

# The Edenic Theory of Reference

Elmar Unnsteinsson

POSTPRINT. PLEASE CITE PUBLISHED VERSION:

*Inquiry* (2019) 62(3):276–308

doi:10.1080/0020174X.2018.1446050

## Abstract

I argue for a theory of the optimal function of the speech act of referring, called the edenic theory. First, the act of singular reference is defined directly in terms of Gricean communicative intentions. Secondly, I propose a doxastic constraint on the optimal performance of such acts, stating, roughly, that the speaker must not have any relevant false beliefs about the identity or distinctness of the intended object. In uttering a singular term on an occasion, on this theory, one represents oneself as not having any confused beliefs about the object to which one intends to refer. This paves the way for an intentionalist theory of reference that circumvents well known problems, which have not been adequately addressed before in the literature.

## 1 Introduction

What would a plausible and detailed intention-based theory of singular reference look like? In this paper, I motivate and lay the groundwork for a proposal of this kind, called the edenic theory of reference. For some time it has been assumed by a number of theorists that Gricean intention-based approaches to singular reference face immediate and fatal objections. In particular, it is thought that various cases of speaker confusion, ignorance, indifference, and deliberate deception make it difficult to sustain a strongly intentionalist perspective.<sup>1</sup>

My aim is to develop and argue for a theory of the metaphysics of reference that takes the common core of these objections into account. First, I introduce the notion of ‘optimality’ as it occurs in Grice’s explication of speaker meaning, arguing that this notion is of help in identifying the core causal factors to be posited in explaining singular communication. The basic idea, here, is that our theory ought

---

\*(✉) [elmar.geir@gmail.com](mailto:elmar.geir@gmail.com)

<sup>1</sup>See, for example, Gauker (2008); King (2014); Lewis (2013); Perry (2012); Reimer (1992); Speaks (2016).

to specify the optimal or proper mental state involved in the communicative speech act of singular reference. This optimal state, I argue, is defined by the absence of specific kinds of false beliefs about the object to which the speaker intends to refer on a given occasion. Finally, then, we have a notion of idealized, so-called ‘edenic’ reference which will do fundamental explanatory work. I present three independent arguments for the resulting theory and address an objection due to Eliot Michaelson. Finally, I compare the edenic theory favorably to a proposal defended by Jeffrey King, which is similar in some respects.

## 2 Edenic reference

The Gricean program in semantics aims to explain the fact that speakers and hearers can communicate their beliefs and other mental states, so efficiently and effortlessly, by means of linguistic utterances. To this purpose, Griceans propose to explicate basic semantic and intentional notions. Linguistic meaning is explicated in terms of notions like *speaker meaning*, *intention*, and *belief*. *Saying* is, normally, explicated in terms of the coincidence or compatibility of speaker meaning and linguistic meaning on an occasion of utterance. And so on. The foundational notion of *speaker meaning* is usually explicated in terms of audience-directed and language-independent communicative intentions.<sup>2</sup>

In his 1982 paper, ‘Meaning revisited,’ Paul Grice writes:

[T]o say what a word means in a language is to say what it is in general *optimal* for speakers of that language to do with that word, or what use they are to make of it; what particular intentions on particular occasions it is *proper* for them to have, or optimal for them to have. Of course, there is no suggestion that they *always* have to have those intentions: it would merely be optimal, *ceterus paribus*, for them to have them. As regards what is optimal in any particular kind of case, there would have to be a cash value, an account of why this is optimal. (Repr. in 1989: 299, first two emphases added)

He then says that there can be a whole range of different accounts of this ‘cash value,’ mentioning that the usage might be conventional or it might be laid down by the inventor of an artificial language, but “what we get in every case, as a unification of all these accounts, is the optimality or propriety of a certain form of behavior.”

This is the spirit in which I propose to take Gricean explications generally. For instance, the complex hierarchy of intentions required for speaker meaning—however it is described in the final theory—constitutes the optimal mental state of

<sup>2</sup>See, e.g., Bach & Harnish (1979); Carston (2002); Grice (1989); Neale (2005); Schiffer (2014); Wilson & Sperber (2012).

a speaker who wishes to communicate something by an utterance. Explications of this sort have many virtues, two of which I mention here. First, they hold the promise of a non-circular and naturalistic account of linguistic meaning.<sup>3</sup> Second, they form the basis of detailed, influential, and fruitful theoretical frameworks in pragmatics and the related study of language evolution, language acquisition, and more (e.g., Sperber & Wilson 1995; Tomasello 2008; Scott-Phillips 2015).

With an eye to these strengths, I propose to argue for a new kind of explication of *speaker reference* within a Gricean framework. Although the theory is informed by prior work in this tradition (esp. Bach 1987: 51–53; Neale 2016; Schiffer 1981: 69–76), it is novel in construing the task of explication in terms of specifying optimal or proper functions for types of speech acts. Narrowing the present inquiry even further, the view to be defended here is a theory of the optimal function of the speech act type of referring singularly to an object in linguistic communication. Non-communicative and non-linguistic acts of reference—if such exist—will be left to one side for the most part.<sup>4</sup>

On a Gricean picture, both linguistic and non-linguistic acts of referring ought to be grounded in the prior notion of speaker meaning, which is metaphysically more fundamental. On this view, the act of referring consists, in the first instance, in an act of speaker meaning a singular proposition. So, interestingly, reference is not really a speech act in its own right, it only occurs in the course of expressing a kind of proposition, namely a proposition with an object dependent truth condition. Contrast this with another tradition, more Strawsonian in spirit, according to which expressing a singular proposition consists in having two even more basic communicative intentions. One is the intention to refer to a particular object, the other the intention to express or refer to a property or concept which is then, by way of combination of the two basic intentions, applied to the object referred to.<sup>5</sup> My assumption here is that, on the contrary, speaker meaning a singular proposition is explanatorily prior to these two types of intentions or communicative acts.

With these caveats firmly in mind, I propose to work with the following definition of speech acts of linguistically referring to an object in communication.

<sup>3</sup>See, for example, Grice (1969); Harris (2016); Loar (1986); Neale (1992, 2005, 2016); Schiffer (1972, 1982: 128–137, 2003, 2006, 2014). For a more skeptical view, see, e.g., Lance & O’Leary Hawthorne (1997: 287–297).

<sup>4</sup>It should be noted, however, that in the Gricean project of explicating linguistically encoded meaning, the notion of non-linguistic acts of reference—in communication—must play a foundational role. And, further, non-communicative and audienceless uses of language must be given a separate treatment. These issues are not to be resolved in this paper.

<sup>5</sup>For example, Strawson (1950: 17–19). See Bertolet (1987: 202–206); Schwarz (1976: 67–68); Stine (1978: 323). The Strawsonian tradition is recently enjoying resurgence, see, especially, Hanks (2011, 2015); Soames (2010, 2014: 105). Soames speaks of the act types of *targeting* and *predicating*, combinations of which give rise to full propositional act types.

### Referring with (RW)

Speaker *S* refers to object *o* with *e* in uttering *U* on a given occasion iff *U* is an utterance of sentence type  $\sigma$ , *e* is an expression type contained in  $\sigma$ ,  $(\exists H)(\exists p)$  such that the truth condition of *p* is *o*-dependent and *S* utters *U* intending

- (1) to produce thereby in *H* the belief *that p*;<sup>6</sup>
- (2) *H* to use the *e*-part of *U* as direct evidence that the belief *S* intends to produce in *H* is about *o*;
- (3) *H* to recognize *S*'s intentions (1) and (2).<sup>7</sup>

To illustrate, suppose I utter 'Gill snores' primarily intending, thereby, to produce in you the belief *that Gill snores*. There is, then, a proposition *p* with a *Gill*-dependent truth condition such that I intend for you to believe *p*—or to believe that *I believe that p*.<sup>8</sup> My whole utterance is an event—involving me emitting a sequence of sounds or inscribing a sequence of letters—encompassing the subutterance associated with the expression type 'Gill.' At its core, then, my *referential intention*, on this occasion, consists in my intention that you use the 'Gill'-part of my whole utterance as direct evidence that the belief in question is about *Gill*. The 'Gill'-part is that temporal part of my act which is intended by me to represent the expression type 'Gill.' RW only appeals to expression types—sentence-size or smaller—and acts of uttering expression types, without committing to a notion of 'token' expressions (cf. Unnsteinsson 2014). Note that there is always a co-occurring informative intention (1), so performing the act defined in RW never occurs without the expression of a proposition (cf. Bach 1987: 51). And, further, when I perform an act

<sup>6</sup>I focus on beliefs only for simplification. Clearly, any other propositional attitude would do just as well. Note also that when there is no hearer, we stipulate that  $S = H$  (Schiffer 1972: 76–80).

<sup>7</sup>Following Sperber & Wilson (1995: ch. 1), among others, the definition of RW skips Grice's third condition for speaker meaning, stating roughly that *H*'s recognition that *S* intends (1) should function, in part, as a *reason* for (1). The condition may well be on to something important about the nature of speaker meaning and, derivatively, speaker reference, but it brings with it particular complications that I wish to sidestep here. See in particular García-Carpintero (2001: 102–103); Neale (1992: 547–549); Wharton (2009: ch. 2).

<sup>8</sup>This proviso is added because, as is well known, Grice (1989: 106–112) and others gave counterexamples to (1) as it occurs in the explication of speaker meaning. In an oral history exam, the examiner asks the student when the Battle of Waterloo was fought. Student replies '1815' *saying* thereby, it seems, *that it was fought in 1815*. But of course the student does not intend to produce this belief in the examiner, for he (the student) knows that she already knows *that it was fought then*. Thus, Grice distinguished between *exhibitive* and *protreptic* utterances. For the latter (1) is left unchanged, for the former it becomes: to produce thereby in *H* a belief *that S believes that p*. McDowell (1980: §5) and Neale (1992: 545–547) point out that this makes our (exhibitive) utterances about our own mental states rather than about the world itself, which appears to be implausible. Based on recent work in relevance theory on hearers' epistemic vigilance—especially Sperber et al. (2010)—I incline to the view that this is entirely acceptable. All of our utterances could be exhibitive but we could still explain how communication provides speaker/hearers with knowledge of the world. I keep the protreptic form here only for simplification.

of referring *with* I have both the usual informative-communicative intention with respect to (1) and the added referential-communicative intention with respect to (2), and this is made explicit in the third and final clause. The act could consist in uttering only a subsentential expression, if the speaker does so in order to communicate a full proposition, which is plausibly the case when names are used to answer questions.

What does it mean, however, for the speaker to intend the hearer to use some utterance-part as *direct* evidence? Well, it seems like speakers can intend their audience to use almost any part of their utterances as evidence for almost any part of what they mean. Suppose, for example, that I say to you

(1) Bill is kissing Barry,

adding disdainfully that Bill is married to someone else. Suppose, further, that I know that you do not know Bill's name and that I refer to Barry precisely to help you figure out who Bill is, namely the one conspicuously kissing Barry.<sup>9</sup> Even if I know that you do not know Bill's name I can intend the 'Bill'-part of my act of uttering (1), in the context, as giving direct evidence that the belief I intend to produce in you is about Bill. This is because I can make use of available contextual cues, such as salience, gestures, and demonstrations, giving more structure to this part of uttering (1). Further, if such cues are completely lacking, speakers will be more disposed to use cleft constructions, 'It is Bill who ...', indicating, perhaps, that a name is being explicitly introduced. But a problem remains, it would seem, for if I can still intend the utterance of 'Barry' as, in some sense, the most immediate reason for you to infer the identity of Bill, RW predicts that I refer to Bill with 'Barry' as well as Barry himself in uttering (1). This is not exactly right, however. True, the fact that the belief is about Barry can be intended as evidence that it is also about Bill but it does not follow that the mere act of uttering 'Barry' is so intended. Call this temporally extended part of the act of uttering (1) /barry/. Clearly, /barry/ is only intended as evidence that the belief is about Bill *via* the fact that it is evidence that the belief is about Barry. So, /barry/ can only be indirect evidence that the belief is also about a man named Bill. The only possible way, it seems, for /barry/ to be direct evidence that the belief is about Bill is when I happen to believe, rightly or wrongly, that Bill is sometimes called 'Barry'. Anyway, RW is only concerned with utterance-parts that are intended *directly* as evidence that a given proposition is about a particular object.

Finally, note that (1) in RW must be the speaker's *primary* communicative intention, in the sense that the utterance will count as a minimally successful act of communication if (1) is indeed communicated, i.e. recognized by the intended audience. The speaker may have secondary intentions to convey other propositions, for example by implicature, which won't have an associated referential intention

<sup>9</sup>I thank an anonymous referee for this journal for providing this example and pointing out the issues it raises.

of the kind specified in (2). Thus, the import of (2) can be captured as follows: the speaker intends the utterance of the singular term as direct evidence that the primary proposition expressed is about some particular object.

Now, as already mentioned, there is a long and lively tradition of arguing that intentionalist theories of this general sort are bound to fail. A number of theorists—starting perhaps with David Kaplan’s (1978: 239) Carnap/Agnew case—have argued that while the speaker’s referential intention on the occasion of utterance may determine one referent, the *actual* referent may be something quite different and, so, the intentionalist makes the wrong prediction in many cases. The speaker may be ignorant, misinformed, or have accidentally misspoken, but the result is always supposed to be that there is a fundamental distinction between what speakers actually refer to and what they merely intended or had in mind.<sup>10</sup>

Although, to my mind, this objection has been adequately addressed in the literature,<sup>11</sup> there is a related worry that has yet to be taken care of. As Jeff Speaks (2016: 33–36) recently pointed out, this is the problem of internally conflicting intentions. The problem is this: What can a theorist who endorses a definition like RW say about cases where the speaker’s referential intention conflictingly determines two objects at once, or no object at all, because of underlying confusion in the speaker’s beliefs? This is a problem, more precisely, because it certainly appears as if speakers who are confused about the identity of an object can still very well refer to that object in communication. Thus, it may appear, confused speakers’ referential intentions do not actually determine what is referred to on such occasions.

To be more concrete, assume the following model of the mental state of confused identity: ‘confusing  $x$  and  $o$ ’ is defined as *either believing falsely that  $o = x$ , or believing falsely that  $o \neq x$* , for any object  $x$ . Believing falsely that two objects are one is called ‘combinatory confusion’ and believing falsely that one object is two is called ‘separatory confusion’. For example, Peter, in Kripke’s well known example, is separatorily confused in that he believes falsely *that Paderewski<sub>1</sub>  $\neq$  Paderewski<sub>2</sub>*, where the subscripts represent the one he takes to be a famous pianist and the one he takes to be a statesman, respectively. Further, when I confuse Bill and Biff, who are identical twins, thinking of them as a single person with a single name, I mentally combine two objects into one and believe falsely, on this theory, *that Bill is Biff*. We can assume that the two Paderewskis and the Bill/Biff-fusion are non-actual possible objects.<sup>12</sup>

<sup>10</sup>See, for example, Reimer (1992, 2004); Perry (2012); Kripke (1977); Wettstein (1984); King (2014).

<sup>11</sup>Åkerman (2010); Bach (1987: 183–186, 1992); Neale (2005: 184–185); Strawson (1974: 52); Unnsteinsson (2017).

<sup>12</sup>The latter is a possible gerrymandered object in the actual world. Whether the former is a possible object at all is controversial but we need not make metaphysical assumptions of this sort here. For convenience I sometimes assume that the Superman who is not Clark Kent is a non-actual possible object, but it can also be assumed that this is a non-actual *impossible* object. Other models



Importantly, the model allows for the idea that the false belief in question could be implicit and nonconscious. Speakers are confused in the requisite manner if either type of false belief—combinatory or separatory—must be attributed to them, by the theorist, to explain their reasoning and behavioral dispositions. It might very well be, for example, that the confused speaker is unable, in virtue of the confusion, to mentally represent precisely the belief *that Bill is Biff*.<sup>13</sup> The only additional constraint we need, to make sense of implicit false beliefs involving object-identity, is that there are at least some explicit and object-dependent propositional attitudes the speaker bears to the relevant individuals. For example, the speaker may have demonstrative and perception-based thoughts about Bill and Biff on different occasions, while implicitly and nonconsciously believing that they are a single individual.<sup>14</sup> Furthermore, to take Gareth Evans' (1973: 19) example, if there is a group of mathematicians using 'Ibn Kahn' to refer to the one who proved some theorems on recently excavated manuscripts, although that name, as it occurred on the manuscript, was being used for the scribe, their use is only confused if they can plausibly be said to have propositional attitudes towards the scribe as well (we'll come back to the example in §2.3 below).

Finally, the proposed model must abstract away from global and collective confusions of identity, such as when scientists are entirely incapable, over a long stretch of time, of distinguishing two natural properties by using any of the means available to them.<sup>15</sup> Such cases present their own set of questions and difficulties. Local and individualistic confusions, on the other hand, are always such that the confused person is epistemically blameworthy, even if only to a very small degree. For example, even if the deception is quite elaborate and well executed, Lois ought to be able, in principle, to disabuse herself of her confusion about Superman.

Now, as 'confusing *x* and *o*' has been defined it seems clear that a speaker could satisfy the definition and still succeed in communicating reference to *o* with a linguistic expression on a given occasion of utterance. Clearly, one would like to say, Peter could refer to the individual called 'Paderewski' by pointing to him and using the singular term 'Paderewski' on a given occasion. But this can be turned into to an apparent counterexample to any intentionalist definition like RW. Consider the

---

of the mental state of confused identity could have been used or appropriated here with similar results—e.g., those of Millikan (2000), Camp (2002), Lawlor (2007), or Recanati (2012). For a full defense and elaboration of the model itself, see Unnsteinsson (2016).

<sup>13</sup>On this basis, Camp (2002: 33) and Millikan (1994: 97) both insist that combinatory confusion cannot be described in terms of belief in a proposition, because the relevant singular representations—of Bill as separate from Biff, and vice versa, in this case—are unavailable to the speaker.

<sup>14</sup>Note that I assume, like the majority of philosophers working on these issues, a distinction between relational or object-dependent and satisfactorial or descriptive contents; confused identity only arises for properly relational attitudes, i.e. attitudes whose content is a proposition with an object dependent truth condition (see, e.g., Bach 1987; Recanati 2012).

<sup>15</sup>For discussion, see, for example, Camp (2002: ch. 2); Evans (1982: ch. 11); Field (1973).

following case. Peter just saw a concert performance by Paderewski<sub>1</sub> the famous pianist with two of his friends. During recess, the three of them meet Paderewski (no subscript) and Peter believes they are meeting Paderewski<sub>2</sub> the statesman. And, of course, Peter falsely believes that Paderewski<sub>1</sub>  $\neq$  Paderewski<sub>2</sub>. Peter says to them, lifting his glass slightly as if he were openly celebrating Paderewski<sub>2</sub>, ‘Paderewski is a great pianist.’ Now, Peter would never say such a thing about Paderewski<sub>2</sub> the statesman because he doesn’t believe he is a musician. Peter’s audience, including Paderewski, take him to be intending to refer to the man standing right in front of him. Further, some theorists would like to conclude that this is the man Peter *actually* referred to. It is perfectly accurate to describe Peter’s mental state such that he intended to refer to an object *o*, such that  $o = \text{Paderewski}_1 \neq \text{Paderewski}_2$ , although this is certainly not how Peter would tend to represent his intention to himself. But he surely intends *not* to refer to the Paderewski standing in front of him, by uttering ‘Paderewski,’ just like he doesn’t intend to refer to either of his two friends. RW would either predict that Peter referred to Paderewski<sub>1</sub> the pianist, or nothing at all, because the conflicting intention determines an impossible, non-self-identical individual. Either way, the prediction is different from what many would take to be intuitive, namely that he referred to the statesman standing in front of him.

Note that I have stated the problem only in terms of separatory confusion. The combinatory case is simpler. Consider a case where I confuse Bill and Biff, calling both ‘Biff’. Standing in front of Biff and a common friend, I might say, ‘Biff was wearing a different shirt yesterday,’ thinking about when I met Bill the day before. It seems like I intend to refer to Bill—that’s the person causally responsible for my memory image—but it still seems like I actually refer to Biff. Another possibility is that the intention conflictingly determines the two individuals simultaneously. But either way the intuitive prediction is different from what we would get out of RW.

The edenic theory of reference is designed to sidestep this problem. We define a notion of *edenic* reference, e-reference for short, thereby placing strong doxastic optimality constraints on competent performances of acts of singular reference. The idea being, roughly, that when speakers e-refer their act constitutes an optimal or proper performance of the communicative and language-involving speech act defined in RW above.



### Edenic reference

Speaker *S* e-refers to *o* with *e* in uttering *U* at time *t* if and only if *S* refers to *o* with *e* in uttering *U* and there is no object *x* such that, at *t*, *S* confuses *x* and *o*.

For the speaker to stop short of e-reference, the mental state of confusion and the performance of the speech act defined in RW must be simultaneous in that the relevant false identity beliefs need to be true of the speaker on the occasion of utterance. The speaker must lack the relevant beliefs, explicitly or implicitly, at the time of utterance to count as an ‘edenic’ referrer. The label is supposed to evoke the highly idealized notion of pure, initial conditions, where speakers could refer directly and without the corrupting effects of confused beliefs (cf. Chalmers 2006). Thus, acts of e-reference require a complex doxastic-intentional mental state which, in words reminiscent of Grice’s own, is the *optimal state* with respect to referring, or if you like, to referring to an object with some linguistic expression in communication (1989: 302, on the optimal state with respect to *communicating*). Grice himself was concerned with postulating the absence of certain sneaky and deceptive intentions so as to avoid infinite regress in the explication of speaker meaning (Neale 1992: 550). The notion of e-reference similarly posits the absence of certain kinds of false beliefs in explicating singular reference.<sup>16</sup> In many cases, of course, ordinary speakers will not realize exactly this complex state, and so they will not e-refer to an object with a linguistic expression on those occasions. Much of the time, such failures do not matter for the practical purposes of everyday communication. The notion of e-reference still captures the mental state proper to the speech act of referring singularly. Or so I argue.<sup>17</sup>

Traditionally, the most important doxastic constraint proposed by Griceans has been the Humpty Dumpty constraint on intention formation, which says, roughly, that speakers cannot rationally intend what they believe to be impossible (see, e.g., Donnellan 1968). So, if I believe it is impossible to utter ‘Blarph schmumpf’ on

<sup>16</sup>A full statement of the edenic theory also involves a related explication of speech acts of singular *coreference* but, for reasons of space, that’s a story for another occasion.

<sup>17</sup>Some may think that the edenic theory is more or less similar to Jerry Fodor’s view in *The Elm and the Expert* (1994) and, so, is subject to all the objections already mounted against the arguments in Chapter 2 of that book. See, especially, Arjo (1996); Aydede & Robbins (2001); Segal (1997); Wakefield (2002). Schneider (2005) defends the view.

Fodor argues, unsuccessfully by most accounts, that Frege cases are *ceteris-paribus* exceptions to psychological laws; pathological cases which can be idealized away. Richard Heck (2012: 136n6) reports that Fodor has abandoned the view. The edenic theory is not in the business of arguing that identity confusion gives rise to exceptions to intentional generalizations and, further, it does not claim that Frege cases are rare; in fact, it is assumed here that they are ubiquitous. The theory is also exclusively concerned with a certain specific communicative speech act, defined partly in terms of speakers’ mental states. Certainly, however, the big picture motivation is very similar to Fodor’s: developing a strongly externalist or information-theoretic account of reference while conceding as little as possible to Frege-type puzzles.

a given occasion to communicate to my interlocutor the proposition *that Trump is president*, I cannot form the intention to do so. Notice, however, that the constraint can be formulated both as a constraint on the formation of communicative intentions and on the performance of a particular speech act. More generally, if I believe that it is impossible for me to *do x by V-ing* I cannot perform the act of *V-ing with the intention of x-ing*. For example, although I can surely flap my arms, I am not able to do so *with the intention* of thereby flying to the Moon.

The notion of e-reference is best understood as placing doxastic constraints on the performance of the speech act of singular reference, as defined in RW. Thus, speakers can form intentions to refer to or think about an intentional object *o* even if they have false identity beliefs about *o*.<sup>18</sup> If Millikan (1997, 2000) is right, this manifests a special kind of cognitive and conceptual corruption, but the very formation of these corrupt thoughts or intentions is not precluded. Confusion only precludes the performance of a specific communicative act, the act of e-reference. The most significant difference between the edenic constraint on reference and the Humpty Dumpty constraint, however, is that confused speakers are completely unaware that they have these false beliefs. They don't realize that they are combinatorily or separatorily confused about the object(s) to which they wish to refer in speech. So, the fact that the confused speaker is unable to e-refer, on an occasion of utterance, is an utterly external matter, determined by whether or not particular beliefs about identity or distinctness are in fact true or false.

At this juncture, it is important to note that this explication is, as Gricean explications generally, intended as a stipulative definition of utility in the construction of a metasemantic theory and not as a piece of conceptual analysis (Schiffer 1981: 68; Neale 2016). Just like what has recently been called 'conceptual ethics' or 'conceptual engineering,' the central motivation is to find the best, or optimal, concept of reference to do some particular theoretical work (e.g., Burgess & Plunkett 2013a, 2013b). In this case, it's the work of explaining aspects of success in linguistic communication. Also, in light of this construal of the project, judgments arrived at by consulting one's untutored intuitions about various cases involving reference are far from dispositive in assessing the merits of the notion. Ordinary language or unreflective intuition may, as noted above, incline one to say that the speaker certainly *referred*, but our theory will still predict that they failed to pull off an act of e-reference. And e-reference is a technical notion about which we wouldn't seem to have reliable intuitions.

Some will argue that this makes e-reference too idealized or simplistic and, indeed, many theorists have argued more generally that a Gricean model of communication suffers from the same flaws.<sup>19</sup> The model assumes that speakers routinely

<sup>18</sup>So, the edenic theory is not about reference 'in thought,' unless, perhaps, if it assumed that we think in a language and that this must be some natural, public language. This is not assumed here.

<sup>19</sup>A group of theorists unsympathetic to the Gricean program label it the 'Lockean' or 'commu-

have complex audience-directed intentions when communicating, although many theorists think they themselves don't have such intentions and, more pressingly, argue that young children, people with autism, and perhaps even some species of nonhuman animals, can communicate easily without them. This is not the place to enter that debate (see, e.g., Breheny 2006; Thompson 2014). But the model is proposed as an idealization necessary to adequately explain linguistic communication. More specifically, the model should be thought of as identifying core causal factors involved in such an explanation. As Strevens (2016), Weisberg (2007), and others have argued, scientific explanations can essentially involve falsifications which serve to highlight difference-making factors in the causal production of some phenomenon. Factors that are causally irrelevant in explaining the phenomenon in question should be isolated, flagged explicitly, and idealized away. Such idealizations, the thought goes, provide us with a better grasp of a correct explanation and are not justified as mere pragmatic, short-term fictions, to be eliminated with the advance of science (Weisberg 2007: 645).<sup>20</sup>

The simple model characterizes communication when all goes according to design; when the mental states of speaker and hearer are optimal with respect to communicating contents. The hypothesis on the table is that the notion of e-reference specifies the core causal factor involved in a correct explanation of our capacity to express and communicate singular or object-dependent information so efficiently by using certain linguistic devices.<sup>21</sup> So, the aim here is to use idealization to capture the underlying mechanism of reference which is thought to undergo external interference through the mental state of identity confusion. As theorists, we can understand speech acts of reference partly in terms of how they deviate from the ideal. The ideal is a core causal factor because it describes an important aspect of the goal of the speaker. Normally, speakers assume that their reference will be edenic and when they are right their act is to that extent optimal.

This is not necessarily so for what can pass for a hearer's understanding of what the speaker referred to on a given occasion. Here it may matter that one find some referent or other, the most plausible one, that will suffice for the purposes of the conversation. As I argue below (Section 4), the hearer's strategy may sometimes

---

nicative' conception of language, arguing that it is naïvely mentalistic and in need of significant revision or replacement. See, in particular, Bar-On (2013); Carruthers (1996); Gauker (2003); Green (2007). These authors are, at least in those publications, united in the belief that utterances express mental states directly, rather than being intended to produce some cognitive effect in an audience.

<sup>20</sup>See also Elgin (2007); Mizrahi (2012); Dellsén (2016).

<sup>21</sup>Stephen Neale (2004: 71–72, 2016) argues that this ought to be considered the 'master question'—i.e., the question in terms of which theoretical terms are to be justified—in philosophy of language: "How rich of an explanation can we provide of our capacity to express and sharpen our thoughts, and to communicate information about the world and about our beliefs, desires, plans, commitments, hopes, fears, and feelings so efficiently—so quickly, systematically and consistently—using various noises, marks, and gestures?" My aim is to answer this question for singular reference in particular.

be to suppose that a given act is edenic when it is not. It is as if the hearer is asking: What would the speaker have said if we exchange the false identity belief for a true one? As Millikan has argued, in a similar context, this often does not result in any obvious answer. How exactly would our overall belief system change upon being told, by an eminent historian, that Cicero and Tully were different people? (Millikan 1997: 508). But sometimes the answer is absolutely clear, so clear indeed that we sometimes say that a speaker who confuses identical twins *is actually talking about* the one she would have named if she were not confused. “You’re talking about Bill. Biff is actually his identical twin!” Making a decision about the so-called actual referent in this way is not necessary to have full theoretical understanding of acts of reference, deviant or otherwise. The folk notion of reference is quite indeterminate anyway and tends to give rise to conflicting intuitions. The edenic theory makes no explicit commitment about what non-edenic speakers actually refer to, but it insists that their acts are best understood, from a theoretical perspective, in terms of their deviation from the edenic ideal.

Let’s clarify this further by seeing how, exactly, the edenic theory is supposed to sidestep the Paderewski puzzle discussed above. Peter’s referential intention, in uttering ‘Paderewski,’ according to RW at least, either determines no referent or determines Paderewski<sub>1</sub> as referent, but does not determine Paderewski<sub>2</sub>. I propose that this makes the act defective as an act of singular reference. This is because part of his goal, in making the utterance, was to avoid this very defect and he failed. The edenic theorist avoids making a commitment about what constitutes the actual referent in this case—nothing, Paderewski, or one of the non-actual possible objects Paderewski<sub>1</sub> and Paderewski<sub>2</sub>—insisting that it is indeterminate and that making the choice is not needed for purposes of explanation. Assigning a particular referent is causally irrelevant in explaining Peter’s act in this case. All we need is a satisfactory description of his confused mental state at the time of utterance and of his goals and intentions in speaking as he did. As we will see below, his goal, but also the task of the audience in reaching an interpretation, will be described partly in terms of edenic reference.

Richard Grandy and Richard Warner, in their development of Grice’s remarks on ideal or optimal mental states, explain the point with a helpful analogy:

There is in sailing an optimal setting for the sails—a setting that maximizes forward thrust. Any reasonably complete text on sailing will explain at least some of the relevant aerodynamic theory. Now this optimal setting is difficult if not impossible to achieve while actually sailing—given continual shifts in wind direction, the sudden changes of direction caused by waves, and the difficulty in determining air-flow patterns by sight. To deal with these practical difficulties, the text supplies numerous rules of thumb which are relatively easy to apply while sailing. Why not just drop the aerodynamic theory altogether and just provide the reader/sailor with the rules of thumb? Because

they are *rules of thumb*. They hold (at best) *other things being equal*. To spot exceptions and resolve conflicts as well as to handle situations not covered by the rules, one needs to know what the aerodynamic optimum is. This optimum plays a crucial role in guiding the use of the rules of thumb. (1986: 25–26, emphases in original)

The notion of edenic reference, as it occurs in the Edenic maxim introduced in the next section, is to be understood as such an ‘optimal setting’ for a certain form of behavior, namely communicative acts of singular reference. Even if this maxim may be buried in a clutter of communicative procedures that are easier to apply directly, as a speaker, to a given situation, implicit knowledge of the maxim still feeds into utterance planning and design as well as utterance interpretation. In the three subsections that follow, I present three arguments for thinking that identity confusion ought to be idealized away, in the sense explained, in developing an intention-based theory of reference.

## 2.1 Edenic reference is part of pragmatic competence

Competent speakers implicitly believe, in virtue of their linguistic competence, that confusion ought to be avoided if one is to communicate successfully. E-reference thus captures a feature of our pragmatic or communicative competence. An adequate theory of the speaker/hearer’s linguistic competence will include at least three sub-theories; (i) a theory of syntactic competence, (ii) a theory of semantic competence, and (iii) a theory of pragmatic competence. This last part describes the mechanisms in virtue of which speakers are able to plan their utterances in light of their beliefs about the mental states of the audience and, correlatively, how hearers are able to correctly attribute communicative intentions to speakers on the basis of the evidence provided by their utterances. Pragmatic competence is thus a rich system of inferential mechanisms and capacities for mindreading.<sup>22</sup>

Patterns of ordinary criticism provide some evidence for edenic reference being part of pragmatic competence. The argument that follows takes the form of an inference to the best explanation. I consider examples where confused speakers’s acts of reference are justifiably corrected. The proposal is that if non-edenic acts of reference are defective qua acts of reference, in virtue of the speaker’s identity confusion, this would help explain the patterns of correction that we can observe. The following section is tasked with establishing that non-edenic acts are as such defective or ‘corrupted,’ in a specific sense.

Now consider the following example. Lois says to *H*,

- (2) Superman loves his<sub>↓</sub> mother.

<sup>22</sup>As Chomsky (1980: 225) puts it, “pragmatic competence places language in the institutional setting of its use, relating intentions and purposes to the linguistic means at hand.” See also Carston (1998: 5–7, 2002: 10–11); Neale (2005: 188–189).

And she points to Clark Kent—i.e. Superman wearing disguise—intending ‘his<sub>↓</sub>’ deictically, as signalled here by the downward arrow. Add that everyone except Lois now knows the truth and there is no reason at all for *H* to keep her out of the loop. But Lois, on the theory presented above, is separatorily confused about the identity of Superman/Clark Kent in virtue her false belief that they are not identical. In such a case *H* is justified in correcting Lois: ‘You’re confused. His<sub>↓</sub> mother is Superman’s mother. He<sub>↓</sub> is Superman.’

After accepting the correction Lois might say ‘Right, so what I really wanted to say was that (3).’

(3) Superman<sup>1</sup> loves his<sub>1</sub> mother.

Where the number 1 signals some suitable relation of binding between expressions. Note that in some languages (2) and (3) have different translations and that there is good reason to suppose that /his/ in English is really the superficial form of at least two distinct pronouns: reflexive and non-reflexive. Certainly, there are many possible suggestions as to why the hearer’s response seems appropriate. But I would offer the following. The hearer *H* objects to Lois’ use of ‘his’ as a singular, deictic term in (2) solely in virtue of *H*’s knowledge that Lois believes falsely *that Superman ≠ Clark Kent*. To see this, consider the very same context where nothing at all is changed, except that, now, Lois is privy to the truth and this is mutually known by her and her interlocutor. If Lois were then to utter (2)—again pointing to Superman dressed as a reporter—she would be doing something a bit unusual, perhaps making a subtle joke, but her act of referring with ‘his’ is not objectionable in the same manner as before. Thus, it is reasonable to suppose, it is in virtue of her identity confusion that the act of reference is subject to correction (see §3 for more discussion of examples of this kind).

I do not pretend that this kind of example is conclusive, especially since it is fairly obvious that not all non-edenic acts of referring are normally subject to criticism as such. The occurrence of ‘Superman’ in (2) is a case in point. The edenic theory implies, in this context, that Lois failed to e-refer to Superman by uttering ‘Superman,’ and it is unlikely that ordinary speakers would find this particularly problematic. But this just happens to be a context in which the infelicity has no practical upshot and, so, is not noticed at all. The defect or deviance of a mechanism does not always give rise to practical problems and should, thus, not be defined in terms of doing so. If I tend to pronounce English ‘w’ as /v/ instead of /w/ my mechanism of ‘w’-articulation is defective. This defect will, however, usually not be noticed by hearers—people are not good at detecting subtle idiolectal differences—and when it is noticed it will usually not matter for practical purposes. It may matter, though, when I want to be taken as referring to my favorite wines but not my favorite vines (Unnsteinsson 2017). As theorists, however, we explain the relevant action partly in terms of the way in which it deviates from the ideal or optimal function of the underlying mechanism, which specifies a significant part



of the action's goal. Edenic reference is the intended target for acts of referring just like the sound /w/ is, normally, the intended target for an act of articulating English 'w'.

Roughly, the defect of non-edenic acts of referring consists in the fact they cannot, in principle, fulfill their directly evidential function. Even Lois' utterance of 'Superman' in (2) is intended, by her, to give direct evidence that the belief she intends to produce in her hearer, is about someone who is not identical to Clark Kent (see next section). But this aspect of her intention can be more or less salient and be more or less important for the purposes of a given conversational exchange. This particular utterance of 'Superman' does not give rise to practical problems even if the posited defect is still lurking in the background. Imagine, for example, a case where Lois is lead to believe that the man she knew as Clark Kent died five years ago and now she only ever thinks about that man in the role of Superman. Ultimately, he mental state would not be usefully described as one of confusion at all, for there must be some mental or linguistic representation or other, that she is at least disposed to activate, such that it is intended to refer to an *o* such that  $o = \text{Clark Kent} \neq \text{Superman}$ .

More simply, consider Rose, who combinatorily confuses Bill and Biff, calling both 'Bill.' Even in a context where the contextual cues make it clear that she is using 'Bill' to refer to *Bill*, because, for example, he's standing right in front of her and Biff is nowhere to be seen, her act of referring may be subject to polite correction. Imagine, for instance, that she says: 'Bill was wearing a different shirt yesterday.' She may very well be right, in the sense that Bill indeed was wearing something different then. But she in fact saw Biff wearing a different shirt yesterday and that's the reason for her utterance. Her interlocutor, discovering Rose's confusion, may correct her: 'You're saying that because you saw *Biff* yesterday. You're confusing Bill with Biff.'

This is evidence that the speech act of referring is subject to, as we might call it, the rule or maxim of edenic reference. Mastering the act of referring involves, among other things, becoming sensitive to the maxim. Of course, competent speakers do not consciously represent the maxim, but they behave as if in accordance with it, particularly when judging some acts of reference to be proper and others improper or somehow faulty. We can now formulate the maxim as follows.

### **Edenic maxim**

Try to refer to *o* with *e* only if you thereby refer edenically to *o*.<sup>23</sup>

To be clear, the claim here is only that speakers who flout the maxim are, by definition, unable to perform the act of e-reference, but they could still successfully

<sup>23</sup>This should remind one of Timothy Williamson's (2000: 243) so-called knowledge norm of assertion. I do not, however, endorse his view that the knowledge norm is constitutive of assertion—but it is part of speakers' semantic or pragmatic competence. Similarly, if the reference rule is constitutive of anything it is merely constitutive of the theoretical notion of RW, which, on my view, also captures an important aspect of pragmatic competence with singular terms.



perform an act of reference, in an ordinary sense of the word. It is just that the act is deficient in that it breaks this implicitly accepted rule. By uttering a singular term—especially if it is a proper name—one normally, but only implicitly, *represents oneself* to one's interlocutor as not having any false identity beliefs about the object referred to. One represents oneself as producing an utterance that is causally grounded in the referent.<sup>24</sup> And of course one can be wrong in so representing oneself.

## 2.2 Confusion is corruptive

Call a tokening of singular term *e* confused if the speaker's act of referring with *e*, on that occasion, is not edenic. In cases where a speaker utters a confused singular term *e*, I will say that their referential intention is *conflicting* and the evidence provided by uttering *e* is *corrupt*. More specifically, the intention is conflicting in that either (i) *e* is equally causally anchored in two objects while uttering *e* is supposed to indicate only one object, or (ii) *e* and another expression *e'*, also in the speaker's idiolect, are equally causally anchored in a single object while uttering *e* is supposed to indicate only something distinct from what would be indicated by *e'*. The evidence provided by the utterance is corrupt in (i), i.e. combinatory confusion, for it optimally functions to pick out a single object while picking out two objects, or their fusion. In (ii), i.e. separatory confusion, the evidence functions optimally to pick out a single object while picking out none, for, in the relevant cases, nothing is picked out by *e* while not being picked out by *e'* as well.<sup>25</sup>

For example, Lois Lane suffers from separatory confusion and, the thought goes, has no conception of Superman which is not at the same time a conception of something which is not Superman, i.e. not Clark Kent. This is because she believes of Superman that he is not identical to Clark Kent; so, whenever she intends to refer to Superman, she intends to refer to something which is Superman and not identical to Clark Kent. But there is no such object. This makes her utterances of this proper name sub-optimal as evidence for a singular referential intention, as the underlying intention itself is in conflict.

For combinatory confusion, consider this example from Marga Reimer (1992). Suppose I share an office with a friend and I forget my keys on the desk while on my way out. Suddenly, I realize this and return to my office. I see my own keys and, right next to them, my colleague's keys. Looking elsewhere I mistakenly grab *his* keys while saying, 'These are mine.' According to Reimer, the Gricean view would predict that I referred to *my own keys* with 'These.' But surely I in fact referred to my colleague's keys—I'm holding them in my hand—and spoke falsely. Thus Reimer, and others, take such examples to show that the Gricean view

<sup>24</sup>Cf. Evans (1982: 310); Williamson (2000: 252n6).

<sup>25</sup>The sense in which confusion is corruptive is explained and argued for in more detail in Unnsteinsson (2016: 219–224).

is wrong. According to the edenic theory, however, my false identity belief that *the-keys-I-just-saw* are identical to *the-keys-in-my-hand* makes my utterance of ‘These’ evidentially corrupt in the context. The right thing to say is that Griceanism predicts that I referred (confusedly) to *both* sets of keys—or, maybe, to their fusion—it’s just that the hearer is *much* more likely to pick up on one referent rather than the other in this particular context. And such confused intentions are not optimal with respect to singular term utterances.

Let me clarify the point at issue with an analogy. Consider Tyler Burge’s (1979) arthritis case. The protagonist *S* believes many things about *arthritis*: *that arthritis is uncomfortable, that she has arthritis in her wrists*, etc. But *S* also has a few false beliefs. *S* believes

- (4) *that she has arthritis in her thigh,*

and, accordingly, *that arthritis can be something other than specifically an inflammation of joints*. In Burge’s original story, *S* reports to a doctor her belief (4) and is corrected. But imagine, instead, that *S* simply utters

- (5) Arthritis is annoying.

For sake of argument, assume, first, that *S*’s reference to *arthritis* is not somehow ‘deferential’—to community, or to experts<sup>26</sup>—and, second, that in (5) *S* refers to her own personal *arthritis*-concept, which is just like the concept of everyone else in *S*’s linguistic community, except that its application is not restricted specifically to *joints*. *S*’s concept properly applies to certain ailments in the thigh, for instance.

Call *S*’s deviant concept *arthritis\** and suppose, then, that by uttering (5) she speaker-means (6).

- (6) *that arthritis\* is annoying.*

Holding the various assumptions in place, it seems that if (6) indeed specifies what is said by *S* in uttering (5) on a given occasion, then (5) is misleading as a piece of linguistic evidence. Granted, if false beliefs always engender this sort of corruption it is awfully common. Furthermore, it seems like the practical upshot would often be negligible. The fact remains that in *S*’s linguistic community (5) is an ideal piece of evidence for the content in (7) and *not* for (6).

- (7) *that arthritis is annoying.*

It is easy, however, to concoct cases where conceptual mismatch of this sort would be of considerable consequence. A doctor who hears *S* talk of her annoying arthritis would probably be misled into looking in the wrong places. Surely this can create serious problems in individual cases.

<sup>26</sup>For a healthy dose of skepticism about the notion of deferred reference, see Greenberg (2014).

I will not argue that cases like this should be idealized away like cases involving confused identity. As I argue in the next subsection, there is a significant difference between false beliefs about the identity or distinctness of objects and other false beliefs. Only the former, on my account, corrupts the proper function of singular term utterances. Surely, however, it is possible to describe even the arthritis case as involving a false identity belief about *arthritis* and *arthritis\**. This is, however, not the best description of the speaker's mental state. Rather, the speaker has a number of false beliefs about arthritis which don't involve the identity relation. Furthermore, the idea of deferred reference is more plausible for the arthritis case than for the others. And that would block any more substantive analogy between the two kinds of case.

The purpose of the current exercise has just been to establish a similar case where the corruption of the linguistic evidence can be seen more clearly, given a few well known, even if controversial, assumptions. I conclude that putative speech acts of singular reference are in principle corrupted by a speaker's confused mental state.

## 2.3 Edenic reference is explanatorily basic

For now, let's focus only on proper names as devices of reference. I assume that what I will say can be extended to other expressions such as demonstratives and indexicals, but this would involve obvious complications. With a respectful nod to Ruth Millikan's (1984, 1989a, 1989b) work on proper functions of linguistic devices, I propose the following thesis about how to explain the successful introduction and perpetuation of the practice of referring with a proper name in a population of speakers.

### Normal names

If the speech act-type *A* of uttering proper name *n* to refer to object *o* is to have the proper function of reliably providing evidence for the intention to refer to *o* in group *G*, the normal explanation of the stabilization of *A* in *G* must specify a historical series of all and only the edenic *n*-utterances in *G*.

Names are normal to the extent to which they have a normal explanation of this sort. On this view, the notion of e-reference specifies part of the normal condition for a proper name. This means that a normal explanation of the proper function of the name will only mention those historical occasions on which this particular condition of normality was in fact satisfied. So, whenever a speaker makes a confused utterance of a name, the utterance malfunctions, since it is not part of the historical series of normal utterances of the name.

But why should we believe that e-reference is part of the normal condition of a proper name in the first place? Well, if every single proper name utterance in some population of speakers were confused—so that there would be no edenic

occurrences of names in the historical sequence of such utterances—there would not be much point in having names at all. At the very least, their point would be something radically different and the practice of uttering names, as we know it, would not have emerged and stabilized.<sup>27</sup>

An abnormal name, i.e. a name which lacks a normal explanation, will fail to perform the proper function of evidencing a singular referential intention. This may appear too obvious to bear spelling out. But let me explain why the possibility of an abnormal proper name ought to be surprising to theorists of a broadly externalist persuasion. One major insight of externalism in philosophy of language and mind was that linguistic reference could be secured even in the face of quite extreme types of speaker ignorance and error (e.g. Donnellan 1970; Kripke 1980; Putnam 1973; Wittgenstein 1953: §79). And the insight constitutes a compelling objection to some simple varieties of descriptivism about names.<sup>28</sup> To illustrate, even if some of the ancients seem to have believed falsely that the Earth was flat and at the center of the universe we have no problem attributing true beliefs about the Earth to them. They presumably believed *that the Earth was below their feet*, where the object of this belief is identical to the object of latter-day Earth-beliefs (Devitt 1991: 159–160).

Thus there is a distinction between extreme cases of confused identity and other cases of substantial ignorance and error. A group of three speakers where the majority of each speaker's Earth-beliefs are false and no two of them share the same false Earth-beliefs would still be able to communicate with each other about the Earth. If we picture the same situation except that the beliefs in question are all false beliefs about the identity or distinctness of the Earth the practice of using the name wouldn't get off the ground—the name wouldn't function properly in interpersonal communication.<sup>29</sup> Therefore, I suggest, even on broadly externalist assumptions there is a type of error, namely confused identity, that can disrupt the evidential function of singular terms. As in §2.2 above, however, we must assume that the confused speaker can't be saved here by some notion of deferred reference.

Invoking the Millikanian idea of normal explanation, we can say that the normality of name *n* of *o* in group *G* is explained by the historical series of edenic *n*-utterances in *G* such that uttering *n* on those occasions provided good evidence for an intention to refer to *o*. Past *n*-utterances made by confused members of *G* are excluded from the series and thus don't contribute to explaining the stabilization

<sup>27</sup>Compare what Millikan says about the imperative mood: "If no token of the imperative mood ever effected more than an abortive attempt or intention to comply with it, it is clear that speakers would soon cease to use the imperative forms at all or to use them as they now do" (1984: 56). Note, however, that I am only endorsing this as a legitimate *form* of explanation and I don't necessarily accept Millikan's particular explanation of the imperative mood. Much care is needed in formulating normal conditions for utterance types. See Origgi & Sperber (2000) for some guidance on to how to do this.

<sup>28</sup>But apparently not to others, see, e.g., Bach (1987: 157–159).

<sup>29</sup>Michael Devitt (1974: 201, 1981) makes a very similar point.

and perpetuation of the *n*-using practice. This is the sense in which e-reference pinpoints an explanatorily basic category of referring with a name; speakers must refer edenically if the name is to perform its proper evidential function and its continued existence is to be explained. E-reference is the normal condition of names, in Millikan's sense of the term. The normal condition of proper names is, then, a core causal factor, or difference-maker, in explaining why speakers are so successful in communicating object dependent propositions by using names.<sup>30</sup>

Consider, however, the following objection. Take Evans' (1973: 19) 'Ibn Kahn' example, mentioned above. This is a case where, arguably, a single proper name has a causal-historical chain of communication with a segment of faulty sections, where reference was non-edenic for a while before returning to edenic reference. Shouldn't an explanation of the practice of using the name mention those non-edenic referential acts? Well, of course, there is a grain of truth in the question, but let's look at the example more carefully.

Simplifying considerably, the name 'Ibn Kahn' goes through three relevant historical stages: (i) it refers to the scribe in his own lifetime, (ii) it refers to the person who both signed the manuscript and authored the mathematical proofs and, finally, (iii) it refers to the author of the proofs. Since the person who signed the manuscript is different from the author of the proofs, stage (ii) is the non-edenic section of the chain. It is assumed here that stage (iii) is mostly edenic, since the mathematicians will have no reason to even consider the identity of the scribe. Further, it is only when speakers have disabused themselves of any lingering confusion from stage (ii) that the name becomes a reliable indicator of an intention to refer to the mathematician.

Now, the response to the objection also has three parts. First of all, notice that stage (ii) in the alleged causal-historical chain of 'Ibn Kahn' is not necessary. Contemporary speakers could simply have started using the name to refer to the mathematician immediately without any confusion (e.g., by being generally uncertain whether names occurring alongside texts in the relevant period signify their authors, but deciding to use it thus anyway). That is to say, the situation does not rationally compel a period of confused reference. Secondly, stages (i) and (iii) should not be considered stages of the same proper name on anyone's account, even if the first point is not accepted. In (i) the name is intended to refer to the scribe but in stage (iii) it has developed fully into a name intended to refer only to the mathematician. The names are the same in name only. Thirdly, and finally, if (i) and (iii) were stages of the same proper name—i.e., they either both refer to the scribe (by some lucky coincidence) or both refer to the mathematician (changing the example significantly)—then the definition of normal names has no problem

<sup>30</sup>So, even if e-reference is partly defined in terms of a prior notion of 'referring with,' the claim is that e-reference is the notion of reference that is *explanatorily* fundamental in theorizing about communication. To be blunt, the hypothesis is that a notion of reference that abstracts away from Frege-type puzzles will play the reference-role in our final theory.

with the case at all. For then stage (ii) will simply be a period in which lots of confused, non-edenic acts of reference occurred. This tends to increase the probability that the name-using practice collapses since the relevant speech acts won't serve as reliable evidence for the referential intention in question but, still, the normal explanation will pick out all and only the edenic parts of the causal-historical chain.<sup>31</sup> Similarly, Millikan would argue, malfunctioning hearts are part of the causal-historical chain—they are *hearts*—but the conditions of their reproduction cannot have been normal and, so, those conditions are not involved in explaining the continued existence of biological devices with the proper function of pumping blood.

### 3 Objection from deliberate obfuscation

Speakers who intentionally and consciously aim to deceive their hearers seem to present unique problems for views like the edenic theory. Eliot Michaelson presents the objection as follows (personal communication).<sup>32</sup> Consider a variant of one of Kripke's (1977: 111) much discussed examples. You and I are taking our usual early morning stroll. We see a man in the distance. I know it is Smith, our mutual friend. I know also that you have not recognized the man yet. I decide to deceive you into believing that the man in the distance is not Smith but Jones. Jones is another acquaintance of ours. So I utter

(8) Jones is raking the leaves.

Assuming that I succeed, you come to believe

(9) *that Jones is raking the leaves*

By now you are truly confused, since you believe that Jones is the man over there while the man over there is in fact Smith. Thus, if I succeed, I also induce in you another belief, one which I, as the speaker, would describe as the belief

(10) *that Smith is raking the leaves.*

It seems to follow that, on the edenic theory, if you were then to utter (8) in the same context your referential intention will probably be in conflict and the utterance evidentially corrupt. Your act will be non-edenic.

But what about my own original use of the name in (8)? Is my act of referring edenic, or confused? I do not believe that Jones is Smith. I know full well that I am referring to Smith in uttering (8). But I am, surely, and knowingly, also referring to Jones, because I intend to produce in you the belief (9). Thus we appear to have a

<sup>31</sup>Similar arguments would apply to Evans' (1973: 11) 'Madagascar' example.

<sup>32</sup>Michaelson explores some similar examples in his (2013) and 'Speaker's Reference, Semantic Reference, Sneaky Reference' (ms.).



case where my referential intention is in conflict although I have no relevant false beliefs involving identity or distinctness. Furthermore, there is nothing to suggest that my performance is pragmatically or semantically incompetent: I am using language in a somewhat unconventional manner, in order to deceive, that's all. It seems to follow, then, that confused or conflicting referential intentions pose no particular problem as such and should not be marked as explanatorily irrelevant in an account of communication via devices of singular reference.

All of this is compatible with the theory on offer here since, as should be clear, only the combination of certain kinds of false *belief* and an act of reference is excluded by the notion of e-reference. The case at hand is interestingly different. Speakers who utter (8) in such a deliberately obfuscating way are doing at least two things at the same time. First, they intentionally but implicitly introduce a new name for an individual, i.e. Smith. In John Perry's (2012: §6.3) terms they instigate a permissive convention; the interlocutors are now allowed to use 'Jones' as a name for the man raking leaves. This succeeds only if the convention is endorsed—implicitly or explicitly—by the hearer. If the hearer's confusion persists, they will not have realized that a new name-using practice has in effect been introduced. Since the speaker knowingly and deliberately uses the name in a manner that has not been established in any relevant idiolect it is undeniable that she is introducing a new communicative convention. It was not customary, for the speaker and hearer, to utter 'Jones' intending thereby to refer to Smith.

This sort of thing happens all the time. Just picture yourself listening to someone who is an expert in a field you know nothing about. The expert will knowingly make utterances by which they both introduce a new name into your idiolect and say something about the object thus named. A botanist may point to a plant and say, "*Pinguicula vulgaris* here is a carnivorous plant." In this case, of course, the practice was already part of the speaker's idiolect. Importantly, however, introducing new expressions can be more or less explicit, more or less misleading, deceptive, obfuscating, and so on.

This brings us to the other thing the speaker is doing by uttering (8) in this particular context. The speaker deliberately induces the hearer to believe that 'Jones' is being used in a way that is entirely familiar rather than novel, namely that it is being used to refer to Jones. A convenient and helpful way of representing this—but by no means the only way—is to posit two homophonic names 'Jones<sub>1</sub>' and 'Jones<sub>2</sub>'.<sup>33</sup> 'Jones<sub>1</sub>' is the name just introduced for the man over there, i.e. Smith, and 'Jones<sub>2</sub>' is the name of our friend Jones in our shared language. In these terms, the second aim of the speaker is to utter 'Jones<sub>1</sub> ...' while producing in the hearer the belief that 'Jones<sub>2</sub> ...' was actually uttered.

If this is right, the example is quite atypical, but has more in common with novel

<sup>33</sup>Talking in terms of two distinct names is only a simplification here. It would perhaps be more accurate to talk of one name with two bearers. The speaker would still be introducing a new convention, namely the convention that Jones' name can now also be used to refer to Smith.



or creative uses of language than with confused reference, as here understood. It was, of course, never my contention that atypical reference is only engendered by confused speakers. One advantage of the foregoing analysis is that it captures a range of appropriate responses to (8). For example, if the hearer is not fooled and realizes exactly what the speaker was trying to do, they might say many different things, and on my account, these are both felicitous:

(11) Yeah, sure, *Jones<sub>1</sub>* over there is raking the leaves.

(12) That's Smith. *Jones<sub>2</sub>* is in Luang Prabang.

In (11) the hearer is implicitly endorsing the new convention, at least temporarily, and with a hint of irony—as in: *you didn't fool me*. The speaker might even say right away, with a suggestive intonation on the name: '*Jones<sub>1</sub>* has really changed since I last saw him.' By uttering (12) the hearer rejects the new name.

Consider the expert botanist again. Call him Boaz. He spots a beautiful *Lamp-rocaphnos spectabilis* in the distance and says to you,

(13) I see a bleeding heart on the ground there.

Clearly, Boaz can utter (13) intending 'bleeding heart' as the name of a plant while, in the same breath, intending for you to believe—if only for a second—that there is an actual animal heart there. This is intuitively described in terms of two different linguistic expressions. To be clear, however, such a description is not compulsory. If the use of the 'new name' is a one-off thing—e.g., because the deceiver corrects the error immediately—the case is not usefully described as involving the partly covert introduction of a new convention. If so, the case can be described more simply as one where a single utterance, like (8), is intended to produce two object-dependent beliefs that differ only in their respective objects, like (9) and (10), although this fact is not explicitly recognized by the hearer.

So, there is one crucial difference between genuine confusion cases and deliberate obfuscation of this kind. In the latter, one can always identify an intentionally obfuscating aspect of the referential intention. In (13) Boaz's obfuscation consists in intending to produce in you the false belief that there is a bloody heart there. In the Smith-Jones case, I intended to produce the false belief (9). Confused utterances are more simple, in that neither deceptive intentions nor the intention to use a word in a new or creative way need be involved. Using 'Bill' as a name for both Bill and Biff can surely be creative, and novel, but it is not intentionally so in the mouth of a confused speaker.

## 4 Comparison to King's account

The edenic theory can be summarized roughly as follows. When an edenic speaker *S* utters a singular term *e* on some occasion the referent of *e* is the object *S* intends

to refer to with *e* on that occasion. If *S* is confused with respect to the term and object, the theory doesn't predict that there is a unique *e*-referent, but there will invariably be more or less useful interpretations of the speaker's utterance from the hearer's point of view—such as taking the speaker to refer to the object that an edenic speaker *would have* referred to by making the same type of utterance in a similar context. Depending on the facts about the confused speaker's mental state, the intended referent will either be two distinct objects or no object at all.

Let's pause to explain this thought a little better. If the speaker is confused, their referential intention will determine either two objects or their possible fusion (Bill and Biff) on the one hand, or a non-actual possible object or no object at all (the Superman who is not Clark Kent) on the other. Whatever is true metaphysically, none of these options would, presumably, be intended or expected by the speaker and this is why the associated act of reference is defective. But the hearer is always, to some extent, helped by implicit knowledge of the Edenic maxim. To understand Lois Lane, for example, her addressee may implicitly take her, on occasion, to be referring to a person who is both Superman and Clark Kent, because that is what an edenic speaker would have done. Similarly, considering Rose again, who confuses Bill and Biff, the Edenic maxim may sometimes lead her addressee to construe her as referring only to Bill or only to Biff, because that is what an edenic speaker would have done. There is, then, a specific search-and-rescue strategy involved in the charitable interpretation of non-edenic speakers. Rose is not automatically taken, say, to believe that proper names are shared between identical twins, or that similar people can be referred to with the same name willy-nilly (but people sometimes seem to adhere to the latter procedure, e.g., with 'Einstein').

This theory is usefully compared to a view recently defended by Jeffrey King (2013, 2014), which is very similar to ideas originally proposed by Stephen Neale (2005: 184–185). King thinks of it as a theory of demonstrative reference, calling it the 'coordination account'. The idea is, roughly, that an utterance of a singular term on a given occasion has a referent *o* if and only if two conditions are satisfied. (1) The speaker intends to refer to *o* by uttering the expression on that occasion. (2) A rational, attentive, reasonable hearer *would take o* to be the object intended. Thus the reference must be such that an ideal—i.e., rational, attentive, reasonable, etc.—interpreter would come up with the correct interpretation. The two conditions are necessary for reference to occur, so, if either one fails in a given case, the demonstrative gets no object assigned as a referent. Neale's original proposal is slightly different, as he argued more directly that there is no need for a notion of what the utterance actually referred to, as distinct from what the speaker intends to refer to or what a rational interpreter would take the speaker to refer to. In explaining referential communicative success, Neale argued, we only need (1) and (2) and the idea of their possible convergence. But much of the motivation for the edenic theory could possibly be marshalled in support of this kind of account and vice versa. Both proposals argue for ways in which speech acts of referring ought

to be idealized.

I will not argue against the account here, only note the most significant point at which the two theories differ, namely in their respective conceptions of successful communication. On King's construal of the view, but not on Neale's, the second condition in effect defines the goal of communication in terms of an ideally rational and well-informed hearer. This is because the utterance has a referent if and only if such a hearer *would* latch on to the object intended. And, presumably, the ordinary hearer's task is to latch on to that actual referent. On the edenic theory, however, the goal of the hearer is to discover the speaker's communicative intention, full stop. It is not to discover how a reasonable hearer *would* interpret the speaker on the occasion of utterance. Consider an example. Suppose *B* is not competent, rational or reasonable. Suppose also that *A* knows all of this about *B*. And, then, *A* utters  $\sigma$  intending to refer to object *o*. It should follow that, in such a case, this interpretation is *not* one which a rational, reasonably well-informed hearer would come up with. The utterance is specifically designed by *A*, on grounds of *A*'s beliefs about the incompetence of *B*, for a hearer that's not so constituted. Of course, however, communication succeeds and *B* understands that by uttering  $\sigma$ , *A* said something about *o*. And, further, *o* is objectively what *A* referred to—and possibly e-referred to—on this occasion.

The account, as here understood, seems to make a different prediction. For presumably it is possible, on this view, that a rational, well-informed hearer would have interpreted *A* as referring to  $o'$ , where  $o \neq o'$ , even if *A* happens not to be speaking to such a hearer. The account then predicts that the utterance of the expression has no referent on this occasion. If King's proposal is understood differently, allowing him to insist that of course the ideal hearer would have been aware of the speaker's intention to exploit the actual hearer's incompetence, it would tend to make the second condition superfluous. The second condition would, then, tend to state merely that the ideal hearer is one who always recognizes the speaker's actual intention. The idealization really only consists in saying that the hearer understands correctly, whichever object the speaker intends. The edenic theory is, in this regard, quite different. The idealizing constraint is only on the speaker's side of the equation, skipping the ideally rational hearer completely; except when it comes to the epistemology of interpretation or strategies of intention-recognition. More specifically, the edenic theory follows Grice's advice in 'Meaning revisited,' where he writes that a full account of speaker meaning calls for the *absence* of certain 'sneaky' intentions in the speaker (1989: 302–303). According to the theory false identity beliefs give rise to sneaky—although not deliberately so—referential intentions, the absence of which needs to be posited in an adequate explanation of how singular communication works.

## 5 Conclusion

Edenic reference is defined as the speech act of referring to an object in speaker-meaning a proposition without the speaker having, at the time of utterance, any false beliefs about the identity of the object, such that those beliefs result in the speaker's referential intentions being confused or in conflict. I have argued, first, that this notion of e-reference is part of what a speaker implicitly knows in knowing a language or, more specifically, part of their pragmatic competence. Second, that confused referential intentions have characteristic breakdown effects on the normal and proper functioning of linguistic communication. Third, that our final theory of meaning and communication should make e-reference, or something very much like it, play the role of the reference relation: it specifies the relation holding between a speaker, expression, and object, when the mechanism of communication performs in accordance with its optimal and proper function. The resulting view is what I call the edenic theory of reference and, as promised in the beginning, it constitutes a completely speaker-based, intentionalist theory of the reference relation. The theory does not avoid idealization, but it does avoid idealizations that are not based directly on actions and propositional attitudes of the speaker on the occasion of utterance. At the very least, this is why the proposal is a novel one, and worth serious consideration.

## Acknowledgements

This paper has benefited from discussion, suggestions, and encouragement from Brian Ball, Daniel Harris, Cressida Gaukroger, Dominic Alford-Duguid, Ben Phillips, François Recanati, Stephen Neale, Gary Ostertag, Michael Devitt, Deirdre Wilson, Eliot Michaelson, Matt Moss, Finnur Dellsén, Thomas Hodgson, Rachel McKinney, Daniel Deasy, Maria Baghramian, Hrafn Ásgeirsson, and Indrek Reiland. Thanks also to audiences at PLM 3 in Oslo, SPE 8 in Cambridge, The M&L Seminar at Institut Nicod, at the 2016 Pacific APA in San Francisco, and thanks to several anonymous referees. Finally, many thanks to Alexandru Radulescu, who provided me with very helpful and engaging comments at the Pacific APA. I gratefully acknowledge the generous support of the Icelandic Centre for Research (163132-051) and the Irish Research Council (GOIPD/2016/186) while working on different parts of this paper.

## Bibliography

Åkerman, J., 2010. "Communication and Indexical Reference." *Philosophical Studies*, 149(3):355–366.

- Arjo, D., 1996. "Sticking Up for Oedipus: Fodor on Intentional Generalizations and Broad Content." *Mind and Language*, 11(3):231–45.
- Aydede, M. & Robbins, P., 2001. "Are Frege Cases Exceptions to Intentional Generalizations?" *Canadian Journal of Philosophy*, 31(1):1–22.
- Bach, K., 1987. *Thought and reference*. Clarendon.
- , 1992. "Paving the Road to Reference." *Philosophical Studies*, 67(3):295–300.
- Bach, K. & Harnish, R., 1979. *Linguistic communication and speech acts*. MIT Press.
- Bar-On, D., 2013. "Origins of meaning: Must we 'go Gricean'?" *Mind & Language*, 28(3):342–375.
- Bertolet, R., 1987. "Speaker reference." *Philosophical Studies*, 52(2):199–226.
- Breheny, R., 2006. "Communication and Folk Psychology." *Mind and Language*, 21(1):74–107.
- Burge, T., 1979. "Individualism and the mental." *Midwest Studies in Philosophy*, 4:73–121.
- Burgess, A. & Plunkett, D., 2013a. "Conceptual Ethics I." *Philosophy Compass*, 8(12):1091–1101.
- , 2013b. "Conceptual Ethics II." *Philosophy Compass*, 8(12):1102–1110.
- Camp, J. L., 2002. *Confusion: A study in the theory of knowledge*. HUP.
- Carruthers, P., 1996. *Language, thought and consciousness*. CUP.
- Carston, R., 1998. "The semantics/pragmatics distinction: a view from relevance theory." *UCL Working Papers in Linguistics*, 10:1–30.
- , 2002. *Thoughts and utterances*. Blackwell.
- Chalmers, D., 2006. "Perception and the fall from Eden." T. Gendler & J. Hawthorne (eds.), *Perceptual experience*, OUP, pp. 49–125.
- Chomsky, N., 1980. *Rules and representations*. CUP.
- Dellsén, F., 2016. "Scientific Progress: Knowledge Versus Understanding." *Studies in History and Philosophy of Science Part A*, 56:72–83.
- Devitt, M., 1974. "Singular terms." *The Journal of Philosophy*, 71(7):183–205.
- , 1981. *Designation*. CUP.
- , 1991. *Realism and truth*. Princeton University Press, 2 edn.
- Donnellan, K., 1968. "Putting Humpty Dumpty together again." *The Philosophical Review*, 77(2):203–215.
- , 1970. "Proper names and identifying descriptions." *Synthese*, 21(3-4):335–358.
- Elgin, C., 2007. "Understanding and the Facts." *Philosophical Studies*, 132(1):33–42.
- Evans, G., 1973. "The Causal Theory of Names." *Aristotelian Society Supplementary Volume*, 47(1):187–208. Repr. in Evans (1985), pp. 1–24.
- , 1982. *The varieties of reference*. Clarendon.
- , 1985. *Collected papers*. Clarendon.
- Field, H., 1973. "Theory change and the indeterminacy of reference." *The Journal of Philosophy*, 70(14):462–481.
- Fodor, J., 1994. *The elm and the expert*. MIT Press.

- García-Carpintero, M., 2001. "Gricean rational reconstructions and the semantics/pragmatics distinction." *Synthese*, 128(1-2):93–131.
- Gauker, C., 2003. *Words without meaning*. MIT Press.
- , 2008. "Zero tolerance for pragmatics." *Synthese*, 165(3):359–371.
- Grandy, R. & Warner, R., 1986. "Paul Grice: A view of his work." R. Grandy & R. Warner (eds.), *Philosophical grounds of rationality: Intentions, categories, ends*, Clarendon Press.
- Green, M. S., 2007. *Self-expression*. OUP.
- Greenberg, M., 2014. "Troubles for content I." A. Burgess & B. Sherman (eds.), *Metasemantics: New essays on the foundations of meaning*, OUP, pp. 147–168.
- Grice, P., 1969. "Utterer's meaning and intention." *The Philosophical Review*, 78(2):147–177.
- , 1982. "Meaning revisited." N. Smith (ed.), *Mutual knowledge*, Academic Press, pp. 223–243. Repr. in Grice (1989), pp. 283–303.
- , 1989. *Studies in the way of words*. HUP.
- Hanks, P., 2011. "Structured propositions as types." *Mind*, 120(477):11–52.
- , 2015. *Propositional content*. OUP.
- Harris, D. W., 2016. "Intentionalism versus The New Conventionalism." *Croatian Journal of Philosophy*, 16(47):173–201.
- Heck, R., 2012. "Solving Frege's Puzzle." *Journal of Philosophy*, 109(1/2):132–174.
- Kaplan, D., 1978. "Dthat." *Syntax and Semantics*, 9:383–400.
- King, J. C., 2013. "Supplementives, the coordination account, and conflicting intentions." *Philosophical Perspectives*, 27(1):288–311.
- , 2014. "Speaker intentions in context." *Noûs*, 48(2):219–237.
- Kripke, S., 1977. "Speaker's reference and semantic reference." *Midwest Studies in Philosophy*, 2:6–27.
- , 1980. *Naming and necessity*. HUP.
- Lance, M. N. & O'Leary-Hawthorne, J., 1997. *The grammar of meaning: Normativity and semantic discourse*. CUP.
- Lawlor, K., 2007. "A notional worlds approach to confusion." *Mind & Language*, 22(2):150–172.
- Lewis, K. S., 2013. "Speaker's reference and anaphoric pronouns." *Philosophical Perspectives*, 27(1):404–437.
- Loar, B., 1986. *Mind and meaning*. CUP.
- McDowell, J. H., 1980. "Meaning, communication, and knowledge." *Philosophical subjects: Essays presented to P. F. Strawson*, Clarendon Press, pp. 117–39.
- Michaelson, E., 2013. *This and that: A theory of reference for names, demonstratives, and things in between*. Ph.D. thesis, UCLA.
- , ms. "Speaker's Reference, Semantic Reference, Sneaky Reference."
- Millikan, R. G., 1984. *Language, thought, and other biological categories: New foundations for realism*. MIT Press.
- , 1989a. "Biosemantics." *The Journal of Philosophy*, 86(6):281–297.

- , 1989b. "In defense of proper functions." *Philosophy of Science*, 56(2):288–302.
- , 1997. "Images of identity: In search of modes of presentation." *Mind*, 106(423):499–519.
- , 2000. *On clear and confused ideas: An essay about substance concepts*. CUP.
- Mizrahi, M., 2012. "Idealizations and Scientific Understanding." *Philosophical Studies*, 160(2):237–252.
- Neale, S., 1992. "Paul Grice and the philosophy of language." *Linguistics and Philosophy*, 15(5):509–559.
- , 2004. "This, that, and the other." M. Reimer & A. Bezuidenhout (eds.), *Descriptions and beyond*, Clarendon, pp. 68–182.
- , 2005. "Pragmatism and binding." Z. G. Szabó (ed.), *Semantics versus pragmatics*, Clarendon, pp. 165–285.
- , 2016. "Silent reference." G. Ostertag (ed.), *Meanings and other things: Themes from the work of Stephen Schiffer*, OUP, pp. 229–344.
- Origgi, G. & Sperber, D., 2000. "Evolution, communication and the proper function of language." P. Carruthers & A. Chamberlain (eds.), *Evolution and the human mind*, CUP, pp. 140–169.
- Perry, J., 2012. *Reference and reflexivity*. CSLI Publications. 2nd ed.
- Putnam, H., 1973. "Meaning and reference." *The Journal of Philosophy*, 70(19):699–711.
- Recanati, F., 2012. *Mental files*. OUP.
- Reimer, M., 1992. "Three views of demonstrative reference." *Synthese*, 93(3):373–402.
- , 2004. "What malapropisms mean: A reply to Donald Davidson." *Erkenntnis*, 60(3):317–334.
- Schiffer, S., 1972. *Meaning*. OUP.
- , 1981. "Indexicals and the theory of reference." *Synthese*, 49(1):43–100.
- , 1982. "Intention-based semantics." *Notre Dame Journal of Formal Logic*, 23(2):119–156.
- , 2003. *The things we mean*. Clarendon.
- , 2006. "Two Perspectives on Knowledge of Language." *Philosophical Issues*, 16(1):275–287.
- , 2014. "Meaning and formal semantics in generative grammar." *Erkenntnis*, 80(1):61–87.
- Schneider, S., 2005. "Direct Reference, Psychological Explanation, and Frege Cases." *Mind and Language*, 20(4):423–447.
- Schwarz, D. S., 1976. "Referring, singular terms, and presupposition." *Philosophical Studies*, 30(1):63–74.
- Scott-Phillips, T., 2015. *Speaking our minds: Why human communication is different, and how language evolved to make it special*. Palgrave Macmillan.
- Segal, G. M. A., 1997. "Content and Computation: Chasing the Arrows, a Critical Notice of Jerry Fodor's the Elm and the Expert." *Mind and Language*, 12(3&4):490–



- 501.
- Soames, S., 2010. *What is meaning?* Princeton University Press.
- , 2014. “Cognitive propositions.” J. C. King, S. Soames, & J. Speaks (eds.), *New thinking about propositions*, OUP, pp. 91–124.
- Speaks, J., 2016. “The role of speaker and hearer in the character of demonstratives.” *Mind*, Online First.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D., 2010. “Epistemic vigilance.” *Mind & Language*, 25(4):359–393.
- Sperber, D. & Wilson, D., 1995. *Relevance: Communication and cognition*. Blackwell, 2 edn.
- Stine, G. C., 1978. “Meaning other than what we say and referring.” *Philosophical Studies*, 33(4):319–337.
- Strawson, P. F., 1950. “On referring.” *Mind*, 59(235):320–344.
- , 1974. *Subject and predicate in logic and grammar*. Ashgate Publishing, 2 edn. Pub. 2004.
- Strevens, M., 2016. “How Idealizations Provide Understanding.” S. Grimm, C. Baumberger, & S. Ammon (eds.), *Explaining Understanding: New Essays in Epistemology and the Philosophy of Science*, Routledge.
- Thompson, J. R., 2014. “Meaning and Mindreading.” *Mind and Language*, 29(2):167–200.
- Tomasello, M., 2008. *Origins of human communication*. MIT Press.
- Unnsteinsson, E., 2014. “Compositionality and sandbag semantics.” *Synthese*, 191(14):3329–3350.
- , 2016. “Confusion is Corruptive Belief in False Identity.” *Canadian Journal of Philosophy*, 46(2):204–227.
- , 2017. “A Gricean Theory of Malaprops.” *Mind and Language*, 32(4):446–462.
- Wakefield, J. C., 2002. “Broad Versus Narrow Content in the Explanation of Action: Fodor on Frege Cases.” *Philosophical Psychology*, 15(2):119–33.
- Weisberg, M., 2007. “Three Kinds of Idealization.” *Journal of Philosophy*, 104(12):639–659.
- Wettstein, H. K., 1984. “How to bridge the gap between meaning and reference.” *Synthese*, 58(1):63–84.
- Wharton, T., 2009. *Pragmatics and non-verbal communication*. CUP.
- Williamson, T., 2000. *Knowledge and its limits*. OUP.
- Wilson, D. & Sperber, D., 2012. *Meaning and relevance*. CUP.
- Wittgenstein, L., 1953. *Philosophical investigations*. Wiley-Blackwell, 4 edn. Trans. by G. E. M. Anscombe, P. M. S. Hacker, J. Schulte (2007).