclusions, even if correct, are not forced upon us. The character of 'I' provides the acknowledged privileged perspective, whereas the analysis of the content of particular occurrences of 'I' provides for (and needs) no privileged pictures. There may be metaphysical, epistemological, or ethical reasons why I (so conceived) am especially *important* to myself. (Compare: why *now* is an especially important time to me. It too is presented in a particular and primitive way, and this moment cannot be presented at any other time in the same way.)⁵⁷ But the phenomenon noted by Frege—that everyone is presented to himself in a particular and primitive way—can be fully accounted for using only our semantical theory.

Furthermore, regarding the first conclusion, I sincerely doubt that there is, for each of us on each occasion of the use of 'I', a particular, primitive, and incommunicable Fregean self-concept which we tacitly express to ourselves. And regarding the second conclusion: even if Castor were sufficiently narcissistic to associate such self-concepts with his every use of 'I', his twin, Pollux, whose mental life is qualitatively identical with Castor's, would associate the *same* self-concept with *his* every (matching) use of 'I'.⁵⁸ The second conclusion would lead to the absurd result that when Castor and Pollux each say 'I', they do not thereby distinguish themselves from one another. (An even more astonishing result is possible. Suppose that due to a bit of self-deception the self-concept held in common by Castor and Pollux fits neither of them. The second conclusion then leads irresistibly to the possibility that when Castor and Pollux each say 'I' they each refer to a third party!)

The perceptive reader will have noticed that the conclusions of the sloppy thinker regarding the pure indexical 'I' are not unlike those of the Fregean regarding true demonstratives. The sloppy thinker has adopted a *demonstrative theory of indexicals*: 'I' is synonymous with 'this person' [along with an appropriate *subjective* demonstration], 'now' with 'this time', 'here' with 'this place' [each associated with some demonstration], etc. Like the Fregean, the sloppy thinker errs in believing that the

sense of the demonstration is the sense of the indexical, but the sloppy thinker commits an additional error in believing that such senses are in any way necessarily associated with uses of pure indexicals. The slide from privileged perspective to privileged picture is the sloppy thinker's original sin. Only one who is located in the exact center of the Sahara Desert is entitled to refer to that place as 'here', but aside from that, the place may present no distinguishing features.⁵⁹

The sloppy thinker's conclusions may have another source. Failure to distinguish between the cognitive significance of a thought and the thought itself seems to have led some to believe that the elements of an object of thought must each be directly accessible to the mind. From this it follows that if a singular proposition is an object of thought, the thinker must somehow be immediately acquainted with each of the individuals involved. But, as we have seen, the situation is rather different from this. Singular propositions may be presented to us under characters which neither imply nor presuppose any special form of acquaintance with the individuals of the singular propositions. The psychological states, perhaps even the epistemological situations, of Castor and Pollux are alike, yet they assert distinct singular propositions when they each say 'My brother was born before me'. Had they lived at different times they might still have been situated alike epistemologically

⁵⁹ So far, we have limited our attention to the first three sentences of the quotation from Frege. How are we to account for the second part of Frege's remarks?

Suppose Dr. Lauben wants to communicate his thought without disturbing its cognitive content. (Think of trying to tell a color-blind person that the green light should be replaced. You would have to find another way of communicating what you wanted to get across.) He can't communicate *that* thought with *that* significance, so, he himself would have to attach a nonstandard significance to 'I'. Here is a suggestion. He points at his auditor and uses the demonstrative 'you'. If we neglect fine differences in perspective, the demonstration will have the same character for all present and it certainly will have the same demonstratum for all present, therefore the demonstrative will have the same *character and content* for all present. The indexical 'now' will certainly have the same character and content for all present. Thus 'the person who is speaking to you [points] now' will have a common character and content for all those present. Unfortunately the content is not that of 'I' as Dr. Lauben standardly uses it. He needs a demonstrative like 'dthat' to convert the description to a term with a fixed content. He chooses the demonstrative 'he', with a relative clause construction to make clear his intention. Now, if Dr. Lauben uses 'I' with the nonstandard meaning usually attached to 'he who is speaking to you [points] now' he will have found a way to communicate his original thought in a form whose cognitive significance is common to all. Very clever, Dr. Lauben.

[Perhaps it is poor pedagogy to join this fanciful interpretation of the second part of the passage with the serious interpretation of the first part.]

⁵⁷ At other times, earlier and later, we can know it only externally, by description as it were. But now we are directly acquainted with it. (I believe I owe this point to John Perry.)

⁵⁸ Unless, of course, the self-concept involved a bit of direct reference. In which case (when direct reference is admitted) there seems no need for the whole theory of Fregean self-concepts. Unless, of course, direct reference is limited to items of direct acquaintance, of which more below.

while asserting distinct singular propositions in saying 'It is quiet here now'. A kidnapped heiress, locked in the trunk of a car, knowing neither the time nor where she is, may think 'It is quiet here now' and the indexicals will remain directly referential.⁶⁰

E. Corollary 2 *Ignorance of the referent does not defeat the directly referential character of indexicals.*

From this it follows that a special form of knowledge of an object is neither required nor presupposed in order that a person may entertain as object of thought a singular proposition involving that object.

There is nothing inaccessible to the mind about the semantics of direct reference, even when the reference is to that which we know only by description. What allows us to take various propositional attitudes towards singular propositions is not the form of our acquaintance with the objects but is rather our ability to manipulate the conceptual apparatus of direct reference.⁶¹

The foregoing remarks are aimed at refuting *Direct Acquaintance Theories of direct reference*. According to such theories, the question whether an utterance expresses a singular proposition turns, in the first instance, on the speaker's *knowledge of the referent* rather than on the *form of the reference*. If the speaker lacks the appropriate form of acquaintance with the referent, the utterance cannot express a singular proposition, and any apparently directly referring expressions used must be abbreviations or disguises for something like Fregean descriptions. Perhaps the Direct Acquaintance theorist thought that only a theory like his could permit singular propositions while still providing a solution for Frege's problem. If we could *directly* refer to a given object in nonequivalent ways (e.g., as 'dthat[Hes]' and 'dthat[Phos]'), we could not—so he thought—explain the difference in cognitive significance between the appropriate instances of ' $\alpha = \alpha$ ' and ' $\alpha = \beta$ '. Hence, the objects susceptible to direct reference must not permit such reference in inequivalent ways. These objects must, in a certain sense, be wholly local and completely given so that for any two *directly* coreferential terms

⁶⁰ Can the heiress plead that she could not have believed a singular proposition involving the place *p* since when thinking 'here' she didn't *know* she was at *p*, that she was, in fact, unacquainted with the place *p*? No! Ignorance of the referent is no excuse.

⁶¹ This makes it sound as if an exact and conscious mastery of semantics is prerequisite to having a singular proposition as object of thought. I will try to find a better way to express the point in a succeeding draft.

α and β , ' $\alpha = \beta$ ' will be uninformative to anyone appropriately situated, epistemologically, to be able to use these terms.⁶² I hope that my discussion of the two kinds of meaning—content and character—will have shown the Direct Acquaintance Theorist that his views are not the inevitable consequence of the admission of directly referential terms. From the point of view of a lover of direct reference this is good, since the Direct Acquaintance theorist admits direct reference in a portion of language so narrow that it is used only by philosophers.⁶³

I have said nothing to dispute the epistemology of the Direct Acquaintance theorist, nothing to deny that there exists his special kind of object with which one can have his special kind of acquaintance. I have only denied the relevance of these epistemological claims to the semantics of direct reference. If we sweep aside metaphysical and epistemological pseudo-explanations of what are essentially semantical phenomena, the result can only be healthy for all three disciplines.

Before going on to further examples of the tendency to confuse metaphysical and epistemological matters with phenomena of the semantics of direct reference, I want to briefly raise the problem of *cognitive dynamics*. Suppose that yesterday you said, and believed it, "It is a nice day today." What does it mean to say, today, that you have retained *that* belief? It seems unsatisfactory to just believe the same content under any old character—where is the *retention*?⁶⁴ You *can't* believe

⁶² For some consequences of this view with regard to the interpretation of demonstratives see "Bob and Carol and Ted and Alice," appendix VII.

⁶³ There is an obvious connection between the fix in which the Direct Acquaintance Theorist finds himself, and *Kripke's problem*: how can ' $\alpha = \beta$ ' be informative if α and β differ in neither denotation nor sense (nor, as I shall suggest is the case for proper names, character)?

⁶⁴ The sort of case I have in mind is this. I first think, "His pants are on fire." I later realize, "I am he" and thus come to think "My pants are on fire." Still later, I decide that I was wrong in thinking "I am he" and conclude "His pants were on fire." If, in fact, I *am* he, have I *retained* my belief that my pants are on fire simply because I believe the same content, though under a different character? (I also deny that content under the former, but for change of tense, character.) When I first thought "My pants are on fire," a certain singular proposition, call it 'Eek', was the object of thought. At the later stage, both Eek and its negation are believed by me. In this sense, I still believe what I believed before, namely Eek. But this does not capture my sense of *retaining a belief*: a sense that I associate with saying that some people have a very rigid cognitive structure whereas others are very flexible. It is tempting to say that cognitive dynamics is concerned not with retention and change in what is believed, but with retention and change in the characters under which our beliefs are held. I think that this is basically correct. But it is not obvious to me what relation between a character under which a belief is held at one time and the set of characters under which beliefs are held at a later

that content under the same character. Is there some obvious standard adjustment to make to the character, for example, replacing *today* with *yesterday*? If so, then a person like Rip van Winkle, who loses track of time, can't retain any such beliefs. This seems strange. Can we only *retain* beliefs presented under a fixed character? This issue has obvious and important connections with Lauben's problem in trying to communicate the thought he expresses with 'I have been wounded'. Under what character must his auditor believe Lauben's thought in order for Lauben's communication to have been successful? It is important to note that if Lauben said 'I am wounded' in the usual meaning of 'I', there is no one else who can report what he said, using *indirect* discourse, and convey the cognitive significance (to Lauben) of what he said. This is connected with points made in section VIII, and has interesting consequences for the inevitability of so-called *de re* constructions in indirect discourse languages which contain indexicals. (I use 'indirect discourse' as a general term for the analogous form of all psychological verbs.)

A prime example of the confusion of direct reference phenomena with metaphysical and epistemological ideas was first vigorously called to our attention by Saul Kripke in *Naming and Necessity*. I wish to parallel his remarks disconnecting the *a priori* and the *necessary*.

The form of a *prioricity* that I will discuss is that of logical truth (in the logic of demonstratives). We saw very early that a truth of the logic of demonstratives, like "I am here now" need not be necessary. There are many such cases of logical truths which are not necessary. If α is any singular term, then

$$\alpha = \text{dthat}[\alpha]$$

is a logical truth. But

$$\Box(\alpha = \text{dthat}[\alpha])$$

is generally false. We can, of course, also easily produce the opposite effect.

time would constitute retaining the original belief. Where indexicals are involved, for the reasons given below, we cannot simply require that the very same character still appear at the later time. Thus the problem of cognitive dynamics can be put like this: what does it mean to say of an individual who at one time sincerely asserted a sentence containing indexicals that at some later time he has (or has not) *changed his mind* with respect to his assertion? What sentence or sentences must he be willing to assert at the later time?

$$\Box(\text{dthat}[\alpha] = \text{dthat}[\beta])$$

may be true, although

$$\text{dthat}[\alpha] = \text{dthat}[\beta]$$

is not logically true, and is even logically equivalent to the contingency,

$$\alpha = \beta$$

(I call ϕ and ψ logically equivalent when ' $\phi \leftrightarrow \psi$ ' is logically true.) These cases are reminiscent of Kripke's case of the terms, 'one meter' and 'the length of bar x '. But where Kripke focuses on the special epistemological situation of one who is present at the dubbing, the descriptive meaning associated with our directly referential term $\text{dthat}[\alpha]$ is carried in the semantics of the language.⁶⁵

How can something be both logically true, and thus *certain*, and *contingent* at the same time? In the case of indexicals the answer is easy to see.

E. Corollary 3 *The bearers of logical truth and of contingency are different entities. It is the character (or, the sentence, if you prefer) that is logically true, producing a true content in every context. But it is the content (the proposition, if you will) that is contingent or necessary.*

As can readily be seen, the modal logic of demonstratives is a rich and interesting thing.

⁶⁵ A case of a seemingly different kind is that of the logical equivalence between an arbitrary sentence ϕ and the result of prefixing either or both of the indexical operators, 'it is actually the case that' (symbolized 'A') and 'it is now the case that' (symbolized 'N'). The biconditional ' $\phi \leftrightarrow AN\phi$ ' is logically true, but prefixing either '□' or its temporal counterpart can lead to falsehood. (This case was adverted to in footnote 28.) It is interesting to note, in this case, that the parallel between modal and temporal modifications of sentences carries over to indexicals. The foregoing claims are verified by the formal system (sections XVIII and XIX, see especially Remark 3). Note that the formal system is constructed in accordance with Carnap's proposal that the intension of an expression be that function which assigns to each circumstance, the extension of the expression with respect to that circumstance. This has commonly been thought to insure that logically equivalent expressions have the same intension (Church's Alternative 2 among principles of individuation for the notion of sense) and that logically true sentences express the (unique) necessary proposition. Homework Problem: What went wrong here?

It is easy to be taken in by the effortless (but fallacious) move from certainty (logical truth) to necessity. In his important article "Three Grades of Modal Involvement,"⁶⁶ Quine expresses his scepticism of the first grade of modal involvement: the sentence predicate and all it stands for, and his distaste for the second grade of modal involvement: disguising the predicate as an operator 'It is necessary that'. But he suggests that no new metaphysical undesirables are admitted until the third grade of modal involvement: quantification across the necessity operator into an open sentence.

I must protest. That first step let in some metaphysical undesirables, falsehoods. All logical truths are analytic, but they can go false when you back them up to '□'.

One other notorious example of a logical truth which is not necessary,

I exist.

One can quickly verify that in every context, this character yields a true proposition—but rarely a necessary one. It seems likely to me that it was a conflict between the feelings of contingency and of certainty associated with this sentence that has led to such painstaking examination of its 'proofs'. It is just a truth of logic!

Dana Scott has remedied one lacuna in this analysis. What of the premise

I think

and the connective

Therefore ?

His discovery was that the premise is incomplete, and that the last five words

up the logic of demonstratives

had been lost in an early manuscript version.⁶⁷

XVIII. The Formal System

Just to be sure we have not overlooked anything, here is a machine against which we can test our intuitions.

The Language LD

The *Language* LD is based on first-order predicate logic with identity and descriptions. We deviate slightly from standard formulations in using two sorts of variables, one sort for positions and a second for individuals other than positions (hereafter called simply 'individuals').

Primitive Symbols

Primitive Symbols for Two Sorted Predicate Logic

0. Punctuation: (,), [,]
1. Variables:
 - (i) An infinite set of individual variables: \mathcal{V}_i
 - (ii) An infinite set of position variables: \mathcal{V}_p
2. Predicates:
 - (i) An infinite number of m - n -place predicates, for all natural numbers m, n .
 - (ii) The 1-0-place predicate: Exist
 - (iii) The 1-1-place predicate: Located
3. Functors:
 - (i) An infinite number of m - n -place i -functors (functors which form terms denoting individuals)
 - (ii) An infinite number of m - n -place p -functors (functors which form terms denoting positions)
4. Sentential Connectives: $\wedge, \vee, \neg, \rightarrow, \leftrightarrow$
5. Quantifiers: \forall, \exists
6. Definite Description Operator: the
7. Identity: =

⁶⁶ *Proceedings of the XI International Congress of Philosophy* 14, 65–81; reprinted in W. V. Quine, *The Ways of Paradox* (New York: Random House, 1966).

⁶⁷ Again, it is probably a pedagogical mistake to mix this playful paragraph with the preceding serious one.

Primitive Symbols for Modal and Tense Logic

8. Modal Operators: \Box, \Diamond
9. Tense Operators:
 F (it will be the case that)
 P (it has been the case that)
 G (one day ago, it was the case that)

Primitive Symbols for the Logic of Demonstratives

10. Three 1-place sentential operators:
 N (it is now the case that)
 A (it is actually the case that)
 Y (yesterday, it was the case that)
11. A 1-place functor: $dthat$
12. An individual constant (0-0-place i -functor): I
13. A position constant (0-0-place p -functor): Here

Well-formed Expressions

The *well-formed expressions* are of three kinds: formulas, position terms (p -terms), and individual terms (i -terms).

1. (i) If $\alpha \in \mathcal{V}_i$, then α is an i -term
(ii) If $\alpha \in \mathcal{V}_p$, then α is a p -term
2. If π is an m - n -place predicate, $\alpha_1, \dots, \alpha_m$ are i -terms, and β_1, \dots, β_n are p -terms, then $\pi\alpha_1 \dots \alpha_m\beta_1 \dots \beta_n$ is a formula
3. (i) If η is an m - n -place i -functor, $\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n$ are as in 2., then
 $\eta\alpha_1 \dots \alpha_m\beta_1 \dots \beta_n$ is an i -term
(ii) If η is an m - n -place p -functor, $\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n$ are as in 2., then
 $\eta\alpha_1 \dots \alpha_m\beta_1 \dots \beta_n$ is a p -term
4. If ϕ, ψ are formulas, then $(\phi \wedge \psi), (\phi \vee \psi), \neg\phi, (\phi \rightarrow \psi), (\phi \leftrightarrow \psi)$ are formulas
5. If ϕ is a formula and $\alpha \in \mathcal{V}_i \cup \mathcal{V}_p$, then $\forall\alpha\phi$ and $\exists\alpha\phi$ are formulas

6. If ϕ is a formula, then
(i) if $\alpha \in \mathcal{V}_i$, then the $\alpha\phi$ is an i -term
(ii) if $\alpha \in \mathcal{V}_p$, then the $\alpha\phi$ is a p -term
7. If α, β are either both i -terms or both p -terms, then $\alpha = \beta$ is a formula
8. If ϕ is a formula, then $\Box\phi$ and $\Diamond\phi$ are formulas
9. If ϕ is a formula, then $F\phi, P\phi$, and $G\phi$ are formulas
10. If ϕ is a formula, then $N\phi, A\phi$, and $Y\phi$ are formulas
11. (i) If α is an i -term, then $dthat[\alpha]$ is an i -term
(ii) If α is a p -term, then $dthat[\alpha]$ is a p -term

Semantics for LD

LD Structures

Definition: \mathfrak{A} is an LD structure iff there are $\mathcal{C}, \mathcal{W}, \mathcal{U}, \mathcal{P}, \mathcal{T}$, and \mathcal{I} such that:

1. $\mathfrak{A} = \langle \mathcal{C}, \mathcal{W}, \mathcal{U}, \mathcal{P}, \mathcal{T}, \mathcal{I} \rangle$
2. \mathcal{C} is a nonempty set (the set of contexts, see 10 below)
3. If $c \in \mathcal{C}$, then
(i) $c_A \in \mathcal{U}$ (the *agent* of c)
(ii) $c_T \in \mathcal{T}$ (the *time* of c)
(iii) $c_P \in \mathcal{P}$ (the *position* of c)
(iv) $c_W \in \mathcal{W}$ (the *world* of c)
4. \mathcal{W} is a nonempty set (the set of *worlds*)
5. \mathcal{U} is a nonempty set (the set of all *individuals*, see 9 below)
6. \mathcal{P} is a nonempty set (the set of *positions*, common to all worlds)
7. \mathcal{T} is the set of integers (thought of as the *times*, common to all worlds)

8. \mathcal{I} is a function which assigns to each predicate and functor an appropriate *intension* as follows:

- (i) If π is an m - n -predicate, \mathcal{I}_π is a function such that for each $t \in T$ and $w \in \mathcal{W}$, $\mathcal{I}_\pi(t, w) \subseteq (\mathcal{U}^m \times \mathcal{P}^n)$
- (ii) If η is an m - n -place i -functor, \mathcal{I}_η is a function such that for each $t \in T$ and $w \in \mathcal{W}$, $\mathcal{I}_\eta(t, w) \in (\mathcal{U} \cup \{\dagger\})^{(\mathcal{U}^m \times \mathcal{P}^n)}$ (Note: \dagger is a completely alien entity, in neither \mathcal{U} nor \mathcal{P} , which represents an 'undefined' value of the function. In a normal set theory we can take \dagger to be $\{\mathcal{U}, \mathcal{P}\}$.)
- (iii) If η is an m - n -place p -functor, \mathcal{I}_η is a function such that for each $t \in T$ and $w \in \mathcal{W}$, $\mathcal{I}_\eta(t, w) \in (\mathcal{P} \cup \{\dagger\})^{(\mathcal{U}^m \times \mathcal{P}^n)}$

9. $i \in \mathcal{U}$ iff $(\exists t \in T)(\exists w \in \mathcal{W})(\langle i \rangle \in \mathcal{I}_{\text{Exist}}(t, w))$

10. If $c \in \mathcal{C}$, then $\langle c_A, c_P \rangle \in \mathcal{I}_{\text{Located}}(c_T, c_W)$

11. If $\langle i, p \rangle \in \mathcal{I}_{\text{Located}}(t, w)$, then $\langle i \rangle \in \mathcal{I}_{\text{Exist}}(t, w)$

Truth and Denotation in a Context

We write: $\models_{cftw}^{\mathfrak{A}} \phi$ for ϕ , when taken in the context c (under the assignment f and in the structure \mathfrak{A}), is true with respect to the time t and the world w .

We write: $|\alpha|_{cftw}^{\mathfrak{A}}$ for The denotation of α , when taken in the context c (under the assignment f and in the structure \mathfrak{A}), with respect to the time t and the world w

In general we will omit the superscript ' \mathfrak{A} ', and we will assume that the structure \mathfrak{A} is $\langle \mathcal{C}, \mathcal{W}, \mathcal{U}, \mathcal{P}, T, \mathcal{I} \rangle$.

Definition: f is an assignment (with respect to $\langle \mathcal{C}, \mathcal{W}, \mathcal{U}, \mathcal{P}, T, \mathcal{I} \rangle$) iff:

$$\exists f_1 f_2 (f_1 \in \mathcal{U}^{\mathcal{V}_i} \& f_2 \in \mathcal{P}^{\mathcal{V}_p} \& f = f_1 \cup f_2)$$

Definition: $f_x^\alpha = (f \sim \{ \langle \alpha, f(\alpha) \rangle \}) \cup \{ \langle \alpha, x \rangle \}$

(i.e., the assignment which is just like f except that it assigns x to α)

Definition: For the following recursive definition, assume that $c \in \mathcal{C}$, f is an assignment, $t \in T$, and $w \in \mathcal{W}$:

1. If α is a variable, $|\alpha|_{cftw} = f(\alpha)$
2. $\models_{cftw} \pi \alpha_1 \dots \alpha_m \beta_1 \dots \beta_n$ iff $\langle |\alpha_1|_{cftw} \dots |\beta_n|_{cftw} \rangle \in \mathcal{I}_\pi(t, w)$
3. If η is neither 'I' nor 'Here' (see 12, 13 below), then

$$|\eta \alpha_1 \dots \alpha_m \beta_1 \dots \beta_n|_{cftw} = \begin{cases} \mathcal{I}_\eta(t, w)(\langle |\alpha_1|_{cftw} \dots |\beta_n|_{cftw} \rangle), \\ \text{if none of } |\alpha_j|_{cftw} \dots |\beta_k|_{cftw} \\ \text{are } \dagger; \\ \dagger, \text{ otherwise} \end{cases}$$
4. (i) $\models_{cftw} (\phi \wedge \psi)$ iff $\models_{cftw} \phi \& \models_{cftw} \psi$
 (ii) $\models_{cftw} \neg \phi$ iff $\sim \models_{cftw} \phi$
 etc.
5. (i) If $\alpha \in \mathcal{V}_i$, then $\models_{cftw} \forall \alpha \phi$ iff $\forall i \in \mathcal{U}, \models_{cf_i^* tw} \phi$
 (ii) If $\alpha \in \mathcal{V}_p$, then $\models_{cftw} \forall \alpha \phi$ iff $\forall p \in \mathcal{P}, \models_{cf_p^* tw} \phi$
 (iii) Similarly for $\exists \alpha \phi$
6. (i) If $\alpha \in \mathcal{V}_i$, then:

$$|\text{the } \alpha \phi|_{cftw} = \begin{cases} \text{the unique } i \in \mathcal{U} \text{ such that } \models_{cf_i^* tw} \phi, \text{ if} \\ \text{there is such;} \\ \dagger, \text{ otherwise} \end{cases}$$
 (ii) Similarly for $\alpha \in \mathcal{V}_p$
7. $\models_{cftw} \alpha = \beta$ iff $|\alpha|_{cftw} = |\beta|_{cftw}$
8. (i) $\models_{cftw} \Box \phi$ iff $\forall w' \in \mathcal{W}, \models_{cftw'} \phi$
 (ii) $\models_{cftw} \Diamond \phi$ iff $\exists w' \in \mathcal{W}, \models_{cftw'} \phi$
9. (i) $\models_{cftw} F \phi$ iff $\exists t' \in T$ such that $t' > t$ and $\models_{cft'w} \phi$
 (ii) $\models_{cftw} P \phi$ iff $\exists t' \in T$ such that $t' < t$ and $\models_{cft'w} \phi$
 (iii) $\models_{cftw} G \phi$ iff $\models_{cf(t-1)w} \phi$
10. (i) $\models_{cftw} N \phi$ iff $\models_{cf c_T w} \phi$
 (ii) $\models_{cftw} A \phi$ iff $\models_{cf t c_W w} \phi$
 (iii) $\models_{cftw} Y \phi$ iff $\models_{cf(c_T-1)w} \phi$

11. $|dthat[\alpha]|_{cftw} = |\alpha|_{cf_{CTCW}}$
12. $|I|_{cftw} = c_A$
13. $|Here|_{cftw} = c_P$

XIX. Remarks on the Formal System

Remark 1: Expressions containing demonstratives will, in general, express different concepts in different contexts. We call the concept expressed in a given context the *Content* of the expression in that context. The Content of a sentence in a context is, roughly, the proposition the sentence would express if uttered in that context. This description is not quite accurate on two counts. First, it is important to distinguish an *utterance* from a *sentence-in-a-context*. The former notion is from the theory of speech acts, the latter from semantics. Utterances take time, and utterances of distinct sentences cannot be simultaneous (i.e., in the same context). But to develop a logic of demonstratives it seems most natural to be able to evaluate several premises and a conclusion all in the same context. Thus the notion of ϕ being true in c and \mathfrak{A} does not require an utterance of ϕ . In particular, c_A need not be uttering ϕ in c_W at c_T . Second, the truth of a proposition is not usually thought of as dependent on a time as well as a possible world. The time is thought of as fixed by the context. If ϕ is a sentence, the more usual notion of the proposition expressed by ϕ -in- c is what is here called the Content of $N\phi$ in c .

Where Γ is either a term or formula,

we write: $\{\Gamma\}_{cf}^{\mathfrak{A}}$ for The Content of Γ in the context c
(under the assignment f and in the
structure \mathfrak{A}).

Definition:

- (i) If ϕ is a formula, $\{\phi\}_{cf}^{\mathfrak{A}} =$ that function which assigns to each $t \in T$ and $w \in W$, Truth, if $\models_{cftw}^{\mathfrak{A}} \phi$, and Falsehood otherwise.
- (ii) If α is a term, $\{\alpha\}_{cf}^{\mathfrak{A}} =$ that function which assigns to each $t \in T$ and $w \in W$, $|\alpha|_{cftw}$.

Remark 2: $\models_{cftw}^{\mathfrak{A}} \phi$ iff $\{\phi\}_{cf}^{\mathfrak{A}}(t, w) = \text{Truth}$. Roughly speaking, the sentence ϕ taken in context c is *true with respect to t and w* iff the proposition expressed by ϕ -in-the-context- c would be true at the time t if w were the actual world. In the formal development of pages 544, 545, and 546, it was smoother to ignore the conceptual break marked by the notion of *Content in a context* and to directly define *truth in a context with respect to a possible time and world*. The important conceptual role of the notion of Content is partially indicated by the following two definitions.

Definition: ϕ is true in the context c (in the structure \mathfrak{A}) iff for every assignment f , $\{\phi\}_{cf}^{\mathfrak{A}}(c_T, c_W) = \text{Truth}$.

Definition: ϕ is valid in LD ($\models \phi$) iff for every LD structure \mathfrak{A} , and every context c of \mathfrak{A} , ϕ is true in c (in \mathfrak{A}).

Remark 3: $\models (\alpha = dthat[\alpha])$; $\models (\phi \leftrightarrow AN\phi)$; $\models N(\text{Located } I, \text{Here})$; $\models \text{Exist } I$. But, $\not\models \Box(\alpha = dthat[\alpha])$; $\not\models \Box(\phi \leftrightarrow AN\phi)$; $\not\models \Box N(\text{Located } I, \text{Here})$; $\not\models \Box(\text{Exist } I)$. Also, $\not\models F(\phi \leftrightarrow AN\phi)$.

In the converse direction (where the original validity has the form $\Box\phi$) we have the usual results in view of the fact that $\models (\Box\phi \rightarrow \phi)$.

Definition: If $\alpha_1, \dots, \alpha_n$ are all the free variables of ϕ in alphabetical order then the closure of $\phi = AN\forall\alpha_1 \dots \forall\alpha_n \phi$.

Definition: ϕ is closed iff ϕ is equivalent (in the sense of Remark 12) to its closure.

Remark 4: If ϕ is closed, then ϕ is true in c (and \mathfrak{A}) iff for every assignment f , time t , and world w , $\models_{cftw}^{\mathfrak{A}} \phi$.

Definition: Where Γ is either a term or a formula, the Content of Γ in the context c (in the structure \mathfrak{A}) is Stable iff for every assignment f , $\{\Gamma\}_{cf}^{\mathfrak{A}}$ is a constant function (i.e., $\{\Gamma\}_{cf}^{\mathfrak{A}}(t, w) = \{\Gamma\}_{cf}^{\mathfrak{A}}(t', w')$, for all t, t', w , and w' in \mathfrak{A}).

Remark 5: Where ϕ is a formula, α is a term, and β is a variable, each of the following has a Stable Content in every context (in every structure): $AN\phi$, $dthat[\alpha]$, β , I , $Here$.

If we were to extend the notion of Content to apply to operators, we would see that all indexicals (including N , A , Y , and $dthat$) have a Stable Content in every context. The same is true of the familiar logical constants although it does not hold for the modal and tense operators (not, at least, according to the foregoing development).

Remark 6: That aspect of the meaning of an expression which determines what its Content will be in each context, we call the *Character* of the expression. Although a lack of knowledge about the context (or perhaps about the structure) may cause one to mistake the Content of a given utterance, the Character of each well-formed expression is determined by rules of the language (such as rules 1–13 on pages 545 and 546, which are presumably known to all competent speakers. Our notation ' $\{\phi\}_{cf}^{\mathfrak{A}}$ ' for the Content of an expression gives a natural notation for the Character of an expression, namely ' $\{\phi\}$ '.

Definition: Where Γ is either a term or a formula, the *Character* of Γ is that function which assigns to each structure \mathfrak{A} , assignment f , and context c of \mathfrak{A} , $\{\Gamma\}_{cf}^{\mathfrak{A}}$.

Definition: Where Γ is either a term or a formula, the *Character* of Γ is *Stable* iff for every structure \mathfrak{A} , and assignment f , the Character of Γ (under f in \mathfrak{A}) is a constant function (i.e., $\{\Gamma\}_{cf}^{\mathfrak{A}} = \{\Gamma\}_{c'f}^{\mathfrak{A}}$, for all c, c' in \mathfrak{A}).

Remark 7: A formula or term has a Stable Character iff it has the same Content in every context (for each \mathfrak{A} , f).

Remark 8: A formula or term has a Stable Character iff it contains no essential occurrence of a demonstrative.

Remark 9: The logic of demonstratives determines a sublogic of those formulas of LD which contain no demonstratives. These formulas (and their equivalents which contain inessential occurrences of demonstratives) are exactly the formulas with a Stable Character. The logic of demonstratives brings a new perspective even to formulas such as these.

The sublogic of LD which concerns only formulas of Stable Character is not identical with traditional logic. Even for such formulas, the familiar Principle of Necessitation (if $\models \phi$, then $\models \Box\phi$) fails. And so does its tense logic counterpart: if $\models \phi$, then $\models (\neg P\neg\phi \wedge \neg F\neg\phi \wedge \phi)$. From the perspective of LD, validity is truth in every possible *context*. For traditional logic, validity is truth in every possible *circumstance*. Each possible context determines a possible circumstance, but it is not the case that each possible circumstance is part of a possible context. In particular, the fact that each possible context has an agent implies that any possible circumstance in which no individuals exist will not form a part of any possible context. Within LD, a possible context is represented by $\langle \mathfrak{A}, c \rangle$ and a possible circumstance by $\langle \mathfrak{A}, t, w \rangle$. To any $\langle \mathfrak{A}, c \rangle$, there corresponds $\langle \mathfrak{A}, c_T, c_W \rangle$. But it is not the case that to every $\langle \mathfrak{A}, t, w \rangle$ there exists a context c of \mathfrak{A} such that $t = c_T$ and $w = c_W$. The result is that in LD such sentences as ' $\exists x$ Exist x ' and ' $\exists x \exists p$ Located x, p ' are valid, although they would not be so regarded in traditional logic. At least not in the neotraditional logic that countenances empty worlds. Using the semantical developments of pages 543–46, we can define this traditional sense of validity (for formulas which do not contain demonstratives) as follows. First note that by Remark 7, if ϕ has a Stable Character,

$$\models_{cftw}^{\mathfrak{A}} \phi \text{ iff } \models_{c'ftw}^{\mathfrak{A}} \phi$$

Thus for such formulas we can define,

$$\phi \text{ is true at } t, w \text{ (in } \mathfrak{A}) \text{ iff for every assignment } f \text{ and every context } c, \models_{cftw}^{\mathfrak{A}} \phi$$

The neotraditional sense of validity is now definable as follows,

$$\models_T \phi \text{ iff for all structures } \mathfrak{A}, \text{ times } t, \text{ and worlds } w, \phi \text{ is true at } t, w \text{ (in } \mathfrak{A})$$

(Properly speaking, what I have called the neo-traditional sense of validity is the notion of validity now common for a quantified S5 modal tense logic with individual variables ranging over possible individuals and a predicate of existence.) Adding the subscript 'LD' for explicitness, we can now state some results.

- (i) If ϕ contains no demonstratives, if $\models_T \phi$, then $\models_{LD} \phi$
- (ii) $\models_{LD} \exists x$ Exist x , but $\not\models_T \exists x$ Exist x

Of course ' $\Box \exists x \text{ Exist } x$ ' is not valid even in LD. Nor are its counterparts, ' $\neg F \neg \exists x \text{ Exist } x$ ', and ' $\neg P \neg \exists x \text{ Exist } x$ '.

This suggests that we can transcend the context-oriented perspective of LD by generalizing over times and worlds so as to capture those possible circumstances $\langle \mathcal{A}, t, w \rangle$ which do not correspond to any possible contexts $\langle \mathcal{A}, c \rangle$. We have the following result:

- (iii) If ϕ contains no demonstratives,
 $\models_T \phi$ iff $\models_{LD} \Box (\neg F \neg \phi \wedge \neg P \neg \phi \wedge \phi)$.

Although our definition of the neotraditional sense of validity was motivated by consideration of demonstrative-free formulas, we could apply it also to formulas containing essential occurrences of demonstratives. To do so would nullify the most interesting features of the logic of demonstratives. But it raises the question, can we express our new sense of validity in terms of the neotraditional sense? This can be done:

- (iv) $\models_{LD} \phi$ iff $\models_T AN\phi$

Remark 10: Rigid designators (in the sense of Kripke) are terms with a Stable Content. Since Kripke does not discuss demonstratives, his examples all have, in addition, a Stable Character (by Remark 8). Kripke claims that for proper names α, β it may happen that $\alpha = \beta$, though not a priori, is nevertheless necessary. This, in spite of the fact that the names α, β may be introduced by means of descriptions α', β' for which $\alpha' = \beta'$ is not necessary. An analogous situation holds in LD. Let α', β' be definite descriptions (without free variables) such that $\alpha' = \beta'$ is not a priori, and consider the (rigid) terms $dthat[\alpha']$ and $dthat[\beta']$ which are formed from them. We know that:

$$\models (dthat[\alpha'] = dthat[\beta'] \leftrightarrow \alpha' = \beta').$$

Thus, if $\alpha' = \beta'$ is not a priori, neither is $dthat[\alpha'] = dthat[\beta']$. But, since:

$$\models (dthat[\alpha'] = dthat[\beta'] \rightarrow \Box (dthat[\alpha'] = dthat[\beta']))$$

it may happen that $dthat[\alpha'] = dthat[\beta']$ is necessary. The converse situation can be illustrated in LD. Since $(\alpha = dthat[\alpha])$ is valid (see Remark 3), it is surely capable of being known a priori. But if α lacks a Stable Content (in some context c), $\Box(\alpha = dthat[\alpha])$ will be false.

Remark 11: Our 0-0-place i -functors are not proper names, in the sense of Kripke, since they do not have a Stable Content. But they can easily be converted by means of stabilizing influence of 'dthat'. Even $dthat[\alpha]$ lacks a Stable Character. The process by which such expressions are converted into expressions with a Stable Character is 'dubbing'—a form of definition in which context may play an essential role. The means to deal with such context-indexed definitions is not available in our object language.

There would, of course, be no difficulty in supplementing our language with a syntactically distinctive set of 0-0-place i -functors whose semantics requires them to have both a Stable Character and a Stable Content in every context. Variables already behave this way, what is wanted is a class of constants that behave, in these respects, like variables.

The difficulty comes in expressing the definition. My thought is that when a name, like 'Bozo', is introduced by someone saying, in some context c^* , "Let's call the Governor, 'Bozo'", we have a context-indexed definition of the form: $A =_{c^*} \alpha$, where A is a new constant (here, 'Bozo') and α is some term whose denotation depends on context (here, 'the Governor'). The intention of such a dubbing is, presumably, to induce the semantical clause: for all c , $\{A\}_{cf}^{\mathcal{A}} = \{\alpha\}_{cf}$. Such a clause gives A a Stable Character. The context-indexing is required by the fact that the Content of α (the 'definiens') may vary from context to context. Thus the same semantical clause is not induced by taking either $A = \alpha$ or even $A = dthat[\alpha]$ as an axiom.

I think it is likely that such definitions play a practically (and perhaps theoretically) indispensable role in the growth of language, allowing us to introduce a vast stock of names on the basis of a meager stock of demonstratives and some ingenuity in the staging of demonstrations.

Perhaps such introductions should not be called 'definitions' at all, since they essentially enrich the expressive power of the language. What a nameless man may express by 'I am hungry' may be inexpressible in remote contexts. But once he says "Let's call me 'Bozo'", his Content is accessible to us all.

Remark 12: The strongest form of logical equivalence between two formulas ϕ and ϕ' is sameness of Character, $\{\phi\} = \{\phi'\}$. This form of synonymy is expressible in terms of validity.

$$\{\phi\} = \{\phi'\} \text{ iff } \models \Box[\neg F\neg(\phi \leftrightarrow \phi') \wedge \neg P\neg(\phi \leftrightarrow \phi') \wedge (\phi \leftrightarrow \phi')]$$

[Using Remark 9 (iii) and dropping the condition, which was stated only to express the intended range of applicability of \models_T , we have: $\{\phi\} = \{\phi'\}$ iff $\models_T(\phi \leftrightarrow \phi')$.] Since definitions of the usual kind (as opposed to dubbings) are intended to introduce a short expression as a mere abbreviation of a longer one, the Character of the defined sign should be the same as the Character of the definiens. Thus, within LD, definitional axioms must take the unusual form indicated above.

Remark 13: If β is a variable of the same sort as the term α but is not free in α , then $\{\text{dthat}[\alpha]\} = \{\text{the } \beta \text{ AN}(\beta = \alpha)\}$. Thus for every formula ϕ , there can be constructed a formula ϕ' such that ϕ' contains no occurrence of 'dthat' and $\{\phi\} = \{\phi'\}$.

Remark 14: Y (yesterday) and G (one day ago) superficially resemble one another in view of the fact that $\models (Y\phi \leftrightarrow G\phi)$. But the former is a demonstrative whereas the latter is an iterative temporal operator. "One day ago it was the case that one day ago it was the case that John yawned" means that John yawned the day before yesterday. But "Yesterday it was the case that yesterday it was the case that John yawned" is only a stutter.

Notes on Possible Refinements

1. The primitive predicates and functors of first-order predicate logic are all taken to be extensional. Alternatives are possible.
2. Many conditions might be added on \mathcal{P} ; many alternatives might be chosen for \mathcal{T} . If the elements of \mathcal{T} do not have a natural relation to play the role of $<$, such a relation must be added to the structure.
3. When K is a set of LD formulas, $K \models \phi$ is easily defined in any of the usual ways.
4. Aspects of the contexts other than c_A , c_P , c_T , and c_W would be used if new demonstratives (e.g., pointings, You, etc.) were added to the language. (Note that the subscripts A, P, T, W are external parameters. They may be thought of as functions applying to contexts, with c_A being the value of A for the context c.)

5. Special continuity conditions through time might be added for the predicate 'Exist'.
6. If individuals lacking positions are admitted as agents of contexts, 3(iii) of page 543 should be weakened to: $c_P \in \mathcal{P} \cup \{\dagger\}$. It would no longer be the case that: $\models \text{Located } I, \text{ Here}$. If individuals also lacking temporal location (disembodied minds?) are admitted as agents of contexts, a similar weakening is required of 3(ii). In any case it would still be true that $\models \text{Exist } I$.

XX. Adding 'Says'

[This section is not yet written. What follows is a rough outline of what is to come.]

The point of this section is to show, in a controlled experiment, that what Quine called *the relational sense* of certain intensional operators is unavoidable, and to explore the *logical*, as opposed to epistemological, features of language which lead to this result.

I have already mentioned, in connection with Dr. Lauben, that when x says 'I have been wounded' and y wishes to report in indirect discourse exactly what x said, y has a problem. It will not do for y to say ' x said that I have been wounded'. According to our earlier remarks, it should be correct for y to report x 's *content* using a character appropriate to the context of the report. For example, accusingly: 'You said that you had been wounded', or quantificationally: ' $(\exists z)(Fz \wedge x \text{ said that } z \text{ had been wounded})$ ' where x alone satisfied ' Fz '. I will try to show that such constructions are the inevitable result of the attempt to make (third person) *indirect discourse* reports of the first person *direct discourse* sayings when those sayings involve indexicals.

The situation regarding the usual epistemic verbs—'believes', 'hopes', 'knows', 'desires', 'fears', etc.—is, I believe, essentially similar to that of 'says'. Each has, or might have, a *direct discourse* sense in which the character which stands for the cognitive significance of the thought is given (he thinks, 'My God! It is *my* pants that are on fire.') as well as an *indirect discourse* sense in which only the content need be given (he thinks that it is *his* pants that are on fire).⁶⁸ If this is correct, and if indexicals are featured in the language of thought (as suggested

⁶⁸My notion of 'indirect discourse' forms of language is linked to Frege's notion of an 'ungerade' (often translated 'oblique') context. My terminology is intended to echo his.

earlier), then any *indirect* discourse reports of someone's thought (other than first person on the spot reports) must contain those features—*de re* constructions, referential occurrences, quantification in, relational senses—that have so puzzled me, and some others, since the appearance of “Quantifiers and Propositional Attitudes.”⁶⁹

What is special and different about the present approach is the attempt to use the distinction between direct and indirect discourse to match the distinction between character and content. Thus when you wonder, ‘Is that me?’, it is correct to report you as having wondered whether you are yourself. These transformations are traced to the indexical form of your inner direct discourse rather than to any particular referential intentions. The idea is that the full analysis of indirect discourse includes mention of the suppressed character of the direct discourse event which the indirect discourse reports, thus:

$\exists c, C$ [c is a context $\wedge C$ is a character $\wedge x$ is the agent of c
 $\wedge x$ indirect-discourse-verb C at the time t of $c \wedge$ the content
 of C in c is that...]

approximates a full analysis of

x indirect-discourse-verb that ... at t .

Rather than try to include all these semantical ideas in an object language which includes the direct discourse forms of the verbs, the object language will include, *as is usual*, only the indirect discourse forms. The information about the character of the direct discourse event will provide the metalinguistic data against which the truth of object language sentences is tested.⁷⁰

⁶⁹Quine, in his “Reply to Kaplan” in *Words and Objections*, ed. D. Davidson et al. (Dordrecht: Reidel, 1969), raises the question—in the idiom of “Quantifiers and Propositional Attitudes” (*Journal of Philosophy* 53 (1956); reprinted in Martinich, op. cit.)—which of the names of a thing are to count as exportable? My point here is that the indexical names must be exportable, not because of some special justification for the transformation from a *de dicto* occurrence to a *de re* occurrence, but because indexicals are devices of direct reference and have no *de dicto* occurrences. I am reminded of the Zen ko-an: How do you get the goose out of the bottle? Answer: It's out!

⁷⁰If this analysis is correct, the suppressed character should wreak its mischief in cases of suspension of belief (I believe, ‘that man's pants are on fire’ but at the moment neither assent to nor deny ‘my pants are on fire’) as does its counterpart in section XI of “Quantifying In.” Burge, in “Kaplan, Quine, and Suspended Belief,” *Philosophical Studies* 31 (1977): 197–203, proposes a solution to the problem of section XI which he believes is in the spirit of Quine's formulations. A similar

What is not yet clear to me is whether all directly referential occurrences of terms within the scope of indirect discourse epistemic verbs are to be justified *solely* on the basis of a like (though generally distinct) term in the direct discourse event or whether in some cases the English idioms which we symbolize with quantification in (for example, ‘There is someone whom Holmes believes to have shot himself’) involve some element of *knowing-who* or *believing-who*. To put the question another way: are all the cases that Quine describes, and others similar, which irresistibly suggest the symbolic idiom of quantification in, accounted for by the semantics of direct reference (including indexicals and possibly other expressions as well) as applied to the (putative) direct discourse events? “Quantifying In” suffers from the lack of an adequate semantics of direct reference, but its explicandum includes the epistemological idea of knowing-who, which goes beyond what can be analyzed simply in terms of direct reference. When Ingrid hears someone approaching through the fog and knows ‘Someone is approaching’ and even knows ‘That person is approaching’, is it justified to say that there is someone whom Ingrid knows to be approaching? Or must we have, in addition to the indexical ‘that person’, *recognition* on Ingrid's part of who it is that is approaching? My present thought is that the cases which irresistibly suggest the symbolic idiom of quantification in involve, in an ambiguous way, two elements: *direct reference* (on which we are close to getting clear, I hope) and *recognition*.⁷¹ (The latter is my new term

proposal in the present context would seem starkly inappropriate. But there has been a shift in task from “Quantifying In” to the present attempt. In large part the shift is to a course outlined by Burge in the last two pages of the above-mentioned article and urged by him, in conversation, for several years. The point only began to sink in when I came on it myself from a different angle.

⁷¹There is another form of common speech which may be thought to suggest formalization by quantification in. I call this form the *pseudo de re*. A typical example is, “John says that the lying S.O.B. who took my car is honest.” It is clear that John does not say, “The lying S.O.B. who took your car is honest.” Does John say ‘ δ is honest’ for some directly referential term δ which the reporter believes to refer to the lying S.O.B. who took his car? Not necessarily. John may say something as simple as, “The man I sent to you yesterday is honest.” The reporter has simply substituted his description for John's. What justifies this shocking falsification of John's speech? Nothing! But we do it, and often recognize—or don't care—when it is being done. The form lends itself to strikingly distorted reports. As Church has shown, in his *Introduction to Mathematical Logic* (Princeton: Princeton University Press, 1956), on page 25, when John says “Sir Walter Scott is the author of *Waverley*” use of the *pseudo de re* form (plus a quite plausible synonymy transformation) allows the report, “John says that there are twenty-nine counties in Utah!” I do not see that the existence of the *pseudo de re* form of report poses

for knowing-(or believing)-who.) The term is chosen to reflect the idea that the individual in question is identified with respect to some prior or independent information—*re-cognition*—not immediately connected with the current attribution.) Of the two elements the former is semantical; the latter, frankly epistemological. The English idiom ‘There is someone such that Ingrid indirect-discourse-propositional-attitude-verb that ...he ...’ always implies that a singular proposition is the object of Ingrid’s thought (and thus that some directly referential term α occurred in her inner direct discourse) and may sometimes imply (or only suggest?) that Ingrid recognized, *who α is*. I offer no analysis of the latter notion.⁷²

In the first paragraph, I referred to a controlled experiment. By that I mean the following. Accepting the metaphor of “inner direct discourse events” and “indirect discourse reports” in connection with the usual epistemic verbs, I want to examine the logical relations between these two. But the study is complicated by at least three factors which obscure the issues I wish to bring to light. First, there is no real syntax to the language of thought. Thus, even in the case of the simplest thoughts the relation between the syntax of the sentential complement to the epistemic verb and the structure of the original thought is obscure. Second, in containing images, sounds, odors, etc., thought is richer than the language of the report. Might these perceptual elements play a role in determining logical relations? Third, thought ranges from the completely explicit (inner speech) to the entirely implicit (unconscious beliefs which explain actions) and through a variety of occurrent and dispositional forms. This makes it hard to pin down the whole direct discourse event. These three factors suggest taking as a paradigm of the relation between direct and indirect discourse—direct and indirect discourse!

Even when reporting the (outer) discourse of another, at least three obscure irrelevancies (for our purposes) remain. First, if Christopher speaks in a language different from that of the report, we have again the problem of translation (analogous to, though perhaps less severe than,

any issues of sufficient theoretical interest to make it worth pursuing.

⁷²There is a considerable literature on this subject with important contributions by Hintikka, Castañeda and others. In connection with the proposal that ‘ a knows who α is’ can be symbolized ‘ $\exists x(a \text{ knows that } x = \alpha)$ ’, it should be noted that a ’s knowledge of the logical truth ‘ $\text{dthat}[\alpha] = \alpha$ ’ leads, simply by the semantics of direct reference, to ‘ $\exists x(a \text{ knows that } x = \alpha)$ ’. This shows only that a *recognition* sense of knowing a singular proposition is not definable, in the obvious way, in terms of a purely *direct reference* sense of knowing a singular proposition.

that of translating the language of thought). We control this by assuming the direct discourse to be in the language of the indirect discourse report. Second, as Carnap once pointed out to me, if Christopher’s discourse had the form ‘ $\phi \wedge \psi$ ’ even the strictest court would accept as true the testimony, ‘Christopher said that $\psi \wedge \phi$ ’. What logical transformations on the original discourse would be allowed in the report? (If Christopher says ‘ $\exists x$ x is round’, may we report him as saying that $\exists y$ y is round?) We control this by allowing no logical transformations (we are explicating *literal* indirect discourse). Third, if in saying ‘The circle can’t be squared’ Christopher thought that ‘can’t’ was synonymous with ‘should not’ rather than ‘cannot’, should he be reported as having said that the circle can’t be squared? We control this by assuming that our speakers make no linguistic errors.

What then remains of the logic? Is the move from direct discourse to literal indirect discourse not simply the result of disquotation (and decapitalization) plus the addition of ‘that’, as in:

Christopher says ‘the world is round’

\therefore Christopher says that the world is round

?

But how then are we to report Dr. Lauben’s saying, ‘I have been wounded’? Certainly not as, ‘Dr. Lauben says that I have been wounded’!

Even in this highly antiseptic environment, the logic of *says* should provide us with a full measure of that baffling and fascinating *de re* versus *de dicto*, notional versus relational, etc., behavior. And here, using the conceptual apparatus of the semantics of direct reference, we may hope to identify the source of these antics.

[I also hope to distinguish, in discussing reports of self-attribution, *x says that x is a fool*, from *x says-himself to be a fool*.]

XXI. Russell on Egocentric Particulars and Their Disposability

In chapter VII of *Inquiry Into Meaning and Truth*,⁷³ Russell gives a series of atrocious arguments for the conclusion that “[indexicals] are not needed in any part of the description of the world, whether physical or psychological.” This is a happy no-nonsense conclusion for an argument that begins by remarking “A physicist will not say ‘I saw a table’, but like Neurath or Julius Caesar, ‘Otto saw a table’.” [Why Julius Caesar would be provoked to say ‘Otto saw a table’, is unexplained.]

⁷³Bertrand Russell (London: Allen & Unwin, 1940).

Let us examine Russell's conclusion without prejudice to his argument. [What follows is an outline.]

In brief, there are essentially two points. First: if we have both the indexicals and an unlimited supply of unused directly referential proper names, and we can do instantaneous dubbing, then in each context *c* for any sentence ϕ containing indexicals we can produce a sentence ϕ^* whose character is fixed and whose content is the same as that of ϕ in *c*. In this sense, if you can describe it with indexicals you can describe it without.⁷⁴ There are problems: (i) things can change fast and dubbings take time, (ii) the indexicals retain a kind of epistemic priority.

The second point is: given any *prior* collection of proper names, there will be things, times, places, etc., without a name. How do I say something about these unnamed entities? (E.g., how do I tell you that your pants are on fire—now? It may be that nothing in sight, including us, and no nearby time has a name.)

There are two cases. It seems most likely that without indexicals some entities cannot even be uniquely *described*. In this case we are really in trouble (unless Russell believes in the identity of indescribables—objects lacking uniquely characterizing descriptions) because without indexicals we cannot freely introduce new names. If every entity *can* be uniquely described, there is still the problem of not presenting the right content under the right character required to motivate the right action (recall the discussion on pages 532–33). The proposition expressed by 'the pants belonging to the *x* *Fx* are on fire at the *t* *Gt*' is not the proposition I want to express, and certainly does not have the character I wish to convey.⁷⁵

XXII. On Proper Names

[Some thoughts on proper names from the perspective of the formal system are contained in Remark 11, page 551. What follows is the most hastily written section of this draft. I sketch a view that is mainly

⁷⁴I assume here that proper names are not indexicals. I argue the point in section XXII.

⁷⁵Some interesting arguments of a different sort for the indispensability of indexicals are given by Burge in "Belief De Re," *Journal of Philosophy* 74 (1977): 338–62, and by Bar-Hillel in his pioneering work, "Indexical Expressions," *Mind* (1954). In connection with the arguments of Burge and Bar-Hillel it would be interesting to check on some related empirical issues involving linguistic universals. Do all languages have a first person singular form? Do they all have all of the standard indexicals?

negative, without including much supporting argumentation (several of the omitted arguments seem both tedious and tendentious). My current inclination is to drop this whole section from the final draft.]

A *word* is an expression along with its meaning. When two expressions have the same meaning, as with "can't" and "cannot", we call the two words *synonyms*. When two meanings have the same expression, we call the two words *homonyms*. In the latter case we also say that the expression is *ambiguous*. (Probably we would say that the *word* is ambiguous, but accept my terminology for what follows.) In a disambiguated language, semantics can associate meanings with expressions. Even in a language containing ambiguities, semantics can associate a set of meanings with an expression. But given an utterance, semantics cannot tell us what expression was uttered or what language it was uttered in. This is a presemantic task. When I utter a particular vocable, for example, the one characteristic of the first person pronoun of English, you must decide what *word* I have spoken or indeed, if I have spoken any word at all (it may have been a cry of anguish). In associating a word with my utterance you take account of a variety of features of the context of utterance that help to *determine* what I have said but that need not be any *part* of what I have said. My egotism, my intonation, my demeanor, may all support the hypothesis that it was the first person pronoun of English. But these aspects of personality, fluency, and mood are no part of any semantic theory of the first person pronoun. The factors I have cited are not, of course, *critical* for the use of the first person pronoun. What are the criteria? What would definitively settle the question? I don't know. I think this is a very difficult question. But among the criteria there must be some that touch on the utterer's intention to use a word in conformity with the conventions of a particular linguistic community. For proper name words, in part because they are so easily introduced, this aspect of the presemantic determination is especially important.

According to the causal chain or chain of communication theory, there are two critical intentions associated with the use of the proper name word. One is the intention to use the word with the meaning given it by the person from whom you learned the word. The other is the contrary intention to create (and perhaps simultaneously use) a proper name word to refer to a given object irrespective of any prior meanings associated with the expression chosen as a vehicle. One who uses a proper name word with the first intention generally (but not always) believes that someone originated the word by using it with the

second intention, and—according to the causal chain theory—intends to refer to the given object.⁷⁶

In "Bob and Carol and Ted and Alice," appendix IX, I introduce the notion of a *dubbing* for what I took to be the standard form of introduction of a proper name word. That notion has been mistakenly taken to imply—what I deliberately sought to evoke—a formal public ceremony. What I actually had in mind was a use of a proper name word with the second intention: the intention to originate a word rather than conform to a prior usage. Thus a fleeting "Hi-ya, Beautiful" incorporates all the intentional elements required for me to say that a dubbing has taken place. I believe that my notion here is closely related to Donnellan's notion of a *referential use* of a definite description. Donnellan's distinction between referential and attributive uses of definite descriptions is easily and naturally extended to referential and attributive uses of proper names. When the intention to conform to a preestablished convention is absent we have the pure referential use. In this case, when a proper name is in question, I take it that an internal, subjective, dubbing has occurred. When a definite description is in question, again the speaker does not intend to give the expression its conventional meaning although he may intend to *make use* of the conventional meaning in conveying who it is that is being referred to or for some other purpose associated with

⁷⁶There is disagreement as to how the given object must be given to one who introduces a proper name word with the second intention. Must he be acquainted with the object, directly acquainted, *en rapport*, perceiving it, causally connected, or what? My liberality with respect to the introduction of directly referring terms by means of 'dthat' extends to proper names, and I would allow an arbitrary definite description to give us the object we name. "Let's call the first child to be born in the twenty-first century 'Newman 1'." But I am aware that this is a very controversial position. Perhaps some of the sting can be removed by adopting an idea of Gilbert Harman. Normally one would not introduce a proper name or a dthat-term to correspond to each definite description one uses. But we have the means to do so if we wish. Should we do so, we are enabled to apprehend singular propositions concerning remote individuals (those formerly known only by description). Recognizing this, we refrain. What purpose—other than to confound the skeptics—is served by direct reference to whosoever may be the next president of Brazil? The introduction of a new proper name by means of a dubbing in terms of description and the active contemplation of characters involving dthat-terms—two mechanisms for providing direct reference to the denotation of an arbitrary definite description—constitute a form of cognitive restructuring; they broaden our range of thought. To take such a step is an action normally not performed at all, and rarely, if ever, done capriciously. The fact that we have the means—without special experience, knowledge, or whatever—to refer directly to the myriad individuals we can describe, does not imply that we will do so. And if we should have reason to do so, why not?

the act of utterance (as in "Hi-ya, Beautiful"). What is important here is that the speaker intends to be creating a meaning for the expression in question rather than following conventions. Dubbings, whether aimed at introducing a relatively permanent sense for the expression or only aimed at attaching a nonce-sense to the expression, are unconventional uses of language. Dubbings create words.

In many, perhaps most, uses of definite descriptions there is a mixture of the intention to follow convention with the intention to refer to a preconceived individual. The same mixture of 'attributive' and 'referential' intentions can occur with a proper name. If I introduce a name into your vocabulary by means of false introduction ("This is Jaakko Hintikka", but it isn't), you are left with an undiscriminated tangle of attributive (to refer to Jaakko Hintikka) and referential (to refer to the person to whom you were introduced) intentions associated with your subsequent uses of the expression 'Jaakko Hintikka'. There are several ways in which one might attempt to account for these mixed intentions in a general theory of language. First, we might distinguish two notions: speaker's-reference and semantic-reference. The presence of an attributive intention justifies giving the expressions a conventional meaning and thus allows us to claim that preexisting words were used. Whereas the presence of a referential intention (not just a *belief* that the semantic referent is the given object, but an independent intention to refer to the given object) justifies the claim that the speaker is referring to the given object independent of any particular interpretation of the expressions he used as words and independent of whether the utterance has an interpretation as words. A second way of accounting for mixed intentions of this kind is to assume that one of the two intentions must be dominant. If the referential intention dominates, we regard the utterance, on the model of "Hi-ya, Beautiful," as an apt (or inept, as the case may be) introduction of a proper name word (or phrase). Thus, as essentially involving a dubbing. On this way of accounting for mixed intentions, a referential use of an expression would endow the expression with a semantic referent identical with the speaker's referent.⁷⁷

⁷⁷This is not an unnatural way to account for the use of the proper name word in the false introduction case, but it does seem a bit strange in the case of a definite description. In that case it involves hypothesizing that the speaker intended the description expression to have a meaning which made the given object its semantic referent, and only *believed* that the conventional meaning would do this, a belief that he is prepared to give up rather than acknowledge that the semantic referent of his words was not the given object. Something like this seems to happen when descriptions grow capitals, as in 'The Holy Roman Empire', and in other cases as

My aim in the foregoing is to emphasize how delicate and subtle our analysis of the context of utterance must be for the presemantic purpose of determining what words, if any, were spoken. I do this to make plausible my view that—assuming the causal chain theory of reference—proper names are not indexicals. The contextual feature which consists of the causal history of a particular proper name expression in the agent's idiolect seems more naturally to be regarded as determining what word was used than as fixing the content of a single context-sensitive word. Although it is true that two utterances of 'Aristotle' in different contexts may have different contents, I am inclined to attribute this difference to the fact that distinct homonymous words were uttered rather than a context sensitivity in the character of a single word 'Aristotle'. Unlike indexicals like 'I', proper names really are ambiguous. The causal theory of reference tells us, in terms of contextual features (including the speaker's intentions) which word is being used in a given utterance. Each such word is directly referential (thus it has a fixed content), and it also has a fixed character. Therefore, in the case of proper name words, all three kinds of meaning—referent, content, and character—collapse. In this, proper name words are unique. They have the direct reference of indexicals, but they are not context-sensitive. Proper name words are like indexicals that you can carry away from their original context without affecting their content. Because of the collapse of character, content, and referent, it is not unnatural to say of proper names that they have no meaning other than their referent.

Some may claim that they simply use 'indexical' in a wider sense than I (perhaps to mean something like 'contextual'). But we must be wary of an overbroad usage. Is every ambiguous expression an indexical because we look to utterer's intentions to disambiguate? Indeed, is every expression an indexical because it might have been a groan?

If the character and content of proper name words is as I have described it (according to the causal theory), then the informativeness of ' $\alpha = \beta$ ', with α and β proper names, is not accounted for in terms of differences in either content or character. The problem is that proper names do not seem to fit into the whole semantical and epistemological scheme as I have developed it. I claimed that a competent speaker knows the character of words. This suggests (even if it does not imply) that if two proper names have the same character, the competent speaker

knows that. But he doesn't. What is perhaps even more astounding is that I may introduce a new proper name word and send it on its journey. When it returns to me—perhaps slightly distorted phonologically by its trip through other dialects—I can competently take it into my vocabulary without recognizing it as the very same word! Shocking!

In earlier sections of this paper I have tried to show that many of the metaphysical and epistemological anomalies involving proper names had counterparts involving indexicals, and further that in the case of indexicals these wonders are easily explained by an obvious theory. Insofar as I am correct in regarding the anomalies as counterparts, the theory of indexicals may help to break down unwarranted resistance to the causal chain theory. It may also suggest the form of a general semantical and epistemological scheme comprehending both indexicals and proper names. This is not the place to attempt the latter task; my purpose here is simply to show that it is not trivial.⁷⁸ Those who suggest that proper names are merely one species of indexical depreciate the power and the mystery of the causal chain theory.

⁷⁸The issues to be resolved by "a general semantical and epistemological scheme comprehending ... proper names" are such as these. Is the work of the causal chain theory presemantic, as I have claimed? Do proper names have a kind of meaning other than reference? Does the causal chain theory itself constitute a kind of meaning for proper names that is analogous to character for indexicals (but which, perhaps, gives all proper names the same meaning in this sense)? Are proper names words of any particular language? Is there synonymy between proper names that are expressed differently (as there is between 'can't' and 'cannot')? How should we describe the linguistic competence of one who does not know that Hesperus is Phosphorus? Is he guilty of linguistic error? Should we say he does not know what words he speaks? Does he know that 'Hesperus' and 'Phosphorus' are different words? Are they? Is it really possible, as I claim, to account for the semantics of indexicals without making use of the full conceptual resources required to account for the semantics of proper names? I raise these issues—and there are others—within the framework of a hypothetical acceptance of the causal chain theory. There are other issues, of a quite different kind, involved in trying to fill out some details of the causal chain theory itself. For example, if one who has received some particular proper name expression, say, "James", hundreds of times, uses that expression attributively as a proper name, and has in mind no particular source, how do we decide which branch to follow back? The first set of issues seems to me to be largely independent of the details of the relevant causal chains.

well, for example Russell's 'denoting phrases' which do not denote. But it still seems strange.

Afterthoughts

David Kaplan¹

Demonstratives is now being published, after all these years, in the form in which it was written and circulated for all these years.² It is manifestly unfinished. It still retains bracketed metacomments like “[My current inclination is to drop this whole section from the final draft.]” So why have I not cleaned it up and finished it?

Two reasons: a small one and a big one. First and least, I don’t know exactly how to fix some of the sections that now seem wrong, and I don’t yet see exactly how to connect my current thinking, about propositional attitudes and proper names, with indexicals. Last and most, the spirit

¹ © 1989 by David Kaplan.

I am deeply grateful to John Perry, Howard Wettstein, and Joseph Almog, not only for their efforts in planning and executing the conference that resulted in the present volume, but for their patient encouragement of the publication of *Demonstratives* and their good-natured tolerance of the time it has taken me to gather my afterthoughts. Throughout my life, I have had the uncommonly good fortune to fall under the influence of persons of great intelligence, good humor, and tolerance. Principal among these are my wonderful parents, Martha and Irv Kaplan, my inspiring teachers, Rudolf Carnap and Donald Kalish, and my remarkable wife, Renée Kaplan, the *ne plus ultra* of all three qualities.

² I have made the following changes to the circulated text of draft #2. Bibliographical references have been added and the footnotes renumbered. In a few places, a word or a bit of punctuation has been added or a phrase has been moved. I have also corrected a few typographical errors. None of the philosophical errors have been touched. (Thanks to Edward Zalta for his logician’s help with the corrections, and thanks to Ingrid Deiwiks for her typographical skills.)

of the work—the enthusiasm, the confidence, the hesitations—has an integrity that I regard fondly. It reflects its time, the time described in the preface. My own concerns have moved to other topics. I have even felt a resurgence of atavistic Fregeanism. For me to revise *Demonstratives* now would be the intrusion of a third party between the author and his audience.

I had thought of responding to criticisms, of which there have been many over the past decade, several in this very volume, and some quite technically challenging. Unfortunately, I do not have the space to agree in detail with all of them. So instead I have decided to try to look more closely at a few of *Demonstratives*' central concepts.

My reflections are divided into four sections, each of which is intended to be more or less coherent (though I must confess that tangent avoidance has never been my strong suit). The separate sections are somewhat disconnected, as one's afterthoughts tend to be.

Table of Contents

I.	WHAT IS DIRECT REFERENCE?	568
	How does rigid designation come in?	569
	The paradigm of the variable	571
	Taxonomy: semantics and metasemantics	573
	How do the two roads figure in names?	576
	A generic argument for transparency	577
	The notion of Content is central to my account	578
	Are dthat-terms directly referential?	578
	Two interpretations of the syntax and semantics of "dthat"	579
	If "dthat" is an operator	580
	If "dthat" is a demonstrative surrogate	581
	The operator interpretation is more 'natural' for the formal system	581
II.	DO DEMONSTRATIONS COMPLETE DEMONSTRATIVES?	582
	Occurrences	584
	Problems with occurrences of true demonstratives	585
	A refined conception of Context for true demonstratives	585
	The semantic role of directing intentions	587
	Linking true demonstratives	588
III.	WHAT IS CONTEXT?	591
	Context provides parameters	591
	An assignment of values to variables is the parameter needed to determine the referent of a variable	591
	The rule of Necessitation fails for free variables	593
	The actual-world as an aspect of Context	594
	Why the deviant logic?	596
	A word for cognitive value	597
IV.	WHO CAN SAY WHAT?	599
	What we can't do with words:	
	the Autonomy of Apprehension	599
	Subjectivist semantics	600
	Consumerist semantics	602
	A role for language in thought: Vocabulary Power as an epistemological enhancement	603
	Does a name put us in causal contact with the referent?	604
	Naming the nonexistent	607
	Logically proper names	610

I. What is Direct Reference?

Demonstratives was written against my own Fregean upbringing, as was its progenitor "Dthat".³ I aimed to challenge several tenets of Fregean semantics. In particular, I argued that Fregean *Sinn* conflates elements of two quite different notions of meaning. One, which I called *character*, is close to the intuitive idea of linguistic meaning (and perhaps of cognitive content). Another, which I called *content*, is what is said or expressed by an expression in a particular context of use. The *content* of an utterance of a complete sentence is a truth-bearing proposition. Where indexicals are involved, the difference between character and content is quite clear. The *content* of the sentence "Today is my birthday" will vary with speaker and day of utterance. The *character* of the sentence is the common meaning which each language user can deploy to speak of himself and of the day of utterance. It is this common character that determines how the content adapts in the varying contexts of use.

The idea of Content—the what-is-said on a particular occasion—is central to my account. It is this notion that I saw, and continue to see, as the primary idea behind Frege's *Sinn*.⁴ For what I call *directly referential* expressions, among which are indexicals and demonstratives, I argue that the Fregean picture of the relation between *Sinn* (content) and *Bedeutung* (referent) is entirely wrong.

Directly referential expressions are said to refer directly without the mediation of a Fregean *Sinn*. What does this mean? There are two things it might mean. It might mean that the relation between the linguistic expression and the referent is not mediated by the corresponding propositional component, the content or what-is-said. This would be directly contrary to Frege, and it is what I meant. But it also might mean that *nothing* mediates the relation between the linguistic expression and the individual. So stated, this second interpretation is a wildly implausible idea. And it is contrary to the development of the notion of character which occurs in the text. This is *not* what I meant.⁵

³ "Dthat" was written and read in 1970, published in *Syntax and Semantics*, vol. 9, ed. P. Cole (New York: Academic Press, 1978); and reprinted in *The Philosophy of Language*, ed. A. P. Martinich (Oxford: Oxford University Press, 1985).

⁴ My own analysis of the notion, however, is closer to Russell's *signification*, than to Frege's *Sinn*. I have written more recently on the difference between the semantics of Russell and Frege in section VII of "Opacity" (in *The Philosophy of W. V. Quine*, ed. L. E. Hahn and P. A. Schilpp (Illinois: Open Court, 1986)).

⁵ Nor did I mean that whatever mediation takes place is nondescriptional. The question whether some sort of description can be fashioned to give the correct reference for a term is not decisive for direct reference (but see footnote 24 below).

The "direct" of "direct reference" means unmediated by any propositional component, not unmediated *simpliciter*. The directly referential term goes directly to its referent, *directly* in the sense that it does not first pass through the proposition. Whatever rules, procedures, or mechanisms there are that govern the search for the referent, they are irrelevant to the propositional component, to content. When the individual is determined (when *the reference is fixed*, in the language of Saul Kripke⁶), it is loaded into the proposition. It is this that makes the referent prior to the propositional component, and it is this that reverses the arrow from propositional component to individual in the Direct Reference Picture of the Preface to *Demonstratives*.

How does rigid designation come in?

If the individual is loaded into the proposition (to serve as the propositional component) before the proposition begins its round-the-worlds journey, it is hardly surprising that the proposition manages to find that same individual at all of its stops, even those in which the individual had no prior, native presence. The proposition conducted no search for a native who meets propositional specifications; it simply 'discovered' what it had carried in. In this way we achieve rigid designation. Indeed, we achieve the characteristic, direct reference, form of rigid designation, in which it is irrelevant whether the individual exists in the world at which the proposition is evaluated. In *Demonstratives* I took this to be the fundamental form of rigid designation.

So certain was I that this *was* the fundamental form of rigid designation, that I argued (from "systematic considerations") that it must be what Kripke had *intended* despite contrary indications in his writing.⁷

It was not. In a letter (asking that I take his remarks into account in these afterthoughts), Kripke states that the notion of rigid designation he intended is that "a designator *d* of an object *x* is *rigid*, if it designates *x* with respect to all possible worlds where *x* exists, and *never designates an object other than x with respect to any possible world*." This definition is designed to be neutral with regard to the question

⁶ Saul Kripke, "Naming and Necessity," in *Semantics of Natural Language*, ed. G. Harman and D. Davidson (Dordrecht: Reidel, 1972); revised edition published as a separate monograph, *Naming and Necessity* (Oxford: Basil Blackwell, 1980). References are to the revised edition. Also see Saul Kripke, "Identity and Necessity," in *Identity and Individuation*, ed. M. K. Munitz (New York: New York University Press, 1971).

⁷ Footnote 16, *Demonstratives*.

whether a designator can designate an object at a world in which the object doesn't exist. It was motivated, he says, by the desire to avoid getting bogged down in irrelevant discussions of the existence question.⁸

My own discussion of rigid designation was motivated by the desire to highlight the features of rigidity that are associated with direct reference. In the first draft of *Demonstratives* I had actually used the expression "rigid designation" where I now use "direct reference". I thought of my work as delving into the phenomena identified by Donnellan, Putnam, Kripke, and by me in "Dthat". Direct reference was supposed to provide the deep structure for rigid designation, to underlie rigid designation, to explain it. It would never have occurred to me to be 'neutral' about existence.⁹ Existence problems would simply disappear

⁸The view I thought of as manifest in his texts, what I called "the more widely held view," is stated on page 146 of "Identity and Necessity" (I&N) in the words, "In a situation where the object does not exist, then we should say that the [rigid] designator has no referent and that the object in question so designated does not exist." Kripke asserts that this view should not be attributed to him and that it occurs nowhere, explicitly or implicitly, in *Naming and Necessity* (N&N). Regarding the statement in I&N, he writes that it would be somewhat odd if "there was a mysterious change of position between my explicit view in *Naming and Necessity* and 'Identity and Necessity', delivered a month or so later." (This was the reason I used the remark in I&N to resolve the uncertainties of N&N.) He then questions the accuracy of the language of I&N (quoted above), writing "It is also possible, I think, that the sentence is mistranscribed from the tape of the talk. A simple change of 'and' to 'or' in the sentence would make it entirely consistent with what I said in *Naming and Necessity*.... The corrected version would read even better if 'so' were changed to 'though' (an easy mistake in the transcription of an oral presentation)."

It is good to know his mind on this matter, and I regret misrepresenting his views. I cannot, however, feel embarrassed by my reading of the textual evidence. In the course of my discussion of rigid designation in *Demonstratives*, I was careful to cite all the relevant passages. The neutral definition he intended, containing the clause "and never designates an object other than x," does not occur in N&N, I&N, or the new preface to N&N written ten years after the lectures were given. I continue to think that 'the more widely held view', now seen not to be Kripke's view, is the more widely held view.

Proper names are the main topic of N&N. Regarding the rigid designation of proper names, Kripke tells us in the new preface that "a proper name rigidly designates its referent even when we speak of counterfactual situations where that referent would not have existed." It is this view of rigid designation that I had thought he intended all along.

⁹That is, to be neutral on such questions as whether a designator can designate an object at a world in which the object doesn't exist or whether a name from fiction such as "Pegasus" might designate a merely possible object that exists in another possible world. I had stated my views strongly on these issues in appendices X and XI of "Bob and Carol and Ted and Alice," in *Approaches to Natural Language*,

when the underlying, direct reference structure was seen. How could rigid designation not be based on some deeper semantical property like direct reference? It couldn't be an *accident* that names were rigid and descriptions were not.¹⁰

It all seemed of a piece to me: the singular propositions, the direct reference, the rigid designation. And all of it could be illustrated by the case of indexicals, in which the mechanism of direct reference was understood. When I set out to revise the section distinguishing Kripke's notion from mine, I realized that it is easier to explain the difference between A and B if they are not both named "A". I therefore determined to introduce a new expression, and so coined the phrase "direct reference".

If we call a designator that designates the same object at all worlds, irrespective of whether the object exists there or not, an *obstinately* rigid designator,¹¹ then in the usual modal semantics, all directly referential terms will be obstinately rigid (though not every obstinately rigid term need be directly referential).¹² It is obstinate rigidity that I took as the fundamental form of rigidity in *Demonstratives*.

The paradigm of the variable

This conception of direct reference takes the variable under an assignment of value as its paradigm.¹³ In evaluating " Fx " at a world w , we do not ask whether its value exists in w , we only ask what value was *assigned* to the variable before the process of evaluation at w began. Until a value is *assigned* we have nothing to evaluate.¹⁴ Furthermore, and this is important, it is irrelevant *how* " x " gets its value, *how* the

ed. J. Hintikka et al. (Dordrecht: Reidel, 1973).

¹⁰It should be noted, of course, that even an accidental difference between the modal behavior of names and descriptions is sufficient to establish that names are not simply abbreviated descriptions.

¹¹Following a suggestion of Nathan Salmon in *Reference and Essence* (Princeton: Princeton University Press, 1981) p. 34.

¹²An example of an obstinately rigid designator that is not directly referential is given in *Demonstratives*, section IV. It has the form:

The $n[(P \wedge n^2 = 9) \vee (\sim P \wedge 2^2 = n + 1)]$.

¹³See paragraph 3 of the Preface to *Demonstratives*.

¹⁴Until a value is assigned, the entity that is to be evaluated at the possible worlds, whether it be thought of as an open formula or as the content of an open formula, is incomplete. There may not yet be enough information available for it to bear a truth-value.

assignment is made, how the value of "*x*" is *described* when it is assigned to "*x*". All that matters to the evaluation is that "*x*" *has* a particular value.

Pronouns in natural language have often been analogized to variables. Pronouns are lexically ambiguous, having both an anaphoric and a demonstrative use.¹⁵ An anaphoric use of a pronoun is *syntactically bound* to another phrase occurring elsewhere in the discourse. In meaningful discourse, a pronoun not used anaphorically is used demonstratively. As I saw the matter, a demonstrative use of a pronoun was simply a *syntactically free* use. Like a free occurrence of a variable, it requires something extralinguistic, a *demonstration* as I then termed it, to *assign* it a value. Demonstrative and anaphoric occurrences of pronouns can thus be seen to corresponded to free and bound occurrences of variables. What I want to stress is that the difference between demonstrative and anaphoric uses of pronouns need not be conceptualized primarily in terms of lexical ambiguity; it can also be seen in terms of the syntactical distinction between free and bound occurrences of terms. I saw the analogy between variables and pronouns as even closer than had been thought.

I believe that the case of the free pronoun, the demonstrative, can take a lesson from the case of the free variable. As in the case of the free variable, the mechanism by which a value is assigned to a demonstrative, *how* a particular demonstration demonstrates its object, is extralinguistic and thus off-the-record, so to speak. It should not figure in the *content* of what was said. (This, of course, still leaves open the possibility that it might figure in the *cognitive value* of the utterance.) All that matters to the evaluation of what is said (content) is that the demonstrative has a particular value.

Thus my vivid talk about loading the referent into the proposition comes down to this: when using a directly referential term, the *mode of presentation of the referent* (if you will allow a lapse into the Frege idiom) is no part of what is said. Only the referent itself figures in content. Directly referential expressions are *transparent*.¹⁶ Though there may be

¹⁵In "Nomoto inscribed his book" and "Each author inscribed his book," we would ordinarily take "his" to be syntactically bound to "Nomoto" and "Each author". Such syntactically bound uses of pronouns are called *anaphoric*. The same form of words can be used with "his" occurring as a demonstrative, for example, if we were to point at a third party when uttering "his".

¹⁶The sense of transparency I wish to evoke has nothing to do with the contrast between Quinean opacity and Russellian transparency (for which see footnote 30 of my "Opacity"). Rather, it is that of the well-designed computer program in

a complex semantical mechanism that mediates the connection between linguistic expression and referent, that mechanism is unseen in what is said.

Taxonomy: semantics and metasemantics

The inspiration for direct reference was, as reported in "Dthat", the true demonstratives. One does feel initially that in the use of a true demonstrative, not only is one trying to put the object itself into the proposition (direct reference), but that the connection between demonstrative and object, call this *reference*, is also extraordinarily direct as compared with the connection between a definite description and its denotation. Demonstratives are transparent, whereas descriptions are visibly at work, searching, searching, searching. Despite this, there is an elaborate theory of reference for demonstratives in *Demonstratives*.

How should we organize our total semantical theory so as to take account of the mechanisms of direct reference? Some have questioned whether these mechanisms even belong to semantics. I think that it is quite important to get clear on this and certain related taxonomic questions if we are to improve our understanding of the relation of semantics to thought.¹⁷ And I am quite unclear on the subject.

There are several interesting issues concerning what belongs to semantics. The fact that a word or phrase *has* a certain meaning clearly belongs to semantics. On the other hand, a claim about the *basis* for ascribing a certain meaning to a word or phrase does not belong to semantics. "Ohsnay" means *snow* in Pig-Latin. That's a semantic fact about Pig-Latin. The *reason* why "ohsnay" means *snow* is not a semantic fact; it is some kind of historical or sociological fact about Pig-Latin. Perhaps, because it relates to how the language is *used*, it should be categorized as part of the *pragmatics* of Pig-Latin (though I am not really comfortable with this nomenclature), or perhaps, because it is a

which the commands are 'obvious' and the user need not take account of, indeed is usually unaware of, *how* a command is executed. He knows only that to delete you press "Delete". What else?

¹⁷On my understanding of the controversy between Donnellan and Kripke, just such a taxonomic question is one of the central points at issue. See Keith Donnellan, "Reference and Definite Descriptions," *Philosophical Review* 75 (1966): 281-304; reprinted in Martinich, op. cit.; Saul Kripke, "Speaker's Reference and Semantic Reference," in *Contemporary Perspectives in the Philosophy of Language*, ed. P. French, T. Uehling, Jr., and H. Wettstein (Minneapolis: University of Minnesota Press, 1977); Keith Donnellan, "Speaker Reference, Descriptions, and Anaphora," also in *Contemporary Perspectives in the Philosophy of Language*.

fact about semantics, as part of the *Metasemantics* of Pig-Latin (or perhaps, for those who prefer working from below to working from above, as part of the *Foundations of semantics* of Pig-Latin). Again, the fact that "nauseous" used to mean *nauseating* but is coming to mean *nauseated* is a historical, semantic fact about contemporary American English. But neither the reason why the change in semantic value has taken place nor the theory that gives the basis for claiming that there has been a change in meaning belongs to semantics. For present purposes let us settle on *metasemantics*.

Does the historical chain theory (or 'picture' as some are wont to say) of what determines the referent of a proper name belong to semantics or to metasemantics? The critical question seems to be: does the theory state a semantic value of proper names, or does it rather tell us the basis for determining a semantic value for a proper name. Those who believe that the semantic function of a name is completely exhausted by the fact that it *has* a particular referent will regard the historical chain theory as a part of metasemantics. Those who believe that a name *means* something like *the individual who lies at the other end of the historical chain that brought this token to me* will regard the historical chain theory as a part of semantics, as *giving* the meaning rather than as telling us how to discover it. In general, if a referent is all the meaning a name has, then any information used to *fix* the referent is metasemantical. If names have another kind of meaning, another kind of semantic value (mere cognitive value, if not identified with *Sinn* or with *character*, won't do), then the fact that certain information is used to fix the referent may well belong to semantics.¹⁸

Now what about the mechanisms of direct reference? In the case of an indexical, it seems clear that the rule that tells us how the referent varies from one context of use to another, for example the rule that tells us that "yesterday" always refers to the day before the day of utterance, is a part of the meaning of the indexical. It is this kind of meaning that I call *character*. To argue that character belongs to metasemantics, one would have to regard indexicals as systematically ambiguous and as having no meaning at all outside a particular context of use. This is a view that seems reasonable for *generic names*, the kind of name that all us Davids have in common. But it is decidedly implausible for indexicals.

There is also the fact that there is a *logic* of indexicals, a logic whose semantically valid arguments deviate from the classically valid. This in itself seems to argue that the mechanisms by which directly referential expressions determine their referents belong to semantics?¹⁹

Demonstratives seem to me a less certain case, perhaps because my views about their semantics is less certain. However, I do think that the indexical model—a common meaning for all uses of, say, "you", which then determines a referent in a particular context of use—is closer to the truth than the generic name model according to which "you" would be a meaningless symbol available to use in dubbing whoever one addresses.

This suggests a related reason for wanting to place the mechanisms of direct reference outside of semantics. It is the analogy between these mechanisms, which determine the referent of expressions that already bear meaning, and the methods available to *create* meaningful expressions from empty syntactical forms, by dubbings, definitions, and the like. Especially in the case of a true demonstrative, one may feel—wrongly, I believe—that one is *assigning* a meaning to an otherwise empty form. If content were all there is to meaning, then, since the mechanisms of direct reference do determine content, it would be reasonable to claim that such mechanisms belong to metasemantics. But in general, it is incorrect to equate meaning with content, and it is certainly incorrect in the case of indexicals.²⁰

So, as between semantics and metasemantics, I remain of the view that the theory of the mechanisms of direct reference, at least as that theory is developed in *Demonstratives*, in terms of character and content, belongs to semantics.

A second interesting question is whether to call the theory of these mechanisms *semantics* or *pragmatics*. The central role of the notion *context of use* in determining content might incline one to say that the theory of character is semantics, and the theory of content is pragmatics. But *truth* is a property of contents, and one wouldn't want to be caught advocating a pragmatic theory of truth. The problem is that on my analysis, the mechanisms of direct reference operate *before* the familiar semantical notions of truth and denotation come into play. If I continue

¹⁹Or does it? What *does* the fact that there is an interesting logic of indexicals tell us about the taxonomic place of character? If there is no interesting logic of names, does that tell us something?

²⁰It may be correct in the case of proper names, though even there I would be more inclined to equate meaning with referent and to say that referent determines content. I will return to the distinction between the assignment of meaning and the evaluation of meaning in the final section.

¹⁸It is interesting to note that historical chains also have a use in what we might call *metasyntax*. They give the basis for saying that various utterances are utterances of the same word. I will return to historical chains in section IV.

to think, as Carnap taught me,²¹ that the overall theory of a language should be constructed with syntax at the base, semantics built upon that, and pragmatics built upon semantics, I am faced with a dilemma. The mechanisms of direct reference certainly are not *postsemantical*. But equally surely they are not syntactical. Thus I put them in the bottom layer of semantics.²²

Whether semantics or pragmatics, it is important to emphasize that there are two roads from singular terms to individuals. The road through what is said, through the propositional component, through content. And the direct road, outside of what is said, outside content. Both roads belong to the *rules* of the language, and not to the vagaries of individual difference among language users. Both connect language to the world.

How do the two roads figure in names?

In *Demonstratives* I inquire into the semantic mechanisms whereby indexicals and demonstratives are connected to their referents. How might an analogous discussion of names proceed? Without prejudice to any ultimate issues of semantics versus metasemantics, we might begin with a frankly metasemantical inquiry into naming (what I elsewhere²³ call "dubbing") and the process by which a given name can change its referent over time (if, as seems to be the case, it can). These are matters on which, in theory, Fregeans and Direct Reference theorists might agree.

There is a second question: Does the mechanism whereby the referent of a name is determined belong to semantics, as does character, or to metasemantics, as does the mechanism of meaning change? And if the answer is "semantics", there is the third question: Is the mechanism a part of what is said when the name is used? Or, are names transparent so that only the referent itself figures in what is said? It is on this question that direct reference theorists confront Fregeans.²⁴

²¹ *Introduction to Semantics* (Cambridge, Mass.: Harvard University Press, 1942), p. 9.

²² The time may have come to rethink what I think Carnap taught me.

²³ In "Bob and Carol and Ted and Alice."

²⁴ Note that the outcome of the initial discussion may prejudice this tertiary question. Even if the mechanism by which a name is connected to its referent is taken to be a part of semantics, if the mechanism characterizes the referent from the perspective of the context of use, as does the character of an indexical, rather than from a world perspective, it may not be suitable to play the role of propositional constituent. Thus the result of the first inquiry may argue for a direct reference answer to the third question.

Finally, there is the question: Is the expression a rigid designator? This again is a matter on which we may all agree.

In this last connection it is important to see, as I earlier did not consistently see, that even one who believes that a name is connected to its referent by a description that the speaker associates with the name and who further believes that this description is *included as part of what is said when the name is used* can achieve rigidity, even obstinate rigidity, through the use of rigidifying operators. Thus, a Fregean who takes the name "Aristotle" to have as its sense *the pupil of Plato and teacher of Alexander the Great* need only add something like *actuality* to the content in order to account for the rigidity of proper names. We then have something like *the actual pupil of Plato and teacher of Alexander the Great* as the propositional component. Rigid designation without direct reference.²⁵ Well ... not quite *entirely* without direct reference, since the rigidifying operator seems to involve some form of direct reference. But certainly the *name* has not come out directly referential.

But are names merely rigid and not transparent? I, of course, believe not. In some cases arguments that have been given for rigidity can be shown actually to support the stronger claim of transparency, but I will not take up those arguments here.

A generic argument for transparency

There is, however, one generic argument for transparency which seems to apply in many cases of alleged direct reference. It is not a *decisive* argument. Rather, it is a challenge to those who maintain a contrary view.

Many users of the so-called directly referential expressions lack a real understanding of the exact mechanism or rule of reference by which the referent is determined. Though we act *in conformity* with some such rule, we do not invariably know the rule in the sense of being able to articulate it.²⁶ If one could articulate all the cultural rules one conformed

²⁵ I think that this form of rigidity, logical rather than mathematical or metaphysical, falls under what Kripke now calls *de jure* rigidity, which he describes as "the reference of a designator [being] stipulated to be a single object, whether we are speaking of the actual world or of a counterfactual situation" (*Naming and Necessity*, footnote 21 to the new Preface). Note that such descriptions can be used to stipulate the constituents of a possible world, as in "Suppose that the actual author had plagiarized the actual plagiarizer."

²⁶ This is contrary to my claim in *Demonstratives* that the character of pure indexicals is known to every competent speaker. There I claimed that Character = Linguistic Meaning. I still believe that Character captures an important sense of

to, anthropology would be a much easier discipline. In the case of syntax, it is even more obvious that we act in accordance with a complex set of rules which most of us could not even begin to articulate. Children certainly master the use of indexicals, demonstratives, and proper names well before they develop the rather sophisticated conceptual apparatus needed to undertake explicit semantical investigations. If we don't know what the semantical rule is, how could it be part of what we say when we use the relevant expression?

So long as we were able to cling to the illusion that words like "I" and "Aristotle" abbreviate simple descriptions that are immediately available to introspection, we could think that anyone who used such an expression knew how it secured its reference and might express this knowledge in using the word. But who still thinks that nowadays?

The notion of Content is central to my account

To recapitulate: the issue is not whether the information used to determine the referent is descriptive or not. It is rather whether the relevant information, of whatever form, is a part of what is said. Opening an alternative semantic road to reference, one that does not run through content but may nevertheless play a role in the analysis of cognition (belief, knowledge, etc.), may in the end help us all, Fregean and non-Fregean alike, to reach a deeper understanding of the puzzling phenomena that challenged Frege.

As is apparent, the notion of content is central to my way of explaining direct reference. I know that there are some who reject the notion of content. I can't prove that my way of organizing the theoretical apparatus is indispensable. Surely it isn't. But there are observations, intuitions if you will, both in the text of *Demonstratives* and in the formal logic, for which every theory must account. This is indispensable.

Are dthat-terms directly referential?

Some semi-technical meditations on dthat-terms may help to illuminate the notions of content and of direct reference.

As parents soon realize, any worthwhile creation quickly becomes autonomous. Recently I have found myself bemused by my own uses of "dthat".

Linguistic Meaning, but I have become more sceptical about the competence of competent speakers and about our access to what our words mean.

Two interpretations of the syntax and semantics of "dthat"

The penultimate paragraph of section IV of *Demonstratives* warns that the possible world semantics of the formal system in section XVIII obscures the distinction between direct reference and rigid designation. The representation of content as a function from possible worlds does not allow us to distinguish between a directly referential expression and one that is merely obstinately rigid. Both cases are represented by the same function, a constant function. There are two separate reasons for this. First, in this representation the content of a syntactically complex expression does not reflect that complexity. I call this the problem of *multiplying through*, as when the content of " $4 \times (5+4) + 8 \times (7-2) + 6$ " is represented by a constant function to 82. Second, even for syntactically simple expressions, the functional representation captures only the obstinately rigid designation, there is no further distinction among obstinately rigid designators that marks the directly referential ones.²⁷

The representation in possible world semantics tempts us to confuse direct reference and obstinately rigid designation.²⁸ Could anyone have confused them after the clear warning of section IV? Could I have? Yes.

This is very unfortunate, because I coined the term "direct reference" just in order to keep the distinction clear. I find the confusion most evident in connection with dthat-terms, about whose syntax and interpretation I seem to equivocate. On one interpretation, "dthat" is a directly referential singular term and the content of the associated description is no part of the content of the dthat-term. On another interpretation, "dthat" is syntactically an operator that requires syntactical completion by a description in order to form a singular term.²⁹

²⁷ If, as some have hypothesized, an expression is directly referential if and only if it is syntactically simple and obstinately rigid, then the second problem is spurious.

²⁸ If so, why use it? First, because the functional representation is *sufficient* to do the work of *Demonstratives*, namely to show that character and content must be distinguished and to develop a coherent theory within which some unconventional claims about logic, belief, and modality could be grounded. Second, because it is a precise and reliable tool, within the scope of its representational limitations.

²⁹ Properly speaking, since descriptions are singular terms rather than formulas, "dthat" would be a functional expression rather than an operator. But I wish I had made "dthat" into an operator for *this* usage. I wish I had made it into a variable binding operator for which I would write "dthat $x Fx$ " instead of writing "dthat[the $x Fx$]". Then there would have been a much clearer distinction between the two uses of "dthat", and I would not have been led into temptation.

In *Demonstratives* dthat-terms are eliminable in favor of definite descriptions plus the Actually and Now operators (Remark 13, section XIX). It should be noted that this result is not fundamental. It is dependent on the possibilist treatment

If "dthat" is an operator

If "dthat" is an operator, and if the description, which constitutes the operand and thus syntactically completes the singular term, induces a complex element into content, then the correct way to describe "dthat" is as a rigidifier. Complete dthat-terms would be rigid, in fact *obstinately* rigid. In this case the proposition would not carry the individual itself into a possible world but rather would carry instructions to run back home and get the individual who there satisfies certain specifications. The complete dthat-term would then be a rigid description which induces a complex 'representation' of the referent into the content; it would not be directly referential. The operator "dthat" might still be regarded as *involving* direct reference, though its own referent would not be the individual denoted by the complete dthat-term, but, like that of all operators, would be of an abstract, higher-order functional type.³⁰

of variables in the formal semantics. The variables range over all possible individuals, and a primitive predicate of existence is introduced to represent the varying domains of the different possible worlds. This form of language is more expressive than one in which at each world, the variables range only over the individuals of that world and ' β exists' is expressed by ' $\exists x x = \beta$ '. I now incline toward a form of language which preserves the distinction between what *is* (i.e., what the variables range over) and what *exists*, but which does not automatically assume that all possible individuals have *being* (i.e., does not assume that the variables range over all possible individuals).

³⁰The operators "it is actually the case that" and "it is now the case that" could also be thought of as rigidifiers on this model. In all three cases I am somewhat uncomfortable calling the operator directly referential, though they certainly seem to contain a directly referential *element*. Perhaps, in view of the highly abstract nature of their content, the content should be thought of as a complex, only one part of which is induced by direct reference. The operator "it is now the case that" would then be seen as a syntactically complex application of the grammatical formative, "it is the case at — that" to the directly referential term "now". And similarly for the operator "it is actually the case that", which would be seen as a syntactical combination involving application of the same grammatical formative to the term "actuality". Such a treatment would comport better with the suggestion that only names, including "now", "actuality", etc., are directly referential.

Nathan Salmon points out that if one wished to treat species names like "horse" as directly referential, and as having the species *Equus caballus* as referent, it would be required to adopt a similar device regarding the predicate "is a horse", treating it as a syntactically complex application of the grammatical formative "is a" (a kind of copula) to the directly referential term "horse". Salmon is sceptical, but to me this seems natural. The content of the predicate "is a horse" would then be a complex formed of copulation with the species *E. caballus*.

The desire to treat a variety of lexical items as directly referential requires more attention to the distinction between grammatical formatives and those 'pure'

If "dthat" is a demonstrative surrogate

The operator interpretation is not what I originally intended. The word "dthat" was intended to be a surrogate for a true demonstrative, and the description which completes it was intended to be a surrogate for the completing demonstration. On this interpretation "dthat" is a syntactically complete singular term that requires no *syntactical* completion by an operand. (A 'pointing', being extralinguistic, could hardly be a part of syntax.) The description completes the *character* of the associated occurrence of "dthat", but makes no contribution to content. Like a whispered aside³¹ or a gesture, the description is thought of as off-the-record (i.e., off the *content* record). It determines and directs attention to what is being said, but the manner in which it does so is not strictly *part* of what is asserted. The semantic role of the *description* is prepositional; it induces no complex, descriptive element into content. "Dthat" is no more an operator than is "I", though neither has a referent unless semantically 'completed' by a context in the one case and a demonstration in the other. The referent of "dthat" is the individual described (rather than an abstract, higher-order function). It is directly referential.

The operator interpretation is more 'natural' for the formal system

The predominant interpretation of "dthat" in the text seems to be as demonstrative surrogate except, I am sorry to say, in the formal system. There, the natural interpretation is as rigidifying operator. The reason for this is that the 'completing' description has a syntactical reality within the formal language. It plays an essential role in the logic, for example in the theorem of Remark 13 showing that dthat-terms are eliminable. Although Frege claimed that the context of use was part of

lexical items that might be regarded as naming an abstract object, like a species or a color. I would treat "is a bachelor" in the same way as "is a horse". While acknowledging the *metaphysical* differences between a species and *bachelorhood*, the syntactical unity of "horse" and "bachelor" suggests an analogous *semantical* treatment. Keith Donnellan makes this point in "Putnam and Kripke on Natural Kinds," in *Knowledge and Mind*, ed. C. Ginet and S. Shoemaker (Oxford: Oxford University Press, 1983), pp. 84–104, especially section III. Also, I would go further in syntactical decomposition and first form the complex denoting phrase "a horse" (with appropriate content) before forming the predicate "is a horse".

³¹This is how Kripke characterized the description which completes a dthat-term in his lecture at the conference.

"the means of expression" of a thought,³² he never, to my knowledge, attempted to incorporate "the pointing of fingers, hand movements, glances" into logical syntax. Can an expression such as the description in a *dthat*-term appear in logical syntax but make no contribution to semantical form? It would seem strange if it did. But there is, I suppose, no strict contradiction in such a language form.

If there are two different interpretations of "*dthat*" in *Demonstratives*, they seem to be run together in footnote 72. But maybe there aren't. *Probably* there aren't. Probably, I was just farsighted in envisioning yet-to-be-realized forms of formal semantics. I earlier held that my views were inconsistent. I now deny that my views are inconsistent!³³

II. Do Demonstrations Complete Demonstratives?

In *Demonstratives* I took the demonstration, "typically, a (visual) presentation of a local object discriminated by a pointing," to be criterial for determining the referent of a demonstrative. While recognizing the teleological character of most pointing—it is typically directed by the speaker's intention to point at a perceived individual on whom he has focused—I claimed that the *demonstration* rather than the *directing intention* determined the referent.³⁴

I am now inclined to regard the directing intention, at least in the case of perceptual demonstratives, as criterial, and to regard the demonstration as a mere *externalization* of this inner intention. The externalization is an aid to communication, like speaking more slowly and loudly, but is of no semantic significance.³⁵

³²Gottlob Frege "The Thought: A Logical Inquiry," *Mind* 65 (1956): p. 296. Original German publication in *Beiträge zur Philosophie des Deutschen Idealismus* (1918–19).

³³Thanks to Nathan Salmon and Joseph Almog for help with this section.

³⁴This view goes back to the case, discussed in "Dthat", of Carnap's picture. I now regard this as a rather complex, atypical case.

³⁵I contrast *no* semantic significance with the fundamental idea of direct reference: that there are matters of semantic significance which do not appear in content. In my earlier treatment, I regarded demonstrations as off-the-record in terms of content, but as semantically relevant in determining character. I now regard them as totally off-the-record in regard to the semantics of demonstratives. I now see demonstrations as playing the same role for true demonstratives as does pointing at oneself when using the first-person pronoun.

We might think of the demonstration on the model of a term in *apposition* to the demonstrative. Such a term appears to duplicate the demonstrative syntactically, but its semantic contribution is to a subordinate, side remark; its semantic contribution to the main clause seems to be only to hold targets for anaphora. (I know

I had rejected this view earlier, in part because it seemed to confound what Donnellan might call the *referential* and the *attributive* uses of a demonstrative. It seemed to me that this should not happen in a proper semantical theory. I recently realized that the distinction still held. In the case of a perceptual demonstrative, the directing intention is aimed at a perceived object. This object may or may not be the object the speaker has in mind. We can distinguish between Donnellan's kind of having-in-mind and perceptual focus.³⁶

A benefit of the view that the demonstration is a mere externalization of the perceptual intention, which determines the referent, is that it offers a new perspective on one of Donnellan's most compelling cases of referential use.

Suppose someone is at a party and, seeing an interesting looking person holding a martini glass, one asks, "Who is the man drinking a martini?" If it should turn out that there is only water in the glass, one has nevertheless asked a question about a particular person.³⁷

Because of the importance of the perceptual element it is tempting to think of this case in terms of demonstratives. Here the directing intention is aimed at the interesting looking person seen holding a martini glass. Had the speaker pointed and said "Who is that man?", the case would have raised no question of referential use. But suppose, having been taught that it is rude to point at people, the normal mode of externalizing the intention is unavailable. What to do? He cannot simply say, "Who is that man?" with *no* externalization. This would baffle his auditor, who would say, "Which man?". To which the original speaker would have to reply, "The man with the martini." So he shortens the dialogue and uses the description "the man with the martini" as a substitute for the demonstration. Here the speaker might equally well have said, "Who is that man with the martini?" or, "Who is that?" followed by an appositive, parenthetical, whispered "(the man with the martini)."

Now according to my new view of what determines the referent of a demonstrative, the demonstration (here, the description) is there only to

of no well-developed semantics of apposition; it seems a topic worth pursuing.)

³⁶Just as it is possible to *misdescribe* a perceived object, for example, as a martini when it is really only water in a martini glass, so it is also possible to *misrecognize* one. For example, I may have you in mind, and believing that it is you whom I see hiding under the bed, begin berating you. Even if it was not you under the bed, might it not still be you whom I criticized?

³⁷Keith Donnellan, "Reference and Definite Descriptions".

help convey an intention and plays no *semantical* role at all. We might sum up the case by saying the speaker had a demonstrative intention and, constrained by the conventions of polite behavior, substituted a description for the usual pointing.³⁸ The slight misdescription has no more effect on the determination of the referent of the tacit demonstrative than would a slight error in aim have had on the determination of the referent of a vocalized demonstrative accompanied by a pointing. In both cases the referent is properly determined by the perceptual intention. In neither case is anything semantical at stake in the description or the pointing. All that is at stake is the accuracy of *communicating* what was said.

What makes this analysis especially intriguing is that this classical case of the referential use of a description can be seen as an *attributive* use of a tacit perceptual demonstrative.

Not all of Donnellan's cases can be accounted for in this way. And in any case, as I have already stated, I believe the distinction between referential and attributive uses is fundamental. But still the idea of finding a role for nonsemantic, communication facilitators, and accounting for referential uses of definite descriptions in this way, is appealing. The theory of direct reference, with its prepropositional semantics, seems especially open to such off-the-record elements in language.

Occurrences

As I carefully noted in *Demonstratives*,³⁹ my notion of an *occurrence* of an expression in a context—the mere combination of the expression with the context—is not the same as the notion, from the theory of speech acts, of an *utterance* of an expression by the agent of a context. An occurrence requires no utterance. Utterances take time, and are produced one at a time; this will not do for the analysis of validity. By the time an agent finished uttering a very, very long true premise and began uttering the conclusion, the premise may have gone false. Thus even the most trivial of inferences, *P* therefore *P*, may appear invalid. Also, there are sentences which express a truth in certain contexts, but not if uttered. For example, "I say nothing." Logic and semantics

³⁸A quite different summary would deny the demonstrative element and say that the conventions of polite behavior constrain the speaker to use descriptions and not to use demonstratives. This yields Donnellan's original analysis. Accept my summary. (Is there a basis in the speaker's intentions for claiming that a description is, or is not, being used in apposition to a tacit demonstrative?)

³⁹Section XIII.

are concerned not with the vagaries of actions, but with the verities of meanings.⁴⁰

Problems with occurrences of true demonstratives

On the theory of true demonstratives in *Demonstratives*, a demonstration accompanies every demonstrative and determines its referent. On my current view, the referent of a true demonstrative is determined by the utterer's intention. But if occurrences don't require utterances, how can we be sure that the requisite intention exists in every possible context? We can't!

A version of this problem already existed in a proposal considered in *Demonstratives* for the formal treatment of "you".⁴¹ The idea is that the context simply be enriched by adding a new feature, which we might call the *addressee*. But suppose there is no addressee. Suppose the agent intends no one, e.g., Thomas Jefferson, dining alone, or surrounded by friends but not *addressing* any of them. Or, suppose the agent is hallucinatory and, though addressing 'someone', no one is there.⁴² The problem is that there is no *natural* addressee in such contexts, and thus no natural feature to provide within a formal semantics.

A refined conception of Context for true demonstratives

There are really two problems here, calling for separate solutions. The first is the case of the absent intention. In this case one would want to mark the context as *inappropriate* for an occurrence of "you", and rede-

⁴⁰I am unclear even as to what arguments *ought* to come out as utterance-valid (as opposed to occurrence-valid). There are different notions of utterance-validity corresponding to different assumptions and idealizations. With no idealizations, the rules of repetition and double negation become invalid. This seems hopeless. Should we assume then that utterances take no time? (We might imagine writing out premises and conclusion ahead of time and holding up the paper at the moment of assertion.) Should we assume that the agent knows the language? Should we assume that the agent *asserts* the premises and conclusion, that he *believes* them? This last is related to the question: should "*P*, but I don't believe it" (Moore's paradox) come out to be an utterance-contradiction? It certainly is not an occurrence-contradiction.

⁴¹Possible Refinement #4 of section XIX combined with the 'indexical theory of demonstratives' of section XVI. The idea is considered, not advocated.

⁴²I have in mind the classic hallucination involving an imagined person, not a hallucination of an actual person who happens not to be present.

fine validity as truth-in-all-appropriate-possible-contexts.⁴³ The second is the case of the hallucinatory agent. Here the context seems appropriate enough, the agent is making no *linguistic* mistake in using "you". But the occurrence should be given a 'null' referent.⁴⁴

Another proposal I have heard is just to *impose* an intention on the agent whether he has it or not. Put more gently, this is a logician's proposal; just *assign* a referent. There are two problems with this. First, if it is *possible* for the agent to intend the proposed addressee, there will already be a possible context in which he does. So nothing is lost by ignoring the context in which he doesn't. And if it is *not* possible for the agent to intend the proposed addressee, the imposition seems much too heavy handed. (We don't want an impossibility to come out true.) Second, if we are impatient with intention and just want to *assign* away and get on with the logic, we could formulate the expression with free variables instead of demonstratives. And we should. Why pretend that real demonstratives are nothing more than free variables? If the logic of real demonstratives turns out to be identical with the logic of free variables, well ... that's something that should *turn out*. It shouldn't be presupposed.⁴⁵

We must make one further refinement in our conception of a context for a true demonstrative. The same demonstrative can be repeated, with a distinct directing intention for each repetition of the demonstrative. This can occur in a single sentence, "You, you, you, and you can leave, but you stay", or in a single discourse, "You can leave. You must stay." Such cases seem to me to involve an exotic kind of ambiguity, perhaps unique to demonstratives (see below). Where different intentions are associated with different syntactic occurrences⁴⁶ of a true demonstrative, we would want to use distinct symbols in our formal language in order to avoid equivocation.

Why do we not need distinct symbols to represent different syntactic occurrences of "today"?⁴⁷ If we speak slowly enough (or start just

before midnight), a repetition of "today" will refer to a different day. But this is only because the context has changed. It is a mere technicality that utterances take time, a technicality that we avoid by studying expressions-in-a-context, and one that might also be avoided by tricks like writing it out ahead of time and then presenting it all at once. It is no part of the *meaning* of "today" that multiple syntactic occurrences must be associated with different contexts. In contrast, the meaning of a demonstrative requires that each syntactic occurrence be associated with a directing intention, several of which may be simultaneous. And if it happened to be true that we never held more than one such intention simultaneously, *that* would be the mere technicality. In fact, it is not true. In the aforementioned cases ("You, you, you, and you ..."), in which there is simultaneous perception of all addressees, I think it correct to say that are several distinct, simultaneous, directing intentions, indexed to distinct intended utterances of the demonstrative "you" (which are then voiced one at a time).

The basic fact here is that although we must face life one *day* at a time, we are not condemned to perceive or direct our attention to one *object* at a time. (If we were, the language of thought would be monadic predicate logic.)

Thus within the formal syntax we must have not one demonstrative "you", but a sequence of demonstratives, "you₁", "you₂", etc., and within the formal semantics the context must supply not a single addressee, but a sequence of addressees, some of which may be 'null' and all but a finite number of which would presumably be marked *inappropriate*.

We will need to be able to formulate sentences of the formal language in which different intentions are associated with different syntactic occurrences of a demonstrative, if we are to face the looming challenge of Frege's Problem, in which one who is simultaneously perceiving two parts of what may or may not be a single object asserts, "That₁ is that₂".⁴⁸

The semantic role of directing intentions

What should we think of as the contextual feature relevant to the evaluation of a demonstrative? In the formal semantics, it may be taken to

⁴³The idea, once broached, of defining validity in terms of *appropriate* contexts might also be used to approach utterance-validity.

⁴⁴There are several ways to accommodate this in a formal semantics. I am imagining a treatment along the lines of my use of † in section XVIII of *Demonstratives*.

⁴⁵There are morals to be drawn from these arguments. I urge the young author of *Demonstratives* to take them to heart if he wishes to do serious work.

⁴⁶I say *syntactic* occurrence to differentiate from my expression-in-a-context sense of "occurrence".

⁴⁷I choose "today" rather than "now" to avoid the distracting issue of the vagueness of "now".

⁴⁸Consider, for example, a magician performing the 'sawing a woman in half' illusion. The audience sees someone's head sticking out of one end of a box and what appear to be someone's feet sticking out of the other end. "Is that person really that person?" they wonder.

be the demonstratum. But at the preformal level, I think of it as the *directing intention*. The directing intention is the element that differentiates the 'meaning' of one syntactic occurrence of a demonstrative from another, creating the *potential* for distinct referents, and creating the *actuality* of equivocation.⁴⁹ It also seems critical for the 'cognitive value' of a syntactic occurrence of a demonstrative, at least for the speaker. Note however that it is neither character, content, nor referent. In the case of the pure indexicals, "today", "here", etc., the relevant contextual feature is always the referent, and there doesn't seem to be any role, let alone a semantic role, for a comparable entity. Curiouser and curiouser!

In *Demonstratives* I accepted "tentatively and cautiously" what I called the *Fregean theory of demonstrations*. The demonstration—a 'manner of presentation' of an individual that was separable from any particular context and could be evaluated at other contexts and circumstances—supplied the character for the associated demonstrative.⁵⁰ A reason why I favored the Fregean theory of demonstrations was that the need for a completing demonstration distinguished the true demonstratives from the pure indexicals. A second reason was that the Fregean idea that *that very demonstration* might have picked out a different demonstratum, an idea that depended on the separability of a demonstration from a particular context, seemed to track very closely the cognitive uncertainties of "that₁ is that₂". This cognitive value appears in character, and thus as an aspect of meaning.

The need for a directing intention to determine the referent of a demonstrative still allows us to distinguish the true demonstratives from the pure indexicals. The parameters for the latter are brute facts of the context, like location and time. But if directing intentions are not separable and evaluable at other points (perhaps they are), the cognitive uncertainties of "that₁ is that₂" may no longer be an aspect of meaning. Should they be?

Linking true demonstratives

It is interesting to note that in natural language every new syntactic occurrence of a true demonstrative requires not just a referent-determining intention, but a *new* referent-determining intention. When two syntactic occurrences of a demonstrative appear to be linked to a single intention,

⁴⁹I regard it as an equivocation whenever a new directing intention is involved, even if it directs a second syntactic occurrence of a demonstrative to the same referent.

⁵⁰See sections XV and XVI of *Demonstratives*.

at least one must be anaphoric. When we wish to refer to the referent of an earlier demonstrative, we do not repeat the demonstrative, we use an anaphoric pronoun, "He [pointing] won't pass unless he [anaphoric pronoun] studies." The fact that demonstrative and anaphoric pronouns are homonyms may have led to confusion on this point. The case is clearer when the demonstrative is not homonymous with the anaphoric pronoun. Contrast, "This student [pointing] won't pass unless he [anaphoric pronoun] studies" with "This student [pointing] won't pass unless this student [pointing a second time at what is believed to be the same person] studies." The awkwardness of the second, shows that the way to *secure* a second reference to the referent of a demonstrative, is to use an anaphor.

This implies that it is impossible to utter an instance of the rule of Double Negation using a premise containing a demonstrative, "You stay. Therefore, it is not the case that you do not stay." We have a Hobson's choice. We can intend the "you" of the conclusion as anaphoric across the sentential barrier to the "You" of the premise (something we readily do in ordinary discourse, but are ill-prepared to do in formal logic).⁵¹ In which case, the argument is valid, but not really an instance of Double Negation (at least not as we know and love it). Or, we can concentrate, try not to blink, and try to hold our attention on the same addressee, in the hope that we will succeed in targeting the same individual with the second demonstrative. (Can we ever be *certain* that they haven't pulled the old switcheroo?) In this case, the form of argument is really something like, "You₁ stay. Therefore, it is not the case that you₂ do not stay", and hence not valid. Even if we idealize the speed of speech, so that we are certain that they haven't pulled a switcheroo, the *form* of the argument is still not that of Double Negation because of the equivocation involved in the use of a second demonstrative.

Perhaps we should give up on Double Negation, and claim that the argument is a valid enthymeme with the implicit premise "You₁ = you₂", the premise we strove to make true by fixing our attention. "All right," said the tortoise to Achilles, "repeat the argument and this time remember to utter the additional premise."

The source of the difficulty is the principle, the correct principle,

⁵¹It would be good if our formal language allowed variables to be bound to arbitrary terms both within the sentence and across the sentential barrier in the way in which anaphoric reference takes place in natural language. The problem of how to do this in a suitably smooth way seems quite interesting.

that every new syntactic occurrence of a demonstrative (one that is not a disguised anaphoric pronoun) requires its own determining intention. The problem, in a nutshell, is that where demonstratives are involved, it doesn't seem possible to *avoid* equivocation. There is an understood, harmless, systematic equivocation built into the semantics of demonstratives in natural language. It is this that I termed "an exotic kind of ambiguity, perhaps unique to demonstratives."

For purposes of logic, on the other hand, it seems essential both to avoid equivocation and to allow any well-formed expression to have multiple syntactic occurrences (in antecedent and consequent, or in premise and conclusion) *without changing its semantical analysis*. The validity of the sentence "If you stay, you stay" (with no anaphors) depends on using the same intention to determine the referent of both occurrences of the demonstrative "you". Just as multiple occurrences of "now" in a single argument must be referenced to the same time parameter, so multiple occurrences of the same demonstrative must be referenced to the same directing intention. Otherwise the language would suffer the same systemic equivocation that natural language does, and there would be no logic, at least none with Double Negation and Repetition and the like. Using the refined conception of context described above, it is easy to write semantical rules that give the same analysis to recurrences of the same demonstrative (what is hard is to write rules that don't). It seems certain that this is how we ought to proceed.

But does it leave our logic vulnerable to a charge of misrepresentation? What is it that we hope to learn from such a logic? I don't think we can regard this as an idealization comparable to that involved in referencing all occurrences of "now" to a single instant. To assume that one intention can drive two occurrences of a demonstrative seems more falsification than idealization.

I hope that there is a key to this problem in my earlier remark that logic and semantics are concerned not with the vagaries of actions, but with the verities of meanings. There is something I'm not understanding here, and it may be something very fundamental about the subject matter of logic.

III. What is Context?

Context provides parameters

Some directly referential expressions, most notably the indexicals, require that the value of a certain parameter be given before a determinate element of content is generated. Context of Use is this parameter. For example, the content of the word "today" is a function of the time of the context of use. If we think of the formal role played by context within the model-theoretic semantics, then we should say that context *provides* whatever parameters are needed.⁵² From this point of view, context is a package of whatever parameters are needed to determine the referent, and thus the content, of the directly referential expressions of the language.

An assignment of values to variables is the parameter needed to determine the referent of a variable

Taking context in this more abstract, formal way, as providing the parameters needed to generate content, it is natural to treat the assignment of values to free occurrences of variables as simply one more aspect of context. My point is taxonomic. The element of content associated with a free occurrence of a variable is generated by an assignment. Thus, for variables, the assignment supplies the parameters that determines content just as the context supplies the time and place parameters that determine content for the indexicals "now" and "here".

The assignment, as I am arguing we should conceive of it, is not 'evaluating' the variable at a world, rather it is generating an element of content, and it is the content which is then evaluated at a world.⁵³ Content is generated at a context, and each context is associated with a particular possible world.⁵⁴ The agent, time, and place are all drawn from that world. Similarly, an assignment associated with a particular

⁵² This, rather than saying that context *is* the needed parameter, which seems more natural for the pretheoretical notion of a *context of use*, in which each parameter has an interpretation as a natural feature of a certain region of the world.

⁵³ I know, I know! There are other ways to treat assignments, but they obscure my point. Having returned to the semantics of free variables, it may seem that I am obsessed with the topic, but bear with me.

⁵⁴ When I revert to the standard "possible worlds" nomenclature rather than the "possible circumstance of evaluation" terminology of *Demonstratives*, it is in order to connect certain points I wish to make with the standard literature. I use the two phrases synonymously.

context may be taken to assign only values that exist in the world of the context. Once such a value is assigned, that is, once a content is determined, the content can, of course, be evaluated at worlds in which the value does not exist.

In arguing that assignments of values to variables play a theoretical role analogous to contexts, I harp upon my theme that free variable can be taken as paradigms of direct reference. Though the theme was stated in *Demonstratives*, I did not then recognize how thoroughgoing it was, because I did not then think of free variables in the robust way I now do, as demonstrative uses of pronouns. Not as real demonstratives, which require a directing intention from the agent of the context, but as a kind of *faux demonstrative*, one which looks real until you check into the origin of its value.

As remarked above, free occurrences of pronouns in meaningful discourse are demonstratives. But a free occurrence of an *anaphoric* pronoun would literally be meaningless. In our logical formalisms, variables play the anaphoric role. Thus a free occurrence of a variable is the mark of an incompletely interpreted expression. The case we are dealing with here is the free occurrence of a variable in a premise or conclusion of an argument. Do not confuse this case, the case with the interpretational gap, with the case in which a *bound* occurrence of a variable *appears* free because we are focusing attention on a subformula. It is the second case, the case of *bound* variables, for which the Tarski apparatus of *satisfaction* and *assignments* was originally designed. In that case there is no interpretational gap; it is the *quantifier* (or other variable binder) that is being interpreted, and we must get it right. So the rules for evaluating *bound* occurrences of variables are another story entirely, and an irrelevant one.

That which is interpretively unconstrained is available for office, and those familiar with logic will be aware that authors of deductive systems have chosen varying paths in their treatment of free variables. Some prohibit them entirely. Some treat them as if they were bound by invisible, outer, universal quantifiers, what is sometimes called the *generality* interpretation. Some treat them as if they were individual constants. My own treatment uses the familiar idea of an *assignment*, taken from the Tarski apparatus for the treatment of bound variables. I even confine the values of the variables to the domain of quantification (assuming the domain of quantification consists of what exists). This seems natural enough. But it does, as will be seen, have surprising consequences.

The discussion of parameters completes the analogy between free

variables and indexicals. From an abstract formal point of view, they are highly analogous. Both are *parametric*, their content varies as the parameter varies. If we package all parameters under the heading *context*, an odd but interesting thing to do, we could even claim that content varies with context, the mark of indexicality. (Note that not all directly referential expressions are parametric; proper names are not.)⁵⁵

These formal analogies should not cause us to lose sight of the fundamental difference between free variables and indexicals.⁵⁶ Indexicals are real, meaning-bearing elements of language. Free variables are not; they are artifacts of our formalism. Assignments are *stipulative*; they have no fact-of-the-matter parameter as do the pure indexicals and true demonstratives. Indexicals are *perspectival*, their content is dependent on the speaker's point of view, the context of utterance. Free variables are not *perspectival* in any but the most attenuated metaphorical sense. It is for these reasons that I use the term *parametric* for what indexicals and free variables have in common.⁵⁷

The rule of Necessitation fails for free variables

One of the things that delighted me about indexicals was the convincingly deviant modal logic. As shown in *Demonstratives*, the rule of Necessitation:

If ϕ is valid, then $\Box\phi$ is also valid.

fails in the presence of indexicals.⁵⁸ The same rule also fails in the presence of free variables. If our assignments to free variables draw their values from the domain of quantification, then

$\exists y y=x$

is valid, but if the domain of quantification varies from possible world to world,

⁵⁵Not, at least on my interpretation. One who thought of proper names as *generic* (as standing for any individual so named) until set into a context of use would be thinking of them as parametric.

⁵⁶It should be clear that I am exploring the notion of a content-generating parameter, not insisting on one way of developing the semantics of free variables.

⁵⁷Perhaps the closest analogy is that developed above (in the subsection: "The paradigm of the variable") between the free variable, the 'free' pronoun, and the demonstrative, whose referent must be stipulated by a directing intention. Even in this case, however, there remains the puzzling problem of the seeming semantic role of the directing intention. In the case of an assignment, it is surely only the value that matters.

⁵⁸For example, take ϕ to be "I am here now" or "I exist."

$$\Box \exists y y=x$$

is not valid.⁵⁹

Harry Deutsch points out a related feature of the logic of free variables. On the present interpretation, although the basic *quantifier* logic for variables is classical, a free logic is simulated within the scope of the necessity operator. Thus, although

$$(\forall x Fx \rightarrow Fy)$$

is valid,

$$\Box(\forall x Fx \rightarrow Fy)$$

is not. An additional antecedent that is characteristic of free logic is required within the scope of \Box :

$$\Box((\exists x x=y \wedge \forall x Fx) \rightarrow Fy).$$

The failure of the rule of Necessitation in the presence of free variables results from the play between context (if the assignment parameter is taken as part of context) and point of evaluation. I view it as indicating that a parametric expression, likely to be directly referential, is at work.⁶⁰

The actual-world as an aspect of Context

The world of the context of use—what is taken for model-theoretic purposes to be the *actual-world*—plays a dual role in the logic. It is the parameter that the context provides for the indexical operator “it is actually the case that.” It is thus a *generation parameter* required to fix

⁵⁹Using a domain of quantification that varies from world to world deviates from the formulation in *Demonstratives*. As noted earlier, in *Demonstratives* I used a fixed domain, thought of as including all ‘possible’ individuals, along with a predicate “exists” whose extension could vary from world to world.

⁶⁰There is another, more sceptical, way to view failures of the rule—as an indicator of unclarity regarding the interpretation of free variables. This may be Kripke’s outlook in his pellucid discussion of the Barcan formula in “Semantical Considerations on Modal Logic,” *Acta Philosophica Fennica* (1963): 83–94. His analysis assumes the generality interpretation of free variables (on which the rule of Necessitation does in fact hold). He then shows that an apparent counterinstance to the rule is based on an incorrect formulation of the rule in this environment. As a corrective he proposes to formulate the system of derivation in a way that prohibits free variables in asserted formulas. He does not question the validity of the rule. I, being familiar with other counterinstances to the rule, have no difficulty with an interpretation of free variables that simply makes the rule invalid.

a determinate content for sentences containing the indexical operator. At the same time, and quite independently, it is also an *evaluation parameter* that plays a special role in the notion of validity. The latter is its more fundamental role, a role that would be required even if the language contained no indexicals for which the actual-world was needed as a generation parameter.⁶¹

Validity is truth-no-matter-what-the-circumstances-were-in-which-the-sentence-was-used. As I would put it, *validity* is universal truth in all *contexts* rather than universal truth in all possible worlds. Where indexicals are involved we cannot even speak of truth until the sentence has been set in a context. But it may appear that for a modal language *without* indexicals, without expressions that require a parameter, the notion of a context of use has no bearing. This is not correct. Truth in every model means truth in the ‘designated’ world of every model. This ‘designated’ world, the world at which truth is assessed, plays the role of actual-world. It is all that remains of context when the generation parameters are stripped away. But it does remain.

Perhaps this is more easily seen if we add the indexical operator “it is actually the case that” to the language. It is then apparent that the ‘actuality’ referenced by this operator is what we have become accustomed to refer to as the “designated” world.

The notion of the *actual-world* can be obtained in either of two ways. As I did, by starting from a full-blown language containing indexicals, deriving the notion of a context of use from its role in the semantics of indexicals, and then recognizing that truth, absolute truth in a model, is assessed at the world-of-the-context, i.e., the actual-world; or alternatively, by starting from a modal language *without* indexicals, recognizing that truth, absolute truth in a model, is assessed at the ‘designated’ world, and noticing that if we *were* to add the actuality operator this designated world would be the actual-world. Briefly, we can come upon the notion either in its guise as ‘world of the context of use’ or in its guise as ‘designated world’. On either approach, the notion of actual-world plays a special role in validity. It is the indispensable residue of the notion of *context*.

The terminology “context of use” evokes agents and utterances; the terminology “it is actually the case that” does not. There is, however, this common, underlying idea, one which I continue to think of as

⁶¹Within the formal system of *Demonstratives*, a content is evaluated at both a world and a time. Within that system, what is said here of the world of the context also holds of the time of the context.

perspectival—the actual world is where we actually are ... now. Recognizing that there are these two faces to the one notion makes me want to differentiate the possible worlds that can play the role of actual-world from those that are ‘merely’ possible, for example, by requiring that the former but not the latter not be empty; but not all will agree that there should be such differentiation. It is, in the end, a question of what you want to *do* with your logic.

Why the deviant logic?

The intuitive distinction between the actual-world, in which the content is generated, and all those possible-worlds in which the content can be evaluated,⁶² lies at the heart of such interesting logical phenomena as the failure of Necessitation. Any feature of a possible world which flows from the fact that it contains the context of use may yield validity without necessity. Such features need not depend on the contingent existence of individuals. For example, in the actual-world, the speaker, referred to by “I”, must be located at the place referred to by “here” at the time referred to by “now”. Hence “I am here now” is valid. But this requirement holds *only* in the actual-world, the world in which the content is expressed. Hence, what is expressed by the sentence need not be necessary. No ‘existence questions’ cloud this case.

I find it useful to think of validity and necessity as *never* applying to the same entity. Keeping in mind that an actual-world is simply the circumstance of a context of use, consider the distinction between:

- (V) No matter what the context were, ϕ would express a truth in the circumstances of that context

and:

- (N) The content that ϕ expresses in a given context would be true no matter what the circumstances were.

The former states a property of sentences (or perhaps characters): validity; the latter states a property of the content of a sentence (a proposition): necessity.

The nonstandard logic of *Demonstratives* follows from two features of the semantics of context and circumstance. The first is the possibility

that a given sentence might have a different content in different contexts. It is this that makes “I am here now” a valid *sentence*. And the second is the fact that not every possible circumstance of evaluation is associated with an (appropriate) possible context of use, in other words, not every possible-world is a possible actual-world. Though there may be circumstances in which no one exists, no possible context of use can occur in such circumstances. It is this that makes “Something exists” a valid sentence. Even if no indexical occurs in the language, the second feature puts bite into the notion of the actual-world.

These two features correspond to two kinds of a priori knowledge regarding the actual-world, knowledge that we lack for all other possible worlds. Corresponding to the first feature, there is our knowledge that certain *sentences* always express a truth regarding the world in which they are expressed. Corresponding to the second feature, there is our knowledge that certain *facts* always hold at a world containing a context. The latter is independent of the indexical resources of the language.⁶³

A word for cognitive value

The contexts of *Demonstratives* are metaphysical, not cognitive. They reach well beyond the cognitive range of the agent. Any difference in world history, no matter how remote, requires a difference in context.⁶⁴

In *Demonstratives* I tried to get at cognitive value through the notion of character.⁶⁵ When the twins, Castor and Pollux, each sincerely say, “My brother was born before I was,” they are said to be in the same cognitive state but to believe different things.⁶⁶ Though the utterances of the twins have the same cognitive value (same character), they do not bear the same truth-value (nor have the same content). I found it attractive to follow Frege in using a strictly semantical concept (character), needed for other semantical purposes, to try to capture his idea of cognitive value.⁶⁷

⁶³ The preceding material of this section resulted from a conversation with Harry Deutsch and Kit Fine.

⁶⁴ As noted, the entire world history is an aspect of context; it is the parameter for the indexical “Actually”.

⁶⁵ I have been told that “cognitive” is not the right word for what I have in mind. (I have also been told that what I have in mind is not the right idea for what I am trying to do.) I am not committed to the word; I take it from Frege (who probably never used it.)

⁶⁶ As indicated in *Demonstratives*, my views on this have been influenced by John Perry.

⁶⁷ Even granting that we cannot *articulate* the rules of character for all directly

⁶² Joseph Almog emphasizes this distinction in “Naming without Necessity,” *Journal of Philosophy* (1986): 210–42.

As in the case of content, the possible-worlds style of formal semantics in *Demonstratives* represents character as a function, in this case as a function from possible contexts of use. I continue to believe that proper names are not parametric, i.e., the *same name*⁶⁸ does not vary in referent from context to context.⁶⁹ Thus, the characters of two distinct proper names of the same individual would be represented by the same constant function, and thus, under the functional interpretation, coreferential names would not differ in character. Since it is indisputable that distinct proper names have distinct cognitive values,⁷⁰ the project of discriminating cognitive values of proper names by character is immediately defeated.⁷¹

Lately, I have been thinking that it may be a mistake to follow Frege in trying to account for differences in cognitive values strictly in terms of *semantic* values. Can distinctions in cognitive value be made in terms of the message without taking account of the medium? Or does the medium play a central role? On my view, the message—the *content*—of a proper name is just the referent. But the *medium* is the name itself.⁷²

referring expressions, we may still recognize a difference in cognitive value when presented with a pair of terms of different character; there may still be a correlation between distinct characters and distinct cognitive values. Joseph Almog suggests that we might express the point by saying that cognitive value *supervenes* on character.

⁶⁸ A less obvious notion than may appear.

⁶⁹ A proposed counterinstance: If a name can change its referent over time, as "Madagascar" is said to have done, then would not that very name have had one referent in an early context and another in a recent context? (For a partial response see the discussion below of logically proper names.)

⁷⁰ It is on the rock of distinct cognitive values for distinct names that Frege erected his gossamer theory. Note that Frege's initial argument makes use only of the uninterpreted forms " $a=a$ " and " $a=b$ ". The distinction between repetition of a single name and the use of two distinct names is already sufficient to make the points about *cognition* even before any examples (or even the notion of *Sinn*) are introduced.

⁷¹ One could, of course, argue that distinct names do differ in character and abandon the idea that character represents only the *parametric* determination of reference, i.e., how content varies from context to context. The fact that *indexicals* are parametric, that their character can be represented as a function from possible contexts, would then be regarded as a special case. The danger of trying to find characterological differences in distinct proper names is that the notion of *character* either will slip over from semantics to metasemantics or will become an ad hoc pastiche. In either case the dignified reality of character as the fundamental semantical value for indexicals would be seriously diluted.

⁷² In the case of indexicals, the character, which I took to represent cognitive value in *Demonstratives* may also be thought of as the medium by which content is generated (though character is semantic rather than syntactic in nature).

There are *linguistic* differences between "Hesperus" and "Phosphorus" even if there are no *semantic* differences. Note also that the syntactic properties of "Hesperus" and "Phosphorus", for example, their distinctness as *words*, are surer components of cognition than any purported semantic values, whether objectual or descriptive.

If words are properly individuated, by their world histories rather than by their sound or spelling, a name might almost serve as its own Fregean *Sinn*. The linguistic difference between "Hesperus" and "Phosphorus"—the simple difference between thinking of Venus qua *Hesperus* and thinking of it qua *Phosphorus*—may be all the difference in mode of presentation one needs in order to derive the benefits of sense and denotation theory. Words are undoubtedly denizens of cognition. If, through their history, they also provide the worldly link that determines the referent, then except for serving as content, they do all that Fregean *Sinn* is charged with. But they do it off-the-record, transparently and nondescriptively.⁷³

IV. Who Can Say What?

To complete my afterthoughts regarding the semantics of direct reference, I must address certain issues on the border between metasemantics and epistemology.⁷⁴ My reflections were driven by a puzzle about Russellian 'logically proper names'. In the end I concluded that the puzzle has a simple answer (to which I will return in the end).⁷⁵ But it prompted thoughts on the more controversial issue of constraints on what an agent in a particular epistemological situation can express.

What we can't do with words: the Autonomy of Apprehension

As I understand Frege and Russell, both believed that the realm of propositions accessible to thought, i.e., those capable of being *apprehended*, is independent of and epistemologically *prior* to the acquisition of language. In using language we merely encode what was already

⁷³ Here I echo an idea urged by Felicia (then Diana) Ackerman in "Proper Names, Propositional Attitudes, and Nondescriptive Connotations," *Philosophical Studies* 35 (1979): 55–69.

⁷⁴ I am indebted to Keith Donnellan for several formative discussions of this material.

⁷⁵ It has at least one simple answer; it also has several less simple answers.

thinkable.⁷⁶ Therefore, whatever can be expressed using language was already, prelinguistically, an available object of thought.⁷⁷

I see this view of the *autonomy of apprehension* in Russell's claim that

in every proposition that we can apprehend (i.e., not only in those whose truth and falsity we can judge of, but in all that we can think about), all the constituents are really entities with which we have immediate acquaintance.⁷⁸

Perhaps it accounts for the feeling one has in reading Russell on logically proper names, and even more so in reading Frege, that, like Humpty Dumpty, everyone runs their own language. When we speak, we *assign* meanings to our words; the words themselves do not *have* meanings. These assignments are, in theory, unconstrained (except by whatever limitations our epistemic situation places on what we can apprehend). In practice, it may be prudent to try to *coordinate* with the meanings others have assigned, but this is only a practical matter.⁷⁹

Subjectivist semantics

We may term this view, *subjectivist* semantics. Although the *entities* which serve as possible meanings may be regarded as objective, in the sense that the same possible meanings are accessible to more than one person,⁸⁰ the *assignment* of meanings is subjective, and thus the *semantics* is subjective. Since each individual user must *assign* meanings rather than receiving them with the words, each user's semantics is autonomous. What the language community does make available to each

of its members is a syntax, an *empty* syntax to which each user must add his own semantics.

The individual can express only those propositions that were already available to him as thoughts before receiving the benefits of linguistic communion. We cannot enlarge the stock of possible meanings that are available to us by drawing on the total stock of meanings extant in the language community. In this sense there is no semantic sharing. What each user can express is independent of the resources of other members of the language community, and in this sense what each user can express is *independent of language*.

There are differences between Frege and Russell in the way in which one's epistemic situation is seen to influence the propositions one can apprehend. Frege suggests that all mankind has access to the same thoughts. Thus that differences in our experience, our location in space and time, our culture (including in particular our linguistic community), do not affect what propositions we can apprehend.⁸¹

Russell's view was plainly different. He believed that our idiosyncratic experiences *do* affect what propositions we can apprehend. For Russell one can apprehend a proposition containing an individual *x* as a component if and only if one is directly acquainted with *x*. And it is clear that what one is directly acquainted with is a function of one's experience.⁸²

A fixed point of all such Russellian theories is that we may be so situated as to be able to *describe* a certain individual *x* but not to *apprehend* it; whereas a friend may be able to *apprehend* that selfsame individual. The friend can dub *x* with a logically proper name *n*, and try to communicate his thought using *n*. No use. We cannot just accept *n* with his meaning, we must assign it our own meaning, and in this case his meaning (namely, *x*) is not available to us for assignment. Sigh!⁸³

⁸¹In "The Thought: A Logical Inquiry," his discussion of the first-person pronoun indicates some ambivalence regarding this view. His suggestion that context of use is a partial determinant of the *Sinn* of an indexical may also indicate ambivalence if it implies (what I believe to be true) that persons in different contexts have access to different (indexical) thoughts.

⁸²Let different views of how direct *direct acquaintance* must be reflect different *Theories of Apprehension*. Russell suggests in the beginning of "On Denoting" that we may be acquainted with other people's *bodies* though we are not acquainted with other people's minds "seeing that these are not directly perceived." This suggestion does not accord with Russell's later views, and some think that this was not his *true* view even at the time of "On Denoting".

⁸³This is the situation in which we are forced to assign a descriptive meaning to the word our friend used as a name. Bad coordination, but unavoidable accord-

⁷⁶Here we may have the foundation for the view that meaning is all in the head, or at least all already directly accessible by the head.

⁷⁷Language, of course, aids communication, and also makes it easier, perhaps even possible, to *reason* using very complex thoughts. But the *manipulation* of thoughts is not what I am getting at here. My interest is in what can be apprehended and what can be expressed.

⁷⁸Bertrand Russell, "On Denoting," *Mind* 14 (1905): 479-93.

⁷⁹Prudential considerations of this kind will not, of course, affect a free spirit like H. Dumpty. An analogy: the concept of driving a car in traffic does not imply obedience to the conventions (sometimes called "rules" or "laws") whereby the movement of different drivers is coordinated. But it is usually (often?, occasionally?) prudent to so act. Dumpty's friend Dodgson appears to have shared his views. See Lewis Carroll (with an Introduction and Notes by Martin Gardner), *The Annotated Alice* (Cleveland: World Publishing Company, 1963), especially the notes on pages 268-69.

⁸⁰This was certainly the view of Frege and sometime the view of Russell.

Consumerist semantics

Contrast the view of subjectivist semantics with the view that we are, for the most part, language *consumers*. Words come to us prepackaged with a semantic value. If we are to use *those words*, the words we have received, the words of our linguistic community, then we must defer to *their* meaning. Otherwise we play the role of language *creators*.⁸⁴ In our culture, the role of language creators is largely reserved to parents, scientists, and headline writers for *Variety*; it is by no means the typical use of language as the subjectivist semanticists believe. To use language as language, to express something, requires an intentional act. But the intention that is required involves the typical consumer's attitude of compliance, not the producer's assertiveness.⁸⁵

There are two senses of "naming": dubbing and referring. To the consumerist, subjectivist semanticists have not adequately distinguished them.

To some, subjectivist semantics will seem a right and proper conservatism: Practice self-reliance—there is no such thing as a free thought! But it should be recognized that the view is incompatible with one of the most important contributions of contemporary theory of reference: the historical chain picture of the reference of names.

The notion of a historical chain of acquisition by which a name is passed from user to user, was first used to facilitate abandonment of the classical, description theory of proper names found in Frege and Russell.⁸⁶ The notion of a historical chain does this by offering an al-

ing to Russell. Frege's theory of apprehension seems to *permit* perfect coordination, which he urges for scientific discourse while recognizing that we don't always achieve it in ordinary discourse.

⁸⁴We may, like the prudent subjectivist semanticist, always attempt to give a known word the same meaning as that *commonly given* to it. We would still be playing the role of language creators, though without the creativity of someone like H. Dumpty.

⁸⁵I would like to formulate the relevant intention as one to use the word with *its* meaning, rather than with the meaning *assigned* by the person from whom the consumer heard (first heard?) the word. The immediate source from which the word was received seems to me to be primarily relevant to question of which word it is (among homonyms), rather than to the question of what meaning it has.

⁸⁶The idea, and its use in the argument against description theory, first appears in print in Keith Donnellan's "Proper Names and Identifying Descriptions," *Synthese* 21 (1970): 335–58; reprinted in *Semantics of Natural Language*, ed. D. Davidson and G. Harman (Humanities Press, 1972). It then appears in Kripke's *Naming and Necessity*, which, coincidentally, was first published in the same collection in which Donnellan's article is reprinted (*Semantics of Natural Language*). Kripke notes, "the historical acquisition picture of naming advocated here is apparently

ternative explanation of how a name in local use can be connected with a remote referent, an explanation that does not require that the mechanism of reference is already in the head of the local user in the form of a self-assigned description. In determining the referent of the name "Aristotle", we need not look to the biography's *text*, instead we look to its *bibliography*.

A role for language in thought: Vocabulary Power as an epistemological enhancement

There is another, possibly more fundamental, use of the notion: to tilt our perspective on the *epistemology* of language away from the subjectivist views of Frege and Russell and toward a more communitarian outlook.⁸⁷ The notion that a referent can be carried by a name from early past to present suggests that the language itself carries meanings, and thus that we can *acquire* meanings through the instrument of language. This frees us from the constraints of subjectivist semantics and provides the opportunity for an *instrumental* use of language to broaden the realm of what can be expressed and to broaden the horizons of thought itself.

On my view, our connection with a linguistic community in which names and other meaning-bearing elements are passed down to us enables us to entertain thoughts *through the language* that would not otherwise be accessible to us. Call this the *Instrumental Thesis*.⁸⁸

The Instrumental Thesis seems to me a quite important, though often tacit, feature of contemporary theories of reference, and one that distinguishes them from many earlier views. It urges us to see language,

very similar to views of Keith Donnellan" (addenda to *Naming and Necessity*, p. 164).

⁸⁷The two uses of the notion of a historical chain of communication are related. It is hard to see how to avoid some version of a description theory of proper names, at least for names of individuals we are not acquainted with, if one maintains a subjectivist semantics. Thus the attack on description theory (by which I mean not just the attack on *classical* description theory but the claim that descriptions are not even required as reference fixers) is *a fortiori* an attack on subjectivist semantics.

⁸⁸Given the wide acceptance of some version of the historical chain explanation for the mechanism of reference for proper names, it is surprising that there has been so little explicit discussion of the epistemological issues to which the Instrumental Thesis is addressed. A notable exception is the discussion of Leverrier's original use of "Neptune" in Keith Donnellan's "The Contingent *A Priori* and Rigid Designation," in *Contemporary Perspectives in the Philosophy of Language*, ed. P. French, H. Uehling, and H. Wettstein (Minneapolis: University of Minnesota Press, 1979).

and in particular semantics, as more autonomous, more independent of the thought of individual users, and to see our powers of apprehension as less autonomous and more dependent on our vocabulary.⁸⁹

Contrary to Russell, I think we succeed in thinking about things in the world not only through the mental residue of that which we ourselves experience, but also vicariously, through the symbolic resources that come to us through our language. It is the latter—*vocabulary power*—that gives us our apprehensive advantage over the nonlinguistic animals. My dog, being color-blind, cannot entertain the thought that I am wearing a red shirt. But my color-blind colleague can entertain even the thought that Aristotle wore a red shirt.

One need not fall in love to speak of love. One need not have grieved to speak of grief. The poet who has never felt or observed love may yet speak of it *if he has heard of it*. The fact that the language to speak of it and to enable us to have heard of it exists may show that *someone* once felt love. But it need not be the poet. And as with love, so with Samarkand (and red, and Aristotle). Our own individual experience may play a dominant role in providing the conceptual resources with which we address the world, but it does not play the whole role.⁹⁰

So how shall I apprehend thee? Let me count the ways. I may apprehend you by (more or less) direct perception. I may apprehend you by memory of (more or less) direct perception. And finally, I may apprehend you through a sign that has been created to signify you.

Does a name put us in causal contact with the referent?

I should add that I do not believe that the third category can be subsumed under the first. Apprehension through the language is not a very indirect form of perception that yields a very indirect form of acquaintance—like hearing a scratchy recording of Caruso or perhaps viewing his letters to his manager. Names are not, in general, among the causal effects of their referents. Perhaps a name should be regarded as among the causal effects of the person who *dubbed* the referent, but only in unusual cases will this *be* the referent.

⁸⁹How could Putnam have apprehended the dismaying thought that he couldn't tell a Beech from an Elm, without the help of his linguistic community? Could one have such a thought without having the words?

⁹⁰My grand instrumentalist views regarding red and love go beyond a more cautious version of the Instrumental Thesis that would be limited to names like "Aristotle" and "Samarkand". I note this at the urging of friends who characterize the cautious view as "persuasive" and my view as "shocking".

Even if we granted the referent a causal role in a typical dubbing by ostension, we can introduce a name by *describing* the referent (e.g., as the ratio of the circumference of a circle to its diameter). Such names are still directly referential and, in my view, still have the capacity to enlarge what we can express and apprehend. If we were to discover that Aristotle had been predicted and dubbed one year before his birth, or had been dubbed "Aristotle" only in medieval times, the name, like " π ", would still be a *name*, with all its attendant powers.⁹¹

I recognize that some will find my tolerance for nonostensive dubbings unacceptable, and may insist that the mere reception of a name *is* the reception of a causal signal from the referent. The name is likened to a lock of hair, a glimpse of one far distant, uninformative, but evocative. If names were like this, if there were a simple, natural (i.e., nonintentional) relation between name and named as there is between hair and behaired, the theory of reference for proper names would be a simple thing ... and it isn't.⁹²

On my view, acquisition of a name does not, in general, put us *en rapport* (in the language of "Quantifying In") with the referent. But this is not required for us to use the name in the standard way as a device of direct reference. Nor is it required for us to apprehend, to believe, to doubt, to assert, or to hold other *de dicto* attitudes toward the propositions we express using the name.⁹³

The *de dicto* hedge reflects my current view that *de dicto* attitudes, even those toward propositions expressed using directly referential terms, cannot easily be translated into *de re* attitudes.⁹⁴ The reason for this lies in part with the problems that led to my original claim that we need to be *en rapport* with those toward whom we hold *de re* attitudes and in part with technical problems involving reflexivity.⁹⁵

⁹¹Howard Wettstein points out that whereas dubbing by ostension has a special Russellian flavor, dubbing by description seems the paradigm for Frege. Since both adhere to subjectivist semantics, *they* believe that their dubbings are strictly for home use and will never go on the open market. (Did either have children?)

⁹²Those who see names as among the causal effects of the thing named seem to me to be insufficiently appreciative of Grice's distinction between nonnatural and natural meaning. H. P. Grice, "Meaning," *Philosophical Review* 66 (1957): 377–88.

⁹³It is required, however, that we *use* the name. I would suppose that with some very exotic names we might forbear their use in favor of their mention, and conceive of the referent only as *the referent of that name*.

⁹⁴This represents a change from the view expressed in footnote 69 of *Demonstratives*.

⁹⁵The first sort of problem involves understanding the conditions under which we correctly ascribe to Holmes, for example, the *de re* attitude that there is someone whom he believes to have committed the murder. It seems clear that the mere fact

The proponents of connectivity urge that although the language enables us to *express* contents that would otherwise be inaccessible (thus contradicting subjectivist semantics), something more, something like being *en rapport* with the components of the content, is required to *apprehend* the content (and thus to hold attitudes toward it).⁹⁶ I think of the proposal as a requirement that we have *knowledge of* the components. This certainly does not require direct acquaintance with the components, but it may require a natural connection to the components that is stronger than that provided by a name introduced into the language by one who did not himself have *knowledge of* the object (for example, a name now introduced for the first child to be born in the twenty-first century, or for the next president of Brazil, whosoever that may be).⁹⁷ The suggestion seems to be that all names (including perhaps names of colors, natural and unnatural kinds, etc.), however introduced, carry their referent as meaning; but not all names carry *knowledge of* their referent. Those names that were properly introduced, by ostension or based on some other form of knowledge of the referent, carry and transmit the requisite epistemic connection. But in a tiny fraction of cases the connection is absent—semantics (or metasemantics) does not require it—and in these cases we have direct reference, and expressibility, but no apprehension.⁹⁸

In theory, this is a dramatic weakening of the Instrumental Thesis, since it urges that more than a *semantic* connection needs to be established between a name and its referent before a name can attain its full powers. In practice, because only a tiny fraction of our vocabulary would lack the requisite connection, it may be almost no weakening at

that the murderer has given himself a *nom de crime* and leaves a message using this name should not suffice. (In fact, I suspect that there are no fixed conditions, only conditions relative to the topic, interests, aims, and presuppositions of a particular discourse.)

The second sort of problem is discussed in "Opacity", appendix B: The Syntactically *De Re*.

⁹⁶ A version of this view can be found in my "Quantifying In," *Synthese* 19 (1968): 175–214; reprinted in A. P. Martinich, op. cit. Others have espoused more sophisticated versions.

⁹⁷ The second example shows that what is required is that the *knowledge of* the individual play a special role in the dubbing. It must be intended to dub the individual *as known*. If someone I know well were to turn out, to my astonishment, to be the next president of Brazil, that would not qualify. Donnellan might say that in a dubbing by description, the description must be used referentially to dub an individual that one *has in mind*.

⁹⁸ A name may later take on the required epistemic connection when the referent appears upon the scene and is recognized as the named object.

all.⁹⁹

I am not entirely unsympathetic to this view.¹⁰⁰ We do distinguish knowledge from belief in part by the way in which we are connected to the object of knowledge. And thus insofar as one needs to *know what it is* that one apprehends, to *know what it is* that one believes, doubts, asserts, etc., the demand for epistemic connection may seem reasonable in analogy to that demanded for knowledge of facts (knowing-that). Note that on this view what gives us *knowledge of* the content of a name is *just* the connection, not any (new) beliefs. In fact, in this sense of knowing-what-we-apprehend, no beliefs at all are involved, only a well-connected name. In any case, a caveat must be added. To know what one apprehends is not to be able to *individuate* it. The Babylonians knew what Hesperus was, and knew what Phosphorus was, but didn't know that they were the same. Similarly, one might apprehend the proposition that Hesperus is a planet, and apprehend the proposition that Phosphorus is a planet, without knowing that they are the same proposition (if they are).

Naming the nonexistent

There are certain categories of objects which clearly have no causal effects upon us. If such objects can be given names, the view that names are among the causal effects of their referents cannot be correct. I have in mind future individuals and merely possible individuals. Such putative entities are *nonexistent*.¹⁰¹

If we can give a name to the person who *once* occupied this body ("John Doe #256"), why should we not be able to give a name to the person who *will*, in fact, arise from this fertilized egg? And if we possess an actual knock-down lectern kit, containing instructions for assembly

⁹⁹ My own hesitations regarding *de re* attitudes (the *de dicto* hedge) can also be seen as a limit on the scope of the Instrumental Thesis, a limit comparable to that proposed by those who suggest that an epistemic connection is required. If those who demand an epistemic connection identify *de dicto* attitudes toward propositions expressed using names (singular propositions) with *de re* attitudes (as I did in *Demonstratives*), it may even be that their qualms are really qualms about *de re* attitudes. But I had better not speak for others' qualms.

¹⁰⁰ Not *entirely*, though I do still maintain the view of footnote 76 of *Demonstratives*.

¹⁰¹ We certainly can't get *en rapport* with such individuals. Past individuals are also, in my view, nonexistent, but they do affect us causally. Some abstract objects, like numbers, do not, I think, affect us causally (in the appropriate sense), and they surely can be given names. I do not consider them because of qualms about the objectivity of such objects.

and all materials (form and matter), why should we not be able to name the unique, merely possible lectern that *would have* been assembled, if only we had not procrastinated until the need was past.

The sceptics, who take the position that an individual cannot be dubbed until it comes into existence, would insist that there is no naming the baby until the end of the first trimester (or whenever the current metaphysical pronouncements from the Supreme Court may indicate). One may, of course, express an intention to dub *whatever first satisfies certain conditions* with a particular name. Perhaps one may even *launch* the dubbing before the referent arrives. But the naming doesn't *take*, the name doesn't *name* it, one cannot *use* the name to refer to it (at least not to refer directly to it in the way names are said to refer by direct reference theorists) until the referent comes into existence.

A difficulty in the sceptical position is that in planning and in other forward-looking activities, we often wish to speak *about* such unnameables, perhaps through the use of descriptions.¹⁰² In my experience, those who protest the possibility of *naming* the first child to be born in the twenty-first century often accept the view that the description is—how shall I put it—not *vacuous*.

Perhaps they accept quantification over such entities and just object to the practice of introducing *names* on the basis of arbitrary descriptions (for names they want connectivity). It would then be natural to add a *narrow existence predicate* to distinguish the robust being of true local existents, like you and me, from the more attenuated being of the nonexistent.

If such quantification is *not* accepted, the position seems odd. Is it assumed that there are clever ways to reformulate any sentence in which such descriptions occur so as to 'eliminate' those that appear outside the scope of a temporal operator?¹⁰³ It is not obvious to me how to do this. How would the *de dicto* sentence, "Katie owes her first- (to be) born child to Rumpelstiltskin" be reformulated?¹⁰⁴

¹⁰²Or other 'denoting phrases' as Russell termed them.

¹⁰³I note that if there is such a method, then there is probably a similar method for eliminating descriptions of past individuals that no longer exist.

¹⁰⁴Using "*Fy*" for "*y* is a first-born child of Katie", and "*Ox*" for "Katie owes *x* to Rumpelstiltskin", we might try the following 'elimination' of the definite description from what is roughly "*O(the x)Fx*" (ignoring the 'if any' aspect of the description "her first-born"),

Future $\exists x(Always \forall y(Fy \leftrightarrow y = x) \wedge Now Ox)$.

This symbolization would be correct for "Katie *will give* her first-born child to

What sounds like scepticism with regard to naming the nonexistent, may merely be the quite different concern that the description of the intended dubbe is insufficiently specific to select a unique nonexistent individual. Such may be the case of the possible fat man in the doorway.

Insufficient specificity seems to be Kripke's qualm in *Naming and Necessity* regarding the merely possible species Unicorn and a merely possible referent for "Sherlock Holmes".¹⁰⁵ However, his discussion of what he calls "the epistemological thesis" (that the discovery that there were animals with all the features attributed to Unicorns in the myth does not establish that there were Unicorns) suggests an entirely different argument, namely that *the way in which these particular names arose* (from pure myth and pure fiction) makes it impossible for *them* to name merely possible entities.¹⁰⁶ This argument is independent of the degree of specificity in the myth or in the fiction.¹⁰⁷

Rumpelstiltskin", but not for "*owes*". The problem is that "*owes*" (like "*needs*" and "*seeks*") is an intensional verb with respect to its grammatical object. Even if it turns out that Katie's first-born child is her ugliest child, *Always* $\forall y(Fy \leftrightarrow Uy)$, she does not now owe Rumpelstiltskin her ugliest child. (However, if she *will give* her first-born child, then she will give her ugliest child.) The 'elimination' of the definite description transforms the predication from *de dicto* to a quantification in. And this leads to incorrect results for intensional verbs. (Note that the same sort of 'elimination' occurs automatically whenever we use first-order logic to symbolize a sentence with an indefinite description as grammatical object. Compare the symbolizations of "Katie owes a bushel of gold to Rumpelstiltskin" and "Katie will give a bushel of gold to Rumpelstiltskin". The interesting problem about indefinite descriptions as grammatical objects of intensional verbs is how to 'uneliminate' them.)

So long as there are no intensional verbs, the eliminations are not plainly incorrect. Intensional operators, so long as they are *sentential* operators, do not create a problem, because definite and indefinite descriptions can be eliminated from predicates while remaining within the scope of the operator.

Some think that "*owes*" can be paraphrased to produce a sentential complement where the grammatical object of "*owes*" appears, for example, as "Katie is now obligated that at some future time she gives her first- (to have been) born child to Rumpelstiltskin". This allows a tense operator ("at some future time") to be inserted between the new sentential operator and the old grammatical object of "*owes*". If you are of this view, try "Katie is thinking about her first (to be) born child", and read appendix A: Paraphrasing Into Propositional Attitudes from "Opacity".

My aim here is to indicate that there is a substantial technical problem faced by those who hope to achieve the effect of quantification over future individuals through the use of temporal operators.

¹⁰⁵Addenda, pp. 156–58.

¹⁰⁶As Harry Deutsch puts it, *reference is no coincidence*.

¹⁰⁷In lecture, Kripke has made the intriguing suggestion that there are abstract but actual (not merely possible) *fictional individuals* that serve as the referents of

Neither insufficient specificity nor the objections concerning extant names from fiction or myth apply to the case of the first child to be born in the twenty-first century or to the case of the possible lectern, in both of which a frank attempt is made to dub what is recognized as a nonexistent object.

Logically proper names

The question that prompted all my thoughts on subjectivist semantics, the Instrumental Thesis, and vocabulary power is this: How should Russellian 'logically proper names' be accommodated in the semantics of Context and Circumstance?

Using "name" for what he sometimes called a "logically proper name," Russell writes,

a name ... is a simple symbol, directly designating an individual which is its meaning, and having this meaning in

names like "Sherlock Holmes". The admission of such entities might be accompanied by a narrow existence predicate to distinguish the fictional from the non. I am not aware of Russell's views on future individuals, but he expressed himself in opposition to fictional entities in *Introduction to Mathematical Philosophy*,

If no one thought about Hamlet, there would be nothing left of him; if no one thought about Napoleon, he would have soon seen to it that someone did. The sense of reality is vital in logic, and whoever juggles with it by pretending that Hamlet has another kind of reality is doing a disservice to thought.

Despite Russell's rhetorical power, I must confess to having been persuaded by Kripke's analysis. (As Joseph Almog points out, it is not clear that Russell's insistence that Hamlet does not have "another kind of reality" would apply to what I take to be Kripke's view that Hamlet, though not a *person*, exists as a *fictional character in our reality*.)

If Kripke is correct, it would seem to settle the case in which an author creates a fiction 'out of whole cloth' but specifies one of the characters, which he names "Woody", to have particular characteristics which, though nothing does in fact *have* the characteristics, our favorite theory of essentialism tells us that there is exactly one possible object that *could* have them (e.g., the characteristic of having been assembled from a certain lectern kit). "Woody" would name an actual fictional entity, not a merely possible nonfictional entity.

Or should we say instead that the author made up a story *about* a particular merely possible nonfictional entity? The fairly plain distinction between an individual, *x*, having the *properties* of a character in a story and the story being *about x*, grows dim when *x* is merely possible. And if we add the difficulties of the distinction between being *about x* and being *modeled on x* (a hard enough distinction for real *x*), I lose discriminability.

its own right, independently of the meanings of all other words.¹⁰⁸

It is hard to resist the idea that for Russell, such names are directly referential. However, his ideas about the existence predicate are baffling. He continues,

The proposition "the so-and-so exists" is significant, whether true or false: but if *a* is the so-and-so (where "*a*" is a name), the words "*a* exists" are meaningless. It is only of descriptions—definite or indefinite—that existence can be significantly asserted; for, if "*a*" is a name, it *must* name something: what does not name anything is not a name, and therefore, if intended to be a name, is a symbol devoid of meaning.

His claim that it is meaningless to predicate existence of a logically proper name is plainly a mistake.¹⁰⁹ Far from being meaningless, such propositions are required as the objects of what Russell called *propositional attitudes*, "I regret that this pain exists", "I am pleased that Nixon exists" (taking "Nixon" and "this pain" to be logically proper names). These assertions are by no means either trivial or meaningless.

The requirement that a logically proper name *name* something seems to have the result that "*a* exists" ("*a*" a logically proper name) cannot be used to express a proposition that is false. But unless "*a*" names a necessary existent, the proposition expressed would not be necessary. Thus we have a seeming failure of the rule of Necessitation. This, along

¹⁰⁸ From chapter 15 of *Introduction to Mathematical Philosophy* (London: Allen & Unwin, 1919), reprinted in Martinich, op. cit.

¹⁰⁹ I do not understand why Russell did not recognize that the intolerable existence predicate could be defined by forming the indefinite description, "an individual identical with *a*", and then predicating existence of the indefinite description in the way Russell finds so commendable, " $\exists x x = a$ ".

The problem with empty names should not have dissuaded him. If such names are taken to be (disguised) definite descriptions, as he usually claimed they were, then (where *a* is now a definite description), " $\exists x x = a$ " is again equivalent to "*a* exists" according to Russell's own theory of descriptions. (As a sidelight, it is interesting to note that even if an empty name is taken to be "a symbol devoid of meaning," it is possible to develop a rigorous semantics according to which " $\exists x x = a$ " is again equivalent to "*a* exists". Russell was not aware of this.)

It is not my claim that the notion of *existence* is captured by the existential quantifier; variables can have any domain. My argument is *ex concessis*. Insofar as existence can be "significantly asserted" of indefinite descriptions, it can be significantly asserted of names.

with Russell's epistemological ideas, which emphasize the special situation of the agent who uses the name, is highly reminiscent of my analysis of indexicals.

These reflections made logically proper names seem a natural topic for the apparatus I had developed in *Demonstratives*, and this drew me in deeper.

When I attempted to apply the apparatus, I was surprised by the results. I was faced with a puzzle. The principles governing logically proper names seemed to imply that a logically proper name *must* name something that exists in its context of use, but need not name a *necessary* existent. But if the referent is not a necessary existent, then there must be a world and time at which it does not exist, and if *c* is a context of use in such a world at such a time, what would be named by an "occurrence" of the name in the context *c*? Briefly, how can every possible occurrence of a name have an existent referent, if the referent isn't a necessary existent?

To make things definite, consider the puzzling case of Nixon. Suppose that I name a certain pain with which I am directly acquainted, "Nixon". We agree that Nixon does not have necessary existence. So there must be a happier world (or time) in which Nixon does not exist.¹¹⁰ If I were to utter "Nixon" in this happier circumstance, what existent would I be referring to? If "Nixon exists" cannot be used to express a proposition that is false, an occurrence of "Nixon" in such circumstances must name something that exists there. This cannot be Nixon, *ex hypothesi*. What could it be?

Be clear that I am not raising questions about how to *evaluate* at the happier circumstance what is expressed by an occurrence of "Nixon exists" in the painful context of dubbing. No problem there; it's false (again, *ex hypothesi*). The question is: What is *expressed* by an occurrence of "Nixon exists" in a context in the happier circumstance? And how can it be true there?

So what is the referent of "Nixon" when it occurs in a context in a world and time in which Nixon doesn't exist?

We can be certain that names do not enter vocabularies through a trans-world chain of communication. If the world is one in which Nixon never exists, how is the agent of the context able to use the term

¹¹⁰For example, the next day, when Nixon has subsided into nonexistence. Or, if you think that pains like Nixon never cease to 'exist' (in some sense) once they appear, take the day *before* Nixon came into existence. Or, better yet, take some possible world in which Nixon *never* comes into existence.

"Nixon"; was the name introduced there to dub a merely possible entity? Not likely.

The solution to the puzzle is, I think, independent of all the issues surrounding subjectivist versus consumerist semantics. As was emphasized earlier, our notion of an *occurrence* of an expression in a context does not require an *utterance* of the expression nor even that the agent of the context have the use of the expression. The apparatus of Context, Character, and Circumstance is designed to help articulate the semantics of an *interpreted* language, one for which meanings, *however derived*, are already associated with the expressions. It takes account of what the meanings are, not of how they came about. Given an interpreted language, a sentence is valid if it expresses a truth in every context, including those contexts in which the language doesn't or couldn't exist, or doesn't or couldn't have that interpretation. Thus the objection that certain meanings *could not arise* or *could not be used* in certain contexts is, strictly speaking, irrelevant to our issue: What is the content in such contexts of an expression which already carries a certain meaning?

So the answer is: Nixon. (Just as you knew all along.) The intuition that "Nixon exists" must be logically valid whenever "Nixon" is a logically proper name, is in error in tacitly assuming that to *evaluate* our language in a foreign context, the language, with its interpretation, must exist there.¹¹¹

I see here a reaffirmation of the importance of a central distinction that I have tried to build into my very nomenclature, the distinction between what *exists* at a given point and what can be 'carried in' to be *evaluated* at that point, though it may *exist* only elsewhere. My 'Circumstances of Evaluation' evaluate contents that may have no native existence at the circumstance but can be expressed elsewhere and carried in for evaluation. What is crucial to the puzzle about "Nixon" is that my 'Contexts of Use' are also points of evaluation, they evaluate characters (meanings) that may have no native existence at the context but can also be created elsewhere and carried in for evaluation.

Where within the formal theory do I take account of the locus of *creation of character*, the *assignment* of meanings that is presupposed

¹¹¹This, however, suggests that there may be another interesting analysis of the puzzle about logically proper names in terms of utterance-validity. And another using the notion, from the discussion of contexts for demonstratives, of a context *appropriate* for a particular expression. These considerations may throw light on a kind of metasemantical analyticity, not the usual: truth solely in virtue of what the meaning is, but instead: truth in virtue of having come to have that meaning.

in the notion of an interpreted language? Where within the formal theory do I take account of such metasemantical matters as constraints on the kinds of dubbings allowed? I do not.¹¹²

¹¹²In addition to assistance specifically acknowledged, I have been much helped (provided one includes expressions of dismay as *help*) by Joseph Almog, Harry Deutsch, Keith Donnellan, Kit Fine, John Perry, Elisabetta Fava, Nathan Salmon, and Howard Wettstein.