Two Types of Rigid Designation

Iris Einheuser
Department of Philosophy
Harvard University
Emerson Hall
Cambridge, MA 02139-4307
Email: ieinheus@fas.harvard.edu

[Published in *Dialectica* Volume 59, Issue 3 (2005)]

The notion of a rigid designator was originally introduced with respect to a modal semantics in which only one world, the world of evaluation, is shifted. Several philosophical applications employ a modal semantics which shifts not just the world of evaluation, but also the world considered as actual. How should the notion of a rigid designator be generalized in this setting? In this note, I show that there are two options and argue that, for the currently most popular application of two-dimensional modal semantics, proper names ought to be treated as rigid relative to the world considered as actual.

I. Conceptual truth and actuality. As is well-known, the ordinary notion of metaphysical necessity does not capture the sense in which *conceptual*, or *a priori*,[1] truths expressed by sentences containing 'the actual F'-locutions are necessary. While the conceptual truth

(Z) If a unique person invented the zip then the inventor of the zip invented the zip

is metaphysically necessary, the conceptual truth

(Z′) If a unique person actually invented the zip then the actual inventor of the zip invented the zip

is not, as someone other than Whitcomb Judson, the actual inventor of the zip, might have invented the zip.

[1] Here, I will use 'conceptual truth' and 'a priori truth' interchangeably.

Z′ is one of many examples which show that the metaphysical notion of necessity and the epistemological notion of apriority are not, as had traditionally been believed, two sides of the same coin.[2] But even though Z′ is not metaphysically necessary, it is *conceptually necessary* in the sense that someone competent with the concepts involved will be in a position to know that it is true. To capture this notion in a familiar framework, the ordinary modal framework has been extended to allow for variations in what world is *considered as actual* (Davies and Humberstone, 1980).

There are various construals of what it is to consider a world as actual,[3] but the core idea relevant for present purposes is this: The actual world plays a role in determining the referents of an interpreted language's expressions. By considering a world other than the *truly actual world*–that is, *our* world—as actual, we imagine this other world to determine the extensions of our expressions. A linguistically competent speaker may be ignorant of the extra-linguistic facts that help fix the referents of her language. That is, she may not know which world is actual as far as the relevant extra-linguistic facts are concerned.[4] For a sentence to be conceptually necessary, then, it must not depend for its truth on the extra-linguistic facts that competent speakers may be ignorant of. Therefore, a sentence should be deemed conceptually necessary just in case it is true no matter how the reference-fixing facts turn out, or in other words, no matter what world is considered as actual. So while *metaphysical* necessity is truth in all worlds considered counter*factually*, *conceptual* necessity is truth in all worlds considered counter*actually*:

($\Box_c$-NECESSITY) $\Box_c S$ is true if and only if for every possible world w, S is true at w if w is considered as actual.

Similarly, we can define S to be conceptually possible just in case $\Diamond_c S$, i. e. just in case there is a world w so that S is true in w if w is considered as actual.

[2]Other examples that make the point have been put forward by philosophers like Saul Kripke (Kripke, 1972), Gareth Evans (Evans, 1979) and David Kaplan (Kaplan, 1989). The notions of metaphysical necessity and apriority, or conceptual necessity, continue to be controversial. The kind of two-dimensional framwork discussed in this paper is meant to shed some light on the notions and the way in which they are related.

[3]For a useful discussion of what it may be to consider a world as actual see (Chalmers, forthcoming) and (Stalnaker, 2001).

[4]Note that no speaker knows what world is actual, unless she is omniscient. But even non-omniscient speakers may know what world is actual as far as the extra-linguistic facts are concerned which determine a given utterance's content.

II. Actuality and two types of rigidity. A singular term is rigid just in case it refers to the same individual in every possible world (and does not refer in a world if that individual does not exist at that world). Equivalently, a singular term is rigid just in case the object it actually refers to is relevant for evaluating a sentence containing the term at any world. In a setting in which worlds can be considered either counterfactually or counteractually, we can distinguish between two notions of rigidity for singular terms. On the one hand, there is the notion of

(ABSOLUTE RIGIDITY) A singular term t is *absolutely rigid* just in case there is an individual i in the truly actual world such that for every world w, t refers to i in w.

In other words, a term is absolutely rigid just in case there is an individual in the *truly* actual world which is relevant for the evaluation of any sentence containing that term at a world regardless of what world is *considered* as actual. On the other hand, there is the notion of

(@-RIGIDITY) A singular term t is *@-rigid*, just in case for all worlds w, if w is considered as actual, then for all i, t refers to i in w (where i is in w) iff t refers to i in all worlds.

In other words, a term is @-rigid just in case for all worlds w, if w is considered as actual, then the individual relevant for the evaluation of a sentence containing that term at a world considered as counterfactual exists at w.[5]

Intuitively, @-rigidity relativises the ordinary notion of rigidity to the world considered as actual, whatever it may be: The world *considered* as actual—rather than the truly actual world—fixes the referent of an @-rigid term. In a context in which no actuality-shifting operators (like $\Box_c$ and $\Diamond_c$) are present, @-rigidity coincides with the ordinary notion of rigidity, which turns out to be a special case of @-rigidity.

Note that the very same distinction can be drawn for predicates. In recent literature, an important class of predicates, the natural kind terms, has been implicitly treated as @-rigid. Thus, for instance, 'water' refers in all worlds considered as counterfactual to H_2O. But if a world had been actual in which xyz, rather than H_2O is the substance filling the lakes and falling from the sky as rain, then 'water' would have referred, in all worlds considered as counterfactual, to xyz.

[5]The next section provides a more detailed semantics for simple languages with @-rigid singular terms.

For a class of singular terms which are @-rigid, consider 'the actual F'-constructions. Their point is to fix on a referent relative to the world *considered as actual*: If a world is considered as actual in which Ben Franklin invented the zip, then 'the actual inventor of the zip' refers, in all possible worlds, to Ben Franklin. If a world is considered as actual in which Marie Curie invented the zip, then 'the actual inventor of the zip' refers, in all possible worlds, to Marie Curie. And so on.[6] Let us symbolize the locution 'The actual F is G' as $(@x : Fx)(Gx)$. Here, for any predicate F, the term $(@x : Fx)$ refers to the unique F in the world considered as actual. Then 'the actual F'-constructions interact with the modal operators as follows

$\Box_c[(@x : Fx)(Gx)]$ is true if and only if in every world w the unique individual that is F in w is G.

$\Box[(@x : Fx)(Gx)]$ is true if and only if in every world w, the unique individual that is F in the actual world is G in w.

And similarly for $\Diamond_c$ and $\Diamond$. Before considering an application of the suggested distinction, let us see what the semantics of languages with rigid terms of either kind looks like.

III. The Semantics. We assume that the only non-logical constant of the language are predicates and constants. The structures over which discourse containing conceptual necessity- and actuality-operators are interpreted are of the form $\langle D, W, w^*, d, I\rangle$, where D is a collection of individuals, (the 'domain'), W is the set of points (the 'possible worlds'), w^* is a designated element from W (the 'truly actual world'), d is function (the 'domain function') which assigns to each $w \in W$ a subset of D (the individuals that exist in w), and I a function which assigns semantic values to the non-logical constants of the language, such that for every n-place predicate P^n,[7]

$$I(P^n) \in \{f \mid W \to \mathcal{P}(D^n)\}^8$$

[6]Note, however, that just as some definite descriptions—such as 'the square root of 4'—are *de facto* rigid, some 'the actual F'-constructions are *de facto* absolutely rigid, such as 'the actual number 2'. Thanks to one of the referees for pointing this out.

[7]Since the emphasis here is on the semantics of singular terms, we treat predicates as absolutely rigid. But the basic framework can be straightforwardly modified to allow for @-rigid predicates like natural kind terms.

[8]That is, the semantic value of an n-ary predicate is a function which maps a possible world to a set of n-tuples of individuals. In other words, the semantic values of n-ary predicates are predicate-intensions.

and for every individual constant c,

$$I(c) \in \{f \mid W \to \{g \mid W \to D\}\}^{9}$$

Relative to structures of this type, satisfaction relative to a variable-assignment s can be defined recursively in the familiar way:[10]

$$\begin{array}{rcl} \models^{w'}_{w,s} P(t_1,\ldots,t_n) & \text{iff} & \langle r(t_1),\ldots,r(t_n)\rangle \in I(P)(w) \text{ where } r(t)=s(t) \text{ if } t \text{ is} \\ & & \text{a variable and } r(t)=I(w')(w) \text{ if } t \text{ is an individual} \\ & & \text{constant} \\ \models^{w'}_{w,s} \forall x\varphi & \text{iff} & \models^{w'}_{w,s'} \varphi \text{ for all assignments } s' \text{ which differ from } s \\ & & \text{at most in what they assign to } x \text{ and whose range} \\ & & \text{is a subset of } d(w) \\ \models^{w'}_{w,s} \Box\varphi & \text{iff} & \forall w'' \in W \models^{w'}_{w'',s} \varphi \end{array}$$

Clauses for the truth-functional connectives, the existential quantifier and the possibility-operator ($\Diamond$) are the usual ones. The only noteworthy clauses in the present setting are the clauses concerning the conceptual modal operators

$$\models^{w'}_{w,s} \Box_c\varphi \quad \text{iff} \quad \forall w'' \in W \models^{w''}_{w'',s} \varphi$$

(similarly for $\Diamond_c$), and the clause concerning ($@x:\varphi(x)$)-constructions:[11]

$$\begin{array}{rcl} \models^{w'}_{w,s} (@x:\varphi(x))(\psi(x)) & \text{iff} & \exists a \in d(w') \text{ s.t. } \models^{w'}_{w',s} \varphi[a] \\ & & \text{and } \forall b \in d(w')\Big[\Big(\models^{w'}_{w',s} \varphi[b]\Big) \Rightarrow a=b\Big] \\ & & \text{and } \models^{w'}_{w,s} \psi[a] \end{array}$$

Aa statement is true in $\langle D, W, w^*, d, I\rangle$ just in case it is satisfied relative to all variable-assignments at the truly actual world when it is considered as actual:

$$\begin{array}{rcl} \models \varphi & \text{iff} & \models^{w^*}_{w*,s} \varphi \text{ for all variable assignments } s \\ & & \text{whose range is a subset of } d(w^*) \end{array}$$

[9] That is, the semantic value of a constant is a function which maps a possible world (considerd as actual) to a function which maps a possible world (considered as counterfactual) to an object in that world. In other words, the semantic value of an individual constant maps possible worlds (considered as actual) onto individual concepts.

[10] It is assumed that all worlds are mutually accessible so that '$\Box$' is governed by the S5-axioms.

[11] '$\models^{w'}_{w,s} \varphi[a]$' is to be read as 'object a satisfies $\varphi(x)$ under variable-assignment s at world w with w' considered as actual'.

Finally, a statement is *valid* just in case it is true in all structures $\langle D, W, w^*, d, I\rangle$ satisfying the above conditions.

This framework allows us to treat individual constants either as absolutely rigid or as @-rigid. To treat them as absolutely rigid, we require their semantic values to be *constant* functions from worlds to constant functions from worlds to individuals. To treat them as @-rigid, we allow that their semantic values be (*non-constant*) functions from possible worlds to constant functions from worlds to individuals. That is, we let the world *considered* as actual (rather than the truly actual world) fix the reference of the term.[12]

IV. An application. I will now apply the distinction drawn above to a recent argument for the conclusion that the metaphysical modalities are not governed by S5 because they do not obey Brouwer's axiom ($p \rightarrow \Box\Diamond p$). Crucial to this argument is the intricate interplay between metaphysical modal operators (which are taken to shift the world of evaluation), conceptual modal operators (which are taken to shift the world considered as actual), rigidly referring proper names and actuality-operators. We will see that the argument is fallacious in that it rests on a confusion between the two types of rigid designation. The argument, due to Yannis Stephanou, starts with the observation that

(1) It is necessarily the case that if someone is Aristotle, then he is the actual Aristotle.

(2) If Aristotle did not exist, then it is impossible that someone should be the actual Aristotle.

are necessary in the sense that they are conceptual, or a priori, truths (Stephanou, 2000).[13]

Applying the machinery introduced in above, the conceptual truth of (1) and (2) can be expressed as

(1′) $\Box_c\Box\forall x[(x = \text{Aristotle}) \rightarrow (@y : y = \text{Aristotle})(x = y)]$

[12]Even if we decide in general to treat proper names as @-rigid when modelling conceptual possibility and necessity, some such names will turn out to be *de facto* absolutely rigid. For instance, arguably the numeral refers, as a matter of conceptual necessity, to the number 2. See footnote 6.

[13]I read 'the actual Aristotle' as synonymous with 'the actual individual identical with Aristotle'. Stephanou's observation is based on the Kripkean intuition that 'Aristotle' is a rigid term, so that in any world w, 'Aristotle' refers in w only if it refers in w to what it refers to in the actual world. So if 'Aristotle' fails to refer in the actual world, it refers in no world at all.

(2′) $\Box_c[\neg\exists x(x = \text{Aristotle}) \rightarrow \neg\Diamond\exists x(@y : y = \text{Aristotle})(x = y)]$

where (1′) is the $\Box_c$-necessitation of (1) and (2′) that of (2).[14] From this, Stephanou infers

(3) $\Box_c[\Diamond\exists x[(x = \text{Aristotle}) \rightarrow \Diamond\exists x(@y : y = \text{Aristotle})(x = y)]$ from 1′

(4) $\Box[\Diamond\exists x(x = \text{Aristotle}) \rightarrow \exists x(x = \text{Aristotle})]$ from 2′ and 3

Then, $\Box$–distribution and appeal to Brouwer's axiom ($p \rightarrow \Box\Diamond p$) allows us to conclude

(5) $\exists x[(x = \text{Aristotle}) \rightarrow \Box\exists x(x = \text{Aristotle})]$

As this consequence is clearly unacceptable, Stephanou concludes that we ought to reject Brouwer's axiom and with it S5. So much for the argument.

What is puzzling in the above argument is the transition from (2′) and (3) to (4), as (4) is an ordinary $\Box$-necessitated formula, while the premises are all $\Box_c$-necessitations. In what follows I will show that the argument does not go through because the premises (1′) and (2′) are true only if 'Aristotle' is treated as absolutely rigid, while the intermediate conclusion (4) is true only if 'Aristotle' is treated as @-rigid.

Suppose first that 'Aristotle' is absolutely rigid. Then both premises, (1′) and (2′) would be false: Consider w and w^* such that a is the Aristotle of w^* (the truly actual world) and no Aristotle exists at w. Then if w is considered as actual,

(1″) $\forall x[x = \text{Aristotle} \rightarrow (@y : y = \text{Aristotle})(x = y)]$

is false at w^*, as the Aristotle of w^* is not the 'actual' Aristotle (that is, the Aristotle of w). So (1′) is false. And similarly for (2′): Let w^* be as before with 'Aristotle' referring absolutely rigidly to a. Let w'' contain an individual i such that 'the actual Aristotle' refers to i if w'' is considered actual. Then

(2″) $\neg\exists x(x = \text{Aristotle}) \rightarrow \neg\Diamond\exists x(@y : y = \text{Aristotle})(x = y)$

is false at w'' if w'' is considered actual, since a, the referent of 'Aristotle', does not exist at w''. So if 'Aristotle' is treated as absolutely rigid, the argument against S5 is unsound.

[14] Stephanou uses different notation. But his remarks indicate that this is how he intends to be understood.

Suppose next that 'Aristotle' as @-rigid. This would rule out the previous scenario because the name 'Aristotle' would refer, in every world (actual or counterfactual) to the same individual as 'the actual Aristotle', so that if 'Aristotle' has no referent in the world considered as actual, it has no referent in any other world either. So on this assumption, both $(1')$ and $(2')$ are true. However, (4) is false: At w (considered as counterfactual) it is possible that Aristotle should exist, yet Aristotle does not exist in w.

Once the different possible construals of the interaction between conceptual modal operators, rigid designators and actuality operators are brought into the open, it should be easier to avoid the kind of confusion driving this argument. The question remains what type of rigidity proper names display. I will argue that when the two-dimensional apparatus is meant to capture conceptual truth, they should be treated as @-rigid.

V. Proper names are @-rigid. First note that rigid designation begins at home. We cannot refer rigidly to an individual that is merely possible. Here is why. There are in principle two ways of introducing a rigidly referring name. First, it may be introduced via an act of direct reference. For instance, we can (either mentally or physically) point to an object and declare 'Let NAME refer to *that*'. This requires that the introducer of the name be appropriately related to the target referent. Current candidates for what the relevant kind of relation might be include causal relations and various non-causal acquaintance-relations. It appears that whatever one takes the required relation to consist in, if the target referent does not exist but is merely possible, then no one stands in that relation to it, so this way of introducing a rigid name is out. The second way of introducing a rigidly referring name is by description, where the description in question *fixes the reference* but does not *give the meaning* of the name. That is, the description cannot be thought of as determining with respect to an arbitrary counterfactual world that the referent in that world is the object which uniquely satisfies a certain description. Rather, the description has to pick out, in the actual world, the object which is to be the referent of the name in every world in which that object exists. But if the object we want to refer to does not exist, this way of fixing the name's reference is not available either.

The idea that rigid designation begins at home relativizes as we relativize 'home' to worlds considered as actual. Consider

(A) If 'Aristotle' refers, then Aristotle is the actual individual who is Aristotle,

or, translated to our symbolism

(A′) Refers('Aristotle')$\rightarrow$ ((@$x : x$ = Aristotle) = Aristotle)

Since proper names, if they refer at all, refer to individuals which are not merely possible, (A) is a conceptual truth. If proper names were absolutely rigid, statements like (A) would, relative to the framework proposed, not be classified as conceptual truths, because they would not be true in all possible worlds considered as actual: Suppose 'Aristotle' refers absolutely rigidly to Aristotle, the teacher of Alexander the Great. Consider w, a world in which Aristotle does not exist. Then in w considered as actual, 'Aristotle' refers to Aristotle, but 'the actual individual who is Aristotle' is empty. Therefore, in w considered as actual (A) is false. So in order for the $\Box_c$-account of conceptual necessity to be adequate, names have to be treated as @-rigid.

In summary, care has to be taken when generalizing the notion of a rigid designator in the presence of actuality-shifting operators. If we want shifts in what world is considered as actual to capture the notion of conceptual truth, we need to construe proper names as having their reference fixed relative to worlds *considered as actual* rather than relative to the truly actual world. That is not to say that names always have to be treated as @-rigid. There may well be applications of the 2-dimensional framework relative to which treating names as absolutely rigid is more appropriate.[15]

References

[1] Chalmers, D. 1997, *The Conscious Mind*, Oxford University Press

[2] Chalmers, D. forthcoming, 'The foundations of two-dimensional semantics', in M. García-Carpintero and J. Macià, eds., *Two-Dimensional Semantics: Foundations and Applications*, Oxford University Press

[3] Davies, M. and Humberstone, L. 1980, 'Two notions of necessity', *Philosophical Studies*, vol.38, pp. 1–30

[4] Evans, G. 1979, 'Reference and contingency', *The Monist*, pp. 161–189

[5] Jackson, F. 2000, *From Metaphysics to Ethics*, Oxford University Press

[6] Kaplan, D. 1989, 'Demonstratives', in J. Almog, J. Perry, and H.Wettstein (eds.), *Themes from Kaplan*, Oxford University Press

[15] I wish to thank Alex Byrne, Oscar Cabaco Olmedo, Phillip Keller, Bob Stalnaker, Steve Yablo and an anonymous referee for helpful discussion and comments on earlier drafts.

[7] Stalnaker, R. 2001 'Considering a world as actual', *Proceedings of the Aristotelian Society*, Supplement

[8] Stephanou, Y. 2000, 'Necessary beings', *Analysis* 60, pp. 188–93