Sentences, strings, and truth[*]

Benj Hellie

February 17, 2013

A hallmark of the specialist literature on truth is a profusion of conflicting intricate hypotheses about the semantic properties of sentences like (1):

1. 'Snow is white' is true

The aim of this literature is to articulate those semantic properties: to make explicit what the sentence means. The intricacy and conflict in the literature suggests a widespread agreement that very little concerning those semantic properties is at all obvious upon superficial examination: that only the highly trained expert is in a position to make explicit what they are.

Why so? The alleged trouble stems from 'liar sentences':

2. (2) is not true

For surely the sentence (2) either is or isn't true. If (2) is true, then, as (2) says, it is not true. So then it both is and isn't true. On the other hand, if (2) is not true, then—well, that's what (2) says, so (2) is true. So then it both is and isn't true. Either way, then, (2) is true and (2) isn't true. Contradiction! But contradictions entail everything—and the world explodes.

The world, of course, has not exploded. Evidently, concludes the literature, we have misunderstood something about what (2) means; and presumably the problem is then that we have misunderstood what 'is true' means. Perhaps when truth is involved, 'fact gaps' are OK? Or perhaps when truth is involved, contradictions are OK? Or perhaps a sentence can say something that is so without being true? Or be true even though what it says is not so? Either way, something mysterious has gone on: cue the experts.

[*]Discussion with Jessica Wilson in the course of composing this paper was extremely helpful. A more long-range influence on the ideas here has been Peter Ludlow.

This mysterianism about the semantic properties of (1) and (2) stands in stark contrast with the literature's attitude toward the *orthographic* properties of these sentences: which sentences are constituted by the underlying strings; what the constituents of those sentences are; how they compose syntactically. These orthographic properties are assumed to be *luminous*, evident upon superficial examination of their constituting strings. That there is widespread agreement on this luminosity thesis is strongly suggested by the standardization of a technical notation for representing the orthography of these sentences.

The luminosity thesis is a crucial presupposition at the first stage of the liar paradox. For suppose that, despite superficial appearances, the string displayed in (2) does not in fact constitute a sentence. If there is no sentence (2), the statement with which the derivation of the paradoxical conclusion commences—'the sentence (2) either is or isn't true'—is semantically defective: falsely presupposing the existence of a topic, and therefore not a legitimate starting point for further argumentation.

In short: if the orthographic properties of the string displayed in (2) are not luminous, that string might fail to constitute a sentence; if there is no sentence, there is no liar paradox; and without a liar paradox, the semantic properties of a sentence like (1) might after all turn out to be more or less what we would have initially thought.

This article provides a theory of the orthographic properties of sentences like (1) according to which these properties are typically *not* luminous. I call it the *String-to-Sentence Theory of Truth*—the 'String Theory', for short—in order to highlight the theory's use of the crucial distinction between *strings*, which are linearly structured sequences of glyphs, and *sentences*, which are syntactically structured entities composed of words and morphemes and constituted by strings. Strings sometimes but not always constitute sentences; and when a string does constitute a sentence, which sentence is so constituted is not always obvious. According to String Theory, when a string apparently constitutes a sentence containing the truth-predicate, which sentence is in fact constituted is quite frequently not obvious at all; and in the case of (2) and other liar sentences, no sentence is

constituted. And as a result, there is no need for a distinctive semantics of truth.

Plan: section 1 discusses a range of examples to undermine the luminosity of orthography; section 2 presents String Theory; section 3 applies the theory to a range of nonparadoxical truth-ascriptions; section 4 discusses the liar paradox; section 5 concludes.

1 Orthographic externalism

What is the source of our customary asymetry in attitudes toward semantics and orthography? Perhaps it appears sustainable because, while discussion of 'externalist' doctrines about semantic properties has been philosophical meat and potatoes for decades, there has been no comparable attention to externalism about orthographic properties. This can make orthographic externalism easy to ignore. But we shouldn't ignore it. Once we start looking for cases in which the user of a string is not in a position to make its orthographic properties fully explicit, they can seem to pop up everywhere.

(A) Bad handwriting

Dr. Cart scribbles on a prescription pad

- Rp.
  Pentobarbitali natrici 3
  Morphiae sulphas 2
  Chlorali hydrati 15
  Saccharum ad 50
  M.f.plv.
  Div. in doses aeq. No XXX (triginta)
  D.S. For sleep: one sachet to be taken at bedtime

Fred has no idea how to decode Dr. Cart's handwriting. In that sense, Fred has no idea which expressions are written on the prescription pad. And

yet, in Fred's view, those expressions are the description of the medicine he wants.

Fred shows the prescription to Pharmacist Sessler, saying 'make me some of this stuff, please'. The stuff in question is the medication Dr. Cart described. If, as is plausible, Fred manages to refer to the stuff, he has quite plausibly used the words Dr. Cart wrote down to do so. If so, Fred has used words he is not in a position to make at all explicit—say, by reading them aloud or copying them over.

(B) Semaphor

As the ship enters the dock, the captain orders Fred to the bow with instructions to wave the flags stacked up on the deck in the order in which they are stacked. Fred is color-blind, and does not understand the semaphor code anyway, but does as commanded. The flags mean 'I am carrying, loading, or discharging dangerous goods' and 'this ship will be in quarantine for 17 days'. Perhaps Fred has asserted these sentences. If so, he will nevertheless be quite dismayed once he learns what he has communicated!

(C) E-type anaphora

According to some (Evans 1977), the occurrence of 'it' in 'if you have a Labrador, it will follow you everywhere' simply *is* the expression 'that Labrador'. More generally, an 'E-type' a discourse-anaphoric pronoun disguises its orthographic character: it is not in fact a pronoun, as superficial appearance might suggest, but is rather a repetition of its antecedent. Suppose this is true. Suppose Fred enters the conversation late, hearing '... and then he started smiling broadly'. Fred cuts in: 'was he happy?' Sam responds: 'he certainly was'. If Fred's question was a successful discourse anaphor, his string 'was he happy' constitutes the sentence 'was Franklin Roosevelt happy?'—despite Fred's being in no position to make that fact explicit.

(D) Semantic individuation of expressions

Some natural deduction systems have the following 'identity-introduction' rule: $\vdash \nu = \nu$. Suppose that this rule genuinely is valid as a characterization of the meaning of ordinary language uses of '=' and singular terms. If so, any ordinary competent true assertion of 'John isn't John' (meaning, for example, that John McCain isn't John Howard) must involve occurrences of distinct names flanking the identity sign. What distinguishes them? According to some (Kaplan 1990), perhaps nothing but their distinct referents. But if so, when Peter asserts 'Paderiewski isn't Paderiewski', he asserts a sentence of form $\neg(\nu = \nu)$—the negation of a theorem. Because Peter isn't insane, he would not knowingly assert a sentence he explicitly knows to be the negation of a theorem. And so he must be unaware that that is what he has done—and therefore mistaken about which sentence he has asserted in producing its constituting string.

(E) Ditto marks

Mo and Ro attend the community round table. At the front are sign-in forms on which those at the meeting are expected to write down their contact information. Ro writes in her contact information first. Mo and Ro live at the same address, so after writing down his name, Mo writes ditto marks under Ro's inscriptions in the remaining boxes. Mo seems not to have disobeyed the instruction to write down his contact information. But if so, he has written down his address and his phone number. Because no address is spelled the same way as any phone number, the expressions Mo has written down are of different orthographic types. But that is not superficially apparent: the strings Mo wrote in the *phone* and *address* boxes are both just ditto marks.

(F) Big dittos

Mark from Michigan finally gets on the air with Rush. While Mark was on hold, he was busying himself with oiling his gun, and was not really paying much attention to the show. Rush begins the conversation with Mark: 'so, whaddaya think about these crazy fools, Mark?' Mark replies: 'hey, big

dittos over to you Rush!'

Later that afternoon at the gun shop, Mark's friend Bert says 'Mark, you were totally right on when you said that liberals should be assigned to corvee labor'. Mark replies 'I said that?' Bert explains: 'yeah, that was Rush's statement to which you gave big dittos'. Mark rejoins 'ah—yep, big dittos indeed!'

Bert seems to be thinking of Mark's act of big-dittoing was a way of re-asserting the sentence Rush had asserted. If Bert is correct in this—and he is not obviously mistaken—then in producing the strin 'big dittos over to you, Rush', Mark asserted 'liberals should be assigned to corvee labor' unknowingly.

(G) Schemata

Elizabeth writes down the following schema:

(Δ) Either it is determinate that φ or it is determinate that $\neg\varphi$.

In writing down (Δ), she intends to assert every sentence resulting from the substitution of the same declarative sentence of English into each occurrence of φ in (Δ). But there are plenty of sentences of English Elizabeth has never heard of and would never be in a position to explicitly pronounce. Examples might include:

(a) 'San Sebastian is a very very very ... very very tasty town', where the ellipsis is filled in with 10^{100} occurrences of 'very'

(b) 'Fred likes the taste of Ethyl 3-{[(2-{[(4-{N′-hexyloxycarbonyl carbamimidoyl} phenyl)amino]methyl}-1-methyl-1H-benzimidazol-5-yl)carbonyl] (pyridin-2- yl-amino)propanoate}'

So while Elizabeth asserts the sentences which are constituted by the substitution of the string (a) for φ in the string (Δ) and of the string (b) for φ in the string (Δ) instances of (Δ), she is not in a position to make explicit that she has done so.

(H) Lagadonian dancing qualia

According to some (Chalmers 2003), in '*this* is what it's like for me', the occurrence of 'this' is 'Lagadonian': partly constituted, and orthographically individuated, by that quale that is its referent. Suppose that Fred thinks he might be suffering from 'dancing qualia': though things are normal with him 'functionally', his qualia cycle back and forth at short but unpredictable intervals between 'normal' and 'inverted' qualia. If the structure of Fred's reasoning is secured at the functional level, then the Lagadonian theory predicts that Fred—looking at a red thing throughout—is uncertain whether his earlier and a later tokening of the string '*this* is what it's like for me' constitute the same sentence. (Compare Hawthorne 2006.)

(I) Ordinary quote marks

According to Quine, the expression 'the number of planets = 8' is a singular term with referent the string between its quote-marks. If so, most users of ordinary quote marks seem to be unaware of this orthographic fact: many of our students (and many of us!) find Quine's theory very hard to use. So if Quine is right, most users of ordinary quote marks are not in a position to make explicit which expressions are constituted by their uses of strings flanked by quote-marks.

(J) Chomskian syntax

Are the various disambiguations of the string 'John knows how many pictures of himself Bill took' readings of the *same sentence*? Chomskian syntacticians say no. This idea has spawned a rich and intricate research program in which the aim is to make explicit which expressions ordinary speakers use. Evidently Chomskian syntax presupposes that, if Chomskian syntax is difficult, ordinary speakers are typically not in a position to make explicit which expressions are constituted by the strings they produce. Chomskian syntax is difficult indeed. So, if the presupposition is correct, ordinary speakers are typically not in a position to make explicit which expressions are constituted by the strings they produce.

(K) A stipulation

> Ro stipulates that in her language, henceforth, the sound 'yep' will constitute whatever sentence was asserted five sentences prior in the conversation to which she is a party. Suppose that in a certain case, all she knows about the sentence asserted five sentences back is that Mo asserted it and Fred denied it. Because she trusts Mo and mistrusts Fred, Ro intervenes in the dispute on Mo's behalf by making the sound 'yep'. When she does so, by her convention, she asserts the sentence Mo asserted—despite her inability to make explicit which sentence is constituted by the sound she produces.
>
> Ro's convention has a use: a somewhat 'boutique' use is a use nevertheless. And what would prevent her from introducing this convention? In general nothing (or nothing but consistency, perhaps) prevents us from making any sort of *semantic* stipulation we might choose to make. I may stipulate that 'fnord' is to express the necessary proposition just if goats eat cans and the impossible proposition otherwise, may I not? If I may, why should the case of orthographic stipulation be any different?

Some of these examples are doubtless more compelling than others. But our aim with these examples is not to establish that there *is* any orthographic externalism. It is merely to establish that orthographic externalism is *comprehensible*. And the pervasive use of schemata and the popularity of Quine's theory of quotation in philosophy, as well as the very existence of the Chomskian research program, do seem strongly to suggest that orthographic externalism is comprehensible. We should therefore be quite skeptical about the luminosity thesis: perhaps it is not, after all, so obvious which sentences are constituted by strings like (1) and (2).

2 The String-to-Sentence Theory of Truth

The String Theory falls into the family of *deflationary* theories of truth (Field 1994).[1] According to competing *inflationary* theories, a sentence like (1) asserts

[1]Roughly, String Theory is an 'E-type' 'prosentential' theory (Grover *et al.* 1975): truth-sentences are anaphors, and anaphors are contextually saturated substitutional variables. Many

the holding of some relation between the sentence 'snow is white' and the world: perhaps that the sentence 'corresponds with the facts'. Deflationary theories deny this. As an alternative, deflationists state that (1) is, in effect, a roundabout way of expressing 'snow is white'. But this roundabout mode of expression is an artifact of a valuable practice within which we use the truth-predicate: to abbreviate, to be nonspecific, to quantify, to gesture at agreement with sentences past, distant, forgotten, unheard, in multitudes—this can be helpful. We wish to commit ourselves to re-expressing every sentence Carnap asserted, though in English rather than German. We may do so by saying 'everything Carnap said is true'. This is worth it—despite the cost of allowing us to say longwinded things like ' 'snow is white' is true'.

A range of broadly deflationary theories have been presented. Why another one? The advantages of String Theory are, or so I will argue, these: it is precisely specifiable; cleanly accommodates a very wide range of data; and dispels the liar paradox.

Here is the very rough idea: a truth-ascription is an instruction for the hearer to re-say the sentence under discussion for herself.[2]

Now a slightly less rough characterization. If you hear me produce (1), the sentence constituted for you by the sounds I produce is whatever sentence *you* use to mean the same thing as the sentence *I* refer to as 'snow is white'. So if you use 'snow is white' to mean the same thing as that sentence, my sounds constitute the sentence 'snow is white' for you; if you use 'la neige est blanc' to mean the same thing as that sentence, my sounds constitute the sentence 'la neige est blanc' for you; if you use 'schnee ist weiss' to mean the same thing, my sounds constitute 'schnee ist weiss' for you; and so forth.

The affinity to deflationism should be obvious upon this rough statement. There can be no distinctive word-world relation described by (1), because (1) does

theorists regard deflationism as committed to the doctrine that 'all we know about' truth is captured in the validity of the T-schema. As we will see below, String Theory regards the T-schema as not universally valid, and derives the T-schema (in those cases in which it is valid) from deeper principles.

[2]Thanks to Jessica Wilson for this formulation.

not contain any words that pertain to any word-world relation. Whenever anyone finds words in the string (1), those words pertain solely to snow and whiteness.

Let us now build up to a sharp statement of String Theory. I proceed in stages, because the situation generates ample opportunities for confusion.

The heart of String Theory is an interaction among three (not necessarily distinct) language-users: the user of the *object* language; the user of the *meta*language; and the *assessor* of the metalanguage. When you hear me produce (1), you are the assessor and I am the user of the metalanguage. Who is the user of the object language? Whoever that person is whose sentence I refer to as 'snow as white': perhaps that person is me; perhaps that person is a 'typical' speaker of our common language; perhaps that person is Fred, about whom we are gossiping. Note that language-users will be understood as *temporally local stages*, rather than 'whole life' persons.

With the three language-users come three linguistic objects: Fred's sentence 'snow is white'; whatever sentence you use to mean the same as Fred's sentence—if you are a monolingual English speaker, that sentence would be 'snow is white'; and my string (1). Adding to the complexity is the fact that my string contains a singular term, the quote-name ' 'snow is white' '—which, confusingly enough, closely resembles Fred's sentence.

Next, String Theory embeds a distinction between a pair of varieties of linguistic object: the strings and the sentences. Strings, to reiterate, *constitute* sentences. Sentences are the primary bearers of orthographic properties; we may speak of the orthographic properties of a string, when it is safe to do so, and mean by this the orthographic properties of the sentence *constituted* by the string—but in many cases it will not be safe to do so.

Next, String Theory individuates expressions (including sentences) in terms of the languages to which they belong: the word 'gift' in Carnap's language is distinct from the word 'gift' in Anscombe's language. We will moreover allow ourselves to individuate languages at arbitrarily fine levels of grain: we think of languages not (or not solely) as entities like *English*, but sometimes as entities like *Anscombe's idiolect* of English; and sometimes even as entities like Anscombe's

idiolect of English *at t*.

Finally, String Theory employs several relations among linguistic objects and between linguistic and nonlinguistic objects. We use the relation a language-user bears to a sentence when the sentence belongs to the user's language; the semantic valuation function, mapping an expression to an entity representing its meaning; and the relation one (language-bound) sentence bears to another (bound, potentially, to a distinct language) when they translate one another.[3]

In order to keep track of all this without lapsing into cumbersome verbosity, we will need to introduce some slightly intricate notation:

- *Language-users:*
  - The object-language user: o
  - The metalanguage user: m
  - The assessor: a
- *Linguistic objects:*
  - The object-language sentence referred to in a truth-ascription: O
  - The metalanguage singular term referring to that object-language sentence: μ
  - The assessment-language sentence the truth-ascription encodes: A
  - The string functioning in the metalanguage as a 'truth predicate': T
- *Relations:*
  - The expression E belongs to x's language: $\Lambda(E, x)$
  - The semantic value of the expression E in x's language is j: $[\![E : x]\!] = j$

[3]We set aside questions about what it is for a pair of sentences to translate one another. We intend by this the usual meaning: they mean the same thing, or close enough. Our deflationism is about *truth*, and not necessarily about meaning.

- The translation from x's language into y's language of the expression E is the expression F: $[E : x] \rightarrow [F : y]$
- The string s as assessed by x is the expression E: $\langle s : x \rangle = E$

- *Language-bound quote marks:*
  - A way of making explicit that the metalanguage quote-name 'E' is referring to an expression in x's language: ${}^{\mathsf{x}}E^{\mathsf{x}}$

- *Intralinguistic constraints:*

  Whenever:

  - $[E : x] \rightarrow [F : y]$; or
  - for some j, $[\![\sigma : x]\!] = j$; or
  - for some s, $\langle s : x \rangle = E$:

  $\Lambda(E, x)$ and $\Lambda(F, y)$; and $\Lambda({}^{\mathsf{x}}\sigma^{\mathsf{x}}, x)$.

  - Namely:
    - Any expression that translates out of x's language is an expression of x's language;
    - Any expression that is a translation into y's language is an expression of y's language;
    - Any expression that has a sentence value in x's language is an expression of x's language;
    - Whenever x assesses a string as being an expression, that expression is of x's language.

With this notation in hand, we can state our theory:

The String-to-Sentence Theory of Truth

Let $\Lambda(O, o)$, $\Lambda(\mu, m)$, $\Lambda(A, a)$; $[\![\mu : m]\!] = O$; $[O : o] \rightarrow [A : a]$: then

– $\langle \mathsf{T}(\mu) : a \rangle = A$.

Or, in words:

Let O be a sentence of o's language, μ be a singular term of m's language, and A be a sentence of a's language; and let the semantic value of the term μ be the sentence O; and let A be the translation into a's language of O: then

– The string constituting the expression μ concatenated with any string serving for m as a truth-predicate, as produced by m, is assessed by a as constituting the sentence A.

Or, more efficiently:

– $\langle \mathsf{T}(\mu) : a \rangle =$

$$(\iota E)\Big([[\![\mu : (\iota x)(\Lambda(\mu, x))]\!] : (\iota y)(\Lambda([\![\mu : (\iota x)(\Lambda(\mu, x))]\!], y))] \rightarrow [E : a]\Big)$$

(Where, as is standard, $(\iota x)(\Phi(x))$ abbreviates 'the Φ')

Or, in words, more efficiently:

– The string constituting the expression μ concatenated with any string serving for the user of μ as a truth-predicate, as produced by the user of μ, is assessed by a as constituting the sentence translating the semantic value of μ into a's language.

Or, jointly optimizing efficiency and comprehensibility:

Where $m = (\iota x)(\Lambda(\mu, x))$—the user of μ—and $o = (\iota y)(\Lambda([\![\mu : m]\!], y))$—the user of the semantic value of μ:

– $\langle \mathsf{T}(\mu) : a \rangle =$

$$(\iota E)\Big([[\![\mu : m]\!] : o] \rightarrow [E : a]\Big)$$

(In words: a hears m's utterance of 'μ is true' as expressing a's sentence A, where A is the sentence translating the sentence O to which μ refers)

This final statement of String Theory is the one we we will employ in practice.

The predictive power of String Theory will be enhanced if we add the following clauses about lexical meaning as auxiliary hypotheses:

Devaluation:

$[\![^{\mathsf{x}}\varphi^{\mathsf{x}} : y]\!] = {}^{\mathsf{x}}\varphi^{\mathsf{x}}$

(The quote-name of an expression refers to the embedded expression)

Uncomplementing:

$[\![\mathsf{comp}^\frown\varphi : y]\!] = {}^{\mathsf{y}}\varphi^{\mathsf{y}}$

(A complementized sentence (when concatenated with a 'truth predicate') refers to the embedded sentence)

Idempotence:

$[E : x] \rightarrow [E : x]$

(Every language-bound expression is its own translation)

This ends our presentation of String Theory.

String Theory requires the revision of certain fairly deeply rooted doctrines in contemporary philosophy of logic and language:

3. (a) As we will see in the next section, String Theory predicts the existence of widespread orthographic externalism.

   (b) In addition, String Theory evidently requires 'assessment-sensitivity' in orthography: which sentence Fred's utterance constitutes 'for me' can be distinct from which sentence it constitutes 'for you' or for Fred, without any of us being somehow mistaken in finding our respective sentence so constituted.

   The existence of assessment-sensitivity in *semantics* (MacFarlane 2003) is taken with increasing seriousness in contemporary philosophy of language. In combination with the doctrine that some expressions are

individuated semantically (D), this may yield orthographic assessment-sensitivity as well; and as we will see, understanding the liar requires seeing orthographic assessment-sensitivity as resulting in semantic assessment-sensitivity, through the stipulation that a name refer to a sentence constituted by a string containing a truth-predicate.

(c) Perhaps most deeply rooted is our thinking of orthography/syntax, semantics, and pragmatics as strongly 'autonomous'. On this picture, an orthographic structure is fully computed by an 'encapsulated module', which is then sent off to another 'module' for semantic interpretation; the result is then sent on to another 'module' for pragmatic interpretation. Interpretation never goes in the opposite direction: there is no 'feedback' from pragmatics into semantics, or from semantics into orthography.

String Theory conflicts with this picture. There is feedback from semantics into orthography: which sentence is constituted by an utterance of $\mathsf{T}(\mu)$ depends on the semantic value of μ. And, if we think of assessment as part of 'pragmatics', there is also, evidently, feedback from pragmatics into orthography.

But this picture, too, is subject to revision. The tradition of 'formal pragmatics' (Lewis 1979) draws only a highly porous line between semantics and pragmatics. It is obviously the case that how one understands an expression influences how it sounds:[4] evidently there is no strict *causal* one-way sign among these domains. In our case, the dependence is *constitutive* rather than causal; still, the presence of one suggests the coherence of the other.

In general, philosophy has not really investigated orthographic matters with much gusto. So it would seem premature to regard these deeply rooted opinions as somehow principled ones. If we benefit sufficiently from doing so, we should be ready to uproot them.

[4]Example: for a lark, I would from time to time pronounce the digit 'eight' as the word 'ape'. I don't think my audience ever remarked on this piece of whimsy.

As I will now argue, the benefits available from String Theory are really quite extensive. This may suggest that philosophy would do well to clear out the cobwebs, blow off the dust, and delve back into questions orthographic.

3 String Theory in action

Let us work String Theory out on some examples. Afterward, we will collect our observations, highlighting several aspects of the explanatory power of the view.

3.1 Examples

E1. **Idiolectic homophony:**

Let μ be '$^{\mathsf{o}}$snow is white$^{\mathsf{o}}$'; let $a = m = o$ be me (the author) now (at the time of writing).

Then, by String Theory:

$$\langle \mathsf{T}(^{\mathsf{o}}\text{snow is white}^{\mathsf{o}}) : a\rangle = (\iota E)([[\![^{\mathsf{o}}\text{snow is white}^{\mathsf{o}} : m]\!] : o] \to [E : a])$$

(which, because $a = m = o$)

$$= (\iota E)([[\![^{\mathsf{a}}\text{snow is white}^{\mathsf{a}} : a]\!] : a] \to [E : a])$$

(which, by Devaluation)

$$= (\iota E)([^{\mathsf{a}}\text{snow is white}^{\mathsf{a}} : a] \to [E : a])$$

(which, by Idempotence)

$$= {}^{\mathsf{a}}\text{snow is white}^{\mathsf{a}}.$$

So, in my assessment, the string (1) constitutes the sentence 'snow is white'.

E2. **Public language homophony:**

Let μ be '^{o}snow is white^{o}'; let $m = o$ be me (the author) now (at the time of writing); let a be you, the reader, at the time of reading.

Then, by String Theory:

$$\langle \mathsf{T}(^{o}\text{snow is white}^{o}) : a\rangle = (\iota E)([[\![^{o}\text{snow is white}^{o} : m]\!] : o] \rightarrow [E : a])$$

(which, because $m = o$)

$$= (\iota E)([[\![^{m}\text{snow is white}^{m} : m]\!] : m] \rightarrow [E : a])$$

(which, by Devaluation)

$$= (\iota E)([^{m}\text{snow is white}^{m} : m] \rightarrow [E : a])$$

(which, assuming that you and I use 'snow is white' to mean the same thing)

$$= {}^{a}\text{snow is white}^{a}.$$

So, in your assessment, the string (1) as I produce it constitutes the sentence 'snow is white'.

E3. **'Propositional' application of 'is true':**

Let μ be 'that snow is white'; let m be me (the author) now (at the time of writing); let a be you, the reader, at the time of reading.

Then, by String Theory:

$$\langle \mathsf{T}(\text{that snow is white}) : a\rangle = (\iota E)([[\![\text{that snow is white} : m]\!] : o] \rightarrow [E : a])$$

(which, by Uncomplementizing)

$$= (\iota E)([^{m}\text{snow is white}^{m} : o] \rightarrow [E : a])$$

(which, by the intralinguistic constraints)

$$= (\iota E)([^{m}\text{snow is white}^{m} : m] \rightarrow [E : a])$$

(which, assuming that you and I use ‘snow is white’ to mean the same thing)

$$= {}^{a}\text{snow is white}^{a}.$$

So, in your assessment, the string ‘that snow is white is true’ as I produce it constitutes the sentence ‘snow is white’.

E4. **Cross-linguistic ascription I:**

Let μ be ‘^{o}schnee ist weiss^{o}’; let o be Carnap in 1932; let $a = m$ be me (the author) now (at the time of writing).

Then, by String Theory:

$$\langle \mathsf{T}({}^{o}\text{schnee ist weiss}^{o}) : a\rangle = (\iota E)([[\![{}^{o}\text{schnee ist weiss}^{o} : m]\!] : o] \rightarrow [E : a])$$

(which, because $a = m$)

$$= (\iota E)([[\![{}^{o}\text{schnee ist weiss}^{o} : a]\!] : o] \rightarrow [E : a])$$

(which, by Devaluation)

$$= (\iota E)([{}^{o}\text{schnee ist weiss}^{o} : o] \rightarrow [E : a])$$

(which, given my knowledge of German)

$$= {}^{a}\text{snow is white}^{a}.$$

So, in my assessment, the string ‘$^{\text{Carnap}}$schnee ist weiss$^{\text{Carnap}}$ is true’ as I produce it constitutes my sentence ‘snow is white’.

E5. **Cross-linguistic ascription II:**

Let μ be ‘^{o}snow is white^{o}’; let o be me (the author) now (at the time of writing); let $a = m$ be Carnap in 1932.

Then, by String Theory:

$$\langle \mathsf{T}(^{o}\text{snow is white}^{o}) : a\rangle = (\iota E)([[\![^{o}\text{snow is white}^{o} : m]\!] : o] \rightarrow [E : a])$$

(which, because $a = m$)

$$= (\iota E)([[\![^{o}\text{snow is white}^{o} : a]\!] : o] \rightarrow [E : a])$$

(which, by Devaluation)

$$= (\iota E)([^{o}\text{snow is white}^{o} : o] \rightarrow [E : a])$$

(which, going by my knowledge of German)

$$= {}^{a}\text{schnee ist weiss}^{a}.$$

So, in Carnap's assessment, the string '^{me}snow is white^{me} ist wahr' as he produces it constitutes his sentence 'schnee ist weiss'.

E6. **Cross-linguistic and 'propositional':**

Let μ be 'dass schnee ist weiss'; let m be Carnap in 1932; let a be me now.

Then, by String Theory:

$$\langle \mathsf{T}(\text{dass schnee ist weiss}) : a\rangle = (\iota E)([[\![\text{dass schnee ist weiss} : m]\!] : o] \rightarrow [E : a])$$

(which, by Uncomplementizing)

$$= (\iota E)([^{m}\text{schnee ist weiss}^{m} : o] \rightarrow [E : a])$$

(which, by the intralinguistic constraints)

$$= (\iota E)([^{m}\text{schnee ist weiss}^{m} : m] \rightarrow [E : a])$$

(which, going by my knowledge of German)

$$= {}^{a}\text{snow is white}^{a}.$$

So, in my assessment, the string ‘es ist wahr, dass schnee ist weiss’ as Carnap produces it constitutes the sentence ‘snow is white’.

E7. **Three distinct language-users:**

Let μ be ‘^{o}snow is white^{o}’; let o be me now; let m be Carnap in 1932; let a be Derrida in 1985.

Then, by String Theory:

$$\langle \mathsf{T}(^{o}\text{snow is white}^{o}) : a\rangle = (\iota E)([[\![^{o}\text{snow is white}^{o} : m]\!] : o] \rightarrow [E : a])$$

(which, by Devaluation)

$$= (\iota E)([^{o}\text{snow is white}^{o} : o] \rightarrow [E : a])$$

(which, going by my knowledge of French)

$$= {}^{a}\text{la neige est blanc}^{a}.$$

So, in Derrida’s assessment, the string ‘^{me}snow is white^{me} ist wahr’ as Carnap produces it constitutes Derrida’s sentence ‘la niege est blanc’.

E8. **Three distinct language-users plus deixis:**

Let μ be ‘the thing Carnap said’; let o be Carnap in 1932; let m be me now; let a be Derrida in 1985.

Then, by String Theory:

$$\langle \mathsf{T}(\text{the thing Carnap said}) : a\rangle = (\iota E)([[\![\text{the thing Carnap said} : m]\!] : o] \rightarrow [E : a])$$

(which, in light of the meaning of my expression ‘the thing Carnap said’, and assuming Carnap said ^{o}schnee ist weiss^{o})

$$= (\iota E)([^{o}\text{schnee ist weiss}^{o} : o] \rightarrow [E : a])$$

(which, going by my knowledge of French and German)

$$= {}^{a}\text{la neige est blanc}^{a}.$$

So, in Derrida's assessment, the string 'the thing Carnap said is true' as I produce it constitutes Derrida's sentence 'la niege est blanc'.

E9. **Embedded occurrences:**

Let μ be '$^{o}\mathsf{T}(^{\text{lewis}}\text{snow is white}^{\text{lewis}})^{o}$'; let o be me now; let m be Carnap in 1932; let a be Derrida in 1985; let the 'sub-object' language user be David Lewis in 1985.

Then, by String Theory: $\langle \mathsf{T}(^{o}\mathsf{T}(^{\text{lewis}}\text{snow is white}^{\text{lewis}})^{o}) : a\rangle$

$$= (\iota E)([[\![^{o}\mathsf{T}(^{\text{lewis}}\text{snow is white}^{\text{lewis}})^{o} : m]\!] : o] \rightarrow [E : a])$$

(which, by Devaluation)

$$= (\iota E)([^{o}\mathsf{T}(^{\text{lewis}}\text{snow is white}^{\text{lewis}})^{o} : o] \rightarrow [E : a])$$

(which, by the intralinguistic constraints)

$$= (\iota E)([\langle ^{\text{lewis}}\text{snow is white}^{\text{lewis}} : o\rangle : o] \rightarrow [E : a])$$

(which, by String Theory)

$$= (\iota E)\Big(\Big[(\iota F)([[\![^{\text{lewis}}\text{snow is white}^{\text{lewis}} : o]\!] : \text{Lewis}] \rightarrow [F : o]) : o\Big] \rightarrow [E : a]\Big)$$

(which, by Devaluation)

$$= (\iota E)\Big(\Big[(\iota F)([^{\text{lewis}}\text{snow is white}^{\text{lewis}} : \text{Lewis}] \rightarrow [F : o]) : o\Big] \rightarrow [E : a]\Big)$$

(which, because Lewis and I use a common language)

$$= (\iota E)\Big([^{o}\text{snow is white}^{o} : o] \rightarrow [E : a]\Big)$$

(which, going by my knowledge of French)

$$= {}^{a}\text{la neige est blanc}^{a}.$$

So, in Derrida's assessment, the string '$^{\text{me lewis}}$snow is white$^{\text{lewis}}$ is true$^{\text{me}}$ ist wahr' as Carnap produces it constitutes Derrida's sentence 'la niege est blanc'.

E10. **Indexicality I:**

Let μ be '^{o}my pants are on fire^{o}'; let $m = o$ be me, the author, at the time of writing; let a be you, the reader, at the time of reading.

Then, by String Theory:

$$\langle \mathsf{T}({}^{o}\text{my pants are on fire}^{o}) : a \rangle = (\iota E)([[\![{}^{o}\text{my pants are on fire}^{o} : m]\!] : o] \rightarrow [E : a])$$

(which, because $m = o$)

$$= (\iota E)([[\![{}^{m}\text{my pants are on fire}^{m} : m]\!] : m] \rightarrow [E : a])$$

(which, by Devaluation)

$$= (\iota E)([{}^{m}\text{my pants are on fire}^{m} : m] \rightarrow [E : a])$$

(which—going out on a limb)

$$= {}^{a}\text{the author's pants are on fire at the time of writing}^{a}.$$

So, in your assessment, the string '$^{\text{me}}$my pants are on fire$^{\text{me}}$ is true' as I produce it constitutes your sentence 'the author's pants are on fire at the time of writing'.

E11. **Indexicality II:**

Let μ be '^{o}your pants are on fire^{o}'; let $m = o$ be me, the author, at the time of writing; let a be you, the reader, at the time of reading.

Then, by String Theory:

$\langle \mathsf{T}(^{o}\text{my pants are on fire}^{o}) : a\rangle = (\iota E)([[\![^{o}\text{my pants are on fire}^{o} : m]\!] : o] \rightarrow [E : a])$

(which, because $m = o$)

$= (\iota E)([[\![^{m}\text{my pants are on fire}^{m} : m]\!] : m] \rightarrow [E : a])$

(which, by Devaluation)

$= (\iota E)([^{m}\text{my pants are on fire}^{m} : m] \rightarrow [E : a])$

(which—going out on a limb)

$= {}^{a}\text{my pants were on fire at the time of writing}^{a}.$

So, in your assessment, the string '^{me}your pants are on fire^{me} is true' as I produce it constitutes your sentence 'my pants were on fire at the time of writing'.

E12. **Indexicality, three distinct language-users:**

Let μ be ' 'my pants are on fire' '; let o be David Kaplan in 1975; let m be me (the author) at the time of writing; let a be you (the reader) at the time of reading.

Then, by String Theory:

$\langle \mathsf{T}(^{o}\text{my pants are on fire}^{o}) : a\rangle = (\iota E)([[\![^{o}\text{my pants are on fire}^{o} : m]\!] : o] \rightarrow [E : a])$

(which, by Devaluation)

$= (\iota E)([^{o}\text{my pants are on fire}^{o} : o] \rightarrow [E : a])$

(which—going out on a limb)

$= {}^{a}\text{David Kaplan's pants are on fire in 1975}^{a}.$

So, in your assessment, the string ‘$^{\text{kaplan}}$my pants are on fire$^{\text{kaplan}}$ is true’ as I produce it constitutes your sentence ‘David Kaplan’s pants are on fire in 1975’.

E13. **Embedding under connectives:**

Outside of the scope of T, however, everything behaves as normal. Returning to the stipulations in case (E4), consider the sentence $\neg\mathsf{T}$(‘schnee ist weiss’). As we saw, in that case, the string I produce embedded under $\neg$ constitutes the sentence ‘snow is white’. So the full sentence of my language is $\neg^\frown$‘snow is white’—or, in English rather than Loglish, ‘snow is not white’.

More generally, we may establish the following rule, which will prove useful in discussion of the liar:

Exportation:
$\langle\neg\mathsf{T}(\mu) : a\rangle = \neg^\frown\langle\mathsf{T}(\mu) : a\rangle$

Proof:

$$\langle\mathsf{T}(\mu) : a\rangle = (\iota E)\Big([[\![\mu : m]\!] : o] \rightarrow [E : a]\Big);$$

so

$$\neg^\frown\langle\mathsf{T}(\mu) : a\rangle = \neg^\frown(\iota E)\Big([[\![\mu : m]\!] : o] \rightarrow [E : a]\Big);$$

and any correct translation of the negation sign translates a negated sentence as the negation of the translation of the complement of the negation sign, so

$$\langle\neg\mathsf{T}(\mu) : a\rangle = \neg^\frown(\iota E)\Big([[\![\mu : m]\!] : o] \rightarrow [E : a]\Big);$$

assembling parts, we have our rule.

E14. **Quantification-in:**

Because T contains an argument place, we would anticipate being able to quantify into it, as in $(\forall s$: Carnap said $s)(\mathsf{T}(s))$. What would that mean, as uttered and assessed by me?

Consider a Tarski-type theory of quantification on which 'everyone eats', $(\forall x)(Ex)$, expresses the proposition containing w just if, for every assignment of a thing to x, w is contained in the proposition expressed by Ex relative to that assignment.

By analogy, $(\forall s$: Carnap said $s)(\mathsf{T}(s))$ expresses the proposition containing w just if, for every assignment of a sentence Carnap spoke to s, w is contained in the propposition expressed by $\mathsf{T}(s)$ relative to that assignment.

Suppose that Carnap spoke two sentences: 'schnee ist weiss' and 'wahrheitswerte sind satzbedeutungen'. Then the proposition expressed contains w just if, for the assignment of 'schnee ist weiss' to s, w is contained in the proposition expressed by $\mathsf{T}(s)$ relative to that assignment, *and* for the assignment of 'wahrheitswerte sind satzbedeutungen' to s, w is contained in the proposition expressed by $\mathsf{T}(s)$ relative to that assignment.

In that case, the proposition expressed contains w just if w is contained in the proposition expressed by T('schnee ist weiss') *and* w is contained in the proposition expressed by T('wahrheitswerte sind satzbedeutungen').

In accord with the discussion at (E4), as assessed by me, T('schnee ist weiss') has the orthographic form 'snow is white', while T('wahrheitswerte sind satzbedeutungen') has the orthographic form 'the extensions of truth values are sentences'.

So the proposition expressed by 'everything Carnap said is true' contains w just if w is contained in the proposition expressed by 'snow is white' *and* w is contained in the proposition expressed by 'the extensions of sentences are truth values'.

Which is, finally, to say that the proposition expressed contains w just if, in w, snow is white and the extensions of sentences are truth values.

As these examples show, String Theory predicts extensive orthographic externalism: in (E6), we may suppose, Derrida does not know what the thing Carnap said was; nevertheless, though Derrida would not be in a position to make this explicit, my string as assessed by Derrida is the sentence 'la neige est blanc'.

3.2 The allure of String Theory

The predictions of String Theory are attractive:

4. (a) (E1) shows that, in purely idiolectic discourse (when $a = m = o$), a string ascribing truth to a quote-named sentence constitutes the embedded sentence: the sentence constituted by the string ' 'snow is white' is true' just *is* the sentence 'snow is white'.

   This, of course, yields something resembling the 'T-schema':

   (T) $\varphi \dashv\vdash \mathsf{T}(\ulcorner\varphi\urcorner)$

   For after all, given that the left-hand sentence and the right hand sentence are *identical*, of course they are jointly entailing.

   Discussion in the literature of the T-schema is problematic. As is widely recognized, examples like (E4) and (E10)—not to mention (2)—require the T-schema to be somehow restricted. But exactly how this is to be done is never, to my knowledge, made explicit: a *purely idiolectic* restriction seems to strike many as unattractive, perhaps in light of cases like (E2). For the T-schema is often treated as the basic and inexplicable fundament of our understanding of the truth-predicate. And if we only know how to apply the truth-predicate in purely idiolectic cases, how are we supposed to apply it outside purely idiolectic cases? (For that matter, what did people do before the invention of the quote-mark?)

   String Theory walks us out of this conceptual morass. There is nothing fundamental, according to String Theory, about the T-schema. As the discussion in (E1) shows, String Theory derives the T-schema in the purely idiolectic case from Devaluation and Idempotence.

Conversely, String Theory allows us to see also just when the T-schema might fail: in order to call Idempotence, the assessor and the user of the object-language must be identical. Otherwise, when $a \neq o$, appeal to the T-schema requires special circumstances, of the sort exhibited in (E2) but not in an array of other cases—including (E4), (E5), (E7), (E10), (E11), and (E12).

(b) String Theory provides a unified treatment of truth-ascription via quote-names, 'that'-clauses (E3, E6), deixis (E8), and quote-named embedded truth-ascriptions (E9). To my knowledge, extant theories of truth do not provide such a unified treatment; in particular, attempts to incorporate *deictic* truth-ascription are nonexistent.

(c) String Theory integrates naturally with a well-understood theory of quantification, as is exhibited in (E14). To my knowledge, no deflationary theory of truth accommodates quantification-in: no small defect, if our primary purpose with truth-discourse is to enable discussion of statements the exact nature of which is immaterial.

4 String Theory and the liar

Perhaps the principal allure of String Theory is the attractive resolution it offers of the liar paradox. Surely—contra **?**—the ordinary truth-predicate is not paradoxical: we manage to use it all the time without the world exploding. As we will now see, String Theory provides a tidy explanation of how that can be.

Note that our intention is to save *ordinary-language* truth discourse from the liar paradox. String Theory is an empirical hypothesis about this natural phenomenon. Let there be no doubt that artificial liar-paradoxical truth-predicates can be defined. The moral should be obvious: don't define them. Problem solved! By contrast, the related moral about ordinary truth-discourse—stop using it—might be too stiff medicine. The security of this discourse is a matter about which we have some non-derivative concern; if String Theory provides that security, that is reason enough to care.

Our solution here is, 'paradoxically', somewhat akin to Tarski's. The liar paradox goes away because there isn't any liar sentence. Unlike Tarski, our concern is natural rather than artificial language. And unlike Tarski, we both postulate no hierarchy of languages and *explain* why there can't be any liar sentence. The reason is that the stipulation of the name of the liar sentence is semantically defective. And this reason falls out of String Theory almost immediately.

4.1 The structure of the liar paradox

We follow the presentation of the liar paradox by **?**. The paradox begins with a stipulative definition: let $L = \neg\mathsf{T}(\ulcorner L \urcorner)$. The definition carries with it standard existence and uniqueness presuppositions: in order for 'L' to refer, there must be exactly one sentence with the property $(\lambda S)(S = \neg\mathsf{T}(\ulcorner S \urcorner))$. Setting aside the uniqueness question, 'L' fails to refer if the *existence* presupposition is unmet: if no sentence has that property. And unless 'L' refers, discourse using 'L' is semantically defective.

To assess the impact of such defectiveness, let us grant for the moment that the existence and uniqueness presuppositions of the definition are met: that there is exactly one sentence with the property $(\lambda S)(S = \neg\mathsf{T}(\ulcorner S \urcorner))$. We then reason as follows to a contradiction.

First, grant the validity of both directions of the T-schema:

Capture: $A \vdash \mathsf{T}(\ulcorner A \urcorner)$

Release: $\mathsf{T}(\ulcorner A \urcorner) \vdash A$

Then we argue by cases:

(P1) $\mathsf{T}(\ulcorner L \urcorner) \vee \neg\mathsf{T}(\ulcorner L \urcorner)$ [LEM]

(P2) Case One:

(a) $\mathsf{T}(\ulcorner L \urcorner)$

(b) L [2a: release]

(c) $\neg\mathsf{T}(\ulcorner L\urcorner)$ [2b: definition of L]

(d) $\neg\mathsf{T}(\ulcorner L\urcorner) \wedge \mathsf{T}(\ulcorner L\urcorner)$ [2a, 2c: adjunction]

(P3) Case Two:

(a) $\neg\mathsf{T}(\ulcorner L\urcorner)$

(b) L [3a: definition of L]

(c) $\mathsf{T}(\ulcorner L\urcorner)$ [3b: capture]

(d) $\neg\mathsf{T}(\ulcorner L\urcorner) \wedge \mathsf{T}(\ulcorner L\urcorner)$ [3a, 3c: adjunction]

(P4) $\neg\mathsf{T}(\ulcorner L\urcorner) \wedge \mathsf{T}(\ulcorner L\urcorner)$ [1–3: dilemma]

Contradiction.

We will argue that the definition of 'L' is semantically defective, because its existence presupposition fails. The effect on this line of argument were either presupposition of the definition of 'L' to fail would be catastrophic. For the LEM premiss would then fail to be in any way meaningful, as each disjunct would embed a nonreferring singular term. And if so, the argument to contradiction cannot move forward. Problem solved.

This resolution of the liar paradox would also be immune to 'revenge' in the deepest possible way. For consider a sentence defined as follows: $L = \neg(\mathsf{T}(\ulcorner L\urcorner) \wedge$ Nondefective($\ulcorner L\urcorner$)): under the supposition that the definition of 'L' *is* semantically defective, the T-schema has no hope of releasing L from $\mathsf{T}(\ulcorner L\urcorner)$—because there is nothing to be released.

So if the definition of 'L' can be shown to fail, we would be in a position to put the liar paradox—along with its offspring and their perpetual vendetta—at last to rest.

4.2 Every sentence is the truth-teller

We begin by discussing the slightly more straightforward case of the 'truth-teller'.

Let us define a metalinguistic singular term naming a sentence that is itself a part of the metalanguage, as follows: $\tau = \mathsf{T}(\tau)$. τ is supposed to be our truth-teller.

But to be perfectly secure, we should move slowly, and we should use the 'formal mode'. How do do this? We want to stipulate a semantic value for the metalinguistic singular term τ. Should we say that $[\![\tau : m]\!] = \mathsf{T}(\tau)$? That is not enough: our aim is to define a *sentence*, and String Theory says that the truth-predicate is only ever *part* of a string—though that string may *constitute* a sentence. Since our aim is for τ to refer to a *sentence*, we must perhaps rather stipulate its semantic value as follows: $[\![\tau : m]\!] = \langle \mathsf{T}(\tau) : a \rangle$.

But this is still not quite enough, for our assessment-sensitive orthography ramifies to assessment-senstive semantics. To highlight this, let us purify the discussion of the confusing involvement of self-reference. Then consider this semantic stipulation: $[\![\rho : m]\!] = \langle \mathsf{T}(^{\mathsf{o}}\text{snow is white}^{\mathsf{o}}) \rangle$ for specified $m = o$ set to me. For varying a, String Theory predicts that $\langle \mathsf{T}(^{\mathsf{o}}\text{snow is white}^{\mathsf{o}}) \rangle = (\iota E)([[\![^{\mathsf{o}}\text{snow is white}^{\mathsf{o}} : m]\!] : o] \rightarrow [E : a]) = (\iota E)([^{\mathsf{o}}\text{snow is white}^{\mathsf{o}} : o] \rightarrow [E : a])$, the assessment-language sentence into which my sentence 'snow is white' translates. And of course, as we recall from (E1), (E5) and (E7), when a is me, that sentence is a sentence of English; while when a is Carnap, it is a sentence of German; and when a is Derrida, that sentence is a sentence of French.

Accordingly, the semantic stipulation of the referent of 'ρ', which includes no stipulation that the assessor variable is indexed to the metalanguage-user variable, requires 'ρ' to have an assessment-sensitive semantic value: assessed by me, it refers to a sentence of English; assessed by Carnap, it refers to a sentence of German; assessed by Derrida, it refers to a sentence of French. (It follows as well that the sentence 'ρ contains the word 'snow'' has an assessment-sensitive truth-value: true as assessed by me, false as assessed by Carnap or Derrida.)

If so, let us recognize this semantic assessment-sensitivity explicitly by indexing our semantic valuation brackets to the assessor: $[\![\tau : m]\!]^a = \langle \mathsf{T}(\tau) : a \rangle$. Finally, for convenience, let us grant for the moment that $\langle \mathsf{T}(\tau) : a \rangle$ is some assessment-language sentence A. (Fully explicitly: for assessor a, if some sentence is uniquely $(\lambda S)(S = \langle \mathsf{T}(\tau) : a \rangle)$, $[\![^{\mathsf{a}}A^{\mathsf{a}} : m]\!]^a = (\iota S)(S = \langle \mathsf{T}(\tau) : a \rangle)$.)

Which sentence, then, for a particular assessor, is A? To find out, let us apply String Theory:

5. **The truth-teller:**

   Let μ be 'τ'. Note that $a = o$: because 'τ' is to refer to the assessment-language sentence $\langle \mathsf{T}(\tau) : a \rangle$, the object-language and the assessment-language can be identified. And note that $a = m$: because '$\mathsf{T}(\tau)$' is an assessment-language string using a singular term of the assessment-language, 'τ' must be an expression of the assessment-language; and we have already stipulated that 'τ' is to be an expression of the metalanguage.

   Then, by String Theory:

$$A = [\![\tau : a]\!]^a = \langle \mathsf{T}(\tau) : a \rangle = (\iota E)([[\![\tau : a]\!]^a : o = a] \rightarrow [E : a])$$

   (which, by definition)

$$= (\iota E)([\langle \mathsf{T}(\tau) : a \rangle : a] \rightarrow [E : a])$$

   (which, by stipulation)

$$= (\iota E)([A : a] \rightarrow [E : a])$$

   (which, by Idempotence)

$$= A.$$

Not very helpful, I fear! So let us consider some examples of sentences A might be. Suppose the assessment-language is English. Then, picking something arbitrarily, perhaps A is 'snow is white'. Then the semantic value of τ is 'snow is white'. It is of course the case that $\langle \mathsf{T}(\tau) : a \rangle$, under this supposition, is 'snow is white'—so nothing in that aspect of the definition is violated. Anything else at odds with this supposition? Not that I can see. But that supposition was chosen arbitrarily: could we have just stumbled upon the right value of A so easily? Perhaps we should try an alternative. Picking something else arbitrarily, suppose A is 'snow is pink'. It is of course then the case, under this supposition, that $\langle \mathsf{T}(\tau) : a \rangle$ is then 'snow is pink'. Nothing else in our stipulations seems to conflict with *this* supposition, either. And obviously the same is true for any other sentence we

might have named.

What that all shows is that our efforts at definition have failed. Our semantic stipulation giving the meaning of τ is *underconstrained*. The *uniqueness* presupposition behind that stipulation—that there is a unique such meaning—is false. Therefore no sentence is *the* truthteller sentence: rather, *every* sentence is *a* truthteller sentence.

Is this a surprise? From the standpoint of String Theory, hardly at all. Put informally, the stipulation is that τ is to name **that sentence the assessor to mean the same thing as itself.** That ain't gonna get us nowhere!

4.3 No sentence is the liar

We now turn to the liar. The astute reader will have anticipated that its difficulty is dual to that of the truth-teller. The stipulation of the liar sentence is *overconstrained*: the *existence* presupposition of the stipulation fails.

Again, in cooking up the liar sentence, we want to stipulate a semantic value for a metalinguistic singular term ℓ. Should we say that $[\![\ell : m]\!] = \neg\mathsf{T}(\ell)$? That is not enough: our aim is to define a *sentence*, and String Theory says that the truth-predicate is only ever *part* of a string—though that string may *constitute* a sentence. Since our aim is for ℓ to refer to a *sentence*, better to stipulate its semantic value as follows: $[\![\ell : m]\!] = \langle\neg\mathsf{T}(\ell) : a\rangle$. But again, we should make explicit the semantic assessment-sensitivity: our official stipulation should be $[\![\ell : m]\!]^a = \langle\neg\mathsf{T}(\ell) : a\rangle$. Finally, for convenience, let us grant for the moment that $\langle\neg\mathsf{T}(\ell) : a\rangle$ is some assessment-language sentence A. (Fully explicitly: for assessor a, if some sentence is uniquely $(\lambda S)(S = \langle\neg\mathsf{T}(\ell) : a\rangle)$, $[\![^{\mathsf{a}}A^{\mathsf{a}} : m]\!]^a = (\iota S)(S = \langle\neg\mathsf{T}(\ell) : a\rangle)$.)

Which sentence, then, for a, is A? To find out, we turn again to String Theory:

6. **The liar:**

   Let μ be 'ℓ'. As above, $a = m = o$.

Then, by Exportation:

$$A = [\![\ell : a]\!]^a = \langle \neg \mathsf{T}(\ell) : a \rangle = \neg^\frown \langle \mathsf{T}(\ell) : a \rangle$$

(which, by String Theory)

$$= \neg^\frown (\iota E)([[\![\ell : a]\!]^a : a] \rightarrow [E : a])$$

(which, by definition)

$$= \neg^\frown (\iota E)([\langle \neg \mathsf{T}(\ell) : a \rangle : a] \rightarrow [E : a])$$

(which, by stipulation)

$$= \neg^\frown (\iota E)([A : a] \rightarrow [E : a])$$

(which, by Idempotence)

$$= \neg^\frown A = \neg A.$$

So if A is the liar sentence, $A = \neg A$. But no sentence of my idiolect is equivalent to its own negation. The stipulation that ℓ is to name, in the assessor's language, whatever sentence it is that she uses to mean the same as the negation of her sentence that ℓ names is incoherent: for it requires, impossibly, that the sentence named by ℓ be its own negation.

So again, our efforts at definition have failed. Our semantic stipulation giving the meaning of ℓ is *overerconstrained*. On the one hand ℓ is to name an assessment-language sentence. On the other hand, that sentence is to mean the same as the negation of the assessment-language sentence ℓ names. Both constraints on the referent of ℓ can't be satisfied simultaneously. So the *existence* presupposition behind the stipulation of the meaning of ℓ—that there is any such meaning for ℓ to have—is false. Therefore no sentence is *the* liar sentence: and moreover, *every* sentence is *not* a liar sentence.[5]

[5]Interestingly, the stipulations on defining a liar sentence require it to be entertainable solely in the purely idiolectic context. One consequence of this is that if there were any such thing as the liar

A more complex case is the 'contingent liar' discussed by Kripke (1975). Suppose Nixon says 'Haldeman's statement is not true' and Haldeman says 'Nixon's statement is true'. Let ν be 'Nixon's statement' and η be 'Haldeman's statement'. Let us take these statements separately.

Nixon produces the string $\neg\mathsf{T}(\eta)$; so $[\![\nu : h]\!]^a = \langle\neg\mathsf{T}(\eta) : a\rangle = \neg^\frown\langle\mathsf{T}(\eta) : a\rangle$. By String Theory, this semantic value is $\neg^\frown(\iota E)([[\![\eta : n]\!]^a : a] \rightarrow [E : a])$. Haldeman, meanwhile, produces the string $\mathsf{T}(\nu)$; so $[\![\eta : n]\!]^a = \langle\mathsf{T}(\nu) : a\rangle$. Plugging in, $[\![\nu : h]\!]^a = \neg^\frown(\iota E)([\langle\mathsf{T}(\nu) : a\rangle : a] \rightarrow [E : a])$. And by Idempotence, $[\![\nu : h]\!]^a = \neg^\frown\langle\mathsf{T}(\nu) : a\rangle$—which is, as we have just seen, a defective semantic valuation clause. Accordingly, at least one of Nixon and Haldeman's strings did not constitute a sentence. Which? We do not need to choose: each is a sentence only if the other is.

5 Conclusion

Frege (1918/1956) recognized that 'it is true that I smell the scent of violets' and 'I smell the scent of violets' say the same thing. This moved Frege toward deflationism. But he was moved toward inflationism by his observation that when the scientist is in a position to assert 'my conjecture is true', that is a significant achievement. But the case for inflationism doesn't work. The scientist's achievement is to be in a position to assert her conjecture: or, more precisely, to assert any sentence expressing her conjecture—to utter, in an appropriate way and at an appropriate time, any string encoding such a sentence.

Or at least that is what we should say if we find ourselves in a position to assert that String Theory is true.

---

sentence, we would have no hope of responding to the liar paradox by weakening the T-schema. Another is that, if there were any such thing as the liar sentence, it would not show 'English' to be inconsistent—not, anyway, if a sentence is part of 'English' only if it can be entertained by more than one person.

References

Chalmers, David J., 2003. 'The Content and Epistemology of Phenomenal Belief'. In Quentin Smith and Alexander Jokic, editors, *Consciousness: New Philosophical Essays*. Oxford: Oxford University Press.

Evans, Gareth, 1977. 'Pronouns, Quantifiers, and Relative Clauses (I)'. In *Collected Papers*, volume 7, 76–152. Clarendon Press.

Field, Hartry, 1994. 'Deflationist Views of Meaning and Content'. *Mind*, 103:249–85.

Frege, Gottlob, 1918/1956. 'The Thought'. *Mind*, 65:289–311.

Grover, Dorothy L., Joseph L. Camp, and Nuel D. Belnap, 1975. 'A Prosentential Theory of Truth'. *Philosophical Studies*, 27:73–125.

Hawthorne, John, 2006. 'Direct Reference and Dancing Qualia'. In Torin Alter and Sven Walter, editors, *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford: Oxford University Press.

Kaplan, David, 1990. 'Words'. *Aristotelian Society Supplementary Volume*, 64:93–119.

Kripke, Saul A., 1975. 'Outline of a Theory of Truth'. *Journal of Philosophy*, 72:690–716.

Lewis, David, 1979. 'Scorekeeping in a Language Game'. *Journal of Philosophical Logic*, 8:339–359.

MacFarlane, John, 2003. 'Future Contingents and Relative Truth'. *Philosophical Quarterly*, 53:321–8211.