Representation without Thought

Confusion, Reference, and Communication

Elmar Geir Unnsteinsson

Thesis submitted to

The City University of New York, Graduate Center

in partial fulfilment for the award of the degree of

DOCTOR OF PHILOSOPHY

April 2015

When I first read it, I thought it was an exercise in irony. Then a very skilful parody of a certain attitude. Then I realised it was serious—it was at the moment I searched my memory and rooted out certain fantasies of my own. But what seemed to me important was that it could be read as parody, irony or seriously. It seems to me this fact is another expression of the fragmentation of everything, ... linked with what I feel to be true about language, the thinning of language against the density of our experience.

DORIS LESSING

Preface

There appears to be a never-ending list of questions and issues about language, meaning, and representation which are both intrinsically worthy of investigation and, at present, seemingly intractable, not to say mystifying. I do not, however, endorse Chomsky's well-known edict that one ought to put mysteries to rest and focus only on allegedly solvable problems. Such a position tends to obscure the larger, more foundational concerns that should be driving our inquiry in the first place. Still, I do not wish to jettison the distinction between so-called mysteries and problems, only to propose that tackling the latter always requires acknowledgment of the former.

The foundational questions, which got me interested in philosophy of language to begin with, include such luminaries as, What is linguistic meaning? How are semantic facts grounded in other more basic facts? How is it that, by and large, we can communicate highly complicated contents so effortlessly and efficiently by the use of language? How and why did natural languages evolve? How are they learned?

One would think that the philosopher is charged with keeping these questions alive and relevant, if anyone is. A major lesson of this dissertation, however, is that even the philosophers have lost touch with foundational concerns in the study of language. I argue that philosophers of language have pursued semantic puzzles—specifically ones involving speakers who are confused about the identity of some object—as if such a practice needed no justification at all. But my point here is not entirely negative: solutions to the puzzles in question just need to be explicitly and strongly related to more foundational issues if they are to be evaluated in terms of widely accepted theoretical virtues like descriptive and explanatory adequacy.

I have been helped along the way by suggestions and encouragement from a great number of people and I am pleased to have the opportunity to thank them here. Thanks to Aðalsteinn Hákonarson, Ásgeir Berg Matthíasson, Elisabeth Camp, Robyn Carston, David Chalmers, Bill Child, Lars Dänzer, Dan Dennett, Matti Eklund, Cressida Gaukroger, Geir Þórarinsson, Peter Godfrey-Smith, Dan

Contents

Introduction

A fundamental question in the philosophy of language is, 'What determines the reference of a singular term on an occasion of utterance?' I argue that the best answer is that reference is determined by the speaker's referential intention. Thus, if the speaker's plan in uttering or inscribing a singular term—say a name or indexical—on a given occasion is to direct the hearer's or reader's attention to a particular object *o* then it follows, simply, that *o* is what the speaker refers to on that occasion. Referential intentions are defined in terms of Gricean communicative intentions more generally, i.e. intentions to have specific cognitive effects on the hearer by way, partly, of the hearer's recognition of that very intention.

Although the answer is simple, intuitive and, most importantly, an integral part of a very successful theory of meaning and communication, it faces significant problems and objections. One kind of problem or consideration is so pervasive in philosophy that it marks its own independent tradition of thought, or so I argue. The problem can be stated as follows. Speakers are often—either momentarily or persistently—confused about the identity of the objects to which they intend to refer but still, even in such cases, it seems perfectly possible that they succeed in referring to one object rather than another. I may think the keys in my hand are in fact my own, when they are not, so I am confused, but it seems like I can still, unproblematically, refer to them with some public language expression, say a demonstrative. Thus, the thought continues, reference occurs despite any internal conflict in the corresponding intention and so it must be determined by something else.[1]

This type of objection springs from what I call 'puzzle-driven' theorizing. Large swaths of philosophy are driven by puzzles, but none more so, it may seem, than philosophy of language. Since Frege theorists have, implicitly or explicitly, set themselves the task of building coherent systems of description designed specifically to capture as many possible cases of confused identity as

---

[1] For a recent statement of exactly this problem, and its relevance to intention-based theories of meaning and reference, see Jeff Speaks (forthcoming).

humanly conceivable—'Frege cases,' 'Paderewski cases,' 'Twin-Earth cases,' and so on. I take a stab at this tradition, arguing that it is not theoretically motivated and that its dismal success rate is reason for pessimism. The debate has been badly served by an emphasis on the apparent failure of substitutivity in belief-reports—but, as I hope to show, the more basic phenomenon is the mental state of confused identity.[2] Further, I show in some detail that the objection is wrong-headed and that the intention-based approach to meaning in general and to singular terms in particular stands untouched.

J. L. Austin made a distinction between doing something by mistake and doing something by accident. As I understand it, it is simply a distinction between doing something wrong because of a false belief on the one hand and because of a failure in performance on the other. When I shoot the wrong donkey because I believe, even if just momentarily, that it is identical to the donkey I was supposed to shoot, I do it by mistake. When I take aim at the right donkey but, being a terrible shot, kill the donkey standing next to it, I did the wrong thing by accident. Counterexamples to the Gricean picture of reference, and the intention-based notion of 'what is said' more generally, fall into three categories.

First, there are mistakes, where a speaker either believes falsely that a single object is two objects or believes falsely that two objects are single object and utters a singular term intended to refer to 'that' object.

Second, there are accidents. This is the category of malapropisms or speech errors more generally. The speaker has some expression as her target but makes a slip along the way and happens to utter some different expression. Theorists have used such examples, just like cases of confused identity, to argue that speakers often refer to objects or make statements which form no part of their communicative intention.

Third, there is pretense. Taking that notion very broadly, speakers often pretend to say things or refer to things, while actually meaning or intending something else. On the Gricean picture, irony, metaphor, overstatement, and figurative speech in general, are explained by holding that the speaker makes as if to say one thing—i.e. merely pretends to say it—while in fact meaning something quite different. So the speaker doesn't necessarily *say* anything. Many theorists disagree with this description and take it that figurative speech supports the view that one can say and mean something without it being part of any communicative intention. So I can say and mean that the weather is nice by uttering 'Nice weather,' while looking at the blizzard outside, trying to be

---

[2]Jennifer Saul (2007) has, to my mind, already made inroads in establishing this point by suggesting examples where substitutivity fails in sentences containing none of the standard 'opacity'-producing expressions.

ironical. It is just that I also mean that the weather is dreadful.

My focus is mostly on the first two categories, so I comment on them first. I argue that both should be treated as pragmatic performance errors in theorizing about the semantics and pragmatics of natural language. In place of puzzle-driven semantics, I propose an 'explanation-driven' approach. Specifically, philosophy of language ought to be driven by the basic task of explaining how it is possible for humans to communicate and express ideas, beliefs, desires so efficiently and successfully by uttering the particular sounds or making the particular inscriptions they do.[3] Any approach of this kind—Gricean intention-based or not—can offer compelling reasons for idealizing away from performance errors in its definition of communicative success. Of course, this does not mean that such errors are irrelevant, they just form no part of the primary explanandum of the final theory.

Pretense calls for a separate treatment altogether but, fortunately, the arguments and proposals made here are neutral on the vexed issue of figurative speech. But something akin to pretense comes up twice in the dialectic. First in the context of intentional malaproisms in Chapter 3 and secondly in the context of deliberately obfuscatory speech acts in Chapter 4. In both cases I argue that the cases in question do not provide occasion for disagreement between different explanatory theories.

The positive theory that emerges, called the 'edenic' theory of reference, places cognitive constraints on the proper performance of the speech act of referring with an expression. In Grice's own terms these are 'optimality' conditions on the mental states involved in a given piece of linguistic behavior. There are two such constraints. The first states roughly that singular term utterances fulfill their proper function only when the speaker is free from any corrupting confusion of identity. The second is a constraint on coreference and states, roughly, that in uttering a sentence containing more than one singular term the speaker must either have a coreferential or a noncoreferential intention. I argue, among other things, that the constraints characterize speakers' basic pragmatic or communicative competence.

A word or two on orientation before giving quick overviews of each chapter. The argument and methodology on display here are thoroughly theory-driven and anti-intuitionistic. As custom demands, I discuss a host of examples or cases where the intuitions of theorists are drawn in different, inconsistent directions. Many philosophers take such cases as evidence for or against a theory and argue accordingly. But, I argue, the plausibility of so proceeding varies considerably from case to case and there are particular problems involved when pragmatic

---

[3]The formulation is based on Stephen Neale's (2004: 71–72, forthcoming).

performance errors, due to confused identity, are taken as primary data. My approach is, however, to put the theory in the driver's seat and try to accept its predictions in particular cases. And if the cases can be redescribed—without undue contortions—in terms of the theory then all is well as far as that goes. But, I argue, atypical cases of this sort, considered in isolation, do not provide independent or direct evidence for the theory.

I am profoundly pessimistic about theory-neutral, intuition-based arguments in philosophy, so I prefer to state a problem which is clearly worth solving and follow through on the predictions of an explanatory theory which seems to have some prior probability.

In Chapter 1, I develop a new theory of confused identity as a representational mental state. Ruth Millikan has argued that such confusion is (i) an 'error of its own kind,' which (ii) 'corrupts' the basic function of cognition, and (iii) cannot be defined in terms of false beliefs. To illustrate, she asserts that if I thoroughly confuse Bill and Biff and treat them as if they were a single individual, I have a corrupt singular concept of Bill/Biff. The concept tracks two objects while it is 'designed' to track only one. Moreover, the belief that Bill is identical to Biff cannot be attributed to me because then, as she puts it, I "should have to have a thought of Bill and another of Biff, which thoughts I was disposed to coidentify. But a thought of Bill that is other than my thought of Biff is exactly what I do not have" (Millikan 1994: 97).

I show in detail that although (iii) is wrong, (i)-(ii) are correct. Specifically, I show that false identity beliefs must be attributed to confused agents if their linguistic behaviour is to be explained at all. On my alternative view, confused identity is defined as a mental state which manifests itself in two basic ways. Either the agent implicitly and falsely believes that two objects are identical ('combinatory confusion') or she implicitly and falsely believes that a single object is two objects ('separatory confusion').

In Chapter 2, I put this theory to work to define the tradition of 'puzzle-driven' semantics. I show how this tradition has dominated philosophy of language and mind since Frege. Theorists in this tradition describe contexts in which confused speakers utter singular terms and refer thereby to an object about which they have false identity beliefs. Cases of this kind have been thought to pose various semantic puzzles, such as: What did the speaker, in that context, *actually* refer to? Philosophers wrongly suppose it is an important theoretical task to answer questions of this sort.

Indeed, it is doubtful that semantic claims about confused speakers can be generalized to speakers and hearers in any 'normal' context. By analogy, cognitive scientists don't assume that claims that are true of the mechanism of face-perception in prosopagnosiacs are also true about such mechanisms in

normal humans. But many of the most influential arguments in the philosophy of language fail, or so I argue on a case-by-case basis, because the speaker's identity confusion disrupts the evidential proper function of his utterance of a corresponding singular term.

In Chapter 3, I start developing the view that semantics should be explanation-driven and intention-based. I present the general outlook of Griceanism and introduce two arguments in its favor. The first is based on recent research in pragmatics, especially on cases where the intended meaning of the speaker is radically underdetermined by the linguistically encoded meaning of the sentence in the language. The second uses experimental data from phonetics and phonology to show that intentionalists provide clear and convincing explanations of the nature of malapropisms and various types of speech error.

Chapter 4 then proposes, as described above, doxastic constraints on any Gricean notion of the act of referring with a singular term, such that problems arising from conflicting referential intentions are completely avoided. I argue that the constraints—i.e. the 'edenic' and the 'cognizance' constraints—characterize a notion of reference which plays a fundamental explanatory role in the final semantic theory.

Finally, in Chapter 5, I argue that the edenic theory of reference calls for a change in the way we think of the logical form of coreferential constructions. It is natural, on the intention-based view proposed here, to think of the linguistic meaning of sentence types in terms of non-propositional blueprints or templates. Such blueprints encode very general instructions for message construction. In the case of simple sentences containing only one singular term and a predicate, the application of the edenic theory is quite straightforward.

Coreference is different. I present four arguments for thinking that edenic and cognizant utterances of polyadic coreferential sentences really express monadic contents. Uttering the polyadic form communicates a monadic property which is applied to the object referred to, even if it is referred to multiple times in the same clause. This is called the 'monadic' thesis. If the monadic thesis is true, and I argue that it is, it gives rise to a new kind of structural mismatch between linguistic meaning and propositional content. I show how intention-based semantics gives a better explanation of this kind of mismatch than other theories, for example minimalism and hidden indexicalism.