INTERPRETATIVE MODESTY*

Suppose a speaker is confused about the meaning of a word. Should we take their uses of it at face value—as expressing beliefs they have? Many will say No, reasoning that one consequence of their confusion is that in them there are no mental items having the roles that the beliefs in question would have in their cognitive life. But what of an interpreter who does this anyway—should we take *their* words at face value, when they attribute beliefs in such cases? Some who would answer No to the first question will think that they should answer No to this second one as well, on the grounds that only a confused interpreter could fail to treat a speaker's confusion as such.

Saul Kripke raises the issue of what he calls "linguistic or conceptual confusion" in his discussion of some famous examples given by Benson Mates—examples concerning beliefs *about beliefs*. Kripke's discussion raises the above questions in a particularly clear and, I think, tractable way. In his discussion of such cases Kripke gives a firm No to our first question, and a tentative No to our second. By examining the principles to which he appeals, we can see why the second No lacks the plausibility of the first. Tempting though it is to assimilate the two cases, doing so is a mistake. The mistake is, at bottom, to think that a competent user of a word must not only *satisfy* the conditions on using it competently, she must also *believe that those are* the conditions on using it competently. Avoiding this mistake requires acknowledging what I think is a kind of *modesty* in interpretation, whereby one allows for others configurations of commitments that one does not allow for oneself.

My position on our two questions conflicts with a claim made by prominent writers in the theory of concept possession. The claim is that one can think that someone thinks something, only if it is something one can oneself think. If that is right, then there can be no *de dicto* attribution of a thought that no one can have (due, say, to linguistic or conceptual confusion). But the claim has been little supported by argument, and in light of the discussion of our two opening questions, seems to involve the same failure to acknowledge interpretative modesty.

Kripke brings up Mates's examples in his discussion of Millianism and belief ascription in "A puzzle about belief".[1] Millianism is a claim about proper names: that "a name does not describe its bearer as possessing any special identifying properties." Kripke takes it to entail that "coreferential proper names should be interchangeable *salva veritate* in belief contexts."[2] Notoriously, such interchanges lead in many cases to counterintuitive results. But when we look carefully at how to derive those results, Kripke argues, we find that there are other principles needed in addition to Millianism, and those principles *themselves* lead to similar results: Millianism's implications are "exactly paralleled by" those of these other principles.[3] The main one appears to Kripke to be "a self-evident truth":

Disquotation If a normal English speaker, on reflection, sincerely assents to 'p', then he believes that p.[4]

Kripke points out that along with a rationality assumption, Disquotation lets us infer *lack* of belief; I will apply it as if it incorporates this auxiliary assumption. So it can take us from a speaker's assenting to not-*p*, to their not believing that *p*.[5]

Strictly speaking, Kripke's purpose in his paper does not require him to discuss the interchangeability of any expressions other than proper names; and his focus in his paper is on their occurrence in very simple sentences. But he clearly, and reasonably, takes there to be some natural generalization of Millianism that has implications concerning interchangeability more generally. It is here that Mates's[6] examples become relevant, for two reasons. First, they are more complex than the examples on which Kripke focuses; second, they concern the interchangeability of expressions of many grammatical kinds, not just proper names.

What are those examples? Mates presented them schematically,[7] but we can work with the instance that Kripke discusses:

---

[1] Saul Kripke, "A puzzle about belief," in Avishai Margalit, ed., *Meaning and Use* (Dordrecht: D. Reidel, 1979), pp. 239–83.

[2] *Op. cit.*, pp. 240–41.

[3] *Op. cit.*, p. 269.

[4] *Op. cit.*, pp. 248–49.

[5] As Kripke explains, this is not an invocation of the (rather less plausible) biconditional principle, which licenses inferences from a ("normal," and so on) speaker's failure to assent to "p," that he does not believe that p (*op. cit.*, pp. 248–50).

[6] Benson Mates, "Synonymity," *University of California Publications in Philosophy*, XXV (1950):201–26. Reprinted in Leonard Linsky, ed., *Semantics and the Philosophy of Language* (Chicago: University of Illinois Press), pp. 111–36.

[7] The schemata are:

Nobody doubts that whoever believes that *D*, believes that *D*.

Mates argued that such a sentence as (*) 'Some doubt that all who believe that doctors are happy believe that physicians are happy,' may be true, even though 'doctors' and 'physicians' are synonymous, and even though it would have been false had 'physicians' been replaced in it by a second occurrence of 'doctors.'[8]

It is not my concern here to decide whether Kripke's overall project requires him to take any position on Mates's examples. What interests me is the position that he does in fact take on them. Kripke claims that "Mates's argument involves issues even more delicate" than those arising from his own showcase examples.[9] His discussion is somewhat complex and unfolds over several extraordinarily long footnotes to his paper. Central to it is the issue of the relationship between disquotational belief ascription and what Kripke calls "conceptual or linguistic confusion."[10] We should first understand why one might think that Disquotation supports Mates's claims about his examples; then we will look at why Kripke rejects those claims, despite his endorsement of that principle.

II. THE DISQUOTATIONAL SCENARIO

It is testimony to their effectiveness that there are multiple ways in which one might find Mates's examples plausible.[11] One is by thinking of something like what I will call the Disquotational Scenario:

Nobody doubts that whoever believes that *D*, believes that *D'*. (*Op. cit.*, p. 215.)

Here, "*D*" and "*D'*" are placeholders for sentences that are intensionally isomorphic in Rudolf Carnap's sense, since it was against Carnap that Mates aimed the examples (*Meaning and Necessity* (Chicago: University of Chicago Press, 1947)).

[8] *Op. cit.*, p. 281n.46. Kripke here uses "doctor" with what is arguably one of its meanings, *medical* doctor.

[9] *Op. cit.*, p. 282n.46

[10] *Op. cit.*, p. 276n.23

[11] One way is by being impressed by the difference in logical form between the complements (see above, n. 7): one is a valid form, the other is not. One could take this difference to explain on its own a difference in interchangeability (Hilary Putnam, "Synonymity, and the analysis of belief-sentences," *Analysis*, XIV (1954): 114–22). However, not all examples that involve multiple embeddings need exhibit such differences in logical form. Moreover, the difference in logical form can be just as plausibly claimed to matter for interchangeability when *singly* embedded within attitude verbs (see for example Tyler Burge, "Belief and synonymy," *Journal of Philosophy*, LXXV (1978): 119–38, at pp. 125–26; Jerry Fodor, "Review of *A Study of Concepts* by Christopher Peacocke," *London Review of Books*, XV, 19 (1993), reprinted in *In Critical Condition* (Cambridge, MA: MIT Press, 1998), pp. 27–34, at pp. 31–32). It is simply an independent aspect of Mates's examples as originally presented. (It would be too much of a digression to defend this claim here. I do so in a separate paper.) The aspect of Mates's examples that matters for Kripke, and that will matter for us, is their involving multiple embedding, not their involving complements that differ in logical form.

Another way of finding them plausible is by finding it plausible that there could be *generally inferentially hyper-cautious* belief ascribers: ones who infer *nothing* non-trivial about

> A speaker *S* assents to "doctors are happy" and "physicians are not happy." An observer *O* believes that *S* is a normal speaker, speaking sincerely and reflectively. For these reasons *O* interprets *S* disquotationally. *O* expresses these disquotational interpretations of *S* by assenting to "*S* believes that doctors are happy" and "*S* does not believe that physicians are happy."

In this scenario, it seems, *O* takes *S* to be a counterexample to the generalization that everyone who believes that doctors are happy, believes that physicians are happy. Which is to say that *O* makes the following claim true.

DoubtDP Some doubt that whoever believes that doctors are happy, believes that physicians are happy.

But nothing in the scenario is a reason to think that *O* doubts that whoever believes that doctors are happy, believes that doctors are happy; nothing, that is, is a reason to think that *O* doubts a logical truth. So the apparent truth of *DoubtDP*, in our scenario, is consistent with the falsity of this variant:

DoubtDD Some doubt that whoever believes that doctors are happy, believes that doctors are happy.

This scenario, then, seems to support the the configuration of truth values that Mates claimed for his examples: *DoubtDP* is true while *DoubtDD* is false, despite their differing only in the replacement of "doctors" by "physicians" in one occurrence. What seems to make *DoubtDP* true is the mental state that *O* is in as a result of his drawing a disquotational inference about *S*. That inference is licensed by Disquotation, for *O* believes that *S* is a normal speaker, speaking sincerely and reflectively. And our description of *O*'s beliefs is also licensed by Disquotation, if when contemplating the scenario we believe that *O* is himself a normal speaker, speaking sincerely and reflectively.

What the Disquotational Scenario seems to do, then, is to make Mates's claim plausible in a way seemingly in line with Kripke's take on our ordinary belief-ascriptive practice. Nonetheless, Kripke does not endorse Mates's claim. I turn now to his treatment of these cases.

---

the beliefs of others. Burge gives this as a reason to find Mates's claim plausible (*op. cit.*, p. 121).

III. KRIPKE'S PRINCIPLE ABOUT SYNONYMY

Kripke begins his discussion of Mates's examples by claiming that they have "relatively little force" against the Fregean claim that identity of sense suffices for interchangeability.[12] This is surprising, since that claim is clearly one of Mates's targets. The complements in our sentences *DoubtDP* and *DoubtDD* differ only in the substitution of one synonym for another. But for Mates, his point does not hinge on exactly what relation obtains between such variants: he thinks that his examples work equally well against an interchangeability claim formulated using "any adequate explication of synonymity."[13] Identity of Fregean sense is one such candidate.

Right after making this claim, Kripke sets out a principle about synonymy:

> Mates's puzzle in no way militates against some such principle as: [**Synonymy**] If one word is synonymous with another, then a sufficiently reflective speaker subject to no linguistic inadequacies or conceptual confusions who sincerely assents to a simple sentence containing the one will also (sincerely) assent to the corresponding sentence with the other in its place. (275n.15)

By "simple sentence" Kripke seems to mean one that is not complex logically and does not itself involve attitude verbs. At least, his examples all fit this description. So I will take it that that is what "simple sentence" means in Synonymy.[14] Put contrapositively, Synonymy says that a speaker who assents to a simple sentence, but not to a synonym-variant of it, is *not* "sufficiently reflective [and] subject to no linguistic inadequacies or conceptual confusions." (For simplicity's sake I will assume that the same goes for assenting to a negation of the synonym-variant; this is a form of failure to assent to the simple, unnegated sentence.) The language Kripke uses here is very similar to that which he uses in the antecedent of Disquotation. So it's reasonable to take Synonymy, together with Disquotation, *not* to license disquotational belief attribution concerning a speaker who assents to some simple sentence but not to a synonym-variant of it. And that is indeed how Kripke sees things: "if someone assents to 'Doctors are happy,' but refuses assent to 'Physicians are happy,' *prima facie* disquotation does not apply to him since

---

[12] *Op. cit.*, p. 275n.15

[13] *Op. cit.*, p. 215.

[14] Kripke does not see Synonymy as applying to his own examples, in which a speaker takes conflicting attitudes to simple sentences (or, in the Peter case, towards different occurrences of simple sentences) involving proper names (276n.23, see also 275n.15). Given that the sentences are simple, the implication is that for Kripke, synonymy cannot obtain between proper names.

he is under a linguistic or conceptual confusion" (282n.46). (Kripke's label "linguistic or conceptual confusion" can be taken to cover a wide range of cases, but he focuses on cases of the sort described, in which a speaker assents to some sentence while not assenting to a synonym-variant of it. From here I will use "conceptual confusion" as a short-hand.)

Since Kripke sets out Synonymy immediately after claiming that Mates's examples are no threat to the Fregean claim, one could be for-given for expecting him to go on to explain how that principle undercuts their plausibility. (Their force against the Fregean claim corresponds, after all, to that plausibility.) In particular, one might expect him to ex-plain how Synonymy conflicts with our treatment of the Disquotational Scenario, which culminates in just the claims that Mates would make about the case.

But the principle does not do this. Synonymy conflicts with no claims we made in setting out, or interpreting, the Disquotational Scenario. What Synonymy entails concerning S is that she is conceptually con-fused. We saw that for Kripke, this entails that S does not satisfy the an-tecedent of Disquotation: she is not "a normal English speaker, [speak-ing] on reflection." But nowhere in our description of the scenario is it stated that S is such a speaker. What is stated is that O *believes* that she is. Synonymy entails that that belief of his is false. But nowhere in our description of the scenario, or our interpretation of it, is it stated or implied that it is true. We might even maintain that O *errs* in inter-preting S disquotationally: that too is consistent with the scenario and our interpretation of it. What matters for the Matesian conclusion is the *possibility* of O's interpreting S so as to arrive at the doubt that makes *DoubtDP* true, not the *correctness* of his doing so. Nor does Synonymy have implications concerning our interpretation of O. The reason is that the sentences to which O assents are not within its scope. They are complex belief-ascriptive sentences, whereas Synonymy concerns only simple sentences.

So if Kripke is to reject Mates's claim, it must be on the basis of a principle that, unlike Synonymy, has implications about what beliefs O *can* have about S, rather than about which ones he *should* have about her. Kripke recognizes that these are distinct issues:

> [O, in our scenario] may realize that "doctors" and "physicians" are syn-onymous; but he applies disquotation to [S] who assents to "Doctors are happy" but not to "Physicians are happy," ignoring the caution of the pre-vious paragraph [that is, ignoring what Synonymy entails about S].[15]

---

[15] *Op. cit.*, p. 282n.46

So it is no surprise that Kripke continues his treatment of Mates's examples by introducing another principle.

The next principle is one that does have implications about *O*. Kripke says of *O* (as named in our scenario) that he "[appears] to be under a deep conceptual confusion (misapplication of the disquotational principle). Perhaps, it may be argued, he misunderstands the 'logic of belief'."[16] Kripke does not elaborate further on this thought, and for now it is enough to work with the claim that such a speaker is conceptually confused. (Kripke's remarks do suggest a stronger, more specific claim: that such a speaker's conceptual confusion relates to the concept BELIEF. I consider that claim below, at the end of §v. The reason for working with the weaker claim is that it is enough, when developed, to undercut Mates's claim.) Let us state this new claim as follows.

Transmission If only the conceptually confused assent to a simple sentence *p* while dissenting from a synonym-variant of it, then only the conceptually confused assent to "*S* believes that *p*" while dissenting from its corresponding synonym-variant.

In our case, the synonym-variant simple sentences are "Doctors are happy" and "Physicians are happy." *S* assents to the first and dissents from (that is, assents to the negation of) the second; and *O* disquotes her on this basis. We saw earlier that Synonymy entails that *S* is conceptually confused. Transmission kicks in now, concerning *O*. It entails that *he* is conceptually confused, in virtue of assenting to his belief-ascriptive sentences about *S*. *O* thus fails to satisfy the antecedent of Disquotation; so, that principle does *not* license the disquotation-based attribution to *O* of the beliefs that *S* believes that doctors are happy, and that *S* does not believe that physicians are happy. This undercuts our interpretation of the scenario, which involved the ascription of those very beliefs to *O* on the basis of the *S*-interpreting sentences he uttered. (Strictly speaking it does not *conflict* with our interpretation, since the conclusion of an unlicensed inference might simply happen to be true. But Disquotation would seem to be the only principle to which one might appeal, in defense of our interpretation of *O* in the Scenario. In this sense, the interpretation is undercut by *O*'s not satisfying its antecedent.) In short: Disquotation, Synonymy and Transmission combine to entail that that scenario does *not* support the Matesian claim.

---

[16] *Op. cit.*, p. 282n.46

So one might think that with Transmission, we have a claim whose own plausibility undercuts that of Mates's examples. Later I will assess how much plausibility that is. (Not much, I'll argue.) But for now—pursuing Kripke's treatment of the example—I will explain how even this new claim of Kripke's is of no use against Mates.

Transmission does undercut our interpretation of *O* in the Disquotational Scenario. But once again, Mates's examples prove resilient. For by involving multiple embeddings, they point to the possibility of occurrences of words that are *more than doubly* embedded in the scopes of attitude verbs. And the claim that these new, higher-level Mates instances exhibit just the behavior that supports Mates's claim is not undercut at all by Transmission, which Kripke introduces in order to undercut instances of Mates's original schema.

The point is by now familiar. To say—appealing to Disquotation and Synonymy and Transmission—that *O*'s disquotational attributions to *S* *should* not be disquoted by someone contemplating the scenario, is not to say that they *cannot* be. And again, it is the possibility that matters. Even if Kripke convinces his readers that *O should not* be credited with the relevant beliefs about *S*'s beliefs, it is enough for Mates's purpose if those readers think that it's *possible* for someone else, not so convinced, to do so. (Indeed they could themselves do so, while recognizing that it is unjustified—exhibiting a kind of weakness of inferential will. But I'll stick with bringing out the problem by imagining the reader's imagining some other reader.) Such an imagined contemplator of the Disquotational Scenario, *C*, *believes that O doubts that everyone who believes that doctors are happy, believes that physicians are happy*. But they do not *believe that O doubts that everyone who believes that doctors are happy, believes that doctors are happy*, for *C* is not imagined to have a reason to believe that *O* doubts a logical truth. *C* is then a counterexample to a *higher-order* Mates instance, and Mates's claim is supported once again.

Of course, there can be readers of Kripke who do not acknowledge the possibility of other readers' arriving, via disquotation of *O*, at beliefs about *O*'s beliefs about *S*'s beliefs. What commitments would preclude such acknowledgement? The obvious candidate, given the story so far, is commitment to an *every-order* strengthening of Transmission:

Transmission+ If only the conceptually confused assent to *p* (any sentence, simple or not) while dissenting from a synonym-variant of *p*, then only the conceptually confused assent to "*S* believes that *p*" while dissenting from its corresponding synonym-variant.

The only difference is the removal of Transmission's requirement that *p* be a simple sentence. We are applying the idea of Transmission now to

all sentences exhibiting "conceptual confusion," not just simple ones.[17] This is hugely consequential.

Transmission+ entails that our imagined contemplator of the Disquotational Scenario, *C*, is conceptually confused, and should not have beliefs attributed to him on the basis of his disquotational belief ascriptions about *O*. Moreover it entails that anyone who does this nonetheless is conceptually confused as well, and should not have beliefs ascribed to *them* disquotationally. All this traces back to the conceptual confusion that *S* manifests with the simple sentences "Doctors are happy" and "Physicians are not happy." The Transmission+-based reasoning can be applied through any number of embeddings of the original ascriptions; there is *no* level at which any kind of Mates example is allowed.

Our finding that Transmission+ is the principle to which Kripke must appeal, in order to reject Mates's examples, allows us now to assess another claim by Kripke:

> Mates's problem would not arise in a world where no one ever was under a linguistic or a conceptual confusion, no one ever thought anyone else was under such a confusion, no one ever thought anyone ever thought anyone was under such a confusion, and so on.[18]

If by "linguistic and conceptual confusion" here Kripke means what *S* exemplifies in the Disquotational Scenario, then these conditions do *not* suffice to rule out plausible instances of Mates's schema. The first stipulation does entail that there cannot be an *S* such as *O* believes there to be: a normal, sincere and reflective speaker who assents to "Doctors are happy" and to "Physicians are not happy." But our reasoning towards the Matesian conclusion would have gone through just as well on the basis of a variant scenario in which *O mishears S*, for it is *O*'s *beliefs* about *S*, not the facts about *S*, that matter to the truth and falsity of *DoubtDP* and *DoubtDD*. The further stipulation, that "no one ever thought anyone else was under such a confusion," does not conflict at all with the Scenario as we presented it, because the scenario does not involve the claim that *O thinks that S is conceptually confused*. So there can be Mates-supporting scenarios that violate none of the conditions Kripke here states. In this passage Kripke states something *close* to what

---

[17] I'm grateful to a referee for pointing out that if Kripke's reason for endorsing Transmission is the thought that a speaker of the sort it describes must have a problem with the concept BELIEF, then he might take himself to have no less reason to endorse Transmission+, the idea being that that problem would be there, or not there, regardless of the complexity of the complements in the belief-ascriptive sentences the speaker asserts. Below (§v), however, I argue that the "must" in that thought is incorrect.

[18] *Op. cit.*, p. 282n.46

Transmission+ says, but it is not quite right, and does not do the work that Kripke claims for it.

With Transmission+, then, we have finally arrived at Kripke's true bulwark against Mates's examples. This principle undercuts their plausibility in proportion to its own plausibility. It is time now to ask what that degree of plausibility is—indeed, what the plausibility of the much weaker principle, Transmission, is. Our negative assessment of it will apply even moreso to the stronger variant.

V. THE PLAUSIBILITY OF TRANSMISSION

It must be admitted that Transmission is initially plausible. If failure to treat simple sentences p and q the same (as far as assent goes) amounts to "conceptual confusion," then, one might wonder, what else but conceptual confusion could explain failure to treat "S believes that p" and "S believes that q" the same? The natural thought is that the cases stand or fall together. But as I will now explain, that natural thought neglects an important point about speakers' perspectives on their own uses of words.

Return to our Disquotational Scenario. Recall that O disquotes S's assents to p and to not-q (where q is a synonym-variant of p, a simple sentence), arriving thereby at his own assents to "S believes that p" and "S does not believe that q." We will now add one supposition to our previous description of the Scenario, so that we can use it to examine the natural thought just described. The supposition is that O is competent by the lights of Synonymy with p and q. That is, O never assents to one while failing to assent to the other.

What does this entail now about O? It is true that Disquotation and Synonymy, conjoined with the claim that p and q are synonym-variant simple sentences, do not license O in disquoting S. So we could charge O with irrationality—disquoting with no support from the relevant principles—*if O endorses all those claims.* Let us assume, so as not to invite the charge that O is "conceptually confused" about belief (a point to which I'll return), that O endorses Disquotation and Synonymy. That leaves the third claim. Does O, in our present case, endorse the claim that p and q are synonym-variants?

Not necessarily. Our just-added supposition about O is that he is disposed never to assent to one of p and q but not to the other; this is all that Synonymy requires of him, if he is not to suffer from "linguistic or conceptual confusion." But this does not entail that O *believes that p and q are synonym-variants*. And if he does not so believe, then he cannot be charged with irrationality in endorsing Disquotation and Synonymy and assenting to "S believes that p" but not to "S believes that q." This gap—between *using* words in a way that reflects their standing in a

certain semantic relation, and *believing that* they stand in that relation—is the gap which prevents Transmission from inheriting the plausibility of Synonymy.

The reason the aforementioned entailment fails was nicely explained in a paper by Steven Rieber.[19] Rieber pointed out that someone could be disposed to use words interchangeably in simple sentences, while also having a belief that stands in the way of judging them to be synonymous, or even coextensive. (So the issue is not that the speaker fails to ask herself whether the words are synonymous, or that she lacks the very concept of synonymy.) The belief is that *there may be* counterexamples to the claim that the words are synonyms, or even coextensive. Having such a belief does not require having any particular case in mind. (This is why having it does not conflict with the dispositions regarding simple sentences.) One way to acquire this belief is by experience, concerning other pairs of words, of clever philosophers presenting cases bringing out a previously unremarked-on difference in meaning or extension. The example Rieber uses is of a speaker, Joan, who uses the English words "bet" and "wager" interchangeably:

> However, Joan is sceptical about apparent synonymies. In the past, she often thought that certain pairs of words were synonymous, then discovered (or was shown) slight differences in meaning. Perhaps she is an analytic philosopher humbled by many counterexamples to claims about co-extensiveness. In any case, she now doubts that "bet" and "wager" are synonymous.[20]

So there can be a speaker who, *without irrationality*, endorses Synonymy and Disquotation and is competent by the lights of Synonymy with synonym-variant sentences p and q, yet who assents to "S believes that p" while dissenting from "S believes that q." This is possible because a speaker's using words in accordance with a given semantic relation between them does not entail their believing that they stand in that relation. This is an important fact about a speaker's perspective on their uses of words: in general a speaker can (without irrationality) fail to fully commit themselves, in general terms, to the propriety of their own word-use dispositions.

[19] Steven Rieber, "Understanding synonyms without knowing that they are synonymous," *Analysis*, LII (1992): 224–28.

[20] *Op. cit.*, p. 226. One is reminded here of Burge's point (cited above, n. 11), that someone might instantiate Mates's claim not on the basis of a belief that some particular pair of words differ in extension or meaning, but on the basis of a general caution about making inferences among variant belief ascriptions. Rieber's claim that Joan *doubts* that the words are synonymous is stronger than I need for my purpose here; for my point, it is enough that her case illustrates the possibility of *failing to believe* that they are synonymous.

I have argued directly against a claim that Kripke tentatively en-
dorses. He writes:

> Now suppose someone [*O*, in our presentation] assents to "Not all who
> believe that doctors are happy believe that physicians are happy." What is
> the source of his assent? If it is failure to realize that "doctors" and "physi-
> cians" are synonymous (this was the situation Mates originally envisaged),
> then he is under a linguistic or conceptual confusion, so disquotation does
> not clearly apply. Hence we have no reason to conclude from this case that
> [the Matesian result obtains]. Alternatively, he may realize that "doctors"
> and "physicians" are synonymous; but he applies disquotation to a man
> who assents to "Doctors are happy" but not to "Physicians are happy," ig-
> noring the caution of the previous paragraph [that is, ignoring Synonymy].
> Here he is not under a simple linguistic confusion (such as failure to re-
> alize that "doctors" and "physicians" are synonymous), but he appears
> to be under a deep conceptual confusion (misapplication of the disquota-
> tional principle). Perhaps, it may be argued, he misunderstands the "logic
> of belief." Does his conceptual confusion mean that we cannot straightfor-
> wardly apply disquotation to his utterance, and that therefore we cannot
> conclude from his behavior that [the Matesian result obtains]? I think that,
> although the issues are delicate, and I am not at present completely sure
> what answers to give, there is a case for an affirmative answer.[21]

My stipulation that *O* endorses both Synonymy and Disquotation is
meant to head off one of the options Kripke here identifies, which is that
O has a problem with the concept BELIEF. (Recall that we had earlier
(§IV) mentioned that some of Kripke's remarks suggest that he would
say that that is *O*'s problem in the Disquotational Scenario. The point I
am developing is that there is another, easily overlooked way in which
the Scenario might obtain.)

The other option Kripke considers is *O*'s "failure to realize that 'doc-
tors' and 'physicians' are synonymous." But there is a crucial ambiguity
here. What matters for the rationality of *O*'s disquotation of *S* (against
the background of *O*'s other commitments) is whether *O* *believes* that
those words are synonymous, not whether he is disposed to handle them
properly in simple sentences. But as far as Synonymy is concerned it is
the disposition, *not* the belief, that matters to whether he is "under a
linguistic or conceptual confusion," so presumably it is the disposition
that matters in this passage, given the prominence Kripke gives to Syn-
onymy. My argument has been, contra Kripke, that that is *not* enough
to secure the result that Kripke wants, which is that *O* has some short-
coming in rationality or linguistic competence, the upshot of which is
that we should not disquote his disquotational ascriptions of beliefs to
S. In our story *O* has no such shortcoming.

---

[21] *Op. cit.*, p. 282n.46

VI. PERSISTENCE WORRIES

One might worry that my argument against Transmission would prove too much: that it implausibly entails a complete decoupling of competence with a word in the complements of attitude-ascriptive ascriptive sentences, from competence with it in simple sentences. To formulate this misgiving, it will help to have a couple of bits of terminology; and to answer it, it will help to explain why our focus in this discussion should be not only on synonymy but also on entailment relations. I'll take the latter point first, as it will help us to formulate the terminology.

Synonymy is a specific version of a more general idea: that there are certain relations among words, such that being a competent user of a word requires being disposed to use it in conformity with those relations. It is a claim about a necessary condition on competence; the claim is not that it is a sufficient condition. The claim is initially plausible. Linguists, for example, do often say that a semantic theory should explain relations such as synonymy and entailment.[22] On its own, this does not suffice to motivate Synonymy. What is needed additionally is the thought that *competence requires* using words in ways that respect such relations. It is not my concern now to evaluate this thought; it is merely to note that it is no more tied to synonymy than it is to entailment. So although our discussion has been about Synonymy because that is the principle that Kripke introduced into his discussion, there is no reason to accept that principle while *not* accepting a corresponding principle about the entailment relations that linguists seem to think of as mattering to semantics in the same way that synonymy does. So, I shall do so in this discussion.

Now for terminology. First, I'll say that a speaker is **competent with** a word as it occurs in a given set of sentences, if their dispositions to assent and dissent to them reflect nothing in the way of "linguistic or conceptual confusion." This is not a notion of across-the-board competence; it is applied relative to this or that set of sentences in which the word occurs.

Second, we will need a concept about a *relation* between competence with a word as it occurs in one set of sentences, and competence with

---

[22] In their semantics textbook Richard Larson and Gabriel Segal give "logicosemantic relations such as contradiction, implication, and synonymy" as part of the "pretheoretical domain of semantics" (*Knowledge of Meaning* (Cambridge, MA: MIT Press, 1995)); in a recent paper Jeff King, summarizing Chomsky's program in *New Horizons* (*New Horizons in the Study of Language and Mind* (Cambridge: Cambridge University Press, 2000)), mentions "certain sorts of entailments" as among the core "semantic features" of words ("W(h)ither semantics!(?)," *Noûs*, LII (2018):772–95, at p. 777). The view is not universal. In their influential semantics textbook, Irene Heim and Angelika Kratzer mention only truth conditions as what semantics is to explain (*Semantics in Generative Grammar* (Oxford: Blackwell, 1998), at p. 1).

it as it occurs in some other set of sentences. Suppose that competence with a word in one set of sentences requires treating it as interchangeable with (in the case of synonymy), or as one-way replaceable by (in the case of an entailment), some other word in such sentences. I'll say that such a condition **persists** from that set of sentences to another set of sentences, if competence with the word in the latter set requires treating them correspondingly. In stating this definition I'm simply helping myself to the notion of one-way replaceability as the counterpart, for entailment, of interchangeability. Formulating this notion precisely would be complicated, since we would need to take account of negative polarity environments' effects on entailments. But in example cases of the sort that will concern us, the idea is clear enough. Consider the "bachelor"-"male" entailment. Suppose that competence with "bachelor" in sentences such as "John is a bachelor" requires being disposed never to assent to one of them while dissenting from its "male" variant. To say that this relation persists to belief-ascriptive sentences is to say that competence with it in sentences such as "Susan believes that John is a bachelor" requires being disposed never to assent to one of them while dissenting from *its* "male" variant.

In this terminology my argument's conclusion can be simply stated: synonymy does not always persist, from occurrences in simple sentences to occurrences in Mates-style sentences.

Now one might ask: are there any relations that do persist? It would be odd if none does. For that would mean that (as far as respecting synonymy and entailments is concerned) there is *nothing* about competent use of a word in attitude ascriptions that carries over from competent use of it in simple sentences.

One might think that the Matesian line is that no relations persist, because it seems that there are no pairs of synonyms that are interchangeable when occurring within Mates-style sentences. But that thought neglects the peculiarity of synonymy among the semantic relations at issue. Identity is, by definition, the finest-grained of all the relations there are. Synonymy, as identity of meaning, is the finest-grained of all the meaning relations there are. Entailment relations are in general far more coarse-grained: there can be entailments that are exhibited by many words. "Brother" and "sister" both entail "sibling"; "triangular," "square," and "round" all entail "shaped"; every Arabic number expression is related to "number"; and so on. Now the fact that there can be reasonable doubts about the obtaining of the finest-grained relation there is, shows nothing about the possibility of reasonable doubts about the obtaining of much coarser-grained relations. We find it plausible that someone—Joan, in Rieber's example—reasonably believes,

of two synonyms, that there could be people who see a very fine difference of meaning between them that she fails to discern. This does not commit us to finding it plausible that in *every* case in which someone's uses of a word respect an entailment relation, they can *reasonably* believe that there might be people of finer discernment who see that it fails to obtain—who see that some round things have no shape, say, or that 16 is not a number. And indeed that is *not* plausible. So Mates's examples do nothing to undercut the idea that there are many semantic relations that persist from simple sentences to attitude-ascriptive ones, even though they do undercut the idea that *synonymy* persists. All this means is that facility with synonyms matters less for overall linguistic competence than facility with entailments does. I think that this is something we should have thought anyway though, given the relative scarcity of synonyms in natural languages.

VII. A PRINCIPLE ABOUT THINKABILITY

There is another misgiving one might have about my argument. It arises from another question that we can ask about relations between competence with a word as it occurs in simple sentences, and competence with it as it occurs in the complements of attitude-ascribing sentences. The question is this. Does the latter competence require the former? Or does it require only that one's uses of the word in simple sentences respect those relations that persist to attitude-ascriptive ones? One might think that my argument supports the idea that semantic relations that *do not* persist are explanatorily idle concerning competence with a word in the latter sentences. Recall Joan. Her disposition to treat "bet" and "wager" as interchangeable in her statements about the world does not generate a disposition to treat them that way in her statements about others' beliefs. Nor does it generate a rational obligation to do so. So what could be the point of nonetheless insisting that her having that first disposition, concerning occurrences of those words in simple sentences, is a necessary condition on competence with them in attitude-ascriptive ones? We would be insisting on a condition that does no explanatory work. So the argument puts in doubt the claim that competence with a word as it occurs in the complements of attitude-ascribing sentences requires competence with it as it occurs in simple sentences.

This claim is closely related to one that has been made in the theory of concept possession. The claim is that for every sentence *p*: necessarily, anyone who is able to think the thought *that S thinks that p* is able to think the thought *that p*. Call that the Thinkability Claim. The relations between that claim and the one described in the previous paragraph, about competence with words, are not straightforward, primarily because there are differences on the question of whether it is possible to

think that *p* without having learned words with which to express that thought. For some theorists—most famously, Jerry Fodor[23]—there are many thoughts that any non-infant human is able to think, regardless of whether they have learned words with which to express them. But to a philosopher who maintains that it is at least very often the case that the ability to think a thought is acquired *by* acquiring the ability to express it in words, putting in doubt the above-described claim about word-competence means putting in doubt very many instances of the Thinkability Claim.

Yet there are prominent writers on the subject of concept possession who have presented the Thinkability Claim as an *obviously true* principle about the nature of thought. Thus, Tyler Burge:

> It is clear that in attributing a way of thinking to someone in the that-clause fashion, one must understand that way of thinking. That is, in making reference to a way of thinking in an unembedded that-clause-type attribution, one must have a capacity to think with the way of thinking (the customary sense) that one attributes.[24]

And Christopher Peacocke:

> It is intuitively plausible that the ability to think about any Thought as the Thought that *p* requires the ability to employ that same Thought *p* in first-order thinking about the world.[25]

Why might these writers think of such claims as "clear" or "intuitively plausible"?[26] It is certainly not because they follow from a more

---

[23] *The Language of Thought* (New York: Thomas Y. Crowell Company, 1975).

[24] Tyler Burge, "Postscript to 'Frege and the hierarchy'," in *Truth, Thought, Reason: Essays on Frege* (Oxford: Oxford University Press, 2005), pp. 167–210, at p. 175.

[25] Christopher Peacocke, *Truly Understood* (Oxford: Oxford University Press, 2008), at p. 287. I'm grateful to a referee for pointing out that Peacocke might not want to endorse this fully general claim, given his claim about the ascription of "I"-thoughts: "You can still know *what* [someone having an 'I'-thought] judges, even if you cannot judge it yourself" (*A Study of Concepts* (Cambridge, MA: MIT Press), at p. 221). If "knowing *what*" entails being able to frame a *de dicto* ascription, then this claim about "I" thoughts entails a class of exceptions to the Thinkability Claim.

[26] I do not mean to suggest that Burge or Peacocke offer nothing in support of the quoted claims. Burge mentions facts about early childhood development of conceptual abilities (175). But all that they show is that for *some* concepts—plausibly, those of "core cognition"—the ability to ascribe follows the ability to think. (On "core cognition" see Susan Carey, *The Origin of Concepts* (Oxford: Oxford University Press, 2009), among many other entries in the psychological literature.) They could not show that this is the case for *all* concepts. Indeed one might argue that their satisfying the former description is what makes them worth specially designating as "core." For his part Peacocke does offer an account of concept possession based on the idea of a thinker's having "tacit knowledge" of a concept's extension-determining principle. (This is a development of the theory first presented in his 1992 book (*op. cit.*).) The basic idea is that the thinker employs the

general principle that is itself obviously true—for example, that being able to describe someone as doing something entails being able to do it oneself. On the contrary, that principle is obviously false. Perhaps it is plausible if one is used to thinking of "understanding a word" as an all-or-nothing matter: that competence with it in some class of its occurrences necessarily brings along with it competence in every other class of occurrences. (This claim is much stronger, and much less plausible, than the Generality Constraint on concept possession articulated by Gareth Evans.[27]) But that is an idea that Mates's examples challenge, in the way that I have tried to bring out. So while the principle that Burge and Peacocke present as obvious does conflict with the position that is naturally motivated by the analysis of Mates's examples that I've presented, it is far from obvious that the former is so incontrovertible that we should reject that analysis by appeal to it.

VIII. CONCLUSION

My purpose here has not been to argue for the very possibility of conceptual confusion, or to argue for the very idea that it undercuts the attribution of belief. Rather it has been to argue against what I think is a too-simple way of working with these ideas.

Many have followed Donald Davidson in maintaining that interpretative practice involves a sort of charity.[28] What we have found, I think, is a sort of *modesty* that it involves. One can be free of conceptual confusion without being very confident that one is. This diffidence in oneself has as a consequence a modesty in the interpretation of others: we allow for others uses of words that we would not allow for ourselves. I believe that the impressive intuitive plausibility of Mates's examples—Hilary Putnam called them "extremely powerful"[29]—shows that each of us implicitly acknowledges this.

This interpretative modesty has been under-theorized, however: philosophers have wanted to work with conceptions of word-competence, or concept-possession, on which being a competent *practitioner* with a

---

concept, in observations or in inferences, only in ways that that principle licenses. (This does not require *processing a representation of* the principle.) Peacocke applies this idea to the question of possessing a concept *of a concept*, as when one ascribes to someone the belief that *p*, thereby (on his view) employing concepts of the concepts figuring in *p*. He does claim that this "Leverage Account" entails the claim that I quoted, but the account is complex and it is not clear to me exactly how it does that.

[27] *The Varieties of Reference*, edited by John McDowell (Oxford: Oxford University Press, 1982), at p. 104.

[28] Donald Davidson, "Radical interpretation," *Dialectica* XXVII (1973): 313–28. Reprinted in *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 1984), pp. 125–39.

[29] *Op. cit.*, at p. 117.

word amounts to being a competent *judge* of its uses by others. Our social life with words does not have this presupposition built into it. There is play, sufficient to make trouble for principles such as Transmission, between first- and second-order competence. This does not mean that there is no relation; it just means that its principles are yet to be fully understood.

MARK MCCULLAGH

Department of Philosophy, University of Guelph