Referential Opacity

The problem of referential opacity is manifested primarily by the failure of a certain inference rule of classical logic to produce intuitively acceptable results when applied to ascriptions of mental states. The rule is known variously as *Leibniz's Law*, or as the *Substitutivity of Identicals*, or, in some logical systems, as *Identity Elimination*. However, the term 'referential opacity' for the phenomenon in question is somewhat tendentious, since it embodies a debatable diagnosis of the problem. We shall often use the more neutral 'substitution failure' instead ('apparent substitution failure' would be even more neutral). And we reserve "Leibniz's Law" for a principle about objects and properties, which says

(1) If *x* and *y* are the same object, then *x* and *y* have the same properties.

For example, if Istanbul and Constantinople are the same city, hence same object, and Istanbul straddles Europe and Asia, then Constantinople also has the property of straddling Europe and Asia.

For an embodiment of (1) as a rule of inference, we want a schema that tells us what the premises of an application of the rule should be, and what we can legally infer from those premises by applying the rule. If we take it that inference manipulates sentences, then the rule should say what sentence or sentences we may infer when we apply the rule to appropriate sentences.

Recasting our example, then, the rule of Identity Elimination, abbreviated "=E",

applies to two sentences: an *identity sentence*, in our example 'Istanbul is identical to Constantinople', which is called the *major* premise, and a *minor* premise which uses the name 'Istanbul' at one or more points; our minor premise is 'Istanbul straddles Europe and Asia'. By =E, we may then infer any sentence that results from replacing at least one occurrence of 'Istanbul' in the minor premise with an occurrence of 'Constantinople', making no other changes. In our example, the only sentence that we can obtain this way is 'Constantinople straddles Europe and Asia'. However, if the minor premise had been 'Part of Istanbul is in Europe and part of Istanbul is in Asia', there would have been three legal options, varying over which occurrences of 'Istanbul' get substituted – the first only, the second only, or both. This rule is called Identity Elimination because we move from a major premise containing an occurrence of 'is identical to' to a conclusion in which that occurrence is not present. It is called Substitutivity of Identicals because it involves substitution of one name for another on the basis of an identity sentence.

The problem of referential opacity is simply to explain what is going wrong when the rule produces implausible results. Throughout this article we shall draw examples mainly from the Superman fiction, treating it as fact. For readers unfamiliar with the legend: a child from the planet Krypton, Kal-El, is sent to Earth, where physical conditions cause him to acquire superpowers. Wearing specific clothing, including a red cape and some kind of blue jumpsuit, Kal-El performs heroic deeds of preventing disasters; rescuing endangered innocents; and foiling would-be perpetrators of crimes (often Lex Luthor). People refer to Kal-El as 'Superman' when talking about Kal-El's actions of this kind.

But Kal-El also takes a day-job as a reporter, using the name 'Clark Kent'. One of

his co-workers, Lois Lane, treats him with relative indifference in the office, but has a pronounced crush on, as she would put it, Superman, unaware it is the same individual (we are supposed to suspend our disbelief that when Kal-El dons glasses and a business suit he is unrecognizable as Superman).

The problematic examples we shall consider mostly involve ascriptions of mental states to Lois (or occasionally Lex), arrived at by applying the rule =E to the major premise 'Superman is Clark Kent' and a carefully chosen minor premise. For example, as we just said, Lois has a crush on Superman (minor premise), so given that Superman is Clark Kent, we should be able to infer by =E that Lois has a crush on Clark Kent. But this latter seems false, and would certainly be rejected by Lois herself. There is no end of similar examples: Lois believes that Superman can fly, but doesn't seem to believe that Clark can; she hopes to see Superman again soon, but seems not much to care when she next sees Clark; she'd like a date with Superman, but apparently has no interest in one with Clark; and so on.

The problem may seem small-scale; yet it turns out to be anything but, and solutions to it line up along major divides in philosophy of language. It was introduced into contemporary discussion by Quine (1956, 1961). Important early contributions include Marcus (1961, 1962, 1975) and Smullyan (1948). More recent influential engagements with Quine are Kaplan (1986) and Fine (1989). However, the essential problem was raised by Frege in the paper that initiated philosophy of language as a separate branch of philosophy, (Frege 1892). Here Frege asks how the astronomical discovery that Hesperus is the same heavenly body as Phosphorus can be a valuable extension of our knowledge, while one wouldn't say the same thing of proving that Hesperus is the same heavenly body as Hesperus

by one step of Identity Introduction (the rule that every sentence of the form $c = c$ is a theorem). If Frege is right, this is another failure of =E, for if knowing that Hesperus is Phosphorus is the same thing as knowing that Hesperus is Hesperus, the one piece of knowledge can't be a valuable extension of our knowledge unless the other is. For this reason, the problem of referential opacity is also known as Frege's Puzzle. To find out about Frege's solution, read on.

Table of Contents

1. Identity Elimination and its Misuses

A little more formally, the rule of inference =E can be stated as:

Identity Elimination Schema

Major: $t_1 = t_2$

Minor: $\phi(t_1)$

Conclusion: $\phi(t_2)$

Here t_1 and t_2 are expressions which refer to entities (in our examples so far, proper names of people or places). $\phi(t_1)$ is a sentence containing at least one occurrence of t_1, and $\phi(t_2)$ is a sentence that results from replacing at least one occurrence of t_1 in $\phi(t_1)$ with an occurrence of t_2. Recurring t_i presumes that t_i is used in the same way throughout, and recurring ϕ presumes that the sentential context ϕ is the same throughout. If these *uniformity* conditions are not met, then the inference scheme is being misapplied, and it is no wonder that false conclusions are derivable. Note that the rule is direction-sensitive: it is the left term in the major premise that is replaced in the minor premise by the right term. A non-directional version is easy to state, but is derivable from the directional version in view of the latter's entailing the *symmetry* of identity, that $t_1 = t_2$ entails $t_2 = t_1$ ($t_1 = t_2 \vdash t_2 = t_1$). This follows from the theorem $t_1 = t_1$, using $t_1 = t_2$ as major premise to replace the first t_1 in $t_1 = t_1$ with t_2. In this case the minor premise is also an identity sentence, and in any use of =E it's important to be clear about which premise is major and which is minor. The reader can confirm that the *transitivity* of identity, that $t_1 = t_2$ and $t_2 = t_3$ entail $t_1 = t_3$ ($t_1 = t_2$, $t_2 = t_3 \vdash t_1 = t_3$), is quickly proved with one choice of major premise for =E but not the other.

For the kind of problematic use of =E we discussed in the Introduction, we will take the following instance to be our paradigm:

(2) a.	Superman is Clark Kent	Major
b.	Lois believes that Superman can fly	Minor
c.	∴ Lois believes that Clark Kent can fly	a,b =E

It's important to see that it is *not* a solution to the problem of referential opacity to say that when we apply the rule in an instance like (2), the flaw is that the major premise is one that Lois does not realize is true. No doubt her ignorance explains psychologically why she does not *draw the conclusion* that Clark can fly, in those very words, but it doesn't go any way to explaining semantically how the inference rule can carry us from two truths to a seeming falsehood: 'Lois realizes (2a) is true' is not itself a premise for the application of the rule in (2), so its falsehood is irrelevant to what is wrong with the application. Indeed, the rule enables the inference that Lois *does* realize (2a) is true: simply change the minor premise of (2) to 'Lois realizes Superman is Superman', surely a trivial truth once she has acquired the name 'Superman' from watching Kal-El perform heroic deeds.

Some terminology is commonly encountered in discussions of cases like (2). Mental-state ascriptions like (2b) and (2c) are called *attitude* ascriptions, since the subject is being ascribed a mental attitude towards something. When the something is specified by a 'that'-clause, or another complementizer such as 'if' or 'whether', the ascription is called a *propositional* attitude ascription, since the 'that'-clause is standardly taken to specify a proposition, the one expressed by the sentence which 'that' prefixes (but see, e.g., Davidson 1969, Bach 1997, and

Moltmann 2003, for rejection of this). So (2b) says that Lois has the attitude of belief towards the proposition that Superman can fly. The sentence following the 'that' in (2b) and (2c) is called the *content* sentence, though in English 'that' can often be dropped, as it can be from (2b) and (2c).

a. Quotation

There is mileage to be gained from the idea that the reason we get counterintuitive instances such as (2) is that the rule of =E is being misapplied in some way, or, relatedly, that the rule as formulated isn't a faithful reflection of the motivation provided by Leibniz's Law, as stated in (1) – a better formulation *would* have to be misapplied to get (2). There are some well-known cases of misapplication of the rule which motivate critiques of (2) as a relevantly similar misapplication. One sort of case, emphasized by Quine (1961), is

(3) a. Istanbul is Constantinople

b. 'Istanbul' has eight letters

c. ∴ 'Constantinople' has eight letters.

This is a misapplication of =E because the name 'Istanbul' does not occur uniformly in (3). In the major premise it is used in the normal way to refer to a certain city. But in the minor premise it is not used to refer to that city (perhaps it is not used to refer at all). Rather, it occurs as part of a complex expression, the quotation-name "'Istanbul'", referring to the name 'Istanbul' rather than the city Istanbul (this is a Tarskian rather than Fregean account of quotation – see further Richard 1986, Washington 1992, Saka 2006 – but the non-uniformity objection to (3) holds on either). (3b) correctly predicates 'has eight letters' of the word

'Istanbul', as opposed to unintelligibly predicating 'has eight letters' of the city Istanbul. So (3) has no more force than a variant in which the minor premise reads 'the first name used in (3a) has eight letters' and the conclusion reads 'the second name used in (3a) has eight letters', which doesn't even pretend to be an application of =E. At best it seems to presume the absurd principle that if two names refer to the same thing then they have the same number of letters.

Quine thought examples like (3) instructive. The position of 'Istanbul' in (3b) is not open to substitution, like the position of 'Superman' in (2b), and 'Istanbul' doesn't seem to be referring normally in (3b), so perhaps the same should be said of 'Superman' in (2b): the position 'Superman' occupies in (2b) is *referentially opaque*, hence the terminology. But it is unclear how instructive (3) really is. Quine suggests (1956:186) that we should give 'serious consideration' to construing mental state ascriptions such as (2b) as involving quotation. (2b) so-construed would say that Lois believes-true 'Superman can fly' as a sentence of English. But he immediately hedges by adding that this 'is not to suggest that the subject speaks the language of the quotation, or any language…We may treat a mouse's fear of the cat as his fearing-true a certain English sentence.' Unfortunately, we are left in the dark about what it is to believe-true or fear-true a sentence of a language when one does not know the language (isn't "believes-true '…'" simply short for "believes that '…' is true"?). Quine then admits that the quotational construal of mental state ascriptions will only yield a 'systematic agreement in truth-value…and no more'. But even that seems dubious. The mouse notwithstanding, a monolingual Frenchwoman may believe that Superman can fly without believing that 'Superman can fly' expresses a truth in English (she may not even have heard of English); conversely, she may believe that 'Superman

can fly' is an example of a sentence that expresses a truth in English, because she has been told so by a trusted informant; clearly, this doesn't mean she believes Superman can fly. (See Church 1950 for a famous discussion of quotational accounts, and Schweizer 1993 for a technical investigation of quotational accounts of modal logic.)

A quotational account that does rather better, Quine notes, is that (2b) says that Lois believes the *meaning* of 'Superman can fly', which avoids the problem of the monolingual Frenchwoman. But then, of course, it is not really the presence of quotation that is blocking substitution. For if this new quotational account is correct, then (2) is valid reasoning if 'Superman can fly' and 'Clark can fly' mean the same on account of (2a). So, to explain the fallacy in (2), we'd need a theory of meaning on which these two sentences do not mean the same, despite (2a). Since the only difference between them is the name in subject position, this suggests we'd want an account of the meaning of names on which names can be coreferential yet differ in meaning, and indeed, some accounts to be considered below pursue this. But quotation is of no real relevance to this line of thought.

b. 'so-called'

Quine has another example of misapplication of =E, but one which tends to undermine the thought that there is something referentially peculiar about the position occupied by the substitution-resistant name (though he appears to regard the example as *supporting* this idea). His well-known 'Giorgione' case (Quine 1961:17) is as follows:

(4) a. Giorgione is Barbarelli

b. Giorgione is so-called because of his size

c. ∴ Barbarelli is so-called because of his size.

In (4) there is nothing unusual about the way in which any of the names is used: in each use there is simply reference to a certain artist. The reason the inference fails to be a legal application of =E is that the sentential context 'is so-called because of his size' does not occur uniformly, since the reference of 'so' changes in moving from (4b) to (4c): in (4b) 'so' refers to the name 'Giorgione', but in (4c), it refers to the name 'Barbarelli'. The supposed application of =E is therefore a simple fallacy of equivocation, brought about by the substitution having a hidden truth-condition-altering side-effect (altering the reference of 'so'). But it may be an instructive fallacy, if anything like a covert 'so' is present in attitude ascriptions. (For other examples of non-uniformity, see Fine 1989:222–36; for more on 'so-called', Forbes 2006:154–7.)

c. Modality

Our last example of misuse of =E involves *intensional* operators, which are operators which do not allow interchange within their scope of *accidentally coextensive* expressions (two predicates are *coextensive* if and only if (iff) they actually apply to exactly the same things, and *accidentally* coextensive if they are coextensive but there *could* have been something to which one applies and the other doesn't; two sentences are accidentally coextensive iff they have the same actual truth-value but could have differed in truth-value). The standard cases of intensional operators are modal operators such as 'it is necessary that', 'it is possible that' and 'it is contingent that' ('it is contingent that *S*' means that *S* is

actually true but could have been false), though in logic, tenses are also treated as intensional operators.

To illustrate how intensional operators induce failure of substitution of accidentally coextensive predicates, suppose I have in my garage three cars, all Bentley racing cars from the 1920's, and that these are the only three in existence (the only three Bentley ever built). Then for any *x*, *x* is a car in my garage iff *x* is a Bentley racing car. But it's metaphysically possible that some car in my garage is not a Bentley, in the sense that there's a way things could have gone as a result of which a car from a different manufacturer ends up in my garage. By contrast, it's not possible that a Bentley racing car is not a Bentley. The problem is that the two predicates '*x* is a car in my garage' and '*x* is a Bentley racing car' are only accidentally coextensive, while intensional operators are sensitive to what might be called the 'modal profile' of expressions within their scope: the array of semantic values they have, sets in the case of predicates, across ways things could have gone, or 'possible worlds'. '*x* is a car in my garage' and '*x* is a Bentley racing car from the 1920s' would have the same modal profile iff at each world, the set of things the first applies to is the same set as the set of things the second applies to. But as we have observed, there is a possible world *w* where the set of things one predicate applies to is different from the set of things the other applies to, since there is, say, a Bugatti in my garage in *w*. As the example shows, attempts to substitute predicates which aren't *necessarily* coextensive within the scope of a modal operator easily go awry, resulting in absurdities such as a Bentley that isn't a Bentley: within the scope of 'possibly' or 'it could have been that', 'car in my garage' cannot be replaced by the accidentally coextensive 'Bentley racing car' in the sentence 'a car in my garage isn't a Bentley'.

The same can happen with expressions which are accidentally *coreferential*. Fix a definition of 'planet' on which there are nine planets in our solar system, and suppose that this is a contingent fact: there could have been more or fewer planets (on that definition of 'planet'). Then the following use of =E derives a false conclusion from true premises:

(5) a. The number that numbers the planets = 3^2

b. It's contingent that the number that numbers the planets = 9

c. ∴ It's contingent that $3^2 = 9$

The conclusion is false because true mathematical identities such as '$3^2 = 9$' are the paradigm cases of necessary truths: there is no way things could have gone in which the number 9 fails to be the outcome when the number 3 is multiplied by itself.

(5) differs from previous examples in that one of the terms in the major premise, 'the number that numbers the planets', is not a proper name, but rather what is called a *singular definite description:* 'definite' because 'the' coupled with a singular nominal (e.g., 'number') implies exactly one, and 'description' because the expression picks out whatever individual, if any, is the unique satisfier of the descriptive condition 'F' in 'the F', in this case 'number that numbers the planets'.

Definite descriptions can be classified in at least two ways. One option is that they are treated as belonging to a unitary semantic category of *singular terms*, together with other grammatical categories such as proper names, demonstratives and indexicals: expressions of all these types 'designate' objects. The classification of definite descriptions along with names goes back to (Frege 1892).

The other approach classifies a definite description 'the F' as a first-order quantifier, like 'some F', 'each F', 'no F' and so on (the apparent structural similarity between 'the F is G' and '{some/each/no} F is G' is seen as genuine). A quantifier like 'some F' is a combination of a *det*(*erminer*)'some' with a predicate F, that then combines with a second predicate. In '(*det* F is G)', 'F' is the *restriction*, or restrictor, in the quantifier '*det* F', and 'is G' is the quantifier's *scope*. In symbols, to take a simple example, 'no dog barked' would be represented as '(no *x: x* is a dog)[*x* barked]' and so by parallelism, 'the dog barked' would be '(the *x: x* is a dog)[*x* barked]': as in the English, only *det* changes as we formalize 'the dog barked', 'each dog barked', 'some dog barked' and so on (for further discussion, see Davies 1981:149–52). (Russell's Theory of Descriptions (1905) is also a quantificational account, in the looser sense that Russell took 'the F' to be an apparent singular term in need of analysis by standard quantifiers.)

Only the singular term account of descriptions raises the problem of referential opacity, for if the descriptions in (5a) are quantifiers rather than singular terms, they are not referential and =E could not be applied in the first place (the same is true on the 'predicate' account of descriptions in Fara 2001): the major premise is not of the form $t_1 = t_2$, but is rather '(the *x: Fx*)[(the *y: Gy*)[*x* = *y*]]'.

However, even if descriptions are singular terms, they may be a special case semantically, which could make (5) not very illuminating about (2). Assuming the singular term analysis, definite descriptions other than mathematical ones are *non-rigid designators:* they do not pick out the same object at all possible worlds (Kripke 1972, 1980:48*ff*). For example, the number nine is the unique satisfier of 'number that numbers the planets' at the actual world, but in some other possible

world, a different (natural) number is the unique satisfier, or, perhaps, there is no unique satisfier because there are no planets. '3^2' is the less common case, a *rigid* definite description: the notation abbreviates 'the product of the number three with itself', and nine uniquely satisfies 'product of the number three with itself' at every possible world, since numbers exist in every possible world and the product operation is the same at every possible world. (There are other ways of cooking up rigid descriptions, using modal operators; see Davies and Humberstone 1980. For further discussion of non-rigidity, see Tichy 2004.)

According to Kripke (1972), proper names, by contrast with typical descriptions, are rigid designators, which is to say that they denote the same object with respect to every possible world. This is either because they have a world-relative reference, but the same referent at every world, or (a different account) because they have an absolute reference, and their reference at any world is a pseudo-relativization of their absolute reference. To see the case for rigidity, suppose we say that the planet Jupiter could have failed to exist. Here we are talking about a specific heavenly body which in the actual world orbits the Sun between Mars and Saturn, but which, we might say, in certain other possible worlds is simply never formed, because of different behavior on the part of the original protoplanetary disk, or because a physical universe never comes into existence, or for whatever possible reason. When we say that Jupiter doesn't exist in such circumstances, we mean to be talking about our relatively familiar planet, the third-brightest object in the night sky, and saying that *it* doesn't exist. So 'Jupiter' denotes Jupiter at each possible world *w*, no matter what happens in *w*, even failure of the denotation of 'Jupiter' to exist (see further Salmon 1981:32–40).

It's crucial to problematic uses of =E in the style of (5) that at least one of the singular terms in the major premise be non-rigid. For if they are both rigid and also codesignate, then the minor premise and the conclusion will have to agree in truth-value. So we might propose a restriction on =E that makes the application in (5) illegal. The weakest restriction motivated by the failure of (5) is that t_1 and t_2 must have the same modal profile: for each w, either t_1 designates the same thing as t_2 at w, or neither designates anything at w. A slightly stronger restriction is that t_1 and t_2 have the same modal profile and at each w, each designates something. Note that such restrictions aren't an aspect of uniformity, which is concerned with whether the *same* expression is used univocally in all its occurrences in an inferential step. Rather, we are simply proposing a *sui generis* addition to the constraints that correct application of =E in modal languages must meet, a constraint that is required because we are treating definite descriptions as singular terms.

Allowing application of =E in formal modal languages only if the terms in the major premise have the same modal profile isn't workable, however, since 'same modal profile' is relative to an interpretation. Nor is there a syntactic surrogate of this condition, again because of relativity to an interpretation. So the standard approach is (i) to decree that =E is only applicable when t_1 are t_2 are individual constants (the formal counterpart of proper names), and (ii) set up the semantics so that all individual constants are rigid designators, no matter the interpretation. (Some might object that it's illegitimate to sneak semantics into the statement of an inference rule, as the combination of (i) and (ii) does.)

Using '□' for 'necessarily', we can legally prove

(6) $c = d \vdash \Box(c = d)$

simply using =E once, with the minor premise '$\Box(c = c)$', which is a theorem and therefore doesn't need to be mentioned on the left of (6). But (using '$\exists!$' for 'there exists exactly one') we will not be able to prove even

(7) $\textit{the } F = \textit{the } G \vdash \Box([(\exists!x)Fx \,\&\, (\exists!x)Gx] \rightarrow (\textit{the } F = \textit{the } G))$,

much less with the unconditional version of the conclusion, '$\Box(\textit{the } F = \textit{the } G)$'. The restriction in =E to individual constants blocks anything like a proof of (7) analogous to that of (6) just mentioned, and restrictions on the rule of $\Box$-Introduction will prevent a derivation of the conclusion given sound rules which (subject to controls) permit individual constants to function as proxies for descriptions in subproofs. Since individual constants are the formal counterparts of natural-language proper names, then, we can classify (5) as a misuse of =E, since in (5a) at least one term is not a proper name.

[A more technical aside for those who prefer the treatment of 'the F' as a quantifier. 'The F is G', that is, '(the x: Fx)[Gx]' can be handled in inference by rules in the style of $\exists$-Elimination and $\exists$-Introduction. For actualist 'the' in a modal language intended for interpretation on varying world-domains, the assumption for 'the'-Elimination would be '$\text{exists}(t) \,\&\, Ft \,\&\, Gt \,\&\, (\exists!x)Fx$', where t is an individual constant. For possibilist 'the', delete '$\text{exists}(t)$' and replace $\exists$ with Σ, the possibilist existential quantifier. The reader who pursues this will find that, e.g., the quantifier version of (7)'s premise doesn't allow derivation of the quantifier version of the conclusion. =E can be used unproblematically in the course of an attempted proof, applied to constants introduced by assumptions for use of

'the'-Elimination. But there will be no way to introduce the final '□', ultimately because of the restriction that □-Introduction can only be applied to a formula if the premises that formula depends are all *fully modalized* (all predicates and quantifiers are in the scope of a modal operator), which the premise in question is not. And there is no indirect proof which gets round this.]

The relevant question for us is whether there is anything in our discussion to justify the claim that the definite description 'the number that numbers the planets' occurs *opaquely* in (5b). As already noted, the approach to descriptions on which 'the F' is a quantifier, not a singular term, would have to be rejected before the question whether descriptions are *referentially* opaque in modal contexts could even arise, since quantifiers are not referential. So for *opaque* to be a possible characterization of the occurrence of 'the number that numbers the planets' in (5b), we must take a side, not necessarily the most plausible side, on the singular term/quantifier dispute.

Even granting that definite descriptions are singular terms, it's hard to see why 'the number that numbers the planets' should be thought to be functioning deviantly in (5b), or in some other way that merits the term 'opaque'. In an extensional language, the designation of a definite description in given circumstances is calculated following the semantic structure of the description. For example, 'the man who first set foot on the Moon' will designate the unique entity, if there is one, that satisfies both 'is a man' and 'first set foot on the Moon'. To satisfy 'first set foot on the Moon' such an entity must be the first satisfier of 'set foot on the moon', which in turn has further semantic structure. This evaluation procedure, of following the structure to arrive at a unique object (if there is one),

does not change when we move to an intensional language; it is simply that in interpreting an intensional language there are typically varying circumstances with respect to which an expression can be evaluated. A conjunction A & B may have different truth-values in different circumstances, but no-one would accuse '&' of being problematic on account of this. Similarly, the fact that 'the F' can have different designations in different circumstances is hardly a cause for concern.

Of course, (5) may seem to indicate a problem; but then, so may the sequent

(8) $A \leftrightarrow B, \Diamond(A \,\&\, C) \nvdash \Diamond(B \,\&\, C)$

(here '$\Diamond$' means 'possibly'; consider the case where $C = \neg B$). From (8) we learn that substitution on the basis of accidental coextensiveness of sentences doesn't work in modal languages and we must constrain any substitution rule to require necessary coextensiveness. In the same way, from (5) we learn that substitution on the basis of accidental codesignation is invalid in modal languages, and we must constrain =E to allow its application only if the codesignation is necessary. This is exactly what we have done, by restricting the singular terms of the major premise to individual constants, whose semantics requires them to be rigid designators.

Is there an analogous restriction on =E that we could employ to make the rule acceptable for languages with attitude verbs like 'believe'? That $t_1 = t_2$ be rigid designators is insufficient, as (2) shows. And we want a condition that doesn't make it a matter of mere *mental compulsion* that any subject in the minor-premise's propositional attitude comes to be in the conclusion's propositional attitude: it has to be *logically* guaranteed. Nothing weaker than identity of proposition determined by the two 'that'-clauses satisfies this demand. So if we

agree that a difference in the semantics of the two names would result in the two content-sentences in (2) expressing different propositions, we will have to say that the two names in a use of =E in the likes of (2) must be *synonymous*.

But it's not clear what it means to apply 'synonymous' to a pair of names. Names are not usually found in dictionaries, so the normal notion of synonymy, on which, say, 'attorney' and 'lawyer' are synonyms in virtue of having the same dictionary definition, won't help. There is also a more serious objection, due to Mates (1952), to the effect that even substitution of dictionary synonyms in attitude ascriptions can produce results not much more comfortable than (5). For example, (9a) below may well be false, yet it seems (9b) could still be true:

(9) a. I suspect that many people doubt that everyone believes all lawyers are lawyers.

b. I suspect that many people doubt that everyone believes all lawyers are attorneys.

One moral we might draw from 'Mates cases' like this is that searching for a criterion which allows substitution of t_2 for t_1 in attitude reports is likely to be futile. (For further discussion of attitude reports differing by a synonym, see Burge 1978 and Kripke 1979:160–61.)

To summarize, we have considered three incorrect uses of =E, (3), (4) and (5), in the hope that our better understanding of why they go wrong will help us pierce the fog surrounding (2). But (3) turned out not to be so useful, given the drawbacks to quotational accounts of attitude ascriptions. (5) suggests trying to modify =E by limiting use of =E to some favored class of singular terms, but Mates

cases cast doubt on whether this line will be productive (see also Kaplan 1969, Section xi). This leaves (4), which shows how a substitution can have a hidden truth-condition-altering side-effect, a paradigm to which we will return.

For the moment, we note a distinction which emerges from the unhelpfulness of (5). (5) illustrates difficulties for =E which arise from the intensionality of certain vocabulary, primarily modal operators, difficulties resolved by a more careful statement of the rule that imports essentially semantic considerations. Since the difficulties for =E illustrated by (2) don't seem to be resolved by semantics for intensional expressions, the problem manifest in (2) is said to arise from the *hyper*intensionality (or *fine-grained* intensionality) of psychological vocabulary such as attitude verbs. However, even hyperintensional semantics doesn't necessarily legitimize a qualified version of =E. (For a version of hyperintensional semantics that takes propositions as primitive, see Thomason 1980, and for discussion, Muskens 2005; for a longer study of alternatives, see Fox and Lappin 2005.)

2. The *De Re*/*De Dicto* Distinction

It's possible to get oneself into a frame of mind according to which there is no such thing as hyperintensionality, and the reasoning of (2) is not flawed at all. For if Lois believes Superman can fly, then, since Superman is Clark, she just *does* believe that Clark can fly, even though she wouldn't *put* it that way. What you believe is one thing, which words you're inclined to use when stating your beliefs is another, and if you are ignorant of an identity, you may disprefer or even reject particular wording that nevertheless correctly reports what you believe. So even if Lois would laugh if someone suggested to her that Clark has superpowers, she

may still believe it.

One view about this argument in favor of (2) is that it is essentially correct. We shall return to this so-called *Russellian* position later. But a second view is that it exploits an ambiguity that is present in (2b), 'Lois believes that Superman can fly', and in (2c), 'Lois believes that Clark can fly'. According to this view, an attitude ascription such as (2b) can be read in a way that permits substitution and in a way that does not. Normally, we understand such ascriptions in the way that does not, which is why we reject (2), but if cajoled enough ('look, she *does* believe Clark can fly, she just wouldn't say it like that') we may switch to a reading that allows substitution. In the usual terminology, this is called the *de re* reading, contrasting with the more common *de dicto* reading, which disallows substitution. Other terminology for this reading is *relational*, contrasting with *notional*; *transparent*, contrasting with *opaque*; and *wide scope*, contrasting with *narrow scope*. We turn now to explaining what difference is being putatively marked here.

a. Defining the distinction

None of the above terminology is entirely happy. It's unclear in what sense the substitution-resistant reading of (2b) is any less 'about the thing' ('*de re*') than a putative substitution-permitting reading, nor is it clear why the truth of (2b) understood in a substitution-resistant way makes the subject of the ascription any less *related* to the object the attitude is about (Lois believes Superman can fly because she has *seen* him do it). And 'transparent/opaque' employs the notion of opacity, which, if it isn't just a synonym for 'substitution resisting', suggests failure to refer in the normal way, an idea we have yet to find a justification for.

But 'wide scope/narrow scope' is more useful. The rationale for 'wide scope' is

the thought that a substitution-permitting reading of (2) can be brought out by a formulation in which the crucial name is moved to a position in front of the attitude verb (it has *wide scope* with respect to the verb), as illustrated in

(10) a. Superman is such that Lois believes that he can fly.

b. Superman is someone who Lois believes can fly.

The step from (2b) to (10a) or (10b) is called *exportation*, and it's intuitively plausible that the exported forms permit substitution: if Superman is someone Lois believes can fly and if Superman is Clark, then indeed Clark is someone Lois believes can fly. So if we read the minor premise and conclusion of (2) in the exported way, we have an explanation of why someone might under pressure find (2) acceptable after all. For (2a) and either (10a) or (10b) entail the exported variant of (2c). Note, however, that we are not saying that exportation is *valid*, e.g., that (2b) entails (10a) (though it seems to – for worries about existential commitment of the kind raised in Donnellan 1974, see Forbes 1996:357–62, and more generally Kvart 1984). The point here is just that (2b) and (2c) could be understood *straight off* in the style of (10), which would explain why (2) might be swallowed.

One advantage of the wide-scope/narrow scope terminology is that it marks a distinction whose existence is not in doubt, insofar as it is simply a syntactic distinction, manifested in the contrast between, say, (2a) and (10a) or (10b). But of course, there is a question whether the syntactic distinction marks any interesting semantic difference. One argument that there is a semantic difference is that the same syntactic distinction arises with definite descriptions and (other?) quantifiers, where a semantic difference is hard to deny. For example, we have

(11) a. Lois believes that the extraterrestrial who works at The Daily Planet likes her.

b. Lois thinks that no extraterrestrial is in this conference room.

c. Lois hopes that someone born on Krypton will come to her aid.

If the quantifiers are given narrow scope, that is, if the examples in (11) are interpreted respecting word-order, (11a) is false, (11b) is (say) true and (11c) is false. (11a) is false because Lois does not think that there are any extraterrestrials who work at The Daily Planet, so would not use 'The extraterrestrial who works at The Daily Planet likes me' to express any belief of hers. (11b) is true even though Clark is in the conference room along with Lois and she can see him. But since Lois presumes none of her colleagues is an extraterrestrial, she will happily use 'No extraterrestrial is in this conference room' to say what she thinks about the planetary origins of those in the room. And (11c) is false because (let's suppose) Lois has never heard of the planet Krypton; therefore she will not think or say 'Would that someone born on Krypton comes to my aid!'. At least, these are the commonsense verdicts about the examples in (11), based, as is evident, on maintaining a close connection between the content of mental states and their verbal expression by the subject (on which see Burge 1978, p.132;).

However, these judgements of truth-value reverse themselves when we consider the exported forms:

(12) a. The extraterrestrial who works at The Daily Planet is someone who Lois believes likes her.

b. No extraterrestrial is someone who Lois thinks is in the conference room.

c. Someone born on Krypton is such that Lois hopes he will come to her aid.

(12a) is true because Clark is the extraterrestrial who works at The Daily Planet and Lois believes Clark likes her; (12b) is false because Clark is an extraterrestrial and Lois thinks Clark is in the conference room; and (12c) is true because Superman was born on Krypton and Lois hopes Superman will come to her assistance.

Not only does this contrast between (11) and (12) indicate that exportation makes a semantic difference, it also indicates what that difference is. The false cases in (11) are false because they make attitude attributions to Lois using concepts that either she lacks ('born on Krypton'), or thinks empty ('extraterrestrial who works at the Daily Planet') and so wouldn't employ positively in any belief she has; while the true case, (11b), is true precisely because 'no extraterrestrial' is used to specify the content of her belief. In (12), on the other hand, problematic material is kept out of the specification of Lois's mental states, which allows (12a) and (12c) to be true, while in (12b) we get a falsehood precisely because 'no extraterrestrial' functions simply as an objectual quantifier, without characterizing the content of her belief. So in propositional attitude attributions with wide-scope material binding into the content sentence, the content sentence only *partially* characterizes the attitude, while if there is a 'closed' content sentence within the scope of the attitude verb, that is, if there is no exported material, the content sentence *fully* characterizes the attitude. And we can then, if we like, resurrect

the '*de re*/*de dicto*' terminology and use it in the same way as 'wide scope/narrow scope'. The hallmark of a *de re* attribution is not that it says that the subject of the attribution stands in a special relation to the thing the attitude is about, but that the attribution designates or characterizes that thing in a way the ascriber chooses irrespective of whether the subject would accept the characterization, and the subject's resisting the characterization is not even *prima facie* reason to think the attribution false; while a contested *de dicto* attribution is *prima facie* false. (See further Yalcin 2015: 210–13; also see Marcus 1962 and Kazmi 1987 on the interpretation of exported quantifiers.)

This gives us a non-tendentious way of using '*de re*/*de dicto*', aligned with 'wide scope/narrow scope', that justifies our proposed diagnosis of any inclination to say that (2) passes muster: the diagnosis is that such judgement relies on construing the minor premise and conclusion *as if* they were in exported form, that is, construing them as *de re* attributions in the just explained sense. Still, it is worth observing that on this account we are equating the '{permits/resists} substitution' distinction in the examples in question with a scope ambiguity. This may be too strong: there may be a substitution-permitting reading of, say, (2b), 'Lois believes that Clark can fly', which isn't to be explained as involving a wide-scope reading for 'Clark'. We will return to this point later, in connection with hidden-indexical semantics.

b. Scepticism about the distinction

We have arrived at an apparently defensible way of understanding the *de re*/*de dicto* distinction, however this distinction is to be employed. We must therefore note that there are a number of expressions of scepticism about the distinction in

the literature, for example (Dennett 1982), (Richard 1990:128 – 31), (Sosa 1970) and (Taylor 2002), whose main points have not been addressed here. So let us briefly consider a selection.

Taylor points out that even if using a definite description would provide an accurate characterization of what a subject *J* believes or doubts, in the sense that the content-sentence containing the description echoes the sentence *J* would produce to express *J*'s attitude, an ascriber will in certain cases resist using the description. These are cases where the ascriber thinks that the definite description is *improper* (a singular definite description *the F* is improper iff it's not the case that there is exactly one *F*). Thus, on seeing Smith's dismembered corpse, Jones may leap to the conclusion that he was murdered and say "Smith's murderer must be insane"; this is a 'whoever that is' use of a description (Donnellan 1966; I am assuming "Smith's murderer" is a form of 'the murderer of Smith'). But if Black knows or believes that Smith was in fact savaged to death by an escaped tiger, she will not make ascriptions like "Jones thinks Smith's murderer is insane" or "Jones expects the police to capture Smith's murderer quickly", which is puzzling if we have the practice of making *de dicto* ascriptions to reflect the content of the subject's attitudes and there is no reason to doubt that Jones's statement "Smith's murderer must be insane" expresses what he sincerely believes.

This reluctance to ascribe may be a result of pragmatic considerations. One reason to think so is that even in the circumstances of the case, it seems that Jones can properly *self*-ascribe notionally with "I believe Smith's murderer is insane". If Black asserts "Jones believes Smith's murderer is insane" just before realizing she shouldn't, and if "believe Smith's murderer is insane" is univocal between

Black's ascription and Jones's self-ascription, the difference in acceptability most probably has to do with the shift in context of utterance, specifically the shift in speaker. One might flesh this out in terms of 'the' being a presupposition-trigger, entailing, even when in the scope of normally entailment-cancelling operators such as negation, that the restriction is uniquely satisfied, which in our case means that exactly one person murdered Smith. And since Black knows that Smith wasn't murdered, she will not say anything that entails that he was. Non-factive attitude verbs are often said to suppress the triggering ('projection') of presuppositions (see Kadmon 2001:116), but in view of Taylor's examples, this may be wrong, or at least too simple.

One problem with this presuppositional theory is that if Black's belief-ascription entails the existence of a murderer, and is therefore not true, then Jones's self-ascription "I believe Smith's murderer is insane" is not true for the same reason, even though it seems true. An alternative approach would be to explore weaker pragmatic accounts of the difference between Black's ascription and Jones's self-ascription, along the lines that using a definite description in a belief-ascription conveys (merely) that the ascriber grants or takes the description to be proper; and cooperative speakers who know this don't use descriptions they think improper. The question would then be how this implicature arises.

So far as undermining the idea that there are *de dicto* or notional ascriptions goes, one might say that the use of presupposition-triggers in the content sentence creates a principled exception. For the phenomenon noted above seems to recur with other triggers. Jones may say 'I think I will manage to save enough money', but Black will not report 'Jones thinks that he will manage to save enough money'

unless Black grants Jones's presupposition that saving enough money will be difficult. For if Black knows that the sum is small and that Jones can easily afford it, she will not want to use 'manage', unless ironically.

There is also a question about how manifest the phenomenon that Taylor isolates is with other quantifiers. If Jones says 'everyone who attacked Smith will be brought to justice' (he now thinks there were multiple killers), would Black, who knows about the tiger, happily report 'Jones thinks everyone who attacked Smith will be brought to justice', even though Jones says so? *If* the report seems infelicitous, that may be a point in favor of a pragmatic account if it is combined with a presuppositional account of 'every F' in 'every *F* is *G*'. According to such an account, the restrictor *F*, in this case 'person who attacked Smith', is presupposed to be non-empty (see Heim and Kratzer 1998:159–172). If we have found a mechanism that makes the use of a description in an ascription's content sentence convey that the *ascriber* takes the description to be proper, the same mechanism might well make the use of 'every *F*' in the content sentence convey that the ascriber takes *F* to be non-empty.

(Sosa 1970) has an interesting example which tries to undercut the *de re*/*de dicto* distinction by suggesting that there are no hard and fast limits on exportability and so no substantial cognitive relation which must hold between the subject *S* and the thing (*res*) *S*'s attitude is about (*de*). In an extreme case (Sleigh 1968), if *S* believes there are spies, but only finitely many, and no two have the same height, *S* may infer and come to believe 'the shortest spy is a spy', and Sosa would allow the exported ascription 'the shortest spy is someone *S* believes is a spy'. So if Phil Kimbly is the shortest spy, Phil Kimbly is someone *S* believes is a spy (strangely, *S*,

though the most upright of citizens, never even thinks of contacting the FBI).

The argument for this *laissez-faire* stance about exportation is that there are many examples where it is perfectly natural. For instance (Sosa 1970:890), the Commanding Officer may say to the captain, 'tomorrow I want the shortest platoon member to go first' or 'I {believe/doubt} the shortest platoon member should go first tomorrow'. The CO has no idea who the shortest platoon member is, but in fact it's the unfortunate Smith again (this is before he meets the tiger). The captain knows Smith is the shortest, and says to the sergeant, 'The CO wants Smith to go first tomorrow' / 'The CO {thinks/doubts} Smith should go first tomorrow', or to Smith, 'The CO wants *you* to go first tomorrow'. It seems perfectly natural for the captain to say such things, yet it looks as if his ascriptions are arrived at by first exporting a description which was used by the CO in a whoever-that-is way, and then substituting a name or pronoun. But shouldn't we object to the exporting, on the grounds that the CO doesn't have a desire or belief or doubt *about* Smith, that such-and-such? His desire that the shortest platoon-member go first seems to be no more *about* Smith than *S*'s belief that the shortest spy is a spy, arrived at as described, is *about* Phil Kimbly. But why then is 'The CO wants Smith to go first tomorrow' so natural?

According to Kripke (2008:348), examples like these are 'toy duck' cases: a child in a toy store points at a stuffed animal, asking his mother if it's a goose, and she replies 'No, it's a duck.' Kripke implies that what the mother says, no matter how natural, can't really be true: "no dictionary should include an entry under 'duck' in which ducks…may not be living creatures at all" (346). Another example might be that you and I go to an exhibition of the work of a famous forger who specialized

in analytic cubism. Pointing at one of his forgeries on the wall, I ask 'Is that a Picasso?', to which you reply, 'No, it's a Braque'. This is a very natural conversation, but the painting isn't really a Braque, and we shouldn't explain the use of artists' names as predicates of their works in a way that permits an NN not to be by NN.

Of course, the simplest explanation of the naturalness of these dialogues is that the remarks "It's a {duck/Braque}" *are* true, even though the duck's made of artificial fibres and Braque had nothing to do with the Braque (see Partee 2003 for how this could be). So if we follow Kripke in rejecting that explanation, we need to find another. Fortunately, at least in Sosa's case of 'The CO wants Smith to go first tomorrow', it's not hard to see what the naturalness consists in: Smith is the person whose going first tomorrow will satisfy the CO's desire that the smallest platoon-member, whoever he is, go first tomorrow; and Smith is the person whose going first tomorrow would realize the quantified eventuality the CO believes or doubts should obtain. Rather than leave it up to the sergeant to find out who the relevant individual is, the captain just tells him, and rather than do so by some laborious step-by-step reasoning about how to satisfy the CO's desire, the captain makes an attitude ascription that's strictly false, but serves both his and the sergeant's interests in seeing that the CO's order is obeyed; for to obey the order, an individual has to be identified. By contrast, the Phil Kimbly ascription seems unnatural because there is no surrounding context to give it a rationale. Perhaps we could invent one, but doing so wouldn't turn an incorrect exportation into a correct one, and nor does it in Sosa's example. An ascription can be well-motivated and promote efficiency in communication, but still be literally false.

c. The *de re* and Leibniz's Law

Assuming that the *de re*/*de dicto* distinction survives sceptical attack, there is one more issue we can address with its aid. At the start of this essay we distinguished Leibniz's Law, 'if *x* and *y* are the same object, then *x* and *y* have the same properties', from the inference rule of Identity Elimination. Problem cases for the rule might suggest that the Law itself is dubious. Why have we not considered this possibility?

The reason is that the Law is formulated in terms of objects and properties, and to regard examples like (2)–(5) as threats to it, we would have to construe these inferences as specifying properties of objects in their minor premises; but when we do this, we see that the apparent threat to the Law fades, as follows.

(3) is a 'wrong object' case, for (3b) ascribes a property to a word, but in (3a) the objects *x* and *y* are cities. (4) is a case of failure to specify a property of an object: (4b) seems to involve the property *being so-called because of its size*, but the italicized phrase fails to specify a property, because of the uninterpretability of 'so' as it occurs in it: 'so' needs a context, linguistic or otherwise. There is certainly at least one property of objects in the offing, that of having a name which was endowed on the basis of size. But in conformity with the Law, *that* property is shared with Barbarelli, and the sentence attributing it, 'Giorgione has a name endowed on the basis of his size', falls short of what (4b) says. There is also the property *being called 'Giorgione' on account of size*, also shared with Barbarelli; see Forbes 2006:154–7 on how ascription of this property relates to (4b).

As for (5), there is certainly a reading of (5b) in terms of properties of objects: the property of contingently being 9 is ascribed to the number that numbers the

planets. But then (5b) is false, since the number that numbers the planets is 9, and 9 isn't *contingently* 9. In other words, this property-of-objects construal requires a *de re* reading of (5b), with the description 'the number that numbers the planets' exported, resulting in a falsehood.

Another property-of-objects construal of (5b) is one where the property is contingency and the object is the proposition that the number that numbers the planets is 9. On this reading, (5b) is true. But this turns (5) into another wrong object case, since in the major premise the objects are numbers, not propositions. And if we change (5a) to make it about propositions, it would have to say that the proposition *that the number that numbers the planets is 9* is the same proposition as the proposition *that* 3^2 *is 9*. If (5) is reformulated this way, it is clearly a correct use of =E, but the falsity of the conclusion, that the proposition that 3^2 is 9 is contingent, tells us that the rewriting of the major premise to state an identity between propositions produced a falsehood.

So what of the original (2)? Here the property-of-objects construals of the minor premise are parallel to those in (5), but we don't want to say quite the same things about them. One property-of-objects reading of (2b) is that Superman has the property of being believed by Lois to be able to fly. (2a) is an identity involving Superman, so certainly we can use =E, in this case to infer that Clark has the property of being believed by Lois to be able to fly. This is just a slightly different formulation of the way of understanding the argument that we identified above as underlying any inclination there might be to say that (2) is valid: the crucial point is that the names that are syntactically in the scope of 'believes' are interpreted semantically to be exported from its scope. But we don't arrive at (2c), understood

as false: that would require *importation* of 'Clark' back into the scope of 'believes', and the fact that (2c) is by default understood as false shows that importation is invalid.

As with (5), we can reconstrue the minor premise and conclusion of (2) to be specifically about propositions. (2b) would then say that the proposition that Superman can fly is believed by Lois, and (2c) would say that the proposition that Clark can fly is believed by Lois. To prevent this just being another wrong-object case, (2a) would then have to be changed to an identity between propositions. Specifically, it would assert that the proposition that Superman can fly is the same proposition as the proposition that Clark can fly. The =E inference is then entirely in accord with Leibniz's Law. The problem, of course, is that the asserted identity between the propositions is false.

Perhaps we should say, then, that (5) is *partially* instructive as regards (2), in that there are parallel property-of-objects readings. What (5) doesn't help with is the formulation of a restriction on the terms used in =E that allows syntactically unstructured individual constants to be substituted in formulations like those actually used in (2), and there seems to be no way to do this.

3. Frege's Theory of Substitution-Resistance

According to the framework for semantics of natural language sketched in (Frege 1892), every meaningful phrase of natural language has potentially two sorts of meaning, a reference ('Bedeutung') and a sense ('Sinn', a cause of many puns in article titles – see, e.g., Dummett 1975, Ch. 17, Burge 1979, Forbes 1990, Salmon 1990; for issues about the translations of the German words, see the discussion

and references in Kripke 2001:254, n.1). A meaningful expression *e*, or a use of *e*, *expresses* a sense. Its sense determines its reference (if it has a reference) by virtue of being a *way of thinking* (or 'mode of presentation') of that reference, but whether there *is* a reference can depend on how things are in the world. In the case of a singular term, the reference is the thing it designates, and this reference is determined by the sense of the term. For example, the sense of the name 'Aristotle' might be articulated by 'the pupil of Plato who tutored Alexander and wrote the Nicomachean Ethics'. Whether or not the name 'Aristotle' has a reference then turns on whether or not there was such a person.

The same is true of sentences. A sentence expresses a *thought*, or, in contemporary jargon, a *proposition*, and a proposition with a reference refers to a truth-value, TRUE or FALSE (the idea that sentences refer is somewhat odd, but see Dummett 1975:180–86). For example, the proposition *that Aristotle was a philosopher* is a way of thinking of a truth-value: this proposition is the proposition that the pupil of Plato who tutored Alexander and wrote the Nicomachean Ethics was a […] (here readers should substitute their favorite explanation of 'philosopher' for the ellipsis, but please, not 'one who philosophizes'). Assuming that there was such a person, then this proposition is presumably a way of thinking of TRUE. However, if 'Aristotle' lacks a reference because there was no such person, then the sentence 'Aristotle was a philosopher' will lack a reference because it has a part that lacks a reference.

It's an important point about this apparatus that the calculation of the reference of the whole sentence proceeds through the reference of the parts. In the case of 'Aristotle was a philosopher', the reference of the whole sentence is obtained by

composing the references of 'Aristotle' and 'was a philosopher', as determined by their senses, in a way which produces a truth-value as output. So it is convenient to think of the reference of 'was a philosopher' as a function, which applied to an object, produces a truth-value. Then if 'Aristotle' provides an object, we will get a truth-value. But if there was no such person, this computational process will hang or crash, which motivates the verdict that in case the name is empty, the sentence is neither true nor false.

a. The sense/reference distinction applied to attitude ascriptions

The sense-reference distinction suggests that we may be able to explain how (13a) below can be true while (13b) is false:

(13) a. Lois hopes Superman is nearby.

b. Lois hopes Clark is nearby.

Assuming that the names have different senses (perhaps 'the red-caped superhero who flies' vs. 'the mild-mannered Daily Planet reporter with a crush on Lois Lane'), (13a) and (13b) will express different propositions because their embedded content sentences do, and so (13a) and (13b) at least potentially may refer to (i.e., have) different truth-values. But in fact the truth-values of 'Superman is nearby' and 'Clark is nearby' cannot differ just because they *express* different propositions. Superman *is* Clark, so, even though the propositions expressed by 'Superman is nearby' and 'Clark is nearby' are different, these sentences must have the same truth-value, given a fixed context to determine what counts as 'nearby'.

We can then scotch the idea that we are approaching an account of how (13a) and (13b) might differ in truth-value by the following reasoning. The references of

(13a) and (13b) are calculated from the references of their three main constituents, (i) 'Lois', referring to Lois, (ii) 'hopes', referring to the hoping relation, and (iii) 'Superman is nearby' and 'Clark is nearby' respectively, which, as we have just noted, refer to the same truth-value. Since (i) and (ii) are common to (13a) and (13b), (13a) and (13b) must also have the same reference, i.e., same truth-value, even if they express different propositions by virtue of having content subsentences that express different propositions.

We can explain reference for expressions other than singular terms and sentences in terms of functions, that is, input-output operations. For example, 'hopes' refers to a function *f* that takes a truth-value as input, say the truth-value of 'Superman is nearby', and produces another function, *g*, which is the reference of the verb-phrase 'hopes Superman is nearby'. *g* takes the referent of the name 'Lois' as input and produces the truth-value of (13a) as output. The problem is then that 'Superman is nearby' and 'Clark is nearby' present the same truth-value to *f*, which must therefore output the *same* function *g* as the referent of the two verb-phrases 'hopes Superman is nearby' and 'hopes Clark is nearby'. Thus Lois is mapped to TRUE by both verb-phrase functions, or to FALSE by both, since they are both the function *g*; and so (13a) and (13b) are equivalent.

The source of the difficulty is clear: we have taken the reference of 'hope' to be a function of the *truth-values* of content sentences that follow it. This is not arbitrary, for the calculation of the reference of any complex phrase uses the references of its constituent phrases along the way, and the content sentence of the ascription does indeed refer to a truth-value, at least when asserted in isolation, or more broadly, when it *occurs extensionally*, not in an intensional

or hyperintensional context. But this leads to a very unintuitive account of the reference of 'hope'. The thing that the attitude of hoping is taken towards is surely a *proposition*, not a truth-value: the *proposition* that Superman is nearby is what Lois hopes to be true, not the proposition's truth-value, whatever hoping *it* to be true would mean.

So, on the one hand we want 'hope' to take the reference of its complement sentence as its input, because reference is computed from referents. On the other hand, we want 'hope' to take the proposition expressed by its complement sentence as input, because it is propositions whose truth we hope for. But the proposition is the *sense* of the content sentence, not the reference.

To solve this conundrum, Frege made a move of what Kaplan describes as 'brilliant simplicity' (Kaplan 1969:117): we attribute to attitude verbs the property of switching the reference of the material that follows in the ascription *from* the 'customary' reference of that material *to* a different reference, namely, the customary sense (sometimes known as the 'indirect' reference). So in (13a), the (customary) reference of 'hopes Superman is nearby' is obtained by applying the (customary) reference of 'hope' to the reference 'Superman is nearby' has in (13a), which is its indirect reference, that is, its customary sense. Thus the reference of 'hope' is provided with the proposition that Superman is nearby as input, as we wanted. This means the concept of reference is relativized to linguistic context of occurrence. If 'Superman is nearby' occurs extensionally, it refers to its truth-value. But if 'Superman is nearby' is the S-part of a complex phrase V+S, where V is an attitude verb, 'Superman is nearby' refers to its sense, the proposition that Superman is nearby.

On this account, 'hope' refers not to a function that takes a *truth-value* and produces, as the meaning of the verb phrase 'hopes Superman is nearby', a function that takes individuals (such as Lois) to truth-values. Rather, it refers to a function which takes a *proposition* as input, for example the proposition that Superman is nearby, though it still produces, as the meaning of the verb phrase 'hopes Superman is nearby', a function which maps some individuals, like Lois, to TRUE, and others, like Lex Luthor, to FALSE. However, since we have already agreed that 'Superman is nearby' and 'Clark is nearby' express different propositions (when occurring extensionally, as we would now add) because of the different senses of 'Superman' and 'Clark', this means that the input to the reference of 'hope' in (13a) is different from its input in (13b): two different propositions, rather than the single truth-value which is all that's available in the absence of reference-switch. Consequently, it's possible that the verb-phrases 'hope Superman is nearby' and 'hope Clark is nearby' refer to different functions; perhaps 'hope Superman is nearby' refers to a function which maps Lois to TRUE, while 'hope Clark is nearby' refers to a function which maps Lois to FALSE. This is Frege's account of how (13a) and (13b) can differ in truth-value, and is the first example of what's nowadays called 'switcher semantics'(Gluer and Pagin 2012).

The reference-switch thesis has immediate application to the question of what is wrong with (2). The Fregean answer is that (2) is a fallacy of equivocation. In (2a), 'Superman' and 'Clark Kent' have their customary referents, namely, Kal-El. But in (2b), 'Superman' refers to its customary sense, the concept of being the red-caped superhero who flies; 'is nearby' also refers to its customary sense. If we say that in any speaker's mouth, 'is nearby' when it occurs extensionally has the sense 'is close to me', then in (2b), 'Superman is nearby' refers to the proposition

Lois would express there and then with 'the red-caped superhero who flies is close to me' (an alternative where it's the ascriber's location that matters is also available). Similarly, 'Clark is nearby' refers to the proposition that Lois would express with 'the mild-mannered Daily Planet reporter with a crush on Lois Lane is close to me'. As the example shows, identity of customary reference does not justify substituting one singular term for another in the content sentence of an attitude attribution, since identity of customary reference falls far short of identity of indirect reference (identity of sense).

Indeed, Frege's theory predicts that it will be hard to find any non-trivial sound arguments in the style of (2), even if we change the major premise to be of the form 'the sense of t_1 = the sense of t_2'. For then the major premise is true only if two different names have the same sense, and it's not clear under what circumstances that would happen. Perhaps it might be self-evident in the circumstances of acquisition that the names refer to the same person: the speaker introduces herself to *x* with 'Hi! My name is Kimberley, but people call me Berry.' But even if *x* correctly recalls this, Mates cases can be constructed: *x* may coherently wonder if everyone knows that Kimberley is Berry. Perhaps we should say that for *x*, for a while, the two names have the same sense, but *x* envisages that others may use the names with different senses, and the semantics of 'everyone knows that Kimberley is Berry' allows, one way or another, for this possibility. (See also Schiffer's discussion of the individuation of senses, Schiffer 1992:502–3.)

b. The hierarchy problem

There are problems of detail with Frege's theory. Perhaps the best-known is the 'infinite hierarchies' problem. As we have already seen with Mates sentences, an

attitude ascription can be embedded within an attitude ascription. A simple case is:

(14) a. Kal-El wonders if Lois has begun to notice that Clark is never around when Superman is.

b. Lois has begun to notice that Clark is never around when Superman is.

c. Clark is never around when Superman is.

According to Frege, (14a)'s 'Lois has begun to notice that Clark is never around when Superman is' refers in (14a) to the sense it expresses in (14b), since it is within the scope of 'wonders' in (14a). And (14b)'s 'Clark is never around when Superman is' refers in (14b) to *its* customary sense, the sense it expresses in (14c) (curiously, (14c) also seems to manifest substitution-resistance, despite the lack of attitude verbs; we will return to this in our discussion of 'simple sentences'). These sentence-senses are obtained systematically from the senses of their constituent words. So in (14b), 'Clark' refers to the way of thinking of Kal-El it expresses in (14c), which we label m_1. But *whenever* a word refers, it does so by expressing a way of thinking of that reference. So 'Clark' in (14b), referring as it does to m_1, must express a way of thinking of m_1, which we label m_2. Plausibly, m_2 cannot be m_1 over again, for (i) $m_2 = m_1$ would require the same way of thinking to be of both a person, Clark, and a way of thinking of that person, m_1; and, (ii), $m_2 = m_1$ means that m_1 is a way of thinking of itself, an idea not breathtaking in its intelligibility (see further Peacocke 2009:162–3; but see also Dummett 1973:264–9 for an attempt to get by with just m_1). So these considerations motivate the thought that in (14b), 'Clark' expresses a way of thinking m_2 which is of m_1 and not identical to m_1.

Now, (14b) occurs in (14a) within the scope of the hyperintensional 'wonders', so

its reference in (14a) and the referents of its constituent words in (14a) must switch; they switch from the referents they have in (14b), to the senses they express in (14b). This means that in (14a), 'Clark' refers to m_2. But then, 'Clark' in (14a) must express a sense which is a way of thinking of m_2, since this is the only way 'Clark' could refer to m_2. Call this sense m_3. As before, it is implausible that m_3 is the same as m_2, since, first, it would have to be a way of thinking of itself, and second, it would have to be both a way of thinking of m_2, but also, since *ex hypothesi* it *is* m_2, would have to be a way of thinking of m_1. m_3, then, appears to be something new.

And so we are off. We can make (14a) the content sentence of a new attitude ascription, say

(15) Lex suspects that Kal-El wonders if Lois has begun to notice that Clark is never around when Superman is.

Now the sense (14a) expresses becomes the reference of (14a) as it figures as the content sentence of (15), and the words of (14a) will express new senses in (15), ways of thinking of the senses they express in (14a); e.g., in (15), 'Clark' will express m_4, a way of thinking of m_3, so that 'Clark' in (15) can refer to m_3. Since there is no principled restriction on how deeply attitude verbs may be embedded within other attitude verbs, we have, apparently, an unending sequence of senses. In particular, 'Clark' can express infinitely many ways of thinking, hardy any of which are intelligible beyond the first. Some Frege scholars have developed formal models of sense and reference which embody such hierarchies; see, e.g., (Church 1951), (Anderson 1980). However, others have tried, in effect, to stop at m_2; see especially (Parsons 1981, 2009); also (Forbes 1987).

c. The semantic innocence objection

But problems of detail aside, there are two main objections to Frege's account of substitution-failure which have emerged in the last few decades, the *semantic innocence* objection and the *no-such-thing-as-senses* objection. We describe the former first.

The semantic innocence objection is so-called because of its famous statement by Davidson (1969:172): 'If we could recover our pre-Fregean semantic innocence… it would seem to us plainly incredible that…words [in the content sentences of attitude attributions] mean anything different, or refer to anything else, than is their wont when they come in other environments.' This is, admittedly, simply an appeal to intuition, but it is a powerful one (see also Loar 1972:43). It is indeed very difficult to detect a switch in the reference of 'Superman' if Lois remarks 'Superman is nearby, with any luck' versus if she remarks 'I hope that Superman is nearby'. The reference-switch thesis also causes problems in other areas, for example, over the reference of anaphoric pronouns. In 'Galileo thought that the Earth moves, and he knew what he was talking about, so it moves', it's undeniable that the 'it' refers to the Earth. But then the pronoun does not directly inherit its reference from its antecedent (see further Segal 1989). No doubt we can find ways round this, but it's questionable whether the road is worth going down, given the lack of intuitive support at its starting point.

d. Do senses exist anyway?

An even more damaging objection to Frege's account of substitution-failure is that the entities which play the crucial role, senses or ways of thinking, are chimerical. That Fregean senses don't exist is the core argument of (Kripke 1972). Briefly

(Kripke's arguments are very well-known), suppose that 'Aristotle' did express a reference-determining sense captured by the singular definite description 'the pupil of Plato who tutored Alexander and wrote the Nicomachean Ethics'. One possibility is that this description articulates the meaning of the name in much the way that a dictionary might articulate the meaning of 'philosopher'. Then it should be both necessary and *a priori* that Aristotle tutored Alexander. But it is neither. Aristotle could have been killed in an Athenian traffic accident in his early teens, so it is not necessary that he tutored Alexander; and that he in fact did so is clearly an empirical claim, which only historical evidence can confirm or disconfirm. As these points show, not even 'if Aristotle existed, he tutored Alexander' or 'if Aristotle and Alexander existed, the former tutored the latter' is necessary or *a priori*.

A somewhat weaker thesis is that the reference of 'Aristotle' is *fixed* by the description, without being in any way synonymous with it. But even merely this would predict, of some perfectly intelligible claims, that they are impossible to understand. For example (based on Kripke's 'Gödel case', 1972, 1980:83–5), suppose that someone claims on a fake-news website to have found documents showing that Aristotle wasn't a pupil of Plato, didn't tutor Alexander and didn't write the Nicomachean Ethics. The first two items he deliberately falsified on his c.v. in order to attract students, and though he *published* the Nicomachean Ethics under his own name, that was after stealing the manuscript from the true author, whom he murdered to ensure his silence. And as time passed, the false claims became firmly lodged in popular lore about Aristotle.

If it went viral, this story about Aristotle would outrage historians of philosophy.

But the very fact that they would be outraged shows that they understand the story well enough. Yet, if the reference of the name were *fixed* by the description, the story should make no sense: *Aristotle* could not have lied about tutoring Alexander, for 'Aristotle' by its very semantics refers to someone who did tutor him, if it refers at all. At best, the author of the story should be claiming to have discovered that there was no such person as Aristotle, not that there was and he behaved reprehensibly. But no historian would contest the fake-news story on the grounds that it is incoherent: the debate would be over the existence or trustworthiness of the documents that the story is based on.

Perhaps this just shows that we picked the wrong description as the sense of 'Aristotle'. It should have been 'the person *commonly thought* to be a pupil of Plato who tutored Alexander and wrote the Nicomachean Ethics', or more carefully, since the sense-giving descriptions are being used in a 'whoever' way, 'whoever it is who is the person commonly thought to be a pupil of Plato who tutored Alexander and wrote the Nicomachean Ethics'. Kripke shows that this is vulnerable to counterexamples involving subjects who haven't kept up with what's commonly thought about whom (1980:88). And he raises a circularity objection (*loc. cit.*). The new description identifies Aristotle as the person commonly thought to be thus-and-so, that is, as the person to whom certain thoughts refer. In Frege's system, those thoughts refer to Aristotle only if their subject sense is a way of thinking of Aristotle. But on the current proposal, this sense simply refers back to some of the same thoughts. There seems to be no way to break out of this loop and give a substantial answer to the question, what makes it the case that *this* individual *x* rather than *that* one *y* is the reference of the sense?

Kripke also points out that we manage to refer easily enough even when there are no identifying descriptions we could cite. He gives the example of 'Richard Feynman', a name many people use without having an associated definite description (1980:81 – this was before Feynman's testimony at the Challenger disaster inquiry). An associated indefinite description might be 'a famous physicist at Caltech who won the Nobel Prize'. But 'a' cannot be strengthened to 'the', since Murray Gell-Mann is also a famous physicist at Caltech who won the Nobel Prize. And if we insert 'not identical to Gell-Mann' into the description, we make it impossible to refer to Feynman without having heard of Gell-Mann (not to mention the looming indeterminacy problem).

e. Alternative accounts of the sense of a name

If Kripke's arguments show that Fregean senses do not exist, then the Fregean solution to the problem of opacity collapses, rather like a well-worked-out theory of human behavior in which demonic possession plays a large and crucial role. However, it would be fair to say that Kripke's counterexamples tell mainly against 'famous deeds' descriptivism and some modifications of it involving qualifiers like 'commonly thought'. It's reasonable to focus on famous-deeds descriptions, since Frege says that everyone who uses the name expresses a reference-determining sense with it, and so to guarantee that there is such a sense one looks to information that is commonly associated with the name by those who have learned it. But perhaps there are other options for the content of name-senses besides famous deeds.

One alternative, developed by Chalmers, is *two-dimensional* sense. A two-dimensional sense is an ordered pair consisting in an *epistemic* sense and a

subjunctive sense. For a name, the epistemic sense is a function from 'scenarios' to individuals, and the subjunctive sense is a function from possible worlds to individuals (Chalmers 2011:596–99). A scenario is something like a coherent total description of how things might have turned out to be, and the epistemic sense of a name may be a non-rigid function on such items: in one scenario, a name may refer to *x*, while in another it may refer to a distinct *y*. But subjunctive senses are rigid: they denote the same object in any two worlds. The idea is then that epistemic operators are sensitive to the epistemic sense, and modal operators to the subjunctive sense, which, since it's a rigid function, may be identified with the object it stably refers to (2011:597, T4, T5).

If epistemic senses are just famous deeds descriptions or their like, Kripke's objections arise over again. And it would certainly be unfortunate if epistemic and subjunctive senses came apart over actual reference, since then statements like "it's *a posteriori* that Aristotle was a philosopher" and "it's contingent that Aristotle was a philosopher" would be about different people. However, Chalmers has a proposal on which this difficulty and certain others won't arise.

Asking what might replace a famous-deeds descriptivist account of how names refer, Kripke suggested what's come to be known as the 'historical chain' account: "someone, let's say a baby, is born; his parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by a chain… it's in virtue of our connection with other speakers in the community, going back to the referent himself, that we refer to a certain man" (1972; 1980:91–4). The same idea was advanced by Geach: "for the use of a word as a proper name there must in the first

instance be someone acquainted with the object named…But …the…name…can be handed on from one generation to another…Plato knew Socrates, and Aristotle knew Plato, and Theophrastus knew Aristotle, and so on in apostolic succession down to our own times. That is why we can legitimately use 'Socrates' as a name the way we do" (Geach 1970:288–9). One thing required for *x* to refer to Socrates with 'Socrates' nowadays, then, is that *x* belong to a linguistic community in which there's an apostolic succession from Socrates to *x* through which the name 'Socrates' is passed along. (Following Kripke, *x* also has to intend to defer in *x*'s use of the name to those from whom *x* acquired it – if *x* decides that "Socrates" would be a fine name for *x*'s pet turtle, that doesn't count.)

Kripke mentions that Nozick once remarked to him that if any theory of reference is correct, some descriptivist theory is immune to counterexamples that are in the style of *Naming and Necessity*. This would be a descriptivist theory on which the descriptions are 'theory-laden': they incorporate the reference-determining conditions the correct theory articulates (Kripke 1980:88, n.38). Chalmer's idea is to exploit this option: taking the historical chain theory as a highly plausible account of reference-determination, he suggests that the epistemic sense of a name NN might just be 'the object NN refers to in the mouths of those from whom I acquired it' or its like (Chalmers 2002:641). This will be a non-rigid function, since in some scenarios, indeed, some possible worlds, the apostolic succession for 'Socrates' will lead to contemporary users but start from an individual *x* who isn't Socrates.

Since the description suggested above involves the term 'refer', there is an obvious circularity worry if the sense is to be reference-determining. Chalmers

argues (2002:641–3) that there is no reason to worry, since the evaluation of one person's epistemic sense takes us back to other people, and their epistemic senses will carry us back to even earlier people, until we arrive at the 'initial baptism' with which the name was introduced. The question would then be whether the concept of reference is ineliminably invoked at this point, as in 'we are introducing the name NN to refer to this child', and how significant a problem that would be.

A second question is whether epistemic senses are *otiose* as far as determining reference is concerned. Is the reason why I can use 'Socrates' to refer to Socrates not simply that I belong to a community in which there is a chain of uses of 'Socrates' linking me to Socrates in the way the historical chain theory describes, and I have added the name to my repertoire with the intention to use it in a way that preserves the reference of those from whom I acquired it? Perhaps adding the name to my repertoire with such a deferential intention is the very same thing as attaching a theory-laden sense to it. But if not, the postulation of an epistemic sense is redundant: the reference of the name in my mouth is already determined by the historical facts, and if I express a certain epistemic sense with it, that is just a private epiphenomenon.

Indeed, even if the epistemic sense determines the same object as the working historical chain, there may be reasons not to postulate it besides avoiding redundancy. For example, one might subscribe to the principle that an explanation of the epistemic sense of NN should not employ concepts that a competent user of NN need not possess, and the notion of reference is precisely such a concept. Nevertheless, the example of theory-laden sense does show that

the array of options beyond famous-deeds descriptivism is not empty.

A second member of the array is what we might call 'cognitive descriptivism', since it is based on a (somewhat metaphorical) hypothesis about cognitive architecture. The idea is that we organize our information about what we take to be separate objects that we have encountered into separate mental files, or dossiers. This seems to have first been proposed by Grice (1969:141–44), and was used in an account of the senses of names in (Forbes 1990). The neo-Fregean idea is that the sense of a name NN for *x* is 'the subject of *this* dossier', where the mental demonstrative 'this dossier' refers to the dossier labelled NN by *x* in *x*'s mental filing system.

Clearly, questions about circularity and redundancy arise much as they do for two-dimensional sense (see Fine 2007:67–8). If what makes *x* the subject of the dossier labelled NN is that *x* is the referent of the name NN, then we have circularity. But if being the subject of the dossier labelled NN consists in – to use the causal theory of (Evans 1973) – being the dominant causal source of the information in the dossier, why not cut out the detour through dossiers and just say that the reference of a name NN is the dominant causal source of information that would be expressed in statements of the form 'NN is…'? Such issues are pursued in (Recanati 2012), and are far from settled in the literature. But it's clear from these examples that famous deeds descriptivism is not in sole possession of the field as an elaboration of Frege's notion of the sense of a name.

However, whatever viable theory of sense may ultimately be produced, the semantic innocence objection will have to be dealt with. (Thomason 1980) is unmoved by it, but we shall next consider accounts of senses that may be invoked

by attitude ascriptions in a way that explains failure of =E, but allows those senses to have their customary, not indirect, references.

4. Hidden Indexical Semantics

The reference-switch hypothesis is one version of a more general idea that the words used in the content sentence of an attitude ascription have a special role that they don't play in other contexts. If we don't think that the special role *displaces* their normal role, we arrive at Loar's idea of a *dual contribution* (Loar 1972:52–3). As Davidson insists, the words of the content sentence play their normal role. But there is another semantic mechanism at work in which they are also complicit. There is a wide range of such accounts in the modern discussion of opacity, perhaps starting with (Loar 1972). (Field 1978) has the content sentence invoking a sentence of the 'language of thought'. (Bealer 1993) proposes an ambiguity theory, on which the content sentence of an ascription introduces both an entity composed of the referents of the words, thereby explaining the innocence intuition, and an entity like a Fregean proposition, thereby accounting for the intuition of substitution-resistance in the likes of (2). (Larson and Ludlow 1993) develops a semantics on which a propositional attitude is an attitude to an 'interpreted logical form', which is a tree-structure in which a node is occupied both by the reference of the expression at that node and the expression itself. Consequently, 'Superman can fly' and 'Clark can fly' are different ILF's simply in virtue of 'Superman' and 'Clark' being different names.

a. Two kinds of hidden-indexical theories

Some versions of the dual contribution approach are known as 'hidden indexical' accounts (Schiffer 1979 – but perhaps 'hidden indexicality' would

be better), because of the role context-dependence plays in determining the second contribution of the content sentence, or because there actually is an indexical expression postulated to occur covertly in the ascription. For example, in (Crimmins and Perry 1989) and (Crimmins 1992), belief ascriptions are said to be made true by items supplied by the context in which the ascription is made, items called 'unarticulated constituents' (because there is no expression in the ascription responsible for their intrusion into the truth-condition). Different but coreferential names may be associated with different *normal notions* of the same object, and an inference like (2) fails because the substitution changes which normal notion of Kal-El is, in their technical sense, 'involved'. Similarly, in Richard's theory (Richard 1990), the content sentence of a belief-ascription invokes a 'Russellian annotated matrix', which, like an interpreted logical form, is an item that contains both Fregean referents and the expressions referring to them, and the truth-condition requires that the RAM in the ascription *correlate* with a RAM believed by the subject of the ascription. What correlates with what is context-dependent, and (2) fails because substitution need not preserve correlation, even though it preserves Fregean reference (Richard 1990:133–41). While in (Forbes 1990, 1996; Recanati 2000:137–63) there is a hidden 'so' in belief-ascriptions, as if 'believes' were 'so-believes', which blocks substitution much as it does in Quine's 'Giorgione' case, (4), since the 'so' refers to the content sentence of the ascription.

One respect in which the above theories differ is over what kind of thing is believed. In Schiffer's general scheme for hidden indexical theories (1992:503–4), what is believed is a proposition of a non-Fregean kind, but the ascription includes as part of its literal meaning that this proposition is believed *under* a

way *w* of thinking of it. Here *w* is something rather like a Fregean proposition, and is specified by the very words used in the content sentence of the ascription. Substitution then has the side-effect of changing the relevant way of thinking, say from the 'Superman can fly'-way to the 'Clark can fly'-way, and this opens the door to change of truth-value. The kind of proposition of which *w* is a way of thinking is termed by Schiffer a 'Kaplan proposition', but is more commonly known as a 'Russellian' proposition, after a famous exchange between Russell and Frege (Frege and Russell 1904). Frege had claimed that Mont Blanc with its snowfields is not itself a component of the thought that Mont Blanc is more than 4,000 meters high, to which Russell replied that 'in spite of all its snowfields Mont Blanc itself is a component part of what is actually asserted…a certain complex'. Accounts of Russellian propositions have been given in some detail (e.g., Crimmins 1992:117–24, but see Jespersen 2003 for critical discussion), and in Schiffer's scheme, attitude ascriptions invoke quasi-Fregean ways of thinking of such complexes, while the attitude itself is to a Russellian proposition.

In the approach of (Forbes 1990, 1996), however, it is a Fregean proposition to which an attitude is held, but one that is specified as the way of thinking of the referent of the content sentence, where this way is determined by that very sentence. The referent is not a truth-value, as Frege would have had it, but rather an abstract state of affairs, which is a structured entity not unlike a Russellian proposition, but seems to fit better into a Fregean scheme. For instance, (2a) becomes

(16) That Superman can fly is so-believed by Lois

or more long-windedly,

(17) Lois believes her so-labelled way of thinking of the state of affairs that Superman can fly

in which 'so' refers to 'Superman can fly', sealing it off from substitution in the same way as it does for 'Giorgione' in (4). (17) requires for its truth that the ascriber's content sentence be a 'linguistic counterpart' of some sentence of Lois's that she would use to express the belief that (17) is attempting to ascribe to her (*cf.* Richard's notion of correlation), a belief which is a way of thinking of the state of affairs that {Superman/Clark/Kal-El} can fly.

Unfortunately, (17) departs from (16) in a rather substantial, if not frequently noticed, way: the 'that'-clause disappears, and the clausal form of 'believes' is replaced by the transitive one (the direct object in (17) is everything following 'believes'). But though there seems to be an equivalence between *believing that*… and *believing the proposition* (*thought, so-labelled way of thinking*) *that*…, it does not generalize to other attitude verbs. For example, fearing that Lex Luthor is involved is not the same thing as fearing the proposition that Lex Luthor is involved (who fears propositions? – Moltmann (2003:82) credits Arthur Prior with first noticing this phenomenon). The same thing occurs, though maybe for different reasons in different cases, with such verbs as 'agree', 'announce', 'anticipate', 'ask', 'boast', 'calculate', 'caution', 'complain', 'conclude', 'crow', 'decide', 'discover', 'dream', 'estimate', 'forget', 'guess', 'hope', 'insist', 'judge', 'know', 'notice', 'observe', 'plan', 'prefer', 'pretend', 'rejoice', 'require', 'see', 'suggest', 'surmise', 'suspect', 'understand' and various cognates of these. The verbs for which the equivalence holds include inference verbs like 'deduce' and 'infer', plus a few other examples like 'doubt', 'establish' and 'verify'. However, it would take us too

far afield to address the issue of what to do for the verbs for which the equivalence fails in order to produce a version of (17) applicable to them (Forbes 2010 presents one way of dealing with the problem).

As the previous paragraph indicates, some hyperintensional clausal verbs that can be used to ascribe propositional attitudes have hyperintensional transitive forms that can be used to ascribe what we might call *objectual* attitudes. These seem to generate failures of =E much as their clausal counterparts do. For example, 'Lex fears Superman' is true, but 'Lex fears Clark' doesn't seem any more plausible than 'Lex fears that Clark will crush him'. The apparatus in (17) can be employed to express a hidden-indexical theory for the transitive verb case: the substitution-resistant reading of 'Lex fears Superman' is 'Lex fears Superman as such', or 'Lex fears Superman so-personified', and the references of the 'such' and 'so' will change if 'Clark' replaces 'Superman', producing the false 'Lex fears Clark {as such/ so-personified}'. A fuller version of the substitution-resisting semantics for 'Lex fears Superman' might be

(18) Lex fears Superman under the way of thinking of him that is so-labelled.

Here 'under' forms an adverbial phrase modifying the whole verb phrase in (18) headed by 'fears' (there is some dispute about how such an 'under' is to be accommodated; see Schiffer 1996, Ludlow 1996, Forbes 2000:152–4).

Hidden indexical theories all accommodate semantic innocence in roughly the same way: there is some entity, whether Russellian proposition or abstract state of affairs, determined by the *customary* referents of the words of the content sentence, so the result is compatible with Davidson's decrying any theory which

claims that words in attitude ascriptions abandon their customary referents for something else. The 'something else' is imported in a different way, as (17) and (18) illustrate.

Hidden-indexical semantics also offers an alternative formal account of the *de re*/*de dicto* distinction. Standardly, the difference is brought out in terms of scope distinctions, as we did in (10). But another possibility is that *de re* readings are those in which a hidden indexical refers only to a part of the content-sentence: if Lois believes that her coworker Mary has gone to St. Petersburg, we may point at Mary and say 'Lois believes that that woman is in St. Petersburg', meaning that she believes some way of thinking of the state of affairs, partially labelled 'is in St. Petersburg'. This would explain why the awkward locutions in (10) are rarely encountered in ordinary speech and writing.

b. Kripke's puzzle

One application of hidden indexical semantics is to Kripke's 'puzzle about belief' (Kripke 1979). Kripke wants to undermine the idea that there is a specific problem of interchange of coreferential names in attitude ascriptions, to be resolved by providing a semantics on which such substitution is evidently fallacious. Rather, he thinks substitutivity problems are a mere symptom of broader anomalies in psychological discourse ('It would be wrong to blame…substitutivity. The reason does not lie in any specific fallacy [in (2)] but rather in the nature of the realm being entered.' 1979:157). So he gives examples intended to bring out anomalies even in the absence of substitution.

The main example, though it appears late in his paper, is that of a subject, Peter, who encounters the same individual under the same name in different contexts

and does not realize it was the same person all the time. Peter goes to a recital by a pianist named Paderewski, and, picking up the name from the recital program, comes to believe on the basis of the performance that Paderewski has musical talent. Later, at an airport, he observes an individual surrounded by reporters, and someone tells him "That's Paderewski, the Polish Prime Minister". Far from connecting the man he sees with the man he heard play, Peter, who believes that no politician has musical talent, remarks out loud, 'Ah, a person of no musical talent, then'. But of course, Ignacy Jan Paderewski, the Prime Minister of Poland after the First World War, was also a celebrated composer and concert pianist.

Kripke wants us to try to answer the question, 'Does Peter, or does he not, believe that Paderewski has musical talent?', and in the course of our attempting to answer it, to realize that no answer can be given, because of 'the nature of the realm being entered'. However, from the Fregean perspective the example is less troubling, as Kripke recognizes (see also Taschek 1988). Peter has two lexical entries for 'Paderewski', in the same way that this writer has three for 'Socrates', one for the Ancient Greek philosopher, another for the late Brazilian footballer, and a third for the former Portuguese Prime Minister (the latter two individuals had different first names, but I do not know what they are, and I don't know if the first individual had any other name; on the individuation of names, see Kaplan 1990). Of course, the difference between Peter and myself is that the names in Peter's lexical entries are coreferential, while the ones in my three are, pairwise, not, unless the footballer, on retiring from the game, moved to Portugal and went into politics.

However, an ascriber A may only have one name for Paderewski (one mental

file so-labelled), which puts A at a certain expressive disadvantage relative to Peter, if the ability to make an accurate report about Peter's beliefs requires A to use names which match Peter's in a manner requiring A to have two names for Paderewski. But there is a very natural way of remedying this (which Kripke uses himself, in n.37): A can simply say that Peter believes that Paderewski the pianist has musical talent, while Paderewski the statesman does not (Forbes 1990:561). From the perspective of a semantics like that of (17), the appositive uses of 'the pianist' and 'the statesman' determine different ways of thinking of the single state of affairs that Paderewski had musical talent. And it is only the way of thinking labelled with Peter's linguistic counterpart of A's 'Paderewski the pianist has musical talent' that he believes: the appositives help us identify which of Peter's ways of thinking of Paderewski we wish to invoke in our ascriptions. The question remains to explain why the major premise that Paderewski the pianist is Paderewski the statesman does not license the inference to 'Peter believes that Paderewski the statesman has musical talent'. This would partly recapitulate our discussion of (2), but the appositives may raise further questions.

It's also conceivable that ascribers in the know about Peter's situation and who take their audiences also to be in the know, can rely on *context* to fix which belief is ascribed to Peter using 'Paderewski has musical talent'; for instance, if the discussion concerns Peter's evaluations of various pianists he has heard, the possessive description "Peter's so-labelled way of thinking" is proper, rather than improper, since the other way of thinking, labelled with Peter's linguistic counterpart of 'Paderewski the statesman has musical talent', won't be in the domain of the context, even if the discussion takes place after the airport encounter.

One can therefore resist Kripke's question whether Peter does or does not believe that Paderewski had musical talent, just as I would resist the question 'Was Socrates, or was he not, a chain smoker?' The footballer was, but (I suppose) the philosopher wasn't, so absent contextual clues I would require disambiguation of the question: 'Are you asking whether Socrates the footballer was a chain-smoker, or Socrates the philosopher?' In the Paderewski case there is no referential ambiguity, but there is still an ambiguity, or maybe indeterminacy, over which way of thinking of the state of affairs in question is being invoked: 'Are you asking whether Peter believes Paderewski the pianist has musical talent, or Paderewski the politician?', would be a perfectly proper question. The explanation why it is perfectly proper is clear enough on hidden indexical theories, but may not be so on others (see also Soames 2002, Chs.2,3).

Of course, this account only works if there is a viable notion of the sense of a name. For those sceptical about the prospects of such a thing, (Fine 2007) offers an alternative. Fine begins with an explanation of the difference between 'Superman is Superman' and 'Superman is Clark', the kind of pair where, as Frege would have said, one member can be a valuable extension of our knowledge, but the other won't be. According to Fine, in 'Superman is Superman', the two names are *coordinated*, but not in 'Superman is Clark'. One manifestation of this is that someone who wonders whether Superman is Superman thereby demonstrates a failure to grasp what is said; while one can wonder whether Superman is Clark without demonstrating any failure of understanding. Since Fine takes the coordinated/uncoordinated distinction to be of semantic import, his view could be regarded as neo-Fregean, since he thinks 'Superman is Superman' and 'Superman is Clark' have different semantics, though his view of how the

difference arises is quite unlike Frege's (nor is his account a hidden-indexical one).

He then argues that the case of Pierre presents us with a puzzle whose solution is to be formulated in terms of this notion of coordination (2007:100–105). The puzzle is that our normal practices of belief dictate that we report Peter as believing that Paderewski has musical talent, and that we also report him as believing that Paderewski has no musical talent. At the same time, according to Fine, we do not want to make a 'composite' report, that Peter believes that Paderewski has musical talent and believes that Paderewski has no musical talent, since this represents Peter as rather unreflective, which is not justified by the description of his predicament (further reflection isn't going to help). Yet the composite report is a simple 'and'-Introduction inference from the acceptable reports. How can it sensibly be resisted?

Fine's suggestion (2007:102–3) is that the composite report is unacceptable precisely because the reporter (who is in the know about Peter's situation) uses 'Paderewski' in a coordinated way across the content sentences of the composite report, while Peter does not use coordinated "Paderewski"'s in giving voice to his two beliefs. But the individual reports are acceptable, taken in isolation. Since there is nothing to be coordinated in the case of an individual report, we can simply take at face value Peter's assertion of 'Paderewski has musical talent', even asserted after he has made both entries in his lexicon, and ascribe such a belief to him. Whereas, for the Fregean, if there is nothing in the context to point towards one of 'Paderewski the pianist' and 'Paderewski the statesman' rather than the other, it will be indeterminate what belief is being ascribed (unless the ascriber's silent intentions settle it). And for the Fregean, the composite report, if it is the

conjunction of two determinate ascriptions, is acceptable. Perhaps it makes Peter sound unreflective; but so does 'GF believes Socrates was a chain-smoker and believes Socrates was not a chain-smoker'.

5. Russellianism

At the beginning of section 2 we noted that there is a possible response to the appearance of substitution failure in (2) according to which the reasoning is not flawed at all: if Lois believes Superman can fly and Superman is Clark, she simply *does* believe that Clark can fly, even though she wouldn't put it that way. The main motivation for this account is the view of propositions advanced by Russell in his previously quoted letter to Frege, according to which Mont Blanc itself, not a way of thinking of it, is the sole constituent the name contributes to the proposition about its height. The *locus classicus* of this theory is (Salmon 1986); other prominent contributions include (Soames 1987), (Saul 1997) and (Braun 1998).

a. Salmon's theory

According to Salmon, belief ascriptions invoke both Russellian propositions and *ways of taking* or of *grasping* those propositions. The apparently two-place attitude relation of belief unfolds into a three-place relation, with a position for a variable over ways of grasping. So for *A believes p* he offers (1986:111)

(19) for some way of grasping propositions w, A grasps p by means of w and BEL(A,p,w).

The correctness of the substitution inference (2) is immediate from this. The proposition that Superman can fly is identical to the proposition that Clark can fly,

and if (2b) is true, Lois has a way of grasping the proposition that Superman can fly under which she believes this proposition. Therefore, she has a way of grasping the proposition that Clark can fly under which she believes this proposition, for it is the same proposition as the former. Ways of grasping may be like Frege's ways of thinking in some respects, but they are not what is believed, and they are not meant to be reference-determining.

Also note that Fine's concern to avoid the composite ascription 'Peter believes Paderewski has musical talent and believes Paderewski has no musical talent' is allayed, since the composite ascription is harmless on Salmon's theory. For it involves two existential quantifiers over ways of grasping: there is some way of grasping the proposition that Paderewski has musical talent under which he believes it (more accurately, BELS it), and some way of grasping the proposition that Paderewski has no musical talent under which he believes *it*. There is no reason to think the second way of grasping is the same as the first but for the injection of negation, so there is nothing that imputes an intellectual deficiency to Peter (Salmon 1986:130–1).

The main question this account raises, of course, is why the verdict that (2) is valid is so hard to swallow. Or rather, since we may allow that there is a *de re* construal of (2) which all parties agree is valid, the question is why it is so hard to resist understanding it in a way on which it's invalid. Salmon proposes to deal with this question by distinguishing between *semantically encoded* and *pragmatically imparted* information (1986:78). As far as what's semantically encoded is concerned, (2b) and (2c) are the same. But they differ over what they pragmatically convey, and those who think (2b) and (2c) can have opposite truth-

value are mistakenly projecting the pragmatic difference onto the semantics. For example, it may be that (2c) (merely) conveys that Lois believes that 'Clark can fly' expresses a truth and that she would assent to it if asked. This would be like the mistake made by students in beginning logic classes when they reject 'all Fs are G' on being informed that some Fs are G. The scalar 'not all' implicature of 'some' obscures their view of the consistency of the two quantified statements.

A different explanation of the apparent falsity of (2c) is provided by Braun (1998). Braun notes that just as 'Superman can fly' and 'Clark can fly' express the same Russellian proposition, so (2b) and (2c) themselves express the same Russellian proposition. But someone judging the truth-values of the two sentences may take this latter proposition in one way when judging (2b) and in another when judging (2c), so it is at least intelligible that they resist the substitution inference.

So there are things the Russellian can say about conversations among the screenwriters for *Superman II*, when they agree that at the start of the movie Lois should be shown beginning to suspect that Clark is Superman, and should then confirm that he is, by tricking him when he is playing the 'Clark'-role into giving himself away. That the screenplay will thereby have Lois beginning to suspect that Clark is Clark and then deceiving him into revealing it is overlooked by the writers, who do not realize that these are the same identity-proposition, taken in different ways.

Russellian propositions are 'coarse-grained' compared to Fregean ones, for the latter are individuated in such a way that the propositions that Clark is Clark and that Clark is Superman are two. But once one accepts the distinction between proposition and way of taking the same, it's not clear what limits there are on

the coarseness of grain that may be tolerated. Perhaps, if our language contains modal operators, we cannot regard 'snow is white' and 'grass is green' as both meaning TRUE while conveying different ways of taking it, but there is no modal obstacle to an unstructured conception of propositions as classes of possible worlds (Lewis 1979; Stalnaker 1984, 1987). The same question about how much coarseness of grain is tolerable arises for hidden-indexical theorists who postulate indexically specified ways of thinking of Russellian propositions.

b. Commonsense psychology

Another question for Russellianism stems from the main purpose we have in ascribing attitudes: to make psychological generalizations ('those who believe Superman is present feel safer', Rupert 2008:83) from which we can arrive by abduction at explanations of behavior. Someone who feels safer if he believes that Superman is present and sees that Clark is present may still behave nervously or flee, which on the face of it is hard to understand if seeing that Clark is present is the same thing as seeing that Superman is present. Similarly, there are general normative principles of rationality such as

(20) A rational person who believes a conditional proposition and comes to believe its antecedent ought to infer its consequent.

This is not to say that such a person ought to *believe* its consequent: once the consequent is inferred, the thinker has various options, such as rejecting the conditional, or its antecedent, as alternatives to accepting its consequent. But a person who, at a minimum, doesn't make the inference, betrays a failure of rationality. However, Lex may believe the proposition that if Superman is nearby then he, Lex, should hide. Lex may then notice that Clark is nearby, and take no

action to conceal himself. But if believing that Clark is nearby is the same thing as believing that Superman is nearby (BEL-ing a certain proposition *via* some way of taking it) it seems that we should convict Lex of a failure of rationality, in that he remains unmoved by his two beliefs and so apparently has failed to use *modus ponens*.

In response to this, Braun (2000) argues that psychological explanation involves *ceteris paribus* (other-things-equal) principles. For example, even if Lex does come to believe he should hide, from the mere fact that he does not hide, or even try to hide, we shouldn't conclude, say, that he does *not* believe he should hide if Superman is nearby. He will only hide, or try to hide, *other things equal*. And if he already knows that he is in a location where there are no hiding-places, his motivation to seek one is thereby overridden.

So far, this is just commonsense psychology. But according to Braun, there is a special way in which things might not be equal: although a conditional and its antecedent are believed, the antecedent as it occurs as minor premise of the *modus ponens* and the antecedent as it occurs as a constituent of the major premise may not be believed in *matching ways* (2000:209). And if they are not, grounds for anticipating the expected behavior are removed. This means the principle stated in (20) is incorrect as it stands: the correct version would require a 'matching ways' restriction. And the reason there is no lapse of rationality on Lex's part when he fails to use *modus ponens* is that, although he believes that he should hide if Superman is nearby and sees that Clark, therefore Superman, is nearby, the constituent corresponding to 'Superman is nearby' of the way he takes the conditional is different from the way he takes the proposition that Superman

is nearby when he comes to believe it once he has noticed that Clark is nearby. Braun admits (2000:234) that he cannot see any other kind of way in which (20) is in need of qualification – though perhaps we need to add something about the subject remaining capable of reasoned thought long enough to make the inference – so there is a whiff of the *ad hoc* about his response; but it does allow for a version of (20) acceptable to Russellians.

c. Saul on simple sentences

Another prominent defense of Russellianism, due to Saul (1997a, 1997b, 1999, 2007), focusses on what she calls 'simple sentences'. These are sentences where we have a strong intuition of substitution-resistance, but there is no sense-invoking expression in the sentence whose semantics might explain the intuition. We have already noted one example, (21a) below. The other examples in (21) also manifest the phenomenon:

(21) a. Clark is never around when Superman is.

b. Clark went into the phone booth and Superman came out.

c. Superman is more successful with women than Clark is.

There is an obvious challenge to the Fregean in these examples. The inference in (2) fails, according to the Fregean, because of the semantics of 'believes', which requires its complement content-sentence to behave in a special way: to switch its reference, to make a double contribution to the truth-condition of the whole ascription, or whatever account of hyperintensionality is on offer. But in the examples in (21), there is no expression which might force analogous behavior on the part of the names. Yet substitution in (21a) and (21c) produces something

impossible, so, despite their apparent truth, (21a) and (21c) must be false. And substitution in (21b) seems to alter the meaning enough that the inference fails to be truth-preserving: (21b) appears to require a change of clothing, but substituting one of the names for the other produces something which does not. These examples show that intuitions of substitution-failure don't depend on the presence of psychological vocabulary. And in the absence of anything else to explain the intuitions, they show that such intuitions about the examples in (21) must be mistaken. Why, then, put any store in corresponding intuitions about (2)?

However, hidden-indexical theorists can justify substitution-failure for the examples in (21) if they are willing to widen the scope of the mechanism of hidden-indexical introduction beyond attitude verbs. For instance, it may be that what we mean by (21b) is something along the lines of 'Clark, so-attired, went into the phone-booth, and Superman, so-attired, came out'. The 'so' of 'so-attired' accounts for substitution-failure as usual, so long as the names can be associated with distinct ways of dressing: dressing in the 'Superman' way, or dressing *as* Superman, where 'as' has special features, *vs*. dressing in the 'Clark' way. For other examples, something more general than ways of dressing is needed, and this affords us an opportunity to make a partial unification of the cases of hyperintensional and simple sentences. A more general concept which is helpful is that of *personification*, and using it, for (21a) we would have

(22) Clark, so-personified, is never around when Superman, so-personified, is.

We have the same element of personification in the explanation of why fear of Superman is *not* the same thing as fear of Clark: to fear Superman, so-personified, is a very different thing from fearing Clark, so-personified (Forbes 2006:166–74).

A possible Fregean view, then, is that (22) or something like it is the literal meaning of (21a). According to (Braun and Saul 2002) however, the intuition that (21a) can be true rests on some kind of *confusion* between it and the likes of (22); the latter is, of course, substitution-resisting, but means something different from the former precisely because of that. Why would we suffer from such a confusion? Here Braun and Saul make use of the mental files metaphor, but they do not regard it as part of an account of difference in *semantic content* (see also Rupert 2008). We put information we'd naturally express with one name in the file labelled with that name, and information we'd naturally express with the other name in the file that other name labels. Then in assessing (21c), say, we compare the romantic history recounted in the entries in one file with that recounted in the other, and this task diverts our attention from the fact that the files concern the same individual. The attention-diverting element then explains why we judge (21c) to be true rather than impossible. Braun and Saul draw an analogy with the 'Moses illusion' (2002:15–16), in which a large majority of subjects, when asked 'How many animals of each kind did Moses take into the Ark?', respond 'Two', partly because the 'how many?' question diverts their attention from their knowledge that in the Bible it was Noah who took animals into his Ark.

But it is difficult to see how such an account could apply to speakers and writers who *produce* sentences like those in (21). For example, in a review of books about Shostakovich, the historian Orlando Figes wrote, 'Shostakovich always signalled his connections to the classical traditions of St. Petersburg, even if he was forced to live in Leningrad' (*The New York Review of Books*, 10th June 2004, p.14). Far from having his attention somehow diverted from the fact that St. Petersburg *is* Leningrad, Figes is consciously writing for an audience aware of the identity,

since only they will appreciate the rhetorical punch of his remark. And he will certainly resist an editor who proposes to replace 'Leningrad' with a second 'St. Petersburg', even though there is nothing hyperintensional about being forced to live somewhere.

Another example comes from an article on the transformation of Eric Blair into George Orwell (*Lingua Franca* vol.9 #9). The writer of the article is hardly *diverted* from the fact that Blair is Orwell, since his topic is exactly how one personification came to be abandoned for another in the same individual: 'Diffident in private, Blair so feared failure in the literary marketplace that he invented a pseudonym for the book he wrote based on his diaries, *Down and Out in Paris and London*. Criticism would be directed at George Orwell, not Eric Blair. But since the book, when published in 1933, was a literary success, Eric Blair became George Orwell.' Perhaps 'criticism would be directed at George Orwell, not Eric Blair' is hyperintensional, but 'Eric Blair became George Orwell' is not, it clearly resists substitution, and it would be absurd to say that the writer only makes the claim because he has allowed himself to lose sight of the fact that Blair and Orwell are the same person.

And, finally, my favorite example: a *New Yorker* cartoon in which Superman, so-personified, is talking to his therapist, and reports, "I'm doing super, but Clark can't find a paper that's hiring." It's unclear who the cartoonist thought would find this funny, but knowing that it's the same person is required to get the joke.

These examples and others (see Forbes 2006:167–8 on 'the Philosopher Stripper') show that examples like those in (21) occur outside fiction, and that those who create them do so in full awareness of the relevant identity. That (21a) means what

(22) does is certainly the most straightforward explanation of why substitution in these examples is hopeless. So the substitution-resistant nature of simple sentences does not provide as great a threat to the claim of substitution-resistance in (2) as might at first seem.

d. Richard's phone-booth

The final argument for Russellianism to be considered here is the well-known phone-booth case in (Richard 1983); I have updated it to cell phones. This example exploits the context-dependence of indexical expressions such as 'I', 'here' and 'now'. We have not discussed the phenomenon of indexicality above, but it was one on which Frege had pronounced views: he wrote about 'I' that "…everyone is presented to himself in a particular and primitive way, in which he is presented to no-one else. So when Dr. Lauben thinks he has been wounded, he will probably take as a basis this primitive way in which he is presented to himself. And only Dr. Lauben can grasp thoughts determined in this way. But now Lauben may want to communicate with others. He cannot communicate a thought which he alone can grasp. Therefore, if he now says 'I have been wounded', he must use 'I' in a sense which can be grasped by others, perhaps in the sense of 'he who is speaking to you at this moment'…" (Frege 1967:25–6). Whatever one thinks of the positive proposal here, and Frege's underlying assumptions about communication, the idea that for each thinker *x*, 'I' can be used by *x* to express a special first-person way of thinking of *x*, is one which has persisted since Frege proposed it, and is of course implicitly present in much of the history of philosophy, for example, in Descartes' *cogito*. (For further discussion of first-person and more generally indexical and demonstrative thought, see Castañeda 1968, Evans 1981, Lewis 1979, Peacocke 1983, 2008 Ch.3, and Perry 1977, 1979.)

An example in (Perry 1979) provides a dramatic illustration. Perry is pushing a grocery cart around the aisles in a store when he comes across a trail of sugar on the floor. He thinks 'that person is making a mess' and sets off in pursuit to let the person in question know that a bag of sugar in their cart has burst ('that person' is an example of 'deferred ostension', referring *via* the sugar-trail to the person whose cart the sugar bag is in; see further Borg 2002). His pursuit brings him back to the same point in the store, and he realizes, '*I* am the one who is making a mess'. This appears to be a new thought, differing from his 'that person is making a mess' in view of the difference between Perry's first-person way of thinking of himself, and his demonstrative way of thinking expressed by 'that person'.

First-person ways of thinking are private in the sense that if *x* and *y* are distinct individuals, *y* cannot employ *x*'s 'I'-way of thinking in *y*'s thoughts, certainly not as a way of thinking of *y*. However, this does not stop *y* from ascribing attitudes to *x* that require *x* to be employing *x*'s first-person way of thinking of *x* (see further Percus and Sauerland 2003). For example, *y* might say that Perry has just realized *he himself* is the one making a mess, which, on a hidden-indexical account, is to make the ascription 'Perry has just so-realized that he himself is the one making a mess'. The ability to describe a proposition as one that is a particular way of thinking of the state of affairs that Perry is making a mess does not imply that the constituents of the proposition are available to the ascriber to use in his or her own thoughts.

But *de dicto* ascriptions may not always be possible. If Perry says of some store employee, 'he knows that I made the mess', he is not ascribing knowledge of the proposition that is his 'I made the mess'-labelled way of thinking of the state

of affairs that Perry made the mess. The most Perry can mean is the *de re* 'I am known by him to have made the mess' (using the scopal account of '*de re*'), since the store employee will probably have identified the culprit demonstratively, '*that* guy is making the mess', after following the sugar trail. Perry cannot even ascribe a *de dicto* demonstrative belief to the employee using 'he believes that guy is making a mess' pointing at his reflection in a mirror. There is a strong intuition that when ascribers use a demonstrative in the content-sentences of their ascriptions, they express their own demonstrative ways of thinking of the relevant object, which means that the ascriptions are *de re* (see further Forbes 1987:13–15).

Let us now return to Richard's example. It involves switching contexts ('context-hopping') and uses Kaplan's (1989) apparatus to manage context-dependence. In Kaplan's semantics, sentences are evaluated in a context and with respect to a possible world, the *circumstances* of evaluation. A context is a sequence of entities which provides referents for the indexicals and demonstratives in a sentence *S* and so determines the Russellian proposition *S* expresses. At a minimum, we would have an agent, a time, a place, an addressee, and an object *x* to be the referent of a demonstrative or demonstrative pronoun (Kaplan uses 'agent' rather than 'speaker' to allow for a sentence such as 'I am not speaking right now' to be true in silent contexts).

When contexts are systematically related, the truth-values of sentences in the contexts are systematically related. For example, suppose that in circumstances *w*, GF is listening to MB at noon Mountain Time, November 16th 2017, and let *c* be a context with GF as its agent, noon Mountain Time, November 16th as its time, and MB as its addressee. Then the sentence 'I am now listening to you' is true in the

context *c* with respect to these circumstances *w*. But if we obtain a new context *c** from *c* by switching agent and addressee, then 'I am now listening to you' is false in *c**, since MB is speaking, not listening, to GF at noon Mountain Time, November 16th. However, 'you are now listening to me' is true in *c** with respect to *w*, since given *w*, 'I am now listening to you' in *c* identifies the same state of affairs as 'you are now listening to me' in *c**, the state of affairs that GF is listening to MB at noon Mountain Time, November 16th.

In the circumstances *w* of Richard's example, a man *a* is in his apartment, talking to a woman *o* on his cell-phone. *a* is also looking out the window onto the street below, where he sees a woman talking on her cell-phone. It never occurs to *a* that the woman he is talking to might be the woman he is watching, but in fact this is so. Then *a* notices another man in the street acting suspiciously, apparently trying to sneak up on the woman *a* is watching (*o*) from behind. In this situation, *a* could use 'she is in danger' to make a sincere assertion to *o* on his phone about what he sees. But *a* would not use 'you are in danger' to make a sincere assertion to *o* on his phone. So in the context *c* with *a* as agent, *o* as addressee, and *o*as the referent of 'she', and taking the facts about what *a* would and would not say at face-value as regards what *a* does and does not believe, the following appear to be true in this context with respect to these circumstances:

(23) a. I believe she is in danger.

b. I do not believe you are in danger.

But Richard argues (1990:117–8) that (23b) is false, in other words, that *a* does have a belief he could express by asserting into his phone 'you are in danger'. For if we

now consider a context *c** in which *o* is agent (and, if we like, *a* is addressee), the truth of (23a) in *c* guarantees the truth of

(24) The person watching me believes I am in danger

in *c**. Consequently, if we switch back to the context *c*,

(25) The person watching you believes you are in danger

is true. But there is a true identity in *c* which entails the negation of (23b), namely,

(26) I am the person watching you.

By =E, we have the anti-Fregean conclusion

(27) I believe you are in danger

now seen to be true in *c*.

By Russellian lights, the reasoning is impeccable. But should it move the Fregean? For the Fregean, attitude ascriptions can be ambiguous between *de re* and *de dicto* construals, and this applies to (27) in particular. Does the derivability of (27) really show that in *c* the protagonist *a* can truly assert a *de dicto* self-ascription,'I believe you are in danger'? Perhaps all that the derivation establishes is the truth of the *de re* reading of (27), 'you are someone I believe to be in danger'. Note that to say that (27)'s *de re* reading is true in *c* is not to say that the agent of *c knows* that it is true. But of course, *a* does have a belief he could express in *c* with either 'you are someone I believe to be in danger' or 'I believe you are in danger': *a* simply has to express his visual mode of presentation of *o* with 'you', in other words, use it as he uses 'she' in Richard's actual set-up. This is not to change context, since

the addressee, *o*, is the same in either case, but that only shows that the Fregean wants to draw distinctions which are hard to capture in Kaplan's framework.

(23a) can be understood *de re*, 'she is someone I believe to be in danger', and if the argument is construed *de re* throughout, the reasoning is correct. But of course the *de re* conclusion is not a problem for the Fregean. A *de dicto* conclusion might well be problematic, but to get one we must at least start with the reading of the premise (23a) on which it is a true *de dicto* self-ascription. Then, if the *de re* but not the *de dicto* reading of (27) is true, there must be some step in which there is a *de dicto* to *de re* switch. The switch appears to occur in moving from (23a) to (24). (24) is relevantly similar to the ascription in Perry's case, 'the store employee knows that I made the mess'. Here Perry is not ascribing knowledge of the proposition that is his 'I made the mess'-labelled way of thinking of the state of affairs that Perry made the mess. By the same token, we should not construe (24) as making an ascription to *a* of belief in the proposition that *o* expresses by 'I am in danger'. For that way of thinking of the state of affairs that *o* is in danger is simply unavailable to *a*, since it involves *o*'s first-person way of thinking of herself. The truth of (24), then, is no more than the truth of 'I am someone who the man watching me believes is in danger', whose truth in *c** is a consequence of (23a)'s truth in *c*. Thus, the *de re* conclusion follows from the *de dicto* starting point, but, to repeat, the *de re* conclusion is acceptable to all parties.

Richard considers this kind of response (1990:128–32; see also 190–96 for his own critique of his earlier argument) and rejects it. This is partly because he thinks the response imputes opacity to subject-position in ascriptions, and partly because he is generally sceptical about the *de re*/*de dicto* distinction. But the

above criticism does not seem to involve any opacity in subject position, that is, a failure of =E when applied to subject-position. Technically, there is no *single* context in which (23a)'s 'I' and (24)'s 'the man watching me' are coreferential, and anyway, the content sentence is different in these two ascriptions. Certainly, the reference of 'I' in *c* is the same as the reference of 'the man watching me' in *c**, but this doesn't threaten the use of =E if the content sentence is fixed and interpreted uniformly, in Fine's sense: 'the man who is agent of *c* believes she is in danger' and 'the man who is watching the agent of *c** believes she is in danger' have the same truth-value if 'she' is unequivocal (in the second ascription, 'she' is not to become anaphoric upon the embedded 'the agent of *c**').

As for general scepticism about *de re*/*de dicto*, the reader may refer to the discussion in section . Relevant examples arise in extensions of Richard's case, where the truth of certain statements is easily explained using the distinction, but those statements are hard to make sense of without it. Suppose that the suspiciously-behaving man turns out to be a harmless drunk who staggers on by. The phone conversation then continues in such a way that *a* soon realizes that the woman he is talking to is the woman he is watching. *a* may then say such things to *o* over the phone as 'so it was you I thought was in danger' or 'I thought you were in danger but didn't say anything because I didn't realize it was you I was watching'. These are perfectly natural remarks which seem to be true along with (23b), and appeal to the *de re*/*de dicto* distinction provides a straightforward explanation of how all these can be true together; while without the distinction, we are back in the Russellian business of endlessly having to explain away the appearance of truth.

WORKS CITED

Almog, Joseph, John Perry, and Howard Wettstein, eds., 1989. *Themes from Kaplan.* Oxford University Press.

Almog, Joseph and Paolo Leonardi (eds). 2009. *The Philosophy of David Kaplan*. Oxford University Press.

Anderson, C. Anthony. 1980. Some New Axioms for the Logic of Sense and Denotation: Alternative (0). *Noûs* 14:217–234.

Bach, Kent. 1997. Do Belief Reports Report Beliefs? *Pacific Philosophical Quarterly* 78:215–241.

Bealer, George. 1993. A Solution to Frege's Puzzle. In *Philosophical Perspectives 7: Language and Logic*, edited by James Tomberlin, 17–60.Ridgeview.

Borg, Emma. 2002. Pointing at Jack, Talking about Jill: Understanding Deferred Uses of Demonstratives and Pronouns. *Mind & Language* 17:489–512.

Braun, David. 1998. Understanding Belief Reports. *The Philosophical Review* 107:555–595.

Braun, David. 2000. Russellianism and Psychological Generalizations. *Noûs* 34:203–236.

Braun, David, and Jennifer Saul. 2002. Simple Sentences, Substitution, and Mistaken Evaluations. *Philosophical Studies* 111:1–41.

Burge, Tyler. 1978. Belief and Synonymy. *The Journal of Philosophy* 75:119–138.

Burge, Tyler. 1979. Sinning against Frege. *The Philosophical Review* 88:398–432.

Chalmers, David. 2002. On Sense and Intension. In *Sense and Direct Reference*, edited by Matthew Davidson, 605–651. McGraw Hill.

Chalmers, David. 2011. Propositions and Attitude Ascriptions: a Fregean Account. *Noûs* 45:595–639.

Church, Alonzo. 1950. On Carnap's Analysis of Statements of Assertion and Belief. *Analysis* 10:97–99. Also in Linsky (ed.) 1971, 168–170.

Church, Alonzo. 1951. A Formulation of the Logic of Sense and Denotation. In *Structure, Method and Meaning: Essays in Honor of Henry M. Sheffer*, edited by P. Henle, H. M. Kallen and S. K. Langer, 3–24. Liberal Arts Press.

Crimmins, Mark. 1992. *Talk About Belief*. The MIT Press.

Crimmins, Mark, and John Perry. 1989. The Prince and the Phone Booth: Reporting Puzzling Beliefs. *The Journal of Philosophy* 86:685–711.

Davidson, Donald. 1969. On Saying That. In Davidson and Harman (eds.), 73–91.

Davidson, Donald and Gilbert Harman (eds). 1969. *Words and Objections: Essays on the Work of W. V. Quine.* Reidel.

Davies, Martin. 1981. *Meaning, Quantification and Necessity*. Routledge.

Davies, Martin, and Lloyd Humberstone. 1980. Two Notions of Necessity. *Philosophical Studies* 38: 1–30.

Dennett, Daniel. 1982. Beyond Belief. In *Thought and Object*, edited by Andrew Woodfield, 1–95. Oxford University Press.

Donnellan, Keith. 1966. Reference and Definite Descriptions. *The Philosophical Review* 75:281–304.

Donnellan, Keith. 1974. Speaking of Nothing. *The Philosophical Review* 83:3–30.

Dummett, Michael. 1973. *Frege: Philosophy of Language*. Duckworth.

Evans, Gareth. 1973. The Causal Theory of Names. *Proceedings of the Aristotelian Society, Supplementary Volume* 47:187–208. Also in Evans 1985, 1–24.

Evans, Gareth. 1981. Understanding Demonstratives. In Evans 1985, 291–321. Oxford University Press.

Evans, Gareth. 1985. *Collected Papers*, edited by Antonia Phillips. Oxford University Press.

Fara, Delia. 2001. Descriptions as Predicates. *Philosophical Studies* 102:1-42.

Field, Hartry. 1978. Mental Representation. *Erkenntnis* 13:9–53.

Fine, Kit. 1989. The Problem of *De Re* Modality. In Joseph Almog, John Perry and Howard Wettstein (eds.) 1989, 197–272.

Fine, Kit. 2007. *Semantic Relationism*. Blackwell.

Forbes, Graeme. 1987. Indexicals and Intensionality: A Fregean Perspective. *The Philosophical Review* 96:3–31.

Forbes, Graeme. 1990. The Indispensability of *Sinn*. *The Philosophical Review* 99:535–563.

Forbes, Graeme. 1996. Substitutivity and the Coherence of Quantifying In. *The Philosophical Review* 105:337–72.

Forbes, Graeme. 2000. Objectual Attitudes. *Linguistics and Philosophy* 23:141–183.

Forbes, Graeme. 2006. *Attitude Problems*. Oxford University Press.

Forbes, Graeme. 2010. Intensional Verbs in Event Semantics. *Synthese* 176:227–242.

Fox, Chris, and Shalom Lappin. 2005. *Foundations of Intensional Semantics*. Basil Blackwell.

Frege, Gottlob. 1892. Uber Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100:25–50. Translated as 'On Sense and Reference', in *Translations from the Philosophical Writings of Gottlob Frege*, edited by Peter Geach and Max Black, 1970. Basil Blackwell.

Frege, Gottlob, and Bertrand Russell, 1904. Selection from the Frege-Russell Correspondence. In Salmon, N. and Scott Soames (eds.), 56–57.

Frege, Gottlob. 1967. The Thought: A Logical Enquiry. In *Philosophical Logic*, edited by Peter Strawson, 17–38. Oxford University Press.

Geach, Peter. 1970. The Perils of Pauline. *Review of Metaphysics* 23:287–300.

Gluer, Kathrin, and Peter Pagin. 2012. General Terms and Relational Modality. *Noûs* 46:159–199.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Basil Blackwell.

Jespersen, Bjorn. 2003. Why the Tuple Theory of Structured Propositions Isn't a Theory of Structured Propostions. *Philosophia* 31:171–183.

Kadmon, Nirit. 2001. *Formal Pragmatics*. Blackwell.

Kaplan, David. 1969. Quantifying In. In Davidson, D. and Harman G. (eds.), 206–242; also in Linsky, L. (ed.), 112–144. Page-references here are to the Linsky reprint.

Kaplan, David. 1986. Opacity. In *The Philosophy of W. V. Quine*, edited by Lewis Edward Hahn and Paul Arthur Schilpp, 229–289. LaSalle: Open Court.

Kaplan, David. 1989. Demonstratives. In Joseph Almog, John Perry and Howard Wettstein (eds.) 1989, 481–563.

Kaplan, David. 1990. Words. *Proceedings of the Aristotelian Society* 64:93–117.

Kazmi, Ali Akhtar. 1987. Quantification and Opacity. *Linguistics and Philosophy* 10:77–100.

Kripke, Saul. 1972. Naming and Necessity. In *Semantics of Natural Language*, edited by Donald Davidson and Gilbert Harman, 252–355. Reidel. Republished with an introduction as *Naming and Necessity* by Saul Kripke, Harvard University Press 1980. Page references here are to the 1980 version.

Kripke, Saul. 1979. A Puzzle About Belief. In *Meaning and Use*, edited by Avishai Margalit, 239–283. Reidel. Also in Salmon and Soames 1988: 102–148, and in Kripke, S. (ed.) 2011, 125–161. Page references in this article are to the last of these.

Kripke, Saul. 2001. Frege's Theory of Sense and Reference. In Kripke, S. (ed.) 2011, 254–291.

Kripke, Saul. 2008. Unrestricted Exportation and Some Morals for the Philosophy of Language. In Kripke, S. (ed.) 2011, 322–350.

Kripke, Saul. 2011. *Philosophical Troubles*. Oxford University Press.

Kvart, Igal. 1984. The Hesperus-Phosphorus Case. *Theoria* 50:1–35.

Lewis, David. 1979. Attitudes *de dicto* and *de se*. *The Philosophical Review* 88:513–543.

Linsky, Leonard, (ed.) 1971. *Reference and Modality*. Oxford University Press.

Larson, Richard K., and Peter Ludlow. 1993. Interpreted Logical Forms. *Synthese* 95:305–355.

Loar, Brian. 1972. Reference and Propositional Attitudes. *The Philosophical Review* 81:43–62.

Ludlow, Peter. 1996. The Adicity of 'Believes' and the Hidden Indexical Theory. *Analysis* 56:97–101.

Marcus, Ruth Barcan. 1961. Modalities and Intensional Languages. *Synthese* 13:303–322. Also in Marcus 1993, pp. 3–35.

Marcus, Ruth Barcan. 1962. Interpreting Quantification. *Inquiry* 5:252–259.

Marcus, Ruth Barcan. 1975. Does the Principle of Substitutivity Rest on a Mistake? In *The Logical Enterprise*, edited by Alan Anderson, Ruth Barcan Marcus and Richard Martin, 31–38. Yale University Press. Also in Marcus 1993, pp. 101–109.

Marcus, Ruth Barcan. 1993. *Modalities*. Oxford University Press.

Mates, Benson. 1952. Synonymity. In *Semantics and the Philosophy of Language*, edited by Leonard Linsky, 111-136. University of Illinois Press.

Moltmann, Friederike. 2003. Propositional Attitudes without Propositions. *Synthese* 135:70–118.

Muskens, R. 2005. Sense and the Computation of Reference. *Linguistics and Philosophy* 28:473–504.

Parsons, Terence. 1981. Frege's Hierarchies of Indirect Senses and the Paradox of Analysis. In *Midwest Studies in Philosophy Vol. VI: The Foundations of Analytic Philosophy*, edited by P. A. French, T. Uehling and H. Wettstein, 37–58. Minnesota University Press.

Parsons, Terence. 2009. Higher-Order Senses. In Almog, J., and Leonardi, P. (eds), 45–59.

Partee, Barbara. 2003. Privative Adjectives: Subsective Plus Coercion. In *Presuppositions and Discourse. Essays offered to Hans Kamp*, edited by Rainer Bäuerle, Uwe Reyle and Thomas Ede Zimmermann. Elsevier.

Peacocke, Christopher. 1983. *Sense and Content*. Oxford University Press.

Peacocke, Christopher. 2008. *Truly Understood*. Oxford University Press.

Peacocke, Christopher. 2009. Frege's Hierarchy: A Puzzle. In Almog, J., and Leonardi, P. (eds), 159–186.

Percus, Orin, and Uli Sauerland. 2003. On the LFs of Attitude Reports. *Proceedings of Sinn und Bedeutung* 7:228–42.

Perry, John. 1977. Frege on Demonstratives. *The Philosophical Review* 86:474–497.

Perry, John. 1979. The Problem of the Essential Indexical. *Noûs* 13:3–31.

Quine, W. V. 1956. Quantifiers and Propositional Attitudes. *The Philosophical Review* 53:177–187. Also in *The Ways of Paradox* by W. V. Quine, 1966, pp. 177–87. Harvard University Press.

Quine, W. V. 1961. Reference and Modality. In *From a Logical Point of View* by W.V. Quine, Harper and Row, 139–57. Also in Linsky, L. (ed.), 1971, 168–170. Page-references here are to the Linsky reprint.

Recanati, François. 2000. *Oratio Obliqua, Oratio Recta*. The MIT Press.

Recanati, François. 2012. *Mental Files*. Oxford University Press.

Richard, Mark. 1983. Direct Reference and Ascriptions of Belief. *The Journal of Philosophical Logic* 12:425–452.

Richard, Mark. 1986. Quotation, Grammar and Opacity. *Linguistics and Philosophy* 9:383–403.

Richard, Mark. 1990. *Propositional Attitudes*. Cambridge University Press.

Rupert, Robert. 2008. Frege's Puzzle and Frege Cases: Defending a Quasi-Syntactic Solution. *Cognitive Systems Research* 9:76–91.

Russell, Bertrand. 1905. On Denoting. *Mind* 14:479–493. Also in *Logic and Knowledge* by Bertrand Russell, edited by R. C. Marsh, 41–56. Allen & Unwin, 1956.

Saka, Paul. 2006. The Demonstrative and Identity Theories of Quotation. *The Journal of Philosophy* 103:452–471.

Salmon, Nathan. 1981. *Reference and Essence*. Princeton University Press.

Salmon, Nathan. 1986. *Frege's Puzzle*. The MIT Press.

Salmon, Nathan. 1990. A Millian Heir Rejects the Wages of *Sinn*. In *Propositional Attitudes*, edited by C. Anthony Anderson and Joseph Owens, 215–247. CSLI Publications.

Salmon, N. and Scott Soames (eds.) 1988. *Propositions and Attitudes*. Oxford University Press 1988

Saul, Jennifer. 1997a. Substitution and Simple Sentences. *Analysis* 57:102–108.

Saul, Jennifer. 1997b. Reply to Forbes. *Analysis* 57:114–118.

Saul, Jennifer. 1999. Substitution, Simple Sentences, and Sex-scandals. *Analysis* 59:106–112.

Saul, Jennifer. 2007. *Simple Sentences, Substitution, and Intuitions*. Oxford University Press.

Schiffer, Stephen. 1979. Naming and Knowing. In *Contemporary Perspectives in the Philosophy of Language*, edited by P. A. French, T. Uehling and H. Wettstein, 61–74. University of Minnesota Press.

Schiffer, Stephen. 1992. Belief Ascription. *The Journal of Philosophy* 89:499–521.

Schiffer, Stephen. 1996. The Hidden-Indexical Theory's Logical-Form Problem: a Rejoinder. *Analysis* 56:92–97.

Schweizer, Paul. 1993. Quantified Quinean S5. *The Journal of Philosophical Logic* 22:589–605.

Segal, Gabriel. 1989. A Preference for Sense and Reference. *The Journal of Philosophy* 86:73–89.

Sleigh, R. C. 1968. On a Proposed System of Epistemic Logic. *Noûs* 2:391–398.

Smullyan, Arthur. 1948. Modality and Description. *The Journal of Symbolic Logic* 13:31–37. Also in Linsky 1971, 35–43.

Soames, Scott. 1987. Direct Reference, Propositional Attitudes, and Semantic Content. *Philosophical Topics* 15:47–87.

Soames, Scott. 2002. *Beyond Rigidity*. Oxford University Press.

Sosa, Ernest. 1970. Propositional Attitudes *De Dicto* and *De Re*. *The Journal of Philosophy* 67:883–896.

Stalnaker, Robert. 1984. *Inquiry*. The MIT Press.

Stalnaker, Robert. 1987. Semantics for Belief. *Philosophical Topics*. Also in *Content and Context* by Robert Stalnaker, 117–129. Oxford University Press, 1999.

Taschek, William. 1988. Would a Fregean be Puzzled by Pierre? *Mind* 97:99–104.

Taylor, Kenneth. 2002. *De Re* and *De Dicto:* Against the Conventional Wisdom. *Philosophical Perspectives* 16:225–265.

Thomason, Richmond. 1980. A Model Theory for Propositional Attitudes. *Linguistics and Philosophy* 4:47–70.

Tichy, Pavel. 2004. The Myth of Non-Rigid Designators. In *Pavel Tichy's Collected Papers in Logic and Philosophy*, edited by Vladimir Svoboda, Bjorn Jespersen and Colin Cheyne.

Washington, Corey. 1992. The Identity Theory of Quotation. *The Journal of Philosophy* 89:582–605.

Yalcin, Seth. 2015. Quantifying In From a Fregean Perspective. *The Philosophical Review* 124:207–253.