It Is Not the Case that [P and 'It Is Not the Case that P' Is True][*]

Elia Zardini

LOGOS, Logic, Language and Cognition Research Group

Department of Logic, History and Philosophy of Science

University of Barcelona

elia.zardini@ub.edu

April 28, 2013

1 A New Semantic Paradox

Heck [2012] has recently offered a new *semantic paradox*, which he suspects forces any reasonable theory of truth to give up truth-theoretic principles that are at least as intuitively compelling as the famous schema:

(T) P iff 'P' is true.

In particular, Heck asks us to consider the intuitively weaker schemas:

(H_0) It is not the case that [P and 'It is not the case that P' is true];[1]

[*]An earlier version of the material in this paper has been presented in 2012 at the LOGOS Semantic Paradoxes and Vagueness Seminar (University of Barcelona). Thanks to the members of the seminar and to two anonymous referees for very stimulating discussions and comments. At different stages during the writing of the paper, I have benefitted from an AHRC Postdoctoral Research Fellowship and from the FP7 Marie Curie Intra-European Research Fellowship 301493 on *A Non-Contractive Theory of Naive Semantic Properties: Logical Developments and Metaphysical Foundations* (NTNSP), as well as from partial funds from the project FFI2012-35026 of the Spanish Ministry of Economy and Competition on *The Makings of Truth: Nature, Extent, and Applications of Truthmaking*, from the project FFI2011-25626 of the Spanish Ministry of Science and Innovation on *Reference, Self-Reference and Empirical Data*, from the project CONSOLIDER-INGENIO 2010 CSD2009-00056 of the Spanish Ministry of Science and Innovation on *Philosophy of Perspectival Thoughts and Facts* (PERSP) and from the FP7 Marie Curie Initial Training Network 238128 on *Perspectival Thoughts and Facts* (PETAF).

[1]Throughout, I'll use square brackets to disambiguate constituent structure in English.

(H_1) It is not the case that [P and 'P' is not true],[2]

and offers a paradox that derives a contradiction from an instance of (H_0) and an instance of (H_1) jointly.

Let T be an object-language truth predicate, and suppose that the language is expressive enough as to contain, for every sentence φ, a canonical name $\ulcorner\varphi\urcorner$ for φ. Suppose also that the language contains a name $\mathfrak{l}$ for the sentence $\neg T(\mathfrak{l})$ (so that $\mathfrak{l} = \ulcorner\neg T(\mathfrak{l})\urcorner$ holds). Let $\mathrm{W}^{\&}$ be the rule $\neg(\varphi \,\&\, \varphi) \vdash \neg\varphi$, which is valid in almost all theories of truth (since almost all these theories validate the full intersubstitutability of φ with $\varphi \,\&\, \varphi$). We can then reason thus:

$$\dfrac{\dfrac{\dfrac{}{\neg(T(\mathfrak{l}) \,\&\, T(\ulcorner\neg T(\mathfrak{l})\urcorner)) \;\vdash\; \neg(T(\mathfrak{l}) \,\&\, T(\ulcorner\neg T(\mathfrak{l})\urcorner))}\,\text{reflexivity}}{\neg(T(\mathfrak{l}) \,\&\, T(\ulcorner\neg T(\mathfrak{l})\urcorner)) \;\vdash\; \neg(T(\mathfrak{l}) \,\&\, T(\mathfrak{l}))}\,\text{indiscernibility of identicals} \qquad \dfrac{}{\neg(T(\mathfrak{l}) \,\&\, T(\mathfrak{l})) \;\vdash\; \neg T(\mathfrak{l})}\,\mathrm{W}^{\&}}{\neg(T(\mathfrak{l}) \,\&\, T(\ulcorner\neg T(\mathfrak{l})\urcorner)) \;\vdash\; \neg T(\mathfrak{l})}\,\text{transitivity}$$

(call this derivation '$\mathfrak{H}_0$'), as well as thus:

$$\dfrac{\dfrac{\dfrac{}{\neg(\neg T(\mathfrak{l}) \,\&\, \neg T(\ulcorner\neg T(\mathfrak{l})\urcorner)) \;\vdash\; \neg(\neg T(\mathfrak{l}) \,\&\, \neg T(\ulcorner\neg T(\mathfrak{l})\urcorner))}\,\text{reflexivity}}{\neg(\neg T(\mathfrak{l}) \,\&\, \neg T(\ulcorner\neg T(\mathfrak{l})\urcorner)) \;\vdash\; \neg(\neg T(\mathfrak{l}) \,\&\, \neg T(\mathfrak{l}))}\,\text{indiscernibility of identicals} \qquad \dfrac{}{\neg(\neg T(\mathfrak{l}) \,\&\, \neg T(\mathfrak{l})) \;\vdash\; \neg\neg T(\mathfrak{l})}\,\mathrm{W}^{\&}}{\neg(\neg T(\mathfrak{l}) \,\&\, \neg T(\ulcorner\neg T(\mathfrak{l})\urcorner)) \;\vdash\; \neg\neg T(\mathfrak{l})}\,\text{transitivity}$$

(call this derivation '$\mathfrak{H}_1$'). $\mathfrak{H}_0$ shows that an instance of (H_0) entails $\neg T(\mathfrak{l})$ and $\mathfrak{H}_1$ shows that an instance of (H_1) entails $\neg T(\mathfrak{l})$'s contradictory $\neg\neg T(\mathfrak{l})$. So, if all of reflexivity, indiscernibility of identicals, $\mathrm{W}^{\&}$ and transitivity hold, on pain of contradiction either (H_0) or (H_1) has to go (call this whole paradox 'paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$').

Paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$ seems to have a *distinctive punch* since, as Heck [2012], pp. 38–39 points out, not only do the *logical* principles used in the paradox seem to be noticeably weaker than those employed in many other semantic paradoxes (for example, no use is made of the law of excluded middle or of *reductio ad absurdum*), but also the *truth-theoretic* principles (H_0) and (H_1) seem to be intuitively no stronger, and in fact intuitively even weaker, than (T) (which is the truth-theoretical principle usually appealed to in developing semantic paradoxes; see fn 9 for some further discussion). I'll discuss in section 4 which theoretical perspective in fact confirms the *prima facie* impression that paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$ has a distinctive punch—until then, I'll take this impression at face value as a reason for looking into which kinds of theories, if any, have the advantage of preserving (H_0) and (H_1) in the presence of paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$.

[2]For simplicity, I've stated (H_0) and (H_1) using as a means to refer to linguistic expressions one (quotation) that is common in natural language. But, clearly, the spirit of (H_0) and (H_1) does not depend on the specific availability of quotation as such a means. It'll thus be natural (and convenient in the application to languages in which quotation may not be available) henceforth to understand (H_0) and (H_1) as admitting the replacement of quotation by any canonical system available in the language in question for referring to linguistic expressions.

In this respect, Heck [2012], p. 39 himself pessimistically concludes that "[. . .] there can be no consistent resolution of the semantic paradoxes that does not involve abandoning truth-theoretic principles that should be every bit as dear to our hearts as the T-scheme once was". As implicitly suggested by this remark of Heck's, one way out of the paradox would be to endorse some form of *dialetheism*—the view, roughly, that some contradictions are true (see Priest [2006] for a well-known presentation).[3] However, I don't think that foes of dialetheism are doomed to reject either (H_0) or (H_1). In the rest of this paper, I'll first briefly sketch a provably consistent, non-dialetheic theory of truth (which, for reasons that will become obvious shortly, I'll call '**IKMT**') that is independently motivated. I'll then show that both (H_0) and (H_1) hold in **IKMT** while $W^{\&}$ fails in it. Since Heck is definitely right that virtually every other non-dialetheic theory must reject either (H_0) or (H_1) (indicative examples include the theories of Kripke [1975]; McGee [1991]; Gupta and Belnap [1993]; Brady [2006]; Field [2008]), insofar as one wishes to uphold the extreme intuitive compellingness of such principles the upshot will be *a new interesting argument in favour of* **IKMT** *over its (non-dialetheic) rivals*. I'll close by arguing that the force of the argument is *amplified* from a *correspondentist* perspective in the philosophy of truth (while possibly at least to some extent mitigated from a deflationist one).

2 A Non-Contractive Theory of Truth

In Zardini [2011], I first offered a provably consistent, non-dialetheic theory of truth (**IKMT**) whose main feature is the rejection of the unrestricted validity of the structural metarule of *contraction*, which allows one to go from $\Gamma, \varphi, \varphi \vdash \Delta$ to $\Gamma, \varphi \vdash \Delta$ and from $\Gamma \vdash \Delta, \varphi, \varphi$ to $\Gamma \vdash \Delta, \varphi$ (more informally, contraction allows one to "contract" multiple occurrences of a premise or conclusion into a single occurrence). For our purposes, the presentation of **IKMT** can be limited to its (multiple-conclusion) *sentential* fragment. This can be defined in sequent-calculus style as the smallest logic containing as axiom the *structural rule*:

$$\frac{}{\varphi \vdash_{\mathbf{IKMT}} \varphi}\ \mathrm{I}$$

and closed under the *structural metarules*:

$$\frac{\Gamma_0 \vdash_{\mathbf{IKMT}} \Delta}{\Gamma_0, \Gamma_1 \vdash_{\mathbf{IKMT}} \Delta}\ \text{K-L} \qquad \frac{\Gamma \vdash_{\mathbf{IKMT}} \Delta_0}{\Gamma \vdash_{\mathbf{IKMT}} \Delta_0, \Delta_1}\ \text{K-R}$$

[3] I take it that by '*consistent*' Heck means *non-dialetheic*; however, for our purposes, it'll be best to deviate from that usage and reserve 'consistent' for *non-trivial*, so that decent dialetheic theories of truth are also counted as "consistent".

$$\frac{\Gamma_0 \vdash_{\mathbf{IKMT}} \Delta_0, \varphi \qquad \Gamma_1, \varphi \vdash_{\mathbf{IKMT}} \Delta_1}{\Gamma_0, \Gamma_1 \vdash_{\mathbf{IKMT}} \Delta_0, \Delta_1}\,\text{S}$$

under the *operational metarules*:

$$\frac{\Gamma \vdash_{\mathbf{IKMT}} \Delta, \varphi}{\Gamma, \neg\varphi \vdash_{\mathbf{IKMT}} \Delta}\,\neg\text{-L} \qquad \frac{\Gamma, \varphi \vdash_{\mathbf{IKMT}} \Delta}{\Gamma \vdash_{\mathbf{IKMT}} \Delta, \neg\varphi}\,\neg\text{-R}$$

$$\frac{\Gamma, \varphi, \psi \vdash_{\mathbf{IKMT}} \Delta}{\Gamma, \varphi \,\&\, \psi \vdash_{\mathbf{IKMT}} \Delta}\,\&\text{-L} \qquad \frac{\Gamma_0 \vdash_{\mathbf{IKMT}} \Delta_0, \varphi \qquad \Gamma_1 \vdash_{\mathbf{IKMT}} \Delta_1, \psi}{\Gamma_0, \Gamma_1 \vdash_{\mathbf{IKMT}} \Delta_0, \Delta_1, \varphi \,\&\, \psi}\,\&\text{-R}$$

$$\frac{\Gamma_0, \varphi \vdash_{\mathbf{IKMT}} \Delta_0 \qquad \Gamma_1, \psi \vdash_{\mathbf{IKMT}} \Delta_1}{\Gamma_0, \Gamma_1, \varphi \vee \psi \vdash_{\mathbf{IKMT}} \Delta_0, \Delta_1}\,\vee\text{-L} \qquad \frac{\Gamma \vdash_{\mathbf{IKMT}} \Delta, \varphi, \psi}{\Gamma \vdash_{\mathbf{IKMT}} \Delta, \varphi \vee \psi}\,\vee\text{-R}$$

$$\frac{\Gamma_0 \vdash_{\mathbf{IKMT}} \Delta_0, \varphi \qquad \Gamma_1, \psi \vdash_{\mathbf{IKMT}} \Delta_1}{\Gamma_0, \Gamma_1, \varphi \to \psi \vdash_{\mathbf{IKMT}} \Delta_0, \Delta_1}\,\to\text{-L} \qquad \frac{\Gamma, \varphi \vdash_{\mathbf{IKMT}} \Delta, \psi}{\Gamma \vdash_{\mathbf{IKMT}} \Delta, \varphi \to \psi}\,\to\text{-R}$$

and under the *truth-theoretic metarules*:

$$\frac{\Gamma, \varphi \vdash_{\mathbf{IKMT}} \Delta}{\Gamma, T(\ulcorner\varphi\urcorner) \vdash_{\mathbf{IKMT}} \Delta}\,T\text{-L} \qquad \frac{\Gamma \vdash_{\mathbf{IKMT}} \Delta, \varphi}{\Gamma \vdash_{\mathbf{IKMT}} \Delta, T(\ulcorner\varphi\urcorner)}\,T\text{-R}$$

(crucially, since contraction fails in **IKMT**, the collections on the left and right of $\vdash_{\mathbf{IKMT}}$ should be thought of as *multisets* rather than as *sets*—that is, as collections that are sensitive to the number of times with which an element occurs in them).

In Zardini [2011]; [2013a]; [2013b]; [2013c], I go through the background-logical and truth-theoretic details of **IKMT** (while proving its consistency *via* a cut-elimination argument), especially in the surprisingly many respects of philosophically interesting strength in which it outperforms other non-classical theories of truth, and I also sketch a metaphysical picture of the states-of-affairs expressed by paradoxical sentences that promises to make philosophical sense of their failure to contract. I refer the interested reader to those works for an in-depth investigation of **IKMT**, and pause here only to make explicit the properties of **IKMT** that are relevant for our purposes. **IKMT**'s theory of truth is *transparent*, i.e. it validates the full intersubstitutability of φ with $T(\ulcorner\varphi\urcorner)$. Moreover, by I, ¬-L, &-L and ¬-R, **IKMT**'s background logic validates the law of *non-contradiction* $\vdash_{\mathbf{IKMT}} \neg(\varphi \,\&\, \neg\varphi)$.[4] Now, it's easy to see that transparency and the law

[4] This fact indicates that claims like "it [...] seems safe to say that no theory that licenses transparency—or, again, even the one direction of it—can validate the law of noncontradiction" (Heck [2012], p. 38), while accurately describing virtually every other non-dialetheic transparent theory of truth, call for some adjustment given the existence of **IKMT**. (Notice that the restriction to non-dialetheic theories is required since, as *aficionados* will know, for better or worse the most influential dialetheic theories actually validate the "law of non-contradiction".)

of non-contradiction jointly entail (H_0) and (H_1). Hence, these principles hold in **IKMT** (and, since, by I and $\rightarrow$-R, $\vdash_{\mathbf{IKMT}} \varphi \rightarrow \varphi$ holds, by transparency (T) holds too).[5]

3 Solving the New Semantic Paradox

IKMT is non-dialetheic, and it validates both (H_0) and (H_1). So how does it block paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$? In fact, **IKMT** also validates reflexivity, the specific version of the indiscernibility of identicals used in the paradox (at least in **IKMT**'s extension to identity, which, to avoid unnecessary complexity, I'm not going into in this paper) and transitivity. Unsurprisingly, however, **IKMT** *does not validate* $W^{\&}$. Indeed, in the context of **IKMT**, $W^{\&}$ implies contraction. This can be seen in two steps. Firstly, it's easy to see that, by ¬-L, ¬-R, I and S, $\neg(\varphi \,\&\, \varphi) \vdash_{\mathbf{IKMT}} \neg\varphi$ holding is equivalent with $\varphi \vdash_{\mathbf{IKMT}} \varphi \,\&\, \varphi$ holding. Secondly, suppose that $\Gamma, \varphi, \varphi \vdash_{\mathbf{IKMT}} \Delta$ holds. Then, by &-L, $\Gamma, \varphi \,\&\, \varphi \vdash_{\mathbf{IKMT}} \Delta$ holds, and hence, if $\varphi \vdash_{\mathbf{IKMT}} \varphi \,\&\, \varphi$ also held, by S $\Gamma, \varphi \vdash_{\mathbf{IKMT}} \Delta$ would hold, which would preclude the possibility of failure of contraction for $\Gamma, \varphi, \varphi \vdash_{\mathbf{IKMT}} \Delta$. Contraposing, failure of contraction for $\Gamma, \varphi, \varphi \vdash_{\mathbf{IKMT}} \Delta$ forces failure of $\varphi \vdash_{\mathbf{IKMT}} \varphi \,\&\, \varphi$, and hence failure of $\neg(\varphi \,\&\, \varphi) \vdash_{\mathbf{IKMT}} \neg\varphi$. Thus, if we reject contraction, we have a principled reason for rejecting $W^{\&}$, and so we solve paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$.

It is worth noting that, while it is not surprising that, *given its specifics*, **IKMT** does not validate $W^{\&}$, it is not the case that, *in general*, *merely because of failure of contraction* a theory of truth can offer the same solution to paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$. For example, a non-contractive transparent theory **IKAT** alternative to **IKMT** diverges from it only with respect to conjunction and disjunction, having them obey the metarules:

$$\frac{\Gamma, \varphi \vdash_{\mathbf{IKAT}} \Delta}{\Gamma, \varphi \,\&\, \psi \vdash_{\mathbf{IKAT}} \Delta}\,\&\text{-L}^0_{\mathrm{A}} \qquad \frac{\Gamma, \psi \vdash_{\mathbf{IKAT}} \Delta}{\Gamma, \varphi \,\&\, \psi \vdash_{\mathbf{IKAT}} \Delta}\,\&\text{-L}^1_{\mathrm{A}} \qquad \frac{\Gamma \vdash_{\mathbf{IKAT}} \Delta, \varphi \qquad \Gamma \vdash_{\mathbf{IKAT}} \Delta, \psi}{\Gamma \vdash_{\mathbf{IKAT}} \Delta, \varphi \,\&\, \psi}\,\&\text{-R}_{\mathrm{A}}$$

$$\frac{\Gamma, \varphi \vdash_{\mathbf{IKAT}} \Delta \qquad \Gamma, \psi \vdash_{\mathbf{IKAT}} \Delta}{\Gamma, \varphi \vee \psi \vdash_{\mathbf{IKAT}} \Delta}\,\vee\text{-L}_{\mathrm{A}} \qquad \frac{\Gamma \vdash_{\mathbf{IKAT}} \Delta, \varphi}{\Gamma \vdash_{\mathbf{IKAT}} \Delta, \varphi \vee \psi}\,\vee\text{-R}^0_{\mathrm{A}} \qquad \frac{\Gamma \vdash_{\mathbf{IKAT}} \Delta, \psi}{\Gamma \vdash_{\mathbf{IKAT}} \Delta, \varphi \vee \psi}\,\vee\text{-R}^1_{\mathrm{A}}$$

(in technical parlance, just as the background logic of **IKMT** is the *multiplicative* fragment of sentential *affine* logics, the background logic of **IKAT** is the *additive* fragment of those logics). It's easy to see that, by I and &-R_A, $\varphi \vdash_{\mathbf{IKAT}} \varphi \,\&\, \varphi$ holds, and hence, by ¬-L and ¬-R, $\neg(\varphi \,\&\, \varphi) \vdash_{\mathbf{IKAT}} \neg\varphi$ holds too (as for the argument at the beginning of this section, contraction can still fail in **IKAT** because the **IKAT**-analogue of &-L fails).

[5]Indeed, given that the full De Morgan laws hold in **IKMT**, the law of non-contradiction is equivalent in **IKMT** with the law of *excluded middle* $\vdash_{\mathbf{IKMT}} \varphi \vee \neg\varphi$, which, by transparency, yields the law of *bivalence* $\vdash_{\mathbf{IKMT}} T(\ulcorner\varphi\urcorner) \vee T(\ulcorner\neg\varphi\urcorner)$. This fact indicates that claims like "[...] all theories known to me that are committed to transparency reject bivalence [...]" (Heck [2012], p. 37), while accurately describing virtually every other non-dialetheic transparent theory of truth, call for some adjustment given the existence of **IKMT**.

Thus, **IKAT** cannot solve paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$ by rejecting $\mathrm{W}^{\&}$. In fact, **IKAT** has in this respect no advantage over other non-dialetheic theories, as it too is forced to reject either (H_0) or (H_1) (in fact, both).[6] In addition to the reasons I've adduced in Zardini [2011]; [2013b], I regard the different treatment of paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$ given by **IKMT** and **IKAT** as a new interesting reason for preferring the former over the latter.[7]

Although thus motivated by failure of contraction and the eminently plausible principle about conjunction &-L, failure of $\mathrm{W}^{\&}$ is open to an important worry. Let's focus on the equivalent but simpler $\varphi \vdash_{\mathbf{IKMT}} \varphi \,\&\, \varphi$. Does the fact that $\varphi \vdash_{\mathbf{IKMT}} \varphi \,\&\, \varphi$ fails not show that conjunction in **IKMT** is objectionably *not extensional*? It would be disappointing if the proposed solution to paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$ relied on understanding conjunction in a rather deviant, not recognisably extensional way. This worry can be addressed by observing that conjunction in **IKMT** does satisfy both the rule of *adjunction* ($\varphi, \psi \vdash_{\mathbf{IKMT}} \varphi \,\&\, \psi$) and the rules of *simplification* ($\varphi \,\&\, \psi \vdash_{\mathbf{IKMT}} \varphi$ and $\varphi \,\&\, \psi \vdash_{\mathbf{IKMT}} \psi$). Satisfaction of these rules[8] arguably shows that, in **IKMT**, *the truth of a conjunction is nothing over and above the truth of both of its conjuncts*—a principle which I take to be what the extensionality of conjunction amounts to. Indeed, since **IKMT** is a theory of truth, it is no surprise that it itself proves (a reasonable formal version of) such principle: $\vdash_{\mathbf{IKMT}} ((T(\ulcorner\varphi\urcorner) \,\&\, T(\ulcorner\psi\urcorner)) \rightarrow T(\ulcorner\varphi \,\&\, \psi\urcorner)) \,\&\, (T(\ulcorner\varphi \,\&\, \psi\urcorner) \rightarrow (T(\ulcorner\varphi\urcorner) \,\&\, T(\ulcorner\psi\urcorner)))$ holds. Thus, we can coherently maintain that $\varphi \vdash_{\mathbf{IKMT}} \varphi \,\&\, \varphi$ and $\mathrm{W}^{\&}$ fail while also maintaining that the truth of a conjunction is nothing over and above the truth of both of its conjuncts. Our solution to paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$ is consistent with the extensionality of conjunction.

4 Truth, Correspondence, Non-Divergence and Non-Contradiction

Among non-dialetheic theories of truth, **IKMT** thus has the virtually unique feature of solving paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$ by rejecting its reasoning rather than by rejecting (H_0) or (H_1). The obvious question confronting us now is to what extent this circumstance can be taken to be an advantage for **IKMT**. As I've mentioned in section 1, from an *intuitive* standpoint (H_0) and (H_1) are extremely compelling, and it is therefore an undeniable cost of virtually every other non-dialetheic theory that it so clashes with intuition.[9] In the

[6] For essentially the same reason, the law of non-contradiction, the law of excluded middle and the law of bivalence all fail in **IKAT**, contrary to what happens in **IKMT** (see fns 4 and 5).

[7] Thanks to an anonymous referee for comments that prompted this paragraph.

[8] Actually, the rules of simplification strictly speaking only say that a conjunction entails the truth of *either* conjunct, and this, at least in a non-contractive framework, may crucially fall short of entailing the truth of *both* conjuncts (to get a rough feel for the gap here, compare with the contrast between 'Show either ticket or passport' and 'Show both ticket and passport'). Fortunately, **IKMT** also validates other principles like &-L which arguably show that the truth of a conjunction entails the truth not only of either, but also of both conjuncts.

[9] (H_0) and (H_1) certainly bear some similarity to the law of non-contradiction (in fact, given transparency, each is equivalent with it), and that law is rejected at least by standard non-dialetheic transparent theories of truth. One may then wonder whether, *from an intuitive standpoint*, (H_0) and (H_1) are any

context of the semantic paradoxes, considerations of intuitive compellingness can however hardly be taken to be conclusive. In fact, if we ascend to a more *theoretical* standpoint, the situation concerning the standing of (H_0) and (H_1) is more nuanced. More in detail, I'm going to argue that, while the adoption of a deflationist perspective in the traditional debate about the *nature of truth* does not seem to provide any theoretical reason for accepting (H_0) or (H_1) additional to those—if there are any—for accepting the law of non-contradiction, the adoption of a correspondentist perspective in that debate does provide a cogent theoretical reason for accepting (H_0) and (H_1) that does not depend on reasons for accepting the law of non-contradiction.[10]

From a *deflationist* perspective, the *raison d'être* of the concept of truth consists in the expressive resources it makes available (as manifested, for example, in so-called "blind generalisations" like 'Something Richard said is true'). From that perspective, the fundamental principle that a truth predicate is required to obey is arguably transparency, for it is transparency that allows the truth predicate adequately to serve the expressive function identified by the deflationist (as argued e.g. by Field [2008], pp. 205–210). Given extremely minimal assumptions about the implication expressed by 'if', transparency entails (T) and, given some other at least natural assumptions about that implication, it is entailed by it.

However, it is unclear whether, from a deflationist perspective, for all their intuitive compellingness there are any cogent theoretical reasons for also including either (H_0) or (H_1) in one's theory of truth. On the one hand, no deflationistically acceptable consideration about truth seems to speak directly to that issue. For the connection between semantic and non-semantic facts required by deflationism is that ''P' is true' be *in some strong sense equivalent* with 'P' (for that is what allows the truth predicate adequately to serve the expressive function identified by the deflationist), and it is transparency which guarantees that—(H_0) *and* (H_1) *do not seem to articulate any additional insight into the*

more compelling than that law. I think that they are, in the sense that there is at least one respect in which the intuitive support for (H_0) and (H_1) does not simply reduce to the intuitive support for the law of non-contradiction. For, on the face of it, (H_0) and (H_1) *merely deny certain intuitively aberrant combinations of polar status* for 'P', ''It is not the case that P' is true' and ''P' is true'. Now, as Heck [2012], p. 39 himself suggests, merely *denying such combinations* is intuitively no stronger, and is in fact intuitively even weaker, than going so far as to *assert the perfect correlation of status* described in (T), especially if the conditional in (T) is read *non-materially* (and it is in fact so read both naturally and by virtually all theories other than **IKMT** that accept (T), including standard non-dialetheic transparent theories). Since (H_0) and (H_1) are thus intuitively even weaker than (T), and since already (T) is intuitively compelling independently of the law of non-contradiction, so are *a fortiori* (H_0) and (H_1). (Notice that, if, contrary to standard non-dialetheic transparent theories, one does envisage the possibility of true contradictions, it becomes less clear that (H_0) is intuitively no stronger than (T). But even in that case it remains clear that (H_1) is intuitively weaker than (T) and so intuitively compelling independently of the law of non-contradiction, and on many theories, including standard non-dialetheic transparent theories, (H_1) is sufficient to generate a variant of paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$.) Having noted all this from an *intuitive* standpoint, in this section I'll rather focus on how (H_0) and (H_1) enjoy remarkable support independent of the law of non-contradiction (also) from a *theoretical* standpoint in the philosophy of truth. Thanks to an anonymous referee for raising this issue.

[10]Thanks to two anonymous referees for comments that led me to investigate the import of paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$ in the wider context of the philosophy of truth.

concept of truth as understood by deflationism. On the other hand, it is true that, given transparency, both (H_0) and (H_1) are equivalent with the law of non-contradiction, so that theoretical reasons for accepting the latter would yield derivative theoretical reasons for accepting the former too. However, rightly or wrongly, a non-dialetheic deflationist that revises classical logic in order to accommodate for transparency is likely to assume that she has good theoretical reasons for rejecting the law of non-contradiction (although this is not the focus of the paper, I air some doubts about this assumption in fn 13). Thus, in general, it remains unclear whether a deflationist has cogent theoretical reasons for accepting (H_0) or (H_1); more specifically, it seems that such reasons—if there are any—will have to be derivative from theoretical reasons for accepting the law of non-contradiction—the deflationist conception of truth does not seem to provide any additional theoretical reason for accepting (H_0) or (H_1).

From a *correspondentist* perspective, however, things look quite different. For the connection between semantic and non-semantic facts required by correspondentism is more varied than that required by deflationism and does in fact include among its components the specific connections articulated by (H_0) and (H_1)—or so I shall argue. Correspondentism holds that truth is a substantial property of sentences (and possibly of other truth bearers), and identifies such property with *correspondence with reality.* Let's assume a rather light-weight and simplified understanding of the latter notion, to the effect that a sentence corresponds with reality iff what it says is the case. Let's also assume that what 'P' says is that P. It follows from these two assumptions that 'P' corresponds with reality iff P, which, by the identification of truth with correspondence with reality, in turn yields (T) (a principle which, as we've seen, is also endorsed on very different grounds by deflationism, for it follows from 'P iff P' and the deflationist claim that ''P' is true' is in some strong sense equivalent with 'P').

But the identification of truth with correspondence with reality arguably licences also a variety of other connections between semantic and non-semantic facts, *including the connections articulated by* (H_0) *and* (H_1). Here is the argument. First, we can reformulate the correspondentist idea that a sentence is true iff what it says is the case as the idea that the property of being true *converges* with reality: it is exemplified by a sentence iff what the sentence says is the case. Next, reflect that the opposite of convergence is *divergence*, and that there are two prominent ways in which a property F of sentences may diverge from reality. Keeping fixed our assumption from the last paragraph that what 'P' says is that P, it may be that:

(D_0) It is not the case that P, but 'P' is F,

in which case the property F diverges from reality in that it "approves" a sentence that does not correspond with reality (or, in other words, it diverges from reality in that it is *unsound* with respect to correspondence with reality). Or, conversely, it may be that:

(D_1) It is the case that P, but 'P' is not F, in which case the property F diverges from reality in that it "disapproves" a sentence that does correspond with reality (or, in other words, it diverges from reality in that it is *incomplete* with respect to correspondence with reality). Now, the identification of truth with correspondence with reality arguably requires the property of being true not to diverge from reality, for that identification requires the property of being true to converge with reality, and a property can only converge with reality if it does not diverge from it. Hence, *a fortiori*, the identification of truth with correspondence with reality requires the property of being true not to diverge from reality in either the (D_0)-way or the (D_1)-way. But that the property of being true does not diverge from reality in either of these ways is precisely what (H_0) (*modulo* the equivalence of a sentence with its double negation) and (H_1) jointly amount to.[11] Hence, while from a deflationist perspective (H_0) and (H_1) *are just a long-winded restatement of the law of non-contradiction*—and should thus be accepted *only because of reasons for accepting the latter*—from a correspondentist perspective they *articulate a specific insight into the nature of truth as correspondence with reality*—and should thus be accepted *independently of reasons for accepting the law of non-contradiction*.[12]

[11] A related argument in favour of (H_0) and (H_1) starts by observing that, together with the by now usual assumption that what 'P' says is that P, they follow from the contrapositives of:

(R_0) If it is not the case [that P] and φ is true, φ does not say that P;

(R_1) If P and φ is not true, φ does not say that P

respectively (*modulo* the equivalence of a sentence with its double negation). Now, in addition to being themselves intuitively compelling, (R_0) and (R_1) are arguably a straightforward consequence of the more informal claims that *true sentences do not say what is not the case* and that *untrue sentences do not say what is the case* respectively. And, given the understanding of the notion of correspondence with reality in terms of saying what is the case, those claims in turn immediately follow from the identification of truth with correspondence with reality, and in particular from the principle that *true sentences are those that correspond with reality* (with the additional assumptions, to get to the first claim, that sentences that say what is the case do not say what is not the case, and, to get to the second claim, that, if true sentences are those that correspond with reality, untrue sentences do not so correspond). One can also derive (R_0) and (R_1) by *antilogism* from two other related principles that similarly spell out aspects of the identification of truth with correspondence with reality:

(C_0) If it is not the case [that P] and φ says that P, φ is not true;

(C_1) If P and φ says that P, φ is true

respectively (*modulo* the equivalence of a sentence with its double negation). Notice, that while antilogism is not unrestrictedly valid at least in standard transparent theories of truth, the two applications needed to go from (C_0) and (C_1) to (R_0) and (R_1) respectively would seem unexceptionable. Thanks to Sven Rosenkranz for drawing my attention to (R_0) and (R_1).

[12] I suppose that a correspondentist friend of some non-dialetheic rival of **IKMT** may reply by saying that a claim that can still be included in her theory of truth is that *it is not determinate* that the property of being true diverges from reality. That may well be correct, but it of course falls dramatically short of what is licensed by the identification of truth with correspondence with reality (and by the idea that the property of being true converges with reality). And the significance of such a weak claim is further problematised by the fact that, on the usual ways of introducing a notion of determinacy into a theory of truth, the reply under consideration will be forced also to accept the rather disconcerting claim that it is not determinate that the property of being true does not diverge from reality. Analogous comments

I stress that the argument in the last paragraph in favour of (H_0) and (H_1) (and so, given paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$, in favour of **IKMT**) only relies on the adoption of a correspondentist perspective and nowhere presupposes the law of non-contradiction. The argument has however an obvious import on that law which should now be made explicit. There has always been a sense of theoretical unease about the fact that standard non-dialetheic transparent theories of truth (such as those of Kripke [1975]; Brady [2006]; Field [2008]) fail to validate the law of non-contradiction (which, as Heck [2012], p. 37 shows, leads to contradiction by a reasoning very similar to that of paradox $\mathfrak{H}_0$–$\mathfrak{H}_1$). One rather obvious theoretical reason in favour of that law is that it would seem *the by far most natural expression of non-dialetheism*. What I would now like to point out is that, given transparency, the argument in the last paragraph in favour of (H_0) and (H_1) yields another theoretical reason in favour of the law. For, given transparency, the law is *entailed* by either of (H_0) and (H_1); since, as I've been at pains to stress, the argument in the last paragraph yields a theoretical reason in favour of (H_0) and (H_1) *without presupposing* the law of non-contradiction, it would also seem to yield a *derivative* theoretical reason in favour of the law itself.[13] Interestingly, this very last point indicates that, from a correspondentist perspective, the relationship between the *truth-theoretic* principles (H_0) and (H_1) on the one hand and the *logical* law of non-contradiction on the other hand is in at least some respect *reversed* in comparison to their relationship from a deflationist perspective: from a correspondentist perspective, the independent theoretical reason unearthed in the last paragraph for accepting (H_0) and (H_1) yields, by transparency, a derivative theoretical reason for accepting the law of non-contradiction too. Correspondence requires non-divergence, and this in turn requires (by transparency) non-contradiction.

Summing up the considerations in this section, the intuitive support for (H_0) and (H_1) is substantially amplified by theoretical considerations made available by the adoption of a correspondentist perspective (while possibly at least to some extent mitigated if one assumes a deflationist perspective). In these respects, the fact that almost all non-dialetheic

apply to the reply that a claim that can still be included in one's theory is that 'Truth diverges from reality' *leads to absurdity* and to the reply that one can still *reject* that truth diverges from reality. These last two replies are also subject to the further charges that, in the first case, the envisaged claim is (a relatively uninteresting) one about the *sentence* 'Truth diverges from reality' rather than about *truth and reality themselves*, and that, in the second case, no new claim at all is included in one's *theory* as the only (relatively uninteresting) addition is rather made at the *psychological* level of what the *theorist* rejects.

[13]No doubt some friends of standard non-dialetheic transparent theories of truth have already learnt to live with the failure of the law of non-contradiction, sweetening the pill by observing (as in fn 5) that, given the intuitionistically unacceptable De Morgan law, the former law entails the law of excluded middle, which is rejected on those theories. For what's worth, I've always been a little sceptical about taking that observation as a good reason for rejecting the law of non-contradiction: it seems to me that, because of the two reasons given in the text in favour of that law, the observation is more naturally taken as an argument in favour of either *accepting the law of excluded middle* or *rejecting the intuitionistically unacceptable De Morgan law*, neither of which is done in standard non-dialetheic transparent theories (the first path would naturally lead to something along the lines of **IKMT** or, alternatively, to dialetheism; the second path would naturally lead to something along the lines of an intuitionist weakening of **IKMT**, which can be obtained, to give a simple example, by restricting $\neg$-R and $\rightarrow$-R to cases where the right-hand side of $\vdash$ has no conclusion or one conclusion respectively). Thanks to Sven Rosenkranz for discussion of these issues.

theories of truth are forced to reject either (H_0) or (H_1) points to a new interesting problem for these theories, while the fact that **IKMT** accepts both (H_0) and (H_1) (indeed, proves them) points to a new interesting advantage that **IKMT** enjoys over virtually every other theory in the non-dialetheic camp.

References

Ross Brady. *Universal Logic*. CSLI Publications, Stanford, 2006.

Hartry Field. *Saving Truth from Paradox*. Oxford University Press, Oxford, 2008.

Anil Gupta and Nuel Belnap. *The Revision Theory of Truth*. MIT Press, Cambridge MA, 1993.

Richard Heck. A liar paradox. *Thought*, 1:36–40, 2012.

Saul Kripke. Outline of a theory of truth. *The Journal of Philosophy*, 72:690–716, 1975.

Vann McGee. *Truth, Vagueness, and Paradox*. Hackett, Indianapolis, 1991.

Graham Priest. *In Contradiction*. Oxford University Press, Oxford, 2nd edition, 2006.

Elia Zardini. Truth without contra(di)ction. *The Review of Symbolic Logic*, 4:498–535, 2011.

Elia Zardini. Getting one for two, or the contractors' bad deal. Towards a unified solution to the semantic paradoxes. In Theodora Achourioti, Kentaro Fujimoto, Henri Galinon, and José Martínez, editors, *Unifying the Philosophy of Truth*. Springer, Berlin, 2013a. Forthcoming.

Elia Zardini. Naive *modus ponens*. *Journal of Philosophical Logic*, 2013b. Forthcoming.

Elia Zardini. Naive truth and naive logical properties. *The Review of Symbolic Logic*, 2013c. Forthcoming.