

Extensionality, Indirect Contexts, and Frege's Hierarchy*

Nicholas Koziolk†

dialectica 70.3 (2016)‡

*I presented a very early version of this paper at the Northwest Philosophy Conference at Pacific University in October 2013, and again at the Northern Graduate Philosophy Conference at Northern Illinois University in November 2013. I also presented related material at various workshops at the University of Chicago. I would like to thank audiences at these talks for their questions and comments. I would also like to thank Michael Kremer, Josef Stern, Malte Willer, and Marko Malink, as well as several anonymous referees, for invaluable feedback on a number of more recent versions.

†Email: nkoziolk@gmail.com | Website: <http://nkoziolk.net>

‡This is the pre-peer reviewed version of the paper. The published version is available at <http://onlinelibrary.wiley.com/doi/10.1111/1746-8361.12146>.

Abstract

It is well known that Frege was an extensionalist, in the following sense: he held that the truth-value of a sentence is always a function only of (its syntax and) the references (the “extensions”) of its parts. One consequence of this view is that expressions occurring in certain linguistic contexts—for example, the that-clauses of propositional attitude ascriptions—do not have their usual references, but refer instead to what are usually their senses. But although a number of philosophers have objected to this result, no one has yet attempted to see what happens to Frege’s views—and, in particular, to his theory of sense and reference—if his extensionalism is abandoned while his other views are (so far as possible) maintained. This paper thus does two things. First, it clears the way for such an attempt, by arguing that recent defenses of Frege’s extensionalism—by Tyler Burge, Saul Kripke, Terence Parsons, and Christopher Peacocke—fall short. Second, it sketches a positive proposal for a non-extensionalist application of Frege’s theory of sense and reference to sentences of indirect discourse and ascriptions of propositional attitude.

Introduction

The essentials of Frege's theory of indirect discourse and propositional attitude ascriptions—his theory of indirect contexts, as I will call it—can be gleaned from the following short passage:

In reported speech one talks about the sense, e.g., of another person's remarks. It is quite clear that in this way of speaking words do not have their customary reference but designate what is usually their sense. In order to have a short expression, we will say: In reported speech, words are used *indirectly* or have their *indirect* reference. We distinguish accordingly the *customary* from the *indirect* reference of a word; and its *customary* sense from its *indirect* sense. The indirect reference of a word is accordingly its customary sense. (Frege 1892: 59, his emphases)

These remarks suggest a certain general view about the senses and references of expressions in indirect contexts. For simplicity, I will focus on just one paradigmatic example of an indirect context, namely the linguistic context occupied by '*p*' in sentences of the form '*S* believes that *p*' (where '*S*' is schematic for the name of a person and '*p*' is schematic for a sentence). Using sentences of this form as an example (which I will continue to do for the remainder of the paper), the view suggested by the quoted passage can be stated in the following way.

Consider the sequence of sentences:

- (1) Socrates is wise,
- (2) Plato believes that Socrates is wise,
- (3) Aristotle believes that Plato believes that Socrates is wise,

and so on. In (1), on Frege's view, 'Socrates is wise' refers to a truth-value and expresses, as its sense, the thought that Socrates is wise. But in (2), he tells us, it *refers* to its customary sense, the thought that Socrates is wise. Moreover, if we follow his suggestion and

distinguish an expression's customary sense from its indirect sense, we must conclude that its *sense* in (2) is *not* the thought that Socrates is wise. It is, rather, a mode of presentation or way of thinking of that thought (Frege 1892: 57).

Similar conclusions about the sense and reference of 'Socrates is wise' in (3) can be derived on the assumption that indirect contexts iterate, i.e., that, because, in (3), 'Socrates is wise' is embedded in the context 'Plato believes that ...', which is itself embedded in the context 'Aristotle believes that ...', the context it occupies there is, as it were, *doubly* indirect. It follows that the reference of 'Socrates is wise' in (3) will be, not the thought that Socrates is wise, but rather the mode of presentation or way of thinking of that thought that it expresses in (2). And its sense in (3) will be a mode of presentation of that mode of presentation.¹

This same reasoning will apply to every sentence in the imagined sequence. The suggested view is thus that, in each sentence in the sequence, the embedded sentence 'Socrates is wise' has a different reference and expresses a different sense. Specifically, the view is that the reference of that sentence in (2) is the sense it expresses in (1), its reference in (3) is the sense it expresses in (2), and so on, where its sense is in each case a way of thinking of its reference. Parallel conclusions follow for each of the significant components of 'Socrates is wise' (i.e., 'Socrates' and 'is wise'). The result is *Frege's hierarchy*: each expression of a language that contains indirect contexts must be associated with an infinite number of distinct senses and references, one for each possible iteration of the expression 'believes that'.

Both Carnap (1947: 130) and Davidson (1965: 14–15) famously objected that any language that is answerable to Frege's theory would have to be unlearnable. For, in order to understand any given expression, one would need to grasp an infinite number of senses,

¹Frege does not explicitly say, in the quoted passage, that indirect contexts iterate, and so the passage does not strictly imply the conclusion of the last paragraph, which is why I have been saying only that the passage *suggests* the view in question. He does, however, appear to commit himself to the view in a famous letter to Russell (Frege 1980: 154).

something that (they contend) is patently impossible. Kripke (2008: 257–58) suggests, however, that the essence of this objection does not rest on the fact that the hierarchy is infinite. On his view, the crucial point is that, even if we consider only singly indirect contexts, Frege's theory already seems to make implausible demands of teachers and learners of languages that contain indirect contexts. As he explains: "Ordinarily, we should think that to teach English to a foreigner one simply teaches some grammar, the vocabulary, and what all the words mean, and then we are through." So, for example, if we teach someone (i) the expressions 'Socrates', 'Plato', 'is wise', and 'believes that', (ii) what each of these expressions means, and (iii) the appropriate English grammar, then she should be able to understand not only (1), but also (2) (and, indeed, if she knows the name 'Aristotle' and what it means, even (3)), without needing to be taught anything further. But, according to Frege's theory, she *should* need to be taught something further: namely, the indirect senses of 'Socrates' and 'is wise'. The fact that we *don't* in fact need to teach her any such thing thus seems to give the lie to Frege's theory.

Recently, however, a number of philosophers—including Kripke himself, in the paper just mentioned (the others are Burge (2004), Parsons (2009), and Peacocke (2009))—have taken it upon themselves to show that, appearances notwithstanding, Frege's theory does not really conflict with the above facts about the transmission of languages that contain indirect contexts. They argue, in particular, that there are general rules that determine an expression's singly indirect sense and reference on the basis of its customary sense, its doubly indirect sense and reference on the basis of its singly indirect sense, and so on, all the way up the hierarchy. All that is needed to save Frege, then, is to add these rules to his theory of sense and reference. The language learner does, of course, need to learn these rules. But they are learned in learning the sense of (for example) the expression 'believes that'. So Frege's theory fits the facts after all.

My objection to this attempt to save Frege will take an unusual, and perhaps surprising, form. For I think that Peacocke (2009), in particular, provides a plausible account of the unique "canonical"

ways of thinking of senses that are meant to serve as the indirect senses posited by Frege's theory of indirect contexts. In other words, I think that we self-conscious thinkers, at least, *are* in a position to grasp the ways of thinking of senses, ways of thinking of ways of thinking of senses, and so on, that make up Frege's hierarchy. Nonetheless, I think that it is a mistake *to introduce the notion of indirect sense—and so the hierarchy itself—into the theory of indirect contexts*, i.e., into the theories of indirect discourse and ascriptions of propositional attitude. In short: there (probably) is a hierarchy of higher-order senses associated with each expression; but grasp of those senses is not required for grasp of sentences of (for example) the form 'S believes that *p*'.

I proceed as follows. In §1, I explain the basic challenge Frege's defenders need to meet. Then, in §2, I argue that the specific proposals developed by Burge (§2.1), Kripke (§2.2), and Peacocke (§2.3) all fail to meet this challenge.² The result of that discussion will be that, in certain important respects, the introduction of indirect senses into the theory of indirect contexts is simply gratuitous. So, finally, in §3, I propose a different way of defending Frege, one that involves not merely supplementing his view, but actually revising it. I argue there that, by altering Frege's formal framework—in particular, by abandoning his extensionalism—we can avoid the need to introduce indirect senses and references at all.

1 The challenge

On Frege's theory of indirect contexts, it is a simple matter to state a principle that determines the indirect *reference* of an expression on the basis of its customary sense, because the reference of an expression in an indirect context just *is* its customary sense. More generally:

²Parsons (2009), for his part, doesn't actually propose a hierarchy-climbing rule of the kind in question. Instead, he *assumes* that such a rule is available, and shows that its addition to Frege's theory would suffice to meet the challenge. The resulting view is, however, subject to the criticism I make to Burge's view.

IND-REF: The reference of an expression in an indirect context of degree n is the sense it expresses in an indirect context of degree $n - 1$.

(An indirect context of degree zero is a transparent context, one in which an expression expresses its customary sense.) IND-REF straightforwardly determines the singly indirect reference of an expression on the basis of its customary sense, its doubly indirect reference on the basis of its singly indirect sense, and so on. The problem, however, is that IND-REF does *not* determine the expression's indirect *sense*. We are thus in need of an additional rule.

Still, it is reasonable to assume that our knowledge of the indirect reference of an expression will be useful in determining its indirect sense. Indeed, it is typically assumed that the concepts of sense and reference are, in general, related by the following principle:

Frege's Principle: The sense of an expression is a way of thinking of its reference.³

And, assuming (per IND-REF) that the indirect reference of an expression is its customary sense, Frege's Principle tells us something, at least, about its indirect sense: the indirect sense of an expression is a way of thinking of its customary sense. Unfortunately, though, we need to know more than that the indirect sense of an expression is *a* way of thinking of its customary sense; we need to know precisely *which* such way of thinking it is. And Frege's Principle cannot tell us that.

Burge (2004) has suggested, however—and he is followed here by Kripke (2008), Parsons (2009), and Peacocke (2009)—that we can answer this remaining question by noting that the singly indirect

³A notable exception is Dummett (1973: 267–268), who suggests that, in indirect contexts, an expression may both express and refer to its customary sense. On such a view, indirect contexts must be treated as exceptions to Frege's Principle. Otherwise, we would be forced to say that the sense of the name 'Socrates', for example, is a way of thinking both of Socrates and of itself. Incidentally, it was Carnap who first recognized the possibility of such a view, though he does not pursue the possibility at any length; see his (1947: 129).

sense of an expression is, plausibly, *the canonical* way of thinking of its customary sense (and, similarly, that its doubly indirect sense is the canonical way of thinking of its singly indirect sense, and so on). We can thus use the general idea of a *canonical way of thinking* to answer our remaining question. For the answer, on this view, is given by the general principle:

IND-SENSE: The sense of an expression in an indirect context of degree n is the canonical way of thinking of the sense it expresses in an indirect context of degree $n - 1$.

But what is a canonical way of thinking?

The notion of a canonical way of thinking is derived from the notion of a canonical (or standard) *name*.⁴ The typical examples of canonical names are numerals, quotation-expressions, and expressions occurring in indirect contexts, which are said to be canonical names of numbers, expressions (that is, expression-types), and senses, respectively.⁵ The canonical way of thinking of something is then simply the sense of its canonical name. So, for example, if the canonical name of the number eighteen is the numeral '18', then the canonical way of thinking of the number eighteen is the sense of the numeral '18', whatever that is.

The central thought behind the notion of a canonical way of thinking is that the way of thinking of an entity that is associated with its canonical name is in some sense unique. It is unlike the ways of thinking of the same entity that are associated with other linguistic devices. Thus, to take an example given by Parsons (2009: 56), the expression ' 9×2 ' is associated with a certain way of thinking of the number eighteen, but that way of thinking is, in a certain sense, less direct than the one associated with the numeral '18'. It is in virtue of its peculiar intimacy with the number eighteen that

⁴See Burge 2004: 170–171, Parsons 2009: 56, Peacocke 2009: 164–165, Kripke 2008: 268–269, and Kaplan 1968: 194–197.

⁵It is generally assumed—though Kripke (2008: 260) seems to be an exception—that there are canonical ways of thinking only of abstract entities. Peacocke (2009: 164) even suggests that there may be canonical ways of thinking only of a certain special *kind* of abstract entity.

the numeral '18' is said to be a canonical or standard name of that number and to express the canonical way of thinking of it, a way of thinking that plays a unique role in our thought about it.

Kripke, building on this idea, says (though he does not employ the language of canonicity) that what is special about the senses of canonical names is that they are *immediately revelatory*, where "a sense is revelatory of its referent if one can figure out from the sense alone what the referent is" (2008: 259) and is *immediately revelatory* if, in addition, "no calculation is required to figure out its referent" (2008: 261). So, for example, '9 × 2' would be revelatory, but not immediately revelatory, whereas '18' would be immediately revelatory. Similarly, Burge holds that the senses of expressions occurring in indirect contexts are special (among senses) in that they must satisfy the Principle for Canonical Names of Senses, which says that "[t]he canonical name of a sense can be understood only if the sense that it names is understood" (2004: 174).⁶ These are both ways of spelling out the idea that a canonical way of thinking is peculiarly intimately related to its object.

We thus know quite a bit about canonical ways of thinking of senses. Assume, for example, that Σ is the customary sense of 'Socrates' and that the canonical way of thinking of Σ is $C(\Sigma)$, where $C(x)$ is a function—Burge (2004: 172) calls it "the canonical sense function"—from senses to canonical ways of thinking of them. Then we know that you can figure out from $C(\Sigma)$ alone, and without calculation, that it's a way of thinking of Σ . And we know that anyone who understands (has a grasp of) $C(\Sigma)$ also understands (has a grasp of) Σ .

The problem is that neither of these features of canonical ways of thinking of senses—neither the fact that they are immediately revelatory, nor the fact that they are subject to the Principle for

⁶Burge also accepts the Stronger Principle for Canonical Names of Senses: "To think the sense of a canonical name of a sense, one must simultaneously think the lowest-level (ultimately, customary) sense in the downward hierarchy associated with the canonical name" (2004: 174). Cf. Peacocke 2009: 166: "it is a distinctive feature of the canonical concept of F that a thinker cannot possess that canonical concept unless he has at least a partial grasp of the concept F itself."

Canonical Names of Senses—tells us what we needed to know. For these features both concern the “downward” relation between sense and reference. That is, they treat of the special way in which the sense of a canonical name determines its reference. But we needed to know about the “upward” relation between reference and sense: we needed to know how the *reference* of a canonical name of a sense determines its *sense*. Or, more precisely, and continuing the example of the last paragraph: we needed to know how someone with a grasp of Σ acquires a grasp of $C(\Sigma)$.

To see the problem more clearly, consider the following consequences of Kripke's remarks about language learning, which I quoted in the Introduction. First, if you understand both

- (1) Socrates is wise

and

- (4) Aristotle is a philosopher,

then you can also understand

- (5) Aristotle is wise.

For you understand each of its significant parts (i.e., you know the relevant vocabulary, and you know what all the words mean), and you understand the way in which they are combined (i.e., you know the grammar), and that is all there is to understanding a sentence. Similarly, if you understand both

- (2) Plato believes that Socrates is wise

and

- (4) Aristotle is a philosopher,

then you can also understand

- (6) Plato believes that Aristotle is wise.

For, again, you understand each of its significant parts—'Plato', 'believes that', 'Aristotle', and 'is wise'—and you understand the way in which they are there combined—for you understand (2), in which different parts are combined in just that way—and that is all there is to understanding a sentence.

The question is how we can entitle ourselves to this second consequence, if we accept Frege's theory of indirect contexts, on which the name 'Aristotle', as it occurs in (6), expresses a different sense than it does in (4). For, while anyone who understands (4) will have a grasp of the *customary* sense of 'Aristotle', what is required to understand (6) is a grasp of its *indirect* sense. And a grasp of the *indirect* sense of 'Aristotle' is not required—or, apparently, (and more importantly) entailed—by an understanding of (4). So, in order to explain the second consequence adduced in the preceding paragraph, we need to show that an understanding of both (2) and (4) suffices for a grasp of the *indirect* sense of 'Aristotle', and thus for an understanding of (6).

The suggestion under consideration—the suggestion proposed, in effect, by Frege's contemporary defenders—is that you have, or can easily enough acquire, a grasp of the indirect sense of 'Aristotle' if (i) you have a grasp of its *customary* sense and (ii) you understand the expression 'believes that'. What we have yet to see, however, is whether there is any way to make this suggestion work. Since Burge, Kripke, and Peacocke each proposes a different way of developing this suggestion into an account of canonical ways of thinking of senses and our grasp of them, I will, in the next section, consider their views separately, taking them in the order just indicated.

2 Indirect contexts and Frege's hierarchy

2.1 Burge

At its core, Burge's account of our grasp of indirect senses is fairly simple. Here is the basic idea:

We understand the indirect sense, and “know what it is” in the sense of comprehending it, if we can both use the that-clause-forming expression (‘that’) and understand the sense of the sentential expression to which it applies. Nothing more is needed. (2004: 172)

More generally, Burge holds that grasp of the sense of an expression in an indirect context of degree n consists entirely in (i) grasp of the expression's customary sense, (ii) grasp of the sense of ‘that’, and (iii) knowledge that the expression in question occurs in an indirect context of degree n .⁷

My objection to Burge's account starts from the observation that it seems to provide no answer whatsoever to the challenge posed in §1. That challenge, again, is to say *how* grasp of an expression's customary sense combines with an understanding of the expression ‘believes that’ to provide an understanding of the expression's indirect sense. Burge's answer to the question “How does it suffice . . . ?” seems to be that *it just does* (“Nothing more is needed”): if you have a grasp of the customary sense and a grasp of the sense of ‘that’, then you have a grasp of the indirect sense.

Burge himself doesn't seem to think that there's a problem here, and in fact he repeatedly draws attention to the very feature of his account presently at issue. For example, after presenting and then summarizing his view, he says that “[m]astering the hierarchy of canonical names *requires and involves* nothing more” (2004: 177, my emphasis) than a grasp of the customary sense and a grasp of the sense of ‘that’. And, in describing the difference between a sense and the canonical way of thinking of it, he says:

These ways of thinking, at different levels of canonical sense, differ [...] *entirely* in the level of embedding of attribution. Thus the differences *are needed by, but are exhausted by*, the levels of logical attribution. (2004: 179, my emphases)

⁷On condition (iii), see Burge 2004: 177–178, where he appeals to “a capacity to keep track of levels of iteration or embedding.”

It seems to me, though, that these differences between senses and canonical ways of thinking of them should be *required* by our theory of indirect contexts only if they are *not* exhausted by the application of the word 'that' (i.e., the "levels of logical attribution"). In other words, it seems to me that an account of our understanding of sentences containing indirect contexts will require the introduction of indirect senses only if our understanding of those sentences *cannot* be explained solely by appeal to our grasp of customary senses and our grasp of the sense of 'that'.

Burge's inability or refusal to say more than he does in answer to the above "how" question thus threatens to undermine the view that expressions in indirect contexts even *have* indirect senses. For it makes it hard to see any relevant difference between sentences that contain indirect contexts and sentences that don't. To understand the sentence 'Socrates is wise', for example, you need only have a grasp of the customary sense of 'Socrates' and the customary sense of 'is wise' and know how the two are combined in the whole. Similarly, on Burge's view, to understand the sentence 'Plato believes that Socrates is wise', you need only have a grasp of the customary senses of 'Plato', 'believes', 'that', 'Socrates', and 'is wise', and know how they are combined in the whole. The question for Burge might then be put this way: in what sense do you *also* need a grasp of the indirect senses of 'Socrates' and 'is wise', if such a grasp is already guaranteed by, or contained in, your grasp of the customary senses of 'that', 'Socrates', and 'is wise'? (And the difference can't just be in the embedding, since we don't need to introduce indirect senses to explain our grasp of the senses of iterable functional expressions like '*x*'s father' or of iterable operational expressions like 'or'.)

It thus seems that, for purposes of providing an account of our understanding of sentences containing indirect contexts, appeal to indirect senses is wholly gratuitous. Importantly, this fact is implicitly acknowledged not only by Burge, but also by Parsons, who says that we can see the hierarchy of senses "as a kind of epiphenomenon of the simple part at the basis of the hierarchy" (2009: 57).⁸ None

⁸The present objection to Burge's view thus applies also to Parsons's view.

of this, of course, entails straightaway that the hierarchy should be abandoned; there may yet be reasons, after all, for introducing indirect senses. But it does suggest that the motivation for the hierarchy lies entirely in Frege's conviction that an expression in an indirect context *refers* to its customary sense and therefore (given Frege's Principle) cannot *express* it, but must instead express a way of thinking of it.

I will return to this issue in §3. First, however, we need to see whether either Kripke or Peacocke has an account of indirect senses that can respect Kripke's points about language learning without rendering indirect senses epiphenomenal.

2.2 Kripke

Having introduced a version of the principle I stated above as IND-REF,⁹ Kripke says:

Words in an indirect context refer to their ordinary senses. But what is the sense in the indirect context? Here we should surely say that the rule [IND-REF] itself, applied to each indirect case, gives the indirect sense. For we have conceded that it determines the reference in each particular case, and whatever determines a reference is a Fregean sense. This consideration can then be applied iteratively, and we appear to have a theory of the entire hierarchy. (2008: 268–269)

The idea, I take it, is that, if we know what sense a word would express in an indirect context of degree n , the rule IND-REF will tell us what sense it would express in an indirect context of degree $n + 1$. An apparent problem with this suggestion is that, as I pointed out in §1, IND-REF itself says nothing at all about the indirect *sense* of any expression. According to Kripke, however,

⁹IND-REF is a generalization of Kripke's rule (β): "When words appear in indirect contexts, that is, 'says that', 'believes that', and so on, they refer to their senses in the clause following the 'that'" (2008: 268).

one does not, when using an expression or introducing an expression, have to specify two things, its reference and its sense. Once one specifies the reference one has specified the sense. (2008: 255)

On Kripke's view, then, the function of IND-REF is indeed to specify an expression's indirect reference. But since to specify its indirect reference *is* to specify its indirect sense, IND-REF also specifies its indirect sense.

The crucial claim here is that to specify an expression's reference is to specify its sense. The example Kripke gives in support of this claim, which comes from Frege's *Grundgesetze der Arithmetik* (1893: sections 31 and 32), is that of the specification of the sense of a *sentence* in terms of a certain sort of specification of its reference, i.e., its truth-value. Thus, take the sentence 'Socrates is wise'. The truth-value of this sentence, on Frege's view, can be described as the result of applying the function $Wise(x)$ to the argument *Socrates*, which result we can write: $Wise(Socrates)$. Now, this specification of the relevant truth-value also, in effect, gives the truth-conditions of the sentence. For to say that 'Socrates is wise' refers to $Wise(Socrates)$ is to say under what conditions the sentence is true: it is true just in case the value of the function $Wise(x)$ for the argument *Socrates* is the True. Finally—and this is the key point for Kripke's purposes—once we have given the truth-conditions of the sentence in this way, we can say what its sense is, namely, the thought that those truth-conditions are fulfilled, or, more fully, the thought that the value of the function $Wise(x)$ for the argument *Socrates* is the True.¹⁰

If these points are to be useful in explaining the hierarchy, however, they must generalize, both to the senses of subsentential, even simple, expressions, and to indirect senses. So, first, consider just the name 'Socrates'. If I say that the name 'Socrates' refers to Socrates, I

¹⁰As a couple of anonymous referees pointed out to me, this last formulation of Frege's idea (if not the idea itself) is questionable: it simply isn't plausible that the thought expressed by the sentence 'Socrates is wise' is about functions, arguments, and truth-values. In §3, however, I provide a way of spelling out Frege's idea—or, at least, something quite close to it—that escapes this objection; see especially §3.5.

have, of course, specified the reference of the name 'Socrates'. And, in doing so, I have myself referred to Socrates in some particular way. The way in which I have referred to Socrates is the sense that Kripke says is specified (because it is expressed) in specifying the reference. That is, the sense of the name 'Socrates' is the way of thinking of Socrates employed in saying that the name 'Socrates' refers to Socrates.

Unfortunately, there is some sloppiness here. Suppose I say that the name 'Hesperus' refers to Venus. Here, again, I have specified the reference of the name 'Hesperus' and, in doing so, I have referred to Hesperus (i.e., to Venus) in some particular way. But, in this case, the way in which I have referred to Hesperus is not the way in which we customarily refer to it in using the name 'Hesperus'. For the names 'Hesperus' and 'Venus', while they have the same reference, differ in sense. So they involve different ways of referring to the same thing. It is thus misleading, at best, for Kripke to say that in specifying an expression's reference I have specified its sense. For I can specify the reference of a particular expression without specifying *its* sense—even if I do, in the relevant sense, specify *a* sense, namely the sense of the expression I use in specifying the mentioned expression's reference.

What we must say, then, is that there is some *particular* way of specifying the reference of an expression that amounts to specifying *its* sense. And so there is: an instance of the disquotational schema

'...' refers to ...

—where the same expression fills both ellipses—will specify not only the reference but also (indirectly, as it were) the sense of the mentioned expression. For example, in

'Socrates' refers to Socrates,

just as the reference of the mentioned expression 'Socrates' is specified (directly) by using the very same expression to refer to it, so the *sense* of the mentioned expression 'Socrates' is specified (*indirectly*) by using the very same expression to *express* it. As Dummett puts it

(in a passage Kripke himself quotes approvingly (2008: 255, note 6)): “we *say* what the reference of a word is, and thereby *show* what its sense is” (Dummett 1973: 227, his emphases). In other words, we know the sense and reference of the mentioned expression by knowing the sense and reference of the used expression; for the sense and reference of the mentioned expression just *are* the sense and reference of the used expression.

The problem is that these remarks cannot be extended in any obvious way to the indirect occurrences of expressions. What we need is a way of completing

‘Socrates’ in ‘Plato believes that Socrates is wise’ refers to ...

where what fills the ellipsis is an expression that refers to the customary sense of ‘Socrates’ and *expresses* its indirect sense. But there are only two ways of filling the ellipsis, and neither of them is satisfactory. First, we might observe that the name ‘Socrates’ itself refers to its customary sense and expresses its indirect sense when it occurs in an indirect context. Unfortunately, however, the second gap in “...’ refers to ...’ is an ordinary, transparent context, where ‘Socrates’ will refer to its customary reference and express its customary sense. So the occurrence of ‘Socrates’ that would fill the ellipsis will have its customary—rather than, as is needed, its indirect—sense and reference. Second, we might introduce a distinct expression to refer to the customary sense of ‘Socrates’ and express its indirect sense. Say we write it ‘Socrates’. Then we have two choices: either the underline expresses a rule that takes us from senses to canonical ways of thinking of them (which is what Kripke’s idea—that to specify the reference of an expression is to specify its sense—was already supposed to do), or the name is semantically primitive. If the former, we are back where we started. If the latter, we are apparently forced to admit that Frege’s theory does, after all, contradict Kripke’s points about language learning.

To be sure, Kripke does make another suggestion, namely, that we interpret Frege as holding that, in every case in which a thinker determines a reference, she is *acquainted* with the way in which

the reference is thus determined, and so is acquainted with the relevant sense.¹¹ To use an expression to refer to its customary sense—as we do when we use it in an indirect context—is thus to be acquainted with its indirect sense, the relevant way of thinking of its customary sense. Without an account of such acquaintance, however, Kripke's account seems to amount to the bare assurance that we are in fact acquainted with higher-order senses. But since Peacocke develops essentially this idea, and does so in a way that makes it quite plausible, I will not press this objection here, but will instead turn immediately to a discussion of Peacocke's view.

2.3 Peacocke

As an account of canonical ways of thinking of senses, Peacocke proposes the following, where $\text{can}(F)$ is the canonical way of thinking (the canonical concept, in Peacocke's way of speaking) of the concept F :

- (*) For an arbitrary concept C to fall under $\text{can}(F)$ is for the fundamental condition for something to fall under C to be that it meets condition R , where R is in fact the fundamental reference rule for the concept F . (2009: 165)

The resulting view, which is grounded in (*) (and iterations of it), is then that “for each concept F , there is a canonical way of thinking $\text{can}(F)$ of the concept F . This canonical way is uniquely fixed [...] by what individuates the concept F itself” (2009: 169).

The general idea here is that the conditions that individuate a *canonical* way of thinking $\text{can}(W)$ of a way of thinking W can be given in terms of the conditions that individuate the way of thinking W itself. In order to apply this general idea, we of course need to have

¹¹Cf. Kripke 2008: 271–272: “Every time we determine a referent, we are introspectively acquainted with how the referent is determined, and that is the corresponding sense. [...] Thus the Fregean hierarchy of indirect senses, doubly indirect senses, and the like is given this way. Each level of the hierarchy is the acquaintance-sense of the previous level.” Mendelsohn (2005: 156), too, suggests that we are acquainted with the senses of expressions we understand.

an account of the individuation-conditions of the ways of thinking that lie at the base of Frege's hierarchy, i.e., those ways of thinking that serve as customary senses. I will assume, however, that these conditions can be given by giving conditions on the *employment* of these ways of thinking.¹² Thus, for example, suppose that Σ is the customary sense of the name 'Socrates'. Then the individuation-conditions of Σ can be given by giving conditions on its employment, as follows: to employ Σ is to think of Socrates *as* Socrates. And, just as we can explain individuation-conditions in terms of employment-conditions, we can explain *possession*-conditions in terms of employment-conditions: thus, to *possess* Σ is to be *disposed* to think, or *capable* of thinking, of Socrates as Socrates—that is, it is to be disposed to employ, or capable of employing, Σ .

We can then provide an account of $\text{can}(\Sigma)$ that parallels Peacocke's (*), and that gives the employment-conditions—and so, as before, the individuation- and possession-conditions—of $\text{can}(\Sigma)$ in terms of the employment-conditions of Σ itself:

- (#) To employ $\text{can}(\Sigma)$ is to think of Σ as that way of thinking to employ which is to think of Socrates as Socrates.

We can thus say, paraphrasing Peacocke, that, for each way of thinking W , there is a canonical way of thinking $\text{can}(W)$ of the way of thinking W . This canonical way is uniquely fixed by what individuates the way of thinking W itself (namely, its employment-conditions).

Now, it might seem that, on the account illustrated by (#), employing a canonical way of thinking $\text{can}(W)$ of a way of thinking W involves making explicit to yourself the employment-conditions on W . And such a view would be too demanding: employing the sense of the sentence 'Plato believes that Socrates is wise'—say, judging that Plato believes that Socrates is wise—does *not* require explicit

¹²This view of ways of thinking is the one that I prefer, but it is also suggested by Peacocke himself, for example when he says that his account is an account of "what it is to be *employing* $\text{can}(F)$ in thought" (2009: 165, my emphasis). It is also an important part of the theory of concepts he develops in Peacocke 1992.

knowledge of the employment-conditions of the sense of the embedded sentence 'Socrates is wise'. Such a view, in other words, would over-intellectualize the requirements on using and understanding others' uses of sentences like 'Plato believes that Socrates is wise'. For using and understanding others' uses of such sentences would require that you explicitly articulate to yourself principles like the above principle for Σ , i.e., that to employ Σ is to think of Socrates as Socrates. And using and understanding others' uses of sentences like 'Aristotle believes that Plato believes that Socrates is wise' would require that you explicitly articulate to yourself principles like (#) itself. And so on. For these reasons, Peacocke suggests that the knowledge of employment-conditions in question—the knowledge required by employment of the relevant way of thinking—is not explicit, but merely *tacit* (2009: 165–166).

My objection to Peacocke's account is that his retreat to merely tacit knowledge doesn't work. Specifically, the principle (#)—or, more generally, (*)—actually *requires* that the knowledge in question be explicit, and not merely tacit.

Here's why. If you have a grasp of the sense of

- (2) Plato believes that Socrates is wise,

then, according to Frege's theory of indirect contexts, you have a grasp of $\text{can}(\Sigma)$, the indirect sense of 'Socrates'. According to Peacocke, you thus have tacit knowledge that Σ is that way of thinking to employ which is to think of Socrates as Socrates. So far, so good: it is plausible that grasp of $\text{can}(\Sigma)$ requires tacit knowledge of (#).¹³ But here's the problem: we need to distinguish mere grasp of a way of thinking from actual *employment* of that way of thinking (just as I distinguished employment- and possession-conditions, above). And, as (#) itself tells us, to *employ* $\text{can}(\Sigma)$, as opposed to merely grasping or possessing it, is *actually to think of* Σ as that way of thinking to

¹³Though there may be reason to doubt this requirement, if young children and (non-rational) animals can represent others' thoughts, since young children and (non-rational) animals presumably do not have even tacit knowledge of principles like (#).

employ which is to think of Socrates as Socrates. So judging that Plato believes that Socrates is wise, for example, involves *actually thinking of* Σ as that way of thinking to employ which is to think of Socrates as Socrates. In other words, it involves *explicit* knowledge of the employment-conditions of Σ .

So the initial objection stands: Peacocke's account is too demanding. For, while it is plausible that mere possession of $\text{can}(\Sigma)$ involves only tacit knowledge of (#), employment of $\text{can}(\Sigma)$ must involve explicit knowledge of (#). To think otherwise is to violate the distinction between mere possession of a way of thinking and employment of it.

Importantly, however, the objection here is an objection only to the conjunction of two distinct claims: (i) that to employ $\text{can}(\Sigma)$ is to have explicit knowledge of the employment-conditions of Σ and (ii) that $\text{can}(\Sigma)$ is the indirect sense of 'Socrates'. For it is only the conjunction of these two claims that entails that the use of the sentence 'Plato believes that Socrates is wise' involves explicit knowledge of (#), and it is only this consequence that is objectionable. If taken on its own, as part of an account of canonical ways of thinking of senses, claim (i) strikes me as plausible. What is *not* plausible is only that grasp of such ways of thinking is a condition on understanding sentences containing indirect contexts. And that is because it not plausible that employment of such ways of thinking is a condition on using sentences containing indirect contexts.

The result is that there are really two separate issues at stake here: (i) the existence of Frege's hierarchy (a hierarchy of canonical ways of thinking of senses) and (ii) the need to draw on the elements of the hierarchy in providing a theory of indirect contexts. My concern here is only with (ii). So far as (i) goes, Peacocke's proposal is attractive. It simply has unacceptable consequences if employed in providing a theory of indirect contexts.

2.4 Conclusion

Before moving on, I want to stress the two central conclusions I have drawn in this section.

First, while Peacocke, in particular, provides a plausible account of canonical ways of thinking of senses (and thus of infinite hierarchies of such ways of thinking), it is not plausible that these ways of thinking serve as the senses of expressions occurring in indirect contexts. Such a supposition would make it out to be much more difficult than it is to understand sentences containing indirect contexts. So, while a theory like Peacocke's might suffice to make plausible the existence of Frege's hierarchy itself, it does not suffice to save Frege's theory of indirect contexts from the sort of objection raised by Carnap, Davidson, and Kripke.

Second, certain aspects of Burge's account of the hierarchy suggest that, at least for the purposes of providing an account of indirect contexts, Frege's introduction of indirect senses is simply gratuitous. For the account suggests that we can understand a sentence like 'Plato believes that Socrates is wise' if (i) we have a grasp of the customary sense of each of its parts and (ii) we know how those parts are combined in the whole. There is thus no need to require a grasp of the indirect sense of 'Socrates is wise'.

Burge's own view, as we saw, is that a grasp of the indirect sense of 'Socrates is wise' *is* needed. And it is indeed needed, given principles Frege himself seems to accept: for the existence of indirect senses is forced on us by the conjunction of *IND-REF* and Frege's Principle. In an attempt to save Frege's theory, Burge concludes that a grasp of the indirect sense is, if required, somehow at the same time *guaranteed* by a grasp of 'that' and a grasp of the customary sense of 'Socrates is wise'. It appears, however, that the guarantee is introduced solely to meet the need. Of course, if the need is real, it may make sense to accept the guarantee (strange as that guarantee may seem). But the present state of affairs should, I submit, make us wonder whether the need is in fact real. It should make us wonder, in particular, whether we can avoid the introduction of indirect senses entirely, by rejecting either *Ind-Ref* or Frege's Principle.

3 Extensionality and indirect contexts

3.1 Motivation

When I introduced Frege's theory of indirect contexts, in the Introduction, I did so without saying anything about his reasons for accepting it. But these reasons are well-known: Frege's theory of indirect contexts is grounded, ultimately, in his extensionalism, i.e., his adherence to the following principle:

the Principle of Extensionality: The truth-value of a sentence is a function only of (its syntax and) the references of its parts.

Given only this principle and facts about the truth-conditions of relevant sentences (i.e., sentences of the form 'S believes that p '), it can be shown that expressions occurring in indirect contexts plausibly *refer* to their customary senses, rather than their customary references. Assuming in addition only the truth of Frege's Principle (i.e., that the sense of an expression is a way of thinking of its reference), it follows that expressions occurring in indirect contexts *express* ways of thinking of their customary senses, rather than their customary senses themselves.

Thus, consider the two sentences

(7) John believes that Hesperus is a planet

and

(8) John believes that Phosphorus is a planet.

If we accept the Principle of Extensionality, and we assume that 'Hesperus' and 'Phosphorus' have their customary references in (7) and (8), then it follows that (7) and (8) have the same truth-conditions. But they do not, because John might not believe (he might even deny) that Hesperus is Phosphorus. So, if we accept the Principle of Extensionality, as Frege does, we are forced to conclude that, in (7) and (8), 'Hesperus' and 'Phosphorus' do not have their customary references. It is this fact that leads Frege to suggest that

they refer instead to their customary *senses*. In other words, it is the Principle of Extensionality that leads Frege to introduce IND-REF and the notion of indirect reference. The introduction of the notion of indirect *sense* then follows from the conjunction of IND-REF and Frege's Principle.

But now why not just treat sentences like (7) and (8) as exceptions to the Principle of Extensionality?¹⁴ After all, if we did so, we could perhaps say both that sentences (7) and (8) have different truth-conditions and that 'Hesperus' and 'Phosphorus' refer, in (7) and (8), to their customary references, and so express their customary senses. The arguments of §2 suggest that such a view, if it is possible, is actually preferable to Frege's own. My aim in the remainder of this essay, then, is to show that that such a view *is* possible, by showing that we can coherently allow such exceptions to the Principle of Extensionality.¹⁵ (I explain my reasons for not rejecting Frege's Principle instead in §3.5.)

3.2 The strategy

The view I propose in what follows employs a semantic framework modeled on Carnap's (1947) "method of intension and extension,"

¹⁴Kripke raises this question as well, but immediately sets it aside, apparently on merely interpretive grounds: "Why shouldn't the reference of the whole be allowed to depend on other features [than the references] of the parts? Indeed, why not just say that exceptions to this principle obviously exist, e.g., in indirect contexts and possibly in quotation? [...] It could be argued that *for Frege* the principle is analytic" (2008: 270, note 46, my emphasis). Kripke may be right about Frege. The question here, though, is whether Frege is right about the principle, i.e., whether the principle really is analytic, or even true.

¹⁵I should mention here that Parsons (2009: 49) actually implicitly rejects (i.e., allows exceptions to) the Principle of Extensionality. He is nonetheless forced to introduce higher-order senses into his theory of indirect contexts, however, because he retains a metaphysical variant of the Principle of Extensionality: namely, the principle that the truth-value of a thought is a function only of (its form or "syntax" and) the objects and functions of which its parts are ways of thinking. I say more about this "Metaphysical Principle of Extensionality," and my strategy for rejecting it, in §3.5 below. Despite this crucial difference between my own view and Parsons's, I have learned a great deal from his paper.

adapting the framework slightly to suit my Fregean purposes. I will begin, then, by explaining briefly how I take this adaptation of Carnap's strategy to work.

Importantly, if we consider just simple sentences, like 'Socrates is wise', Frege's and Carnap's theories are formally quite similar. In particular, where Frege assigns to each expression both a *sense* and a *reference*, and holds that sense determines reference (cp. Frege's Principle), Carnap assigns to each expression both an *intension* and an *extension*, and holds that intension determines extension. And where Frege holds that the truth-conditions of such sentences are fixed by (their syntax and) their references, and indeed that the reference of a sentence is its truth-value, Carnap holds that the truth-conditions of such sentences are fixed by (their syntax and) their extensions, and indeed that the extension of a sentence is its truth-value. Finally, where Frege holds that the sense of a sentence is a thought, namely, the thought it expresses, Carnap holds that the intension of a sentence is a proposition, namely, the proposition it expresses.

As I understand it, the shared idea here is that each expression has two different *semantic values*: *sense* and *reference* in Frege; *intension* and *extension* in Carnap. And, in both cases, one of these semantic values (sense in Frege; intension in Carnap) is connected with what a sentence *expresses*, while the other (reference in Frege; extension in Carnap) is connected with the *truth-conditions* of sentences.

What is interesting about comparing these two semantic frameworks, however, is that, for Carnap, the connection between extension and truth-conditions is, in certain cases, *broken*. Thus, when he comes to a treatment of indirect contexts, what Carnap says is not that the extension of an expression in an indirect context *shifts*—which would be the direct parallel of what Frege says—but instead that the truth-conditions of some sentences depend on their *intensions* rather than on their *extensions*.¹⁶

¹⁶More precisely, what Carnap says is that the truth-conditions of belief-sentences (and other propositional-attitude sentences) depend on intensional isomorphisms,

To allow exceptions to the Principle of Extensionality, then, is to accept the Fregean variant of Carnap's claim that in sentences of the form 'S believes that p ' the embedded sentence ' p ' occurs in a nonextensional context, in the sense that the truth-conditions of the whole sentence depend on the *intension*, rather than the *extension*, of the sentence ' p ' (and, *a fortiori*, on the intensions, rather than the extensions, of the significant parts of ' p '). The Fregean variant of this claim is simply that the truth-conditions of a sentence of the form 'S believes that p ' depend on the (customary) *sense*, rather than the (customary) *reference*, of the embedded sentence ' p '. This is, importantly, a claim that Frege himself accepts. But, because he also accepts the Principle of Extensionality, he takes it to imply that the embedded sentence ' p ' *refers to* its (customary) sense. What the possibility of Carnap's alternative semantic framework shows, then, is precisely that this last move is not forced on us by the nature of semantic theorizing as such; it shows us that we don't need to say that an expression in an indirect context *refers to* its customary sense in order to say that the truth-conditions of a sentence of the form 'S believes that p ' *depend on* the customary sense (rather than the customary reference) of ' p '.

In order finally to assess these claims, however, we need to see what a non-extensionalist Fregean theory of indirect contexts looks like in detail. I proceed, in what follows, by first providing (in §3.3) an appropriate Fregean semantics for a fragment of English that contains no indirect contexts. With that background in place, I show (in §3.4) how to extend that semantics to an expanded fragment of English that does contain indirect contexts.¹⁷

3.3 A sample language

Vocabulary. The vocabulary of the initial language fragment is as follows.

i.e., not just on the intension but on the way in which it is, as it were, constructed (see 1947: 53–59). But, for present purposes, we can ignore this complication.

¹⁷My approach here is modeled on that employed by Parsons 2009.

Table 1: Senses and references of the basic expressions

	Sense	Reference
'Socrates'	$\langle \text{Socrates} \rangle$	Socrates
'Plato'	$\langle \text{Plato} \rangle$	Plato
'Aristotle'	$\langle \text{Aristotle} \rangle$	Aristotle
' ξ is wise'	$\langle \xi \text{ is wise} \rangle$	Wise(x)
' ξ believes ζ '	$\langle \xi \text{ believes } \zeta \rangle$	Believes(x,y)

Terms: 'Socrates', 'Plato', 'Aristotle'

Predicates: ' ξ is wise', ' ξ believes ζ '

The Greek letters employed here anticipate the syntax of the language. In particular, both ' ξ ' and ' ζ ' mark places at which predicates can be completed by terms.

Syntax. There is just one syntactic rule:

SYNT-1: The result of completing a predicate with an appropriate number of terms is a sentence.

The language thus contains my example sentence 'Socrates is wise'. It also of course contains other sentences, but I will continue to use this sentence as my primary example of a sentence that contains no indirect contexts.

Semantics. The senses and references of the simple expressions are given in Table 1. In general, the reference of a term is an object, and the reference of a predicate is a function from objects to truth-values. In accordance with Frege's Principle, the sense of an expression is to be understood as a way of thinking of its reference. So the sense of a term is a way of thinking of the object to which the term refers, and the sense of a predicate is a way of thinking of the function to which the predicate refers. I indicate ways of

thinking by the use of angle brackets, and I employ Greek letters to mark places at which ways of thinking of functions (the senses of predicates) can be completed by ways of thinking of objects (the senses of terms) to form thoughts (the senses of sentences).

The following semantic principles then determine the sense and reference of every sentence of the language, on the basis of its syntax and the senses and references of its parts (the predicate ' $F\xi$ ' in these principles is arbitrarily complex and may contain "gaps" other than that marked by ' ξ ', making the principles recursive):

REF-1: The reference of a complex expression ' Fa ' is the result of applying the reference of the predicate ' $F\xi$ ' to the reference of the term ' a '.

SENSE-1: The sense of a complex expression ' Fa ' is the result of completing the sense of the predicate ' $F\xi$ ' with the sense of the term ' a '.

Using the principle REF-1, we can then derive the reference of the sentence 'Socrates is wise', as follows.

The reference of 'Socrates is wise':

1. The reference of the sentence 'Socrates is wise' is the result of applying the reference of the predicate ' ξ is wise' to the reference of the term 'Socrates'. (By REF-1)
2. The reference of ' ξ is wise' is $\text{Wise}(x)$. (Axiom)
3. The reference of 'Socrates' is Socrates. (Axiom)
4. So the reference of 'Socrates is wise' is the result of applying $\text{Wise}(x)$ to Socrates, i.e., $\text{Wise}(\text{Socrates})$. (From 1–3)

In order to derive the *sense* of this sentence, however, we need to know how the senses of predicates combine with the senses of terms to form the senses of sentences. We need to know, in other words, how ways of thinking of functions from objects to truth-values combine with ways of thinking of objects to form thoughts.

The “syntax” of thoughts. I understand ways of thinking of objects, like $\langle \text{Socrates} \rangle$, and ways of thinking of functions from objects to truth-values, like $\langle \xi \text{ is wise} \rangle$, as referential and predicative *capacities*, respectively. This construal of ways of thinking as capacities is not essential to what follows, but it will make the formulation of the relevant principles rather less cumbersome. What will matter about these capacities, for present purposes, is only that each referential capacity is associated with a single object and that each predicative capacity is associated with a single function from objects to truth-values. There are of course difficult questions that arise here—What are these capacities? How are they associated with the objects and functions of which they are (as here understood) ways of thinking? etc.—but these are issues that need to be treated elsewhere. Here, I will simply assume that the syntax of the *language* provides an image, of sorts, of the “syntax” of the thoughts expressible in it.¹⁸ So, just as the result of completing a predicate with an appropriate number of terms is a sentence, so the result of completing a predicative capacity with an appropriate number of referential capacities is a thought. Again, a proper account of the “syntax” of thoughts will rest on a complete metaphysics of thoughts, which it is not my aim to provide here. Here I am simply assuming, in effect, that the metaphysics of thoughts will bear out the claim that thoughts have a sort of syntax, a basically language-like structure.

With these assumptions in mind, we can complete the semantics for our initial language fragment by introducing the following principle (where $\langle F\xi \rangle$ is, again, arbitrarily complex and may contain “gaps” other than that marked by ‘ ξ ’):

THOUGHT-SYNT-1: The result of completing a predicative capacity $\langle F\xi \rangle$ with a referential capacity $\langle a \rangle$ is the way of thinking $\langle Fa \rangle$.

¹⁸Cf. Frege 1923: 1: “It is astonishing what language can do. With a few syllables it can express an incalculable number of thoughts, so that even a thought grasped by a human being for the very first time can be put into a form of words which will be understood by someone to whom the thought is entirely new. This would be impossible, were we not able to distinguish parts in the thought corresponding to the parts of a sentence, so that the structure of the sentence serves as an image of the structure of the thought.”

We can then derive the sense of the sentence 'Socrates is wise'.

The sense of 'Socrates is wise':

1. The sense of 'Socrates is wise' is the result of completing the sense of the predicate ' ξ is wise' with the sense of the term 'Socrates'. (By SENSE-1)
2. The sense of ' ξ is wise' is $\langle \xi$ is wise \rangle . (Axiom)
3. The sense of 'Socrates' is \langle Socrates \rangle . (Axiom)
4. So the sense of 'Socrates is wise' is the result of completing $\langle \xi$ is wise \rangle with \langle Socrates \rangle , i.e., \langle Socrates is wise \rangle . (From 1–3 and THOUGHT-SYNT-1)

The semantics of thoughts. One final issue I need to mention here has to do with the way in which the truth-conditions of a thought are determined by its component parts. In the case of the thoughts expressible in this initial language fragment, at least (but not, as we will later see, in every case), it would seem that the truth-value of a thought is determined by the objects and functions of which its parts are ways of thinking. So if we say that the *denotation* of a referential or predicative capacity is the object or function of which it is a way of thinking, we can provide a semantics of thoughts by means of the following principle (where, as before, $\langle F\xi \rangle$ is arbitrarily complex and may contain "gaps" other than that marked by ' ξ ')

THOUGHT-TC-1: The denotation of a complex capacity $\langle Fa \rangle$ is the result of applying the denotation of the predicative capacity $\langle F\xi \rangle$ to the denotation of the referential capacity $\langle a \rangle$.

So the truth-value of the thought \langle Socrates is wise \rangle is the result of applying the denotation of $\langle \xi$ is wise \rangle to the denotation of \langle Socrates \rangle , i.e., the result of applying $\text{Wise}(x)$ to Socrates, i.e., $\text{Wise}(\text{Socrates})$. What is important here is that this thought turns out to have the same truth-conditions as the sentence that expresses it, the sentence 'Socrates is wise'. We will see below that it takes some additional work to secure this result for sentences containing indirect contexts. The reason for the difficulty—to anticipate—is that a simple

generalization of THOUGHT-TC-1 is essentially another version (a metaphysical, as opposed to a linguistic, version) of the Principle of Extensionality; and so it, too, must admit of exceptions.

3.4 The theory of indirect contexts

I turn now to the theory of indirect contexts. My strategy here, again, is to introduce an indirect context-creating expression ('that') and to supplement the semantics of §3.3 so as to explain how the sense and reference of every sentence of the newly expanded language fragment is determined by (i) its syntax and (ii) the senses and references of its parts.

Vocabulary. In addition to the vocabulary provided above, the expanded language fragment contains a single operator.

Operators: 'that ρ '

Again, the Greek letter ' ρ ' anticipates the syntax of operators; it marks a place at which an operator can be completed by a sentence.

Syntax. There is just one additional syntactic rule:

SYNT-2: The result of completing the operator 'that ρ ' with a sentence is a term.

The expanded language fragment thus includes the example sentences 'Plato believes that Socrates is wise' and 'Aristotle believes that Plato believes that Socrates is wise'. As before, it also includes many other sentences; but I will continue to use these as my primary examples.

Semantics (Part I): Reference. According to Frege, the truth-value of a thought of the form 'S believes that p ' is a function of the (customary) sense, rather than the (customary) reference, of the embedded sentence ' p '. On Frege's own theory, this idea is captured

by the conjunction of IND-REF and the Principle of Extensionality. On the Carnapian variant of Frege's theory, by contrast, it can be captured instead by the single principle:

REF-2: The reference of a complex expression ' $F(\text{that } p)$ ' is the result of applying the reference of the predicate ' $F\xi$ ' to the sense of the sentence ' p '.

Once we have introduced REF-2, there is no reason to suppose that the operator 'that ρ ' has a reference. For, given only REF-2 and the principles provided in §3.3—but no assignment of a reference to 'that ρ '—we can already derive the reference of the sentence 'Plato believes that Socrates is wise'.

The reference of 'Plato believes that Socrates is wise':

1. The reference of 'Plato believes that Socrates is wise' is the result of applying the reference of the predicate ' ξ believes ζ ' to the reference of 'Plato' and the sense of 'Socrates is wise' (in the appropriate order). (By REF-1 and REF-2)
2. The reference of 'Plato' is Plato. (Axiom)
3. The reference of ' ξ believes ζ ' is Believes(x,y). (Axiom)
4. The sense of 'Socrates is wise' is \langle Socrates is wise \rangle . (Previous result)
5. So the reference of 'Plato believes that Socrates is wise' is the result of applying Believes(x,y) to Plato and \langle Socrates is wise \rangle (in the appropriate order), i.e., Believes(Plato, \langle Socrates is wise \rangle). (From 1–4)

One important consequence of not assigning a reference to 'that ρ ' is that complex terms of the form 'that p ' will not be assigned references either. But I assume that the semantics will be adequate if it assigns an appropriate sense and reference to every *sentence* of the language, even if it does not assign a sense and reference to every *complex expression* of the language.

Since we can make do without assigning a reference to 'that ρ ', I will assume that we can also make do without assigning it a sense. In other words, I will give a thoroughly syncategorematic

treatment of 'that ρ '.¹⁹ To say that 'that ρ ' has no sense, however, is not to say that it is meaningless. On the contrary, its meaning is given by the principle REF-2, which describes its contribution to the truth-conditions of the sentences in which it occurs. What does follow from the claim that 'that ρ ' has no sense is that no thought-component corresponds to it. In other words: there is no capacity that stands to 'that ρ ' as the referential capacity $\langle \text{Socrates} \rangle$ stands to the name 'Socrates', or as the predicative capacity $\langle \xi \text{ is wise} \rangle$ stands to the predicate ' ξ is wise'. So put, however, the claim that 'that ρ ' has no sense does raise an important question: how *are* we to understand the thoughts expressed by sentences that contain the operator 'that ρ '—for example, the thought that Plato believes that Socrates is wise? How, in particular, are the parts of these thoughts combined in the whole?

The semantics of thoughts. I will begin here with the semantics of thoughts, for the simple reason that we already know what the truth-conditions of the thought that Plato believes that Socrates is wise need to be, if the theory I am proposing here is going to be acceptable. For we already know that the truth-value of the sentence 'Plato believes that Socrates is wise' is $\text{Believes}(\text{Plato}, \langle \text{Socrates is wise} \rangle)$, that this sentence expresses the thought that Plato believes that Socrates is wise, and that the latter needs to have the same truth-value as the former. The challenge, then, is simply to say how a thought with the parts $\langle \text{Plato} \rangle$, $\langle \xi \text{ believes } \zeta \rangle$, and $\langle \text{Socrates is wise} \rangle$ —the parts, on the present account, of the thought that Plato believes that Socrates is wise—can have the truth-value $\text{Believes}(\text{Plato}, \langle \text{Socrates is wise} \rangle)$.

The first key to meeting this challenge is the principle REF-2. What REF-2 tells us is that the truth-value of the thought that Plato believes that Socrates is wise is the result of applying the reference of ' ξ believes ζ ' to the reference of 'Plato' and the sense of 'Socrates is wise'. But the reference of ' ξ believes ζ ' is the denotation of $\langle \xi \text{ believes } \zeta \rangle$, the reference of 'Plato' is the denotation of $\langle \text{Plato} \rangle$,

¹⁹My reasons for preferring the syncategorematic treatment are given in §3.5.

and the sense of 'Socrates is wise' is the thought $\langle \text{Socrates is wise} \rangle$. So we can rephrase the relevant application of REF-2 in terms of the relevant thought-components themselves, and say, simply, that the truth-value of the thought that Plato believes that Socrates is wise is the result of applying the denotation of $\langle \xi \text{ believes } \zeta \rangle$ to the denotation of $\langle \text{Plato} \rangle$ and the thought (n.b.: *not* the denotation of the thought) $\langle \text{Socrates is wise} \rangle$.

The second key to meeting the present challenge is to see that thoughts are partly defined by their truth-conditions, and completely defined by *the way in which* their parts *determine* their truth-conditions.²⁰ Thus, the thought that Plato believes that Socrates is wise can be defined as the thought that has all and only the parts $\langle \text{Plato} \rangle$, $\langle \xi \text{ believes } \zeta \rangle$, and $\langle \text{Socrates is wise} \rangle$ and whose truth-value is the result of applying the denotation of $\langle \xi \text{ believes } \zeta \rangle$ to the denotation of $\langle \text{Plato} \rangle$ and the thought $\langle \text{Socrates is wise} \rangle$. If we then mark the fact that the truth-value of a thought is a function of one of its parts itself, rather than of the denotation of that part, by putting the relevant part in square brackets, the thought that Plato believes that Socrates is wise is $\langle \text{Plato believes } [\text{Socrates is wise}] \rangle$. The semantics of the thoughts expressible by the expanded language fragment, which includes 'that ρ ', can then be completed by adding the following principle:

THOUGHT-TC-2: The denotation of a complex capacity $\langle F[p] \rangle$ is the result of applying the denotation of the predicative capacity $\langle F\xi \rangle$ to the thought $\langle p \rangle$.

Together, THOUGHT-TC-1 and THOUGHT-TC-2 give the result that the truth-value of the thought that Plato believes that Socrates is wise, i.e., $\langle \text{Plato believes } [\text{Socrates is wise}] \rangle$, is $\text{Believes}(\text{Plato}, \langle \text{Socrates is wise} \rangle)$. So the thought that Plato believes that Socrates is wise has the same truth-conditions as the sentence 'Plato believes that Socrates is wise'.

²⁰Compare this idea to the idea of Frege's, mentioned in §2.2, that the sense of a sentence is the thought that its truth-conditions are fulfilled.

The “syntax” of thoughts. My introduction of square brackets reflects the fact that there are two ways in which an expression can contribute to the truth-conditions of the sentences in which it occurs, and thus, correspondingly, two ways in which a thought-component can contribute to the truth-conditions of the thoughts in which it occurs. I also said that the way in which a thought-component contributes to the truth-conditions of a thought in which it occurs is definitive of the thought itself. And this latter fact has consequences for the “syntax” of thoughts: it implies that there are two different ways in which a thought-component can occur in a thought. I will thus say that a thought-component marked by square brackets occurs *indirectly* in the relevant thought, and that a thought-component not marked by square brackets occurs *directly* in the relevant thought. Using these notions of direct and indirect occurrence, we can then revise the principle THOUGHT-SYNT-1 and introduce a second principle to complete the syntax of the thoughts expressible in the expanded language fragment, the fragment that includes the operator ‘that ρ ’:

THOUGHT-SYNT-1 (revised): The result of completing a predicative capacity $\langle F\xi \rangle$ *directly* with a referential capacity $\langle a \rangle$ is the way of thinking $\langle Fa \rangle$.

THOUGHT-SYNT-2: The result of completing a predicative capacity $\langle F\xi \rangle$ *indirectly* with a thought $\langle p \rangle$ is the way of thinking $\langle F[p] \rangle$.

Thus, the result of completing the predicative capacity $\langle \xi$ believes $\zeta \rangle$ directly with the referential capacity $\langle \text{Plato} \rangle$ and indirectly with the thought $\langle \text{Socrates is wise} \rangle$ is the thought $\langle \text{Plato believes [Socrates is wise]} \rangle$.

I will say more about the notions underlying my use of square brackets in §3.5, where I respond to two potential objections to my theory of indirect contexts.

Semantics (Part II): Sense. The final step in formulating the theory of indirect contexts is to provide a principle that allows us to derive the sense of a sentence containing the operator ‘that ρ ’ from

its syntax and the senses and references of its parts. This principle must interact appropriately with THOUGHT-SYNT-2, and so must appeal to the notion of the indirect completion of a predicative capacity, meaning that we must also revise the earlier principle SENSE-1. The revised version of SENSE-1, and the new principle SENSE-2, are as follows:

SENSE-1 (revised): The sense of a complex expression ' Fa ' is the result of completing the sense of the predicate ' $F\xi$ ' *directly* with the sense of the term ' a '.

SENSE-2: The sense of a complex expression ' $F(\text{that } p)$ ' is the result of completing the sense of the predicate ' $F\xi$ ' *indirectly* with the sense of the sentence ' p '.

Given these principles, plus the earlier principles, we can now derive the sense of the sentence 'Plato believes that Socrates is wise'.

The sense of 'Plato believes that Socrates is wise':

1. The sense of 'Plato believes that Socrates is wise' is the result of completing the sense of the predicate ' ξ believes ζ ' *directly* with the sense of 'Plato' and *indirectly* with the sense of 'Socrates is wise' (in the appropriate order). (By SENSE-1 and SENSE-2)
2. The sense of ' ξ believes ζ ' is $\langle \xi$ believes $\zeta \rangle$. (Axiom)
3. The sense of 'Plato' is $\langle \text{Plato} \rangle$. (Axiom)
4. The sense of 'Socrates is wise' is $\langle \text{Socrates is wise} \rangle$. (Previous result)
5. So the sense of 'Plato believes that Socrates is wise' is the result of completing $\langle \xi$ believes $\zeta \rangle$ directly with $\langle \text{Plato} \rangle$ and indirectly with $\langle \text{Socrates is wise} \rangle$ (in the appropriate order), i.e., $\langle \text{Plato believes} [\text{Socrates is wise}] \rangle$. (From 1–4, THOUGHT-SYNT-1 and THOUGHT-SYNT-2)

We can also derive both the sense and the reference of the more complex sentence 'Aristotle believes that Plato believes that Socrates is wise'—and, indeed, of any sentence involving further iterations of 'believes that'. But the derivation is straightforward enough that I

will not produce it here. Instead, I will close by considering a couple of important potential objections to the view just introduced.

3.5 Objections and replies

First objection. I said above that a thought is defined, not just by its parts, but by the way in which its parts determine its truth-conditions. And you might wonder whether this view of the nature of thoughts isn't just an *ad hoc* means of introducing the crucial distinction between direct and indirect occurrences of thought-components. So I want to take a moment here to show that the idea has application outside the theory of indirect contexts. That's not to say that the idea doesn't require more explanation and defense; it probably does. But it is to say that my use of it here at least isn't *ad hoc*. And some additional explication of these ideas is undoubtedly in order anyway, since it is plausibly a general failure to acknowledge the possibility of what I have called an indirect occurrence of a thought-component that has prevented views like the one proposed above from being developed and taken seriously by theorists of indirect contexts.

Consider, then, the thought that Romeo loves Juliet and the thought that Juliet loves Romeo. These are thoughts that have exactly the same parts—⟨Romeo⟩, ⟨Juliet⟩, and ⟨ ξ loves ζ ⟩—but different truth-conditions. They are distinct thoughts, I will say, not because they differ in *matter*, but because they differ in *form*. And they differ in form precisely because their truth-conditions depend *in different ways* on their parts (on the *same* parts): the truth-value of the former is the result of applying the denotation of ⟨ ξ loves ζ ⟩ to the denotation of ⟨Romeo⟩ and the denotation of ⟨Juliet⟩, *taken in that order*, whereas the truth-value of the latter is the result of applying the denotation of ⟨ ξ loves ζ ⟩ to the denotation of ⟨Juliet⟩ and the denotation of ⟨Romeo⟩, taken in *that* order.

So the idea is that the order in which the terms occur in a sentence can serve to characterize the sentence's form. Similarly, I am suggesting, the fact that some expressions in a sentence occur within

the scope of the complementizer 'that' can serve to characterize the sentence's form. What ties these two features of sentences together is the idea that the logical form of a sentence, and also of a thought, is simply the way in which its truth-conditions depend on its parts (and, specifically, on their semantic values).

Second objection. On the approach I have adopted here, 'that ρ ' is treated syncategorematically: it has neither a sense nor a reference. Its entire semantics is provided, instead, by the principles REF-2, SENSE-2, THOUGHT-SYNT-2, and THOUGHT-TC-2. But, you might wonder, why not treat it categorically, and say that it refers to, say, the identity function, and expresses a way of thinking of that function?²¹ Such a view would, after all, assign the right truth-conditions to the relevant sentences. And it promises to allow us to eschew the distinction between direct and indirect occurrences of thought-components, and the novel logical form (marked by square brackets) that comes with it.

The answer is that, while the categoric approach does indeed assign the correct truth-conditions to *sentences* containing indirect contexts, it assigns the *wrong* truth-conditions to the thoughts expressed by those sentences.

The basic problem is this. If we allow only (what I have called) direct occurrences of thought-components, the truth-value of any thought will need to depend on the objects and functions of which its parts are ways of thinking. In other words, we will need to accept the following metaphysical variant of the Principle of Extensionality:

the Metaphysical Principle of Extensionality: The truth-value of a thought is a function only of its form (i.e., its "syntax") and the denotations of its parts.

But given (i) the Metaphysical Principle of Extensionality and (ii) the anti-extensionalist assumption that (for example) the thought that Socrates is wise is a component part of the thought that Plato

²¹Cf. Parsons 1982: 325, Pietroski 1996.

believes that Socrates is wise, it follows that the thought that Plato believes that Socrates is wise depends on the *truth-value* of the thought that Socrates is wise. It follows, in other words, that the thought that Plato believes that Socrates is wise is true just in case Plato stands in the belief-relation to the truth-value *Wise(Socrates)*. Supposing that the latter is the True, it follows that the thought that Plato believes that Socrates is wise is true just in case Plato has any true belief whatsoever.²²

So far as I can see, there are only two ways of avoiding this conclusion. The first is to take Frege's view, and conclude that the thought that Plato believes that Socrates is wise does *not* have the thought that Socrates is wise as a component part. The other is the view proposed in §3.4.

Conclusion

I have made three claims in this paper. The first is that, even if Frege's hierarchy exists, it is a mistake to appeal to it in formulating a Fregean theory of indirect contexts, because it is more difficult to grasp the canonical way of thinking of a sense than it is to understand the occurrence of the relevant expression in an indirect context. The second is that—the Principle of Extensionality aside—for purposes of providing a theory of indirect contexts, it would be gratuitous to introduce indirect senses, because we can explain our understanding of sentences containing indirect contexts without so much as mentioning them. The third is that we can, and therefore should, abandon—that is, allow exceptions to—the Principle of Extensionality. In support of this last claim, I have developed, in some

²²This same objection applies to views—like Dummett's (1973: 267–268), an earlier view of Peacocke's (1996), and Pietroski's (1996)—that retain the Principle of Extensionality while allowing exceptions to Frege's Principle. For a detailed explanation of the objection, as applied to Peacocke's earlier view, see Burge 2004: 183ff. The moral is that the only broadly Fregean alternative to Frege's own theory of indirect contexts is a resolutely non-extensionalist one, i.e., one that allows exceptions to both the Principle of Extensionality and the Metaphysical Principle of Extensionality (stated below).

detail, a non-extensionalist neo-Fregean theory of indirect contexts, one that employs Frege's theory of sense and reference within a non-extensionalist semantic framework modeled on Carnap's method of intension and extension. My final contention, then, is that, in the absence of strong arguments in favor of exceptionless adherence to the Principle of Extensionality, the non-extensionalist variant of Frege's theory deserves to be taken very seriously.

So: *are* there strong arguments in favor of exceptionless adherence to the Principle of Extensionality? The only principled reason²³ I have been able to find in Frege's own writings stems from his desire for his *Begriffsschrift* to be fully substitutional. The desire for substitutivity motivates a commitment to extensionality because a fully extensional language is guaranteed to be fully substitutional. But this motivation for extensionalism doesn't seem to me to go very far. Substitutivity does make for easy assessment of the validity of proofs. But it would be neither objectionable nor unmanageable to make a handful of principled exceptions. So the real question is whether there is, somewhere deep in Frege's philosophy of logic, a reason for thinking that fully logical thought and language *are* (must be) fully substitutional. I myself doubt that such a reason, even if one were to be found, would persuade philosophers working today. But, nonetheless, if there is such a reason, it would deepen our understanding of Frege, and of the logic of indirect contexts, if we could lay it bare.

Among Frege's defenders, on the other hand, the most powerful motivation for preserving extensionalism seems to be that "no non-extensionalist theory of belief has provided an alternative that is plausibly superior to the Fregean approach" (Burge 1979: 160). But, if the theory developed in §3 is coherent, this claim is simply false: the non-extensionalist neo-Fregean theory proposed here retains all the benefits of Frege's own, while avoiding the defects that led philosophers like Carnap, Davidson, and Dummett to reject or revise it.

²³ Another reason is *ad hominem*: Frege accepts Leibniz's *salva veritate* principle, applying it to both thoughts and sentences. Cf. Frege 1892: 158.

References

- Almog, Joseph and Paolo Leonardi (eds.) 2009. *The Philosophy of David Kaplan*. Oxford: Oxford University Press.
- Beaney, Michael (ed.) 1997. *The Frege Reader*. Malden, MA: Blackwell.
- Burge, Tyler. 1979. Frege and the hierarchy. *Synthese* 40(2): 265–281. Reprinted in Burge 2005, 153–166. References give the page number of the reprint.
- 2004. Postscript to 'Frege and the hierarchy'. In Burge 2005: 167–210.
- 2005. *Truth, Thought, Reason*. Oxford: Oxford University Press.
- Carnap, Rudolf. 1947. *Meaning and Necessity*. Chicago: The University of Chicago Press.
- Davidson, Donald. 1965. Theories of meaning and learnable languages. In Davidson 2001: 3–15.
- 2001. *Inquiries Into Truth and Interpretation*. Oxford: Oxford University Press.
- Dummett, Michael. 1973. *Frege: Philosophy of Language*. London: Duckworth.
- Frege, Gottlob. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100: 25–50. Translated by Max Black and reprinted as "On Sinn and Bedeutung" in Beaney 1997, 151–171. References give the page number of the original, then the page number of the reprint.
- 1893. *Grundgesetze der Arithmetik*. Vol. I. Jena: Verlag Hermann Pohle. Selections translated by Michael Beaney and reprinted in Beaney 1997, 194–223. References give the page number of the original, then the page number of the reprint.

- Frege, Gottlob. 1923. Gedankengefüge. *Beiträge zur Philosophie des deutschen Idealismus* III: 36–51. Translated by R. M. Stoothof and published in 1963 in *Mind* 72(285), 1–17. References give the page number of the translation.
- 1980. *Philosophical and Mathematical Correspondence*. Ed. by Gottfried Gabriel et al. Trans. by Hans Kaal. Chicago: The University of Chicago Press.
- Kaplan, David. 1968. Quantifying in. *Synthese* 19: 178–214.
- Kripke, Saul. 2008. Frege's theory of sense and reference: some exegetical notes. In Kripke 2011: 254–291.
- 2011. *Philosophical Troubles*. Oxford: Oxford University Press.
- Mendelsohn, Richard L. 2005. *The Philosophy of Gottlob Frege*. Cambridge: Cambridge University Press.
- Parsons, Terence D. 1982. What do quotation marks name? Frege's theories of quotations and that-clauses. *Philosophical Studies* 42: 315–328.
- 2009. Higher-order senses. In Almog and Leonardi 2009: 45–59.
- Peacocke, Christopher. 1992. *A Study of Concepts*. Cambridge, MA: The MIT Press.
- 1996. Entitlement, self-knowledge and conceptual redeployment. *Proceedings of the Aristotelian Society* 96: 117–158.
- 2009. Frege's hierarchy: a puzzle. In Almog and Leonardi 2009: 159–186.
- Pietroski, Paul M. 1996. Fregean innocence. *Mind & Language* 11(4): 338–370.