

Testing Theories of Reference
In *Advances in Experimental Philosophy of Language*
Jussi Haukioja, ed. London: Bloomsbury Academic (2015): 31-63

Michael Devitt
The Graduate Center, The City University of New York

‘Semantics, cross-cultural style’, an important piece of ‘experimental philosophy’ by Edouard Machery, Ron Mallon, Shaun Nichols, and Steve Stich (2004), tested theories of reference against the referential intuitions of undergraduates. Whatever we make of the significance of its results – and there has been lots of discussion of this¹ – the paper is important because it raises a very serious methodological question: *How should we test theories of reference?* I have been much occupied with this question in recent years (2011b,c; 2012a,b). In this paper I bring together some earlier conclusions (particularly in sections 3 and 4) and offer some further thoughts (particularly in sections 5 and 6).

I have argued that Machery *et al* are right to criticize the standard methodology in the philosophy of language, a methodology that simply tests theories of reference against philosophers’ referential intuitions. But Machery *et al* are wrong to propose that we should instead test the theories against the folk’s referential intuitions. The primary goal for experimental semantics should not be testing theories against anyone’s referential intuitions but rather testing them against the reality that these intuitions are about: theories should be tested against *linguistic usage*. The challenge then is to figure out how to do that.

1. ‘Reference’

What do we mean by ‘reference’? The word ‘reference’ and its cognates are ordinary English words with a range of common meanings. When used in philosophy, however, they are technical terms. Many philosophers use ‘reference’ in a quite restricted sense that picks out a relation that holds only between a *singular* term of a certain type – for example, proper names and demonstratives - and one semantically significant object. This restricted use is illustrated in *The Reference Book* (Hawthorne and Manley, 2012) and in the large literature that the book discusses.² But experimental semanticists favor a generic use of ‘refer’. We need some word in semantics to cover the many different semantically significant relations that expressions bear to the world. ‘Refer’ seems as good choice as any for this role. So, on this usage, not only proper names and demonstratives but count nouns, mass nouns, verbs, adjectives, and so on, all refer.

I talk here of ‘reference’ picking out ‘semantically significant relations’ between expressions and the world. Which relations are those? They are, or at least should be, ones identified by the explanatory work the term ‘reference’ does in a theory of language.

¹ Including: Ludwig, 2007; Marti, 2009; Deutsch, 2009; Jackman, 2009; Ichikawa *et al*, 2011; Sytsma and Livengood, 2011; Devitt, 2011b,c, 2012b; Genone and Lombrozo, 2012; Ostertag, 2013.

² I have taken a critical look at the theoretical role of ‘reference’ in this literature in a review article on this book (2014b).

This raises the very important question: What are languages and why are they theoretically interesting? Languages are representational systems that are parts of the natural spatio-temporal world and are of theoretical interest because of their causal roles in that world, particularly their roles in communicative behaviors. Thus, Karl von Frisch won a Nobel Prize for discovering that the bees' 'waggle dance' is a language communicating messages about food sources. Another scientist, C. N. Slobodchikoff (2002), discovered that the 'barks' of Gunnison's prairie dogs form a language that communicates messages about predators. And it is a truism that humans have languages that communicate messages that are the contents of thoughts: 'language expresses thought'. So, just as the bees and the prairie dogs have representational systems used for communicating messages to each other, so do we.³ In light of this, the properties of languages that we need to explain – let's call them *meanings* – are those that enable languages to play their causally significant roles in the lives of the organisms that have them, in particular their roles in communication.

A popular idea, and one that I subscribe to, is that reference, along with syntactic properties, are the central notions in an explanation of meanings: they are the core notions in the theory of language. Thus, consider 'Jack thinks that Fred loves himself'. The idea is that its meaning is largely explained by its syntactic structure and the reference of its expressions ('Jack', 'loves', etc). So the semantically significant relations of expressions to the world that we should pick out as 'reference' are the ones that contribute to explaining meanings. It is partly in virtue of standing in those relations that expressions play their important causal roles. I have argued for this naturalistic view of semantics elsewhere (1981, 1996, 2013b).

So we identify referential relations by their causal roles. Then we need to explain those relations: we need theories of reference.

2. Theories of Reference

Three general types of theory seem possible. (1) According to 'description' theories, the reference of *E* is fixed by certain descriptions that competent speakers associate with *E*; *E* refers to whatever those descriptions, or a weighted most of them, uniquely describe. The received view for decades was that the reference of proper names was to be explained by a description theory. But then came the revolution, led by Saul Kripke (1980).⁴ Description theories of names were seen to have serious problems, particularly the problem of 'ignorance and error': speakers who seem perfectly competent with a name are too ignorant to provide the descriptions of its referent demanded by description theories; worse, speakers are often so wrong about the referent that the descriptions they would provide apply not to the referent but to another entity or to nothing at all. There were similar problems for description theories of some other terms. (2) These problems with description theories stimulated interest in theories that took the reference of *E* to be explained not

³ Strangely, this view of human language is rejected by Chomskians; see, e.g. Chomsky, 1986 and 1996; Dwyer and Pietroski, 1996; Laurence, 2003; Collins, 2008a,b; Antony, 2008. They see a human language as an internal state not a system of external symbols that represent the world. I have argued against this view: 2003, 2006a,c, 2008a,b,c, 2009.

⁴ Devitt and Sterelny, 1999, chs. 3-5, is an account of the revolution.

indirectly via associated descriptions but rather by some direct relation between *E* and the world, presumably some sort of causal relation, historical, reliabilist, or teleological. (3) Finally, there is the possibility of theories that explain reference partly in terms of the associated descriptions of (1) and partly in terms of the direct relations of (2).

I have often emphasized a crucial point about theories of reference: *description theories are essentially incomplete* (e.g., 1996, p. 159). A theory of type (1) explains the reference of *E* by appealing to the referential properties of descriptions associated with *E*: *E* refers to whatever those other expressions jointly refer to; thus, perhaps, ‘vixen’ refers to whatever its associated descriptions, ‘female’ and ‘fox’, jointly refer to. How then are the references of those other expressions to be explained? What explains the reference of ‘female’ and ‘fox’? Perhaps we can use description theories to explain those other references too. This process cannot, however, go on forever: there must be some expressions whose referential properties are not parasitic on those of others, else language as a whole is cut loose from the world. Description theories pass the referential buck, but the buck must stop somewhere. It stops with theories of type (2) that explain reference in terms of direct relations to reality.⁵ Those theories offer, we might say, *ultimate* explanations of reference.

If any expressions refer then some expressions must be amenable to ultimate explanations. So theorists of reference should always be on the lookout for likely candidates for ultimate explanations (and theorists should not have needed problems with description theories to stimulate interest in other theories).

We turn now to our main question: How should we test what theory of reference is right for an expression?

3. The Received View⁶

3.1 The Role of Intuitions

The received view is that we should test theories of reference, indeed any semantic theory, by consulting our intuitive judgments about language, our metalinguistic ‘intuitions’. It would be hard to exaggerate both the apparently dominant evidential role of such intuitions in the philosophy of

⁵ Could the buck stop with theories of type (3)? I doubt it but we would need to see the details of such a theory to be confident in rejecting it. A danger with any such theory is that it will lead to a totally unacceptable semantic holism; see Devitt, 1996, pp. 127-32.

⁶ The brief accounts of the received view in this section, and of the Modest Explanation in the next, are based on earlier works: Devitt, 1994, 1996, 2006d, 2012a. There is a related discussion of intuitions in linguistics in Devitt 2006a,b. This discussion was criticized in Collins, 2006, Matthews, 2006, Miščević, 2006, Rattan, 2006, Rey, 2006, and Smith, 2006. Devitt 2006c is a response. There have been some later critics: Pietroski, 2008, responded to in Devitt, 2008a; Textor, 2009, responded to in Devitt, 2010a; Culbertson and Gross, 2009, which led to the exchange, Devitt, 2010b, Gross and Culbertson, 2011; Miščević, 2009, responded to in Devitt, 2014c; Fitzgerald, 2010, responded to in Devitt, 2010b; Ludlow, 2011, and Rey, 2013, responded to in Devitt 2013a.

language and the agreement among philosophers that these intuitions should have this role.⁷ This emphasis on intuitions reflects, of course, a widely held view about the methodology of ‘armchair philosophy’ in general.⁸

⁷ For evidence of this, see Devitt, 2012a, pp. 554-5.

⁸ This widely held view has recently been challenged by Herman Cappelen in a splendidly iconoclastic book (2012). (An earlier article by Max Deutsch (2009) also challenges the view.) Cappelen mounts an impressively detailed argument against what he calls ‘Centrality’, the thesis that ‘contemporary analytic philosophers rely extensively on intuitions as evidence’ (p. 1). Cappelen’s challenge deserves an argument in response. I have offered one elsewhere (forthcoming). Here is a brief summary of my response:

(1) Cappelen first argues against the support that Centrality seems to get from the fact that philosophers often ‘characterize key premises in their arguments as “intuitive”’ (p. 4). He finds this intuition talk very hard to interpret and claims that under none of the interpretations that he proposes does the talk support Centrality. I present two objections. (i) The talk is in fact easy to interpret: for the most part, ‘intuitive’ and ‘intuition’ are not technical terms and so mean here just what they ordinarily mean. In thinking about this, it is important to note that differing theories of intuitions do not entail differing meanings of ‘intuition’. (ii) Cappelen does not produce convincing reasons for not taking these philosophers at their word in their characterization of their premises. So this characterization does support Centrality (in which ‘intuition’ also has its ordinary meaning).

(2) Cappelen argues next that Centrality gets no support from philosophical practice. He proposes three ‘diagnostics’ to detect the presence of intuitions: intuitions ‘seem true or have a special phenomenology; they are default justified, or can justify other propositions without themselves requiring justification; and they are based solely on conceptual competence’ (p. 111). He then examines many well-known cases in the philosophical literature and finds no evidence of philosophers relying on items satisfying these diagnostics and hence concludes that Centrality is false (ch. 8). I present four objections. (I) The diagnostics are based on philosophers’ *theories* of intuitions (ch. 6), theories that, I have argued (see items cited in note 6), are largely false. So the diagnostics based on them are inappropriate. (II) The thesis of Centrality no more needs to be supported by a theory-based diagnostic for intuitions than, say, the thesis that lonely people tend to have dogs needs to be supported by a theory-based diagnostic for dogs. Testing these theses simply requires *abilities to recognize* intuitions and dogs, abilities that almost everyone, including proponents of Centrality, surely have. (III) If we *must* have a diagnostic for intuitions, the place to look for one is in not in philosophical theories but in dictionaries. Based on a quick look at a few dictionaries, I propose a diagnostic along the following lines: ‘an immediate judgment without much conscious reasoning or inference’. Indeed, Cappelen himself notes something like this ‘interpretation’ in his struggle with intuition talk (pp. 33, 62). (IV) Using our ordinary ability to recognize intuitions, aided by this minimal diagnostic if necessary, Centrality gets ample support from the cases that Cappelen examines.

(3) Finally, I think that Cappelen is insufficiently struck by the need to answer the following question: If philosophers are not really relying on intuitions as evidence in the cases he examines, what are they relying on? He rightly insists that these cases are full of arguments; that is the way of philosophy, as he emphasizes. But arguments need premises and it is hard to see what the premises could be in *these* cases - though I certainly do not say in *all* philosophical cases - other than intuitions. Claiming that philosophers rely on propositions that are ‘pre-

Why is this reliance on intuitions appropriate? Machery *et al* (2004) have noted that the intuitions that play this evidential role are usually those of the philosophers themselves and have questioned the appropriateness of this: Why should theories of reference rest on the intuitions of philosophers rather than those of the folk, for example, on those of the undergraduates in Rutgers and Hong Kong that they tested? I think our objection to the standard philosophical practice should be more radical: the problem is not that of relying on philosophers' intuitions rather than the folk's, the problem is that of relying on intuitions at all (2011b,c). The right response to armchair philosophy is not to move in more armchairs for the folk.⁹

I am not alone in being concerned about the role of intuitions. Thus, Jaakko Hintikka remarks: 'One searches the literature in vain for a serious attempt to provide' a justification for the appeal to intuitions (1999, p. 130). In a similar vein, Timothy Williamson remarks: 'there is no agreed or even popular account of how intuition works, no accepted explanation of the hoped-for correlation between our having an intuition that P and its being true that P.' He describes this as 'a methodological scandal' (2007, p. 215).

3.2 'Cartesianism'

So why do philosophers think that the use of intuitions is appropriate? It clearly would be appropriate if we could be confident that the intuitions reflected *knowledge*. And the received view is that a competent speaker of a language does indeed have knowledge about her language, *propositional* knowledge, 'tacitly' at least, *simply in virtue of being competent* in the language:

It is an undeniable feature of the notion of meaning... that meaning is *transparent* in the sense that, if someone attaches a meaning to each of two words, he must know whether these meanings are the same. (Dummett, 1978, p. 131)

The natural view is that one has *some kind of* privileged semantic self-knowledge. (Loar, 1987, p. 97)

The idea of this sort of privileged access - that we are in a special position to know about our own competence - is an instance of general 'Cartesianism':

Since Descartes, it has seemed undeniable to most philosophers that each of us has a privileged way of knowing about his or her own mental states... whenever we have a thought, belief, intention, or desire, we can in principle come to know *what* we think, believe, intend, or desire just by internal examination, without engaging in an empirical investigation of the external world. (McKinsey, 1994, p. 308)

theoretic' (p. 61) or 'common ground' (p. 119) is only the beginning of an answer. Why are these propositions thought to be evidence if not because they are intuitively true?

⁹ I owe this nice remark to Genoveva Martí who thinks she heard or read it somewhere but can't recall the circumstances.

The idea that we have a Cartesian access to semantic facts seems to be an almost unquestioned part of the semantic traditions of Frege and Russell.

3.3 *A Priori Knowledge?*

Why should we suppose that ordinary competent speakers have this knowledge of semantic facts? Many seem to think that the knowledge is a priori. Thus Jerrold Katz claims: ‘We know sense properties and relations of expressions on the basis [of] the speaker’s a priori linguistic intuitions in clear cases’ (1997, p. 21). And Michael McKinsey thinks that it is ‘fairly clear’ that ‘the principle that the meanings of words are knowable a priori...is taken for granted by most philosophers of language and by many linguists’ (1987, p. 1).

Now, of course, the idea that some knowledge is a priori is widespread in philosophy. Nonetheless, Quine has raised serious doubts about it. The main problem with the idea, in my view, is that we do not have even the beginnings of an account of what a priori knowledge *is*. We are simply told what it *isn't*, namely empirical knowledge. Still, suppose we set such general doubts aside and accept that at least our knowledge of mathematics and logic is a priori, what could be the basis for supposing that our knowledge of *meanings* is too? The meaning of a word is presumably constituted by relational properties of some sort: ‘internal’ ones of the sort described by description theories of type (1); or ‘external’ ones of the sort described by causal theories of type (2); or a combination of internal and external relations of the sort described by theories of type (3) (sec. 2). I have argued that we have no reason to suppose that we have some nonempirical way of forming a justified belief about which of these relations constitute the meaning of a word (1994, 1996, 1998, 2011a, 2014a).

3.4 *Embodied Theory?*

If the view that competent speakers have a priori knowledge of semantic facts does not hold up, what else could justify the ubiquitous reliance on intuitions in the philosophy of language? Perhaps philosophers can take a leaf out of the book of linguists.

The common linguistic view of intuitive judgments is expressed in passages like the following:

it seems reasonably clear, both in principle and in many specific cases, how unconscious knowledge issues in conscious knowledge...it follows by computations similar to straight deduction. (Chomsky, 1986, p. 270)

I have described the common view as follows: linguistic competence, all on its own,

provides information about the linguistic facts....So these judgments are not arrived at by the sort of empirical investigation that judgments about the world usually require. Rather, a speaker has a privileged access to facts about the language, facts captured by the intuitions, simply in virtue of being competent... (2006a, p. 96)

On this view, intuitive syntactic judgments are, ‘noise’ aside, ‘the voice of competence’, and so provide good evidence about the language. Let’s call this thesis ‘VoC.’

The evidence that VoC is the orthodox Chomskian view of linguistic intuitions strikes me as overwhelming (2006a,b, 2010b, 2013a).¹⁰ Indeed, if Chomskians did not hold VoC, they would have no view of the source of linguistic intuitions.

Stich has suggested that philosophers of language may be implicitly embracing VoC as a justification for the authoritative role given to referential intuitions (1996, p. 40; see also Hintikka, 1999 and Williamson, 2007). Philosophers may think that speakers derive their referential intuitions from embodied referential principles. So, just as the true grammar is already embodied in the mind of every speaker, so too, according to this suggestion, are true semantic theories of reference. Referential intuitions, like syntactic ones, are the result of something like a deduction from a represented theory.¹¹

Although VoC is much more promising than the view that we have a priori knowledge of meaning, I have argued that it is wrong (2006a,b, 2010b, 2013a). The main problems with it are, first, that, to my knowledge, it has never been stated in the sort of detail that could make it a real theory of the source of intuitions. Just *how* do the allegedly embodied principles yield the intuitions? We need more than a hand wave in answer. Second, again to my knowledge, no argument has ever been given for VoC until Georges Rey’s recent attempt (2013) which, I argue (2013a), fails. Third, given what else we know about the mind, it is unlikely that VoC could be developed into a theory that we would have good reason to believe.

I have pointed out some other implausibilities of VoC. These are briefly as follows. (i) If competence really spoke to us, why would it not use the language of the embodied theory and why would it say so little? (ii) There would be a disanalogy between the intuitions provided by the language faculty and by perceptual modules. (iii). Developmental evidence suggests that the ability to speak a language and the ability to have intuitions about the language are quite distinct, the former being acquired in early childhood, the latter, in middle childhood as part of a *general* cognitive development.

Perhaps the best reason for rejecting VoC, is that there is a better explanation of intuitions and their evidential role.

4. The Modest Explanation of Intuitions

¹⁰ So I was surprised to find three knowledgeable philosophers rejecting the attribution: John Collins (2008a, pp. 17-19), Gareth Fitzgerald (2010), and Peter Ludlow (2011, pp. 69-71). I have responded to Fitzgerald (Devitt, 2010b, pp. 845-7) and to Ludlow (Devitt, 2013, pp. 274-8). Ludlow’s discussion is notable for its egregious misrepresentation of the evidence. I have also provided more evidence (2013, p. 273) in the works of Barry Smith (2006), Mark Textor (2009), and Georges Rey (2013). I still think that the evidence is overwhelming. But see Jeffrey Maynes and Steven Gross (2013) for a nice discussion of the matter.

¹¹ Interestingly, Chomsky, who holds VoC for grammatical intuitions seems to reject it for semantic ones (1995, p. 24). For discussion, see Devitt, 2012a, pp. 558-9.

If VoC is not the right theory of intuitions, what is? I argue that intuitive judgments about language, *like intuitive judgments in general*, ‘are empirical theory-laden central-processor responses to phenomena, differing from many other such responses only in being fairly immediate and unreflective, based on little if any conscious reasoning’ (2006a, p. 103).¹² Although a speaker’s competence in a language obviously gives her ready access to the *data* of that language, the data that the intuitions *are about*, it does not give her ready access to the *truth* about the data; the competence does not provide the *informational content* of the intuition. In this respect my view is sharply different from VoC. And it is sharply different in another respect: it is *modest*, making do with cognitive states and processes we were already committed to. So, following Mark Textor (2009), I now call it ‘the Modest Explanation’.

According to the Modest Explanation, intuitions about language, like other intuitions, are ‘theory-laden’. This could do with some explanation. First, the view is *not* that these intuitions are theoretical judgments or the result of theorizing. Rather, the intuitions are mostly *the product of experiences* of the linguistic world. They are like ‘observation’ judgments; indeed, some of them *are* observation judgments (2006a: 103) As such, they are ‘theory-laden’ in just the way that we commonly think observation judgments are. The anti-positivist revolution in the philosophy of science, led by Thomas Kuhn and Paul Feyerabend, drew our attention to the way in which even the most straightforward judgments arising from observational experiences may depend on background expertise. We would not make the judgments if we did not hold certain beliefs or theories, some involving the concepts deployed in the judgments. We would not make the judgments if we did not have certain predispositions, some innate but many acquired in training, to respond selectively to experiences.¹³ There is need for some cautionary words about this theory ladenness.

(a) The power of background expertise to influence judgments should not be exaggerated. Thus a person observing the Müller-Lyer arrows will judge that one ‘looks longer’ than the other even though she knows perfectly well that they are the same length. (b) The view is not that we consciously bring the background into play in a way that amounts to *theorizing* about the experience. Surely, we mostly don’t. Nonetheless, the background plays a causal role in the judgment. (c) The view is not that we need to have done a deal of thinking about language before having linguistic intuitions: a thoroughly ignorant person may *learn* to have intuitions in an experimental situation (2006a, p. 114).¹⁴ (d) Finally, the theory ladenness we are discussing is *epistemic*. It should not be confused with *semantic* theory-ladenness, the view that the meaning of

¹² This theory of intuitions could be seen as starting from the minimal dictionary-based diagnostic I proposed in response to Cappelen; see note 8 above. Should Cappelen, indeed anyone, object that the theory is not true, my positive proposals for testing theories of reference would lose nothing, so far as I can see, from replacing ‘intuition’ by ‘intuition*’ in the following discussion and turning my theory of intuitions into a definition of ‘intuition*’.

¹³ So ‘theory’ in ‘theory-laden’ has to be construed *very* broadly to cover not just theories proper but also these dispositions that are part of background expertise.

¹⁴ I claim that this is the way to view intuitions of the ignorant in the ingenious ‘minimal pair’ experiments (2006a, p. 110).

an observation term is determined by the theory containing it. This ‘semantic holism’, also part of the revolution, has little to be said for it in my view (1996, pp. 87-135).

It is not a methodological consequence of the Modest Explanation of intuitions that they should have no evidential role in theorizing about the nature of some area of reality. However, it is a consequence that they should have that role only to the extent that they are likely to be reliable about that area of reality, only to the extent that they are *reliable indicators*. And this reliability needs to be assessed using independent evidence about the reality. But we need that independent evidence anyway:

Although we may often be right to trust an intuition in the short run, it is crucial to see that nothing rests on it in the long run. We can look for more direct evidence in scientific tests. In such a scientific test we examine the reality the intuition is *about*. These scientific examinations of reality, not intuitions about reality, are the primary source of evidence. The examinations may lead us to revise some of our initial intuitions. They will surely show us that the intuitions are far from a complete account of the relevant bit of reality. (2011b, p. 425)

The intuitions in question here are ones identifying objects as having properties *of the sort adverted to in the very theory being tested*; for example, intuitions about fish when testing a biological theory of fish, about money when testing an economic theory of money, and about reference when testing a semantic theory of reference. Such intuitions from people who are reliable about the reality in question, are of course good evidence about the nature of that reality. But I make two points about the intuitions. First, they are only indirect evidence. Second, their reliability needs to be established. Both these points show the need for more direct evidence: *the primary evidence for a theory about a certain reality comes not from afore-mentioned intuitions about the reality but from more direct examinations of that reality*. We don’t rest biology on intuitions about fish and the like, or economics on intuitions about money and the like. No more should we rest semantic theories on intuitions about reference and the like. We should examine linguistic reality more directly. That reality is to be found in linguistic usage.

I shall consider the task of gathering this evidence in section 6.

But first we need to clarify this contrast between indirect and direct evidence by distinguishing different sorts of intuitions. And we need to say something about the *likelihood* that intuitions, particularly referential intuitions, are reliable. All this requires distinguishing among intuitions according to the degree to which they are theory laden and according to the expertise of those who have them.

5. Varieties of Intuitions

5.1 Perceptual Judgments as Intuitions

After introducing the Modest Explanation of intuitions in *Ignorance of Language* I immediately make a clarification:

It may be that there are many unreflective empirical responses that we would not ordinarily call intuitions: one thinks immediately of perceptual judgments like ‘That grass is brown’ made on observing some scorched grass, or ‘That person is angry’ made on observing someone exhibiting many signs of rage. Perhaps we count something as an intuitive judgment only if it is *not really obvious*. I shall not be concerned with this. My claim is that intuitions are empirical unreflective judgments, *at least*. Should more be required to be an intuition, so be it. (2006a, p. 103)

Some perceptual judgments that are ‘not really obvious’ are among my later examples of intuitions. The following one is in the text:

Consider, for example, a paleontologist in the field searching for fossils. She sees a bit of white stone sticking through grey rock, and thinks ‘a pig’s jawbone’. This intuitive judgment is quick and unreflective. She may be quite sure but unable to explain just how she knows. (p. 104)

The next ones are in a note:

other nice examples: of art experts correctly judging an allegedly sixth-century Greek marble statue to be a fake; of the tennis coach, Vic Braden, correctly judging a serve to be a fault before the ball hits the ground. (p. 104n)

However, I also treat some ‘really obvious’ perceptual judgments as intuitions in a discussion of the visual module (pp. 112-3). And I did also in an earlier work (1996).

So I have a generous view of what counts as an ‘intuition’, including even such ‘really obvious’ judgments as that the grass is brown and the person angry. Given my view that all of these perceptual judgments, whether really obvious or not, are immediate, empirical, and theory-laden, it is an uninteresting verbal issue whether I am ‘right’ to be so generous. These perceptual judgments are certainly all *intuitive*. There seems to be no theoretically interesting reason for grouping some but not all together under the term ‘intuition’. This having been said, the difference in the obviousness of these intuitions is interesting. The less obvious an intuition of a certain sort, the more expertise, the more ‘theory’, is required to have reliable ones of that sort. So whereas judging brown grass requires little expertise, judging a pig’s jawbone requires a lot.

I have used (2006a, pp. 104-5) the following quote from the cognitive psychologist, Edward Wisniewski, to demonstrate the importance of expertise:

researchers who study behavior and thought within an experimental framework develop *better* intuitions about these phenomena than those of the intuition researchers or lay people who do not study these phenomena within such a framework. The intuitions are better in the sense that they are more likely to be correct when subjected to experimental testing. (1998, p. 45).

The role of expertise has obvious methodological consequences. Although the intuitions of the inexpert may often be reliable enough about some matter, we should in general prefer the

intuitions of the expert: ‘the more expert a person is in an area, the better the person’s theory, the wider her range of reliable intuitions in the area’ (2010b, p. 860).

On the generous view, all immediate perceptual judgments count as intuitions. So there obviously cannot be any blanket dismissal of intuitions as evidence. For, at bottom, all theories rest on immediate judgments about what is perceived and hence on intuitions, generously conceived. The most direct contact we can have with reality is via the experiences of it that yield perceptual judgments. These judgments at the periphery of our ‘web of belief’ provide the empirical justification for our theories (2011c, p. 30). And the evidential question that concerns us should be not *whether* to use intuitions as evidence but *what* intuitions to use as evidence.

In light of this, return to our contrast between resting on certain intuitions and seeking more direct evidence. The intuitions in question were, as noted, ones that identify objects as having properties *of the sort adverted to in the very theory being tested*. We note now that, on the generous view, the more direct evidence is also to be found in intuitions, albeit different ones that are more basic and less theory-laden; for example, perceptual judgments about animal behavior when doing biology, about human behavior when doing economics, about colors and smells when doing chemistry, about instrument readings when doing physics. I shall bring out this contrast by considering the evidence for a theory of reference.

Suppose that Jill witnesses an utterance by Jack in condition C. Here is a series of immediate perceptual judgments that Jill makes in response to this event, judgments roughly ordered for theory ladenness:

Jack emitted a noise
Jack uttered something
Jack said something in English
Jack said, ‘Einstein was a physicist’
Jack said that Einstein was a physicist
Jack referred to Einstein

All of Jill’s judgments are ‘intuitive’ and, on the generous view, all count as ‘intuitions’. But only the last is a *referential* intuition. The other judgments, perhaps obtained in scientific tests, are examples of ‘more direct evidence’ of linguistic usage on which a theory of reference should ultimately be based. Thus, suppose that a theory of reference for names predicted that, in condition C, Jack would not say ‘Einstein was a physicist’.¹⁵ Then, Jill’s judgment that Jack did say this is likely clear evidence against the theory. And it is *primary* evidence in a way that Jill’s judgment about what, if anything, Jack referred to is not, however reliable Jill is in such referential judgments.

It is an empirical question just how reliable a person’s intuitive judgments are and hence how good they are as evidence. However, supposing that Jill is an ordinarily educated English speaker, we can surely count on her being very reliable in her perceptual judgments of most of the sorts illustrated above; they are at the really obvious end of the scale. Thus, when Jill judges that someone utters a certain English sentence like ‘Einstein was a physicist’, it is very likely that she is

¹⁵ It is sadly difficult to come up with such predictions; see section 6.

right. What about her perceptual judgments *about reference*, like that Jack referred to Einstein? This is the most theory-laden of Jill's judgments. How likely is a judgment of that sort to be true? Provided that Jill's judgments are about reference in ordinary humdrum situations, I predict that they would be quite reliable. In order for the folk to make reliable judgments about reference, they will doubtless have to have reflected on language a bit, aided by some minimal education. But that is what we can expect from normal members of our society. So I predict that if we tested folk referential intuitions about humdrum situations against evidence from usage, we would find the intuitions fairly reliable. Nonetheless, we should prefer the judgments of the experts here for the same reason we should prefer them anywhere: we can expect them to have a wider range of reliable intuitions. So we should prefer the intuitions of philosophers.

So far as I know, perceptual judgments about reference by the folk have not played a significant evidential role with theories of reference. But judgments of this sort by philosophers surely have. I have given this example: my intuitions about names, formed when I first heard Kripke in 1967, 'have been confirmed, day in and day out for forty years, by observations of people using a name to refer successfully to an object that they are ignorant or wrong about' (2011c, pp. 21-2). This sort of approach to referential intuitions is surely common in philosophy. Philosophers who favor truth-referential theories of meaning – and there are many of them – surely find support for their theories in their observations of paradigm instances of reference. Indeed, if they did not make such observations, they would surely not suppose that reference could play a key explanatory role in a theory of meaning.

5.2. Memory Judgments as Intuitions

Suppose that Mary is another witness to Jack's utterance in condition C. She arrives at just the same judgments from this experience as does Jill but she does so the next day, based on her *memory* of John's utterance. All of Mary's judgments, like Jill's, are 'intuitive' and, continuing my generous policy, I count them all as 'intuitions'. Once again, the epistemic status of each judgment depends on the details of Mary's reliability. And there is no basis for a blanket dismissal of them.

Arguably, many of the intuitions of philosophers, led by Kripke, about the reference of names like 'Cicero', 'Catiline', 'Feynman', 'Einstein', and 'Columbus' out of the mouths of the ignorant, intuitions central to the most powerful argument against description theories of names, are memory judgments of this sort (as Genoveva Marti (2014) also points out).¹⁶ Kripke and others likely make many of these judgments in response to *remembered observations of the actual uses of these names, or analogous ones, by the ignorant*.

So, I am suggesting, it is likely that many of the referential intuitions that play a role in testing theories of reference are about *actual* cases, perceived or remembered. It seems to be easy for philosophers to overlook this likelihood, slipping into the view that all the evidence for these theories comes from *thought experiments* about *hypothetical* cases. Machery *et al* are an

¹⁶ In support of Centrality, the thesis rejected by Cappellan (note 8 above), it is *obvious* that Kripke uses referential intuitions as evidence in this argument; for example, in judging that 'the man in the street...uses 'the name "Feynman" as a name for Feynman' (1980: 81). So does just about every theorist of reference.

example, as I noted (2011b): they talk of the theories of reference being ‘assessed by consulting one’s intuitions about the reference of terms in *hypothetical* situations’ (2004, B1; emphasis added).¹⁷

5.3. *Intuitions in Thought Experiments*

Now consider the intuitive judgments expressed by Frank in the following situation. Frank did not witness Jack saying anything but is presented with a *description* of an utterance by Jack in condition C and asked what he would say about it. So Frank is engaged in a thought experiment about a hypothetical case. Frank immediately forms just the same judgments as Jill and Mary, including the judgment that Jack referred to Einstein. Referential judgments of this sort are what philosophers seem mostly to have in mind as the intuitions that provide evidence for or against a theory of reference. I agree these such intuitions do, of course. So, although, I have just claimed, a philosopher’s referential intuition about a humdrum use of a name by the ignorant will sometimes be a perceptual or memory judgment, it will sometimes not be: it will be based not on the experience of an actual use of the name but on a thought experiment. Think, for example, of the judgment I made about a nice case invented by Donnellan:

A child is gotten up from sleep at a party and introduced to a person called ‘Tom’. ‘Later the child says to his parents, “Tom is a nice man”...nothing the child possesses in the way of descriptions, dispositions to recognize, serves to pick out in the standard way anybody uniquely’ ([Donnellan]1972: 364). Yet the child is talking about that very person he was introduced to. (2011b, p. 421 n. 3)

My judgment here is made in a thought experiment and is not based on perception. Still, judgments arising from thought experiments like this about *humdrum* hypothetical situations should be distinguished from those arising from thought experiments about *fanciful* hypothetical situations, cases like Kripke’s one of ‘Gödel’. What makes the humdrum situations humdrum is they are of just the sort that we all, folk and philosophers alike, are perceiving day in and day out, situations of ignorant people using familiar names. So, although the judgments we make in thought experiments like Donnellan’s are not perceptual, they are closely related to perceptual ones, for they are about situations *just like many that we remember in our experience*. So, if we were to seek judgments of this type from the folk, it seems to me *quite likely* that they would be about as reliable as the perceptual judgments that they are related to. They don’t require much expertise. Still, once again, we should prefer the judgments of philosophers.

When we move to intuitions about hypothetical fanciful cases like ‘Gödel’, I have argued that we should forget about the folk (as indeed we had until Machery *et al*): these intuitions require too much expertise (2011b, pp. 420-3). As Stich aptly remarked, with cases like Twin Earth in mind (and long before he was seized by experimental philosophy), ‘nonphilosophers often find such cases so outlandish that they have no clear intuitions about them’ (1983, p. 62n.). For cases like this, the only intuitions worth worrying about are those of philosophers and similar experts.¹⁸

¹⁷ They have since modified their position to include intuitions about actual cases (2013).

¹⁸ I have pointed out that Gödel experiments have put us well on the way to showing that philosophers are indeed more reliable than the folk in their referential intuitions about fanciful

So, we have made some distinctions among the immediate and unreflective empirical judgments that I generously count as intuitions. There are the perceptual ones, memory ones, and ones formed in thought experiments. Among the latter there are ones about humdrum hypothetical cases and ones about fanciful hypothetical cases. And among them all there are differences in the degree to which their reliability depends on expertise. I conjecture that referential intuitions about humdrum cases, whether perceptual or not, are likely to be fairly reliable without much expertise about language. In contrast, referential intuitions about fanciful hypothetical cases probably require a good deal of expertise.

5.4 Consequences for Machery *et al*

The Modest Explanation of intuitions that I have presented has consequences for the methodology of Machery *et al*.

First, they tested the wrong referential intuitions. If we are going to test referential intuitions, we should prefer to test ones about humdrum cases like Kripke's 'Einstein' and 'Columbus', or even Donnellan's sleeping child, rather than ones about fanciful cases like Kripke's 'Gödel'. For, we should expect intuitions about humdrum cases to be more reliable.

Second, they tested the wrong people. If we are going to test referential intuitions we should prefer to test those of philosophers, particularly if the intuitions are about fanciful cases like 'Gödel', because philosophers are more expert (2011b, pp. 425-6).¹⁹ This line of thought yielded an example of what has become known as 'the Expertise Defense' against the findings of Machery *et al*. The Expertise Defense has led to a lively exchange of opinion: Weinberg *et al*, 2010; Machery and Stich, 2012; Machery *et al*, 2013; Machery, 2011; Devitt, 2011c; Machery, 2012; Devitt, 2012b.²⁰

One objection that Machery *et al* have to preferring the referential intuitions of philosophers to those of the folk is that the philosophers' intuitions may be theoretically biased (2013). Indeed, they may be, but that is the sort of epistemic risk that we *always* run in science, since all judgments are theory-laden. And there are two points to make about it. First, we can try

cases (2011c, p. 24).

¹⁹ James Genone and Tania Lombrozo may have misunderstood the similar criticism I made (2010d) in commenting on a draft of their paper (Genone and Lombrozo, 2012). For, they respond by doubting that 'expert intuitions are superior to folk intuitions when it comes to ordinary referential practices, or that there could be specialists in the practice of using names and concepts in general' (2012, p. 734). But my criticism does not rest on the view that philosophers are more expert than the folk at 'using names and concepts in general' but that they are more expert at making intuitive judgments about the reference of those names and concepts. Genone and Lombrozo's experiment elicits the latter referential intuitions not 'ordinary referential practices'. (I am all for eliciting the practices; see section 6.)

²⁰ See also the following exchange arising out of my analogous claim (2006a, pp. 108-9) that we should prefer the grammatical intuitions of linguists over those of the folk: Culbertson and Gross, 2009; Devitt, 2010b; Gross and Culbertson, 2011.

to control for bias, just as we do elsewhere in science. Second, the risk should not be exaggerated. The intuitive judgments that scientists make about their domains tend to be in agreement. For evidence of this among linguists, see Sprouse and Almeida (2013). For evidence among reference theorists, one has to look no further than the response to Kripke's intuitions about names. These intuitions were devastating for the sorts of description theories of names that were the received theories at that time. Yet philosophers who wanted to save description theories did not reject the intuitions, whether about humdrum cases or fanciful Gödel cases, but rather tried to construct novel description theories that were compatible with those intuitions (see Devitt and Sterelny, 1999, sec. 3.5 for discussion). And it is not surprising that, despite theory ladenness, experts tend to share intuitions because those intuitions are not determined simply by theoretical background: they are determined largely, we hope, by experiences of the reality of that domain.

Third, my main disagreement with Machery *et al* is with their focusing experimental semantics on testing anyone's intuitions about reference. I have argued that the focus should be on testing linguistic usage. So I disagree with the following:

that philosophers of language should emulate linguists, who are increasingly replacing the traditional informal reliance on their own and their colleagues' intuitions with systematic experimental surveys of ordinary speakers' intuitions. (Machery and Stich, 2012, p. 495)

The syntactic intuitions elicited by linguists are at best indirect evidence of the nature of the syntactic reality that they are about. Linguists need more direct evidence and that is to be found by examining linguistic usage. The story for philosophers is much the same. Referential intuitions are at best indirect evidence of the nature of referential reality. Philosophers need the more direct evidence that can be found by examining linguistic usage. The focus of experimental semantics should be on that. And that is what I now turn to.

6. Testing Usage

How are we to test theories of reference against usage? I think that we should get inspiration from linguistics. Even though the received methodology in linguistics, like in philosophy, is dominated by attention to the role of intuitions – far too much so, in my view (2006a, pp. 98-100) – linguistics is importantly different from philosophy in that linguistics often acknowledges the role of usage as a source of evidence. Thus evidence for grammars is found in the corpus and in elicited production.²¹ I shall consider these in turn.

6.1 The Corpus

I have elsewhere emphasized that a major source in linguistics of evidence about syntax comes from the corpus, the vast mass of linguistic sounds and inscriptions that competent speakers produce as they go about their business without prompting from linguists. Linguists observe these

²¹ Also in reaction time studies, eye tracking, and electromagnetic brain potentials. Perhaps philosophers can get inspiration from these experiments too, but I have no idea how.

performances and seek answers to questions like: ‘Do people ever say x ?’; ‘How do they respond to y ?’; ‘In what circumstances do they say z ?’ (2006a, p. 98). Philosophers can also look to the corpus for evidence about reference. I have illustrated what a rich source this could be, in principle, with a bit of the corpus provided, ironically enough, by Machery *et al.* In going about their business, which is testing folk referential intuitions, Machery *et al.* use the name ‘Gödel’ many times. These uses are in vignettes presented to their experimental subjects. Machery *et al.* are surely as competent as anyone with ‘Gödel’ and yet, I point out, their use of the name in the following passage is inconsistent with what (standard)²² description theories would predict: ‘Now suppose that Gödel was not the author of this theorem. A man called ‘Schmidt’, whose body was found in Vienna under mysterious circumstances many years ago, actually did the work in question’ (2004, p. B6). If the description theory were true this utterance would be anomalous. (So Machery *et al.*’s experiment is biased against description theories.)

For, if [Machery *et al.*]’s use of ‘Gödel’ refers to that eminent logician in virtue of their associating with it the description, ‘the prover of the incompleteness of arithmetic’, this passage is not something that [they] would be disposed to say. They would not, in one and the same breath, both refer to Gödel and suppose away the basis of that reference. (2011c, p. 28)

I shall now give a similar illustration using a vignette from another, more recent, experiment. This experiment, by James Genone and Tania Lombrozo (2012), was designed to test descriptive and causal theories of reference for natural and nominal kind terms.

A stimulus that they provided to some subjects included the following:

There is a small island in the Indian Ocean called ‘Alpha’. Natives of Alpha, called ‘Alphians’, sometimes catch diseases not found anywhere else in the human population. When this happens, they consult Alphian doctors. One of the diseases on Alpha is called ‘tyleritis’.

Facts about the Alphian disease called ‘tyleritis’:

- Tyleritis affects the muscles and causes muscle pain.
- Tyleritis is only caused by exposure to a rare mineral.
- Tyleritis can be diagnosed with a blood test.
- Tyleritis can be cured by an injection.

Alex is a native Alphian. Alex first heard of tyleritis when his uncle contracted it and he overheard other family members discussing it. Alex knows that tyleritis* is a disease, and that it can cause pain. Alex also has a number of other beliefs about tyleritis*.

Alex’s beliefs about the Alphian disease tyleritis*:

- Tyleritis* affects only the joints and causes joint pain.
- Tyleritis* is caused by a virus.
- Tyleritis* can only be diagnosed with a tissue biopsy.

²² This qualification is necessary to exclude theories that are not along the lines of those offered by Frege, Russell, and Searle. So we are excluding what I call ‘circular descriptivism’ and ‘causal descriptivism’ (Devitt and Sterelny 1999: 60-1). The qualification should be taken as read in what follows.

- Tyleritis* is incurable. (asterisks added)

All of Genone and Lombrozo's uses of the invented term 'tyleritis' that I have marked with asterisks are inconsistent with what description theories of 'tyleritis' would predict.²³ (So this experiment is also biased against description theories.)²⁴

(A) Suppose that a description theory was true of 'tyleritis'. Then its reference would be determined by descriptions drawn from the ones used above to state the 'facts' about tyleritis. Furthermore, competent users of the term must associate those reference-determining descriptions with it. It follows, then, that Alex is not competent with the term. For, rather than associating those descriptions, he associates descriptions nearly all of which are false of tyleritis. So his uses of 'tyleritis' do not refer to tyleritis but to something else or nothing at all.

(B) Next, consider Genone and Lombrozo's uses of 'tyleritis' that I have marked. What do those uses refer to? Given the story, they must refer to the Alphian disease described in the list of 'facts'; i.e. the uses refer to tyleritis. For, when the term is properly used, that's what it refers to and we may assume that Genone and Lombrozo are using it properly: after all, they invented it!

(C) Finally, Genone and Lombrozo have some interesting things to say with the marked uses of 'tyleritis'. These remarks concern Alex's beliefs. I start with a presumption: Genone and Lombrozo are inviting us to suppose that they base their views of those beliefs largely, if not entirely, on *what Alex says* using 'tyleritis'. With this in mind, consider their remarks about Alex's beliefs. (i) One of these is that 'Alex...has a number of...beliefs about tyleritis'. We should see this as reflecting Genone and Lombrozo's referential intuition about Alex's use of 'tyleritis'. Since they are expert enough, we expect their intuition to be right (sec. 5). Yet if 'tyleritis' was covered by a description theory, this intuition would be wrong: Alex's use of 'tyleritis', hence the beliefs he expresses with the term, would not be about tyleritis – see (A). So Genone and Lombrozo's remark is contrary to what the description theory predicts. (ii) Still, I am looking to the corpus for a test *against usage* not against referential intuitions. For this test we must consider Genone and Lombrozo's other remarks. These ascribe beliefs to Alex; for example, they ascribe the belief that Tyleritis is caused by a virus. But, if a description theory was true of 'tyleritis', these are not beliefs that Genone and Lombrozo would be disposed to ascribe. For, their uses of 'tyleritis', hence the beliefs they ascribe to Alex, are about tyleritis –

²³ Interestingly, in another passage, Genone and Lombrozo's usage is inconsistent with what a *causal* theory would predict: 'Bob's beliefs were always identical to Alex's' (2012: 725). Bob is on another island, Brom, and has a number of beliefs about a disease found only on Brom and also called 'tyleritis'. But this disease 'affects only the joints and causes joint pain' and so is a different disease from the one called 'tyleritis' on Alpha. So, even though Bob and Alex express their beliefs in identical words, their beliefs are *not* identical according to a causal theory, because they have different origins. (Similarly, if Putnam is right, when Oscar and Twin-Oscar say 'Water is refreshing', they do not express the same beliefs because one is about H₂O, the other about XYZ.)

²⁴ I pointed this out before (2010d); for a response, see Genone and Lombrozo 2012, pp. 740 n. 24.

see (B) – and *yet Alex's* beliefs would not be about tyleritis – see (A). And, Genone and Lombrozo, as competent users of 'tyleritis', are disposed to ascribe beliefs about tyleritis on the basis of uses of 'tyleritis' only to people whose uses refer to tyleritis. That is part of what their competence with 'tyleritis' amounts to.

The sort of thinking that lies behind the final step could do with elaboration. I shall provide this in the discussion of elicited production to follow. I shall also point to a worrying flaw in the thinking: 'the implicit-scare-quote problem'.

These are two examples of how we can gather evidence about reference from usage in the corpus. But there are well-known difficulties in using the corpus as evidence. First, in this case, one has to note something in the linguistic phenomena that *is* evidence for/against some theory of reference. Then one has to have a record of it, which is problematic if it is spoken rather than written. And one may need to document quite a lot of information about the speaker and circumstances. Still, the examples illustrate what a mass of evidence the corpus provides that *could* be mined scientifically. And it indicates the important role that the corpus plays as informal evidence about reference.²⁵

6.2 Elicited Production

Fortunately, we don't have to rely on the corpus for direct evidence in usage: we can *induce* usage from competent speakers in experimental situations. Consider this description of 'the technique of *elicited production*' in linguistics:

This technique involves children in a game, typically one in which children pose questions to a puppet. The game orchestrates experimental situations that are designed to be uniquely felicitous for production of the target structure. In this way, children are called on to produce structures that might otherwise not appear in their spontaneous speech. (Thornton 1995, 140)

Clearly much direct evidence could be gathered in this way. However, contriving appropriate situations in an experiment is likely to be a laborious business.

I proposed an easier technique of elicited production for linguistics. Instead of *constructing* situations to see what people say and understand in those situations, 'we can *describe* situations and ask people what they would say or understand in those situations' (2006a, p. 99). Note that this method is not the common one of prompting *metalinguistic intuitions* about described situations, yielding indirect evidence about language. Rather it is prompting *linguistic usage* in described situations, yielding direct evidence about language.²⁶ I have suggested recently that this method

²⁵ I made a similar point about the evidential role of the corpus in linguistics as part of a response to the tendency in linguistics to exaggerate the role of speakers' intuitive judgments (2006a, pp. 98-9).

²⁶ I sum up my discussion of linguistic evidence: 'the main evidence for grammars is not found in the intuitions of ordinary speakers but rather in a combination of the corpus, the evidence of what we would say and understand, and the intuitions of linguists' (2006a, 100).

could provide a rich source of evidence about reference. I think that this is the way forward in experimental semantics (2011c, pp. 29-30).

The challenge then was, of course, to design elicited production experiments that would do the job. My earlier discussion of naturalized semantics in *Coming to Our Senses* (1996: ch. 2; also, 1994) prompted an idea (2011b: 430-2). Wesley Buckwalter, Kate Devitt, and I have conducted experiments along the lines of this idea. We started by testing theories of reference for proper names, using the name ‘Beyoncé’ as an example. Here is a summary of the idea for the experiment:

1. Find participants who are ‘experts’ on pop stars in general and Beyoncé in particular.
2. Present two vignettes about a character, Dr. Marcus, in which he uses the name ‘Beyoncé’. In one he associates with that name descriptions that do identify Beyoncé, in the other, he does not, including associating *mis*descriptions. Ask the participants to explain a certain behavior of the character, Dr. Marcus.
3. Assumption. The participants, being ‘experts’, are competent users of the name, ‘Beyoncé’. So if a participant uses ‘Beyoncé’ she will refer to the pop star. And if she uses it (or a pronoun anaphoric on it) in ascribing a thought that Beyoncé is... to Marcus in order to explain his behavior she has understood Marcus as having referred to Beyoncé with the name. So, probably, Marcus did refer to Beyoncé.
4. Descriptivist prediction: the differences in the associated descriptions in the two vignettes will make a difference in the participants’ readiness to use ‘Beyoncé’ to ascribe thoughts to Dr Marcus: they will be much less likely to ascribe such thoughts in the vignette where the descriptions do not identify Beyoncé. Anti-descriptivist prediction: the differences in associated descriptions will make no difference in usage.

More needs to be said about why the description theory predicts that participants would not use ‘Beyoncé’ to ascribe beliefs to Dr. Marcus. First, note that the prediction does not rest on any assumptions about the evidential status of referential intuitions, intuitions deploying the theoretical concept of *reference* (2011c, p. 30). So what does it rest on? The thinking is that it rests simply on *the assumption that participants are competent with the name ‘Beyoncé’*. What does this competence amount to? (a) On a fairly theory-neutral view of thoughts, this competence is the ability to use the inscription ‘Beyoncé’²⁷ to express a part of a thought that refers to the famous singer; and the ability to assign to an inscription ‘Beyoncé’ that refers to the singer in the context of utterance a part of a thought that refers to the singer. (b) Adding more theory by assuming the Representational Theory of the Mind (RTM), we can simplify: the competence is the ability to translate back and forth between mental representations of the singer Beyoncé and ‘Beyoncé’. (c) Adding even more theory by assuming the Language of Thought Hypothesis (LOTH): the competence is the ability to translate back and forth between the mental word <Beyoncé> referring to the singer and ‘Beyoncé’. As a result of this competence, assuming LOTH for convenience, a participant tends, in understanding a ‘Beyoncé’ utterance, to form a <Beyoncé> thought iff the utterance refers to the singer. And, as a result of this competence, her

²⁷ Similarly, of course, the sound /Beyoncé/, etc. Take this addition as read in what follows. I focus on inscriptions because they are what we are dealing with in the experiment.

production of a ‘Beyoncé’ utterance will refer to the singer iff it is produced by a <Beyoncé> thought. *This is simply how her competence with ‘Beyoncé’ is manifested.* So, on the basis of this assumption of her competence, we theorists can then conclude that she will tend to use the name ‘Beyoncé’ in her description of the mental states that explain Dr. Marcus’ behavior iff Dr. Marcus’ use of the name refers to the singer. Our conclusion does not rest on any assumption that the participant thinks any thoughts about the reference of ‘Beyoncé’; she simply thinks whatever thoughts are expressed in the ‘Beyoncé’ utterances she is understanding and producing.

6.3 The Implicit-Scare-Quote-Problem

I had a worry about this line of thought from the beginning. How would an ‘expert’ participant *indicate* that she thinks that a character who associates ‘incorrect’ descriptions with an expression ‘E’ does *not*, on those grounds, believe/hope/wonder whether/etc ...E...? That is, suppose the description theory for ‘E’ was *right*, what would be the predicted response? The prediction is that, because the character does not associate the descriptions that determine the meaning and reference of ‘E’ the participants should treat the character as not using ‘E’ in the conventional way and not referring to E. But the difficult question is: *How would participants indicate this?* Here are some of the ideas we came up with for Dr Marcus-‘Beyoncé’:

- (a) ‘Dr. Marcus doesn’t know who he is talking about with “Beyoncé”’
- (b) ‘Dr. Marcus is not talking about Beyoncé’
- (c) ‘Dr. Marcus is thinking about someone else’
- (d) ‘Dr. Marcus thinks that the invitation is for someone he wrongly calls “Beyoncé”’
- (e) Any explanation where the name ‘Beyoncé’ is in scare quotes.

One concern about these possible responses is that at least two of them, (b) and (c), are expressions of referential intuitions and yet the idea was to test theories against usage not against referential intuitions. (e) raises a more worrying concern. Even where the character’s associated descriptions are false of E and a participant responds that the character, say, ‘believes that...E ...’, this may be consistent with a description theory of ‘E’ *because the participant may, without drawing attention to this, be distancing herself from the usage: there may be implicit scare quotes.*

Stories like the following give a feel for the implicit-scare-quote problem. Suppose that the word ‘vixen’ really is covered by a description theory, as it likely is. Now imagine Harry is in a house where there is talk of a vixen in the garden. It quickly becomes apparent from Harry’s near hysterical remarks using of the word ‘vixen’ that he has a belief that he would express ‘Vixens are tigers’. Harry then rushes violently from the house just as Sam is entering. Might we not explain Harry’s behavior to Sam by saying, ‘Harry thinks that there is a ferocious vixen in the garden’? In so doing, there are implicit scare quotes around ‘vixen’, indicating, as Wikipedia puts it, that the expression ‘may not signify its apparent meaning or that it is not necessarily the way the quoting person would express its concept’; we are describing Harry’s belief the way *he* would even though critical of his usage; we are distancing ourselves from the usage. So our mere use of the term ‘vixen’ does *not* show that we have understood Frank as having referred to vixens, hence does *not* count against the description theory of ‘vixen’. Indeed, if our use does have scare quotes that is evidence *for* the description theory. This goes right against the assumption 3 in the above summary.

Though we were well aware of the implicit-scare-quote problem, our preliminary Beyoncé experiment did not address it. This cast doubt on the apparently decisive support it gave to Kripkean anti-descriptivism.²⁸ To get some idea just how serious this problem was, we decided to complicate our next set of experiments. We would not only test usage by eliciting an explanation of behavior, as in the preliminary Beyoncé experiment, but we would also test usage by offering a choice of explanations, selected with the aim of showing whether there were indeed implicit scare quotes in the elicited usage.

These experiments covered not only proper names but also artifactual kind terms, which strike me as much more interesting than proper names. Philosophers have strong Kripkean intuitions about the reference of proper names, intuitions that favor anti-descriptivist theories. In contrast, we have little in the way of firm intuitions about the reference of artifactual kind terms. We really don't have much idea what sort of theory is right for these terms (Devitt and Sterelny, 1999, pp. 93-100).

The results of these experiments were baffling and we certainly could not draw any interesting conclusions about reference from them. But we did learn one thing: the implicit-scare-quote problem is very real. The results did seem to confirm that we were right to be worried about participants distancing themselves from usage. So this way of testing usage by elicited production is problematic.

These experiments support the idea that whenever a person x ascribes thoughts to another person y , there is a real risk of implicit scare quotes, a risk of x describing y 's thought the way y would do so whilst distancing herself from this usage. This risk can be present not only in getting evidence from elicited production but also in getting it from the corpus. So there is a flaw in the earlier thinking about Genone and Lombrozo's uses of 'tyleritis' in ascriptions of belief to Alex (sec. 6.1). These uses are inconsistent with what description theories would predict only on the assumption that the uses are not implicitly accompanied by scare quotes. That assumption may be false.

Our other example of evidence from the corpus against a description theory was provided by the uses of 'Gödel' by Machery *et al.* Should we be worried about implicit scare quotes here too? I suspect not. These uses are not in ascriptions of thoughts to others. There does not seem to be any reason why Machery *et al.* would want to distance themselves from their own usage.

6.4 The Future

Where do we go from here? Most important of all, we must resist the reactionary response of simply resting our theories of reference on referential intuitions. One problem with that response is that, as the case of artifactual kind terms illustrates, we simply do not have intuitions that could do the job for many terms. But the deeper problem is that resting on referential intuitions is not scientifically respectable. Theories of reference, like any scientific theories, need to be tested directly against the reality that they concern. That reality for theories of reference is to be found in linguistic usage.

²⁸ Nat Hansen pressed this concern in comments on the experiment.

So we have to come up with ways to test theories of reference against linguistic usage. Experiments along the lines of our preliminary Beyoncé experiment have not been a success. Can we come up with a modification of this that solves the implicit-scare-quote problem? To do so we need to come up with a reliable way of detecting when people are distancing themselves from the language used in ascribing thoughts. Perhaps we can discover this by testing description theories that, intuitively, seem likely to be true; for example, a description theory of ‘vixen’. Alternatively, perhaps we should abandon the idea of testing usage in ascriptions of thoughts altogether, given the implicit-scare-quote worry that the idea generates. Then we will need to find some other sort of test of usage that is relatively free of this worry. Can we get inspiration here from the corpus example provided by Machery *et al*?

7. Conclusion

We need theories of reference because, it is widely thought, reference relations are explanatorily important relations in theories of language. This gives rise to the question that has concerned this paper: How should we test theories of reference?

The received view is that we should test them against referential intuitions. And the intuitions used have been those of philosophers. Machery *et al* (2004) wonder why it is appropriate to rely on the intuitions of philosophers rather than those of the folk. I wonder why it is appropriate to rely on referential intuitions at all.

It is common for philosophers to think that their practice of relying on referential intuitions is appropriate because they are a priori. I have argued against this view. Stich has suggested another justification for the practice: philosophers might follow linguists in thinking that linguistic intuitions are ‘the voice’ of our linguistic competence. I have argued against this view too and instead urged ‘the Modest Explanation’ of intuitions

According to the Modest Explanation, referential intuitions are empirical theory-laden central-processor responses to linguistic phenomena, differing from many other such responses only in being fairly immediate and unreflective. So we should rely on the intuitions only to the extent that they are reliable indicators of the nature of linguistic reality. And, at best, they are only indirect evidence. We should be seeking more direct evidence by examining the linguistic reality that these intuitions are about: we need to examine linguistic usage. The results of this direct examination can then also be used to assess the reliability of referential intuitions.

I generously include all immediate perceptual judgments among intuitions with the result that all of our theories rest ultimately on intuitions of some sort. So the contrast between indirect and direct evidence for a theory of reference becomes a distinction among intuitions. The indirect evidence comes from *referential* intuitions, whereas the direct comes from immediate perceptual judgments, particularly ones about usage, that deploy no concept of reference; for example, from the judgment that Jack said ‘Einstein was a physicist’.

Referential intuitions differ among themselves in several significant ways. There are the perceptual ones, memory ones, and ones formed in thought experiments. Among the latter there are

ones about humdrum hypothetical cases and ones about fanciful hypothetical cases. And among them all there are differences in the degree to which their reliability depends on expertise. I conjecture that referential intuitions about humdrum cases, whether perceptual or not, are likely to be fairly reliable without much expertise about language. In contrast, referential intuitions about fanciful hypothetical cases, like the ‘Gödel’ cases tested by Machery *et al*, probably require a good deal of expertise.

Finally, I have addressed the problem of testing theories of reference more directly against the evidence of linguistic usage. One source of this evidence is the corpus. I illustrate this with some material used by Genone and Lombrozo in their experiment testing theories of reference. Many of their uses of the invented term ‘tyleritis’ seem to be inconsistent with what description theories of ‘tyleritis’ would predict. But there are notorious difficulties in using the corpus as evidence. So, I have proposed, philosophers should follow linguistics in using the method of elicited production to test their theories of language. But it has so far proved difficult to come up with a satisfactory experimental test because of the ‘implicit-scare-quote’ problem. The problem is that elicited production, indeed the corpus, provides the evidence we need only if speakers are not implicitly distancing themselves from their usage. We need experiments that control for this worry. The experiments conducted so far have not managed this.²⁹

REFERENCES

- Antony, L. (2008), ‘Meta-Linguistics: Methodology and ontology in Devitt’s *Ignorance of Language*’. *Australasian Journal of Philosophy*, 86, 643-56.
- Barber, A (ed.). (2003), *Epistemology of Language*. Oxford: Oxford University Press.
- Cappelen, Herman. (2012), *Philosophy without Intuitions*. Oxford: Oxford University Press.
- Chomsky, N. (1986), *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger Publishers.
- _____. (1995), ‘Language and nature’. *Mind*, 104, 1-61.
- _____. (1996), *Powers and Prospects: Reflections on Human Nature and the Social Order*. Boston: South End Press.
- Collins, J. (2006), ‘Between a rock and a hard place: A dialogue on the philosophy and methodology of generative linguistics’. *Croatian Journal of Philosophy*, VI, 469-503.
- _____. (2008a), ‘Knowledge of language redux’. *Croatian Journal of Philosophy*, VIII, 3-43.
- _____. (2008b), ‘A note on conventions and unvoiced syntax’. *Croatian Journal of Philosophy*, VIII, 241-7.
- Culbertson, J, and Gross, S. (2009), ‘Are linguists better subjects?’ *British Journal for the Philosophy of Science*, 60, 721-36.
- Deutsch, M. (2009), ‘Experimental philosophy and the theory of reference’. *Mind and Language*, 24, 445–66.
- Devitt, M. (1981), *Designation*. New York: Columbia University Press.

²⁹ The first version of this paper was delivered as the Presidential Address at the Society for Philosophy and Psychology 39th Meeting at Brown University in June 2013. I acknowledge the support of the Spanish Ministerio de Economía y Competitividad. (“Reference, selfreference and empirical data” FFI2011-25626). Finally, many thanks to Mark Phelan for helpful comments on the penultimate version.

- _____. (1994), 'The methodology of naturalistic semantics,' *Journal of Philosophy*, 91, 545-72.
- _____. (1996), *Coming to Our Senses: A Naturalistic Program for Semantic Localism*. Cambridge: Cambridge University Press.
- _____. (1998), 'Naturalism and the a priori'. *Philosophical Studies* 92, 45-65 (reprinted in Devitt, 2010c).
- _____. (2003), 'Linguistics is not psychology', in Barber, 2003, pp. 107-39.
- _____. (2006a), *Ignorance of Language*, Oxford: Clarendon Press.
- _____. (2006b), 'Intuitions in linguistics'. *British Journal for the Philosophy of Science*, 57, 481-513.
- _____. (2006c), 'Defending *Ignorance of Language*: Responses to the Dubrovnik papers'. *Croatian Journal of Philosophy*, VI, 571-606.
- _____. (2006d), 'Intuitions', in V. G. Pin, J. I. Galparaso, and G. Arrizabalaga (eds), *Ontology Studies Cuadernos de Ontologia: Proceedings of VI International Ontology Congress (San Sebastian, 2004)*. San Sebastian: Universidad del Pais Vasco, pp.169-76 (reprinted in Devitt 2010c).
- _____. (2008a), 'Explanation and reality in linguistics'. *Croatian Journal of Philosophy*, VIII, 203-31.
- _____. (2008b), 'A response to Collins' note on conventions and unvoiced syntax'. *Croatian Journal of Philosophy*, VIII, 249-55.
- _____. (2008c), 'Methodology in the philosophy of linguistics'. *Australasian Journal of Philosophy*, 86, 671-84.
- _____. (2009), 'Psychological conception, psychological reality: A response to Longworth and Slezak'. *Croatian Journal of Philosophy*, IX, 35-44.
- _____. (2010a), 'What "intuitions" are linguistic evidence?' *Erkenntnis*, 73, 251-64.
- _____. (2010b), 'Linguistic intuitions revisited,' *British Journal for the Philosophy of Science*, 61, 833-65.
- _____. (2010c), *Putting Metaphysics First: Essays on Metaphysics and Epistemology*. Oxford: Oxford University Press.
- _____. (2010d), 'Comments on Genone and Lombrozo', delivered at the Society for Philosophy and Psychology Conference, Portland (OR), June 2010.
- _____. (2011a), 'No place for the a priori', in M. J. Shaffer and M. L. Veber (eds), *What Place for the A Priori?* Chicago and La Salle: Open Court Publishing Company, pp. 9-32 (reprinted in Devitt 2010c).
- _____. (2011b), 'Experimental semantics'. *Philosophy and Phenomenological Research*, LXXXII, 418-35.
- _____. (2011c), 'Whither experimental semantics?' *Theoria*, 27, 5-36.
- _____. (2012a), 'The role of intuitions', in Russell and Fara, 2012, pp. 554-65.
- _____. (2012b), 'Semantic epistemology: Response to Machery'. *Theoria*, 74, 229-33.
- _____. (2013a), 'Linguistic intuitions are not "the voice of competence"', in Haug, 2013, pp. 268-93.
- _____. (2013b), 'What makes a property "semantic"?' in A. Capone, F. L. Piparo, and M. Carapezza (eds), *Perspectives on Pragmatics and Philosophy*. Cham: Springer, pp. 87-112.
- _____. (2014a), 'We don't learn about the world by examining concepts: A response to Carrie Jenkins', in R. Neta. *Current Controversies in Epistemology*. New York: Routledge, pp. 23-34.
- _____. (2014b), 'Lest auld acquaintance be forgot'. *Mind and Language* 29, pp. 475-84.

- _____. (2014c), ‘Linguistic intuitions: In defense of “Ordinarism”’. *European Journal of Analytic Philosophy* ###.
- _____. (forthcoming), ‘Philosophy without intuitions? A response to Herman Cappelen’. *Inquiry*.
- ____ and Sterelny, K. (1999), *Language and Reality: An Introduction to the Philosophy of Language*, 2nd edn. 1st edn 1987. Oxford: Blackwell Publishers.
- Donnellan, K. S. (1972), ‘Proper names and identifying descriptions’, in D. Davidson and G. Harman (eds), *Semantics of Natural Language*. Dordrecht: D. Reidel, pp. 356-79.
- Dummett, M. (1978), *Truth and Other Enigmas*, Cambridge, MA: Harvard University Press.
- Dwyer, S. and Pietroski, P. (1996), ‘Believing in language’. *Philosophy of Science*, 63, 338-73.
- Fitzgerald, G. (2010), ‘Linguistic intuitions’. *British Journal for the Philosophy of Science*, 61, 123-60.
- Genone, J. and Lombrozo, T. (2012), ‘Concept possession, experimental semantics, and hybrid theories of reference’. *Philosophical Psychology*, 25, 717-42.
- Gross, S. and Culbertson, J. (2011), ‘Revisited linguistic intuitions’. *British Journal for the Philosophy of Science*, 62, 639-56.
- Haug, M. (ed.). (2013), *Philosophical Methodology: The Armchair or the Laboratory?* London: Routledge.
- Hawthorne, J. and Manley, D. (2012), *The Reference Book*. Oxford: Oxford University Press.
- Hintikka, J. (1999), ‘The emperor’s new intuitions’. *Journal of Philosophy*, 96, 127-47.
- Ichikawa, J., Maitra, I., and Weatherson, B. (2011), In defense of a Kripkean dogma. *Philosophy and Phenomenological Research*, 82, 418-35.
- Jackman, H. (2009), ‘Semantic intuitions, conceptual analysis, and cross-cultural variation’. *Philosophical Studies*, 146, 159-77.
- Katz, J. (1997), ‘Analyticity, necessity, and the epistemology of semantics’. *Philosophy and Phenomenological Research*, 57, 1-28.
- Kripke, S. A. (1980), *Naming and necessity*. Cambridge, MA: Harvard University Press.
- Laurence, S. (2003), ‘Is linguistics a branch of psychology?’ in Barber, 2003, pp. 69-106.
- Loar, B. (1987), ‘Subjective intentionality’. *Philosophical Topics*, 15, 89-124.
- Ludlow, P. (2011), *The Philosophy of Generative Linguistics*. Oxford: Oxford University Press.
- Ludwig, K. (2007), ‘The epistemology of thought experiments: First-Person approach vs. third-person approach’. *Midwest Studies in Philosophy*, 31, 128-59.
- McKinsey, M. (1987), ‘Apriorism in the philosophy of language’. *Philosophical Studies*, 52, 1-32.
- ____ (1994), ‘Individuating Belief’, in J. Tomberlin (ed.) *Philosophical Perspectives*, 8: *Logic and Language*, Atascadero, CA: Ridgeview Publishing Company.
- Machery, E. (2011), ‘Expertise and intuitions about reference’. *Theoria*, 27, 37-54.
- ____. (2012), ‘Semantic epistemology: A brief response to Devitt’. *Theoria*, 74, 223-7.
- ____, Mallon, R., Nichols, S., and Stich, S. P. (2004), ‘Semantics, cross-cultural style.’ *Cognition*, 92, B1-B12.
- ____, ____ , ____ , and ____ . (2013), ‘If folk intuitions vary, then what?’ *Philosophy and Phenomenological Research*, 86, 618-35.
- ____ and Stich, S. P. (2012), ‘The role of experiments in the philosophy of language’, in Russell and Fara, 2012, pp. 495-512.
- Martí, G. (2009), ‘Against semantic multi-culturalism’. *Analysis*, 69, 42-48.
- ____. (2014), ‘Empirical data and the theory of reference’, in M. O’Rourke (ed.), *Topics in Contemporary Philosophy, Volume 10: Reference and Referring*. Cambridge, MA: MIT Press.

- Matthews, R. J. (2006), 'Could competent speakers really be ignorant of their language?' *Croatian Journal of Philosophy*, VI, 457-67.
- Maynes, J. and Gross, S. (2013), 'Linguistic intuitions'. *Philosophy Compass* 8/8, 714–30.
- Miščević, N. (2006). 'Intuitions: The discrete voice of competence'. *Croatian Journal of Philosophy*, VI, 523-548.
- _____. (2009). 'Competent voices: A theory of intuitions'. <http://oddelki.ff.uni-mb.si/filozofija/en/festschrift>
- Ostertag, G. (2013), 'The "Gödel" effect'. *Philosophical Studies*, 166, 65–82.
- Pietroski, P. (2008), 'Think of the children'. *Australasian Journal of Philosophy*, 86, 657-69.
- Rattan, G. (2006), 'The knowledge in language'. *Croatian Journal of Philosophy*, VI, 505-21.
- Rey, G. (2006) 'Conventions, intuitions and linguistic inexistent: A reply to Devitt'. *Croatian Journal of Philosophy*, VI, 549-69.
- _____. (2013), 'The possibility of a naturalistic Cartesianism regarding intuitions and introspection', in Haug, 2013, pp. 243-67.
- Russell, G. and Fara, D. G. (eds). (2012), *Routledge Companion to the Philosophy of Language*. New York: Routledge.
- Slobodchikoff, C. N. (2002), 'Cognition and communication in prairie dogs', in C. Allen and G. M. Burchardt (eds), *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*. Cambridge MA: MIT Press, pp. 257-64.
- Smith, B. C. (2006), 'Why we still need knowledge of language'. *Croatian Journal of Philosophy*, VI, 431-56.
- Sprouse, J. and Almeida, D. (2013), 'The role of experimental syntax in an integrated cognitive science of language', in K. Grohmann and C. Boeckx (eds), *The Cambridge Handbook of Bilingualism*,. Cambridge: Cambridge University Press, pp. 181-202.
- Stich, S. P. (1983), *From Folk Psychology to Cognitive Science: The Case Against Belief*. Cambridge, MA: Bradford/MIT Press.
- _____. (1996), *Deconstructing the Mind*, New York: Oxford University Press.
- Sytsma, J., & Livengood, J. (2011), 'A new perspective concerning experiments on semantic intuitions'. *Australasian Journal of Philosophy*, 89, 315–32.
- Textor, M. (2009), 'Devitt on the epistemic authority of linguistic intuitions'. *Erkenntnis*, 71, 395-405.
- Thornton, R. (1995), 'Referentiality and wh-movement in child English: Juvenile D-Linkency'. *Language Acquisition*, 4, 139-75.
- Weinberg, J. M., Gonnerman, C., Buckner, C, and Alexander, J. (2010), 'Are philosophers expert intuiters?' *Philosophical Psychology*, 23, 331-55.
- Williamson, T. (2007), *The Philosophy of Philosophy*. Oxford: Blackwell Publishing.
- Wisniewski, E. J. (1998) 'The psychology of intuitions', in M. R. DePaul and W. Ramsey (eds) *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, London: Rowan and Littlefield Publishers, pp. 45-58.