Michael Devitt

There is much talk of intentions in semantics. Intentions to refer are said to determine reference. The central idea of Gricean “intention-based-semantics” is that speaker meanings are constituted by the speakers’ communicative intentions. Finally, it is common, particularly among Griceans, to believe that there is a constraint on what a speaker can intend to communicate by an utterance. I think that all of this is mistaken.

MISTAKE I: INTENDING TO REFER

Stephen says, “Grice is great”; or, pointing to Grice, “He is great” or “That man is great”. Who does Stephen refer to with these “singular referring terms”? One answer that is very popular among philosophers is that he refers to “whoever he intends to refer to” (#Schiffer, Bach, Neale, King). Another answer that has some popularity among philosophers is that he refers to “the particular person he has in mind” (#Strawson, Donnellan, Devitt, neo-Donnellanians##). Both answers are initially appealing because they are in line with what the folk might say in answer to our question. I think that having-in-mind is indeed a helpful starting point for a good answer, but no more than that. Intending-to-refer is not even that. I shall begin by arguing for this and then make three more objections to the intending-to-refer hypothesis. I conclude that it should have no place at all in a theory of language.

Objection 1: Implausible Starting Point

What is wrong with intending-to-refer as a starting point? Think of what is required for someone doing $A$ to intend to $V$ to $x$; for example, what is required for someone leaving the apartment to intend to bicycle to the Met. This intention is a thought, a “propositional attitude”, about bicycling and so contains the concept bicycling. Similarly, for a speaker uttering a singular referring term in a sentence to literally intend to refer to $x$ with the term, she must have a thought containing the concept reference. So, according to the intending-to-refer hypothesis, she can’t refer without thinking about reference! There is no reason to believe this, no reason to believe that any expression of a thought about something must be accompanied by a further thought about reference. That is far too intellectualized a picture of referring and is psychologically implausible. Referring is a cognitive skill, mere know-how; or so I have argued (1981a, 1996, 2006, 2011). One could refer without even having a concept of reference (1981a: 97). This is not to deny that referring is intentional, in some sense: it’s an action like walking, kissing, etc. (And we would make no theoretical progress by saying that the speaker refers to $x$ in virtue of intentionally referring to $x$: referring is intentional.) And it is not to deny that any normal adult speaker with a minimal education could probably tell you after the event what she was “talking about”. That’s a very easy bit of semantic knowledge. Yet even that easy bit is surely beyond the

---

1 Objections 1 and 2 draw on my 2015: 110-11.
capacity of many organisms that nonetheless refer: for example, humans at the age of three; bees, prairie dogs, vervet monkeys, and other nonhuman species that have what cognitive ethologists call “referential languages”. In the face of this, one could of course restrict the intending-to-refer hypothesis to adult humans, but that seems ad hoc. In any case, to repeat, it is not plausible that such metalinguistic intentions are constitutive of anyone’s referring.

Having-in-mind does not have this problem as a starting point. For, having \( x \) in mind simply requires that the part of the thought expressed that causes the use of the term refers to \( x \).

That is the first objection to including intending-to-refer in a theory of language. There are three more objections, two of which apply also to having-in-mind.

**Objection 2: Incomplete**

Even if intending-to-refer was a plausible starting point it would immediately raise the question: *In virtue of what* did the speaker intend to refer to \( x \) (rather than to \( y \) or \( z \)) in using the singular term? Without an answer to this question, the intending-to-refer hypothesis could not advance our theory of reference significantly: the explanatory problem has simply been moved a short distance from the reference of the utterance to the reference of the intention.

Now, of course, the having-in-mind hypothesis raises an analogous question and so is similarly incomplete: In virtue of what did the speaker have \( x \) in mind? In virtue of what does the part of the thought that causes the use of the term refer to \( x \). What reality is this somewhat vague folk talk getting at?

I have argued that, for the sort of having-in-mind and singular reference exemplified in Stephen’s utterances, the reality is causal (1974, 1981a,b).\(^3\) The part of Stephen’s thought determines that he had \( x \) in mind (in this sense) in virtue of standing in a certain sort of causal relation to \( x \), a relation involving the perceptual grounding of someone’s thought in \( x \) and, perhaps, reference borrowings (Kripke’s wonderful insight).\(^4\) The details of this causal relation

\(^2\) Developmental evidence suggests that the capacity to have metalinguistic thoughts comes later, in middle childhood; e.g. Hakes 1980

\(^3\) I was clear from the start that the vague talk of “having in mind” is but a “stepping stone” to a causal theory of reference, not something that features in the theory (1974: 202). Still, Andrea Bianchi’s criticisms (2019) have made it clear to me that I should have been more careful in my handling of this stepping stone. I should not have claimed, early, to offer an “analysis” (1974: 202) of “having in mind” and, late, “an explanation - better, an explication - of this somewhat vague folk talk” in causal terms (2015c: 111). The folk talk arguably covers more than the causal relation we want; see how it can mislead (Objection 4). So my causal explanation was, in reality, an explanation of a restricted, “technical”, notion of having-in-mind.

\(^4\) Joseph Almog has recently proposed a similar causal explanation of having in mind (2012: 177, 180-2), attributing it to Donnellan without evidence; for discussion, see Devitt 2015: 111 n. 4. As Julie Wulfemeyer has aptly remarked recently: “The grounding cognitive relation went largely unexplained by Donnellan” (2017: 2).
for demonstratives, pronouns, and names will differ a bit, of course.\(^5\) (It follows from this, note, that the reference of the term is determined by a mental state of the speaker. So, contrary to the tradition arising from formal semantics – see, e.g., David Lewis (1983)– the “indices” of the context external to the speaker’s mind play a reference-determining role only to the extent that relations to that context partly constitute that mental state.)

**Objection 3: Redundant**

Of course, the intending-to-refer hypothesis could be completed by an answer to the in-virtue-of-what question even if it seldom is. Thus the hypothesis could adopt a causal answer along the lines I have just sketched: applied to our example, Stephen intends to refer to Grice in using ‘Grice’, ‘he’, or ‘that man’ in virtue of that use being immediately caused by an intention that stands in the specified causal relation to x. Suppose that the intending-to-refer hypothesis is indeed completed by such an answer. We can see then that the hypothesis is redundant, an unhelpful excursion. For, intentions are thoughts (propositional attitudes) and the explanation of in virtue of what Stephen’s intention refers to Grice is an explanation of in virtue of what any of his thoughts refer to Grice. So it is an explanation of in virtue of what his original thought about Grice’s greatness, expressed using ‘Grice’, ‘he’, or ‘that man’, referred to Grice. And that would explain why Stephen referred to Grice in using one of those “singular referring expressions”. We would have solved our original problem without any need to appeal to metalinguistic intentions.

**Objection 4: Misleading**

I have proposed a causal theory for singular referring terms. The tradition proposed description theories for some or all of them. Take a normal use of a term ‘E’ where the speaker refers to x on either theory: so the use is causally connected to x in the way specified by the causal theory and x is uniquely described by the associated reference-determining description, ‘the F’, specified by the description theory. Then, if we are prepared to attribute a referential intention to the speaker at all - setting aside the worry in Objection 1 - we should all agree that the speaker using ‘E’ intends to refer to the F. But now suppose that the speaker not only has the identifying belief that she would express, “E is the F”, but also several others that she would express, “E is the G”, “E is the H”, and so on. In a quite ordinary sense, we should also say that the speaker, in using ‘E’ intends to refer to the G, to the H, etc. What would those intentions show about who she referred to? Nothing at all, on either theory. First, on a causal theory, whatever descriptions she associates with ‘E’ are quite beside the point of what she refers to. Indeed, the resulting intentions may well mislead us. Thus, suppose that the speaker’s belief that E is the G is false; perhaps ‘G’ uniquely describes not x but y. Then her intention to refer to the G would mislead us into thinking that she referred to y not x. Second, on a description theory, only her intention to refer to the F is a reliable guide to her reference for only it is reference-determining. So, her intention to refer to the G would be as misleading as it would be on the causal theory.

\(^5\) And the causal relation determining a proper name’s semantic reference differs from, but is built upon, that determining its speaker reference (2015: 126; 2019: sec. 5 Andrea book##).
The having-in-mind hypothesis can be just as misleading: the speaker in the above scenario might be said to have the \( F \), the \( G \), the \( H \), in mind – “having in mind” is vague enough for that - with the potentiality to mislead us about \( E \)’s reference.

The possibility that talk of intentions to refer might mislead in this way is not an idle one. In their seminal work in “experimental semantics”, Edouard Machery, Ron Mallon, Shaun Nichols, and Stephen Stich (2004; “MMNS”) conducted a Gödel experiment in which they asked subjects who John, a character in the vignette, was “talking about”. This question was criticized as ambiguous. MMNS wanted the participant’s intuition about who the name conventionally refers to, its \textit{semantic} reference. But the participant might take the prompt as asking about who the character means by the name, its \textit{speaker} reference. So the experiment may not have yielded what was wanted.\(^6\) Here is one such criticism:

Examples…suggest that a question of the form, “Who is \( S \) talking about in using \( n \)?” does not have a univocal meaning….In particular, although the communicative intention of the speaker (John) is not made explicit, the vignette may be taken to suggest that he intends to talk about the man who discovered that arithmetic is incomplete since the vignette insists that this is the only piece of information associated with “Godel.” It is thus unclear whether those participants who answer that John is talking about the man who discovered that arithmetic is incomplete are making a genuine descriptivist judgment about the semantic reference of “Godel” or rather making a judgment about speaker’s reference. Consequently, it may be that almost everybody has causal–historical intuitions about semantic reference, but that some participants…report their intuitions about speaker’s reference. (Machery et al. 2015: 67; “MSD”)

The vignette may indeed suggest that John “intends to talk about the man who discovered that arithmetic is incomplete” but, in light of our discussion, we can see that this does not provide cause for concern about the alleged ambiguity in MMNS’s question. For, that intention gives no reason for the subjects to take the speaker reference to be \textit{different from} the semantic reference. On the causal theory, that intention is irrelevant to both the semantic and speaker reference. On the description theory in question – one that has the reference determined by an associated description along the line of ‘the man who discovered the incompleteness of arithmetic’- the intention is relevant to the speaker reference \textit{only because} it is relevant to the semantic reference.

In sum, if taken literally, the intending-to-refer hypothesis is: (1) implausible; (2) incomplete; (3) redundant once completed; (4) misleading.

\section*{MISTAKE II; INTENDING TO COMMUNICATE}

What explains the speaker meaning of a sentential utterance? A central idea of Gricean “intention-based semantics” is that this meaning is constituted by the speaker’s intention \textit{to communicate a certain content to an audience}. As Stephen Schiffer puts it, “Meaning entails audience-directed intentions, and one cannot mean something without intending to be understood” (1992: 515). Stephen Neale expands the idea as follows: “In doing \( x \), \( S \), meant that \( p \)

\footnotesize{\(^6\) I think that the criticism is mistaken (Devitt and Porot 2018: 1579-80 n. 17).}
iff (roughly) for some audience, \(A\), \(S\) did \(x\) intending \(A\) to think that \(p\) via \(A\)’s recognition that \(S\) intended \(A\) to think that \(p\)” (Neale 2016: 281).

No doubt speakers *typically* produce utterances to communicate messages to an audience. And those utterances are intentional acts and might involve intentions of some sort. Furthermore, I’m inclined toward the view that representational systems like human languages owe their very existence, whether they are innate or conventional, to their role of communicating messages to audiences (2006: 130-1). But it does not follow that the *only* meaningful use of language is to convey a message to an audience. More importantly, it does not follow that it is essential to its meaningful use that it involve an intention to communicate. And it is not essential as Noam Chomsky points out:

> Under innumerable quite normal circumstances—research, casual conversation, and so on—language is used properly, sentences have their strict meaning, people mean what they say or write, but there is no intent to bring the audience (not assumed to exist, or assumed not to exist, in some cases) to have certain beliefs or to undertake certain actions. (1975: 62)

Making an utterance is indeed an intentional act but the basic act in question is better described using the popular locution, “expressing a thought”: “there is much to be said for the old-fashioned view that speech expresses thought, and very little to be said against it” (Fodor et al. 1974: 375). Expressing a thought can be done in various ways: by *speaking*, *writing*, or *signing* in some other way. And the point of this expressing may not be to communicate a message to an audience. There may *be* no audience, or the speaker may not think there is. The speaker may be musing (“To be, or not to be”), making notes, trying out a line for a poem, and so on. Even where the speaker is addressing an audience, she may not, for a variety of possible reasons, *care* if the message gets through. She may even be happier if it doesn’t! Thus, it is common – particularly, I think, among the well-educated English – for people who are bored by ordinary small talk to regard conversation as a challenge to be * cleverly and wittily indirect*. Such a person certainly does not *intend* his hearer to grasp his message. Perhaps he will *hope* that the hearer is smart enough to do so, but he may not: he may anticipate the pleasure of watching the hearer fail. In sum, communicating a thought to an audience is just *one form* of expressing that thought: it is expressing the thought with the intention of being understood by the audience.

If this is right, a speaker means that \(p\) in virtue of expressing the thought that \(p\). And it is not necessary that the speaker intends to communicate that thought to some audience.

Now some philosophers will surely be tempted to make moves to save the view that communicative intentions are essential: “when making notes one is communicating with one’s future self”; “when musing, one is communicating to a possible audience”; “when having fun at one’s audience’s expense, one is communicating… - tricky, I’ll have to get back to you on that one”. But, first, is any of this psychologically plausible? That is to say, is it plausible that speakers making utterances that are at least *apparently* non-communicative *must* have some such thought about some audience or other? *Is it really plausible that it is impossible to mean something by an utterance without thinking about an audience?* Second, and more important, what is the point of saving the view? Why do we need to go beyond *expressing a thought* to the
more demanding attempting to communicate that thought to an audience in order to explain speaker meaning? Expressing a thought will do the job. There is no theoretical motivation to require an intention to communicate the thought.

The idea that speaker meaning is constituted by communicative intentions is the fundamental mistake of intention-based semantics. But that mistake is compounded by the accounts given of such communicative intentions. Here is one of Grice’s later ones:

“U meant something by x” is true iff U uttered x intending thereby:

1. that A should produce response r
2. that A should, at least partly on the basis of x, think that U intended (1)
3. that A should think that U intended (2)
4. that A’s production of r should be based (at least in part) on A’s thought that U intended (1) (that is, on A’s fulfilment of [2])
5. that A should think U intended (4). (Grice 1989: 96-7)

Even this definition proved insufficiently complex to rule out all the suggested counter-examples (Schiffer 1972). What are we to make of this? I have just indicated the psychological implausibility of the claim that meaning something by an utterance must be accompanied by a thought about an audience A. What about the claim that it must be accompanied by this thought?! It is hardly plausible that this baroque structure of intentions is psychologically real in any speaker. Grice was sensitive to this concern from the start, disclaiming “any intention of peopling all our talking life with armies of complicated psychological occurrences” (1989: 222). But then if there are no such armies we have no theory: we are left with no explanation of speaker meaning. I know of no satisfactory solution to this problem. I think the program is misguided (Devitt and Sterelny 1999: 149-50).

I have used the ordinary locution, “expressing a thought” to describe the basic act which constitutes speaker meaning. I don’t say that the locution is perfect for the task. We want something that describes the intentional act common to speaking, writing, signaling, emailing, tweeting, and so on, something that abstracts from the differences between these behaviors. Certainly the behaviors are covered by “expressing a thought” but very likely some other ones are too. Thus, we might ordinarily say that a person’s face expresses the belief that her company is boring even if its so doing is unintentional. On the other hand, one might argue that when humans greet (“Hi”), cheer (“Bravo”), abuse (“Get lost”), and the like, they are expressing some mental state but often not a thought. So, English seems not to have an ordinary locution that perfectly describes the action we want to pick out for our theoretical purposes. And that is hardly surprising: the folk are not bent on explaining speaker meaning. Still, “expressing a thought” seems closest to what we want. So I will stick with it, perhaps narrowed to exclude the unintentional and widened to include the likes of greetings.

A further complication should be mentioned. Humans are not the only animals that make meaningful “utterances” in a language; consider, for example, bees and prairie dogs. What

---

7 This discussion of “expressing a thought” has benefited from the criticisms of Elmar Geir Unnsteinnson, Daniel Harris, and Gary Ostertag.
constitutes the “speaker meaning” of a particular bee dance? Not presumably the content of the thought expressed but rather the content of some “lesser” representational state in the bee that causes the dance. So to cover nonhumans perhaps we should take speaker meaning to be something like “the content of the inner representational state causing that use of language”. But I set this complication aside.

So what is it to “express a thought”, in this perhaps slightly technical sense? I have nothing more to say about this. And I don’t think that philosophers should expect to have anything more to say at this time. So, expressing a thought is partly unexplained in the theory, and to some extent a “primitive”. This is not to say that expressing a thought is inexplicable. Psycholinguists studying language production are in effect trying to throw light on it (see, e.g., Vigliocco and Vinson 2003). The point is just that this is now beyond philosophers.

In light of this, in virtue of what does a sentential utterance have its speaker meaning? In virtue of the causal story of its production. Thus the speaker meaning of a word in that utterance is determined by the concept it expresses on that occasion. The “speaker syntax” of the utterance’s sentence is determined by the structure of the underlying thought that the utterance expresses and by the way in which the utterance was produced from that underlying thought.

MISTAKE III: CONSTRAINTS ON INTENTIONS

Griceans standardly hold a belief that there is a constraint on what a speaker can intend to communicate by an utterance, a belief that reflects the belief that there is a constraint on intentions in general. The literature contains many versions of this constraint, some of them astonishingly strong. Here is a sample. Talking of intentions in general, Grice claimed in one place that

(1) a condition on ‘X intends to do A’ is that X “is sure that he will in fact do A” (1971: 266).

Talking of intentions to refer, Stephen Schiffer claimed that

(2) “the speaker cannot intend to refer to a particular female [by ‘she’] unless he expects his hearer to recognize to which female he is referring” (2005: 1141).

Reinaldo Elugardo and Robert Stainton make the weaker, hence more plausible, claim that the speaker’s expectations must be about what the hearer can do rather than what she will do:

(3) “the intentions that a speaker can have are importantly constrained by her reasonable expectations about what the hearer can figure out” (2004: 445).

And Grice, in another place, urges a similarly weaker constraint: talking about intentions to mean something by a hand wave HW, Grice proposes that

(4) a speaker “must (logically) be in a position, when uttering HW, to suppose that there is at least some chance that these intentions will be realized” (1989: 125).

These are all “positive” claims, requiring the person intending to have a certain belief (or “be in a position” to have one). Elsewhere Grice’s constraint is a much weaker “negative” one, requiring the person not to have a certain belief:

---

8 See also Harman 1976; Velleman 1989.
9 See also Brand 1984.
“one cannot in general intend that some result should be achieved, if one knows that there is no likelihood that it will be achieved” (Grice 1989: 101; emphasis added). In an early paper Stephen Neale urges a slightly weaker negative claim:

(6) the formation of genuine communicative intentions by $U$ is constrained by $U$’s expectations: $U$ cannot be said to utter $X M$-intending $A$ to $\varphi$ if $U$ thinks that there is very little or no hope that $U$’s production of $X$ will result in $A \varphi$-ing”. (1992: 552; emphasis added)

Later Neale urges a constraint that is much the same as Grice’s negative one:


We can capture well enough the range of possible constraints as follows, with two positive ones followed by two negative ones, the stronger first in each case:

$X$ cannot intend to $A$ unless:

P1: $X$ believes that if she will $A$. [cf. (1), (2)]
P2: $X$ believes that she can $A$, has some chance of $A$-ing. [cf. (3), (4)]

N1: $X$ lacks the belief that she will not $A$.
N2: $X$ lacks the belief that she cannot $A$, has no chance of $A$-ing. [(5), (6), (7)]

Before assessing these constraints, we should think about the force of “cannot intend”. Perhaps it yields a constitutive constraint: if the constraint is not met, it is not metaphysically possible for any mental state of $X$, whatever else it might be, to be an intention to $A$. That seems very strong. So perhaps “cannot intend” yields a weaker normative constraint: if the constraint is not met, it is not metaphysically possible for $X$ to rationally intend to $A$ (though $X$ might irrationally intend to). I shall focus on the normative constraints. Clearly, anything that counts against a normative constraint will count against the corresponding constitutive constraint.

Consider the positive constraints, P1 and P2. Neither is plausible. First, and simply, $X$ might have no beliefs at all about the likely outcome of trying to $A$. Perhaps she has no relevant evidence; or she has some evidence but is unable or unwilling to assess its significance. Yet she might still think, rationally, that it’s worth trying to $A$ anyway; “I don’t know if I can but I might as well give it a go”, she says to herself. Second, as Keith Donnellan (1968: 212, n. 10) and Stephen Schiffer (1972: 69) bring out, if $X$’s life is at stake she might try anything however bleak she thinks the prospects of success. Suppose, for example, that $X$ is shipwrecked and her only hope is to swim to shore. Although $X$ is too realistic to believe that she will reach shore (P1), or even that she can reach shore (P2), she may still form the rational intention to reach shore. For similar reasons, the negative constraint, N1, is not plausible. Even if $X$ believes that she won’t reach shore however hard she tries, if she is in dire enough circumstances it may not be just possible for her to try to do so, it may be rational for her to try: “What have I got to lose?”, she says to herself.

---

10 See also Mele 1992.
But what about the last and weakest constraint, N2? Could X really intend to swim to shore even if she believes that she has no chance of doing so, that her doing so is impossible?\footnote{Neale recently concluded a subtle discussion of suggested constraints by endorsing this one (2016: 278-80).} Initially it may well seem not. Still, on further thought, even this constraint seems dubious. Suppose X has an appropriate amount of epistemic humility and is well aware of her own fallibility. So she realizes that her belief about the impossibility of her reaching shore may be wrong: Then, given what is at stake, she may still intend to swim to shore in the hope that she is wrong. And so she should.

So, I’m inclined to think that intentions are not subject to any of the constraints contemplated by Griceans, whether construed as constitutive or normative. I suspect that we should settle for a weak psychological, not constitutive or normative, constraint on intentions along the following lines: the less a person believes that she will succeed in A-ing if she tries then, other things being equal, the less likely it is that she will intend to A.

Finally, why is it so common among Griceans to insist on some constraint on what a person can intend? The answer seems to be: the Humpty Dumpty worry. If speaker meaning is simply a matter of what the speaker intends to communicate, then we need some constraints on that intention to avoid the unacceptable consequence that the speaker might mean absolutely anything by an utterance. I think that the above psychological, not constitutive or normative, constraint on intentions is sufficient for the purpose. Thus it is unlikely that Humpty Dumpty would intend to communicate to Alice that that is a nice knock-down argument by saying “There’s glory for you” because he should believe that he is very unlikely to convey that message to Alice by that expression, given the expression’s conventional meanings in English.

REFERENCES


