Saul Kripke on Belief Ascription

Mircea Dumitru
University of Bucharest*

ABSTRACT | The paper critically examines the issue of whether failure of attributing propositional attitudes, such as belief ascriptions, is intrinsically failure to substitute codesignative terms (names) within opaque contexts. One open question is whether Fregean "senses" still have any role to play within semantics. This question is more acute in the aftermath of Kripke's (1980), because one of the most remarkable achievements of the arguments Kripke deploys is the demolition of the Frege-Russell tradition in regard to the semantics of names. The descriptive theory of names, the view that names are disguised definite descriptions, and, in general, the Fregean mechanism of referring through senses that are expressed by definite descriptions associated with names have become obsolete. And there has been a growing consensus among logicians and philosophers of language that as long as we stick to direct discourse, to transparent contexts, and to the modal case, the Millian view on proper names that Kripke has advocated so forcefully cannot be plausibly denied. Still, having in view the peculiar consequences that Kripke's Millian view on names may have for opaque contexts, and particularly for the interpretation of the substitutivity of codesignative singular terms (especially names) within propositional attitude ascriptions, one may wonder, and consequently, one can legitimately ask, whether Fregean "senses" don't have still something to offer. Thus, examining and assessing Kripke's (1994) one may legitimately wonder whether there is still a case for Fregean senses within semantics of belief ascriptions. The paper examines aspects of the dialectics which is going on in this debate and points at some open issues which are worth pursuing in the spirit of Kripke's (1994).

KEYWORDS | Saul Kripke; John Stuart Mill; Gottlob Frege; Bertrand Russell; Belief Ascriptions; Failure of Substitutivity; Semanitcs of Names; Rule of =E

* *Correspondence:* Mircea Dumitru – Faculty of Philosophy, University of Bucharest, Splaiul Independenței no. 204, 060024 Bucharest, Romania. Email: mircea.dumitru@unibuc.ro

Mircea Dumitru

1 Introductory Remarks

As we know, a piece of reasoning such as the following:

> (BC) Lois believes that Superman can fly;
>
> Superman = Clark Kent;
>
> therefore, Lois believes that Clark Kent can fly,

raises the question of why the outcome of applying a sound rule of first-order logic with identity, viz. "= E," within a belief ascription context is an unsound argument. For assuming the story as a fact, and depending upon contingencies regarding Lois' beliefs, she still believes (or at least can believe) that Clark Kent cannot fly, although she believes that Superman can fly, and actually Superman *is* Clark Kent.

Whatever the semantic account of this puzzle, what seems reasonable to resist to is: (i) to say that provided the truth of he premises of (BC) the belief ascriber is entitled to draw the above displayed conclusion, and (ii) to impute to Lois (or in general to whoever is the subject of that belief attribution) contradictory beliefs, if, as it happens, she believes that Clark Kent cannot fly, although Superman *is* Clark Kent, and she believes that Superman can fly.

However, since what is involved in particular in drawing the conclusion of (BC) above seems to be what one can call a principle of substitutivity of coreferential or codesignative names in belief contexts, and since it seems to imply that as a result of its application Lois does have contradictory beliefs, one can construe (BC) as an instance of a *reductio ad absurdum* of that substitutivity principle within belief contexts. Simply put what (BC) shows is that the principle of substitutivity of coreferential names in belief attribution contexts fails.

Then, it seems legitimate to ask that in response to this semantic phenomenon we come up with an explanation of the mechanism that blocks the application of a rule that, save for propositional attitude contexts, quotation contexts and the "so-called" context, is all right.

However, if we look at (BC) as a case in which we cannot ascribe a belief to someone solely on logical grounds precisely because one is not allowed to apply "= E" within belief ascription contexts then one is inclined to think that that substitutivity principle is the heart of the matter when troubles with belief ascriptions occur. And to say this boils down to countenancing the view that in any troublesome belief ascription context a substitutivity principle is essentially involved.

The drawback of such an explanatory route is that it seems to identify *a priori* the class of phenomena of failure of ascribing beliefs on logical grounds with the class of those arguments in which failure of substitutivity of coreferential names

seems to be what is at stake. But as it is well known, Kripke has shown that there are cases in which one has serious problems with attributing to someone beliefs when *no* substitutivity at all is involved. Or at least not in an apparent way.

Thus, it seems to be more cautious not to prejudge the solution to the problem and, accordingly, not to address the problem under the tag "the failure of substitutivity of coreferential names within belief contexts," although "Why we cannot apply such a rule in propositional attitude contexts?"[1] still remains a legitimate question, but to pursue it under the more neutral label of failure of ascribing beliefs solely on logical grounds. This way, we leave open that the latter failure be the outcome of the collapse within beliefs ascription contexts of certain semantical principles, including, but not being reducible to, substitutivity of identicals, whose application outside such contexts is unproblematic.

Thus, keeping in mind the central position that failure of substitutivity of coreferential names within belief ascription contexts has in our topic of belief ascription, one can start addressing the problem of failure of belief ascription solely on logical grounds insisting on the interpretations provided for that failure of applying "= E" in belief ascription contexts. The problem is open, in principle, to several interpretative options of which two are more significant for the semantics of propositional attitude ascriptions.

Either the rule "= E" (also known as Substitutivity of Identicals) fails to apply to such a context, and accordingly any such illegitimate application has as outcome an unsound argument, or else, appearances notwithstanding, the rule is not essentially involved in those cases in which ascriptions of a belief is a *non sequitur* from previously established premises, one of which properly ascribing a belief and the other stating identity between a name which occurs in that ascription and another codesignating name, and thus some other semantic principles might be responsible for failure of ascribing beliefs.

In the former case we have to provide a logical and semantic analysis of that

---

[1] Depending on the construal of Kripke's cases and proposals, one can adopt either a revisionary view according to which, our semantic intuitions notwithstanding, "= E" does apply within belief ascription contexts, but there are other semantic principles whose breaking down within those contexts obscures that fact, or else one can adopt the view that "= E" is illegitimately applied within belief contexts, but nevertheless the core of the phenomenon of failure of belief ascription on logical grounds is not to be identified with failure of substitutivity of coreferential names, on pain of ignoring the existence of puzzling cases of belief ascription contexts in which no substitutivity principle is basically involved. Whatever the interpretation of Kripke's puzzling cases of belief ascription, I believe that the former option, which seems to be embraced by Kripke, still requires an explanation of at least why do we have such a strong intuition that in belief attribution contexts "= E" cannot be applied. It seems to me that we can consistently held that "= E" fails to apply within belief ascription contexts, but nevertheless that the failure of ascribing beliefs on logical grounds is not to be identified with failure of substitutivity of coreferential names within those contexts.

Mircea Dumitru

mechanism which is supposed to be responsible for the phenomenon of failure of substitutivity. It is worth noticing that this option doesn't necessarily boil down to the view that "= E" is unsound. The point is rather that a propositional attitude context is such that an application of the rule "= E" within it is upsetting some other assumptions (premises) than those that license a validity-preserving application of the rule. Hence, it is a *mis*application of the rule (perhaps something similar to a fallacious argument), but the rule is all right. Of course, one may held that an application of "= E" *to a propositional attitude context* is unsound. Likewise, we don't challenge the metaphysical principle known as Leibniz's Law, or the Indiscernibility of Identicals that the formal rule implements.

In the latter case, we face the challenge of showing that, contrary to some semantic intuitions that are shaped by compelling evidence, failure of the application of "= E" to premises embedded within a propositional attitude context is not *the* right diagnosis for what happens when from reporting a belief and from stating identity between two codesignative names, one of which occuring into the belief report, we are not entitled to draw any conclusion whatsoever as to the same content of the ascribed belief save for a substitution of one of the two names for the other. Hence, failure of substitutivity is a misnomer for the problem under consideration, since certain other semantic principles that are unproblematic when they are applied in transparent contexts may bring about paradoxical results in an area of ascription of belief.

The remarks that I put forward in this paper are the results of an exploration of both explanatory alternative options. The main point under which that exploration falls is that a unitary explanation of the failure of ascribing beliefs on logical grounds is desirable, and that the reason for that failure may be found, appearance to the contrary notwithstanding, in some structural features of a belief context which are such that they block the use of a substitutivity principle that outside such a context is sound and unproblematic.

Accordingly, one has to show how arguments such as Kripke's are name-sensitive, i.e., are such that, although they are not based on a principle of substitutivity, they occur in a context which bloks the replacement of an occurence of a name (in a belief sentence) with any codesignative name which the subject of that belief attribution is unaware of fitting in her entertained belief, being thus, on a par with similar arguments resting, however, on a substitutivity principle.

The gain of such a strategy seems to be not only in matters methodological, viz. a unitary frame for explaining diverse failures of belief ascriptions on logical grounds, but in motivating a view as to the reasons of that failure, as well. To my understanding, that is a view that should give a proper due to some intension-based concepts, being as such akin to a Fregean explanation of why codesignative

names can't be substituted in propositional attitude, and hence, belief ascription contexts.

I shall begin with a short diagnosis of the problem. After Kripke's *Naming and Necessity*, no discussion of a philosophical topic concerning sense, reference, and attribution of attitude can possibly be the same as before. Although the gist of the doctrine defended in (Kripke 1980) bears heavily on the semantic mechanism of naming and referring in transparent and modal contexts, its revolutionary import has challenged some of the background suppositions of the theories developed for ascription of belief, as well. Particularly, its main tenet that even within approaches of propositional attitude context "[t]he spirit of […] a Millian line (advocated by Kripke in his earlier *Naming and Necessity*) should be maintained as far as is feasible" (Kripke 1994, p. 360).

Consequently, it seems to me that in Kripke's analysis one can trace two main critical arguments against a Frege-Russell theory of belief ascriptions. The first retains the assumption on which the Frege-Russell account rests, viz. that problems with belief ascriptions are caused by failure of substitutivity of codesignative names, and the argument is that a Fregean analysis of belief contexts falls short of an account of the phenomena to be explained.

The second is more revisionary and destructive because it challenges even the assumption which the Fregean analysis is based upon. To this effect Kripke coins certain examples which purport to show that similar problems with belief reports and ascriptions occur when no substitutivity principle is involved, and that some other principles that are benign when used outside belief ascription contexts, viz. a disquotational principle and a translation principle, may be the cause of our failure of attributing beliefs.[2] Thus, those principles themselves and not a substitutivity principle may break down when we come to such kind of contexts.

The point that my paper is making is that cases like those envisaged by Kripke

---

[2] The Disquotational Principle (DP) is a devise of bridging sentences uttered or otherwise entertained by somebody with the content of her beliefs thereby expressed. It comes in two forms: (i) the weaker form states that the sincerely assent, made on reflection, to a sentence 'p' by a normal English (or, for that matter, of any other natural language) speaker who is not confused by whatever lexical error is a sufficient condition for attributing her a belief with that content which 'p' expresses; (ii) the stronger form makes the former sufficient condition a necessary condition as well.

The Principle of Translation (PT) guarantees the preservation of truth for any true sentence under any translation of it into a sentence of any other language, including the very same language, in which case a homophonic translation of that sentence from one idiolect into the idiolect of the translator is allowed.

However, in order that the DP and PT work in cases in which sentences belong to other natural languages than that of the belief ascriber, a Tarskian disquotational principle for truth (usually left tacit), viz. '*p*' is true iff *p*, where '*p*' inside and outside the quotation marks is to be replaced with any sentence of that language, is also needed.

trigger a name-sensitive context which blocks the use of a substitutivity principle, although no appeal to a substitutivity is overtly at stake in the construction of those cases.

Of course, this is to say neither that Kripke's point against substitution as the heart of the matter in belief ascriptions is misguided, nor that he fails to see how his cases can be construed as being essentially the same as cases in which failure of a substitutivity principle is obvious.

It is only fair to pause here and to notice that Kripke himself is completely aware of the parallelism between his "*Londres/* London" and "Paderewski" cases and that about Jones assenting to "Cicero was bald" and not assenting to "Tully was bald" although Cicero *is* Tully, which is obviously based on a substitutivity principle.[3]

My point, though, is a methodological one, and with a bit of substantive import. Namely that since Kripke's account of cases of failure of belief ascriptions, no matter how instrumental in showing that substitutivity is not the heart of the matter, is not decisive as to what is *that something* which prevents us from drawing conclusions from statements of belief ascriptions, then why not stress as well upon Lois' case where *no failure of such principles as DP and PT is* basically involved more than *the obvious failure of substitutivity*, and, accordingly, why not say that failure of substituting a name with a codesignative name is a salient feature of the contexts in which Kripke' own examples are embedded, coming thus to the idea that failure of substitutivity is *a* paradigmatic (though, maybe, not *the* paradigmatic) case in the understanding of the mechanism of failure of belief ascriptions?

Then, following this line of argumentation, one can try to show how in a context such as Lois' or Peter's, the structural features of the situation which obtains for the subject of the belief ascriptions are such that they prevent her or him, and, consequently, the belief ascriber as well, from proceeding to a substitution of a name which occurs in that context with a codesignative name.

Simply put, Kripke's cases may well not be based on any substitutivity principle, but they are generated within, or are embedded, or at least they trigger a context that blocks substitutivity. And hence, methodologically we are not misguided to address the topic of failure of belief ascriptions through one of the prominent feature of the context in which that failure occurs, viz. through failure of substitu-

---

[3] Kripke is very clear on this matter: "It will intuitively be fairly clear, in these cases, that the situation of the subject is 'essentially the same' as that of Jones with respect to 'Cicero' and 'Tully'. Moreover, the paradoxical conclusions about the subject will parallel those drawn about Jones on the basis of the substitutivity principle, and the arguments will parallel those regarding Jones. Only in these cases, no special substitutivity principle is invoked" (Kripke 1994, p. 364).

tivity of codesignative names within those very contexts.[4]

2 Is Failure of Belief Ascription Intrinsically Failure of Substitutivity?

Whichever way we go, either Frege's or Kripke's, the question of the logical signifi-cance that we attach to "= E" in propositional attitude ascription contexts involves a lot more than a formal problem to be dealt with within the confines of a natural deduction system. For it carries its load over topics in the philosophy of language and of logic such as meaning, reference, quantification, and various kinds of modal locutions.

For the purposes of the present paper, one open question that is of a gen-uine interest is whether Fregean "senses" still have any role to play within seman-tics. This question is more acute now, in the aftermath of (Kripke 1980), because one of the most remarkable achievements of the arguments he deploys in (Kripke 1980) is the demolition of the Frege-Russell tradition in regard to the semantics of names. The descriptive theory of names, the view that names are disguised definite descriptions, and, in general, the Fregean mechanism of referring through senses that are expressed by definite descriptions associated with names have become obsolete. And there has been a growing consensus among logicians and philosophers of language that as long as we stick to direct discourse, to transpar-ent contexts, and the modal case, the Millian view on proper names that Kripke has advocated so forcefully cannot be plausibly denied.

Still, having in view the peculiar consequences that Kripke's Millian view on names may have for opaque contexts, and particularly for the interpretation of the substitutivity of codesignative singular terms (especially names) within propo-sitional attitude ascriptions, one may wonder, and consequently, one can legiti-mately ask, whether Fregean "senses" don't have still something to offer. Thus, isn't there a case for Fregean senses within semantics of belief ascriptions (to name just for convenience one of the propositional attitude ascriptions)?

Kripke's arguments to the effect that there are no Fregean senses develop a robust Millian view on names, and consequently, their target is the Frege-Russell descriptive theory of proper names.

According to the Millian view, the sole semantic function of a name is to name/ refer to its bearer. A name is a tag for the object it names.

---

[4] In the analysis and interpretation that I put forward in my paper I also draw on (Forbes 1994) and (LePore and Loewer 1990).

Mircea Dumitru

A name doesn't refer to its bearer through the senses that are expressed by some definite descriptions associated to it. It directly refers to its bearer. The mechanism of getting a referent for a name consists in two parts: there is an act of dubbing, in which the name is introduced, and then the referent of the name is passed from one speaker to another through a chain that under normal communicational circumstances leaves the referent unchanged, although it may very well change the meanings that normal speakers associate in their idiolects with that name.

Unlike a definite description, a name doesn't describe the object that uniquely satisfies certain identifying properties. Against Frege's doctrine, Kripke argues that a definite description neither gives the meaning of the name that can be put into correspondence with it, nor fixes its referent.

And to finish this sketch of Kripke's view on names, let's add that unlike definite descriptions that can pick up different objects in different counterfactual situations, a name is a rigid designator that refer to the same object in all the counterfactual situations in which its bearer exists.

No matter how well is defended this Millian view on names within transparent and modal contexts, a strong feeling persists that its consequences with respect to contexts involving belief ascriptions are odd. To mention just one of them: names, unlike definite descriptions, are interchangeable within belief context. For, as Kripke puts it, "[w]hether a given subject believes something is presumably true or false of such a subject no matter how that belief is expressed; so if proper name substitution does not change the content of a sentence expressing a belief, coreferential proper names should be interchangeable *salva veritate* in belief contexts" (Kripke 1994, p. 354).

But the problem is that inter-substitutivity of codesignative proper names in belief contexts far from being truth-preserving seems to be at odds with our semantic intuitions.[5] Lois sincerely assents to "Superman can fly," but not to "Clark Kent can fly" even though Superman is Clark Kent. Thus, she believes that Superman can fly, but she does not believe that Clark Kent can fly.

Now, it is obvious that if a strict Millian view is correct the two propositions referred to by the terms "that Superman can fly," and "that Clark Kent can fly" should be identical. For, as Kripke says commenting upon this consequence of the

---

[5] One note on scope disambiguation is in order here. When I speak about failure of substitutivity of codesignative singular terms within belief contexts, I consistently assume, as it is done in the literature, that belief locutions are read as having large scope, or alternatively, although maybe not equivalently that the beliefs are construed *de dicto*. For if we read them as having small scope, or as being *de re* beliefs, then substitutivity of codesignative singular terms succeeds, and accordingly, the arguments whose conclusions are derived by applying "= E" are valid.

strict Millian view on proper names, if "the linguistic function of a proper name is completely exhausted by the fact that it names its bearer, it would appear that proper names of the same thing are everywhere interchangeable not only *salva veritate* but even *salva significatione*: the proposition expressed by a sentence should remain the same no matter what name of the object it uses" (Kripke 1994, p. 353).

However, adapting a kind of Russellian argument, one can say that the two propositions: "that Superman can fly," and "that Clark Kent can fly" are not the same, unless they have all their properties in common. But the proposition that Superman can fly has at least one property that the proposition that Clark Kent can fly lacks, namely that it is believed by Lois.

Of course, this doesn't necessarily impute any inconsistency to Kripke's view on proper names, because he doesn't advocate in his (Kripke 1980) a principle of substitutivity for propositional attitude ascriptions or epistemic contexts. However, it is arguable that Kripke can't be quite happy and comfortable with his Millian view on proper names since propositional attitude ascriptions and epistemic contexts threaten to show that his refutation of Frege is not as total as he might have wanted.

More interesting, though, it is to figure out whether the supposition on which Kripke successfully built the interchangeability of coreferential proper names in modal contexts does not have as a consequence the countenancing of a Fregean position in regard to belief ascriptions. To make his case regarding the thesis of rigidity of proper names from which it follows the interchangeability in modal context of coreferential proper names compelling, Kripke has to make a sharp distinction between metaphysical modalities and epistemic modalities.

For, to hold that "Superman" and "Clark Kent" name the same person only if the sentence "Superman is Clark Kent" *is necessarily true*, despite the fact that Lois *doesn't know it*, one has to support the view that not all necessary truths are known *a priori*, and hence, that there are necessary truths known *a posteriori*. And what Lois can find out is that she kept referring to the same person using, without being aware about that, two different names. But doesn't this support the Fregean view that in referring to one object one can use different "ways of thinking" (or "modes of presentation") of that object?

Roughly speaking, one sensible conclusion to draw is that in contexts in which *our knowledge* of the truth of our sayings about individuals is sensitive to *our* ways of referring to them, there is a place for the modes of fixing the referents of the names we use. And this, on a reasonable reading, is tantamount to accepting a place for Fregean "senses" as modes of fixing the referents of proper names. The supposed compromise sought here would consists in accepting that a name is a

rigid designator, and that its associated definite descriptions are not synonymous with it, but that they are used nevertheless in order to fix its reference.

What will happen, then, to Kripke's doctrine?

> So it appears that even though, according to my view, proper names would be *modally* rigid – would have the same reference when we use them to speak of counterfactual situations as they do when used to describe the actual world – they would have a kind of Fregean 'sense' according to how that rigid reference is fixed. And the divergence of 'sense' (in this sense of 'sense') would lead to failures of interchangeability of codesignative names in contexts of propositional attitude, though not in modal contexts. Such a theory would agree with Mill regarding modal contexts but with Frege regarding belief contexts. The theory would not be *purely* Millian. (Kripke 1994, p. 356)

This diplomatic attitude that urges a compromise between a Millian and a Fregean view, seems to be undermined, according to Kripke, by some semantic facts that point against the Fregean assumptions of the theory regarding belief contexts. Let's take a brief look at them:

One way of looking at the failure of interchangeability of any two codesignative names within belief contexts seems to support the view that its main source consists in the existence of two different associated descriptions,or modes of determining the reference of the names. Were the subject of a belief attribution able to grasp that the two distinct definite descriptions uniquely identify the same object, she would proceed without further ado to the interchangeability of the two names within the belief context. However, this is not something to find out by pure logic, or by merely introspection. Empirical investigations are required.

This move has two salient features: (i) it provides the Fregean theory with a response to the objection that most names don't have conventional senses shared by all the speakers of a linguistic community; however, the price that this response imposes upon those that embrace it is the view that there is no sense shared by all in the community, but only a community-wide reference; (ii) it fares rather well with the Frege-Russell view that names belong to individuals' idiolects, and their senses are dependent upon certain descriptions associated with them.

Now, leaving aside all the objections that have been raised against this twist of the Frege-Russell theory toward a cluster-of-descriptions theory of names, the most serious objection that Kripke opposes to a Freagen analysis of belief contexts is its failure to "account for the phenomena it seeks to explain" (Kripke 1994, p. 356).

The main Fregean dogma that the sense of an expression determines its reference is at odds with a simple linguistic fact, namely that people sometimes do not know about individuals whom they intend to refer to anything that can be used for uniquely identifying them. So, they cannot associate with a name some definite descriptions that uniquely identify the bearer of the name. However, this doesn't hinder them from using names with a particular determinate reference in their minds.

Thus, to cope with the shortcomings brought about by their lack of specific knowledge about individuals whom they intend to refer to, people usually attach to the names they use indefinite descriptions. Now, as Kripke's "Feynman - Gell-Mann" example shows, the senses, if they can be called so, that indefinite descriptions express are identical. Henceforth from the theory it follows that the referent of one name is the same as the referent of the other. But it is obvious that the speakers cannot say whether or not the two names are codesignative only by looking at the senses that the descriptions express. And actually, it is quite plausible that in cases similar to "Feynman - Gell-Mann" example the names refer to distinct individuals.

I am not sure that Frege would have endorsed the view that an associated *in*definite description is successful in *fixing* the reference of a name, let alone in giving its meaning. However, it seems to me that the fact that a definite description is not forthcoming is not sufficient to block our referring to the bearer of a name, whichever, or whoever is that. And this semantic intuition is exploited by Kripke in his "Feynman - Gell-Mann" example against the premise that we consider now of the Frege-Russell theory, viz. that failure of interchangeability of codesignative names is produced by the difference between the descriptions associated with the names.

Hence, although Fregean senses were promising at one step in the investigation of the ascription of belief, it seems now that they are of no real help for us. "So the apparent failure of codesignative names to be everywhere interchangeable in belief contexts, is not to be explained by differences in the 'senses' of these names" (Kripke 1994, p. 359).

Now, having shown that a Fregean theory of names is not after all very helpful in understanding what is happening when we fail to attribute beliefs, Kripke launches a more destructive attack. This time the target is the very idea that the impossibility of substituting a name for another codesignative name in an intensional context lies at the very root of failure of belief ascriptions, regardless what kind of theory intended as an account of failure of substitutivity is considered.

To this effect, Kripke introduces and discusses several cases in which failure of belief attribution is obvious, but no substitution of a name for another name is

apparent. The series of cases is set up in such a way that with each new case the sense-based account of failure of belief attribution and the identification of this latter problem with failure of substitutivity of codesignative names gradually lose their initial credibility they might have had.

Thus, Pierre's "*Londres*/London"[6] case and its variants to be briefly discussed in the sequel purport to show that we are in a predicament if asked what Pierre believes. And this is a puzzle because from his assent to "*Londres est jolie*" by (DP) and (PT) we can infer that Pierre believes that London is pretty, whereas from his assent to "London is not pretty" by the same two principles[7] we are led to the conclusion that Pierre believes that London is not pretty. And the facts are such that Pierre cannot be blamed for any contradiction whatsoever, for as long as he doesn't know that the city he calls "*Londres*" is the same as the city he calls "London" he can't figure out, using logic only, that his beliefs are inconsistent, and thus, that at least one of them must be false.

As Kripke himself points out, "[Pierre] lacks information, not logical acumen. He cannot be convicted of inconsistency: to do so is incorrect" (Kripke 1994, p. 368).

But, then, one can ask how Pierre's position fails to be the same as Lois'?

For what Pierre is in no position to do is to substitute in the translation into English and French, respectively, of either of the two sentences he endorses the English counterpart for "*Londres*" or the French counterpart for "London." And Pierre can't operate this substitution because he lacks precisely that piece of knowledge that "the cities he calls 'London' and '*Londres*' are one and same ..." (Kripke 1994, p. 368).

Couldn't we say, then, that Lois' and Pierre's position don't differ at all, at least with respect to the possibility of making a substitution? And then, isn't Pierre's "*Londres*/London" case a case of failure of attributing belief due to its occurrence in a situation in which a substitutivity of codesignative names fails? I am inclined to see the things in this way.

But be that as it may, Kripke, however, agrees that these paradoxical conclusions about Pierre parallel Lois' predicament. Only that in Pierre's case, no overt recourse to a substitution principle is made.

And on the other hand, as Kripke repeatedly emphasizes, no other equivalent description of Pierre's situation, which is both possible and not conducive to any paradoxes, will do as a solution to the original puzzle. For the puzzle is: Does

---

[6] Cf. (Kripke 1994, 365–372).

[7] Because, for reasons provided by Quine, a homophonic translation from the idiom of the utterer into that of the interpreter should be taken into account when dealing with an interpreter's attempt at understanding what somebody who belongs to the same linguistic community as hers or his is saying.

Pierre believe or does he not believe that London is pretty?

Once again, according to Kripke, the temptation to see this puzzling situation through Fregean lenses, as an outcome of associating two distinct sets of identifying properties with the same name[8], [9]should be resisted for two main reasons.

First, because even if we concede that this Frege-Russellian explanation shows what Pierre's case really boils down to, it is not an adequate answer to the problem we started with, viz., the behaviour of names in contexts of belief ascriptions.

To my understanding, though, the way Kripke emphasizes the problem we started with, viz., "Does Pierre, or does he not, believe that London (not the city satisfying such-and-such descriptions, but *London*) is pretty?" (Kripke 1994, p. 370) – shows his strong penchant toward assimilating names which occur within contexts of belief attributions with rigid designators. But this is question-begging, because this is precisely the point to be argued: are names in contexts of belief attribution behaving like rigid designators (thereby allowing for a sort of substitutivity principle), or like descriptions (and thus we have at least part of the reason why subtitution fails in such contexts)?

Moreover, Kripke's question what does Pierre believe "not [about] the city satisfying such-and-such descriptions but [about] *London*" (Kripke 1994, p. 370), seems to be more appropriately addressed in a context in which, unlike the context of *de dicto* beliefs which is Kripke's own concern, we are concerned with *de re* beliefs.

Second, according to Kripke, the Frege-Russell approach will not do because even if we grant its basic notion that the puzzle arises from the fact that two differ-

---

[8] A more accurate statement is this: two sets of identifying properties which are expressed by two different definite descriptions are associated respectively to each name from a pair of names which belongs in the class of equivalents of all the pairs of names, the former of which is a given name and the latter of which is one of its translation into another language. Of course, there is something odd in saying that a name has a *translation*, but nevertheless this is the case with names of famous people and places. More often than not the translation consists rather in a phonetic variant of the original name due to peculiarities of pronunciation and to the need of matching sounds from one language into more or less equivalent phonemes of other languages than in a genuine translation of the meanings of names. For, if we endorse Kripke's doctrine on names, these lack any meanings which could be rendered either by a synonymous definite description or by a reference-fixing definite description. However, the qualifier "more often than not" is intended as a reminder of those cases in which those which are currently considered as names were in the beginning, when the baptize took place, nicknames. "Richard Lion Heart" or "Vlad the Impaler," which are names in their own right, ask for a matching of the meaning of their descriptive component part (" … Lion Heart", and " … the Impaler") into an equivalent meaning in the target language when they are translated from the original language in which they were coined.

[9] For it seems very naturally to say that had Pierre known that the two distinct sets of identifying properties, which he had learnt in France and in England, respectively, pick up the same city, he would have been in the position to spot an inconsistency within his beliefs using his logical acumen.

ent sets of identifying properties have been associated with "*Londres*" and "London," respectively – an idea that may retain some force regardless the strong arguments against identifying descriptions as being synonymous with, or at least as fixing the reference of, names – and we try to fix the problem by letting only one and the same description be associated with both names, the same type of puzzle will reoccur at another level: the puzzle will involve names which occur at the level of those identifying definite descriptions.

Kripke's point can be interpreted again as being against a substitution based diagnosis of the problem. No constraint along Fregean lines to the effect that Pierre associates the same identifying properties with both "*Londres*" and "London" will do as a solution of the paradox: for, suppose that a definite description phrased in French fixes the referent of "*Londres*" and its translation in English fixes the referent of "London." Now, the same indeterminacy with respect to the referent of a name within a context of belief ascription and of its translation can reoccur, only this time at the level of the names which figure within the uniquely identifying description.[10]

Since the same problem concerning the indeterminacy of the referent of codesignative proper names within contexts of belief attributions can reoccur at any subsequent level at which a "defining" definite description is introduced (because in any such descriptions other names occur, and eventually one would expect demonstratives and indexicals to appear), the hope of a description theorist to provide a Fregean solution to these paradoxical cases is that eventually she can reach an ultimate level where the defining properties are "pure" properties whose expression do not require proper names.[11]

---

[10] For instance if "*Angleterre*" and "*le Palais de Buckingham*" (pronounced "*Bookeengam*") occur in the French definite description which Pierre associates with "*Londres*", then after he learns English by direct method, he can associate with "London" the exact English translation of the French description in which "England" and "Buckingham Palace" will occur throughout in the same places in which "*Angleterre*" and "*le Palai's de Buckingham*" occur in the French description. Now, assuming that Pierre has never seen England and London, but that he was shown some pictures of picturesque parts of the city which is named by a name whose referent is picked up by the French definite description, and then, after he learnt English, he was shown some pictures of unattractive parts of the city which meets the English definite description, then Pierre can consistently believe that London is pretty (because he assents to the French sentence "*Londres est jolie*") and that London is not pretty (because he endorses the sentence "London is not pretty"). And the problem now is the same as before except for the fact that it involves names which occur in the uniquely identifying definite description: Pierre can't figure out that "*Angleterre*" and "England," and "*le Palais de Buckingham*" and "Buckingham Palace" pick up the same referent, respectively.

[11] But in view of the possibility of extending Kripke's puzzling cases to natural kind terms (cf. (Putnam 1975) and (Kripke 1980, 1994)), in the defining descriptions at that presumed ultimate level couldn't occur natural kind terms either.

However, no convincing and plausible proposal as to how such an ultimate level can be reached has been advanced, or, if supposing that it can eventually be reached, how can the properties expressed by such name-free descriptions be still uniquely identifying after all names and related devices for unique referring have been eliminated.

The only way out from this predicament is to endorse a rather extreme philosophical reform which consists in banning the translation of names.

Actually, apart from any point as to the merit of this proposal, to acknowledge that the only solution consists in forbidding translation of names, which in this cases only means to preserve names as they are in their original phonetical guise, fits quite well with my idea that for reasons of not jeopardizing the content of a belief the context within which that ascription of belief is realized should be such that a substitution of a name for a codesignative name is forbidden, unless the subject of that belief attribution is aware of (has knowledge of) the fact that the two names are codesignative.

Although this proposal can be reckoned as being highly implausible in face of our customary practice of learning the names of famous people and places as names within our own language, the proposal, if it is not sentenced before it is judged, looses its *prima facie ad-hoc* character.

For, the practice of translation of names is limited only to names of some famous people and geographical localities, and to ask for a reform, in which we stick to non translated names within contexts of belief ascriptions for the sake of avoiding puzzling cases about belief, is nothing else than to generalize a practice which is in use for the vast majority of names.[12]

Besides, what one gets by translating such a name is a phonetic variant in the target language of that name as it occurs in the original language, and thus, what runs against this proposal is not something pertaining to semantical analysis, but to the weight of tradition.[13]

However, it is not so clear how well this proposal could work. For, suppose that for philosophical reasons concerning reports of beliefs we adopt the policy of non-translatability of names. But once in force, how far does this principle reach? If Putnam's and Kripke's work on the semantics of natural kind terms adequately reveals analogies between proper names and natural kind terms, then one point in case seems to be the similar behavior of both sorts of terms within belief reports. Should we, then, extend a similar policy of non-translatability to

---

[12] As Kripke puts it: "At least the restriction in question merely urges us to mend our ways by doing *always* what we presently do *sometimes*" (Kripke 1994, p. 374).

[13] For the problem raised by the translation of what initially were nickuames, but are currently deemed as names see footnote 7.

natural kind terms? If similar problems ask for similar approaches, then one can't see how to adopt this proposal for one sort of terms, and deny it for the other sort. But this brings the whole idea to an extreme implausibility.

What is typically involved in all these cases of failure of belief ascription is a lack of awareness on behalf of the subject of ascription that two names, or two tokens of the same name in different languages, refer to one individual. For otherwise, if she knew this, she would be able to spot a contradiction within her beliefs by her logical acumen, only.

Is it, then, necessary that in all such cases two distinct tokens of the same name, one of which being the translation of the other, or two distinct tokens of two names occur? Moreover, is it the case that translation from one language into another is essentially involved in Kripke' cases, leaving, thus, open the possibility that the reason of those paradoxically results be that it is something wrong with translation?

Surprisingly not. As the "Paderewski" case shows the same puzzle still arises even if in reporting beliefs we stick to one language only, and even if we confine ourselves to phonetically identical tokens of a single name. In a way "Paderewski" parallels "*Londres* /London" case, and just because no Principle of Translation is needed now,[14] but only the Disquotational Principle, this new version of the puzzle shows more clearly that what is involved in the context of Peter's learning of the referent of the name "Paderewski" is a fundamental ambiguity and that this ambiguity is one source of the puzzle and of our predicament in attributing beliefs to Peter.

For Peter assents to "Paderewski has musical talent," and hence believes that, consecutively to his learning of the name "Paderewski" through a description (definite or indefinite, it doesn't matter) which fixes the referent of that name in the person of the famous pianist.

Then, when he assents to "Paderewski has musical talent," and accordingly he believes that without changing his mind about Paderewski the musician, and without being responsible for any inconsistency, he has introduced the referent of the same name by a referent-fixing description like "the Polish nationalist and Prime Minister."

And because there is no *a priori* guarantee that "Paderewski" names the same individual, for a name in a natural language can denote ambiguously and has more than one denotation[15], it would be simply fallacious for Peter to say and be-

---

[14] Save for those who may make a case for a homophonic translation from Peter's idiolect into the idiolect of the belief ascriber, but even then it is not a translation from one *language* into *another* one.

[15] With respect to this feature, the "two languages" example still works even if we spoke languages in

lieve that whatever holds with respect to the referent of "Paderewski-the-pianist" should also hold with respect to the referent of "Paderewski-the-politician."

To spell the details of this ambiguity think at a certain model which makes true the sentence "Paderewski has musical talent," which Peter assents to, thereby believing it. In the domain of that model we have to put an individual and make "Paderewski" refer to it, and put that individual, as well, in the extension of the predicate "x has musical talent."

Now, "Paderewski has no musical talent" can't be true in the same model, if "Paderewski" uniquely and unambiguously refer to the same individual in the domain. Hence, if we want that Peter do not contradict himself in believing that Paderewski doesn't have musical talent, since we have no recourse to other worlds than the actual world where the individual which "Paderewski" refers to is not in the extension of "x has musical talent," one should fix other parameters of the model such that Peter's beliefs do not get inconsistent.

There may be several options but what seems prominent for the case under discussion is that Peter learns the referent of "Paderewski" which makes the sentence "Paderewski has musical talent," true, say, at a concert. Thereafter, whenever he will utter the sentence "Paderewski has musical talent," he will have in mind *that* "Paderewski" he listened to that concert, and the referent of that Paderewski is fixed by the description "the pianist listened to that concert."

Whereas, Peter learns the referent which he takes for granted that it makes the sentence "Paderewski has no musical talent," true under different circumstances which do not entitle him to believe that the same individual is the referent of "Paderewski" in both sentences that he endorses. Those circumstances are such that they can fix the referent of "Paderewski" though it could be perception again, through a description such as "the Prime Minister of Poland" or "a Polish nationalist."

This seems to be a compelling case for stating that in reporting beliefs names do not refer rigidly, but through description which fix their referent, and this accounts for the fact that names may refer ambiguously just in case one and the same name is associated with two different description which happen not to pick up the same individual.

Coming back to the model which makes the first sentence true, we cannot say that it refutes the second sentence, unless the different descriptions associated with "Paderewski" do pick up the same individual in the domain. And since nobody

---

which all names must denote uniquely and unambiguously. Of course, we never speak languages in which all names denote uniquely, but this being possible obscures the presence of an ambiguity which the construction of the "Paderewski" case rests on.

can know such a fact *a priori* to reason assuming such a supposition would be fallacious.

Let's take a look at the same case from another angle.

Imagine this dialogue between Peter and one of his friends. The latter asks Peter what does he believe about Paderewski, or if we want something more definite, about his musical gift.

Now, supposing that Peter believes that there are two persons with this name, i.e., that the name is used ambiguously and not with a unique referent, he should act upon this belief and proceed first by disambiguating the question. Then, it is very likely that prior to answering his friend's question Peter will ask him: "Paderewski – who? The musician or the politician?"

Now, there are two plausible ramifications of the situation: either Peter's friend is like Peter in not being aware that Paderewski the musician *is* Paderewski the politician, or else Peter's friend is more knowledgeable and he will point out to Peter that "Paderewski" refers in that context to just one person. (Of course, we are not obliged to think that they are using the logicians' jargon.)

Depending on which one is the case, Peter will either reiterate his beliefs about the musician *and* the politician, because for him there might be two different persons, or else will notice that his otherwise prudent attachment to the principle not to infer that one name should refer to one and the same person failed to apply, because this time, the name he learnt in different situation *does* refer to only one person. Accordingly, he will have to give up one of his beliefs, on pain of being accused of logical inconsistency.

So, actually there was no need for disambiguation, but since Peter believed there was, one should retain Peter's use of a description as a referent fixing devise for a name which occurs in a belief report. And besides, if his friend is more knowledgeable and wants to spot Peter's mistaken idea that "Paderewski" is used ambiguously, he still needs the referent-fixing description device in order to tune the two tokens of "Paderewski" that Peter is using. Thus, it is very likely that he will say to Peter that "the famous pianist" and "the Prime Minister" pick up the same individual in the person of Paderewski.

Hence, in both situation, viz. either there is an ambiguity and one needs to control it, or there is no ambiguity and one needs to dispel the mistaken impression that there might be one, the usage of a description for fixing the referent of the name is decisive in restoring the communication.

Consequently, although the "Paderewski" case makes the restriction that names should not be translated, if we want to avoid the puzzle, irrelevant, precisely because the puzzle occurs even though the name is phonetically repeated, it doesn't seem to me that it is as successful in showing that the substitutivity issue is en-

tirely alien and irrelevant for the problem raised by the puzzle.

For, if we grant that Peter's assent to contradictory sentences, without thereby being committed to having inconsistent beliefs, is due to his wrong belief that "Paderewski" is used ambiguously, then his failure of seeing that the two sentences *are* contradictory can be accounted for by letting names within contexts of belief attribution behave like descriptions which are asked for fixing their referents.

Accordingly, one token of a name cannot be substituted for another token of the same name, even though they are codesignative, because their referent is fixed by different descriptions, and there is a very sound presumption that the referent of a token might not coincide with the referent of the other.

To risk a substitution in such cases and to get a true conclusion from a premise, or in Lois' case from two premises, is like winning the lottery. It would be to proceed fallaciously because the reasoning would be based on a coincidence, viz. that the two reference-fixing descriptions pick up the same referent for the name to which they are associated.

For this reason, I venture to say that the description-like behavior of names within contexts of belief attribution is the dual of those rare but, however, existing cases in which definite descriptions behave like names; descriptions, like e. g. "The Holy Roman Empire", where the descriptive conditions don't determine the referent, refer rigidly like names do in modal contexts.

I guess we can't say that Kripke has a fully developed theory of belief attribution and opacity. His main purpose in (Kripke 1994) is to provide reasons in favor of the idea that the substitutivity of identicals is not responsible for getting the paradoxical results in arguments involving ascription of belief, and thus, either "=E" is a perfectly legitimate rule in belief contexts, or at least it is not responsible for failure of belief ascriptions.

What "*Londres*/London" example and "Paderewski" example purport to show is that even though a clear case of substitutivity is absent, still there are problems with belief ascriptions. And other principles that lie behind the normal practice of linguistic communication, such as the disquotational principle and the principle of translation, are perhaps responsible for Kripke's puzzle about belief which any theory of belief and names should deal with. The nature of the problem is still elusive, but a hypothesis can be that the cases envisaged in those examples "lie in an area where our normal apparatus for ascription of belief is placed under the greatest strain and may even break down" (Kripke 1994, p. 379).

In the end I want to reiterate and to emphasize some scattered points I made earlier as reactions to, and comments on, Kripke's very ingenious cases and his subtle proposals.

Mircea Dumitru

The point that my paper is making is that cases like those envisaged by Kripke trigger a name-sensitive context which blocks the use of a substitutivity principle, although no appeal to a substitutivity principle is overtly at stake in the construction of those cases.

Simply put, Kripke's cases may well not be based on any substitutivity principle, but they are generated within, or are embedded within, or at least they trigger a context that blocks substitutivity.

My reading of this proposal is articulated counterfactually, as follows:

Had the contexts in which "*Londres*/London" and "Paderewski" cases, respectively, occur been such that they would have allowed for a substitution of "*Londres*" for "London" and of "Paderewski-the-musician" for "Paderewski-the-politician" within Pierre's/Peter's idiolect, then no Kripkean puzzle about what does Pierre/Peter believe would have obtained. For Pierre/Peter could have used his logical acumen to spot an inconsistency within his set of beliefs.

Thus, although Kripke's cases don't explicitly rest on a substitutivity principle, they seem to be embedded within contexts which do not allow for such a substitution. In this sense, they could be seen on a par with Lois' case.

Accordingly, one has to show how arguments such as Kripke's occur in contexts which are name-sensitive, i.e., are such that, although they are not based on a principle of substitutivity, they occur in a context which blocks the replacement of an occurrence of a name (in a belief sentence) with any codesignative name which the subject of that belief attribution is unaware of fitting in her entertained belief, being thus on a par with similar arguments resting, however, on a substitutivity principle.

Methodologically, then, we are not misguided to address the topic of failure of belief ascriptions through one of the prominent features of the context in which that failure occurs, viz. through failure of substitutivity of codesignative names within those very contexts.

One very difficult question on which I can speculate only is why contexts of belief ascriptions are name-sensitive?

If by dividing and then eventually conquering the problem we can hope for gaining understanding, then I would like to conjecture about the logical part of the issue, and to defer the discussion on the more speculative aspect concerning the *structure* (*architectonic*) of the content of our beliefs for future discussion.

For the logical facet: I begin with the observation that if both Lois an Pierre/Peter substituted a codesignative name or a translation/phonetic variant of a name, respectively, for a given name or for the original name, respectively, then they would proceed fallaciously.

This suggests that a validity preserving strategy requires at least to refrain

from drawing any conclusion from a belief report in which a name occurs, unless the subject of that belief attribution is aware of the sameness of reference of that name and of another name. Only if this is the case, we could draw a conclusion having the same content with that of the belief report save for the substitution of the codesignative name for the name which occurs in the original belief report.

Apparently, if the content of a belief ascription is not preserved under *that particular way* in which it is believed, i.e., if we don't keep the frame of the content of that belief ascription sealed, and we allow for a substitution of *any* codesignative name for *any* given name within that belief ascription, then the arguments we get thereby are no more truth-preserving.

And this suggests, further, that in order to have a truth-preserving argument in a context of belief ascription a necessary condition is not to jeopardize the ascribed content by substituting a codesignative name for a given name, unless the subject of that belief attribution is aware of (has knowledge of) the fact that the two names are codesignative.

Then, it is legitimately to ask more about the semantics of singular terms, in particular of names, and about their behavior, within contexts of belief attribution.

It seems very likely that cases like those considered in this paper show that we cannot assign a fixed semantic role for names, regardless the context of their occurrence.

This remark boils down neither to the view that names are synonymous to definite descriptions, nor to the view that in all contexts the referent of a name gets fixed through a definite description associated with that name.

Rather, it has only in view that the same way in which there are cases in which descriptions behave like rigid designators, there also are cases in which names behave like descriptions, and that, in particular, a name has a description-like behavior within contexts of belief ascriptions.[16] But then, this is the reason of failure of substitutivity, and this accounts at least in part for cases of failure of attribution of beliefs.

Roughly speaking, one sensible moral to draw is that in contexts where *our knowledge* of truth of our sayings about individuals is sensitive to *our* ways of referring to them, there is a place for the *modes of fixing* the referents of the names we use. And this, on a reasonable reading, is tantamount to accepting a place for Fregean "senses" as modes of fixing the referents of proper names.

The supposed compromise sought here would consist in accepting that in most contexts a name is a rigid designator, but that, nevertheless, in belief ascription

---

[16] For instance, in "Paderewski" case, a reference-fixing description is needed in order to disambiguate the use of the name "Paderewski."

contexts they have a description-like behavior, and precisely for reasons which have to do with disambiguation of the discourse, and for the need of keeping under control the indeterminacy of the referent of that name, a referent-fixing description is associated with it.

My primary motivation for discussing the main tenets of Kripke's theory of proper names in relation with belief attribution has been prompted by his powerful attack against intensional entities and their alleged place within semantics.

However, whereas the arguments deployed in (Kripke 1980) leave almost no hope for the friends of intensions, at least in regard to transparent and modal contexts, Kripke's case against intensions is not so compelling with regard to belief ascription and other propositional attitudes. Here, there are no knock-down argument that render the need for intensional entities superfluous.

Thus, to the extent to which nothing definitive, either positive or negative, can be said about substitutivity within belief ascription contexts, the overall moral we can draw from the discussion of this issue is that in the state of our present knowledge there are no serious reasons to give up the exploration of the failure of attributing beliefs through the topic of failure of substitutivity.

References

Forbes, Graeme. 1994. "Introduction to Opacity and Attribution of Attitude." In *Basic Topics in the Philosophy of Language,* edited by R. M. Harnish, 323–337. United Kingdom: Prentice Hall.

Kripke, Saul. 1980. *Naming and Necessity.* Cambridge, MA: Harvard University Press.

———. 1994. "A Puzzle about Belief." In *Basic Topics in the Philosophy of Language,* edited by R. M Harnish, 352–392. Prentice Hall.

LePore, Ernest, and Barry Loewer. 1990. "A Study in Comparative Semantics." In *Propositional Attitudes: The Role of Content in Logic, Language, and Mind,* edited by C. A. Anderson and J. Owens, 91–111. United States: Cambridge University Press.

Putnam, Hillary. 1975. "The Meaning of 'Meaning'." *Minnesota Studies in the Philosophy of Science* 7:131–193.