

Finding Sense in Theories of Truth
Draft II 1/2

Mark Textor

1 Introduction

“[...] we are already committed to tasks of developing life-like, contextually competent theories of truth for natural languages and of refining in application Frege’s intuitive criterion of sameness and difference of sense. We do not need (and we must absolutely dispense with) the traditional proposition in developing such interpretive theories or working out the framework in which to develop them. But this is not to say that, even at the end, such theories cannot deliver such a thing.”

D. Wiggins (1992), p. 90.

Donald Davidson is well-known for his claims (i) that a Tarski-style theory of truth could be developed for a natural language L and (ii) that such a theory of truth could serve as a theory of meaning for L . The second claim is more clearly expressed in the following way: a theory of truth for a natural language L will entail for each sentence of L a so-called T-sentence, a theorem of the form “ s is true $_L$ if and only if p ”.¹ If a theory of truth fulfills certain empirical and formal constraints its T-sentences will specify the meanings of the object-language sentences: we can substitute in these T-sentences “is true $_L$ if and only if” with “means-in- L that”. Davidson thinks that someone who masters such a privileged theory of truth is able to understand utterances of L -sentences.

Although I will discuss the basic thought of Davidson’s theory I am not primarily interested in Davidson’s view. My interest lies in certain applications of Davidson’s theoretical apparatus. John McDowell has argued that Davidson’s theory of meaning allows us to clarify Frege’s notion of sense which seems to many authors rather mysterious.² McDowell proposes to view a Davidsonian theory of truth as a theory which specifies the Fregean sense for each sentence of a natural language.

McDowell’s proposal is not concerned with indexical sentences. But indexical sentences are of special interest for philosophers who are sympathetic to a Fregean approach to natural language.

The first reason for this is that indexical expressions provide the basic motivation for a distinction between *linguistic meaning* and *Fregean sense*. Frege calls the sense of a complete assertoric sentence “thought”. If we take the phenomenon of indexicality seriously, it is more useful to ascribe Fregean sense and reference to

¹The predicate “is true $_L$ ” is a one-place predicate, the subscript “ L ” is a singular term denoting the particular language for which the predicate is defined, cp. D. Davidson (1990), p. 289.

²cp. J. McDowell (1976) and (1980)

utterances of sentences.³ One of Frege's criteria for thought-difference must accordingly be re-formulated thus: If two utterances can diverge in truth, they cannot express the same thought. So my utterance of the indexical sentence "I am hot" and your utterance of the same sentence express *different* thoughts. Yet the utterances have the same linguistic meaning. If the Fregean sense of an expression and its linguistic meaning would coincide, McDowell's whole idea of elucidating the former notion by appealing to a Davidsonian theory of truth would lose much of its interest: it would simply be a *relabelling* of Davidson's theory of linguistic meaning as a theory of sense. But if linguistic meaning and Fregean sense do not coincide, it is a genuine insight that the Fregean sense of an expression (if there is such a thing) could be specified by a Davidsonian theory.

The second reason why indexical expressions are of crucial importance for Fregean theories of language is that it has recently been argued that these expressions resist an integration into the Fregean semantic framework. Therefore Perry and other critics abandon the Fregean framework in approaches to natural language.⁴

Gareth Evans' seminal paper "Understanding Demonstratives" tries to answer the just mentioned criticism by using McDowell's conception of sense. Evans makes a more demanding claim than McDowell: according to Evans a privileged theory of truth for an *indexical* language L can serve as a theory of the *Fregean thoughts* expressed by (indexical) utterances of L sentences. Evans argues that such a view would render Frege's semantic immune against the above mentioned criticism. My paper will investigate whether Evans' idea that a theory of truth can serve as a theory of sense (utterance content or Fregean thought) for an indexical language is true.

The main expositional problem of this paper is that Evans uses in his writings Davidson's much discussed and disputed theory of meaning as a starting-point. As I cannot argue here for or against this theory, I will simply investigate whether it can be of any help to a philosopher who is sympathetic to Frege's views. My conclusion will be mainly negative: a Davidsonian theory of truth cannot serve as a theory of understanding for indexical languages.

The plan for the paper runs as follows: Section 2 to 4 are set the stage for the discussion of Evans' view. Section 2 tries to get a grip on the slippery notion of understanding. I will introduce basic distinctions which must be observed in the critical assessment of the arguments and theories that will be discussed in the subsequent sections. Section 3 introduces the main principles of Davidson's theory of meaning. Section 4 relates Davidson's ideas on meaning to Frege's concept of sense. In section 5 I discuss Davidson's treatment of indexical expressions. This section anticipates problems which will later arise for Evans' application of Davidson's theory. Section 6 and 7 will assess the explanatory power of Evans' theory of sense. In section 9 I will investigate whether the notion of a *conditional* truth-condition can be of any help to Evans' project.

³cp. M. Sainsbury (1996), p. 92.

⁴Cp. J. Perry (1977).

2 Aspects of Understanding

The concept of understanding and the correlative concept of meaning are slippery notions. Therefore some preliminary clarifications are necessary. The first one concerns the pragmatic aspects of understanding. As most writers I will leave out of consideration the pragmatic dimension of understanding. Of course someone who reads Bishop Butler’s famous dictum “Everything is what it is and not another thing” and wonders why someone should honestly assert such a trivial thing has in one sense not understood the inscription. He has not realized what Bishop Butler *conveyed* or *suggested* by this inscription. Yet in another sense he has understood it completely. I will focus on this less demanding concept of understanding according to which one need not grasp the implicatures of an utterance (if it has any) to understand it

Can we say now that one understands an utterance iff one knows its linguistic meaning? Consider a this example: a competent speaker of English, *S*, listens to a tape-recording of the sentences “I am trapped here since yesterday. Please help me!” *S* does neither know when the tape was recorded nor does he recognize the voice of the speaker. Does *S* understand the utterance in these circumstances? Yes and No. He does understand it in the sense that he knows its linguistic meaning: He recognizes that it is an utterance of an English sentence and he knows the semantics of the expressions that compose the uttered sentence. For instance *S* knows the following rule of reference for “I”:

In every assertoric utterance in which the English word “I” is used it refers to the utterer.⁵

But *S* does not understand the relevant utterance in another sense: *S* does not grasp the proposition expressed by the particular utterance. *S* does not know who the speaker of the utterance is, where this person is located and at which day he uttered the sentence in question. This prevents him from grasping, as one might say, what was said by the particular utterance on the tape.

I think this distinction between knowledge of the linguistic meaning and knowledge of what is said by an utterance is central to our understanding of “understanding”. But it is presumably not the only distinction one needs. An important question is how to characterize what *S* must know in order to grasp what was said by an indexical utterance, its *official content*.⁶ This question has a modest answer, given by John Perry, and a demanding answer, given by Gareth Evans.

The modest answer assumes that it is sufficient for knowing what was said by an utterance containing indicators that one has identified the references of the indicators in the utterance-context and knows the linguistic meaning of the non-indexical components of the uttered sentence. Let me clarify this with an example. *S* utters

⁵Of course this formulation needs polishing. We have to take care of mimetic and elliptic “I”-utterances like “I am parked out back” etc. But these considerations are not relevant here.

⁶This is John Perry’s term, compare J. Perry (forthcoming)

“He is tall” and “he” refers in his utterance to John. According to the modest conception of what was said I would count as someone who understands the utterance (in the relevant sense) if I grasped consequently a thought about John to the effect that he is ill. I could re-express what I understood in various ways. Here are some examples:

“You are tall” (said to John).

“He is tall” (said about John).

“That man is tall’ (said while pointing to John).

All these utterances are legitimate expressions of what was said by *S*’s original utterance. What they have in common is that they predicate the same property of the same individual. Therefore it seems natural on this view to identify the object of understanding in this strong sense with an ordered pair consisting of an object-tuple and an *n*-place property.

The demanding answer is motivated by examples like the following one from Brian Loar who pictures a conversation between Smith and Jones on the train:

“Suppose that Smith and Jones are unaware that the man being interviewed on television is someone they see on the train every morning and about whom, in the latter role, they have just been talking. Smith says ‘He is a stockbroker’, intending to refer to the man on television; Jones takes Smith to be referring to the man on the train. Now Jones, as it happens, has correctly identified Smith’s referent, since the man on television is the man on the train, but he has failed to understand Smith’s utterance.”⁷

Examples like this one seem to show that merely knowing the reference of the indexicals of an indexical utterance is not enough for understanding in the full sense. For knowing what was said by an indexical utterance, the proponent of the demanding answer argues, it is not enough to identify the referents of the indicators and to know the linguistic meaning of the non-indexical elements of the uttered sentence, one must identify the referents *in a specific way*. This idea is central to Evans’ account of meaning and understanding. If one accepts the demanding answer it seems natural to identify what is said, the propositional content of an utterance, not just with an ordered pair of an object-tuple and an *n*-place property, but with a different object. The relevant object should be so chosen that different objects correspond to different ways of identifying the reference of the indexicals.

But once this step is taken a problem immediately arises: the speaker’s and the hearer’s way of identifying a referent need not or in some cases cannot even be identical: surely my way of identifying the referent of my “I” utterances is different from the way any other person identifies this object, if this person understands my “I”-utterances.⁸ Which one of these ways of identifying the referent is relevant for the proposition expressed?

⁷B. Loar (1976), p. 357.

⁸If one can speak in the case of “I” at all of identifying the referent.

In this paper I will not try to show that one of these answer to our question about understanding is correct. I will simply assume Evans' demanding answer and investigate whether Evans is right in holding that a Davidsonian theory of truth (or a variant of such a theory) can serve as a theory of Fregean thought or utterance content.

3 Davidson on Truth and Meaning

An account of the semantic competence of a speaker S of a language L would be a detailed specification of the knowledge sufficient for understanding actual and potential utterances of sentences of L . Davidson thinks such an account will rely on a theory of truth for L . Davidson's work is inspired by Tarski's attempt to give a recursive definition of the notion of truth by appealing to the concept of meaning 'in the guise of sameness of meaning, or translation'. Davidson reverses the order of explanation: Davidson takes truth to be "the central primitive concept" and he argues that a recursive definition of a truth-predicate for a language will illuminate the meaning of complex expressions of L .⁹ A rough formulation of Davidson's main idea is: knowledge of a theory of truth for a language L suffices for understanding utterances of L sentences.¹⁰ When we ascribe to a speaker such knowledge, we credit him with a systematic method for interpreting utterances of a language. This systematic method or theory can also be described as a theory of meaning for the language under consideration. Its T-sentences "may be treated as 'giving the meaning' of object language sentences."¹¹

A materially adequate theory of truth will entail for every sentence of L a theorem of the form "s is true $_L$ if and only if p". We get from this schema a T-sentence if we substitute for "s" a canonical designation for sentences of the object-language L and for "p" a sentence of the meta-language.

If a theory of truth is a theory of meaning, as Davidson claims, the right-hand side of a T-sentence should specify the meaning of the sentence of the object-language the T-sentence is about. Now due to the logic of the bi-conditional and the rules of disquotation, a T-sentence is true if and only if the denoted sentence of the object-language and the used sentence of the meta-language have the same truth-value. This reveals the central problem of Davidson's approach to characterize meaning in terms of truth conditions. There is a gap between a theory of truth and a theory of meaning: sameness of truth-value does not guarantee sameness of meaning. A well known example illustrates that.

Imagine a theory of truth whose object-language is German and whose meta-language is English. A correct T-sentence of such a theory could be:

(T1) "Der Mond ist rund" is true $_G$ if and only if hesperus is a planet.

⁹D. Davidson (1984), p. XIV.

¹⁰In the latest formulation of this principle the languages under consideration are idiolects. Cp. D. Davidson (1990), p. 312.

¹¹cp. D. Davidson (1984), p. 150.

(T1) is true, but doesn't specify the meaning of "Der Mond ist rund". A mere theory of truth can't be a theory of meaning. I will call the problem illustrated by this examples following Fodor and Lepore the *extensionality problem*.¹² Which additional properties must a materially adequate theory of truth have in order to qualify as a theory of meaning?

Davidson's answer is (at least) twofold: the privileged theory of truth must satisfy (1) formal and (2) empirical constraints.¹³

(1) *Formal constraints on the theory of truth.* Meaning giving T-sentences must be theorems of a "reasonable simple theory"¹⁴ that assigns materially adequate truth conditions to all the sentences of the language under consideration. In order to be reasonable simple a theory of truth cannot be just a list of T-sentences. A theory that fulfills Davidson's constraint must discern structure in sentences of L . From the perspective of a theory of truth the sentences of L are composed from a finite number of atomic expressions. The theory of truth Davidson envisages will contain for each atomic expression an axiom which states its contribution to the truth-condition of sentences in L . Axioms for singular terms will look thus:

(A1) "Der Mond" refers $_G$ to the moon.¹⁵

The axioms for predicates are modelled on:

(A2) "ist rund" is satisfied $_G$ by x if and only if x is round

The axioms for atomic sentences are of the following form:

(A3) "Fa" is true $_G$ if and only if what "a" refers $_G$ to satisfies $_G$ "F"

Additionally there will be a finite number of axioms for the logical vocabulary of the language.

The theory of truth derives the truth-conditions of a potential infinity of sentences from such a finite list axioms. A reasonable simple theory of this type could not entail a T-sentence like (T1) without also entailing further incorrect T-sentences. For instance this theory would fail to give correct T-sentences for the German sentence "Dies ist rund".¹⁶

In the beginning of this section I ascribed to Davidson the view that the concept of truth is primitive. But does a theory of truth which contains axioms like (A1) and (A2) not suggest the idea that truth can be explained in terms of reference and satisfaction? Davidson answers this query in the following way:

"[...] these notions [reference and satisfaction] we must treat as theoretical constructs whose function is exhausted in stating the truth conditions for

¹²J. Fodor, E. Lepore (1992), p. 62.

¹³cp. D. Davidson (1984), p. 150.

¹⁴cp. D. Davidson (1984), p. 26.

¹⁵The subscript "G" indicates that the reference-relation is defined for German. I will use this notation also for the other clauses.

¹⁶cp. D. Davidson (1984), p. 26 fn1.

sentences.”¹⁷

According to Davidson the notions of reference and satisfaction are conceptually dependent on the notion of truth: they are just invoked to give a recursive definition of truth.¹⁸

(2) *Empirical constraints on the theory of truth.* But even a theory of truth which fulfills Davidson’s formal constraint could fail to generate only meaning giving T-sentences. A reasonable simple theory of truth for German could be non-interpretative because the axioms for the non-logical vocabulary could contain ‘redundant parts’ that would prevent the entailed T-sentences from giving the meaning of the object-language sentences. Here is an example:

“ist rund” is satisfied_G by x if and only if x is round and $2 + 2 = 4$

From a theory of truth which contained this axiom and the T-sentence (T2) could be derived by in a systematic and simple way:

(T2) “Der Mond ist rund” is true_G if and only if the moon is round and $2 + 2 = 4$

(T2) is materially adequate but does not provide the meaning of “The moon is round”.¹⁹

To solve this problem Davidson appeals to empirical constraints.²⁰ Davidson holds that a theory of truth for a language L is an *empirical* theory that interprets the utterances of speakers of a language.²¹ An empirical theory of meaning of the type Davidson envisages will shed light on what a sentence of a language L means only if we can tell whether a T-sentence for a sentence S of L is true, without already knowing what S means in L . Our grasp and application of the notion of truth to utterances must be prior to our understanding of these utterances. Davidson uses frequently the model of radical interpretation to illustrate the implications and difficulties of this approach. How can someone who interprets a foreign language for the first time come to know that the sentence:

S is true_L.

is true without already knowing that S means in L that p? Yet Davidson argues that we can have evidence that a sentence S is true without already knowing the meaning of S .

Davidson’s argument for this relies on the the notion of holding true a sentence. Davidson thinks that we cannot *in general* identify and ascribe beliefs to someone if we have not already interpreted his speech. But Davidson also holds that there is an

¹⁷D. Davidson (1984), p. 223.

¹⁸This thesis is criticised in J. Perry (1995).

¹⁹cp. W.O.V. Quine (1977).

²⁰cp. D. Davidson (1984), p. 26 fn11, p. 135ff. p. 152f.

²¹In the extreme case the language will be the language of one speaker at a certain time.

exception to this rule: the attitude of holding a sentence true (accepting a sentence as true) is according to Davidson a belief we can ascribe to someone without already mastering his language. He thinks this is possible because holding true

“is a single attitude applicable to all sentences, and so does not ask us to be able to make finely discriminated distinctions amongst beliefs. It is an attitude an interpreter may plausibly be taken to be able to identify before he can interpret, since he may know that a person intends to express a truth in uttering a sentence without having any idea *what* truth.”²²

But is it really unproblematic to say that an interpreter can identify that a speaker holds true a sentence S without understanding S or any other sentence the speaker might utter? If the interpreter does not yet understand any of the speaker’s utterances (does not know which proposition they express) he can hardly discriminate between holding a sentence true and other attitudes directed towards a sentence like believing the sentence to be probable (to express a probable truth), believing it to be desirable (represent a desirable state of affairs) and so on.²³

In verifying the theory of truth T-sentences for a subclass of indexical sentences are of central importance.²⁴ The attitude of holding true an indexical sentence like “Es schneit” depends on the circumstances of their utterance. The T-sentences for these sentences are confirmed by correlating attitudes of holding true with the conditions that give rise to them. For instance (T3):

(T3) “Es schneit” is true_G when spoken by x at t if and only if it snows near x at t .

is confirmed by the empirical generalization

(EG) $(\forall x)(\forall t)$ (if x belongs to the German speech community then (x holds true “Es schneit” at t if and only if it is snowing near x at t))

which in turn is based on evidential statements like:

(ES) Kurt belongs to the German speech community and Kurt holds true “Es schneit” on Monday at noon and it snowing near Kurt on Monday at noon.

But how can (EG), a statement about the acceptance behaviour of German speakers, confirm a theory which ascribes truth conditions to (potential) utterances? On the one hand the observed speakers of course make mistakes: sometimes a speaker holds true “Es schneit” at t even if it is not snowing near him at that time. On the other hand the speakers are often unaware of facts about their vicinity: although it snows near me I fail to hold true “Es schneit” because I am totally absorbed watching a football game. So how can Davidson’s empirical constraints lead to meaning-giving theory of truth for the language under consideration?

²²cp. D. Davidson (1984), p. 135.

²³cp. J. Bennett (1985), p. 612.

²⁴The truth-theoretic treatment of indexical sentences is the topic of the next section.

Davidson answers these objections by an appeal to what he calls the “Principle of Charity”. Someone who develops a theory of truth must ascribe truth conditions to sentences of L in a way that “maximizes agreement, in the sense of making Kurt (and others) right, as far as we can tell, as often as possible.”²⁵ If we use the Principle of Charity we can establish on the basis of the conditions under which speakers hold a sentence S true the truth conditions of S : if Kurt holds true “Es schneit” on Monday at noon and it snowing near Kurt on Monday at noon a theory *that maximizes agreement* will ascribe to the sentence “Es schneit” the truth condition specified in (T3).

Davidson holds that the empirical statements that confirm the theory of truth depend on causal relations between speakers of L and the world.²⁶ Kurt holds true the sentence “Es schneit” on Monday *because* it is snowing near him on Monday. If we develop a theory of truth which maximizes agreement we are looking for causal factors which cause speakers to accept sentences. This idea excludes T-sentences like (T2). German speakers don’t accept the sentence “Schnee ist weiß” *because* snow is white and $2 + 2 = 4$. The cause of their acceptance behaviour is that snow is white. Davidson’s methodological rule to ascribe truth-conditions to sentences in a way that reflects the acceptance behaviour of speakers excludes therefore T-sentences like (T2). They cannot be theorems of an empirically adequate theory of truth.

Davidson offers no philosophical argument that these constraints will exclude all non-interpretative truth theories for a language L . But he hopes that the imposed constraints will be adequate.²⁷ But even if they are there is a further problem. John Foster has pointed out that someone may know a theory of truth for the language L that met all empirical and formal constraints but yet couldn’t understand utterances of L sentences. He could simply fail to recognize that the theory he knows met the relevant constraints, i.e. that it’s theorems *interpret* sentences of L . Knowing a theory of truth that meets all the constraints is therefore not sufficient for being a competent speaker of L . In addition the interpreter must know that the theory of truth θ is an empirically and formally adequate theory for the language L . So Davidson arrives at the thesis:

“A T-sentence of an empirical theory of truth can be used to interpret a sentence, then, provided we also know the theory that entails it, and know that it is a theory that meets the formal and empirical criteria.”²⁸

But of course we can use a T-sentence to interpret a single object language sentence S without knowing an adequate theory, we only need to know *that* it is a theorem of an adequate theory (Perhaps a reliable informant has told us that the T-sentence (T3) is a theorem of an adequate theory of truth for German). I think we should therefore read Davidson’s claim as relating to the ability to understand a potential

²⁵D. Davidson (1984), p. 136.

²⁶D. Davidson (1984), XIV

²⁷cp. D. Davidson (1984), p. 74, p. 134.

²⁸D. Davidson (1984), p. 175.

infinity of sentences (I use “D” to abbreviate the predicate expressing the constraints Davidson has imposed on a theory of truth):

If S knows a theory of truth θ for L and S knows that $D(\theta)$, then S has the ability to understand all potential utterances of L -sentences.

A theory of truth that fulfills Davidson’s constraints does not become a theory of meaning, but, if an interpreter possesses the required additional knowledge, he can *use* the theory of truth to interpret sentences of the object-language.²⁹

I have neither discussed all the details nor all the problems of Davidson’s approach to meaning. But its main thought should by now be clear: By narrowing down the range of extensional truth-theories we can get at meaning without using intensional notions in the theory itself. For instance Davidson’s theory will neither contain theorems nor axioms of the form:

x means ...

Or:

The meaning of x = ...

The theory just delivers theorems of the form:

s is true_L if and only if p

These T-sentences have at first sight nothing to do with the notion of meaning. But under certain conditions a theory of truth can fulfill the additional function of serving as a theory of meaning. The theorems of a privileged theory of truth neither state the meaning of object-language sentences nor do they describe or name meanings. But someone who has the right extra-knowledge can use these theorems to interpret utterances of the language. McDowell and Evans focus on this thought in their attempt to develop a theory of Fregean sense. I will now outline their ideas.

4 Truth Conditions and Sense

What makes Davidson’s semantic theory attractive for someone who sympathizes with Frege? A standard complaint is that Frege uses the notion of sense but does not provide an account of it.³⁰ Frege’s main motive for the introduction of the notion of sense is his attempt to solve the so-called “identity puzzle”: how can the true sentences “Hesperus = Phosphorus” and “Hesperus = Hesperus” differ in cognitive value? A semantic theory which only ascribes reference or extension to singular terms seems to have no answer to this question. Frege’s answer is, in short, that the first sentence is informative, because “Hesperus” and “Phosphorus” differ in *sense*. But Frege’s answer seems not to be very helpful. Mark Sainsbury points out:

²⁹cp. J. Foster (1976), p. 19.

³⁰M. Dummett (1981), p. 227.

“One could accuse Frege of treating sense as whatever explains the identity puzzle; in which case the explanation is inadequate, being a mere relabelling of the problem. For Frege’s explanation to be satisfactory, he must give an account of sense which goes beyond characterizing it as whatever explains the puzzle.”³¹

Frege ascribes a sense to all expressions which can be used in meaningful discourse. Frege’s conception of sense for singular terms can be characterized thus:

- (F1) The sense of “a” determines the reference of “a”.
- (F2) The sense of “a” is part of the thought expressed by an assertoric utterance of “Fa”.
- (F3) The sense of “a” is the reference of “a” in indirect contexts like “S believes that”

But (F1) to (F3) hardly amount to a satisfactory account of sense.

For McDowell and Evans Davidson’s theory gives us an idea how a satisfactory account of sense must look. Although Davidson is primarily concerned with linguistic meaning, he himself uses sometimes Fregean terminology to express the following ambitious thesis:

“Frege held that an account of language requires us to attend to three features of sentences: reference, sense, and force. [...] I have argued that a theory of truth patterned after a Tarski-type truth definition tells us all we need to know about sense. Counting truth in the domain of reference, as Frege did, the study of sense thus comes down to the study of reference.”³²

Giving an account of the truth-conditions of a sentence and giving an account of the sense it expresses are not two separable tasks, in fulfilling the first, one already fulfills the second. McDowell and Evans subscribe to this thesis according to which Davidson’s theory of truth can fill the gap in Frege’s account of natural language. Evans writes for instance:

“Although a theory of meaning for a language must give the senses of expressions, we are not to think of the theory of meaning as a separate tier, additional to and independent of the theory of reference. If sense is a way of thinking of reference, we should not expect to be given the sense of an expression save in the course of being given the reference of that expression. Rather than look for a theory quite independent of the theory of reference, we must take one formulation of the theory of reference—the formulation of the theory which identifies the references of expressions in the way in which one must identify them in order to understand the language—and make it *serve* as a theory of sense.”³³

We can think for the purposes of this paper of the theory of reference as that part of a theory of truth which determines the reference of the singular terms of the language.

³¹M. Sainsbury (1995), p. 68. Cp. also: W. Taschek (1995), p. 71.

³²D. Davidson (1984), p. 109.

³³G. Evans (1981), p. 295.

There will be many different truth-theories which assign correct references to the singular terms of the language. But only a theory of reference that is part of a theory that can serve as a theory of linguistic competence (a theory of understanding) can serve as a theory of sense. And a speaker who knows that can use a theory of reference as a theory of sense.³⁴

Evans uses the *Tractatus* metaphor of stating and showing to illustrate this conception of sense: a T-sentence of an interpretative theory of truth *states* the truth-condition of an object-language sentence and it shows its sense.³⁵ McDowell speaks of the Davidsonian *Ersatz* for the Fregean notion of sense.³⁶

In his *Grundgesetze der Arithmetik* Frege seems sympathetic to the truth-conditional view of sense. Frege stipulates in this writing which reference the subsentential expression of the arithmetical language have. The stipulations for the subsentential expressions determine under what conditions a sentence composed of them is true. Someone knows the stipulations for the subsentential expressions can derive from them and the composition rules for the arithmetical language under what conditions the sentence refers to the True, or, in a more familiar vein, the sentence is true. Frege then connects truth-conditions and sense in the following passage:

“It is determined through our stipulations under what conditions [any sentence of the arithmetical language] stands for the True. The sense of this name [of a truth-value, i.e. the sense of this sentence], that is the thought, is the sense or thought that these conditions are fulfilled.”³⁷

While Frege is concerned with the formal language of arithmetic, Davidson is interested in natural languages. But the basic idea seems to be the same: if the truth conditions of a sentence are specified in the *right way*, then a specification of truth conditions is a specification of sense. In Frege’s case “right way” means something like “derived (in a canonical way) from the stipulations for the subsentential expressions of the formal language”. Davidson’s explication of “right way” for theories of meaning for natural languages cannot rely like Frege’s explication for formal languages on stipulations. Davidson must appeal to empirical constraints on a truth-theory for the natural language and the idea of a canonical-derivation of truth-conditions.

The truth-conditional theory of sense demystifies the notion of sense.³⁸ The semanticist only needs a privileged theory of truth and additional knowledge about it in order to use a theory of truth as a theory of sense. So we don’t have to assume that there are two theories, one of sense, one of truth. Strictly speaking it is even false to speak of a theory of sense. There is only the *function* of serving as such a theory.

The truth-conditional theory of sense not only fills the explanatory gap in the Fregean account of natural language, it has an additional attractive feature. Perry

³⁴G. Evans (1982), p. 26.

³⁵G. Evans (1981), p. 295. cp. also: M. Dummett (1981) p. 227.

³⁶cp. J. McDowell (1980), p. 143.

³⁷G. Frege (1893a), Sec. 32.

³⁸G. Evans (1982), p. 26.

and Kaplan have ascribed to Frege the thesis that the sense of a singular term must be the sense of a definite description. Their criticism of Fregean semantics is mainly based on this interpretation. They argue that Frege’s alleged assumption is incompatible with our intuitions about (i) the modal properties of singular statements (Kaplan) and (ii) our epistemic intuitions about indexical utterances (Perry) (More on this in section 6). Evans thinks the Ersatz conception enables Frege to escape the criticism put forth by Kaplan and Perry because it does not force us to equate the sense of a singular term with the sense of a definite description, an individual concept. According to the truth-conditional theory of sense a sentence like:

The referent of “Afla” = Afla

shows the sense of “Afla” without identifying it with the sense of a definite description.

Evans’ work is well-known for the provocative claim that empty singular terms are *meaningless*: a sentence with an empty singular term does not express a Fregean thought. Some passages in Evans’ book suggest that one motive for this thesis is Evans’ adoption of the Ersatz conception of sense. The unmodified Ersatz conception takes Frege’s metaphor of sense as a mode of presentation quite literally. The sense of a singular term is given by an axiom which specifies the terms referent. So it seems that a singular without a referent cannot be ascribed a sense.³⁹ The only exception for Evans are so-called “descriptive names”, names whose reference is fixed by a definite description. The reference-axioms for descriptive names are modelled on:

$(\forall x)$ (The referent of “a” = $x \leftrightarrow x$ is the ϕ)

But there are developments of the Ersatz conception which do not force this conclusion upon us. If the theory of truth contains reference-axioms like:

$(\forall x)$ (The referent of “Pegasus” = $x \leftrightarrow$ [Pegasus] (Pegasus = x))⁴⁰

in which the brackets indicate that the name has narrow scope we can ascribe sense to empty names, even if they are non-descriptive. The above axiom is true although “Pegasus” is empty. Nonetheless the axiom shows the sense of the singular term, without giving a description.

So the truth-conditional conception of sense in its original version can be used to illustrate Evans’ thesis about empty singular terms, but it does force us to accept it. This is not to say that Evans has no independent arguments for his No-Reference-No-Sense thesis.⁴¹ But I won’t scrutinize them in this paper. I will investigate whether Evans’ conception of sense can really do the work it is supposed to do, namely specify what a speaker has to know in order to understand an (indexical) utterance. In order to prepare this topic I will discuss in the next section Davidson’s truth-theoretic treatment of indexicals.

³⁹G. Evans (1982), p. 27 and p. 35.

⁴⁰Cp. T. Burge (1987), p. 45.

⁴¹Cp. G. Evans (1982), chap. 9.

5 “A (very large) fly in the ointment”

When Tarski tried to define a truth-predicate for a language, he considered only so-called “formal languages” that fulfill the condition that “the sense of every expression is unambiguously determined by its form.”⁴² In a formal language different tokens of the same (type-) sentence cannot vary in truth-value.⁴³ In his truth-definitions for formal languages Tarski used therefore a monadic truth-predicate for sentence types.

Natural languages don’t fulfill the form-determines-sense condition. Different tokens of the sentence “I live in Heritage CT. 2, Atherton” can differ in truth-value. An indexical type-sentence like “I live in Heritage CT. 2, Atherton” is therefore neither true or false.

Indexical expressions thus pose a problem for Davidson’s attempt to develop a Tarski-style theory of truth for a natural language like English. Tarski’s method of defining a monadic truth-predicate which applies to sentence-types cannot be extended to natural-languages without modification. If we want to define a truth-predicate for a natural-language L , we can no longer assume that (type-)sentences are true or false simpliciter. But how should we deal then with the truth-predicate in application to a natural language?

Davidson answers:

“We could take truth to be a property, not of sentences, but of utterances, or speech acts, or ordered triples of sentences, times, and persons; but it is simplest just to view truth as a relation between a sentence, a person, and a time. [...] The theory of meaning undergoes a systematic but not puzzling change: corresponding to each expression with a demonstrative element there must in the theory be a phrase that relates the truth conditions of sentences in which the expression occurs to changing times and speakers. Thus the theory will entail sentences like the following:

(T4) ‘I am tired’ is true as (potentially) spoken by p at t if and only if p is tired at t .

(T5) ‘That book was stolen’ is true as (potentially) spoken by p at t if and only if the book demonstrated by p at t is stolen prior to t .⁴⁴

The first interesting observation to be made is that Davidson’s T-sentences provide truth-conditions for *potential* utterances. This raises a problem if one, like Davidson, favours an extensional first-order language as the meta-language of the truth-theory, because one needs more demanding logical notions to spell out the content of T-sentences like (T4) and (T5). Evans for instances uses notions from modal logic like Lewis’ counterfactual-conditional and an actuality-operator to give an explanation of the relation expressed by “ S is true as potentially uttered by x

⁴²A. Tarski (1956), p. 166.

⁴³At least if we hold that sense determines truth-value.

⁴⁴D. Davidson (1984), p. 34. My sentence numbering.

at t ".⁴⁵ I won't go in these problems here, since Davidson now thinks that it is not necessary to use "potentially" in an account of the truth-conditions of indexical sentences. He writes in his "Reply to Wolfgang Kühne":

"The point of "potentially" was to provide truth conditions for possible as well as actual utterances. But if one realizes, as I belatedly did, that T-sentences are natural laws, i.e. generalizations of a lawlike character concerning some individual or community, then we no more need to speak of what is potential here than we do in stating the laws of physics."⁴⁶

The second thing to notice is that Davidson's T-sentences for indexical sentences are universally quantified statements.⁴⁷ The expressions " p " and " t " in (T4) and (T5) are variables ranging over all speakers and times. Davidson relativizes the truth-predicate and the new argument-places show up on the right-hand side of the bi-conditional as variables ranging over speaker and times. In contrast no variables appear in a T-sentence for a non-indexical sentence, which does not itself incorporate variables. For the purposes of further discussion I will call the quantified T-sentences "*generic* T-sentences".⁴⁸ One main problem will be the derivation of the 'right' specific T-sentences for individual utterances from the generic T-sentences.

Now the crucial point concerning T-sentences for indexical object-language sentence is made by Davidson himself:

"In natural languages indexical elements, like demonstratives and tense, mean that the truth conditions for many sentences must be made relative to the circumstances of their utterance. When this is done, the right side of the biconditional of a T-sentence never translates the sentence for which it is giving truth conditions. In general, an adequate theory of truth uses no indexical devices, and so can contain no translations of a very large number and variety of sentences. With respect to these sentences, there is not even the illusion that interpretation depends on the ability to translate. (The 'means that' idiom does no better here.)"⁴⁹

Davidson initially promised to capture the notion of meaning in terms of truth-conditions: he seems to hold that an empirically confirmed truth theory for L will entail theorems which could be used to specify the meaning of sentences in L : a

⁴⁵G. Evans (1985b), S. 360.

⁴⁶D. Davidson (1993), S. 21.

⁴⁷Another point concerns Davidson's relativisation of the truth-predicate. The truth-predicate we use in ordinary language is monadic and not tetradic. One strategy which accepts this fact abandons the view that sentences are truth-value bearers. If we use the notion of a proposition, we could modify Davidson's proposal in a way that incorporated a monadic truth-predicate:

What p said in uttering 'I am tired' at t is true if and only if p is tired at t .

But since there seem to be no decisive reasons to accept ordinary linguistic practice, this is a minor point. Davidson himself seems to hold that the expression "sentence s is true (as English) for speaker u at time t " is simply a *façon de parler* for "the statement expressed by sentence s (as English) by speaker u at time t is true" (D. Davidson (1984), p. 44.)

⁴⁸I take this term from I. Rumfitt (1993), p. 441.

⁴⁹D. Davidson (1984), p. 175.

speaker who knew that an adequate theory of truth entailed the T-sentence of the form:

s is true_L if and only if p

could derive from this knowledge a ‘meaning-sentence’ of the form:

s means_L that p

If the meta-language contains the object-language such a sentence would provide a translation of the object-language sentence in the meta-language.

Davidson thinks that *in general* an adequate theory of truth will not contain indexicals. This thesis is Davidson’s reason for his claim that T-sentences for indexical sentences—in contrast to T-sentences for non-indexical sentences—cannot be used to specify the meanings of these sentences: The right hand-side of a T-sentence doesn’t have the same meaning as the sentence referred to on the left-hand side.

If we accept Davidson’s initial premise that the meta-language of a theory of truth must be non-indexical his conclusion that even an adequate theory of truth that does not qualify as a theory of meaning follows. A discussion of (T5) will show this. Since (T5) is a quantified statement we have on the right-hand side no complete sentence which could be used to give the meaning of the object-language sentence. So in order to derive a T-sentence (T5’) which is a candidate for a meaning-giving T-sentence from (T5) we must instantiate the relevant variables. If the speaker a utters at noon the sentence “That book was stolen” and points to a book a legitimate instantiation of (T5) is:

(T5^I) ‘That book was stolen’ is true as (potentially) spoken by the speaker a at noon if and only if the book demonstrated by the speaker a at noon is stolen prior to noon.

Can (T5^I) be used to specify the meaning of “That book was stolen” as uttered by a at noon? No, because “That book was stolen” as uttered by a at noon does not *mean* that the book demonstrated by the speaker a at noon is stolen prior to noon.⁵⁰

If one needs a justification for this non-synonymy claim one can appeal to an intuitive criterion for difference of meaning: two sentences S (as used on a particular occasion) and S' (as used on a particular occasion) differ in meaning, if somebody who understands S and S' can have different attitudes to them (accept S and reject S' or vice versa). Now I can perfectly understand and accept the sentence “That book was stolen” (as used by a at noon) without eo ipso accepting the sentence “The book demonstrated by the speaker a at noon is stolen prior to noon”. I could simply fail to identify the aspects of utterance context that figure in the definite description. Perhaps I don’t know that the speaker is called “ a ” or that the time of utterance is noon.

⁵⁰cp. J. Bennett (1985), p. 605–606.

But why should someone accept Davidson's initial premise that an adequate theory of truth uses no indexical devices? I could find in Davidson's early articles only a convincing argument against treating indexicals like proper names as terms with constant reference in a theory of truth. Such a treatment would give us T-sentences like

(T4C) 'I am tired' is true if and only if I am tired

that are derived from axioms like:

('I'-A) 'I' refers to me.

Davidson comments on this treatment:

"What suffers in this treatment of demonstratives is not the definition of the truth predicate, but the plausibility of the claim that what has been defined is truth. For this claim is acceptable only if the speaker and circumstances of utterance of each sentence mentioned in the definition is matched by the speaker and circumstances of utterance of the truth definition itself. It could also be fairly pointed out that part of understanding demonstratives is knowing the rules by which they adjust their reference to circumstance; assimilating demonstratives to constant terms obliterates this feature."⁵¹

The truth of the sentence "I am tired" does not depend on the state of one single speaker. So what is defined here is not truth.

The treatment of indexicals as semantic constants has another devastating consequence. Knowledge of a theory of truth which entailed T-sentences like (T4C) could be possessed only by one person, namely the one denoted by "I". And knowledge of such a theory would not be sufficient for understanding utterances of natural language like English. An indexical theory of truth would enable the person who knows it just to interpret his own utterances or utterances in specific contexts. This contradicts the very idea of a theory of truth as theory of the mastery of a natural language.

Although Davidson's argument shows that indexicals cannot be treated as constants in a theory of truth it does not discount the possibility to integrate indexicals in other ways into the theory of truth. For instance indexicals could be used to instantiate the T-sentences (T4) and (T5). If we instantiate (T4) in the following way, we would get the right result without treating indexicals like constants:

(T4^{I1}) 'I am tired' is true as spoken by me now if and only if I am tired.⁵²

The generic T-sentences reflect the context dependence of indexical reference, the indexical instantiations of the generic T-sentences yield meaning-giving T-sentences.⁵³ This reflection vindicates Prior's judgement about Davidson's non-indexical theory of truth. Prior wrote:

⁵¹D. Davidson (1984), p. 34.

⁵²We can drop "now" on the right-hand side of the T-sentence if we hold like Prior the redundancy theory of the present.

⁵³I. Rumfitt (1993), p. 443 fn23.

“If we are to give not only the truth-conditions but the meaning of tensed and first-person utterances, we not only may but must use tensed and first-person languages.”⁵⁴

One can speculate why Davidson resists this conclusion. One idea is that he wants to treat all T-sentences as law-statements and only a universal sentence is a candidate for law-statement. The meaning-giving T-sentences are particular sentences. Therefore they cannot be laws and hence they cannot be sentences of a theory of truth in Davidson’s sense. But even if we accept this line of reasoning there is still room for these ‘T-sentences’. Perhaps the meaning-giving sentences cannot be *part* of the truth-theory but they are, one might say, what results if one applies the theory in order to interpret indexical utterances. We could then say that a theory of truth is a theory of meaning if it yields the right ‘T-sentences’ when applied to interpret (indexical) utterances.

Does a Davidsonian theory of truth then after all qualify as a theory of meaning? No, because nothing in the theory requires that a generic T-sentence is instantiated (i) by indexicals *at all* nor (ii) by the *right* indexicals.

Suppose I utter “I am tired” on January 12th 1995 at 12pm. The truth of (i) is shown by the T-sentence (T4^{I2}):

(T4^{I3}) “I am tired” is true as spoken M.T. at 12pm on January 12th 1995 by if and only if M.T. at 12pm on January 12th 1995 is tired.

(T4^{I3}) is also a legitimate instantiations for (T4) and shows (ii) to be true:

(T4^{I3}) “I am tired” is true as spoken by M.T. at 12pm on January 12th 1995 if and only if this person (M.T. is pointed out) is tired at this moment.

Neither (T4^{I2}) nor (T4^{I3}) give the meaning of “I am tired” as used on this occasion. Yet from the perspective of a theory of truth they are legitimate instantiations of (T4).

The problem of a Davidsonian theory of truth for an indexical language is that the theory requires that the interpreter identify features of the context to interpret utterances of indexical sentences, but the theory does not tell him which of his non-semantic information about utterance context is apt to yield meaning-giving instantiations of generic T-sentences. “[A] good semantic theory” writes Sainsbury, “ought to tell [the interpreter] not to let his non-semantic information surface in the wrong place.”⁵⁵ Davidson’s theory of truth is no good semantic theory in this respect: it places no restrictions on legitimate instantiations of T-sentences for indexical sentences and such restrictions are needed if a theory of truth should serve as a theory of meaning.

The same problem arises for the notion of understanding. Knowledge of generic T-sentences for indexical sentences and knowledge of facts about the utterance of the indexical sentences is not sufficient for understanding indexical utterances at least

⁵⁴A. N. Prior (1970), p. 3.

⁵⁵M. Sainsbury (1996) p. 6.

in the strong sense of understanding. In order to understand my utterance of “The meeting is taking place now” I must think that the meeting is taking place now. But Davidson’s theory does not discriminate between this instantiation of the relevant generic T-sentence and various others which also qualify as something yielding an interpretation of my utterance.

6 Indexical Sense

The last section has shown that Davidson’s original theory of truth for a natural language cannot serve as a theory of understanding. Evans and McDowell think nonetheless that a theory of truth is the best way to clarify Frege’s notion of sense. One advantage of the truth-conditional conception of sense is that it does not force upon Frege the view that the sense of a singular term is an individual concept. Evans’ article “Understanding Demonstratives” (G. Evans (1981)) is an attempt to rebut John Perry’s criticism of Fregean semantics on the basis of the truth-conditional conception of sense. In the next section I will investigate whether Evans’ reply to Perry can solve the problems we already discussed.

I will first rehearse Perry’s criticism of Frege. Perry argues that Frege’s theory cannot incorporate indexical expressions. Consider the sentence:

(S4) “Today is fine”

Let us suppose I utter (S4) on February 23th 1995. Perry asks in his article “Frege on Demonstratives” which sense completes the linguistic meaning of the predicate “is fine” to the Fregean thought expressed by my utterance. Two simple answers to the *completion problem* are easily discarded: neither the reference of “today” in the context of the utterance – the day d – nor the linguistic meaning of the type-expression can serve as the completing sense. The day d is a reference corresponding to indefinitely many senses, but it is not itself a sense. The linguistic meaning of “today” is disqualified as the expressions sense, because it does not vary from occasion to occasion. If the linguistic meaning of “today” were part of the Fregean thought, as Frege’s doctrine requires, the thought could no longer be true or false simpliciter.

Perry describes the only possibility which is in his opinion left to Frege thus:

“How can we extract from a demonstrative an appropriate completing sense? Such a sense, it seems, would have to be intimately related to the sense of a unique description of the value of the demonstrative in the context of utterance [its reference].”⁵⁶

Perry does not give the notion of an ‘intimate relation’ between the sense of a unique description and the sense of a demonstrative any substance. The only relation which is relevant in his texts in this connection is very intimate one indeed. Perry takes the Fregean approach to be based on the idea that the sense of a demonstrative is

⁵⁶J. Perry (1977), p. 485.

identical with the sense of a unique description of the demonstratum. Perry then goes on to show that the sense of a definite description of d cannot be the completing sense Frege needs.

Evans thinks that this argument has loop-hole. And to exploit it he uses a theory of the type we have been discussing in the last sections.

Evans holds that the basic insight of Frege was the thought “that to understand an expression, one must not merely think of the reference that it is the reference, but that one must, in so thinking, think of the reference *in a particular way*. The way in which one must think of the reference of an expression in order to understand it is that expression’s *sense*.”⁵⁷ So “today” as uttered on day d has a completing sense if and only if one must think of d in a specific way in order to understand the relevant utterance of “today”. Is there such a specific way in which one must think of the day d in order to understand an utterance of “Today is fine” on that very day? Evans’ answer to this question is an emphatic:

“[...] of course there is. Even if d is the first day after my last lecture, I shall not have understood the utterance of [(S4) “Today is fine”] if I think of d only as the first day after my last lecture, thereby coming to believe that the utterance is true if and only if the first day after my last lecture is fine, perhaps not realizing that today is the first day after my last lecture. In order to understand [(S4) as uttered on d] I must think of d as the current day, thereby coming to have the thought which I might express with the words: ‘What the speaker said is true if and only if it is fine today.’ ”⁵⁸

Evans’ consideration has convinced me that “today”, as uttered on d , has a completing sense. If the linguistic data force us to ascribe such a sense to “today”, one of Perry’s assumptions must be false. Which one is it?

Evans holds that Perry’s thesis (DD) is wrong:

(DD) The sense of a demonstrative is identical with the sense of a unique non-indexical description of its reference on an occasion of use.

But in his response to Evans Perry points out that (DD) is not essential to his argument:

“The significant point is not that the demonstrative could not be regarded as an abbreviation for a description. It is rather that the sense of a demonstrative cannot be one that determines the reference independently of context. Evans maintains that my argument depends on the assumption that Frege thought any sense of a referring expression could be captured by a definite description. Perhaps it did, but it needn’t. The real assumption is that we would have to identify the sense the demonstrative has temporarily (at some places, for some people, at some times, on some days) with an expression that has that sense permanently.”⁵⁹

⁵⁷Evans (1981), p. 294.

⁵⁸G. Evans (1981), p. 303. I substituted my numbering for Evans’ own.

⁵⁹J. Perry (unpublished), p. 8. Compare also: J. Perry (1993), p. 15 fn4.

Perry argues that the sense of a demonstrative determines its reference only relative to a context. If we understand by “sense” linguistic meaning, there is no need to quarrel with Perry. But if we understand by “sense” what Frege understands by this word, there is genuine disagreement between Frege and Perry.

I take it that Perry has convincingly shown that there is no unique description which determines the reference of the demonstrative independently of context *and* can be considered to specify its sense. But has Perry thereby shown that we cannot specify the sense an indexical has on an occasion of use in no way at all?

Evans’ negative answer to this question uses the idea that a theory of truth can serve as theory of sense. But we must note first that a theory of truth like the one we discussed in the last two sections is of no help to Evans. Such a theory does not generate sense-giving T-sentences. Evans is interested in a theory that specifies the sense an expression has *on an occasion of use*. “Today” for instance has according to Frege different senses on different days. The theory should, as Evans would express it, *show* the sense an expression has on a specific occasion of use. This expression used to specify this sense expresses it also only relative to a context, but the sense expressed in this context determines its reference absolutely. In order to achieve this the theory of truth Evans envisages should contain T-sentences whose right-hand sides expresses the thought one must have in order to understand an indexical (sentence) on an occasion of use.

Evans tries to give a theory which realizes this objective by integrating indexicals in the meta-language of the theory of truth. He writes:

“[...] in order to understand an utterance of [“Today is fine”] made on *d*, one must have, on *d*, the thought one might express in words by:

What the speaker said is true if it is fine *today*

It seems reasonable to say that such a statement is capable of showing the sense which the sentence has on that occasion.”⁶⁰

Evans’ theory of the completing sense an indexical has on an occasion of use is an indexical theory of truth, whose T-sentences are themselves context-dependent. This bold theoretical move seems to solve the problems Perry has pointed out for the Fregean theory. The sense of an indexical expression on an occasion of use is given by an indexical T-sentence and not by an expression which has this sense permanently. In this way Evans realizes Prior’s advice mentioned on page 19. But bold though it may be Prior’s and Evans’ theoretical alternative faces the problems which arose for a theory of truth with an indexical meta-language in the previous section. Let’s see whether Evans has a solution for these problems.

7 Objections to the Truth-Conditional Theory

In the first section of my critical discussion of Evans’ proposal I will assess John Perry’s criticism of Evans’ thesis. Perry has developed two arguments which I will

⁶⁰G. Evans (1981), p. 304. I think we can safely interpret the “if” as “if and only if”.

now present.

1. *Evans' theory is just a partial theory of understanding.* Is Evans' claim that in order to understand an utterance of "Today is fine", made on *d*, one must think the thought expressible with the words "What the speaker said is true if and only if it is fine today" true? Perry's answer is contained in the following example:

"To see that something has gone wrong here, simply imagine receiving a postcard, written on *d* that says "It is fine today". One would not have understood this if one had the thought which one would express in the words "What the writer said is true if and only if it is fine today."⁶¹

From this observation Perry draws the conclusion:

"What Evans seems to be giving us is a theory based on one very special case, that in which the context of utterance and the context of understanding are the same with respect to the factor crucial to the demonstrative in question."⁶²

This conclusion seems to me basically right. But I think also that Evans can concede this point. This becomes clearer, when we relate Evans' proposal to the distinctions made in the introduction of this paper.

Evans is interested in a theory of understanding indexical utterances in the strong sense, namely in the sense in which it is necessary to identify the referent of the relevant indexical on the occasion of use in a *specific way* in order to understand the utterance. In this sense one cannot understand an utterance of "Today is fine" on a day *d* if one is not in a position to think a thought about *d* one could express with the words: "Today is fine, that's what the speaker is saying." Dummett points out that this conception of understanding is also the one Frege was aiming at in his writings:

"What is needed, to grasp the thought expressed by an utterance – to understand it in Moore's and Evans's sense – is to know the meaning of the words, to know the relevant constructions, and to apprehend in the appropriate way the reference of the indexicals and demonstratives; and Frege's whole use of the term 'sense' accords with this. It is certainly wrong to overlook the respects in which his notion of sense diverges from that of linguistic significance: [...] "⁶³

Perry's example is one in which the person who receives the postcard is not in the position to grasp the thought expressed. A necessary condition for understanding in the demanding sense seems to be that the objects indexically referred to in the context of utterance are also indexically accessible in the context of understanding. This condition is not fulfilled in Perry's example: the addressee of the postcard cannot refer to the day *d* with an indexical or at least not with the indexical whose use would show that he has understood the utterance in the demanding sense. Therefore she cannot understand the indexical utterance in the sense outlined above.

⁶¹J. Perry (unpublished), p. 10.

⁶²ibid., p. 11.

⁶³M. Dummett (1984), S. 202.

The sense in which the addressee of the postcard understands the utterance can be accounted for by the notions of linguistic meaning and knowledge about the utterance which is necessary to identify it as an utterance of an English sentence. If the addressee is a competent speaker of English and she recognizes that the inscription on the postcard is an inscription of the English sentence “It is fine today”, she knows thereby that the inscription is true if and only if it is fine on the day of the production of the inscription. As a competent speaker of English she can even know *in which way* the person who produced the inscription thought of the relevant day: the producer thought of the relevant day in the “today”-way, the way of thinking of a day which is expressed by an utterance of “today” on the day of this utterance.

Since Evans is only interested in the demanding (Fregean) concept of understanding, Perry’s first criticism is unsuccessful. But it is helpful in understanding the character of Evans’ theory.

2. *Evans’ theory yields wrong results.* Perry’s second argument points to a more serious problem of Evans’ proposal. Perry’s first argument was based on the observation that one can understand an indexical utterance (in one sense) without being in the same context as the utterer. Perry’s second argument is based on the idea that Evans theory would require that the interpreter of an indexical utterance and the utterer be in the same context, but that this is not in all cases possible.

An example will help to explain this statement. What must someone think according to Evans’ theory, if he is said to understand an utterance of the sentence “I am having a good time today” by *a*? Perry generalizes Evans’ theory. Perry envisages the following statement as “showing” the sense this sentence has as made by *a* on *d*:

“In order to understand the utterance of ‘I am having a good time today’, made by *a* on *d*, one must have on *d*, the thought one might express in words by ‘What the speaker said is true if I am having a good time today’.”⁶⁴

Perry envisages this statement as derived from a theory of truth which incorporates indexicals in its meta-language. More specifically, the reference-axioms of the theory of truth are indexical sentences like:

The referent of “I” = I

The referent of “you” = you

The relevant theorem might “show the sense” the utterance has as made by *a* on *d*. But, Perry comments, the theory “doesn’t work for anything but monologue”. And this is surely a “crazy result”.

Is this a knock-down argument against Evans’ proposal? I think that Perry’s criticism points to difficulties with Evans’ approach, but does not identify them clearly. In the next section I will try to answer Perry’s main arguments.

⁶⁴ibid., p. 10.

8 T-sentences and Quasi-indicators

Section 5 has already provided the building blocks for an answer to Perry’s second objection. Perry’s reference axioms treat indexicals as constant terms. This leads to the ‘wrong results’. But we know now that a theory of truth can adopt a different treatment of indexicals. The theory can deliver ‘specific’ T-sentences like:

(T4^{I1}) ‘I am tired’ is true as spoken by me now if and only if I am tired

if it contains generic T-sentences like the ones discussed in section 5. The specific T-sentences can interpret indexical utterances: they give the ‘right result’. The basic problem with Evans’ Ersatz conception is another one.

But Perry may have a more general objection in mind. T-sentences like (T4^{I1}) interpret indexical utterances only *for a specific interpreter*. For instance (T4^{I1}) only interprets the utterance if utterer and interpreter are identical. This is an unwelcome restriction of the theory.

Perry could argue that a theory of truth which could serve as a theory of understanding must meet the following constraint:

- The T-sentences of the theory of truth must state something which could be known independently of utterance context.

This is a rather strong constraint on a theory of truth. Can it be met?

Hector-Neri Castañeda has claimed that in natural language has a semantic mechanism, Castañeda calls the vehicles of this mechanism “quasi-indicators”, which seems to be designed to solve this problem. If this mechanism can be incorporated in a meaning-giving theory of truth the problem developed above can be solved. Under this assumption we can devise specific T-theorems which state something that (i) can be known independently of specific contexts and (ii) interprets indexical utterances. I will now introduce Castañeda’s basic concepts.

When it comes to specify the uses of expressions like “he” or “then” most writers list three different kinds of uses. These expressions can be used as

- indexicals
- anaphorical pronouns
- bound variables

Castañeda has argued in a series of papers⁶⁵ that the list is incomplete. An example illustrates Castañeda’s point. I can report a belief Dr. Lauben has about Dr. Lauben in the following way:

(S5) Dr. Lauben believed that he was wounded.

The expression “he” functions in (S5) not as a demonstrative or an anaphorical pronoun. Its semantical role is similar to an anaphorical expression in that its

⁶⁵Starting with H. - N. Castañeda (1966) continuing with (1967) and (1968).

reference is dependent on the grammatical antecedent “Dr. Lauben”. But unlike an anaphor the grammatical antecedent cannot be substituted *salva veritate* for the expression in its scope. Dr. Lauben may forget his name and still believe *himself* to be wounded. Therefore the sentence

(S6) Dr. Lauben believed that Dr. Lauben was wounded.

and (S5) can have different truth-values in the same situation. Castañeda holds therefore that in English there are two different words with the same spelling:

“It is only a linguistic freak that “he” in the sense of “he*” looks exactly like the third person pronoun “he”, which occurs, for instance in “Arthur came, but he knew nobody he saw; he left early.”⁶⁶

Castañeda distinguishes the non-deictic “he” from the deictic one either by indexing it with “*” or by adding in parentheses “himself”.⁶⁷ Castañeda names expressions which function in the just described way *quasi-indicators*.

Quasi-indicators and indicators in belief-ascriptions in oratio recta are distinguished by a an interesting property. Consider first:

(S7) John believes that I am wounded.

In (S7) the indicator “I” is ‘in the mouth’ of the speaker who attributes a belief to John. The indicator is used to refer to the speaker and this way of referring to the speaker is surely not one that is available to John to whom the belief is ascribed. This gives rise to general characteristic of indicators in oratio obliqua which Castañeda described in the following way:

“An indexical reference *R* in *oratio obliqua*, [(1)] that is not part of an attribution of self-knowledge (self-belief, self-conjecture, etc.) leaves it by itself wholly unspecified how the person to whom knowledge (belief, etc.) is attributed refers to the person or object to whom *R* is made.”⁶⁸

Using Castañeda’s apt term one can say that an indicator in the content-clause of an oratio obliqua construction is *propositionally opaque*: the indicator does not specify how the person to whom the belief is ascribed refers to the item he believes something about. So the content-clause in (S7) does not reveal the complete content of Johns thinking.

The saving clause (1) is well motivated. My utterance of:

(S8) I believe that I am wounded

specifies in the most precise manner how the person to whom the belief is ascribed refers to the item he believes something about: “I” is *propositionally transparent* in the content clause of a *self*-ascription.

In contrast to indexicals quasi-indicators are according to Castañeda propositionally transparent: a construction like

⁶⁶H.- N. Castañeda (1966), p. 132.

⁶⁷ibid., p. 132.

⁶⁸H.- N.Castañeda (1968), p. 440

(S8) John believes that he* is wounded.

reveals the proposition which is the content of John's belief. It is the proposition John would express with the words "I am wounded" or a synonymous sentence.

This makes an important similarity between anaphorical pronouns and quasi-indicators visible. In its anaphorical use "he" carries extra information about the referent of its grammatical antecedent: The antecedent refers to a male being. Quasi-indicators also carry extra information. "He*" in a context like (S7) carries information about the *way* John conceives of John in the attributed belief.⁶⁹

This characteristic of quasi-indicators seems to be what is needed for Evans' account. In order to show this I rewrite the above criticized statement with quasi-indicators:

"In order to understand the utterance of "I am having a good time today", made by *a* on *d*, one must have on *d*, the thought one might express in words by 'What the speaker said is true if he* is having a good time today'".

This seems to give the right results at least in cases in which I am not the utterer. I understand an utterance of "I am F" in the strong sense if and only if I come to know that it expresses a true thought at the time of utterance if what the speaker said is true if and only if he* is F. "He*" is here a quasi-indicator which is chained to the designation of the utterer. This suggests a modification of Evans' strategy. The general rule should not be "Include indicators in the meta-language of the theory of truth" but "Include *quasi*-indicators in the meta-language of the theory of truth".

This rule takes for granted that there are quasi-indicators. But are there any in a natural language like English? Castañeda's homonymy-thesis is surely not intuitively plausible. Boër and Lycan as well as Perry argue that we can account for Castañeda's data by pragmatic explanations. Therefore it is not necessary to assume that there are words like "he*" which have a distinct semantic role.⁷⁰ Let's take a look at their arguments against Castañeda's view.

If "he" in constructions like "S believes that he is F", where "S" and "he" are co-referential, had the distinctive semantic role Castañeda assigns to this word, then this construction should *imply* the truth of the following sentence:

(I) S is willing to utter "I am F" under appropriate circumstances.

But this implication does not hold. Boër and Lycan remark:

"[S] may not be willing to assert anything, having been handsomely bribed by Boër and Lycan (who will stop at nothing to gain their philosophical ends) never to assert anything again."⁷¹

Boër and Lycan concede that this remark only shows that the potential implicatum has not been described in the right way. They think that the potential implicatum could be:

⁶⁹Cp. R. Brandom (1994), p. 564.

⁷⁰cf. S. E. Boër, W. Lycan (1980) and J. Perry (1983)

⁷¹S. E. Boër, W. Lycan (1980), p. 454.

(I') S tokens a mental sentence which says the same as S' utterance of the public language sentence "I am F".⁷²

But this is no real option for someone who is not willing to accept the language of thought hypothesis.

Yet there is still a third alternative. Peter Geach and Wilfried Sellars have pointed out that oratio recta constructions can be used *metaphorically* to report thoughts. So we can say "S said in his heart 'The weather is fine today'" in order to ascribe a thought to S. Using this way of speaking we can reformulate the alleged implicatum:

(I'') S says in his heart "I am F".

This belief-ascription is metaphorical. Its literal content could be spelled out along the following lines: S has a belief such that: if S would express it, he would express it with an utterance of "I am F".

Does a construction like "S believes that he is F" imply the truth of a sentence like (I')? We must distinguish between implication and implicature. If p implies q, it is not possible that q is false and p true. But if p implicates q, it is possible that q is false and p true. Implicature is best understood as a (in some cases conventionalized) form of suggestion. The sentence "John doesn't make much money but he is happy." implicates (suggests) the truth of the sentence "If someone doesn't make much money, he tends to be unhappy". But of course this utterance doesn't imply the truth of the second sentence.

We have a good reason to suppose that p implicates q, but doesn't imply it, if we can cancel q without contradicting our original statement. Castañeda often uses "himself" in place of "*" to distinguish the non-deictic "he" from the ordinary third person pronoun. Does the construction "S believes that he himself is F" imply (and not just implicate) the truth of the statement "S says in his heart 'I am F' "? Perry gives a negative answers to this question. His argument relies on a counter-example to the implication thesis:

"The dean has been complaining that professors who publish less than ten articles per year on the average are overpaid. He has particular ones in mind, Professors A, B, Q, and Z in the blind sample he has been studying. Then one day he counts the articles he has written and finds only ninety-three articles over the past teen years, agreeing exactly with the figure for Professor Z, which could not be a coincidence.

We say,

The dean was surprised to find that he believed himself to be overpaid."⁷³

Perry's stage-setting makes it natural to cancel the alleged implication: the report does not imply that the dean was surprised to find that he believed in his heart 'I am overpaid'. This is not what he is surprised to discover.

⁷²ibid., p. 455.

⁷³J. Perry (1983), p. 110.

Is this a counter-example to Castañeda implication thesis? Has Castañeda confused implication and implicature? No, Perry presents the case in which a de se report is used to ascribe *two mental episodes* with just *one* sentence. One might say that the example relies on a *condensed report*. If one does not spell out what is condensed in the report one creates like Perry the impression that quasi-indicators lack the alleged special implicatum. But how should one expand the attribution of surprise in Perry's example? Here is a plausible proposal:

(S11) The dean (i) believed that Professor *Z* is overpaid, and (ii) the dean is surprised to learn that he himself is Professor *Z*.

Clause (ii) makes clear what is the object of the dean's surprise, namely the fact the dean would express with the words: "I am Professor *Z*". This way of elaborating Perry's example dispels the impression that one can literally use the quasi-indicator without attributing thereby a de se or ego attitude.

Even if there are perhaps more convincing counter-examples to Castañeda's theory, a semanticist who is attracted to a theory of truth whose meta-language contains expressions like "he himself" can still use these expressions. He can simply stipulate that certain expressions function exactly like Castañeda's quasi-indicators and introduce them into the meta-language of his theory of truth. But now a more serious problem arises: Perhaps we can use quasi-indicators in a theory of truth, but can the theory of truth give the meaning of these expressions by displaying their contribution to the truth-conditions of sentences containing them? Let us take an example.

A T-sentence of the truth-theory enriched by quasi-indicators would look like this:

(T4^{I-Q}) 'I am tired' is true as spoken by M.T. if and only if he* is tired.

If we don't take "he*" to be an expression which is semantically primitive, we must be able to spell out the contribution of "he*" to the truth-condition of "I am tired". How can we do this? As we already noted, quasi-indicators carry extra information. "He*" carries information about the *way* M.T. referred to M.T. If we integrate this information in the truth-condition we get something like:

(T4^{I-Q*}) 'I am tired' is true as spoken by M.T. if and only if M.T. truly attributes to M.T. under the ego-mode of thinking the property of being tired.

If we take this step, we use senses or modes of presentations in the semantic theory. This undermines the very motivation of the truth-conditional theory of indexical sense: the theory should only contain specific assignments of reference, satisfaction and truth, which show the sense of the expressions under consideration, without relying like (T4^{I-Q*}) on the notion of sense or mode of presentation itself. This seems to be a serious problem for the modification of Evans' account. Now one might argue that the modification is not necessary, because the restriction which motivated it is too strong. But even if we accept this, there is another serious problem for the truth-conditional theory.

9 Evans' on the Instantiation-problem

Perry ascribes to Evans a treatment of indexical expressions which is similar to the one Davidson rejected. Indexical expressions are treated as terms with constant reference. But nothing forces Evans' to accept this treatment. Evans can accept that the hearer derives statements like

(T_d) "What the speaker said is true if and only if it is fine today"

from general and non-indexical statements which serve as axioms for the theory of truth. And he comes very near to this view, when he says:

"When we think of principles from which [(T_d)] might be derived, we naturally think of the general statement:

For all days *d*, 'today' as uttered on *d* refers to *d*."⁷⁴

If the meta-language contains indexicals, the speaker can derive from such general and non-indexical statements the indexical T-sentences which interpret specific indexical utterances. In order to do this the speaker must instantiate the variables in the general statements with the right singular terms, namely indexicals.

This limited recognition of indexical expressions in the language of a theory of truth avoids the problems Davidson and Perry have pointed out. The axioms and even the general T-sentences for indexical sentences will be non-indexical. Only their instantiations will be indexical sentences. So the theory of truth can state something which can be common knowledge and can be used by different speakers.

But although this development of the truth-theoretic approach frees us from one serious problem, it is flawed. The problem we encountered with Davidson's original theory is still unsolved. In order to understand an indexical utterance in the strong sense, a speaker must identify the relevant utterance and the relevant elements of the utterance context. He can do this in a variety of ways. And to this variety of ways corresponds a variety of instantiations for the generic T-sentences which yield all correct T-sentences. But only one of these instantiations yields a specific T-sentences which interprets the indexical utterance. And the theory of truth cannot discriminate between the "right" and "wrong" instantiations.

This holds also if we add quasi-indicators to the meta-language of our theory of truth. For instance a theory of truth which can serve as a theory of sense must take into account who is the utterer of an "I"-sentence. If the interpreter and utterer of an "I" utterance are *not* identical the generic T-sentences must according to our proposal be instantiated with an quasi-indicator: in order to understand the utterance of "I am having a good time today", made by *a* on *d*, the interpreter must think on *d* the thought one might express in words by 'What *a* said on *d* is true if he* is having a good time on that day*'. But if interpreter and utterer of an "I"-utterance *are* identical the generic T-sentences cannot be instantiated with an quasi-indicator: in order to understand my utterance of "I am having a good

⁷⁴G. Evans (1981), p. 305.

time today”, made today, I must think today the thought I might express in words by ‘What I said today is true if I am having a good time today’. If I could just instantiate the relevant generic T-sentence with quasi-indicators I would not count as somebody who understood the utterance, because I could fail to notice that I am the utterer. In this situation I would fail to grasp the thought the sentence expressed in my mouth.

Evans thinks the way in which we exposed the problem up to now is misguided. He tries to set things straight in the following passage:

“Speakers do not literally deduce the truth conditions which sentences have from certain universally quantified principles, whose precise form we must endeavour to establish. Speakers judge the truth conditions of particular sentences, and in so doing they exercise complex and interconnected dispositions in which their understanding of the individual atoms of the language may be taken to consist. We are therefore not required to attribute to speakers the general belief that any token of ‘today’ refers to the day on which it is uttered—and then wonder what form that belief takes, or how they derive the right kind of judgement of truth conditions from it.”⁷⁵

This passage must be read bearing in mind Evans’ answer to the question “In which sense has a competent speaker ‘knowledge’ of a theory of meaning for his language?”⁷⁶ When we say that a person ‘knows’ a theory of meaning, we do not mean to say that she holds true each axiom of a theory of meaning. And we don’t think that her understanding of sentential utterances is the result of certain derivations from these axioms. We rather ascribe to a competent speaker a set of dispositions. Each disposition corresponds to an axiom of the theory for an expression. The dispositions are characterized by their joint exercise in judging utterances of sentences to be true under such-and-such conditions. This shows that the characterization of the instantiation-problem given up to know is misleading. A competent speaker of English does not face the problem of deriving the *right* truth conditions for indexical sentences from axioms he somehow believes.

In which way can we positiveley characterize the speaker’s knowledge of the sense of an indexical expression? Evans regards the judgement that an sentential utterance has such-and such truth-conditions as the exercise of two dispositions which cannot be manifested in isolation and which must be interdefined. Let us see how his strategy works in our case. If we assume that we have already defined the concept of tacitly knowing the satisfaction condition of a predicate we can define the connected notion of tacitly knowing the meaning or reference rule of “today” in following way:

U tacitly knows that for all days d , ‘today’ as uttered on d refers to d iff he has a disposition such that:

(II ϕ) and (II ψ) if

(i) U tacitly knows that an object satisfies ϕ iff it is ψ .

⁷⁵G. Evans (1981), p. 305.

⁷⁶cp. G. Evans (1981), p. 328f.

- (ii) U hears an utterance of “Today ϕ ”, then U will judge that the utterance is true iff today is fine.⁷⁷

This definition does avoid the instantiation problem in its original form, but it still does not capture the demanding notion of understanding Evans is after. If our judgements of the truth-conditions of individual “Today is fine” utterances were among others based on the disposition above, then someone who judges on d (and d is the first day after M.T.’s first lecture in Hamburg) upon hearing an utterance of “Today is fine” that the utterance is true if and only if the first day after M.T.’s first lecture in Hamburg is fine, would count as someone who understands the utterance. Now his judgement is of course true, but someone who makes it in these circumstances does not understand the utterance *in the demanding sense* Evans is interested in.

Evans thinks therefore that the reference-rule for “today” does not properly identify the disposition we ascribe when we say that someone tacitly knows the sense of “today”. Evans says:

“We are rather ascribing to speakers a *propensity to form particular beliefs*, of particular tokens of ‘today’, that they refer to the day of utterance, identified in a particular way, the exercise of which yields thoughts of the form of [What the speaker said is true if it is fine *today*]. The inclusion in a theory of reference of a general principle like [For all days d , ‘today’ as uttered on d refers to d] is a gesture in the direction of identifying the relevant propensity, and it certainly needs supplementation.”⁷⁸

The propensity Evans tries to capture is not just a propensity to judge under certain circumstances that a specific token of an expression refers to something, it is a propensity judge under certain circumstances that a specific token of an expression refers to something *identified in a specific way*. How can we, as Evans says, ‘identify’ this propensity properly? It seems that we can only do it if we mention modes of presentation or ways of apprehending a reference in the characterization of the disposition, i.e. in the tacit knowledge of a reference rule. But this of course contradicts the very point of the truth-conditional conception of sense. The truth-conditional conception of sense tries to *show* the sense an expression has by *stating* its reference. If we must *mention* senses in giving the reference of an expression the theory must take them as primitive and cannot specify them by reference axioms.

10 Conditional Truth-Conditions

Our initial question was: Can a theory of truth “show” the sense an indexical has on the occasion of utterance? We have not yet found a theory of truth which allows us to answer this question positively. Perhaps we have followed Davidson too closely in developing a theory of truth for indexical languages. We cannot ascribe an absolute

⁷⁷‘ Π ’ is the universal substitutional quantifier, the variable having the following substitution classes: ϕ , names of the predicate expressions of the (object) language, and ψ : predicate expressions of the (meta) language (Vgl. G. Evans (1981), p. 329.

⁷⁸ibid.

truth-value to an indexical type-sentence. Davidson's reaction to this observation was to relativize the truth-predicate: we no longer have "s is true" but "s is true as (potentially) spoken by p at t ". But there is also an alternative way to deal with indexical expressions in particular and context-dependent expressions in general in a theory of truth. The motivation of the alternative view is sketched by Burge:

"Sentences containing demonstrative constructions are neither true nor false apart from actual use. To evaluate 'That is a dog' as true or false, we need someone to use 'that' in the sentence referentially."⁷⁹

An indexical sentence can be ascribed a truth-value, provided that certain *conditions* are fulfilled. If the speaker has referred to x with "that" in his utterance of "That is a dog", then his utterance is true if and only if x is a dog. This suggests that indexical sentences have conditional truth-conditions.⁸⁰ Does this idea help in devising a theory of truth for an indexical language that can serve as a theory of understanding?

In recent articles James Higginbotham has answered the last question positively. He puts forth the view that a theory of truth that ascribes conditional truth-conditions to utterances can serve as a theory of sense.⁸¹ A theory of linguistic competence will give the truth-conditions of *utterances* according to the following general scheme:

If u is an utterance of Σ by s and $\phi(u, S, s, x, X)$, then

u is true if and only if $\pi(x, X)$

where x and X are first and second order parameters.⁸² Now Higginbotham does not much provide of an explanation of this scheme. But an instance of the schema makes clear what he has in mind:

If u is an utterance of "that is a dog" by s at time t and s refers with "that" in u to x , then:

u is true if and only if x is dog.

Does the conditional T-sentence for "That is a dog" provide a sense for the demonstrative expression? Higginbotham answers:

"In one sense, it appears, they do not; for they enter what is expressed only through the value they determine in an utterance. But in another sense they obviously do; for they require that the speaker has singled out an object to which reference is being made. The sense that the words carry is wrapped up in whatever capacities are required to bring this performance off."⁸³

⁷⁹T. Burge (1974), p. 212.

⁸⁰The strategy of conditionalization goes back to S. Weinstein (1974) and T. Burge (1974).

⁸¹Vgl. J. Higginbotham (198?) and (1994)

⁸²J. Higginbotham (1994), p.93.

⁸³ibid., p. 93.

The end of Higginbotham’s reply suggests that he thinks that conditional T-sentences for indexical utterances somehow “involve” the completing sense of the indexical. The sense is not involved in the right-hand side of the consequence of the T-sentence, which is supposed to give us the conditions under which an utterance of the sentence under consideration is true. But the antecedent involves the completing sense of the relevant indexical. Now, what Higginbotham means by “involve” or “wrap” is not clear. Let us try to clarify it.

Consider the following example. A speaker S refers with the demonstrative “that” to a dog. This act of the speaker is based on the exercise of certain abilities, in this case the capacity to discriminate objects visually. The exercise of this ability, the visual discrimination of the dog, is a way of thinking of the dog. In Higginbotham’s opinion the connection between this way of thinking of an object and the use of the demonstrative seems to be that the competent referential use of “that” requires that one thinks in this specific way of the object one wants to refer to. A competent use of an expressions is for Higginbotham a use with knowledge of the conditions under which it is understood.⁸⁴

The conditional view of the truth conditions for indexical sentences locates the knowledge someone must have for understanding an indexical utterance not in the T-sentence itself, but in the condition prefixed to the T-sentence. The instantiation problem seems therefore not to arise. We can instantiate the generic T-sentence any way we want, as long we instantiate it with singular terms with the right reference.

Higginbotham puts forth than a claim about understanding very similar to Evans’ view:

“You say on day d , ‘Tomorrow is July 4’. Understanding of this utterance, and the competent production of it (that is, production of it with knowledge of the conditions under which it is understood) requires that one think of the day d' (that is, $d + 1$) as the day following $d \dots$ ”⁸⁵

To discuss this claim I switch to the correlated indicator “today”. Higginbotham would assign the following truth-conditions to an utterance of “Today is fine”:

If u is an utterance of ‘today is fine’ by s , and s refers with the utterance of ‘today’ therein to δ , then

$$u \text{ is true} \leftrightarrow \delta \text{ is fine}$$

At first sight we have made no progress in solving the problem of the ‘right instantiation’: a competent speaker of English could derive from Higginbotham’s conditional T-sentence the following specific sentence for an utterance of “Today is fine”: “What the speaker said is true if he has referred to Christmas Eve with “today” and Christmas Eve is fine.” This surely does not interpret the utterance under consideration.

But this criticism seems to be unjust. The sense of the referring expressions should be somehow contained in the antecedent of the conditional T-sentence. Hig-

⁸⁴Cp. J. Higginbotham (1994) about “Tomorrow is July 4” on p. 94.

⁸⁵J. Higginbotham (1994), p. 94.

ginbotham argues that knowing the truth-conditions of an indexical utterance presupposes knowledge of the sense the indexicals have: the interpreter must determine that the condition prefixed to the T-sentence is fulfilled and in doing this he must think of the relevant object in a particular way. Because of this Fregean sense or mode of presentation is connected to understanding. At the same time the sense of an expression does not enter the *unconditional* truth-conditions.

The plausibility of Higginbotham's truth-conditional conception of sense hinges on the assumption that the interpreter must identify the reference of the indexical mentioned in the antecedent of the T-sentence in special way in order to arrive at the unconditional truth-condition. That this is not so is shown by a variation of an example with which we are already familiar. Imagine me listening to a recording of an utterance of Mr. X saying: "today is fine" on July 4. On the tape is written: "Date of recording: July 4". I don't recognize that the utterance is made on the day I hear it. But given that I understand the inscription on the tape and think that it expresses a true belief I know enough about the date of recording to infer from the conditional truth-condition of the utterance that it is true if and only if July 4 is fine. This is surely not the thought required for understanding the utterance in the strong sense and more importantly I have not arrived at the truth-condition *in the right way*. The example shows that an interpreter can determine that the relevant condition is fulfilled *without* thinking of the reference of the indexicals in the way intuitively required for understanding.

Now there is a second reading of Higginbotham's paper which seems to avoid these problems. Higginbotham writes about an conditional T-sentence like the one we have been discussing:

"[...] the semantics [...] of the representation for "Tomorrow is July 4" does after all include a completing sense of the word "tomorrow", since it is this sense that must be grasped by someone *who refers to a day with that word*."⁸⁶

This passage seems to indicate that not the concept of understanding an indexical utterance is central but that of making an indexical reference. Higginbotham's position seems to be that one cannot refer to x with for example "that" without thinking of x in a special way. This fact, Higginbotham argues, is displayed in the antecedent of the T-sentence.

But can the concept of making a reference with a term really bear the weight Higginbotham is putting on it? Must someone, who uses for instance an indexical referentially, think of the referent of the indexical in a special way, must he be, to use David Kaplan's famous phrase, be "en rapport" with the object indicated?

Here is a counterexample. I can, even if I am blind-folded, point out an object and say "This is a dog". In this situation I demonstratively refer to an object, but I cannot fully understand my utterance. This is shown by the fact that any addressee must *perceive* the object I have demonstratively referred to in order to fully understand my utterance. Therefore the concept of making a reference is

⁸⁶J. Higginbotham (1994), p. 95.

not intimately connected to the concept of a mode of presentation. The example illustrates

“the possibility of exploiting an established device of reference in order to manifest the intention to be understood in a certain way, when one is not in a position to understand one’s own words in that way.”⁸⁷

Is the example underdescribed? Higginbotham focuses on *competent* indexical reference. Can we say that in the above example of the “blind” demonstrative reference the expression “that” was only *incompetently* used?

No. I use the expression with knowledge of the conditions under which it is understood. I know that the hearer must perceive the object I demonstrated and think that it is a dog in order to fully understand my remark. I know that this is so, and I exploit this knowledge to in my act of reference. I can intend to refer to the object demonstratively although I cannot perceive it because I know under which conditions my utterance will be understood and that others can fulfill these conditions.

11 Sainsbury on ‘Anaphora-Preserving’ Instantiation

The problem we have been discussing up to now is basically this: in a reasonable simple truth-theory we need universally quantified principles which determine the reference of an expression in dependency from contextual factors. The universal quantifier allows us to derive uninterpretative T-sentences, so the theory of truth does not qualify as a theory of sense.

The question is then: how can we place appropriate restrictions on the instantiations of the universal quantifier without mentioning modes of presentation in the theory? Mark Sainsbury tries to answer this question by using a construction which is not part of standard logic. Sainsbury discusses the problems connected with conditional T-sentences like:

If u is an utterance of ‘today is fine’ by s , and s refers with the utterance of ‘today’ therein to δ , then

$$u \text{ is true} \leftrightarrow \delta \text{ is fine}$$

How can we avoid unwanted instantiations? In oratio obliqua we utilize anaphorical devices to make correct reports. To use an example similar to one we already discussed, consider this. If Dr. Lauben said:

(S12) I am wounded.

I can report his utterance in the following way:

⁸⁷G. Evans (1982), p. 171.

(S13) Dr. Lauben said that he (himself) was wounded.

In (S13) the reference of “he (himself)” depends on the grammatical antecedent “Dr. Lauben”. Sainsbury argues that this way of reporting does not just give information to whom Dr. Lauben was referring in his utterance but also in which way he referred.

According to Sainsbury we can solve the instantiation problem if a theory of truth contains expressions which ‘simulate’ this feature of reported speech:

“We can steer through this dilemma by appealing to a special kind of ‘anaphora-preserving’ instantiation. The δ variable, as it occurs to the right hand of the biconditional, can simulate the kind of anaphoric dependence we find in reported speech if we insist that it be replaced by a pronoun anaphorically dependent upon the first instantiation of the variable. (One could introduce a device in the quantification to mark the need for this style of instantiation, which I believe is needed elsewhere as well, for example by the notation ‘ $\forall xFx: \dots (\text{that } F)_x$ ’).”⁸⁸

The conditional T-sentences of the envisaged theory of truth look like this:

$T_I (\forall x) (\forall y) (x \text{ is an utterance of “I am hungry” \& } y \text{ utters } x, \text{ then:}$

$x \text{ is true} \leftrightarrow \text{that person}_y \text{ is hungry.})$

Or this:

$T_T (\forall x) (\forall y) (x \text{ is an utterance of “Today is fine” on } y, \text{ then:}$

$x \text{ is true} \leftrightarrow \text{that day}_y \text{ is fine.})$

I will supplement this T-sentences with a third one in Sainsbury’s style for second person utterances:

$T_Y (\forall x) (\forall y) (\forall z) (x \text{ is an utterance of “You are a fool” \& } y \text{ addresses } z \text{ in uttering } x, \text{ then:}$

$x \text{ is true} \leftrightarrow \text{that person}_z \text{ is fool.})$

Phrases like “that person_y” and “that day_y” are variables. The subscript shows which phrase of the instantiated T-sentence is their grammatical antecedent.

Do conditional T-sentences like these interpret individual utterances, i.e. do they specify what is said by the utterance? Of course not. For instance an conditional assertion like T_I doesn’t assert something about an individual utterance of mine. But given such a conditional assertion we should be able to infer an unconditional assertion that interprets an utterance. Sainsbury’s rule is: “*instantiate, detach, and finally use what is to the right of the biconditional to interpret u, or to report what was said by u.*”⁸⁹

Let us see, how this proposal works. I utter on March 26th 12 o’clock the sentence “I am hungry”. Let us call this utterance u_1 . From (1)

⁸⁸M. Sainsbury (Ms)

⁸⁹M. Sainsbury (1996), p. 5.

(1) u_1 is an utterance of “I am hungry” and Mark Textor utters u_1 .

and T_I we should now be able to infer a sentence that interprets my utterance. But on Sainsbury’s account we can’t do that. From (1) and T_I we get:

(2) u_1 is true \leftrightarrow that person $_{M.T.}$ is hungry.

This is *not* the conclusion we want and need. We need a sentence that expresses a complete proposition to interpret u_1 and (2) doesn’t express a complete proposition, because the quasi-indicator has lost his grammatical antecedent.

Compare:

(S14) Bertrand knows that $2 + 2 = 4$.

(S15) $2 + 2 = 4$.

And:

(S16) Jeff knows that he* is the new editor of *Noûs*.

(S17) He* is the new editor of *Noûs*.

We need an inference like the one from (S14) to (S15), but we only have got one like the one from (S16) to (S17).

Perhaps there is way to fix this problem, if we are generous in the interpretation of the quasi-indicators: we can think of the reference of the quasi-indicator “that person $_{M.T.}$ ” in the conclusion still anaphorically dependent on the grammatical antecedent “Mark Textor” in the premises. This idea gains support from the fact that analogous referential chains seem to exist in natural language discourse:

John was convinced that he* is the editor of *Noûs*. But then he realized that he* wasn’t.

If the derived T-sentence stays in “anaphorical contact” with the premises, Sainsbury’s idea seems to work. It just requires - so to say - semi-detachability. Let us see, if this *really* works.

The central question for our purposes is: Is someone who knows a theory of truth like the one Sainsbury favours placed thereby in a position to understand indexical utterances in the strong sense we are interested in? We would not say that S had understood an indexical utterance, if S did not identify the referents of indexicals in a specific way. Does Sainsbury’s account capture this dimension of our intuitive concept of understanding?

The answer seems to be *No*. Suppose I hear an utterance of “You are a fool”. Call this utterance u_2 . I know that this is an utterance of an English sentence and I know that it is made by P.T. and addressed to M.T. Now I use Sainsbury’s machinery to interpret and derive a T-sentence. If everything works out allright and the quasi-indicators pick up the right references, I should finally know that the following is true of u_2 :

(T(u_2)) P.T.'s utterance u_2 addressed to M.T. is true \leftrightarrow that person_{M.T.} is a fool.

According to Sainsbury's theory this knowledge enables me to interpret the relevant utterance. But I think it does not. I think it is uncontroversial to say: in the strong sense I understand u_2 only if I realize that I am the one addressed, i.e. only if I think the thought I would express with the words: "I am a fool, that's what this guy is saying". But I could come to know what was said by (T(u_2)) without realizing that I am M.T. and consequently without realizing that I am addressed. So this knowledge does not enable me to interpret the relevant utterance. I could know all the truth-theoretic facts about u_2 without thereby grasping the thought required for understanding it.

If I am right about this, the introduction of quasi-indicators in the meta-language doesn't yield a theory of truth that enables us to interpret indexical utterances in the sense our linguistic intuitions seem to require. An explanation of this failure is easily given: Our linguistic ability is a systematic and general competence. A theoretical representation of this competence must conceive of the T-sentences which shall interpret individual utterances as derived from a finite set of universally quantified principles. But such principles will license the derivation of T-sentences which are not apt for the interpretation of individual utterances. The theory of truth itself cannot discriminate between "right" and "wrong" T-sentences.

References

- Bennett, J., (1985) "Critical Notice: *Inquiries into Truth and Interpretation*", in: *Mind*.
Boër, S. E., Lycan, W., (1980) "Who, Me?", in: *Philosophical Review* 89.
Brandom, R., *Making It Explicit* Harvard 1994
Burge, T., (1974) "Demonstrative constructions, reference, and truth", in: *Journal of Philosophy* 71 (1974).
— (1987) "Truth and Singular Terms", revised version in: L.B. Puntel, *Der Wahrheitsbegriff* Darmstadt.
Castañeda, H.- N., (1966) "'He': A Study in the Logic of Self-consciousness", in: *Ratio*, VIII.
— (1967) "Indicators and Quasi-Indicators", in: *American Philosophical Quarterly* 4.
— (1968) "On the Logic of Attributions of Self-Knowledge to Others", in: *Journal of Philosophy* LXV (1968).
— (1980) "Reference, Reality and Perceptual Fields", in: *Proceedings and Addresses of the APA* 53.
Davidson, D., (1984) *Inquiries into Truth and Interpretation* Oxford.
— (1990) "The Structure and Content of Truth", in: *Journal of Philosophy*.
— (1993) "Reply to Wolfgang Kühne", in: R. Stoecker, *Reflecting Davidson*, Hawthorne 1993.
Dummett, M., (1981) *Frege – Philosophy of Language*, Second Edition, London.
— (1984) "An Unsuccessful Dig", in: C. Wright (ed.) *Frege: Tradition & Influence* Oxford 1984.
Evans, G., (1981) "Understanding Demonstratives", reprinted in: G. Evans (1985a) *Collected Papers* (ed. A. Phillips) Oxford.

- (1982) *Varieties of Reference* (ed. J. McDowell) Oxford.
- (1985b) “Does Tense Logic rest upon a Mistake”, in: G. Evans (1985a).
- Fodor J., Lepore, E., (1992) *Holism: A Shopper’s Guide*. Oxford/Cambridge, Mass.
- Foster, J., (1976) “Meaning and Truth Theory”, in: *Truth and Meaning* (ed. Gareth Evans and John McDowell), Oxford.
- Frege, G., (1893) *Grundgesetze der Arithmetik Bd. I*, Jena.
- Higginbotham, J., (198?) “Contexts, Models, and Meanings”, in: R. Kempson (ed.),
- (1994) “Priorities in the Philosophy of Thought”, in: *Proceedings of the Aristotelian Society Supp.-Vol.*
- Kaplan, D., (1989a) “Demonstratives”, in: J. Almog, J. Perry, H. Wettstein eds. *Themes from Kaplan* Oxford.
- (1989b) “Afterthoughts”, in: *Themes from Kaplan*.
- Loar, B., (1976) “The Semantics of Singular Terms”, in: *Philosophical Studies* 30.
- McDowell, J. (1976) “Truth-Conditions, Bivalence and Verificationism”, in: *Truth and Meaning* (ed. J. McDowell & G. Evans) Oxford 1976,
- (1980) “On the Sense and Reference of a Proper Name”, in: *Mind* 86 (1977). Reprinted in: *Reference, Truth and Reality* (ed. Mark Platts) London 1980.
- Perry, J., (1977) “Frege on Demonstratives”, in: *Philosophical Review* 86.
- (1983) “Castañeda on He and I” reprinted in: J. Perry, *The Problem of the Essential Indexical*.
- (unpublished) “Understanding Evans on Demonstratives”. Unpublished Manuscript.
- (1993) *The Problem of the Essential Indexical* Oxford.
- (1995) “Unfair to Words — Wittgenstein’s Builders and Davidson’s Sentences”, *Proceedings and Addresses of the APA* 67.
- (forthcoming) “Indexicals and Demonstratives”, forthcoming in: Robert Hale & Crispin Wright, eds. *Companion to the Philosophy of Language*.
- Prior, A. N., (1970) “I”, in: *The Jowett Papers 1969-1970* (ed. A. Pineau) Oxford.
- Quine, W.V.O., (1977) “Review of of Evans/McDowell *Truth and Meaning*”, in: *Journal of Philosophy* 74.
- Rumfitt, I., (1993) “Content and Context – The Paratactic Theory Revisited and Revised”, in: *Mind* 102.
- Sainsbury, R. M., (1995) “Philosophical Logic”, in: A.C. Grayling, *Philosophy – A Guide through the Subject* Oxford.
- (1996) “Fregean Indexicals”, in: C.A. Stein, M. Textor eds. *Intentional Phenomena in Context* Hamburg 1996.
- “Indexical Involvement and Detachment” (Ms.)
- Tarski, A., (1956) *Logic, Semantics, and Metamathematics* New York/Oxford.
- Taschek W., (1995) “Belief, Substitution, and Logical Structure”, in: *NOUS* 24.
- Weinstein, S., (1974) “Truth and Demonstratives”, in: *Noûs* 8.
- Wiggins, D., (1992) “Meaning, Truth-Conditions, Proposition: Frege’s Doctrine of Sense Retrieved, Resumed and Redeployed in the Light of Certain Recent Criticisms”, in: *Dialectica* 46.