

# On Indirect Sense and Reference

Lukas Skiba

This is my final draft of an article forthcoming in *Theoria*

## Abstract

According to Frege, expressions shift their reference when they occur in indirect contexts: in ‘Anna believes that Plato is wise’ the expression ‘Plato’ no longer refers to Plato but to what is ordinarily its sense. Many philosophers, including Carnap, Davidson, Burge, Parsons, Kripke and Künne, believe that on Frege’s view the iteration of indirect context creating operators gives rise to an infinite hierarchy of senses. While the former two take this to be problematic, the latter four welcome the hierarchy with open arms. In this paper I argue (against the quartet) that the hierarchy should be avoided and (against all of them) that it can be avoided.

## 1 Frege and the hierarchy

Frege believes that language is extensional in the sense that the referent of any complex referring expression is uniquely determined by the referents of its components and the order in which these components occur in the complex expression. For Frege, of course, not only names and definite descriptions count as referring expressions but so do predicates, which he takes to refer to concepts, and sentences, which he takes to refer to truth-values.

We don’t have to look far to find apparent counterexamples to Frege’s principle of extensionality. Although the names ‘Plato’ and ‘Aristocles’ co-refer, substituting one for the other in a propositional attitude ascription can affect the truth-value of the sentence in question. For example, we may suppose that (i) is true while (ii) is false:

- (i) Anna believes that Plato is wise.
- (ii) Anna believes that Aristocles is wise.

In *On Sense and Reference* Frege famously refuses to acknowledge such cases as counterexamples to his principle of extensionality. According to him, expressions within a that-clause following a propositional attitudes verb, such as ‘Plato’ in (i) and ‘Aristocles’ in (ii), don’t have their ordinary referent but refer to the sense which they express in ordinary, unembedded contexts (cp. 1892, pp. 59, 66 ff.). Let us call this *the doctrine of reference shifts* and let us call the contexts in which expressions are subject to such a reference shift *indirect contexts*. Since the ordinary senses of ‘Plato’ and ‘Aristocles’ are distinct, the two names thus have different referents in (i) and (ii)

according to the doctrine of reference shift. The fact that substituting one for the other can change the truth-value of the attitude ascription therefore no longer contradicts the principle of extensionality. We have to conceive of the reference shift not only as applying to the expressions ‘Plato’ and ‘is wise’ but also as applying to the complex expression ‘Plato is wise’ as it occurs within the that-clause. For if this expression were still referring to its ordinary referent, rather than to the sense (thought) which it expresses in unembedded contexts, the original problem would re-occur since we would be entitled to substitute any true sentence for it without affecting the truth-value of the attitude ascription. Furthermore, Frege also takes the that-clause as a whole to refer to the thought which the embedded sentence expresses in ordinary contexts (cp. 1892, p. 66; 1904, p. 164).

Although the principle of extensionality is an important source of motivation for Frege’s doctrine of reference shifts it is by no means the only one. In fact, as has been argued by Makin (2000, Ch. 5 §2), this doctrine is entirely natural given both Frege’s conception of propositional attitudes and the reason for which he initially introduced the notion of sense. For Frege, Anna’s believing that Plato is wise consists in her standing in the relation of believing to a certain object. Since Anna can believe that Plato is wise without believing that Aristocles is wise, the object of Anna’s belief cannot be determined merely by the *referents* of the expressions in ‘Plato is wise’. For these referents are the same as the referents of the expressions in ‘Aristocles is wise’. Rather, the object of her belief must be determined by the *senses* of the expressions in ‘Plato is wise’ for these are different from the senses of the expressions in ‘Aristocles is wise’. When we report Anna’s cognitive situation by asserting ‘Anna believes that Plato is wise’ it is thus natural from Frege’s perspective to suppose not only that our utterance of ‘that Plato is wise’ refers to the object of Anna’s belief but also that our utterance of ‘Plato’ refers to the sense which partially determines that object.

The doctrine of reference shifts is thus no marginal curiosity in Frege’s philosophy of language, but is what enables him to uphold his principle of extensionality and can be seen as the natural upshot of his theory of sense and reference. This doctrine, however, immediately gives rise to a question that Frege left unanswered and that has since occupied friends and foes of his theory from Carnap to Kripke. The question can be put as follows: given that a single indirect context creating operator, such as ‘Anna believes that ...’, induces a reference shift, what happens if such operators are iterated? Consider the following sentences where ‘Plato’ occurs in what we will respectively call a *direct*, *singly indirect* and *doubly indirect context*:

- (1) Plato is wise.
- (2) Anna believes that Plato is wise.
- (3) Berta believes that Anna believes that Plato is wise.

Several options suggest themselves. On the first option the iteration of indirect context creating operators gives rise to an infinite hierarchy of senses: Frege tells us that ‘Plato’ refers to its ordinary, direct sense in (2). Given that sense determines reference, understood as the principle that two expression tokens with different referents must have different senses, ‘Plato’ can then no longer express its direct sense in (2) but must express a distinct singly indirect sense. It is natural to suppose then that when ‘Plato’ occurs in (3) it refers to this singly indirect sense and expresses a distinct doubly

indirect sense. By the same reasoning, each additional embedding results in a further shift of the sense and reference of ‘Plato’. Since for any number of indirect context creating operators there is a sentence with yet one additional such operator, ‘Plato’ is associated with infinitely many senses. More generally, the infinite hierarchy view can be characterised as follows:

#### **INFINITE HIERARCHY VIEW**

When an expression follows  $n$  indirect context creating operators it expresses its indirect sense of level  $n$  and refers to the distinct sense which it expresses when it follows  $n - 1$  such operators (for  $n \geq 1$ ).

On the second option only the first indirect context creating operator results in a shift of the sense and reference of ‘Plato’, adding further such operators doesn’t have any effect on the sense and reference of ‘Plato’. This view thus makes do with two levels of sense, direct and singly indirect senses, and can be characterised more generally as follows:

#### **TWO-LEVEL VIEW**

When an expression follows  $n$  indirect context creating operators it expresses its singly indirect sense and refers to its distinct direct sense (for  $n \geq 1$ ).

A third option additionally eliminates singly indirect senses by giving up the principle that sense determines reference. The first indirect context creating operator only effects a reference shift of ‘Plato’ but no sense shift. This view thus makes do with one level of sense, direct senses, and can be characterised more generally as follows:

#### **ONE-LEVEL VIEW**

When an expression follows  $n$  indirect context creating operators it expresses its direct sense and refers to the same direct sense (for  $n \geq 1$ ).

These three options by no means exhaust the space of possibilities, but they are the most natural candidates so it is not surprising that attention has converged on them.

Frege himself makes remarks that are incompatible with the last two options. After saying that ‘we distinguish accordingly the *customary* from the *indirect* reference of a word’ (1892, p. 59) he immediately adds ‘and its *customary* sense from its *indirect* sense’ thus contradicting the one-level view. The only place where he considers doubly indirect contexts is in a letter to Russell (1902). There he claims, without argument, that the doubly indirect referent of an expression must be different from its singly indirect referent, thus contradicting the two-level view as well. However, he neither considers further embeddings nor identifies doubly indirect referents with singly indirect senses. He thus falls short of endorsing the infinite hierarchy view or indeed taking any definite position.

In any case, many commentators are convinced that Frege is *committed* to the hierarchy. Carnap (1947, §30) considers the one-level view in passing but deems it unavailable to Frege due to the principle that sense determines reference (cp. 1947, p. 129). He doesn’t consider the two-level view, but goes on to claim that Frege’s commitment to the hierarchy ‘can be shown by an analysis which we shall not describe

here in detail' (1947, p. 131) where 'not in detail' unfortunately turns out to be a euphemism for 'not at all'.

Like Carnap, Davidson (1965) is convinced that Frege's theory inevitably leads to the hierarchy. He considers neither the one-level nor the two-level view, perhaps because he not only targets Frege's own theory but also Church's (1951) development thereof which explicitly embraces the hierarchy. While Carnap rejects the hierarchy, and hence Frege's theory, primarily because of its ontological extravagance, Davidson famously goes on to argue that it would implausibly render our language unlearnable. There are infinitely many senses associated with each expression. We have only finite cognitive resources. So we can't even come to know all the senses of any one expression.

Impressed by Davidson's argument, Dummett (1973 Ch. 9, 1981 Ch. 6) recommends the one-level view even after realizing in his (1981) that it clashes with the principle that sense determines reference. He considers this a small price to pay in order to avoid the hierarchy which, neglecting the two-level view, he sees as the only alternative.

It is not until Burge (1979) that genuine arguments are presented purporting to show that the hierarchy is an inevitable consequence of Frege's theory. These arguments, elaborated and extended in (2004), have convinced Kripke (2008) who agrees 'with Burge [...] that the hierarchy is an actual consequence of Frege's theory' (2008, p. 184, fn. 9). Künne (2010b, 2010a) in turn finds 'Kripke's comments on the hierarchies [...] entirely convincing' (2010b, p. 539). Parsons who originally believed the hierarchy could be avoided (1981) changes his mind in (2009, §2) due to a variant of one of Burge's arguments.

Unlike Davidson and Carnap, however, none of these philosophers takes the hierarchy to be a problem for Frege's theory. In contrast, a consensus has emerged that the hierarchy is not only unavoidable but should be welcomed by the Fregean with open arms.

In this paper I argue against both of these points of agreement. From a Fregean perspective, the hierarchy can and should be avoided. In §2 I examine the best developed attempts to make the hierarchy palatable. These have been provided by Burge (2004) and Kripke (2008). I argue that both attempts as well as their potential combination leave the hierarchy view more problematic than so far appreciated and that we should therefore still aim to avoid it. In §3 I examine Burge's (1979, 2004) arguments to the effect that this is impossible. While I concede that one of his arguments is effective against the one-level view I argue that his arguments don't trouble the two-level view, which therefore emerges as the best extension of Frege's theory to iterations of indirect context creating operators. In §4 I consider and reject two potential objections to my defense of the two-level view.

## 2 Why the hierarchy should be avoided

The plan for this section is as follows. In §2.1 I present Burge's attempt to tame the hierarchy. In §2.2 I point out its problems. In §2.3 I outline Kripke's venture followed by a discussion of its shortcomings in §2.4 In §2.5 I consider a potential combination of their strategies and argue that it still fails to make the hierarchy palatable.

## 2.1 Burge's hierarchy view

Let us follow Parsons (1981) in calling the relation that holds between the sense of an expression and the expression's referent *presentation*. Disregarding the senses of empty expressions, each sense presents exactly one entity, while each entity is presented by many senses. This is no different when the entity presented is itself a sense. For instance, the sense expressed by 'Plato' is presented by the sense expressed by 'the sense of the first proper name mentioned in this paper'. Assuming that John is nursing a deep-seated affection for it, it is also presented by the sense of 'John's favourite sense'. And according to the infinite hierarchy view it is of course also presented by the singly indirect sense of 'Plato'.

Now, Burge believes that indirect senses are in a certain way privileged over their direct brethren. He calls indirect senses *canonical* and the expressions expressing them *canonical names*. Canonicity is then characterised as follows (cf. 2004, p. 174):

### (CS) Canonical Senses

The sense of a canonical name cannot

- be understood without understanding the sense it names.
- be thought without simultaneously thinking the lowest-level (ultimately, customary) sense in the downward hierarchy associated with the canonical name.

Thus, understanding the doubly indirect sense of 'Plato' is sufficient for understanding its singly indirect sense, while thinking it means simultaneously thinking its direct sense.

Burge assumes that each sense is presented by exactly one canonical sense (cp. 2004, p. 179). While the relation between senses and the things they present is many-one, the relation between canonical senses and the things they present is one-one. This assumption is crucial for Burge's solution to Davidson's learnability problem. In order to convince us that 'all the talk of Davidson, Dummett and others [...] of being unable to learn the hierarchy is wayward and unsupportable' (2004, p. 172), Burge sketches a formalisation of natural language which is supposed to be evidently learnable despite incorporating the hierarchy view. His formalisation makes reference shifts explicit by translating one and the same natural language expression by means of different expressions whenever the natural language expression is subject to a reference shift.

Consequently, Burge's formal language contains finitely many primitive expressions - 'Plato<sub>B</sub>', 'Aristocles<sub>B</sub>', 'is wise<sub>B</sub>', 'is tall<sub>B</sub>', ... - which are used exclusively to translate their homophonic natural language counterparts - 'Plato', 'Aristocles', 'is wise', 'is tall', ... - when those occur in direct contexts.<sup>1</sup> In addition to this, Burge's language contains a canonical name for each of the senses expressed by the primitive expressions just introduced: '<Plato>', '<Aristocles>', '<is wise>', '<is tall>', ... . These canonical names are used exclusively to translate the according natural language expressions when those occur in singly indirect contexts. The natural language sentence (2) can hence be formalised as:

(2') Anna<sub>B</sub> believes<sub>B</sub> <Plato><is wise>

---

<sup>1</sup>I've added the subscript to expressions belonging to Burge's formal language in order to avoid confusion.

As Burge acknowledges, the canonical names for direct senses have to be conceived of as primitive expressions themselves (cp. 2004, p. 174). Despite its appearance, ‘<Plato>’ cannot be regarded as the result of applying a functor ‘<...>’ to the name ‘Plato<sub>B</sub>’. Burge doesn’t tell us why but the reason is straightforward. While the sense and reference of the natural language expression ‘Plato’ differ according to what context it occurs in, ‘Plato<sub>B</sub>’ only ever has the direct sense and reference of ‘Plato’. So if ‘<Plato>’ were the result of applying the functor ‘<...>’ to ‘Plato<sub>B</sub>’, then ‘<...>’ would have to be a functor that creates a direct context. But then we would be entitled to exchange ‘Plato<sub>B</sub>’ for a co-referring expression with a different sense, such as ‘Aristocles<sub>B</sub>’, without changing the referent of the complex expressions. Yet then ‘<Plato>’ and ‘<Aristocles>’ would co-refer and hence could not in fact be the canonical names for the respective *distinct* senses of ‘Plato<sub>B</sub>’ and ‘Aristocles<sub>B</sub>’.

Since the canonical names for direct senses are primitive, their addition to the vocabulary doubles the number of primitive expressions in Burge’s language. But there are still only finitely many of them so that a finite being can learn them all.

But what about further embeddings? To formalise a natural language sentence where ‘Plato’ occurs in a doubly indirect context Burge needs a canonical name for its singly indirect sense. The canonical names for direct senses were introduced as primitive expressions. Doesn’t Burge now have to introduce a further layer of primitive canonical names referring to singly indirect senses? If he has to do this for each level of the hierarchy, his formal language will contain infinitely many primitive expressions and thus be unlearnable.

Now Burge’s assumption that each sense is presented by exactly one canonical sense becomes relevant. It allows him to introduce a functor ‘ $C(x)$ ’ which stands for a function that maps every sense to the canonical sense presenting it. So, for example ‘ $C(<Plato>)$ ’ refers the canonical sense that presents the sense referred to by ‘<Plato>’. Thus ‘ $C(<Plato>)$ ’ refers to the sense expressed by ‘<Plato>’, the singly indirect sense of ‘Plato’. Furthermore, Burge stipulates that whenever ‘ $C(x)$ ’ is applied to a canonical name the resulting expression is itself a canonical name for the sense it refers to. Hence ‘ $C(<Plato>)$ ’ is the canonical name for the singly indirect sense of ‘Plato’. The natural language sentence (3) can then be formalised as:

(3)      Berta<sub>B</sub> believes<sub>B</sub> <Anna><believes>  $C(<Plato>)$  $C(<is wise>)$

What about triply indirect contexts? Since ‘ $C(<Plato>)$ ’ is a canonical name, the expression ‘ $C(C(<Plato>))$ ’ is the canonical name of the sense of ‘ $C(<Plato>)$ ’, and thus formalises ‘Plato’ when it occurs in a triply indirect context.

In this way, infinitely many canonical names are generated from Burge’s finite vocabulary by iterations of ‘ $C(x)$ ’. His formal language is learnable with finite resources. In order to master it, all you need to master is a finite number of primitive expressions, including primitive canonical names for direct senses as well as the functor ‘ $C(x)$ ’. Davidson’s learnability problem is avoided.

According to Burge, appealing to the functor ‘ $C(x)$ ’ in formalising natural language belief ascriptions has a further nice feature (cp. 2004: p. 182; & Appendix II). It allows us to account for certain inferences we ought to account for, such as the following:

- (P1) Anna believes that Plato is wise.  
(P2) Berta believes that Anna believes that Plato is wise.  
(C) So,  $\exists x(\text{Anna believes } x \ \& \ \text{Berta believes that Anna believes } x)$

We might call the quantification that occurs in the conclusion *heterogenous* as we are quantifying into indirect contexts of different levels. Prima facie, the hierarchy view threatens to render inferences involving heterogenous quantification illegitimate. After all it treats ‘that Plato is wise’ as referring to the direct sense of ‘Plato is wise’ in (P1) but as referring to the distinct singly indirect sense of ‘Plato is wise’ in (P2). Burge’s appeal to the functor ‘ $C(x)$ ’, however, allows him to account for such inferences. He envisages his formal language to be governed by a composition principle which ensures that ‘ $C(\langle \text{Plato} \rangle)C(\langle \text{is wise} \rangle)$ ’ expresses the same sense as ‘ $C(\langle \text{Plato} \rangle \langle \text{is wise} \rangle)$ ’. The premisses of the argument can then be formalised as follows:

- (P1’)  $\text{Anna}_B \text{ believes}_B \langle \text{Plato} \rangle \langle \text{is wise} \rangle$   
(P2’)  $\text{Berta}_B \text{ believes}_B \langle \text{Anna} \rangle \langle \text{believes} \rangle C(\langle \text{Plato} \rangle \langle \text{is wise} \rangle)$

By quantifying into the functor ‘ $C(x)$ ’ the inference can then be preserved:

- (C’) So,  $\exists x(\text{Anna}_B \text{ believes}_B x \ \& \ \text{Berta}_B \text{ believes}_B \langle \text{Anna} \rangle \langle \text{believes} \rangle C(x))$

## 2.2 Problems for Burge

Burge’s theory deals with Davidson and quietens qualms about quantification. But it also faces two problems both of which arise from the fact that his canonical names for direct senses have to be treated as primitive. To bring out the first problem let us have a closer look at Burge’s conception of canonicity. According to (CS) you cannot understand the singly indirect sense of ‘is wise’ without understanding its direct sense and you cannot think the former without thinking the latter.

Unfortunately, neither the notion of ‘understanding a sense’ nor that of ‘thinking a sense’ is particularly clear, and both are left unexplained by Burge. Suppose ‘is wise’ and ‘is sapient’ express the same sense. You might think that understanding either one of these expressions is sufficient for understanding the sense they express and that something similar applies to thinking their sense.

On this reading of ‘understanding/thinking a sense’ (CS) doesn’t seem to give us the whole story about indirect senses. For surely we want to say that understanding ‘is wise’ in a singly indirect context is sufficient for understanding this particular expression in a direct context, not just some expression or another which expresses the same sense. We can spell out this phenomenon more generally:

### (DT) **Downward Transparency**

Understanding an expression  $E$  following  $n$  indirect context creating operators is sufficient for understanding  $E$  following  $m$  such operators, for  $n > m$ .

A successful account of the hierarchy ought to explain why (DT) holds despite the fact that expressions are supposed to express different senses in the respective contexts. Can Burge’s theory do this? First note that (DT) has been formulated with an eye

to natural language and doesn't apply to Burge's formalization since Burge translates, for example, direct occurrences of 'is wise' as 'is wise<sub>B</sub>' but singly indirect occurrences as '<is wise>'. To accommodate the idea underlying (DT), Burge thus needs to claim that although both 'is wise<sub>B</sub>' and 'is sapient<sub>B</sub>' express the sense canonically named by '<is wise>', understanding '<is wise>' is sufficient for understanding 'is wise<sub>B</sub>' rather than 'is sapient<sub>B</sub>'. But how is this supposed to come about given that '<is wise>' is primitive? As we have seen '<is wise>' isn't really built up from 'is wise<sub>B</sub>' any more than it is built up from 'is sapient<sub>B</sub>'. It contains the string of letters 'is wise' only in the way in which 'Burge' contains the expression 'urge'.

Note that this problem only arises with regard to the step from direct to singly indirect contexts. The canonical name for the doubly indirect sense of 'is wise' - '*C*(<is wise>)' - *does* contain '<is wise>' as a semantically relevant component. Hence it is not surprising that understanding '*C*(<is wise>)' is sufficient for understanding '<is wise>'.

The second problem is almost the converse of the first problem. It is not *exactly* the converse because it is not clear whether the *exact* converse of (DT) is plausible. Burge thinks that at a certain stage of development a child might be able to understand 'Plato is wise' in direct contexts but not in indirect contexts (cp. 2004, p. 175). But in any case a principle very close to the converse of (DT) holds. This is best brought out by considering an example from Kripke (2008, p. 184 f.):

Suppose a German (who, we might add, does understand belief ascriptions in German) is taking her first English lesson. She has been taught that the sense expressed by 'Plato is wise' is the same as that expressed by 'Platon ist weise' and that the sense expressed by 'Anna believes that ...' is the same as that of 'Anna glaubt, dass ...'. At this point she should have no problem to understand 'Anna believes that Plato is wise'. But how can this be if the expressions shift their sense in indirect contexts? Kripke envisages the pupil saying: 'Unfortunately you have not taught me how to understand these 'belief-sentences' [...] For these I need not only to know the senses of the English words, as I have just been taught, but also their indirect senses.' (2008, pp. 184-5)

A successful Fregean theory of indirect contexts has to explain why the pupil's complaint is unjustified. More generally it has to explain why the following holds:

(UT)      **Upward Transparency**

Given one masters indirect contexts in general, being taught what sense an expression *E* expresses following *n* indirect context creating operators is sufficient for understanding *E* following *m* such operators, for  $n < m$ .

Burge's theory can't account for (UT) either. On the contrary, if his formalisation were to reveal the real structure of natural language belief ascriptions, we should expect the student to react just as she does in Kripke's example. Since '<Plato>' and '<is wise>' are primitive and do not in fact contain 'Plato<sub>B</sub>' and 'is wise<sub>B</sub>' as semantically relevant components, understanding the latter pair of expressions doesn't help you to understand the former.

Again, this problem only concerns the step from direct to singly indirect contexts. Burge can explain why a student who is taught the singly indirect sense of 'Plato is wise' ought to understand it in all indirect contexts of a higher level. If you can understand



‘ $C(\langle \text{Plato} \rangle)$ ’ then you can also understand all canonical names of a higher level, such as ‘ $C(C(\langle \text{Plato} \rangle))$ ’, as these result from iterating a functor you already understand.

Burge’s theory thus suffers from a serious defect, as it cannot explain why (DT) and (UT) hold in general. To be clear, my objection at this point is that while Burge’s appeal to the functor ‘ $C(x)$ ’ can explain e.g. why understanding an expression in a doubly indirect context is sufficient for understanding this expression in a singly indirect context and vice versa, this explanation doesn’t transfer to the step from a singly indirect to direct contexts. This is because the formalisation of an expression in a singly indirect context cannot be seen (for the reason given in §2.1) as the result of applying a functor, such as ‘ $C(x)$ ’, to the formalisation of that expression in a direct context, but must be regarded as primitive. My objection is not that it is impossible to combine Burge’s functor theory with an account that might explain why (DT) and (UT) hold for the missing step. Indeed I think that this might well be possible but that such a repair is undesirable for other reasons, as I will argue in §2.5.

### 2.3 Kripke’s hierarchy view

Kripke agrees with Burge that indirect enjoy a certain privilege. However, he characterises this privilege somewhat differently. He calls indirect senses *immediately revelatory* which he spells out as follows (cp. 2008: p. 187-89):

(IR)      **Immediately Revelatory Senses**

Anyone who understands an expression that expresses an immediately revelatory sense must immediately know what this expression refers to.

Unlike Burge, for whom only indirect senses are canonical, Kripke takes some direct senses to be immediately revelatory too. While you can understand ‘the number of planets’ without immediately knowing what number this expression refers to, this is different with the expression ‘eight’ according to Kripke. What exactly does Kripke mean when he says that understanding an expression with an immediately revelatory sense enables you to ‘know what it refers to’? One can hardly understand ‘the number of planets’ without knowing that it refers to the number of planets, but this doesn’t count as knowing what this expression refers to in the sense intended by Kripke. At this point it is helpful to consult Künne who has characterised expressions such as ‘eight’ and ‘that Plato is wise’ very similarly already in his (1983, p. 172). In his (2010b) discussion of Kripke, Künne illustrates the crucial point as follows (cp. p. 539): upon hearing ‘the number of Ursula’s children is seven’ a competent speaker cannot sensibly answer ‘I have no idea which number that is’ while this reply is acceptable if ‘seven’ is replaced with ‘the number of deadly sins’. Similarly, upon being told ‘Anna believes that Plato is wise’ you cannot sensibly answer ‘I have no idea which thought Anna is said to believe’. This is different when ‘that Plato is wise’ is replaced with ‘John’s favourite thought’.<sup>2</sup>

---

<sup>2</sup>Note that in the sentence resulting from the second completion, ‘John’s favourite thought’ has to be regarded as having its direct sense and reference. So the example really brings out a difference between the immediately revelatory indirect sense of ‘Plato is wise’ and the non-revelatory direct sense of ‘John’s favourite thought’. It also shows that Frege must regard ‘believes’ only as creating an indirect contexts when it is followed by a that-clause or an expression that functions as a sentence in direct contexts.

Burge and Kripke's theories differ more drastically when it comes to tackling the learnability problem. Unlike Burge, Kripke draws on the fact that on Frege's theory all expressions exhibit a systematic ambiguity akin to that commonly associated with indexical expressions.

Consider the sentence 'Romney and Obama ran for Presidency and the latter won.' In order to know what 'the latter' refers to, all you need to know is its linguistic context and the following rule:

- (R1) When 'the latter' occurs after some objects have been mentioned it refers to the object mentioned last.

According to Kripke, on Frege's theory *all* expressions work in a similar way to indexicals such as 'the latter'. For in order to know what a particular expression refers to in an indirect context all you need to know is its linguistic context and a general rule which we can formulate as follows (cp. 2008: p. 196):

- (R2) When an expression follows  $n$  indirect context creating operators it refers to the sense which it expresses when it follows  $n - 1$  such operators (for  $n \geq 1$ ).

Note that (R2) is a rule about what words *refer* to in indirect contexts. The learnability problem arises from expressions being associated with infinitely many *senses*. You might wonder how a rule about *reference* can help with that. Here is how Kripke thinks it does:

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 [(R2)] 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. (2008, p. 196)

To see how this works in detail, let us approach the learnability problem via (UT). Kripke's account explains why being taught the direct sense of 'Plato' and 'is wise' is sufficient for understanding these expressions in a singly indirect context. When you are confronted with 'Anna believes that Plato is wise', you can apply the general rule (R2) to determine the referents of 'Plato' and 'is wise' in the process of which you are given their singly indirect sense. So Kripke has no problem to explain why (UT) holds for the step that caused trouble for Burge.

The theory can also explain why (UT) holds in general. Kripke tells us that the rule (R2) is 'intended to be understood iteratively' (2008: p. 196) so that when (R2) is 'recursively applied [it] gives the doubly indirect sense, and so on' (ibid.). Thus, when confronted with 'Berta believes that Anna believes that Plato is wise' a second application of (R2) tells you that the words in 'Anna believes that Plato is wise' refer to the senses they express in a direct occurrence of 'Anna believes that Plato is wise'. Thus 'Anna' and 'believes' refer to their direct sense and 'Plato' and 'is wise' refer to their singly indirect senses, the senses given by the first application of (R2). The second application thus gives you the singly indirect senses of 'Anna' and 'believes' and the doubly indirect senses of 'Plato' and 'is wise'.

By accounting for (UT) Kripke's theory also deals with Davidson's learnability problem. If you understand all expressions in direct contexts and master the general

rule (R2) then you can understand any number of sentences that result from iterating indirect context creating operators.

## 2.4 Problems for Kripke

Accounting for (DT) proves harder than accounting for (UT). Kripke appears to think that postulating indirect senses to be immediately revelatory is sufficient: he makes this postulation immediately after having dismissed Heidelberger's (1975) account of indirect senses on the ground that it fails to account for an instance of (DT).

According to (IR), understanding 'is wise' in a singly indirect context is sufficient for knowing which sense it refers to, where this knowledge is supposed to preempt inquiries of the form 'which sense exactly is that?'. Does the possession of such knowledge entail that one understands 'is wise' in a direct context? If it does, and if Kripke and Kühne are right that we are familiar with immediately revelatory senses from the direct level, then the postulation indeed offers an explanation as to why (DT) holds. A phenomenon that might initially seem mysterious upon introducing indirect senses is explained by their having a familiar property.

Unfortunately, having such knowledge does *not* entail understanding 'is wise' in a direct context. Suppose you (a) know that an indirect occurrence of 'is wise' refers to the direct sense of 'is sapient' while (b) understanding 'is sapient' in a direct context. In these circumstances, it seems, you cannot wonder which sense it is that 'is wise' refers to. Any enquiry of the form 'which sense exactly is that?' would be inappropriate. But you can be in this position without understanding 'is wise'. So simply postulating that indirect senses are immediately revelatory can't explain why (DT) holds.

At least not on its own. Fortunately, Kripke's framework allows us to say more than he actually does. Consider the mechanism by which an occurrence of 'is wise' is supposed to be understood when it expresses its singly indirect sense. When you are confronted with 'Anna believes that Plato is wise' you are given this sense by determining the expression's referent via an application of the rule (R2). Now, this application tells you that 'is wise' refers to the sense that 'is wise' expresses in a direct context. It doesn't tell you that 'is wise' refers to the sense that 'is sapient' expresses in a direct context. So we can cite the fact that the indirect sense of 'is wise' is immediately revelatory *plus* the fact that this sense is given by an application of (R2) which explicitly mentions 'is wise', rather than 'is sapient', in order to account for (DT).

While (DT) can be dealt with in this way, another problem is more recalcitrant. Recall the inference involving heterogenous quantification:

- (P1) Anna believes that Plato is wise.
- (P2) Berta believes that Anna believes that Plato is wise.
- (C) So,  $\exists x(\text{Anna believes } x \ \& \ \text{Berta believes Anna believes } x)$

Burge, remember, accounts for the legitimacy of this inference by formalising 'that Plato is wise' in (P1) and (P2) as '<Plato><is wise>' and ' $C(\text{<Plato><is wise>})$ ' respectively and quantifying into the functor ' $C(x)$ '. Kripke's theory, however, dispenses with the whole idea of there being a hidden functor to be revealed by formalisation. The learnability of infinitely many expressions with different senses

is ensured by the iterative application of the general rule (R2) rather than by the iteration of a functor, such as ' $C(x)$ '. While this allows Kripke's theory to account for (UT) and (DT), it leaves us empty-handed when it comes to accounting for the validity of inferences involving heterogenous quantification. According to Kripke you are given the sense that 'Plato is wise' expresses in (P2) by an application of (R2) which tells you that this expression refers in (P2) to the sense 'Plato is wise' expresses in (P1). This accounts for there being *some* connection between the referents of these words in (P1) and (P2), but none that would justify the validity of the inference.

To summarise, Burge and Kripke both deal with Davidson. While Burge additionally deals with heterogenous quantification he fails to explain why (DT) and (UT) hold. In contrast, Kripke accounts for both (DT) and (UT) but not for heterogenous quantification. Neither can pride himself with having vindicated the hierarchy view in general.

## 2.5 Comparing the options

At this point one might try to somehow combine Burge and Kripke's theories in order to avoid their respective shortcomings. One might claim that the senses of expressions in singly indirect contexts are given by an application of Kripke's rule (R2) while still applying Burge's hidden functor theory from doubly indirect contexts onwards. And one might hope to thereby combine the ability of Kripke's account to deal with the (DT) and (UT) with the ability of Burge's account to deal with heterogeneous quantification.

Such a *combined theory* seems, however, quite unattractive. It tells two rather distinct stories where one would have expected a unified account. At the very least, defending the infinite hierarchy thus proves to be significantly more cumbersome than Burge and Kripke appreciate. This seems even more unfortunate once we realise that the hierarchy not only makes things more complicated but is theoretically pointless when it comes to the original motivation for Frege's theory.

The reference shift in singly indirect contexts is useful because it explains why 'Plato' and 'Aristocles' are interchangeable *salva veritate* in direct contexts but not in singly indirect contexts while leaving Frege's doctrine of extensionality intact. This works because expression with the same direct referent can express different senses and therefore have different singly indirect referents.

In contrast to this, the combined theory will have to maintain that any two expressions that have the same singly indirect referent will also have the same doubly indirect referent, the same triply indirect, and so on. This is because the combined theory still needs to incorporate Burge's hidden functor theory in order to deal with heterogenous quantification. We have seen in §2.1 that this theory requires there to be a one-one relation between indirect senses and the things they present. But this means that any two expressions with the same direct sense have the same singly indirect sense, doubly indirect sense, and so on.

So on the combined theory any two expressions that are interchangeable in singly indirect contexts remain interchangeable in all indirect contexts of a higher level. But then why not simply say that they keep referring to the same sense in all of these contexts?

In order for the additional reference shifts to be useful in the way in which the initial reference shift is useful there would have to be expressions with the same referents in a singly indirect context but different referents in an indirect context of a higher level. Although this is explicitly ruled out on Burge's official theory, he says in a footnote that his theory 'can be conceived in such a way as to accommodate that position' (2004, p. 172, fn. 12). It is not clear how exactly this is supposed to work with regard to Burge's original theory and it is even less clear with regard to the combined theory. Suppose two expressions, 'a' and 'b', have the same direct sense but a different singly indirect sense. The combined theory would have to maintain that applying (R2) to a singly indirect occurrence of 'a' gives you one sense while applying the same rule to 'b' gives you another, despite the fact that each application determines the same referent, the direct sense of 'a' and 'b'. How is this supposed to come about? It cannot be explained by the first application's mentioning a different expression than the second, for then we could not allow any two distinct expressions to have the same singly indirect sense. But in virtue of what are the two applications then supposed to give two different senses rather than the same? In absence of any answer to this question, it seems that the combined theory cannot in fact allow for the additional reference shifts to become theoretically useful.

The two-level view which rejects the spurious reference shifts thus seems clearly preferable. Adopting it, we obviously avoid Davidson's learnability problem. We still need to explain why (DT) and (UT) hold but we can avail itself of a version of Kripke's theory to do so. Of course we need to replace (R2) with its two-level equivalent:

(R3) When an expression follows  $n$  indirect context creating operators it refers to the sense which it express when it occurs in a direct context (for  $n \geq 1$ ).

Burge's hidden functor theory, in contrast, can be dispensed with entirely. Accounting for inferences involving heterogenous quantification becomes utterly straightforward as 'that Plato is wise' now has the same referent in indirect contexts of any order.

What about the one-level view? One might think that this view is preferable even to the two-level view. While it handles the learnability problem and heterogenous quantification with the same ease it additionally makes (DT) and (UT) become trivial and hence doesn't even need to appeal to a modified version of Kripke's theory.

Yet adopting the one-level view comes at a price. First, it requires us to come to terms with the idea of a sense presenting itself. Starting from singly indirect contexts onwards 'Plato' both refers to and expresses its direct sense. This sense thus presents itself.

Second, as has already been mentioned in §1, it forces us to give up the principle that sense determines reference. 'Plato' has the same sense in direct and indirect contexts but different referents. On the one-level view, the sense of an expression determines its reference only in tandem with the expression's linguistic context.

Third, it also requires us to give up the principle that sense determines the logical type of an expression, at least if we disagree with Frege's assimilation of sentences to referring expressions. In a direct context 'Plato is wise' is a sentence. But in indirect contexts it refers to a thought. Yet, according to the one-level view, it has the same sense in both cases. So the one-level view allows expressions with the same sense to

belong to different logical categories.<sup>3</sup> As with reference, the most that can be said is that the sense of an expression determines the expression's logical type in tandem with its linguistic context. Unlike the the second consequence, the third consequence appears to have gone unnoticed by Dummett, the most prominent defender of the one-level view. Dummett calls Frege's lumping sentences and singular terms into the same logical category 'a gratuitous blunder' (1973, p. 184) but doesn't seem to notice the conflict his one-level view creates with a correction of this mistake.

The two-level view has none of these potentially troublesome consequences. While the two-level view is clearly preferable to the hierarchy view, the comparison between the one-level view and the two-level view is not straightforward at this point. The decision would depend on how problematic one regards the three consequences of the one-level view in comparison to the postulation of an additional layer of sense by the two-level view. As it turns out we don't need to make this decision. Burge intends to rule out both the one-level view as well as the two-level view, but, as we shall see, he succeeds only with regard to the former. The two-level view will thus emerge as the sole winner in the competition for the best extension of Frege's theory to iterated indirect contexts.

### 3 How the hierarchy can be avoided

Burge argues that both the one-level view as well as the two-level view lead to absurd consequences when combined with principles that are both 'substantial parts of [Frege's] theory' as well as 'motivated, powerful, and attractive - quite independently of Frege's maintaining them' (2004, p. 167). His direct arguments (2004, Appendix I) attempt to show that rejecting the hierarchy straightforwardly leads to absurdities. His is indirect argument (1979, §I & II; 2004, p. 183 ff.) is more subtle. It attempts to show that unless the hierarchy is accepted for a language containing iterated indirect contexts it is impossible to provide an acceptable truth-theory for that language. In §3.1 I discuss the direct argument against the one-level view and show that it fails. In §3.2 I discuss the direct argument against the two-level view and show that it fails as well. In §3.3 I discuss the indirect argument and show that it is effective only against the one-level view but fails to cause trouble for the two-level view.

#### 3.1 The direct argument against the one-level view

I adapt Burge's notation to the taxonomy of this paper:

- $S('A', D)$ : the sense of 'A' in a direct context
- $S('A', I_1)$ : the sense of 'A' in a singly indirect context
- $S('A', I_2)$ : the sense of 'A' in a doubly indirect context

Burge starts by assuming, for the purpose of a reductio, the characteristic claim of the one-level view:

---

<sup>3</sup>Burge makes a similar point in fn. 27 of the 2004 postscript. He says that from the perspective of the one-level view there is a 'an unclarity about the logical syntax of "believes that p" [...] Insofar as it expresses its customary sense, it is a sentence, insofar as it refers to its customary sense it is a term' (2004, p.187).

$$\text{(One-Level)} \\ S(\text{'Plato'}, D) = S(\text{'Plato'}, I_1)$$

He then considers the following expression:

$$(1) \quad S(\text{'Plato = Plato'}, D)$$

To this expression, Burge intends to apply a principle he calls ‘the decomposability of sense’. More specifically he wants to express that ‘the sense of a sentence occurring in a [direct] context and asserting the self-identity of [Plato] is decomposable into the senses that the component parts of that sentence express in a [direct] context’ (2004, p. 201). To do so he uses the symbol ‘ $\wedge$ ’ that he has stipulated to mean ‘appropriately grammatically composed with’ (cp. 2004, p. 185, fn. 25). Burge then asserts that (1) co-refers with:

$$(2) \quad S(\text{'Plato'}, D) \wedge S(\text{'='}, D) \wedge S(\text{'Plato'}, D)$$

Now Burge claims that, given (One-Level), we can substitute ‘ $S(\text{'Plato'}, I_1)$ ’ for ‘ $S(\text{'Plato'}, D)$ ’ without changing the referent of (2). Thus (2) co-refers with:

$$(3) \quad S(\text{'Plato'}, D) \wedge S(\text{'='}, D) \wedge S(\text{'Plato'}, I_1)$$

Burge is aware that he is thereby assuming that ‘ $\wedge$ ’ creates a direct context so that we can inter-substitute expressions with the same referent but different senses within its scope. Burge proceeds by introducing a name ‘ $a$ ’ that he stipulates to refer to the direct sense of ‘Plato’ in a direct context and allows to express any sense we like as long as it is compatible with this reference.

He then again relies on the assumption that ‘ $\wedge$ ’ creates a direct context in asserting that we can substitute ‘ $a$ ’ for ‘ $S(\text{'Plato'}, I_1)$ ’ so that (3) co-refers with:

$$(4) \quad S(\text{'Plato'}, D) \wedge S(\text{'='}, D) \wedge a$$

Finally, Burge intends to exploit ‘the compositionality of sense’ in order to establish that (4) co-refers with:

$$(5) \quad S(\text{'Plato = } a', D)$$

This would establish that (1) and (5) co-refer, which means that ‘Plato = Plato’ and ‘Plato =  $a$ ’ express the same sense in direct contexts. This is absurd given that ‘ $a$ ’ was stipulated to refer to a sense, not to Plato. Therefore, says Burge, we must reject the assumption (One-Level).

This argument is fallacious. Whatever is responsible for the absurdity it culminates in, it is not the assumption Burge intends to disparage. (One-Level) was only appealed to in justifying the co-referentiality of (2) and (3). But this step is unnecessarily circuitous. Given that Burge assumes ‘ $\wedge$ ’ to create a direct context anyway he could have moved as well from (2) straight to (4) omitting the detour via (3) and thereby any mention of (One-Level). So this argument fails to cause any trouble for the one-level view.

If (One-Level) is not responsible for the absurd conclusion then what is? It’s the final step from (4) to (5) via ‘the compositionality of sense’. Burge explains this

move as follows: ‘the sense (or thought) composed appropriately of the senses of the semantically relevant parts of a sentence is identical with the sense (or thought) expressed by the whole sentence’ (2004: p 201). If ‘appropriately’ combining the senses of the semantically relevant parts of a sentence means combining them in a way that yields the sense of the sentence in question - and I can’t see what else it could mean - then this principle is a truism. But this truism in no way justifies the co-referentiality of (4) and (5). Trivially, the sense denoted by (5), i.e. the direct sense of ‘Plato = *a*’, is the result of appropriately combining the direct senses *expressed by* ‘Plato’, ‘=’ and ‘*a*’. But (4) doesn’t refer to this combination of senses. It refers to the result of combining the senses *referred to* by ‘ $S(\text{‘Plato’}, D)$ ’, ‘ $S(\text{‘=’}, D)$ ’ and ‘*a*’.

Now, the senses referred to by ‘ $S(\text{‘Plato’}, D)$ ’ and ‘ $S(\text{‘=’}, D)$ ’ just are the sense expressed by ‘Plato’ and ‘=’. But ‘*a*’ was stipulated to refer to the direct sense of ‘Plato’ and to express *any sense we like*.

Perhaps, Burge didn’t actually mean to be so tolerant about the sense of ‘*a*’. Perhaps, he really wanted to stipulate that in a direct context ‘*a*’ both refers to and expresses the direct sense of ‘Plato’. If the one-level theorist had to accept this stipulation as coherent she would indeed be committed to infer that ‘Plato = Plato’ expresses the same sense as ‘Plato = *a*’. After all, on this new stipulation ‘*a*’ expresses in a direct context the same sense as ‘Plato’ and we should always be able to inter-substitute co-sensical expressions without changing the sense of a sentence.

So this new stipulation would make Burge’s argument valid and much simpler as well. But unlike the original stipulation the one-level theorist will reject this new stipulation as incoherent. As we have seen in §2.5, qua one-level theorist she will maintain that sense determines reference only in tandem with linguistic context. Thus, according to her, the sense of ‘Plato’ is such that it presents Plato when expressed by an expression in a direct context and presents itself when expressed by an expression in an indirect context. So when ‘*a*’ is stipulated to express in a direct context the direct sense of ‘Plato’, it is stipulated to express a sense that presents Plato in a direct context. From the perspective of the one-level theory Burge is then no longer free to stipulate at the same time that in a direct context ‘*a*’ refers not to Plato but to the direct sense of ‘Plato’. To illustrate this consider the sentence ‘Obama and Romney ran for President and *x* won’. You cannot coherently stipulate that ‘*x*’ expresses the sense of ‘the former’ and then go on to stipulate that it refers to Romney in this linguistic context. But from the perspective of the one-level theory, the new stipulation regarding ‘*a*’ is similarly incoherent. Burge’s direct argument against the one-level view fails.

### 3.2 The direct argument against the two-level view

What about the direct argument against the two-level view? Burge starts by assuming for reductio that an expression expresses the same sense in singly indirect and doubly indirect contexts:

$$\begin{aligned} & \text{(Two-Level)} \\ & S(\text{‘Plato’}, I_1) = S(\text{‘Plato’}, I_2) \end{aligned}$$

Burge now stipulates an expression ‘*B*’ to refer in direct contexts to the direct sense of ‘Plato’ by expressing the singly indirect sense of ‘Plato’. This stipulation need not be



incoherent from the perspective of a two-level view, since this view doesn't hold that sense determines reference only in tandem with linguistic context.

We can now present Burge's argument by first considering the expression:<sup>4</sup>

$$(1) \quad S('B=B', D)$$

This expression denotes the direct sense expressed by the sentence ' $B=B$ ' which says of the direct sense of 'Plato' that it is identical with itself. Since ' $B$ ' is stipulated to express the singly indirect sense of 'Plato' in a direct context, the sense expressed by ' $B=B$ ' must be the result of composing the singly indirect sense of 'Plato' with the direct sense of '=' and again with the singly indirect sense of 'Plato'. Thus (1) co-refers with:

$$(2) \quad S('Plato', I_1) \wedge S( '=', D) \wedge S('Plato', I_1)$$

Now Burge's argument continues in a similar fashion to the previous one. First he points out that given (Two-Level), we can substitute ' $S('Plato', I_2)$ ' for ' $S('Plato', I_1)$ ' in (2) without changing the referent of the expression. Thus (2) co-refers with:

$$(3) \quad S('Plato', I_1) \wedge S( '=', D) \wedge S('Plato', I_2)$$

Then Burge again proceeds by introducing a new name, ' $b$ ', which is stipulated to refer, in a direct context, to the singly indirect sense of 'Plato' and is allowed to express any sense compatible with that assumption. Burge, again relying on the assumption that '^' creates a direct context, asserts that we can substitute ' $b$ ' for ' $S('Plato', I_2)$ ' so that (3) co-refers with:

$$(4) \quad S('Plato', I_1) \wedge S( '=', D) \wedge b$$

Appealing to the 'compositionality of sense' Burge intends to infer that (4) co-refers with (5):

$$(5) \quad S('B=b', D)$$

This would establish that ' $B=B$ ' and ' $B=b$ ' express the same sense which the two-level theorist cannot accept because she takes the former to be true but the latter to be false. ' $B$ ' was stipulated to refer to the direct sense of 'Plato' while ' $b$ ' was stipulated to refer to its singly indirect sense which according to the two-level view are distinct.

It is not hard to see that this argument suffers from very similar defects as the previous one. Given the assumptions about '^' and ' $b$ ' any appeal to (Two-Level) in the derivation is an unnecessary detour. And just as before the real culprit for the unacceptable conclusion is the move from (4) to (5) which is not justified by any reasonable compositionality principle.

---

<sup>4</sup>Burge (2004, p. 202 ff.) sets up the argument slightly differently. He starts by considering the expression ' $S(S('Plato', I_1) = S('Plato', I_1), D)$ ' and asserts that it co-refers with (2). This is both false and unnecessary for his argument. It is false because the sense expressed ' $S('Plato', I_1)$ ' is not the sense referred to by ' $S('Plato', I_1)$ '. The singly indirect sense of 'Plato' is not the direct sense of the definite description 'the sense of 'Plato' in an indirect context'. It is unnecessary because Burge goes on to introduce the expression ' $B$ ' with the stipulation described above. Since ' $B$ ' is stipulated to express the sense denoted by ' $S('Plato', I_1)$ ', ' $S('B=B', D)$ ' does co-refer with (2) and Burge could have started his argument from this point as I have done.

Again, one wonders whether Burge actually meant to stipulate that ‘*b*’ refers, in a direct context, to the singly indirect sense of ‘Plato’ in virtue of expressing that very sense. Given this stipulation the two-level theorist would indeed be committed to assert that ‘ $B = B$ ’ expresses the same sense as ‘ $B = b$ ’. After all ‘*B*’ and ‘*b*’ are now stipulated to express the same sense in a direct context.

As with the first argument, this new stipulation would make Burge’s second argument valid and much simpler. But now the two-level theorist doesn’t have to accept this new stipulation as coherent. For the sense of ‘*b*’ is now stipulated to present itself. As we have seen in §2.5, it is an advantage of the two-level theory that it need not countenance senses that present themselves. So a proponent of the two-level theory is free to reject the idea of senses presenting themselves as incoherent, as incidentally Burge does himself (cp. 2004, p. 190), and to reject the new stipulation on this ground.

### 3.3 The indirect argument for the hierarchy

Burge thinks that his direct arguments rule out both the one-level view and the two-level view. We have seen that in fact they don’t rule out either. But Burge has a further, less direct, argument for the hierarchy. He argues that unless we embrace the hierarchy we will not be able to provide a truth-theory for our language, such that the metalanguage in which this theory is formulated is capable of translating object language sentences containing iterated indirect contexts. The argument consists of two parts. Burge (1979, § III) argues that the metalanguage of a suitable truth-theory must support a certain substitution principle. This argument is presented in §3.3.1. He also argues that a metalanguage that supports such a substitution principle is incapable of translating sentences such as ‘Berta believes Anna believes Plato is wise’ unless the hierarchy is accepted (1979, § II). This argument is presented in §3.3.2. Finally, I show in §3.3.3 how Burge’s reasoning can be resisted.

#### 3.3.1 The substitution principle

Burge sketches a truth-theory that allows us to derive a truth-condition for sentences such as ‘Berta believes Anna believes Plato is wise’. I will follow Burge in rewriting this sentence in function-argument form as (OL1) and in adding the subscript ‘*OL*’ to object language expressions and ‘*ML*’ to metalanguage expressions where disambiguation is required:

(OL1) believes(Berta, believes(Anna, is wise(Plato)))<sub>OL</sub>

Burge assumes that in (OL1) ‘Plato<sub>OL</sub>’ refers to its direct sense, rather than to its singly indirect sense, since the second part of his argument attempts to perform a reductio on this assumption shared by the one-level and two-level view. His truth-theory then contains reference axioms that can be presented as follows, where ‘d.ref’ abbreviates ‘direct referent of’, ‘i.ref’ abbreviates ‘indirect referent of’, ‘sns’ abbreviates ‘sense of’ and ‘*e*<sub>1</sub>’ and ‘*e*<sub>2</sub>’ are variables ranging over object language expressions:<sup>5</sup>

---

<sup>5</sup>I have adjusted the truth-theory to my example sentence and simplified it in some respects. Note also that my labeling of the axioms differs from Burge’s.

- (a) d.ref 'Berta<sub>OL</sub>' = Berta<sub>ML</sub>
- (b) i.ref 'Anna<sub>OL</sub>' = sns 'Anna<sub>OL</sub>'
- (c) i.ref 'Plato<sub>OL</sub>' = sns 'Plato<sub>OL</sub>'
- (d) i.ref  $\ulcorner$  is wise( $e_1$ )<sub>OL</sub>  $\urcorner$  =  $C_2$ (sns 'is wise<sub>OL</sub>', i.ref  $e_1$ )
- (e) i.ref  $\ulcorner$  believes( $e_1, e_2$ )<sub>OL</sub>  $\urcorner$  =  $C_3$ (sns 'believes<sub>OL</sub>', i.ref  $e_1$ , i.ref  $e_2$ )

The last two axioms require explanation. ' $C_2$ (sns 'is wise<sub>OL</sub>', i.ref  $e_1$ )' is supposed to refer to the result of composing the sense of 'is wise<sub>OL</sub>' with the sense which is the indirect referent of the expression  $e_1$ . The result of this composition is of course intended to be the sense expressed by  $\ulcorner$  is wise( $e_1$ )<sub>OL</sub>  $\urcorner$ . This is ensured by the reference axiom for the relevant object language expression plus the following additional axiom:

$$(f) \quad C_2(\text{sns } e_1, \text{sns } e_2) = \text{sns } (\ulcorner e_1(e_2) \urcorner)$$

Similar remarks apply (e) and require the addition of the axiom:

$$(g) \quad C_3(\text{sns } e_1, \text{sns } e_2, \text{sns } e_3) = \text{sns } (\ulcorner e_1(e_2, e_3) \urcorner)$$

Finally, Burge's truth-theory contains an axiom stating the truth-condition for an atomic sentence that results from completing the two-place predicate 'believes<sub>OL</sub>' with two expressions:

$$(h) \quad \text{True}(\ulcorner \text{believes}_{OL}(e_1, e_2) \urcorner) \leftrightarrow \text{believes}_{ML}(\text{d.ref } e_1, \text{i.ref } e_2)$$

With these axioms the following truth-condition for (OL1) can be derived (see appendix A):

$$(TC1) \quad \begin{aligned} &\text{True}(\ulcorner \text{believes}(\text{Berta}, \text{believes}(\text{Anna}, \text{is wise}(\text{Plato}))) \urcorner_{OL}) \\ &\leftrightarrow \text{believes}_{ML}(\text{Berta}_{ML}, \text{sns}(\ulcorner \text{believes}(\text{Anna}, \text{is wise}(\text{Plato}))) \urcorner_{OL}) \end{aligned}$$

Now, crucially, this derivation requires inter-substituting expressions with the same direct referent but different senses in the second place of the metalanguage predicate 'believes<sub>ML</sub>( $x, y$ )'. For instance, at one point we have to substitute the expression

$$C_3(\text{i.ref 'believes}_{OL}', \text{i.ref 'Anna}_{OL}', \text{i.ref 'is wise}(\text{Plato})_{OL}')$$

for the expression

$$\text{i.ref 'believes}(\text{Anna}, \text{is wise}(\text{Plato}))_{OL}'.$$

This means that 'believes<sub>ML</sub>( $x, y$ )', unlike the object language predicate which it translates, cannot create an indirect context in its second place. If it did, the required substitutions would be illegitimate.

In fact Burge goes further in claiming that a suitable metalanguage has to eliminate indirect contexts *tout court* so that all expressions with the same direct referent can be inter-substituted for one another in *any* context. However, all that is required to get his indirect argument going is the assumption that all co-referential expressions can be inter-substituted in the second argument place of 'believes<sub>ML</sub>( $x, y$ )'. We will call this assumption (Sub).

### 3.3.2 From the substitution principle to the hierarchy?

Burge is unhappy with the truth-condition (TC1):

[(TC1)] would not explain the truth conditions of [(OL1)] to someone who understood [the metalanguage] but who did not already understand the expressions of [the object language] (Did not understand what their senses are). Although [TC1] may perhaps describe the truth conditions it does not ‘give’ them. To give the truth conditions we need a sentence on the right hand side of the biconditional which states the relevant truth conditions purely in the terms of [the metalanguage] without mentioning expressions of [the object language]. That is we need a plausible translation of [(OL1)] into [the metalanguage]. From this viewpoint, an adequate theory of truth for [the object language] must satisfy Tarski’s convention T, which requires the metalanguage in which the theory is given provide a translation of the relevant sentence of [the object language]. (1979: p. 165)

Now Burge believes that any attempt to provide such an adequate truth-theory without accepting the hierarchy is doomed to fail. I will first outline how such an attempt might look like and then present Burge’s objection to it.

We will translate direct occurrences of object language expressions - ‘is wise(Plato)<sub>OL</sub>’ - by their homophonic metalanguage counterparts: ‘is wise(Plato)<sub>ML</sub>’. We cannot then go on to translate the sentence ‘believes(Anna, is wise(Plato))<sub>OL</sub>’ simply as ‘believes(Anna, is wise(Plato))<sub>ML</sub>’. Since ‘believes<sub>ML</sub>( $x, y$ )’, unlike ‘believes<sub>OL</sub>( $x, y$ )’, creates direct contexts in both argument places ‘believes(Anna, is wise(Plato))<sub>ML</sub>’ doesn’t assert that the relation of belief holds between Anna and the indirect referent of is ‘wise(Plato)<sub>ML</sub>’, the thought that Plato is wise. If anything, it asserts that the relation of belief holds between Anna and the direct referent of ‘wise(Plato)<sub>ML</sub>’ which is a truth-value.

Whatever our stance on the hierarchy, our metalanguage has to make reference shifts explicit by translating a singly indirect occurrence of ‘is wise(Plato)<sub>OL</sub>’ by means of an expression which in a direct context refers to the direct sense of ‘is wise(Plato)<sub>OL</sub>’ and expresses whatever sense ‘is wise(Plato)<sub>OL</sub>’ expresses in a singly indirect contexts. We have already encountered an expression stipulated to have just these characteristics in §2.1: the canonical name ‘<Plato><is wise>’. We will use this expression, or rather the result of putting it into function-argument form, ‘<is wise>( <Plato> )’, as our metalanguage translation of singly indirect occurrences of ‘is wise(Plato)<sub>OL</sub>’. More generally we will translate all singly indirect occurrences of object language expression by means of the canonical names for their direct senses.

Now, for a reason we encountered in §3.1 a one-level theorist cannot be happy with this. From the perspective of the one-level view a direct occurrence of ‘<is wise>( <Plato> )’ would have to both denote and express the direct sense of ‘is wise(Plato)<sub>OL</sub>’. But since a one-level theorist believes that sense determines reference only in tandem with linguistic context, ‘<is wise>( <Plato> )’ can only express the direct sense of ‘is wise(Plato)<sub>OL</sub>’ if it expresses a sense which in a direct context presents a truth-value, not the direct sense of ‘is wise(Plato)<sub>OL</sub>’.

It is thus not clear at all how a one-level theorist could provide an adequate truth-theory. Burge's indirect argument indeed rules out the one-level view, although perhaps not for the reason he had in mind.

In contrast, Burge's indirect argument doesn't as yet put any pressure on the two-level view. From the perspective of the two-level theory there is no problem in assuming that a direct occurrence of '<is wise>(<Plato>)' refers to the direct sense of 'is wise(Plato)<sub>OL</sub>' and expresses this expression's distinct singly direct sense. Moreover, since a two-level view rejects any further reference shifts we can use the canonical names to translate the corresponding object language expressions not only in singly indirect contexts but in indirect contexts of any level.

Can't we then modify Burge's original truth-theory so that it becomes adequate by his own standards? We replace Burge's reference axioms (b), (c), (d), and (e) by the following axioms which specify the reference of the object language expressions by means of the new metalanguage expressions:

- (b') i.ref 'Anna<sub>OL</sub>' = <Anna>
- (c') i.ref 'Plato<sub>OL</sub>' = <Plato>
- (d') i.ref 'is wise(*e*<sub>1</sub>)<sub>OL</sub>' =  $C_2(\langle \text{is wise} \rangle, \text{i.ref } e_1)$
- (e') i.ref 'believes(*e*<sub>1</sub>, *e*<sub>2</sub>)<sub>OL</sub>' =  $C_3(\langle \text{believes} \rangle, \text{i.ref } e_1, \text{i.ref } e_2)$

We now need an axiom ensuring that the result of composing the senses referred to by '<is wise>' and '<Plato>' is the sense referred to by '<is wise>(<Plato>)'. So we replace (f) with:

$$(f') \quad C_2(e_1, e_2) = \ulcorner e_1(e_2) \urcorner$$

where the variables are now understood as also ranging over the newly introduced metalanguage expressions used to translate object language expressions that occur in indirect contexts. Similarly (g) is replaced with:

$$(g') \quad C_3(e_1, e_2, e_3) = \ulcorner e_1(e_2, e_3) \urcorner$$

The remaining two axioms are left untouched. The resulting truth-theory allows us to derive the following truth-condition (see Appendix B):

$$(TC1') \quad \text{True '(\text{believes}(\text{Berta}, \text{believes}(\text{Anna}, \text{is wise}(\text{Plato})_{OL}))' \leftrightarrow \\ \text{believes}_{ML}(\text{Berta}_{ML}, \langle \text{believes} \rangle(\langle \text{Anna} \rangle, \langle \text{is wise} \rangle(\langle \text{Plato} \rangle)))$$

Unlike (TC1), (TC1') appears to 'give' the truth-condition of (OL1) in the sense required by Burge. As required, its right-hand-side doesn't mention expressions of the object language but is a term for term translation of the sentence mentioned on the left-hand-side.

Burge, however, argues that this translation of (OL1) is in fact unacceptable and that an acceptable translation can only be provided once the hierarchy is embraced. His argument can be put as follows:<sup>6</sup>

<sup>6</sup>In Burge's formulation of the argument in (1979, Section II) he uses a different example as an object language sentence and greek letters as metalanguage translations. My '<believes>', '<Anna>', and '<is wise>(<Plato>)' correspond to his ' $\beta_1$ ', ' $T_1$ ', and ' $\alpha$ '; my 'believes<sub>ML</sub>', 'Anna<sub>ML</sub>' and 'is wise(Plato)<sub>ML</sub>' correspond to his ' $\beta$ ', ' $\Gamma$ ', and ' $\alpha_0$ '.

- (1) The expression ‘believes(Anna, is wise(Plato))<sub>OL</sub>’ denotes its direct sense in (OL1). Hence, its suggested translation ‘<believes>(⟨Anna>, <is wise> (⟨Plato>))’ must also denote this sense.
- (2) The expressions ‘believes<sub>ML</sub>’, ‘Anna<sub>ML</sub>’ and ‘is wise(Plato)<sub>ML</sub>’ respectively express the senses denoted by ‘<believes>’, ‘⟨Anna>’ and ‘<is wise>(⟨Plato>)’.
- (3) Therefore, ‘believes(Anna, is wise(Plato))<sub>ML</sub>’ expresses the sense denoted by ‘<believes>(⟨Anna>, <is wise>(⟨Plato>))’.
- (4) But, given (1), this means that ‘believes(Anna, is wise(Plato))<sub>ML</sub>’ expresses the direct sense of ‘believes(Anna, is wise(Plato))<sub>OL</sub>’.

This conclusion is indeed unacceptable. First, we have already seen that ‘believes(Anna, is wise(Plato))<sub>ML</sub>’ *cannot* express the same sense as ‘believes(Anna, is wise(Plato))<sub>OL</sub>’. Since ‘believes(*x*, *y*)<sub>ML</sub>’ doesn’t create an indirect context in its second argument place, the metalanguage expressions asserts that Anna believes a truth-value if it asserts anything at all. Furthermore, (Sub) allows us to substitute any true sentence for ‘is wise(Plato)<sub>ML</sub>’. Thus ‘believes(Anna, is wise(Plato))<sub>ML</sub>’ entails ‘believes(Anna, is wise(Aristocles))<sub>ML</sub>’. And a variant of the above argument throughout which any occurrence of ‘Plato’ is replaced by ‘Aristocles’ will conclude that ‘believes(Anna, is wise (Aristocles))<sub>ML</sub>’ expresses the same sense as the object language sentence ‘believes(Anna, is wise(Aristocles))<sub>OL</sub>’. Hence it seems that we are once again forced to infer from Anna’s believing that Plato is wise that she believes that Aristocles is wise - just what Frege’s theory was meant to avoid.

Note that if we had gone for the hierarchy view then Burge’s argument wouldn’t apply. On the hierarchy view, ‘Plato<sub>OL</sub>’ doesn’t refer to its direct sense but its singly indirect sense in (OL1). Hence in order to translate (OL1) we could not employ ‘<Plato>’ but would need yet a different metalanguage expression, such as Burge’s ‘C(⟨Plato>)’, which is stipulated to refer to the singly indirect sense of ‘Plato<sub>OL</sub>’ and expresses its doubly indirect sense. On the hierarchy view we would thus translate (OL1) as:

$$\text{believes}_{ML}(\text{Berta}_{ML}, \langle \text{believes} \rangle(\langle \text{Anna} \rangle, C(\langle \text{is wise} \rangle)(C(\langle \text{Plato} \rangle))))$$

It is thus the expression ‘<believes>(⟨Anna>, C(⟨is wise>)(C(⟨Plato>)))’ - rather than ‘<believes>(⟨Anna>, <is wise>(⟨Plato>))’ - which denotes the direct sense of the object language sentence ‘believes(Anna, is wise(Plato))<sub>OL</sub>’. And the expression ‘C(⟨is wise>)(C(⟨Plato>))’ is supposed to denote the singly indirect sense of ‘is wise(Plato)<sub>OL</sub>’ rather than its direct sense. It hence doesn’t denote the sense expressed by ‘is wise(Plato)<sub>ML</sub>’. On the hierarchy view Burge’s indirect argument thus doesn’t get off the ground. Burge therefore takes his indirect argument to establish that the hierarchy must be accepted.

### 3.3.3 Why the indirect argument fails

Burge thinks that his indirect argument relies only on principles that are ‘substantial parts of [Frege’s] theory’ and furthermore ‘motivated, powerful, and attractive - quite

independently of Frege's maintaining them' (2004: p. 167).

I believe this assessment rests on a confusion. What justifies step (3) of Burge's argument? One sometimes gets the impression that apart from the principle of extensionality all Burge takes this step to rely only on is the following principle:

(Pr1) The sense of a complex expression is uniquely determined by the senses of its component expressions.

For instance, in his (2004) recapitulation of the argument he cites the 'functional compositionality of sense' (2004: p. 184, 186) which he spells out as saying that the sense of a complex expression 'is functionally dependent on the senses expressed by the constituent parts of the expression' (2004, p. 186).

(Pr1) is indeed a substantial part of Frege's theory as well as independently attractive. But it doesn't justify step (3). (Pr1) tells us that there is a function from the senses expressed by 'believes<sub>ML</sub>', 'Anna<sub>ML</sub>', 'is wise(Plato)<sub>ML</sub>' to the sense expressed by 'believes(Anna, is wise(Plato))<sub>ML</sub>'. The principle of extensionality tells us that there is a function from the referents of '<believes>', '<Anna>', '<is wise>(<Plato>)', to the referent of '<believes>(<Anna>, <is wise>(<Plato>))'. Even given step (2) it doesn't follow from this that the former complex expression expresses what the latter denotes. For nothing has been said to ensure that we are dealing with the same function in the two cases.

In fact, Burge slips in a further assumption that *does* ensure this. Somewhat in passing, and not without later declaring it as inessential, he assumes the following principle (cp. 1979, p. 159):

(Pr2) Every sense of a predicate is a function that takes the senses of the terms completing the predicate as arguments and maps them to the sense of the resulting sentence.

He thus conceives of the sense expressed by 'believes(Anna, is wise(Plato))<sub>ML</sub>' as the value of the function expressed by 'believes<sub>ML</sub>' for the arguments expressed by 'Anna<sub>ML</sub>' and 'is wise(Plato)<sub>ML</sub>'.

Given (Pr2) and the assumption that '<believes>' denotes the sense of 'believes<sub>ML</sub>' it is very natural to regard the referent of '<believes>(<Anna>, <is wise>(<Plato>))' as the value of the function denoted by '<believes>' for the arguments denoted by '<Anna>' and '<is wise>(<Plato>)'. On this conception, '<believes>(<Anna>, <is wise>(<Plato>))' of course denotes the sense expressed by 'believes(Anna, is wise(Plato))<sub>ML</sub>' as we are plugging the same arguments into the same function.

Adopting this conception thus fills the lacuna in Burge's argument. Of course it also means that Burge is in fact not merely relying on (Pr1) but also on (Pr2). But while (Pr1) indubitably is a substantial and independently attractive principle of Frege's theory, it is far from clear that (Pr2) also deserves this label.

First, it is not clear at all whether Frege himself subscribed to (Pr2). Although he occasionally describes the senses of predicates or functors as 'unsaturated', the same word he also uses to describe functions, he never explicitly identifies them with functions (cf. 1892-5, fn. *B*; 1923, p. 541)

Second, and regardless of whether Frege actually would have endorsed (Pr2) it is even less clear whether this would have been a good idea. Dummett (1981: Ch. 13,

Ch. 15) argues emphatically that it wouldn't. He considers it a serious mistake to think of the sense of any expression as a function. In a nutshell, Dummett's point is that if the sense of a predicate were a function from senses of terms to senses of sentences (i.e. thoughts), then grasping the sense of a predicate would presuppose grasping the thoughts in the range of that function which is its sense. This, Dummett argues, would mean that the understanding of sentences would have to be prior to the understanding of their component predicates in a way in which it can't be.

So it turns out that the crucial principle assumed in Burge's argument is in fact rather contentious and might well be excluded from a charitable reconstruction of Frege's general theory on grounds that are entirely independent from issues regarding the hierarchy. By rejecting it the two-level theorist can block Burge's indirect argument before it reaches step (3) while preserving both the principle of extensionality as well as (Pr1). The two-level theorist can allow 'believes<sub>ML</sub>', 'Anna<sub>ML</sub>' and 'is wise(Plato)<sub>ML</sub>' to express (respectively) the senses denoted by '<believes>', '<Anna>' and '<is wise>(<Plato>)', without being pushed to the conclusion that 'believes<sub>ML</sub>(Anna<sub>ML</sub>, is wise(Plato)<sub>ML</sub>)' expresses the sense denoted by '<believes> (<Anna>, <is wise>(<Plato>))'. For in the absence of (Pr2) there is no longer any pressure to identify the function that maps the senses of the components of 'believes<sub>ML</sub>(Anna<sub>ML</sub>, is wise(Plato)<sub>ML</sub>)' to the sense of this complex expression with the function that maps the referents of the components of '<believes>(<Anna>, <is wise>(<Plato>))' to the referent of this complex expression. Having blocked Burge's argument in this way there is no reason to doubt that the right-hand side of (TC1') is an adequate translation of the sentence it mentions on the left-hand side, and hence no reason to doubt that modified truth-theory that delivers (TC1') is adequate in the sense required by Burge.

Curiously, the way in which Burge formulates his own truth-theory plays into the hands of the two-level theorist. For although we have seen Burge's argument to rely on (Pr2), his own truth-theory in fact harmonizes better with a rejection of this principle than with an acceptance thereof. Given that Burge treats the senses, and hence the indirect referents, of predicates as functions one would expect his truth-theory to equate the indirect referent of 'believes(Anna, is wise(Plato))<sub>OL</sub>' with the result of plugging the indirect referents of 'Anna<sub>OL</sub>' and 'is wise(Plato)<sub>OL</sub>' into a function identified with the indirect referent of 'believes<sub>OL</sub>'. However, we have seen in §3.3.1 that Burge's axiom (e) in fact equates the indirect referent of 'believes(Anna, is wise(Plato))<sub>OL</sub>' with the result of plugging the indirect referents of 'Anna<sub>OL</sub>', 'is wise(Plato)<sub>OL</sub>' and the indirect referent of 'believes<sub>OL</sub>' into a function  $C_3$  - just as one would proceed if one refrained from identifying the sense of 'believes<sub>OL</sub>' with a function. Burge's own treatment can therefore be preserved on our modification of his truth-theory. As we have seen, all the two-level theorist needs to do is to deny that the function  $C_3$  coincides with the function that maps the senses of the components of complex expressions to the sense of the complex expression.<sup>7</sup>

---

<sup>7</sup>Parsons (2009), just like Burge, assumes (Pr2) but, unlike Burge, incorporates this principle explicitly in his truth-theory. Accordingly he also identifies the two functions that our two-level theorist wishes to keep apart. Unsurprisingly, a variant of Burge's argument makes Parsons conclude that his truth-theory requires him to embrace the hierarchy (cp. 2009, § 2).



## 4 Potential objections

After having argued against the hierarchy view in §2, I have defended the two-level view against Burge's arguments in §3. I will now consider two potential objections against my defense of the two-level theory.

### 4.1 From the doctrine of reference shifts to (Pr2)?

My defense of the two-level theory against Burge's indirect argument involves the rejection of the principle (Pr2) according to which the senses of predicates are functions. Now, it could be objected that Frege's doctrine of reference shifts itself provides a reason to accept (Pr2) when combined with the following Fregean principle:

(Pr3) Any complex expression can be analyzed into components at least one of which refers to a function.<sup>8</sup>

According to (Pr3) a singly indirect occurrence of 'Plato is wise' must be analyzable into components at least one of which refers to a function. But in an indirect context all of the components of 'Plato is wise' are supposed to refer to their direct senses. So one of the components must be such that its direct sense is a function. Unless we want to identify the direct sense of the term 'Plato' with a function we have to identify the sense of the predicate 'is wise' with a function.

I think the two-level view can accommodate (Pr3) without having to identify either the direct sense of 'is wise' or that of 'Plato' with a function. To see how let us remind ourselves that natural language belief ascriptions normally contain the expression 'that'. Now we can treat the 'that' in 'Anna believes that Plato is wise' as referring to a function that maps the indirect referents of 'Plato' and 'is wise' to the referent of the entire that-clause, the thought that Plato is wise. This allows us to maintain that the indirect referents of both 'Plato' and 'is wise', i.e. their direct senses, are objects - not functions.

The fact that the occurrence of 'that' is not mandatory in more vernacular parlance shouldn't be seen as a problem. In those cases the linguistic context, in particular the juxtaposition of 'Anna believes' and 'Plato is wise', can be seen as tacitly performing the semantic job otherwise explicitly performed by 'that'.

On this view it is most natural to treat the 'that' of propositional attitude ascription as a multigrade functor, i.e. a functor that can take a varying number of arguments. In 'Berta believes that Anna believes that Plato is wise' the second 'that' takes two arguments, 'Plato' and 'is wise', while the first 'that' takes three arguments: 'Anna', 'believes' and the complex singular term 'that Plato is wise'. Dummett's (1973, p. 45) claim that Frege's conception of incomplete expressions doesn't allow for multigrade expressions has been refuted in Oliver & Smiley (2004, § 3.2).

---

<sup>8</sup>At least one because, according to Frege, two saturated expressions can't combine to a complex expression. Not *exactly one* because e.g. a first-level predicate and a (one-place) second-level predicate, though both regarded as unsaturated, can combine to a complex expression.

## 4.2 A different road to the hierarchy?

Although Burge's arguments are by far the most prominent arguments in favour of the hierarchy, a different, albeit related, argument has been proposed by Boisvert & Lubbers (2003). Boisvert & Lubbers<sup>9</sup> believe that Frege is committed to the following principle:

- (Pr4) The referent of a complex expression in any indirect context is composed of the referents of its component expressions in that context.<sup>10</sup>

Now recall that the two-level theorist maintains that in the sentence

Berta believes that Anna believes that Plato is wise

both 'Anna believes that Plato is wise' as well as 'Plato is wise' refer to the sense they express in direct contexts. It then follows from (Pr4) that the direct sense of 'Anna believes that Plato is wise' contains the direct sense of 'Plato is wise' as a component. B&L take this to be problematic because it conflicts with a principle of sense compositionality which they ascribe to Frege. According to them, Frege is not only committed to the principle (Pr1) which says the sense of a complex expression is uniquely determined by the senses of its component expressions but also to a principle according to which the sense of a complex expression is composed out of the senses of its component expressions.

The exact nature of this composition is left somewhat unclear on B&L's account. Unfortunately, they don't specify a general procedure by which we can determine for the sense of a given expression what senses exactly it is composed of. They do, however, compare the way in which a sense can be decomposed into other senses to the way in which a rectangle can be decomposed into, e.g. either two adjacent rectangles or two triangles (p. 40). This comparison strongly suggests that their notion of the (de)composition coincides with that of Dummett (1981, Ch. 15) who famously illustrates the multiple ways in which a sense can be decomposed with the multiple ways in which a country can be subdivided into different regions (p. 263). Fortunately, Dummett went on to give a precise account of what it means for a sense to be a component of another. Somewhat simplified the account is this:

- (Pr5) A sense  $S1$  is a component of a sense  $S2$  expressed by expression  $E2$  iff  $E2$  contains some expression  $E1$  that expresses  $S1$ .

Whether or not (Pr5) is exactly what B&L have in mind need not concern us because (Pr5) is plausible and sufficient to get their argument going.

Since the two-level theorist agrees that 'Plato is wise' expresses its singly indirect (rather than its direct) sense when it occurs in 'Anna believes that Plato is wise' the two-level theorist will agree that the latter expression doesn't contain any expression that expresses the direct sense of 'Plato is wise'. But this means the two-level view is now forced to the contradictory conclusion that the direct sense of 'Anna believes

---

<sup>9</sup>Henceforth: 'B&L'

<sup>10</sup>This is their principle (9\*\*) on p. 41

Plato is wise' both does contain the direct sense of 'Plato is wise' as a component (via (Pr4)) and that it doesn't (via (Pr5)).

Although B&L don't draw this connection, their argument might also be seen as an alternative way to fill the lacuna in Burge's indirect argument - a way which doesn't rely on (Pr2) and which is therefore not undermined by the rejection of this principle that I recommended to the two-level theorist. Given (Pr4) we should regard the referent of '<believes>(<Anna>, <is wise>(<Plato>))' as the sense composed out of the senses referred to by '<believes>', '<Anna>' and '<is wise>(<Plato>)'. And given (Pr5) we should regard the sense of 'believes<sub>ML</sub>(Anna<sub>ML</sub>, is wise(Plato)<sub>ML</sub>)' as the sense composed out of the senses expressed by 'believes<sub>ML</sub>', 'Anna<sub>ML</sub>' and 'is wise(Plato)<sub>ML</sub>'.

Since the senses referred to by '<believes>', '<Anna>' and '<is wise>(<Plato>)' just are the senses expressed by believes<sub>ML</sub>', 'Anna<sub>ML</sub>' and 'is wise(Plato)<sub>ML</sub>', it would follow once again that the referent of '<believes>(<Anna>, <is wise>(<Plato>))' coincides with the sense expressed by 'believes<sub>ML</sub>(Anna<sub>ML</sub>, is wise(Plato)<sub>ML</sub>)'.

Is B&L's argument against the two-level view successful? As in Burge's case, B&L's argument is only as convincing as the principles on which it relies. And while the two-level theorist might well want to accept a sense compositionality principle along the lines of (Pr5), the case B&L make for (Pr4) is not convincing. B&L are aware that (Pr4) doesn't hold as a general principle for reference: Sweden is not a component of the referent of 'the capital of Sweden' although it is the referent of a component expression. Why, then, should we believe that (Pr4) holds in the special case of indirect reference? B&L give two reasons for doing so (p. 53 ff.).

First, they argue that (Pr4) holds for the even more special case of an indirect occurrence of a complex expression which doesn't itself contain an indirect context creating operator, such as the occurrence of 'Plato is wise' in 'Anna believes Plato is wise' (p. 53 - 58). Here 'Plato is wise', 'Plato' and 'is wise' refer to their direct senses. And the direct sense of 'Plato is wise' is composed out of the direct sense of 'Plato' and the direct sense of 'is wise'. Hence the indirect referent of 'Plato is wise' is composed out of the indirect referent of 'Plato' and the indirect referent of 'is wise' in accordance with (Pr4). B&L think that we should extrapolate from this case and conclude that (Pr4) must hold in general, in particular also for expression containing indirect context creating expressions such as 'Anna believes that Plato is wise'. Otherwise, they claim, Frege's theory would be less 'neat' (p. 58).

This verdict about the neatness that results from accepting (Pr4) might be plausible if it wasn't for the very fact that accepting (Pr4) also leads to the infinite hierarchy by rendering the two-level view unavailable. I think that my discussion in the first half of this paper shows that having to accept the hierarchy would make Frege's theory much less neat than simply having to reject (Pr4) which hasn't as yet been independently motivated.

The second, and, according B&L, more important reason to accept (Pr4) consists in their claim that it would otherwise be impossible for a Fregean to allow for an expression that refers to the function which maps the indirect referents of the components of a complex expression to the indirect referent of that complex expression (p. 58 - 61).

Recall that I suggested in response to the first objection that a Fregean can treat the 'that' of believe ascriptions as doing just this. So, do B&L provide any reason to doubt this?

No. For some reason, the only candidate expressions that B&L consider are expressions that result from extracting singular terms from definite descriptions. They then argue that no such expression can be adequate unless (Pr4) is accepted.

There is no need to consider the details of this argument, for B&L's focus on definite descriptions is unjustified. For Frege every incomplete expression refers to a function. One way of producing an incomplete expression is by extracting one or more singular terms from a more complex singular term which may or may not be a definite description. On the present proposal, 'that Plato is wise' is a complex singular term in 'Anna believes that Plato is wise'. Furthermore its components 'Plato' and 'is wise' are also singular terms since their direct senses (their indirect referents) are objects rather than functions. Therefore, the 'that' of belief ascription can be seen as the result of extracting two singular terms from a more complex one and is thus suitable to perform the job envisaged for it.

Since we have seen no reason to accept B&L's (Pr4), the two-level theorist need not be worried by BL's argument even if she wishes to embrace a sense compositionality principle along the lines of (Pr5).

## 5 Conclusion

I have argued that Burge's and Kripke's combined efforts fail to make the infinite hierarchy view palatable and that we should therefore still try to avoid it. I have then shown that this can be done despite Burge's arguments to the contrary. While Burge's direct arguments neither threaten the one-level view nor the two-level view, his indirect argument is only effective against the one-level view. Providing an adequate truth-theory while maintaining the two-level view only requires rejecting a principle which might well be excluded from a charitable reconstruction of Frege's general theory on grounds that are entirely independent from issues regarding the hierarchy. The two-level view thus emerges as the best extension of Frege's theory of indirect contexts to iterations of indirect context creating operators.

## Acknowledgements

I am grateful to Alex Oliver and Robert Trueman, seminar participants at the University of Cambridge, as well as two anonymous referees for helpful comments. This work has been supported by the Robert Smithson studentship of the Faculty of Philosophy at the University of Cambridge.

## Appendix A

True('believes(Berta, believes(Anna, is wise(Plato)))<sub>OL</sub>') ↔

- (1) believes<sub>ML</sub>(d.ref 'Berta<sub>OL</sub>', i.ref 'believes(Anna, is wise(Plato))<sub>OL</sub>') (h)
- (2) believes<sub>ML</sub>(Berta<sub>ML</sub>, i.ref 'believes(Anna, is wise(Plato))<sub>OL</sub>') (a)
- (3) believes<sub>ML</sub>(Berta<sub>ML</sub>, C<sub>3</sub>(sns 'believes<sub>OL</sub>', i.ref 'Anna<sub>OL</sub>', i.ref 'is wise(Plato)<sub>OL</sub>')) (e)
- (4) believes<sub>ML</sub>(Berta<sub>ML</sub>, C<sub>3</sub>(sns 'believes<sub>OL</sub>', sns 'Anna<sub>OL</sub>', i.ref 'is wise(Plato)<sub>OL</sub>')) (b)
- (5) believes<sub>ML</sub>(Berta<sub>ML</sub>, C<sub>3</sub>(sns 'believes<sub>OL</sub>', sns 'Anna<sub>OL</sub>', C<sub>2</sub>(sns 'is wise<sub>OL</sub>', i.ref 'Plato<sub>OL</sub>'))) (d)
- (6) believes<sub>ML</sub>(Berta<sub>ML</sub>, C<sub>3</sub>(sns 'believes<sub>OL</sub>', sns 'Anna<sub>OL</sub>', C<sub>2</sub>(sns 'is wise<sub>OL</sub>', sns 'Plato<sub>OL</sub>'))) (c)
- (7) believes<sub>ML</sub>(Berta<sub>ML</sub>, C<sub>3</sub>(sns 'believes<sub>OL</sub>', sns 'Anna<sub>OL</sub>', sns 'is wise(Plato)<sub>OL</sub>')) (f)
- (8) believes<sub>ML</sub>(Berta<sub>ML</sub>, sns 'believes(Anna, is wise(Plato))<sub>OL</sub>') (g)

## Appendix B

True('believes(Berta, believes(Anna, is wise(Plato)))<sub>OL</sub>') ↔

- (1) believes<sub>ML</sub>(d.ref 'Berta<sub>OL</sub>', i.ref 'believes(Anna, is wise(Plato))<sub>OL</sub>') (h)
- (2) believes<sub>ML</sub>(Berta<sub>ML</sub>, i.ref 'believes(Anna, is wise(Plato))<sub>OL</sub>') (a)
- (3) believes<sub>ML</sub>(Berta<sub>ML</sub>, C<sub>3</sub>(<believes>, i.ref 'Anna<sub>OL</sub>', i.ref 'is wise(Plato)<sub>OL</sub>')) (e')
- (4) believes<sub>ML</sub>(Berta<sub>ML</sub>, C<sub>3</sub>(<believes>, <Anna>, i.ref 'is wise(Plato)<sub>OL</sub>')) (b')
- (5) believes<sub>ML</sub>(Berta<sub>ML</sub>, C<sub>3</sub>(<believes>, <Anna>, C<sub>2</sub>(<is wise>, i.ref 'Plato<sub>OL</sub>'))) (d')
- (6) believes<sub>ML</sub>(Berta<sub>ML</sub>, C<sub>3</sub>(<believes>, <Anna>, C<sub>2</sub>(<is wise>, <Plato>))) (c')
- (7) believes<sub>ML</sub>(Berta<sub>ML</sub>, C<sub>3</sub>(<believes>, <Anna>, <is wise>(<Plato>))) (f')
- (8) believes<sub>ML</sub>(Berta<sub>ML</sub>, <believes>(<Anna>, <is wise>(<Plato>))) (g')

## References

- Almog J. & Leonardi P. (eds). 2009. *The Philosophy of David Kaplan*. Oxford: Oxford University Press.
- Black M. & Geach P. (eds). 1980. *Translations from the Philosophical Writings of Gottlob Frege*. 3rd edition. Oxford: Blackwell.
- Boisvert D. & Lubbers C. 2003. Frege's commitment to an infinite hierarchy of senses. *Philosophical Papers* 32(1): 31–64.
- Burge T. 1979. Frege and the hierarchy. *Synthese* 40(2): 265–281. Reprinted in Burge (2005). Page references to Burge (2005).
- . 2004. Postscript to 'Frege and the hierarchy'. In *Truth, Thought, Reason - Essays on Frege*, 167–210. Oxford: Clarendon Press.

- . 2005. *Truth, Thought, Reason - Essays on Frege*. Oxford: Clarendon Press.
- Carnap R. 1947. *Meaning and Necessity. A Study in Semantics and Modal Logic*. Chicago: University of Chicago Press.
- Church A. 1951. A formulation of the logic of sense and denotation. In *Structure, Method, and Meaning - Essays in Honor of H. M. Scheffer*, 3–24. New York: The Liberal Arts Press.
- Davidson D. 1965. Theories of meaning and learnable languages. In *Proceedings of the 1964 International Congress for Logic, Methodology, and Philosophy of Science*, 383–393. Amsterdam: North Holland.
- Dummett M. 1973. *Frege: Philosophy of Language*. London: Duckworth.
- . 1981. *The Interpretation of Frege's Philosophy*. London:.
- Frege G. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100: 25–50. Translated by Max Black in Geach, P. and Black, M. (eds.) 1980, pp. 56–78.
- . 1892-5. Ausführungen über Sinn und Bedeutung. Unpublished, reproduced in Hermes et al. 1969, pp. 128–136. Translated by Long and White in Hermes, H. and Kambartel, F. and Kaulbach, F. (eds.) 1979, pp. 118–126.
- . 1902. Letter to Russell (28. 12. 1902). In G. Gabriel, H. Hermes, F. Kambartel, C. Thiel & A. Veraart (eds), *Philosophical and Mathematical Correspondence*, 152–155. Oxford: Blackwell.
- . 1904. Letter to Russell (13. 11. 1904). In G. Gabriel, H. Hermes, F. Kambartel, C. Thiel & A. Veraart (eds), *Philosophical and Mathematical Correspondence*, 160–166. Oxford: Blackwell.
- . 1923. Logische Untersuchungen. Dritter Teil: Gedankengefüge. *Beiträge zur Philosophie des deutschen Idealismus III (1923-1926)* 36–51. Translated by R. Stoothoof in Klemke 1968, pp. 537–558.
- Gabriel G., Hermes H. et al (eds). 1980. *Philosophical and Mathematical Correspondence*. Oxford: Blackwell.
- Heidelberger H. 1975. Dummett on Frege's philosophy of language. *Metaphilosophy* 6: 35–43.
- Hermes H., Kambartel F. et al (eds). 1979. *Posthumous Writings*. Chicago: University of Chicago Press.
- Klemke E.D. (ed). 1968. *Essays on Frege*. Urbana, IL: University of Illinois Press.
- Kripke S. 2008. Frege's theory of sense and reference: Some exegetical notes. *Theoria* 74: 181–218.
- Künne W. 1983. *Abstrakte Gegenstände - Semantik und Ontologie*. Frankfurt/M: Klostermann.
- . 2010a. *Die Philosophische Logik Gottlob Freges*. Frankfurt/M: Klostermann.
- . 2010b. Sense, reference and hybridity. *Dialectica* 64: 529–551.
- Makin G. 2000. *The Metaphysicians of Meaning*. London, Routledge.
- Oliver A. & Smiley T. 2004. Multigrade predicates. *Mind* 74: 609–681.
- Parsons T. 1981. Frege's hierarchies of indirect senses and the paradox of analysis. *Midwest Studies in Philosophy* 6: 37–57.
- . 2009. Higher-order senses. In P. Almog J. & Leonardi (ed), *The Philosophy of David Kaplan*, Oxford: Oxford University Press.