



## King's Research Portal

DOI:

[10.1111/mila.12115](https://doi.org/10.1111/mila.12115)

*Document Version*

Peer reviewed version

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Michaelson, E. (2016). The Lying Test. *Mind & Language*, 31(4), 470-499. <https://doi.org/10.1111/mila.12115>

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

# The Lying Test

Eliot Michaelson  
King's College London

## Abstract

As an empirical inquiry into the nature of meaning, semantics must rely on data. Unfortunately, the primary data to which philosophers and linguists have traditionally appealed—judgments on the truth and falsity of sentences—have long been known to vary widely between competent speakers in a number of interesting cases. The present essay constitutes an experiment in how to obtain some more consistent data for the enterprise of semantics. Specifically, it argues from some widely accepted Gricean premises to the conclusion that judgments on lying are semantically relevant. It then endeavors to show how, assuming the relevance of such judgments, we can use them to generate a useful, widely acceptable test for semantic content.

## 1 Introduction

Semantic theory has a problem: to make progress, it requires data. But there is widespread disagreement about what constitutes the relevant data. At first glance, this might seem to be a relatively minor problem. After all, nearly everyone agrees that truth-conditions are relevant for semantic theorizing.<sup>1</sup> So why not just stick to judgments on truth and falsity while eschewing any more controversial data? Well, because *those* judgments are

---

Draft as of February 27, 2015.

<sup>1</sup>Importantly though, not *quite* everyone agrees here. One exception is Relevance Theorists like Sperber & Wilson (1986) and Carston (2002), who take truth-conditions to properly attach to certain sorts of pragmatic content (e.g. ‘explicatures’). Other dissenters include Bach (2005, 2006b) and Recanati (2012), though for rather different reasons.

notoriously inconsistent.<sup>2</sup> Without some sort of supplemental data, we thus look forced to resign ourselves to being unable to make progress on a variety of interesting semantic issues—since we will be at a loss for which of our varied truth-judgments to take more or less seriously. Unfortunately, the prospects for settling what constitutes the right sort of supplemental data look dim at present. The problem is that disagreements about which are the relevant supplemental data generally reflect deeper disagreements regarding the very nature of semantic inquiry. In other words, the question of what counts as semantic data is bound up with foundational issues regarding just what it is that semantic theories are in the business of explaining.

The present essay constitutes an attempt to navigate this impasse by exploring one previously overlooked, but potentially rich and widely acceptable, source of semantic data: judgments on lying. While constrained in the potential work they can do (more on this below), lies are worth considering as a relevant source of data for several reasons: first, judgments on lying appear to be both relatively stable and uniform in a variety of cases where other potentially semantically-relevant judgments are neither. Second, judgments on lying are in fact regularly employed in high-stakes, real-world situations (e.g. in courts of law) to track something that looks strikingly like semantic theorists' truth-conditional contents. Third, careful attention to one highly plausible way of helping to distinguish between semantics and pragmatics (i.e. via an appeal to Grice's theory of implicatures) naturally suggests that judgments on lying are relevant to semantic theorizing. And, fourth, as it turns out, *if* judgments on lying are relevant to semantic theorizing, then there is reason to believe that they will provide us traction on some very live debates in contemporary semantics.

Even at the outset, however, it should be stressed that attending to judgments on lying will hardly prove a panacea for semantics. At best, these judgments will help us make progress on the semantics of declarative sentences—but not, for instance, on questions or imperatives (at least not directly). One might also question just how robust judgments on lying actually turn out to

---

<sup>2</sup>By way of illustration, consider sentences involving incomplete definite descriptions, specific indefinite descriptions, quantifiers without explicit domain restrictions, conditionals, future contingents, predicates of taste, epistemic modals, or deontic modals. Part of the particular challenge in offering a semantic account of any of these terms or complexes is that competent speakers often disagree about whether specific uses of sentences that include such expressions are true or false in context.

be across a range of circumstances.<sup>3</sup> Finally, and most importantly, one might steadfastly deny that judgments on lying, like judgments on anything else that we *do* with language, are at all relevant to the project of semantic inquiry. My aim here is merely to demonstrate that *if* we have robust judgments about whether speakers have lied in a variety of circumstances in which it is highly contentious whether the sentences they uttered are true or false, and *if* we think that semantic theory should be at least minimally responsive to empirical data, *then* appealing to judgments on lying looks like a very plausible way to make progress in semantics. In particular, I aim to show that, if these two conditions hold, then we can use judgments on lying to construct a test—the ‘Lying Test’—that should help us sort more from less plausible semantic theories.

While the ambitions of this Lying Test are therefore constrained, it is equally important to stress that they are hardly minimal. In particular, as I will outline below, once we add the Lying Test to our semantic tool kit, it quickly becomes apparent how that test can be used to generate a new argument against at least one venerable semantic theory: namely, the Russellian theory of definite descriptions.

## 2 Background: Pure vs. Empirical Semantics

Uncontroversially, semantics studies the meanings of words and phrases, and how these combine to generate the meanings of sentences. Standardly, this is taken to mean that semantics is in the business, *inter alia*, of associating truth-conditions with sentences, or at least with sentences-at-contexts.<sup>4</sup> Ac-

---

<sup>3</sup>While I take the judgments to which I appeal below to be uncontroversial, I am perfectly open to empirical testing of any these claims. There is also a very real, and interesting, question of just how far from the cases discussed below one can stray while maintaining the robustness of these judgments. Since my aim here is simply to make the case that these judgments are at least *potentially* relevant to semantic theory, the lack of empirical work on judgments regarding lying should prove to be no major impediment.

<sup>4</sup>See, among others, Montague (1970, 1973), Davidson (1967), Partee (1975, 1976), Kripke (1977), Kaplan (1989), Heim & Kratzer (1997), Borg (2004, 2012), Cappelen & Lepore (2005), King & Stanley (2005), and Szabó (2006). As mentioned in footnote 1, this claim, while popular, is by no means universally accepted.

As Strawson (1950) long ago pointed out, due to the prevalence of context-sensitive terms and constructions in natural language, it is probably more accurate to talk of associating truth-conditions not with sentences *per se*, but rather with either sentences-at-contexts or utterances-of-sentences. For the purposes of this inquiry, I remain officially

cordingly, judgments on truth and falsity are standardly understood to be one of the primary inputs for semantic theorizing.

One strand of the semantic tradition—call it ‘Empirical Semantics’—takes it that, while judgments on truth and falsity are one important datum to which semantic theories are accountable, these are not the only such datum. In particular, theorists in this tradition have proposed that semantic theories should also explain our judgments on ‘what is said’ (Kaplan 1989) or ‘what is claimed or stated’ (Camp 2007).<sup>5</sup> Other theorists in this tradition, broadly conceived, have proposed that semantic content should explain judgments about when two speakers have ‘agreed’ (Cappelen & Lepore 2005) or when they have ‘disagreed’ (Cappelen & Hawthorne 2009) in particular contexts. What these various theorists have in common is the thought that semantics should, in some way or other, explain not just the conditions under which particular sentences are true or false, but also how we *use* those sentences in particular contexts towards particular ends. In other words, these theorists all accept that there must be some relatively straightforward connection between what sentences mean and how we use those sentences, even if there is little agreement on what exactly that connection is.<sup>6</sup>

In contrast, another venerable tradition in semantics—call it ‘Pure Semantics’—holds that judgments about how we use sentences are irrelevant to the project of semantic inquiry (cf. Salmon 2004, 2005, Bach 2005, 2006b, Soames 2009). Semantics is about the literal meaning of sentences. How we use sentences, on the other hand, is responsive to all manner of things that have little or nothing to do with literal meaning. As such, our judgments about what is said, stated, or claimed, or about when speakers agree or disagree—all of which are uncontroversially judgments about how sen-

---

agnostic with regard to whether semantic truth-conditions attach to sentences-at-contexts or to utterances-of-sentences. Since the latter will typically prove the more natural formulation below, I generally stick to that locution throughout. However, for the reasons offered in Michaelson (2014), I am in fact inclined to endorse the sentences-at-context theory rather than the utterances-of-sentences theory.

<sup>5</sup>While Kaplan (1989) is probably the best-known proponent of the semantic relevance of judgments on ‘what is said’, Davidson (1968, 1976) is plausibly read as advocating a similar connection in the course of his remarks on ‘samesaying’. For further discussion, see Platts (1979), Burge (1986), and McDowell (1987).

<sup>6</sup>Note that the common commitment of this sort of theorist falls well short of constituting an endorsement of the Wittgensteinian claim that “meaning is use.” In slogan form, such theorists would be more appropriately characterized by the significantly more hedged “meaning explains use, at least in a certain range of cases.”

tences are used—are simply neither here nor there from the point of view of semantics.

This debate regarding the nature of semantic evidence is, quite obviously, bound up with debates regarding the proper way to understand the boundary between semantics and pragmatics. Unfortunately, there would appear to be no straightforward way of mapping these two debates onto each other. The problem is that, while some Pure Semanticists (e.g. Bach 2005, 2006b, Soames 2009) hold that semantic contents are very often incomplete and, therefore, *not* truth-evaluable,<sup>7</sup> others have voiced no such suspicions (e.g. Salmon 2004, 2005). Empirical Semanticists, on the other hand, have tended to be more consistently friendly to the claim that semantic contents are complete and truth-evaluable. It is far from clear, however, that this attitude is forced on the Empirical Semanticist by anything about her basic commitments regarding the sorts of evidence pertinent to semantic theorizing.

I will not pretend to be capable of settling the debate between Pure and Empirical Semanticists here, let alone the related debate regarding the proper distinction between semantics and pragmatics. For present purposes, I am happy to grant to Pure Semanticists that explanations for why particular judgments on language use are supposed to be relevant to the project of semantic inquiry have often been frustratingly thin—and sometimes entirely lacking.<sup>8</sup> On the other hand, given the frequency with which informed, reflective speakers (not to mention semantic theorists) genuinely disagree on the truth-conditions of sentences containing particular, semantically interesting terms and complexes, it is hard to countenance giving up the search for *some sort* of additional semantic data, something to help adjudicate debates when judgments on truth simply give out as a guide to meaning. Pure Semanticists offer us little succor here. In practice, adherents of this camp often seem all-too-willing to ignore the fact that their own judgments on ‘strict’ truth and falsity are not universally accepted and their appeals to ‘semantic intuitions’ not universally shared. If that is to be the methodology of semantics, then we run a very real risk of semantic theories amounting to little more than mappings of individual theorists’ own idiolects.

The genuine risk of such a slide into idiolect-mapping should, I think,

---

<sup>7</sup>In this, these Pure Semanticists posit a picture akin to that of Relevance Theory. The primary differences, I take it, amount to (i) the degree to which such incompleteness is posited to obtain in natural language, and (ii) the mechanisms via which it is supposed to be resolved, either in the mind of the listener or for the purposes of truth-evaluation.

<sup>8</sup>See, however, Cappelen & Hawthorne (2009) for a counter-instance.

serve to motivate at least a provisional willingness to pursue the project of Empirical Semantics. But that willingness immediately saddles us with a bevy of difficult questions: how plausible a story can be offered to the effect that one or another sort of judgment regarding language use is semantically relevant? How consistent are these judgements? And how exactly can these judgments be used to judge the relative merits of different semantic proposals? It is with regard to questions like these that Empirical Semanticists have traditionally parted ways from each other, often radically so.

As with the more general debate between Empirical and Pure Semanticists, I will not pretend to be capable of settling this internecine debate between the various branches of Empirical Semantics in the course of the present inquiry. Rather, while I am inclined (along with e.g. Cappelen & Lepore 1997, 2005, Camp 2007, and Soames 2009) to think that ‘said that’-judgments are far too unstable to be of much use to semantics,<sup>9</sup> it strikes me that further work would be required to discern which, if any, of the other extant Empirical Semantic proposals might prove to be viable options.<sup>10</sup> I am not going to undertake that work here. Instead, I am going to focus on making the case for the semantic relevance of judgments on lying, assuming, that is, that semantic contents either are or determine the truth-conditions of sentences-at-contexts. The hope is that such judgments might then provide us with another point from which to triangulate on the semantic contents of particular terms and phrases—and, thereby, another point from which to evaluate each of the various other Empirical Semantic proposals on the table.

A brief caveat before moving on: given that I will be assuming below that semantic content is truth-conditional, it might seem that theorists who deny this will have little use for this inquiry. Such a conclusion would be overly hasty, however. The way the Lying Test is constructed leaves open the possibility that it will ultimately turn out to be pragmatic, rather than semantic, content that explains our judgments on lying. Thus, those who are committed to the non-truth-conditional of much, or even all, semantic content should have no objection to the Lying Test; it provides no evidence either for or against their preferred theories. What I hope the discussion below will serve to highlight, however, is the need for such theorists to provide an explanation of what exactly underwrites our widely consistent judgments

---

<sup>9</sup>In a slightly different context, see also Ziff (1972) for an exploration of the many different ways that the ‘said that’ locution can be used.

<sup>10</sup>For discussion and criticism of the recently-popular strategy of appealing to judgments on agreement and disagreement, see Plunkett & Sundell (2013) and Plunkett (2014).

on lying—should these judgments turn out to be as widely consistent as I suspect they are. In fact, it seems to me that modified versions of the Lying Test might be of great use to non-truth-conditional semantic theorists in sorting out *different sorts* of pragmatic content and assigning those contents to different explanatory roles.<sup>11</sup> For the sake of brevity, however, I will refrain from laying out the details of the necessary modifications in what follows.<sup>12</sup>

With these preliminaries out of the way, here is the plan for the remainder of this essay. In the next two sections (§§3–4), I aim to motivate the semantic relevance of judgments on lying: first, by considering Grice’s proposed ‘Cancelability Test’ for distinguishing conversational implicature from semantic content (§3); and then by showing how we can muster parallel considerations in order to construct a ‘Lying Test’ which will prove superior to Grice’s proposal in various ways (§4). Roughly, the Lying Test claims that semantic theories should prove at least minimally useful (in a sense to be outlined below) to the task of constructing a plausible theory of lying. In the next section of the essay (§5), I buttress the case for the Lying Test by considering several objections to the test and explaining why each is ill-founded. Then, in the final part of the essay (§§6–8), I illustrate the potential usefulness of the Lying Test by considering how a particular, widely popular semantic theory—namely, Russell’s theory of definite descriptions—fares with respect to it. As we will see, Russell’s theory fails the Lying Test, thus suggesting that it offers an incorrect semantic account of definite descriptions. In §9, I conclude by briefly reflecting on how both the Lying Test and judgments on lying fit into the wider project of semantic inquiry.

### 3 Grice’s Cancelability Test

Grice observed that what we communicate is plausibly the result not just of what the sentences we utter literally mean, but also of the way that such

---

<sup>11</sup>Borg (2012) has recently made the case for both the desirability of such a sorting mechanism and the substantial difficulties facing Relevance Theorists and their fellow-travelers in producing one (pp. 37–38). See also Carston (2002, pp. 189–91) for discussion.

<sup>12</sup>For the record, I am myself inclined towards a pluralistic view of the aims of semantics, and thus suspect that *both* a non-truth-conditional and a truth-conditional notion of semantic content will prove necessary to realize the full range of semantics’ explanatory ambitions. Note that such an attitude is not quite as radical as it may initially sound; it is already implicit in the Kaplanian notion of *character*, particularly if we suppose that characters compose into the characters of complexes, and, ultimately, of sentences.



literal meaning interacts with the background assumptions shared by the speaker and listener. Having drawn this distinction between ‘what is said’ (for Grice, a technical notion in the neighborhood of truth-conditional, semantic content) and ‘what is merely implicated’ (other things communicated, often via the interaction between what is said and these shared background assumptions), Grice proposes a method for distinguishing these sorts of contents. In particular, he offers the following criteria for identifying conversational implicatures (Grice 1989a, p. 44):

- CI** A content  $q$  is conversationally implicated by an utterance of P at a context C if:
- (A) this utterance serves (or should serve, at least, assuming that the listener is paying attention) to communicate, *inter alia*,  $q$ ,
  - (B) ‘not Q’ is a standard way of denying  $q$ , and
  - (C) the speaker could have consistently uttered “P, but not Q” instead.<sup>13</sup>

The basic idea here is that if  $q$  is part of what is said, strictly speaking, by an utterance of P, then P entails  $q$ . In other words, in addition to whatever else an utterance of P might entail, it entails whatever is said. That, in turn, means that if one can consistently utter “P, but not Q,” then  $q$  cannot be a part of what is said by an utterance of just plain P. If it were, then  $q$  would have been entailed by the first conjunct (i.e. P), and so the whole of this utterance would have been inconsistent (Ibid., p. 44).<sup>14</sup> According to Grice, in cases like these—where an utterance of P serves to communicate  $q$  and yet  $q$  can be consistently denied—the natural classification of  $q$  is as a

---

<sup>13</sup>While each of these conditions represents one of the characteristics Grice (1989a) associates with conversational implicatures, I present them here in a rather different manner than he does. Regimenting these conditions as above will prove useful for drawing out the similarities between Grice’s Cancellability Test and my own proposed Lying Test. Despite the liberties taken with the presentation, I believe this to be an accurate paraphrase of Grice’s proposal. For more extensive discussion of how to identify conversational implicatures on Grice’s theory, see Blome-Tillman (2013) and Åkerman (2014).

<sup>14</sup>Grice actually considers another type of cancellation to be possible as well: *implicit* cancellation via the context. The idea is that some non-standard contexts can suffice to make it clear that the speaker could not possibly have meant to communicate some content  $q$ . Since such cases tend to be far more complicated than the more standard cases of *explicit* cancellation, I restrict my focus here solely to instances of the latter. For further discussion of implicit cancellation, see Blome-Tillman (2008).

conversational implicature, itself a paradigm instance of pragmatic content.<sup>15</sup>

Less formally, we might think of what is said (in Grice's technical sense) by an ordinary, non-ironic utterance of P (where P is a declarative sentence) as the sub-part of what is communicated (or ought, reasonably, to be communicated) by that utterance to which the speaker is committed on pain of inconsistency.<sup>16</sup> This sub-part of what is communicated thus stands in contrast to conversational implicatures, which the speaker may subsequently deny without either contradicting herself or even implicitly retracting her original claim.<sup>17</sup>

Despite providing only a set of jointly sufficient conditions for some content's being a conversational implicature, **CI** still looks well-situated to help us in the task of sorting linguistic evidence. By way of illustration, consider utterances of the following two sentences:

- (1) Danny and Laura had a kid and got married.
- (2) Danny and Laura had a kid and got married, but not in that order.

(2) is not a contradiction. Yet (2) includes a clause that explicitly denies something that an utterance of (1) alone would ordinarily serve to communicate: namely, that Danny and Laura *first* had a kid *and then* got married. Call this temporal proposition *q*. *q* satisfies all three conditions of **CI**: (i) in most contexts, it will be communicated by an utterance of (1), (ii) 'not in

---

<sup>15</sup>This typology is not exhaustive. Grice also recognizes another type of implicated content, which he calls 'conventional implicatures'. Conventional implicatures, according to Grice, are a type of 'conventional meaning': they are not cancelable, they are aspects of meaning that are paired with terms via convention, and yet they do not contribute to the truth-conditions of an utterance (Grice 1989d, p. 121). Since Grice didn't use the term 'semantic content', it is somewhat contentious how we ought to map this notion onto his preferred typology of meanings. If that term is used to refer to *at least* truth-conditional content, then it will uncontroversially subsume the category of what is said. For present purposes, the important point is just that what is said is semantic and that conversational implicatures are not; we can safely ignore the question of how best to conceive of conventional implicatures.

<sup>16</sup>Though rarely noted, Grice himself appeals to the notion of 'commitment' in order to help introduce his notion of conversational implicature in his (1989c).

<sup>17</sup>This, of course, is not to claim that speakers don't in fact regularly deny what they have earlier said, nor is it to claim that such retractions are always infelicitous. On the contrary, speakers regularly contradict their previous claims on obtaining new information, changing their mind, etc.—and we commonly accept such retractions and revisions of previous claims without issue. Still, I take it that we can ordinarily recognize the difference between a speaker's saying something new and her retracting or revising some prior claim.

that order’ is a standard way of denying  $q$ , and (iii) utterances of (2) are non-contradictory. According to the standard Gricean interpretation of the case, the temporal ordering communicated by an utterance of (1) should therefore be classified as a conversational implicature, rather than part of what is said by that utterance (Neale 1992, pp. 534–35).<sup>18</sup>

While **CI** has been widely deployed to help distinguish semantic from pragmatic content (cf. Neale 1992, Bach 2001a,b, and Horn 2004), what we really want is something stronger: to be able to test the plausibility of particular semantic theories. Here is how we can leverage **CI** into a test for semantic theories, or what I will call the ‘Cancelability Test’ (**CT**):

- CT** If  $q$  is part of the semantic content associated with a sentence  $P$  at a context  $C$ , then:
- (A) One should not be able to consistently utter “ $P$ , but not  $Q$ ” at  $C$ , where
  - (B) ‘not  $Q$ ’ is a standard way of denying  $q$ .<sup>19</sup>

So, to test a particular semantic theory, we first check to see which  $q$  that theory predicts for an utterance of  $P$  at  $C$ . Then we look to see whether we can consistently utter a sentence of the form “ $P$ , but not  $Q$ ” for some  $Q$  that is a standard way of denying  $q$ . If we can, then the semantic theory in question ‘fails’ the Cancelability Test, providing some evidence against that theory. Less formally, semantic theories fail the Cancelability Test if they mistakenly classify some  $q$ , which is in fact a conversational implicature of  $P$ , as part of the semantic content expressed by  $P$  at  $C$ .

The Cancelability Test yields some initially promising results. First, consider:

---

<sup>18</sup>It is worth noting that Relevance Theorists have sometimes proposed to treat the temporal asymmetry of utterances like (1) as *both* pragmatically derived *and* part of the truth-conditional content of the utterance, i.e. as an ‘explicature’ (cf. Carston 2002). For the sake of simplicity, I will ignore this possibility in what follows.

<sup>19</sup>This ‘Cancelability Test’ differs from a related test that has often gone by the same name: that is, a test *for* conversational implicatures. If some  $q$  is communicated by an utterance of  $P$  at  $C$ , but “ $P$ , but not  $Q$ ” can also consistently be uttered, then  $q$  is said to be a conversational implicature in virtue of exhibiting the property of ‘being cancelable’. **CT**, of course, is really just the contrapositive of this other sort of cancelability test—aimed at testing for semantic (i.e. not pragmatic, and hence not cancelable) content rather than for conversational implicature. For problems with this other sort of cancelability test, see below. See also Blome-Tillman (2013) and Åkerman (2014) for related complications.

- (3) John went to France.
- (4) John went to France, but he did not go to Paris.

Given a certain set of background assumptions (common, say, among North American academics), it is overwhelmingly likely that an utterance of (3) will result in the listener believing not just that John went to France, but also that he went to Paris. Yet it would be strange to think that the proposition that John went to Paris is any part of the semantic content of (3). This suspicion is reinforced by the results of the Cancelability Test. As (4) demonstrates, a semantic theory that predicts that the proposition that John went to Paris is a part of the semantic content of (3) will fail the Cancelability Test, since that content can subsequently be denied without contradiction.

Now consider once more the pair we introduced above, (1) and (2). The proper semantics for ‘and’ is standardly assumed to include no information about temporal order. Consider for the moment the converse theory: that utterances of sentences containing ‘and’ are, or at least sometimes are, associated with semantic contents that include a temporal ordering. This theory fails the Cancelability Test, as was illustrated by (2) above. That is, since the relevant temporal ordering can be canceled, the hypothetical theory under consideration violates clause (A) of **CT**.

Unfortunately, this second result is not quite as robust as it might initially appear, and this points to an important limitation of the Cancelability Test. The problem arises from the observation that sentences containing lexically ambiguous terms can often be appended with ‘but not Q’-clauses without giving rise to a contradiction. This is despite the fact that any particular utterance of a sentence containing an ambiguous term plausibly expresses only a single proposition. Consider, for example, an utterance of (5):

- (5) John picnicked on a bank, but he did not picnic on a riverbank.<sup>20</sup>

‘Bank’ is uncontroversially ambiguous. But the first clause of (5) alone cannot be used to communicate both the proposition that John picnicked on a riverbank and the proposition that he picnicked on a financial institution. Rather, the first clause plausibly expresses just one of these propositions—with the role of the second clause being to make clear to the listener which proposition that is. If this is right, then sometimes we can append ‘but

---

<sup>20</sup>This example is adapted from Davis (1998, p. 46).

not Q'-clauses to utterances of sentences even when those utterances cannot possibly conversationally implicate  $q$  (since they don't serve to convey  $q$ ).

Here is how this generates a problem for the Cancelability Test: such observations should push us to amend that test to include a third clause, to the effect that the sentence under consideration is unambiguous. But some of the semantic theories that the Cancelability Test seemed to tell against are, in fact, ambiguity theories. Take for instance Strawson (1952)'s suggestion that 'and' is ambiguous between a temporal and non-temporal reading. If this is the suggestion under consideration, then the fact that one can consistently utter the relevant "P, but not Q"-construction (i.e. (2)) fails to show that the theory violates the Cancelability Test. The problem is that Strawson's suggestion embraces the ambiguity of 'and', thus leaving it open to him to claim that the appended 'but not Q' serves to flag an intended disambiguation rather than indicating that an implicature has been canceled.<sup>21</sup>

None of this is to suggest that the Cancelability Test has no useful role to play in deciding between semantic theories; it is rather to claim that the Cancelability Test *alone* is insufficient to do the job, particularly in cases where ambiguity theories are thought to be viable options.<sup>22</sup> The Cancelability Test plausibly gets quite a bit right, and it seems to be in the vicinity of

---

<sup>21</sup>The difficulty of teasing apart conversational implicature from ambiguity has been noted for some time, and is discussed at length in both Sadock (1978) and Davis (1998).

<sup>22</sup>To be clear, Grice himself never claimed that the Cancelability Test would suffice to argue against ambiguity accounts of various terms and constructions. Rather, his general strategy was to invoke his Modified Occam's Razor whenever a unified semantics, supplemented with an appropriate pragmatics, proved capable of predicting the content communicated by an utterance involving the relevant term or construction (Grice 1989a, pp. 47–49). While this strategy has generally received a warm welcome in philosophy of language, it does little to underwrite the Cancelability Test *qua* test for semantic content. This is because there is no guarantee that a Gricean pragmatic strategy will be available to account for the full communicated content in any particular instance; that must be argued on a case-by-case basis. What's more, Modified Occam's Razor relies on a preference for simpler semantic theories as opposed to simpler pragmatic theories that is typically assumed rather than argued for. Once more, there is no guarantee that this preference will yield the simplest, or even a maximally coherent, overall theory of language in any particular case (for doubts on this from within the Gricean program itself, see Thomason 1990). Such doubts threaten to substantially reduce the usefulness of the Cancelability Test as an isolated, all-purpose test for content, since the results of that test will always need to be supplemented by further argumentation in order to provide evidence for or against any particular semantic theory. All this, however, leaves unaffected the possibility that the Cancelability Test might serve as one useful tool among others for tracing the bounds of semantic content.

something deeply right about the nature of semantic and pragmatic content. What is surprising is that, despite the enduring popularity of the Cancelability Test, little attention has been paid to the question of whether there are other tests in the vicinity that might prove to be more versatile than Grice's original suggestion. The remainder of this essay will be dedicated to investigating the potential usefulness of one such test, the Lying Test.

## 4 The Lying Test

The Cancelability Test ran off of the observation that standard, non-ironic utterances of declarative sentences commit us to certain contents in a particularly strong way: such that we cannot subsequently deny them without contradicting ourselves. The Lying Test essentially runs off of a parallel suggestion: rather than trying to test for such 'indefeasible commitment'—that is, commitment to a certain content on pain of contradiction—directly, why not test for one of the expected side effects of this sort of commitment? Perhaps some such side effects will prove robust enough to avoid the problems with ambiguity that attempts at direct testing, i.e. the Cancelability Test, fell prey to. My suggestion is this: if semantic content is indeed something to which we are committed, on pain of contradiction, by ordinary, non-ironic utterances of declarative sentences, then semantic content looks like the right sort of thing to play a central (though, perhaps, not an exclusive) role in determining whether or not speakers are lying in particular situations. Accusing a speaker of having lied is, after all, typically a high-stakes way of holding her accountable for what she has said.<sup>23</sup>

This last point—that judgments on lying track (at least) what we indefeasibly commit ourselves to by means of our utterances—is crucial for what follows. Let me therefore pause to offer some brief support for this claim. One good place to see this idea at work is in legal reasoning. The U.S. Le-

---

To be clear, none of the preceding is meant to imply that the basic Gricean explanation of (1)–(4) is incorrect. In fact, the basic outline of that explanation strikes me as overwhelmingly plausible.

<sup>23</sup>There are, perhaps unsurprisingly, ways of claiming that someone has lied that do not seem to involve such a holding them to account for what they have said. I can, for instance, use the term as a form of praise in a context where your clever lie has just gotten us out of a sticky situation (e.g. "Thank God for that lie!"). My claim is merely that, paradigmatically, accusing someone of having lied is a way of holding her accountable for what she has said—not that every use of the term is an instance of this function.

gal Code defines perjury as, basically: stating, under oath, something that one believes to be false.<sup>24</sup> The problem is understanding what such ‘stating’ amounts to. In practice, the U.S. Supreme Court has effectively employed a distinction between *lying* under oath and engaging in other sorts of deceptive linguistic behavior under oath in order to mark the boundary between perjurious and non-perjurious behavior. Only lying under oath—understood as requiring that a witness state something under oath that indefeasibly commits her to some content that she believes to be false—has been found to constitute perjury.<sup>25</sup> The best explanation of this practice, it seems to me, is that judges implicitly recognize a distinction between what witnesses indefeasibly commit themselves to under oath and what they only apparently or defeasibly commit themselves to by means of their utterances. And, while we are likely to be more attuned to this distinction in adversarial situations, there is no reason to think that the distinction between defeasible and indefeasible commitments obtains only under oath.<sup>26</sup> Nor is there reason to think that the contents to which we indefeasibly commit ourselves by means of our utterances should somehow lose their tight connection with lying once we step outside of the courtroom.

Now, back to developing the Lying Test. This, it turns out, will require a bit of finesse, largely because there is no one accepted definition of lying on which to draw. The good news is that we don’t actually need one. Rather, we can make do with what, in the literature at least, has proven to be a widely accepted set of sufficient conditions for an utterance’s being a lie:

**LIE** In uttering a sentence P, X lies to Y if:

- (A) X’s utterance of P has semantic content *p*,
- (B) X believes that *p* is false, and

---

<sup>24</sup>The relevant section of the United States Legal Code, 18 U. S. C. §1621, is in fact both rather complicated and opaque. The above gloss does, however, closely match how it has been understood—and, ultimately, how it has been deployed—in actual case law, e.g. *Bronston v. United States* 409 U. S. 352 (1973).

<sup>25</sup>Cf. *Bronston v. United States* 409 U. S. 352 (1973). There, the Supreme Court found an utterance that plausibly carried with it a false relevance implicature to be non-perjurious. For more extensive discussion of *Bronston* and the myriad difficulties that arise in trying to define perjury, see Solan & Tiersma (2005).

<sup>26</sup>Camp (2007) has in fact suggested that adversarial situations can be more broadly leveraged to help isolate truth-conditional content (pp. 206–07). I will not pursue that line of thought here, however—at least not directly.

(C) X intends to deceive Y with respect to  $p$ .<sup>27</sup>

**LIE** is basically a slightly modified version of the sufficiency half of Augustine’s venerable definition of lying (Augustine 1952, pp. 54–57). As a set of necessary and sufficient conditions, Augustine’s definition has come in for intense criticism in recent years. In fact, most philosophers working on the subject now reject Augustine’s proposal as overly restrictive. However, all of these attacks have all been aimed, more specifically, at the necessity half of this definition.<sup>28</sup> Since I propose to make use of only the sufficiency half of Augustine’s definition, we can safely sidestep these worries.

As a set of merely sufficient conditions, **LIE** is in fact quite plausible. Here is the basic case in favor of **LIE**: in most ordinary cases of lying, familiar from everyday life, the liar says something she believes to be false in order to deceive the listener about that thing. Put slightly differently: suppose that I tell you something that I believe to be false, and suppose further that I intend to deceive you about that something. It seems that I’ve just lied to you. This, I take it, is the basic intuition behind Augustine’s proposal—or, at least, behind the sufficiency half of that claim characterized by **LIE**. Crucially, in and of itself, **LIE** in no way entails that we cannot lie via contents that are merely pragmatically conveyed by our utterances, rather than being stated or said. This is because **LIE** only purports to offer a set of jointly sufficient conditions for lying; it makes no claim to exhaust the full spectrum of ways in which we might be able to lie.

The plausibility of **LIE** can be further reinforced by way of example. Consider the following scenario:

**COOKIE THIEF**

You have kindly baked some cookies

---

<sup>27</sup>Some people (for example, Carson 2006) have strong intuitions to the effect that we need to add a fourth condition as well: that  $p$  must actually be false. Thankfully, such individuals generally seem willing to classify utterances that satisfy (A)–(C) but fail to meet this further condition as ‘attempted lies’. Readers with such intuitions are therefore invited to interpret my use of the term ‘lie’ as meaning **LIE** OR **ATTEMPTED LIE** throughout.

<sup>28</sup>The main problem with Augustine’s proposal, as pointed out in Carson (2006) and Sorensen (2007), is that it would appear to mistakenly classify a number of lies as non-lies. This has led to the widespread rejection of (C) as a necessary condition on lying. For recent attempts to generate more adequate definitions of lying, see Carson (2006), Sorensen (2007), Fallis (2009), Saul (2012), and Stokke (2013). Fallis (2010) raises parallel issues with regard to whether the Augustinian definition provides necessary and sufficient conditions for what he calls ‘deceptive lying’. For an attempt to defend a variant of the Augustinian definition, see Lackey (2013).



and brought them into the department. You plan to leave them in the main office for all of our colleagues to share. However, before you have transferred them from your office to the main office, you briefly go to make some copies—leaving your door slightly ajar. Spotting my chance, I steal the cookies and hide them in my desk in a special smell-proof container. Later, exasperated, you ask me if I know where the cookies are. Fully intending to deceive you, I reply “Ralph took the cookies.”

It seems clear that I lied to you here. And **LIE** offers a reasonable explanation for this. Suppose that my utterance should be associated with the semantic content  $p$ : Ralph took the cookies. I both believe  $p$  to be false and intend to deceive you with regard to  $p$ —that is, I want to get you to believe that Ralph took the cookies (and thus that I didn’t). Therefore, I have lied to you with my utterance of “Ralph took the cookies.”

We are now in a position to leverage **LIE** into a ‘Lying Test’ (**LT**) paralleling **CT** above. Basically, what the Lying Test proposes is that if a speaker isn’t lying in uttering  $P$ , despite both believing  $p$  to be false and intending to deceive the listener about  $p$ , then  $p$  cannot be part of the semantic content of  $P$ . More formally:

- LT** If  $p$  is part of the semantic content associated with a sentence  $P$ , as uttered by  $X$  to  $Y$ , then either:
- (A)  $P$  is a lie,
  - (B) it is not the case that  $X$  believes that  $p$  is false, or
  - (C) it is not the case that  $X$  intends to deceive  $Y$  with respect to  $p$ .<sup>29</sup>

To test a particular semantic theory, we must first pick an utterance falling under its purview and check to see whether or not that utterance is a lie. We will be interested only in non-lies, since only those can possibly result in a theory’s failing the Lying Test.<sup>30</sup> Once we have found some appropriate non-lie, we check to see what the semantic theory in question predicts as the

---

<sup>29</sup>Proponents of non-truth-conditional accounts of semantic content are invited to substitute whatever they take to be the minimal truth-conditional pragmatic notion (e.g. explicature, asserted content) in their theory wherever I have ‘semantic content’. What matters is that this content should be (i) truth-conditional, and (ii) content to which the speaker committed by her utterance, on pain of contradiction.

<sup>30</sup>One might be tempted to try deriving a converse of the Lying Test that runs on lies as opposed to non-lies. Conceived of as a test for semantic theories, two problems arise

relevant  $p$ . If the speaker both believes  $p$  to be false and intends to deceive the listener with respect to that  $p$ , then the semantic theory in question ‘fails’ the Lying Test.<sup>31</sup> If, on the other hand, one of (B) or (C) obtain with regard to this  $p$ , then we will say that the theory in question ‘passes’ the test with respect to this particular case.

Of course, we aren’t really interested in whether a semantic theory passes the Lying Test with respect to any particular case. Rather, what we want to know is whether we can find at least one case, at least one non-lie, for which a given theory predicts that neither (B) nor (C) obtain. Supposing that we can, then we will say that the semantic theory in question ‘fails’ the Lying Test in a more general sense. It is failure in this more general sense that will serve as the focus of our attention in what remains. What such failure indicates is that the semantic theory in question is incompatible with either (i) there being a tight connection between semantics and indefeasible commitment, or (ii) there being a tight connection between indefeasible commitment and lying—even assuming only a very weak, and highly plausible, set of jointly sufficient conditions for an utterance’s being a lie. So long as we take each of these connections to be well-motivated, we therefore obtain some evidence against the semantic theory in question.

## 5 Reinforcing the Lying Test

Now I turn to several worries that naturally arise for the Lying Test. As I will demonstrate, none of these will turn out to be terribly compelling—so long as we appropriately bear in mind the restricted aims of the test.

---

for this proposal: first, since we haven’t ruled out that pragmatic content might suffice to make an utterance a lie, all of the various pragmatic options would have to be considered and ruled out as potential sources of that judgment before any conclusions could be drawn; second, while there is plausibly agreement that **LIE** offers one set of sufficient conditions for an utterance’s being a lie, there is little agreement what other conditions, if any, would make an utterance a lie. This means that a converse Lying Test is bound to be far more controversial, since we will need to check a semantic content-prediction pair against *every plausible* set of sufficient conditions for lying in order to see whether, according to at least one of these, the relevant semantic theory can generate the correct prediction. Barring further progress on the nature of lying itself, a converse lying test thus faces significant barriers to implementation and holds very little practical appeal.

<sup>31</sup>Note that, if Carson (2006), Sorensen (2007), and others are right and the Augustinian definition is too restrictive, then we should expect to find cases where *both* (A) and (B) or (C) obtain. Accordingly, the ‘or’ of **LT** should be read as inclusive rather than exclusive.

First, one might worry that judgments on lying will turn out to be every bit as varied as judgments on truth or falsity. To see that this is not the case, consider sincere utterances of the following:

- (6) France is hexagonal.
- (7) Chocolate is tasty.
- (8) There will be a sea battle tomorrow.

I take it that sincere utterances of (6)–(8) are clearly not lies. Yet it is highly controversial which, if any, truth-conditions should be associated with each. Thus, while there are undoubtedly any number of cases where it is unclear whether or not the speaker has lied, these cases are not co-extensive with the set of cases in which it is unclear whether the speaker said something true or false. This lack of complete overlap is all that that is required to preserve the potential usefulness of the Lying Test; what is needed is for there to be cases where it is clear that someone has not lied, even though it is unclear what the semantic content of her utterance is. (6)–(8) aptly demonstrate the existence of such cases. In fact, cases like (6)–(8) seem to suggest that this gap between judgments on lying and truth may be relatively common.<sup>32</sup>

Second, one might worry that judgments about whether or not someone has lied are ultimately driven by judgments about what they said, claimed, or stated. If that were right, then judgments on lying should turn out to be, at best, only be so clear as judgments on these other matters. That, in turn, would strongly suggest that judgments on lying are explanatorily redundant; they could be dispensed with in favor of direct inquiry into what has been said, claimed, or stated. But this second worry fails to present a compelling challenge to the Lying Test for much the same reason as the first. Once more the problem is that, while the question of what is said, claimed, or stated by utterances of (6)–(8) is highly controversial, the question of whether or not sincere utterances of these sentences are lies is not. So, once again, we seem to have good evidence that judgments on lying can be clear and consistent even when judgments on other potentially semantically relevant facts—such as what has been said, claimed, or stated—are not.

Third, one might wonder whether the Lying Test is compatible with a range of widely-accepted semantic theories. If it were to cut against those,

---

<sup>32</sup>These cases deserve more discussion than I can provide here. In particular, as Andreas Stokke has stressed to me, traditional notions of sincerity may prove difficult to square with cases like these.

one natural response would be to think that something must have gone wrong in crafting the test. To assuage this worry, it will prove instructive to compare the results of the Lying Test with those of the earlier Cancelability Test.

Suppose that (1) is uttered in a context in which Danny and Laura first got married and then had a kid, and where the speaker also believes this to be the case.

(1) Danny and Laura had a kid and got married.

I take it that utterances like (1) are plausibly non-lies even when they are intended to deceive listeners about the temporal order of the relevant events (i.e. in cases where the speaker both believes that Danny and Laura got married first and intends to deceive the listener about the order of events). If this is right, then the Lying Test will turn out to be compatible with a unified, non-temporal theory of ‘and’. What’s more, in contrast to the Cancelability Test, the Lying Test may actually tell against an ambiguity account. Allow me to elaborate.<sup>33</sup>

On a unified, non-temporal account of ‘and’, the observation that (1) is a non-lie is explained by noting that clause (B) of the Lying Test will always be fulfilled. This follows from the speaker believing that Danny and Laura had a kid and got married in one order or the other. So unified accounts of the meaning of ‘and’ pass the Lying Test.

Ambiguity accounts, on the other hand, typically make it the speaker’s prerogative to choose how the ambiguous term should be disambiguated. Supposing that the speaker intends to deceive the listener with regard to the temporal order of the events in question, then one natural way for her to intend this—assuming, again, that the ambiguity theory of ‘and’ is in fact correct—would be for her to intend for her use of ‘and’ to be disambiguated temporally. Thus, the relevant  $p$  associated with the speaker’s utterance of (1) would include a temporal ordering of the marriage- and childbirth-events, birth first. And, since the speaker believes this temporal ordering to be false, she must likewise believe the relevant  $p$  to be false. What’s more, she intends to deceive the listener with regard to that ordering, and hence with regard to  $p$ . So, assuming that the speaker intended for her utterance

---

<sup>33</sup>Were the judgment regarding (1) to turn out to be controversial, the Lying Test would still be compatible with a unified, non-temporal theory of ‘and’. It just wouldn’t have the potential to tell between this theory and an ambiguity account. Thus, the Lying Test would be no better and no worse off than the Cancelability Test in this regard.

to be disambiguated temporally, the ambiguity theorists looks forced to say that the speaker intended to deceive the listener with regard to the relevant temporal ordering—and, thus, that she lied. But that seems wrong. Hence, ambiguity theories of ‘and’ look set to fail the Lying Test.<sup>34</sup>

Next, consider an utterance of (3):

(3) John went to France.

In §3, I pointed out that, given a certain set of background assumptions, an utterance of (3) can be used to effectively convey not just that John went to France, but that he went to Paris in particular. Suppose however that while John did go to France, he *did not* go to Paris, and that the speaker is well aware of this fact. Suppose further that the speaker intends to prompt the listener to believe, on the basis of her utterance of (3), that John *did* go to Paris. In other words, the speaker intends to deceive the listener with her utterance of (3). Still, in spite of its being highly deceptive, this utterance clearly falls short of being a lie.

All of this is perfectly compatible with the Lying Test. On standard semantic accounts of (3), utterances of (3) should be associated with the content  $p$ : John went to France. While the speaker here does not believe that John went to Paris, she does believe this  $p$ . This means that the standard semantic account of (3) passes the Lying Test; since condition (B) of **LT** is satisfied, predicting that (3) is a non-lie looks to be wholly unproblematic.<sup>35</sup>

The Lying Test thus offers some initially appealing results: first, like the Cancelability Test, the Lying Test is compatible with the claim that (3) only pragmatically conveys that John went to Paris. Further, both tests

---

<sup>34</sup>This is not to claim that Lying Test will smoothly tell between the propriety of ambiguity and non-ambiguity theories in all instances. For example, the ‘privative oppositions’ described in both Zwicky & Sadock (1975) and Atlas (2005)—cases where one putative disambiguation logically entails the other—prove significantly more troublesome for the Lying Test (as they have for many other tests for ambiguity). Unfortunately, full consideration of these cases will have to wait for another occasion.

<sup>35</sup>Some variants on this case are trickier. Consider, for instance, a situation in which someone asks John’s roommate if he is at home, and John’s roommate responds by uttering (3). In fact, John is both at home and has, at some point in his life, been to France. Perhaps this utterance of (3) should qualify as a lie (I am myself uncertain). However, even if this utterance is a lie, it has little bearing on the plausibility of the Lying Test. Once more, it is crucial to recall that **LIE**, on which the Lying Test is built, only purports to offer one set of sufficient conditions for lying. If the present sort of case is a lie, it is perfectly plausible to suppose that it is a pragmatic lie.

are incompatible with the claim that the semantic content of (3) includes: that John went to Paris. Second, and again like the Cancelability Test, the Lying Test is compatible with the claim that ‘and’ is unambiguous. In contrast to the Cancelability Test, however, the Lying Test may turn out to be incompatible with the claim that ‘and’ is ambiguous, depending on how the speaker’s intentions are best cashed-out in this case. I take it that the right thing to say in each of these cases is: the semantic content of (3) does not include the claim that John went to Paris, and ‘and’ is not ambiguous between a temporal and logical reading.<sup>36</sup> These are thus good cases with which to check the calibration of the Lying Test. So far, so good: the Lying Test does at least as well as the Cancelability Test in terms of according with some widely accepted semantic theories. In fact, the Lying Test might do the Cancelability Test one better; not only is it compatible with a unified theory of ‘and’, it may be incompatible with the ambiguity-theoretic alternative.

As with any new proposal in semantic theory, it is only natural to question the propriety of employing the Lying Test as a test for semantic content. However, that test has so far withstood the most obvious objections to it: it looks well-positioned to provide us with data even where judgments on truth and falsity run out; the judgements on which it relies are not simply judgements on saying, stating, or claiming in disguise; and the test squares with antecedently plausible semantic theories in a range of cases. What remains is to demonstrate that the Lying Test can be used to deliver interesting results in some hard semantic cases. I turn now to the task of showing that it can in fact deliver such results. As mentioned at the start, the Russellian theory of definite descriptions, for instance, fails the Lying Test.

## 6 The Russellian Theory of Definite Descriptions

According to Russell, definite descriptions, such as ‘the F’, have no meaning in isolation. However, when they appear in simple declarative sentences, such

---

<sup>36</sup>I am assuming that the plausibility of these claims can be established independently of the results of either the Cancelability or Lying Tests. In the contemporary literature, both these claims are standardly assumed to be overwhelmingly plausible. To be clear: I am not attempting to give an argument for either of these claims here. Rather, I am assuming that these semantic claims are correct and using that assumption as a helpful data point for gauging the plausibility of the Lying Test.

as “The F is G,” the semantic value of the whole sentence can be represented in first-order logic as:  $\exists x(Fx \ \& \ \forall y(Fy \supset y = x) \ \& \ Gx)$  (Russell 1905, p. 490).<sup>37</sup> In other words, for “The F is G” to be true, there must be one and only one object that satisfies the descriptive content (here, ‘F’), and that object must also satisfy the predicative content (here ‘G’).

Resistance to the Russellian theory has generally been motivated by the claim that it fails to make accurate truth-predictions (that is, it fails to match our judgments on particular cases). Responses on behalf of the Russellian theory have centered on contesting both the consistency and the significance of these data. Let me briefly outline this dialectic.

Donnellan (1966) introduces several cases where, to his mind, the Russellian theory predicts the wrong truth-conditions.<sup>38</sup> Each of these cases involves a situation in which the speaker is mistaken about some aspect of the environment—namely, about the applicability of the descriptive content she uses to the object she intends to talk about. In the paradigm case, Donnellan’s speaker asks “Who is the man drinking the martini?” while trying to learn something about a man who (unbeknownst to the speaker) happens to be drinking water from a martini glass (Donnellan 1966, p. 287). Following Kripke’s later discussion, we can modify the case so as to avoid having to grapple with the semantics of questions: “The man over there drinking champagne is happy tonight” (Kripke 1977, p. 256). Donnellan’s claim is that this utterance can be true even when the man that the speaker has in mind is drinking seltzer-water rather than champagne. All that matters, according to Donnellan, is whether he is happy tonight.<sup>39</sup>

Kripke (1977) defends the Russellian theory against this challenge by invoking Grice’s distinction between ‘speaker meaning’ and ‘semantic meaning’

---

<sup>37</sup>More recent Russellians, such as Neale (1990), have held onto many of Russell’s basic commitments while treating ‘the F’ as a generalized quantifier—usually represented as ‘[the  $x$ :  $Fx$ ]’ (Neale 1990, p. 45). This has the dual advantages of making it easier to incorporate definite descriptions into a compositional semantics and of allowing ‘the’ to be integrated into Barwise and Cooper (1981)’s generalized quantifier theory.

<sup>38</sup>Donnellan was of course by no means the first to criticize the predictions of Russell’s theory. While I omit discussion of it in order to streamline the dialectic, Strawson (1950) also argues at some length that the Russellian theory makes inaccurate predictions regarding truth

<sup>39</sup>While I take this to be an accurate rendition of the standard interpretation of Donnellan’s comments, I do not mean to be endorsing this understanding of Donnellan’s paper. In fact, I take it to be quite plausible that some subtleties of Donnellan’s argument are overlooked on this standard interpretation.

(or, for Grice, ‘what is said, strictly speaking’).<sup>40</sup> Kripke claims that while definite descriptions possess a constant, unitary semantic meaning (given by the Russellian theory) they can nonetheless be felicitously used by speakers with an intention to communicate something about an entity that fails to match the descriptive content of that description. In such cases, like the one Donnellan points to, we can be misled by our tendency to attend primarily to speaker meaning into thinking that the truth-conditions of what the speaker meant just are the truth-conditions of the utterance itself. But, as Kripke points out, it is perfectly coherent to distinguish between these two types of meaning, to claim that truth-conditions are always properly given by semantic meaning, and to offer an alternative explanation of our common intuitions on these cases—specifically, that they are driven by this common psychological tendency rather than by the underlying semantic facts (Ibid., pp. 262–64). So, whereas Donnellan tries to stretch semantic theory to accommodate our intuitions on certain cases, Kripke demonstrates how, by appealing to Gricean pragmatic theory, we can predict those intuitions while hewing to a cleaner semantic picture.<sup>41</sup> I will not evaluate the relative merits of these views here; suffice it to say that the Kripkean defense of the Russellian theory of definite descriptions has found a sympathetic audience in the last thirty-odd years of philosophy of language.

## 7 Running the Test

Now we turn to the task of applying the Lying Test to the Russellian theory. Let us begin with a variation on Donnellan’s classic example:

**MARTINI DRINKER**      Mr. X and Ms. Y are at a party.  
In their shared visual field, there is one and only one man drinking  
from a martini glass. There is another man behind him drinking

---

<sup>40</sup>Technically, Kripke distinguishes between ‘speaker reference’ and ‘semantic reference’. However, since we will be interested in the semantic contents of whole sentences, not just their (possibly) referential sub-parts, we will do better to think about this distinction in terms of the more general notion of meaning rather than the more specific notion of reference.

<sup>41</sup>To be clear: Kripke explicitly fails to endorse the Russellian theory of definite descriptions, despite his spirited defense of it (Kripke 1977, p. 271). It is a complicated exegetical issue whether Kripke’s defense should itself be read as a tacit endorsement of the Russellian view; I mean to take no stand on that here.



from a flask. The man drinking from the glass is drinking water, whereas the man drinking from the flask is actually drinking a martini out of it. Mr. X knows this. Ms. Y does not. The man drinking water from the martini glass also happens to be a friend of Mr. X's, whereas Mr. X rather dislikes the man drinking from the flask. Mr. X wants to start a conversation about the man drinking water from a martini glass, so he says "The man drinking the martini is a friend of mine." Mr. X reasons, correctly, that this utterance will be an effective way to prompt Ms. Y to focus on the person that Mr. X wants to talk about. What's more, Mr. X and Ms. Y are in a context where it is not at all significant whether or not Mr. X's friend is actually drinking a martini (perhaps the friend has even sworn Mr. X to secrecy regarding his habit of drinking water out of martini glasses at parties).<sup>42</sup>

Has Mr. X lied to Ms. Y? It seems to me that he has not.<sup>43</sup> This means that clause (A) of the Lying Test goes unsatisfied. Let us therefore turn to an examination of whether one or both of (B) and (C) might be satisfied.

---

<sup>42</sup>If the reader finds the above example forced, or has trouble eliciting clear judgments about it, she is welcome to consider instead the closely related question-answer pair uttered in relevantly similar circumstances: (a) Ms. Y: "Who is the man drinking the martini?"; (b) Mr. X: "The man drinking the martini is a friend of mine." Since Russellians are committed to a unified semantic account of definite descriptions, surrounding discourse context (such as being part of a question-answer pair) should have no effect on the semantic content of the utterance in question, i.e. Mr. X's utterance. Embedding this utterance in a discourse context *might* have an effect on the pragmatics of the utterance, according to the Russellian, but this is a separate matter. (In fact, it seems to me that the total communicated content of Mr. X's utterance is identical in both cases. If that's right, then it offers us further evidence, given that Mr. X's utterance-tokens are identical, that the semantic content of his utterance is identical across these two situations.)

<sup>43</sup>While this judgment strikes me as both clear and correct, some readers may be tempted to question it. One natural suggestion, I take it, would be that because Mr. X's utterance is only an innocuous falsehood, it fails to register as a lie despite actually being one. Mr. X's utterance would thus constitute a sort of white lie. But note that, ordinarily, white lies are easily recognizable as such; if this case constitutes a white lie, it is a strangely opaque one. The most obvious explanation for this opacity, I take it, is that Mr. X's utterance isn't a lie at all. Similarly, one might worry that, since we don't find Mr. X's utterance morally objectionable, we hesitate to call it a lie. But note that, to the extent that we sometimes find lies to be necessary, nice, or even admirable, we can still recognize them, on reflection, as lies. This yields an important disanalogy to the present case—a disanalogy that gives us at least some reason to think that Mr. X simply hasn't lied.

First, consider (B). The Russellian predicts the following as the relevant  $p$ : there is a unique man drinking a martini (at this party, or in our shared visual field, or whatever), and he is a friend of Mr. X's.<sup>44</sup> Given the situation as described in **MARTINI DRINKER**, Mr. X clearly believes this  $p$  to be false. That is, while Mr. X does believe there to be a unique martini drinker at the party (i.e. the man with the flask), he does not believe that person to be his friend. (B) thus also goes unsatisfied.

Having ruled out either (A) or (B) being satisfied with respect to **MARTINI DRINKER**, the only option left for the Russellian is to claim that (C) is satisfied. In other words, the Russellian must claim that Mr. X does not intend to deceive Ms. Y with respect to  $p$ . Given that this is a fairly natural thing to say about the case, we might wonder how exactly a problem is supposed to arise for the Russellian here. To be clear: I agree that no deceptive intentions are operative in **MARTINI DRINKER**. Surprisingly though, Russellians are forced to reject this claim, at least if they want to follow Kripke in invoking Gricean pragmatics in order to defend their theory against the original Donnellan case.

To see this, consider how the Russellian-cum-Gricean will describe **MARTINI DRINKER**: Mr. X intends for Ms. Y to come to believe, on the basis of his utterance, that the man drinking from the martini glass is a friend of his. How does Mr. X intend for that to happen? Well, first Ms. Y is supposed to update her beliefs with  $p$ , the semantic content of Mr. X's utterance. Then, since Ms. Y also believes that the man visibly drinking from a martini glass is the only relevant man drinking a martini, she will further update her beliefs with  $q$ : *that man* (the man drinking from the martini glass) is Mr. X's friend. Since Mr. X expects for Ms. Y will update her beliefs in this way (in fact, it's precisely what he intends for her to do), it is this second proposition,  $q$ , that he ultimately intends to communicate.

Note that successful communication here, at least on the Russellian-cum-Gricean analysis, relies on a very specific inference: in order for Mr. X's statement to seem relevant to the situation at hand, Ms. Y has to infer which of the men in their shared perceptual vicinity Mr. X was trying to

---

<sup>44</sup>For the purposes of this paper, I am going to assume that the Russellian has some story about domain restriction—that is, about how context restricts the domain of quantification to people at this party, or in our shared visual field, or whatever. Note that this makes the Russellian view a much more serious opponent, since it is almost certain that the property of unique-martini-drinkerness-simpliciter is never instantiated. For discussion, see Neale (1990) and Reimer (1992) and (1998).

talk about. In other words, what Mr. X speaker means is to be conveyed by way of a ‘relevance implicature’. On the standard Gricean analysis, relevance implicatures must be ‘worked out’ on the basis of semantic content (Grice 1989b, p. 31). All of this entails that Mr. X must have some very specific intentions with regard to  $p$ : namely, he must intend for Ms. Y to update her beliefs with it, and then to engage in some further reasoning on the basis of this new belief that will ultimately lead her to believe  $q$ . Best I can tell, that’s just to say that Mr. X intends to deceive Ms. Y with regard to  $p$ ; he intends for Ms. Y to come to believe  $p$  on the basis of his utterance, and despite the fact that he himself believes  $p$  to be false.<sup>45</sup>

This last inference on which I have relied—that, if a speaker intends for her listener to come to believe  $p$  on the basis of her utterance, and if she also believes  $p$  to be false, then the speaker intends to deceive her listener with respect to  $p$ —may strike some as worrisome. There are two main reasons to question the inference: first, there is a class of apparent counterexamples to the principle. If someone holds a gun to my head and tells me that she will fire it unless I get you to believe some  $p$  (which I believe to be false) on the basis of my utterance (perhaps over the telephone, so you cannot see the situation), it is not at all clear that I intend to deceive you when I try to get you to form this belief. The problem with this worry, from our perspective at least, is that nothing in **MARTINI DRINKER** parallels the gunman in this case; in **MARTINI DRINKER**, Mr. X is free to say whatever he wants. Second, it might seem relevant that Mr. X’s intention regarding  $p$  is aimed at a non-deceptive end: getting Ms. Y to believe  $q$ . But, again, Mr. X could have just said ‘Q’, i.e. “That man is a friend of mine.”<sup>46</sup> Supposing

---

<sup>45</sup>Recently, Bach (2006a) has claimed that Grice’s account of implicatures should be understood as a ‘rational reconstruction’ rather than as a psychological theory or cognitive model. Very well. In that case, we might well doubt that Mr. X has anything like the intentions I have attributed to him in **MARTINI DRINKER**. But note that it is perfectly compatible with Bach’s claim to think that we ought to attribute linguistic intentions to speakers not on the basis of their actual psychology, but rather on the basis of a rational reconstruction of their psychological states. In fact, much of what Griceans like Bach commonly claim regarding speaker meaning—in particular, their requirement that, in order to mean something, speakers must have highly complex ‘reflexive intentions’—surely presupposes the latter way of assigning intentions to speakers (cf. Grice 1989d,e Bach 1992, 2005, Neale 1992). So it is unclear why, on Bach’s preferred understanding of the Gricean model, different rules should apply in the present case.

<sup>46</sup>Or, alternatively: “The man drinking from the martini glass is a friend of mine”; “You see that man with the martini glass over there—he’s a friend of mine.”

that the Russellian cum Gricean is right and Mr. X intends to get Ms. Y to believe both  $q$  and  $p$ , why should we think this gets Mr. X off the hook for intending that Ms. Y believe  $p$ ? After all, he has plenty of other options for getting Ms. Y to believe just  $q$ .<sup>47</sup>

If the reasoning here is sound, then the Russellian cannot consistently claim that condition (C) is satisfied. That, in turn, means that the Russellian theory of definite descriptions fails the Lying Test, at least once it is supplemented by a Gricean theory of pragmatics. In other words, standard contemporary Russellianism appears to be incompatible with either the suggestion that (i) semantic content is among the things to which we indefeasibly commit ourselves by means of standard, non-ironic utterances of declarative sentences, or (ii) the contents to which we indefeasibly commit ourselves by means of our utterances play a central role in determining whether or not we have lied. However, several responses are available to the Russellian.

## 8 Objections and Replies

First, the Russellian might claim that the speaker is only ‘making as if to say P’, rather than actually saying P outright (both of these ‘sayings’ being meant in the Gricean sense, where this entails that the speaker also means  $p$ ).<sup>48</sup> The idea here would be that the speaker is only pretending to put her utterance forward in a literal manner in order to convey some other content in a non-literal manner (Grice 1989a, p. 41). The speaker should thus not be construed as committing herself to the content standardly expressed by her utterance. Paradigm cases of making as if to say include both irony and

---

<sup>47</sup>Consider as well what would result if we were to give up on the above inference and instead accept that: a speaker needn’t intend to deceive her listener when she intends for that listener to come to believe  $p$  on the basis of her utterance, where the speaker further believes  $p$  to be false—so long as the speaker further intends for the listener’s coming to believe that  $p$  to prompt her to accept  $q$ , which the speaker does in fact believe to be true. Without some further restrictions on when intentions like these count as non-deceptive, this would allow for instance that a witness who testifies that she saw the defendant commit the crime might not intend to deceive the jury, even when she did not in fact see the defendant commit the crime—so long as she also believes the defendant to be guilty and ultimately intends for her testimony to prompt the jury to come to believe the defendant’s guilt. In other words, giving up on the inference assumed in the main text risks embracing an overly-forgiving standard of deception, one that threatens to mistakenly classify some seemingly paradigmatic cases of deceptive behavior as being non-deceptive.

<sup>48</sup>Thanks to both Daniel Harris and Stephen Neale for pressing this objection.

sarcasm, and in such cases there is no doubt that the speaker has no intention of being taken at her word. Yet she *does* intend to convey something to the listener—namely, the ironic or sarcastic content.

While tempting, there is good reason to refrain from interpreting **MARTINI DRINKER** as a case of making as if to say: namely, the speaker fails to give any indication that his utterance is not to be taken at face-value here. Nor does the context somehow serve to make clear that the utterance is intended to be taken non-literally. This stands in striking contrast to standard cases of irony and sarcasm, where speakers typically make use of a particular intonational pattern or even a distinct syntactic form (as with ‘like’-sarcasm, cf. Camp 2012) to signal to listeners their intention not to be taken literally. Exceptions to this pattern seem to require a very particular sort of context, one that makes the speaker’s utterance, interpreted literally, plainly absurd.<sup>49</sup> Since neither of these conditions obtain in **MARTINI DRINKER**, we lack any substantive reason to suppose that the speaker is only making as if to say in this case. What’s more, in **MARTINI DRINKER**, the speaker may well be indifferent to whether or not the listener ultimately believes *p*, so long as she believes *q*. This stands in contrast to standard cases of sarcasm and irony, where speakers specifically intend that their listeners *not* come to believe the literal meanings of their utterances.<sup>50</sup>

Second, the Russellian might distinguish between intended actions and the knowable side-effects of such actions. With this distinction in hand, she might then claim that the listener’s updating her beliefs with *p* is only a knowable side-effect of Mr. X’s intended act of communicating *q*—and, thus, that Mr. X had no intentions at all, deceptive or otherwise, regarding *p*. Such a distinction is itself not at all unwarranted: consider that when

<sup>49</sup>Deadpan sarcasm would appear to be a striking example of this possibility.

<sup>50</sup>The Russellian might try weakening this reply by suggesting that, even if Mr. X wasn’t making as if to say specifically, he still meant his utterance non-literally in some other manner. Thus, Mr. X once more fails to indefeasibly commit himself to *p*, in virtue of this non-literal use of the definite description. However, we should be wary of this response for reasons that largely mirror those above: while there are certainly any number of ways to use language non-literally—e.g. joking, rehearsing, translating, or even just speaking loosely—we have no independent reason to believe that referential uses of definite descriptions (be they semantically significant or not) are to be counted among these non-literal uses. Without some further reason to believe that referential uses of definite descriptions, which are both perfectly common and which appear to be aimed at the straightforward transmission of object-directed thoughts, should in fact be understood as non-literal uses akin to joking or loose talk, this proposal seems both under-motivated and ad hoc.

Mr. X intends to open a door, he knows that he will inevitably move some air molecules. This hardly means that every time Mr. X forms an intention to open a door, he thereby forms a distinct intention to move some air molecules. However, there is a striking disanalogy between the cases: when Mr. X intends to open the door, any way that he might do so will result in the movement of some air molecules. In attempting to communicate  $q$ , Mr. X has plenty of options for getting Ms. Y to believe  $q$  without requiring that she first update her beliefs with  $p$ . This difference strikes me as sufficient to motivate the claim that, in **MARTINI DRINKER**, Mr. X does intend for Ms. Y to update her beliefs with  $p$ , whereas he need not intend to move some air molecules whenever he opens a door.

Third, the Russellian might claim that, in running the Lying Test, we should consider not semantic content, but rather the content that the speaker *meant*, in a Gricean sense. Sometimes, speakers mean the semantic content of their utterances, but sometimes they do not. This suggestion yields the correct prediction with regard to **MARTINI DRINKER**, since the speaker certainly meant  $q$ , but plausibly did not mean  $p$ . But consider instead:

**CAR SALESMAN**            Mr. X is a used car salesman trying to sell Ms. Y a particular car. Ms. Y asks Mr. X “Does this car have any trouble with overheating?” In fact, the car Mr. X is trying to sell Ms. Y overheats regularly and Mr. X is well aware of this. However, instead of responding “Yes,” Mr. X says “Look, I’ve driven this car across the Mojave Desert without a problem.” As it happens, Mr. X has done just this—though that was five years ago, and well before the car started overheating.<sup>51</sup>

While what Mr. X says here is highly misleading, it strikes me as incorrect to claim that he has lied.<sup>52</sup> Yet this is a paradigm case of the speaker meaning something other than what he, strictly speaking, said. Mr. X seems to

---

<sup>51</sup>The example is adapted from Carson (2006, pp. 284-85). The case also parallels the one discussed in *Bronston v. United States*, in which the Supreme Court found that the analogue of Mr. X had not in fact perjured himself. For discussion of several related examples, see Adler (1997).

<sup>52</sup>Note that if **CAR SALESMAN** constitutes a lie, then pretty much any deceptive use of language should as well. This would therefore threaten to efface the distinction between lies and merely deceptive uses of language that underlies both the Supreme Court’s reasoning in *Bronston v. United States* as well as most extant philosophical inquiries into the nature of lying.

have speaker-meant  $q^*$ : this car doesn't overheat. And, in fact, Mr. X very likely succeeded in communicating  $q^*$ . So, if speaker meaning is the relevant content for **LIE**, then Mr. X should be lying here. But he is not.

It is, however, open to the Russellian to modify this proposal slightly and claim that the proper content for running the Lying Test is indeed semantic content—though only when the speaker also means that content. This move effectively evades the objection derived from **CAR SALESMAN**, since the meant content there is uncontroversially pragmatic. A position such as this one would seem to be in the spirit of the view advocated in Soames (2009), according to which speakers needn't 'assert' (in our terminology, roughly, 'mean') the semantic content associated with their utterances, even when they are speaking "literally and unmetaphorically," i.e. even when they are not merely making as if to say (p. 281).

Note that adopting this modified version of Russellianism requires giving up on the basic Gricean assumption that semantic content is something to which we indefeasibly commit ourselves by means of standard, literal, non-ironic utterances of declarative sentences.<sup>53</sup> This, in turn, forces us to predict that, in certain circumstances, (2') can actually be uttered felicitously:

- (2') Danny and Laura had a kid and got married, but Danny and Laura did not get married.

Specifically, so long as the speaker doesn't commit herself to the first clause in (2'), her utterance of this sentence should be perfectly felicitous. I take this claim to be highly implausible. (2'), it seems to me, can *never* be uttered felicitously. The natural explanation for this would seem to be that uses of declarative sentences indefeasibly commit us to their semantic contents, and that committing oneself to two incommensurate contents simultaneously results in infelicity. So following Soames suggestion and giving up on the basic Gricean assumption that semantic content generates indefeasible commitments might help save Russellianism, but that help comes at a steep cost.

Fourth, the Russellian might contend that, in fact, we only consider utterances to be candidate-lies if their semantic content is itself *relevant* to the purposes of the conversation.<sup>54</sup> But that would seem to entail that one cannot lie discourse-initially in situations where there is insufficient background context to antecedently fix the purposes of the conversation (imagine

---

<sup>53</sup>Soames in fact acknowledges this at Soames (2009, p. 281).

<sup>54</sup>Thanks to Kathryn Lindeman for pressing this objection.

two complete strangers who happen to run into each other on the street), which is false.<sup>55</sup> Moreover, this sort of defense either (i) effectively imposes a pragmatic constraint on when semantic content generates infeasible commitments, or (ii) effectively imposes a pragmatic constraint on when infeasible commitments are relevant to making an utterance a lie. But (i) amounts to a covert severing of the tight connection between semantic content and infeasible commitment assumed by both myself and by more traditional Griceans, whereas (ii) amounts to a covert severing of the tight connection between infeasible commitment and lying for which I argued at the beginning of §4. Thus, in order to avail herself of this type of defensive strategy, it would be incumbent on the Russellian to first offer a positive defense of the acceptability of either (i) or (ii).<sup>56</sup>

Fifth, and finally, the Russellian might claim that, in the case of **MARTINI DRINKER**, the speaker simply doesn't know what she's saying. Presumably, being unable to identify *p* would prevent one from having a determinate intention involving it. However, this strikes me as an extreme solution to the Russellian's impasse; it entails a thoroughgoing skepticism about our knowledge of what we say. I assume that adopting such a skeptical attitude will strike most readers as unappealing.<sup>57</sup>

---

<sup>55</sup>It is important to note that general rules of good conversational conduct—such as “say what you believe to be true”—won't alone suffice to determine standards of relevance for any particular conversation. Recognizing this, one might be tempted to try weakening this sort of relevance-constraint on lying so that the utterance need only be relevant to the purposes of at least one member of the conversation. But note that speakers can lie by intentionally uttering falsehoods for no particular purpose (just to ‘mess with’ their listeners, as it were), and even when those listeners are wholly uninterested in the contents of those utterances. This weakened relevance-constraint, however, incorrectly predicts that such utterances are non-lies, since they are pertinent to no one's aims.

<sup>56</sup>A more nuanced version of this relevance-constraint might attempt to appeal to something like Potts (2003)'s distinction between *at-issue* and *not-at-issue* content. The idea, presumably, would be that the descriptive material in a definite description is generally not-at-issue content. Rather, it is used to guide or comment on the at-issue content expressed by utterance (Potts 2003, pp. 3-4). However, claiming that such content is typically not-at-issue entails giving up on the Russellian theory of definite descriptions—since not-at-issue content, according to Potts' theory, is the wrong sort of thing to figure into the truth-conditions of utterances. This move, therefore, is not open to the Russellian.

<sup>57</sup>Diehard externalists about content might well be tempted by such a view. I will not attempt to address their concerns here, as it strikes me that the plausibility of that view should be decided on other grounds. However, I would note that such skepticism will have serious repercussions for discussions of both linguistic competence and language-learning. For elaboration, see Matthews (2006).



## 9 Conclusion

In the last three sections, I demonstrated how the Lying Test can be used to generate a novel argument against at least one venerable semantic theory. While this is unlikely to serve as the last word on Russellianism, it does represent a serious challenge for that view. Russellians have tended to appeal to Gricean pragmatics in order to explain why many people make truth judgments inconsistent with their preferred theory of descriptions. Now we have seen that the conjunction of Russellianism and Griceanism proves incompatible with taking judgments on lying seriously as a guide to semantic content. The Russellian can thus either jettison her appeal to Gricean pragmatics—leaving her without a response to Donnellan’s original challenge—or else she can contest the semantic relevance of judgments on lying. But if she opts for the latter route, we should ask in all seriousness just what judgments are left to motivate her theory. Already, she has chosen to privilege certain judgments on truth and falsity over others; now she would seem to be opting to ignore a class of relatively consistent judgments, the semantic relevance of which was motivated in a way that parallels how Grice himself suggested we tease apart what is said from what is merely implicated!<sup>58</sup>

Thus, it seems to me that we have at least provisionally vindicated the claim that the Lying Test offers a genuinely useful way of testing semantic theories. While it is undoubtedly incapable of adjudicating every hard problem in semantic theory, the Lying Test proved capable of addressing at least one long-standing debate in semantics. It may well help with others. Still, one might wonder: was it really necessary to go into such detail to develop the Lying Test? Couldn’t we have just appealed informally to intuitions on lying, as philosophers sometimes have, to make much the same point?<sup>59</sup> The short answer is “no,” and the reason is that, as we saw above, what judgments on lying indicate regarding content is importantly, and non-obviously, asymmetric. This is because it remains an open question whether we can lie with pragmatic content. What is clear is that we do in fact lie when we

---

<sup>58</sup>Note that this is not meant as an argument in favor of the traditional alternative to Russellianism, i.e. Donnellianism, or the claim that definite descriptions are ambiguous between referring and non-referring uses. In fact, I suspect that the semantic picture is simply more complicated than either of the traditional views allows for. For further discussion of how we can lie with definite descriptions, see Stokke (2013).

<sup>59</sup>Such appeals have a long, if only very occasional, history in philosophy of language. See, for instance, Frege (1892/1997, p. 160).

intend to deceive our listeners about the semantic contents we express. We should hardly expect this asymmetry to have been implicitly recognized, let alone exploited, in informal appeals to judgments on lying in the course of actually doing semantics.

I conclude with what I take to be the positive lesson of this inquiry: it is at least possible that we have yet to exhaust the criteria by which semantic theories ought to be judged. If we have not, then extant theories of semantic content may well incur costs or exhibit virtues that we have yet to realize, let alone focus on. If I am right, then a semantic theory's ability to explain certain judgments on lying is one such unexpected virtue. My hope, however, is that there will prove to be others—and that attending to these features of language use in sufficient detail will ultimately improve not just our semantic theories themselves, but also our understanding of what it is that those theories are aiming to explain.<sup>60</sup>

## References

Åkerman, Jonas. 2014. Infelicitous cancellation: The explicit cancellability test for conversational implicature revisited. Forthcoming in *Australasian Journal of Philosophy*.

---

<sup>60</sup>Particular thanks are due to Sam Cumming and Mark Greenberg for suffering through a very early draft of this paper, and for encouraging me to continue working on it. For extensive discussion and feedback over many years, thanks are due as well to: Zed Adams, Jonas Åkerman, Matthew Babb, Mel Bervoets, Ray Buchanan, Alexis Burgess, Arudra Burra, Elisabeth Camp, Alejandro Pérez Carballo, Robyn Carston, Jonathan Cohen, Shamik Dasgupta, Kenny Easwaran, Andy Egan, Don Fallis, Ashley Feinsinger, Marina Folescu, David Friedell, Andrew Jewell, David Kaplan, Nathaniel Hansen, Daniel Harris, Jeffrey Helmreich, Sarah Jansen, Rachel Johnson, Brent Kiou, Dustin Locke, Kathryn Lindeman, Chauncey Maher, Jeff Mahon, Rachel McKinney, Lisa Miracchi, Matthew Moss, Stephen Neale, Paul Nichols, Daniel Nolan, Eileen Nutting, Tyke Nunez, Alexi Patsorous, Jessica Pepp, David Plunkett, Gabe Rabin, Jennifer Saul, John Scott-Railton, Ted Shear, Andreas Stokke, Alex Radulescu, Marga Reimer, Andrew Reisner, David Sosa, Tom Ward, and Tamar Weber. Thanks also to my audiences at the 2011 CUNY Graduate Philosophy Conference, the 2011 Pacific APA, and the 2011 Berkeley-Stanford-Davis Graduate Philosophy Conference. Finally, thanks to my two anonymous reviewers at *Mind & Language* for their very helpful comments. An earlier version of this paper is cited in Saul (2012) as 'Lying and Content'. The differences between that draft and this one are immaterial for Saul's purposes.

- Adler, Jonathan. 1997. Lying, deceiving, or falsely implicating. *Journal of Philosophy* 94(9). 435–452.
- Atlas, Jay. 2005. *Logic, meaning, and conversation*. Oxford: Oxford University Press.
- Augustine. 1952. Lying. In Roy Deferrari (ed.), *Treatises on various subjects*, 53–110. New York: Fathers of the Church.
- Bach, Kent. 1992. Intentions and demonstrations. *Analysis* 52. 140–146.
- Bach, Kent. 2001a. Seemingly semantic intuitions. In J. K. Campbell, M. O'Rourke & D. Shier (eds.), *Meaning and truth: Investigations in philosophical semantics*, 21–33. New York: Seven Bridges Press.
- Bach, Kent. 2001b. You don't say. *Synthese* 128. 15–44.
- Bach, Kent. 2005. Context *ex Machina*. In Zoltán Szabó (ed.), *Semantics versus pragmatics*, 15–44. Oxford: Oxford University Press.
- Bach, Kent. 2006a. The top ten misconceptions about implicatures. In Betty Birner & Gregory Ward (eds.), *Drawing the boundaries of meaning: Neo-Gricean studies in pragmatics and semantics in honor of Laurence R. Horn*, 21–30. John Benjamins.
- Bach, Kent. 2006b. What does it take to refer? In Ernest Lepore & Barry Smith (eds.), *The Oxford handbook of philosophy of language*, 516–554. Oxford University Press.
- Barwise, K. Jon & Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(1). 159–219.
- Blome-Tillman, Michael. 2008. Conversational implicature and the cancellability test. *Analysis* 68(2). 156–160.
- Blome-Tillman, Michael. 2013. Conversational implicatures (and how to spot them). *Philosophy Compass* 8(2). 170–185.
- Borg, Emma. 2004. *Minimal semantics*. Oxford: Oxford University Press.
- Borg, Emma. 2012. *Pursuing meaning*. Oxford: Oxford University Press.

- Burge, Tyler. 1986. On Davidson's 'saying that'. In Ernest Lepore (ed.), *Truth and interpretation*, 190–208. Oxford: Basil Blackwell.
- Camp, Elisabeth. 2007. Prudent semantics meets wanton speech act pluralism. In G. Preyer & G. Peter (eds.), *New essays on semantics and pragmatics*, 194–213. Oxford: Oxford University Press.
- Camp, Elisabeth. 2012. Sarcasm, pretense, and the semantics/pragmatics distinction. *Noûs* 46(4). 587–634.
- Cappelen, Herman & John Hawthorne. 2009. *Relativism and monadic truth*. Oxford: Oxford University Press.
- Cappelen, Herman & Ernest Lepore. 1997. On an alleged connection between indirect speech and the theory of meaning. *Mind and Language* 12(3–4). 278–296.
- Cappelen, Herman & Ernest Lepore. 2005. *Insensitive semantics: A defense of semantic minimalism and speech act pluralism*. Oxford: Blackwell.
- Carson, Thomas L. 2006. The definition of lying. *Noûs* 40. 284–306.
- Carston, Robyn. 2002. *Thoughts and utterances*. Oxford: Blackwell.
- Davidson, Donald. 1967. Truth and meaning. *Synthese* 17. 304–323.
- Davidson, Donald. 1968. On saying that. *Synthese* 19(1–2). 130–146.
- Davidson, Donald. 1976. Reply to Foster. In Gareth Evans & John McDowell (eds.), *Truth and meaning: Essays in semantics*, 33–41. Oxford: Clarendon Press.
- Davis, Wayne A. 1998. *Implicature: Intention, convention, and principle in the failure of Gricean theory*. Cambridge: Cambridge University Press.
- Donnellan, Keith. 1966. Reference and definite descriptions. *Philosophical Review* 75. 281–304.
- Fallis, Don. 2009. What is lying? *Journal of Philosophy* 106(1). 29–56.
- Fallis, Don. 2010. Lying and deception. *Philosopher's Imprint* 10(11). 1–22.

- Frege, Gottlob. 1892/1997. Über sinn und bedeutung. In Michael Beaney (ed.), *The Frege reader*, 151–171. Oxford: Blackwell. Max Black, trans.
- Grice, H. Paul. 1989a. Further notes on logic and conversation. In *Studies in the way of words*, 41–57. Cambridge: Harvard University Press.
- Grice, H. Paul. 1989b. Logic and conversation. In *Studies in the way of words*, 22–40. Cambridge: Harvard University Press.
- Grice, H. Paul. 1989c. Presupposition and conversational implicature. In *Studies in the way of words*, 269–282. Cambridge: Harvard University Press.
- Grice, H. Paul. 1989d. Utterer's meaning, sentence-meaning, and word-meaning. In *Studies in the way of words*, 117–143. Cambridge: Harvard University Press.
- Grice, Paul. 1989e. Meaning. In *Studies in the way of words*, 213–223. Cambridge: Harvard University Press.
- Heim, Irene & Angelika Kratzer. 1997. *Semantics in generative grammar*. Oxford: Blackwell Publishers.
- Horn, Laurence R. 2004. Implicature. In Laurence R. Horn & Gregory Ward (eds.), *The handbook of pragmatics*, 3–28. Oxford: Blackwell.
- Kaplan, David. 1989. Demonstratives. In Joseph Almog, John Perry & Howard Wettstein (eds.), *Themes from Kaplan*, 481–563. Oxford: Oxford University Press.
- King, Jeffrey C. & Jason Stanley. 2005. Semantics, pragmatics, and the role of semantic content. In Zoltán Szabó (ed.), *Semantics versus pragmatics*, 111–164. Oxford: Oxford University Press.
- Kripke, Saul. 1977. Speaker's reference and semantic reference. In Peter French, Thomas Uehling & Howard Wettstein (eds.), *Midwest studies in philosophy, vol. 2*, 255–276. Minneapolis: University of Minnesota Press.
- Lackey, Jennifer. 2013. Lies and deception: An unhappy divorce. *Analysis* 73(2). 236–248.

- Matthews, Robert. 2006. Knowledge of language and linguistic competence. *Philosophical Issues* 16. 200–220.
- McDowell, John. 1987. In defense of modesty. In B. M. Taylor (ed.), *Michael Dummett*, 59–80. Dordrecht: Marinus Nijhoff.
- Michaelson, Eliot. 2014. Shifty characters. *Philosophical Studies* 167(3). 519–540.
- Montague, Richard. 1970. English as a formal language. In Bruno Visentini et al. (ed.), *Linguaggi nella società e nella tecnica*, 189–224. Milan: Edizioni di Comunità.
- Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In Jaakko Hintikka, Julius M.E. Moravcsik & Patrick Suppes (eds.), *Approaches to natural language: Proceedings of the 1970 Stanford workshop on grammar and semantics*, 221–242. Dordrecht, Holland: D. Reidel Publishing Co.
- Neale, Stephen. 1990. *Descriptions*. Cambridge: MIT Press.
- Neale, Stephen. 1992. Paul Grice and the philosophy of language. *Linguistics and Philosophy* 15. 509–559.
- Partee, Barbara H. 1975. Opacity and scope. In Milton K. Munitz & Peter Unger (eds.), *Semantics and philosophy*, 81–101. New York: Academic Press.
- Partee, Barbara H. 1976. Some transformational extensions of Montague Grammar. In Barbara H. Partee (ed.), *Montague grammar*, 51–76. New York: Academic Press.
- Platts, Mark. 1979. *Ways of meaning*. London: Routledge and Kegan Paul.
- Plunkett, David. 2015. The role of metalinguistic negotiation in philosophical disputes and how it matters for thinking about the methodology of philosophy. Manuscript, Dartmouth College.
- Plunkett, David & Tim Sundell. 2013. Disagreement and the semantics of normative and evaluative terms. *Philosopher's Imprint* 13(23). 1–37.

- Potts, Christopher. 2003. *The logic of conventional implicatures*: University of California, Santa Cruz dissertation.
- Recanati, François. 2012. *Truth-conditional pragmatics*. Oxford: Oxford University Press.
- Reimer, Marga. 1992. Incomplete descriptions. *Erkenntnis* 37. 347–363.
- Reimer, Marga. 1998. Quantification and context. *Linguistics and Philosophy* 21. 95–115.
- Russell, Bertrand. 1905. On denoting. *Mind* 14. 479–493.
- Sadock, Jerrold M. 1978. On testing for conversational implicature. In Peter Cole (ed.), *Syntax and semantics 9: Pragmatics*, 281–97. New York: Academic Press.
- Salmon, Nathan. 2004. The good, the bad, and the ugly. In Marga Reimer & Anne Bezuidenhout (eds.), *Descriptions and beyond*, 230–260. Oxford: Oxford University Press.
- Salmon, Nathan. 2005. Two conceptions of semantics. In Zoltán Szabó (ed.), *Semantics versus pragmatics*, chap. 7, 317–328. Oxford: Oxford University Press.
- Saul, Jennifer. 2012. *Lying, misleading, and what is said: An exploration in philosophy of language and in ethics*. Oxford: Oxford University Press.
- Soames, Scott. 2009. The gap between meaning and asserting. In *Philosophical essays*, vol. 1, Princeton: Princeton University Press.
- Solan, Lawrence M. & Peter M. Tiersma. 2005. *Speaking of crime: The language of criminal justice*. Chicago: University of Chicago Press.
- Sorensen, Roy. 2007. Bald-faced lies! *Pacific Philosophical Quarterly* 88. 251–264.
- Sperber, Dan & Deirdre Wilson. 1986. *Relevance: Communication and cognition*. Blackwell Publishers.
- Stokke, Andreas. 2013. Lying and asserting. *Journal of Philosophy* 110(1). 33–60.

- Strawson, Peter F. 1950. On referring. *Mind* 59. 320–344.
- Strawson, Peter F. 1952. *Introduction to logical theory*. London: Methuen.
- Szabó, Zoltán G. 2006. The distinction between semantics and pragmatics. In Ernest Lepore & Barry Smith (eds.), *The Oxford handbook of philosophy of language*, 361–390. Oxford: Oxford University Press.
- Thomason, Richmond H. 1990. Accommodation, meaning, and implicature: Interdisciplinary foundations for pragmatics. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in communication*, 325–363. Cambridge: MIT Press.
- Ziff, Paul. 1972. What is said. In Gilbert H. Harman & Donald Davidson (eds.), *Semantics of natural language*, 709–721. Dordrecht: Reidel.
- Zwicky, Arnold M. & Jerrold M. Sadock. 1975. Ambiguity tests and how to fail them. In J. Kimball (ed.), *Syntax and semantics 4*, 1–36. New York: Academic Press.