Reference, Theories of

Michael Devitt, City University of New York, New York, USA

**CONTENTS**
- What is a theory of reference?
- Varieties of theories of reference
- Theories of reference and cognitive science

*Theories of reference are about the relations between representations and the world that are thought to determine (at least partly) the meaning or content of those representations.*

**WHAT IS A THEORY OF REFERENCE?**

Referential relations hold between representations and the world: in particular, between parts of sentences – words – and the world and between parts of thoughts – concepts – and the world. The most obvious example of such a relation is the naming relation, the sort that holds between ‘Winston Churchill’ and the famous statesman. However, it is usual to think of reference as covering a range of semantically significant relations: for example, between the word ‘witty’ and Wittgenstein, and between the concept ‘bachelor’ and all bachelors. These relations are variously described by the terms ‘designate’, ‘denote’, ‘signify’, ‘apply’, ‘satisfy’, ‘instance’, ‘fall under’, and ‘about’.

Reference is important because it is thought to be the core of meaning and content. Thus, the fact that ‘Winston Churchill’ refers to that famous statesman is the core of its meaning, and hence of its contribution to the meaning of any sentence – for example, ‘Winston Churchill is witty’ – that contains it. And the fact that the concept ‘bachelor’ refers to all bachelors is the core of its content and hence of its contribution to the content of any thought – for example, the thought ‘Winston Churchill is not a bachelor’ – that contains it. The usual view is that meaning determines reference.

The first question that arises about the reference of a term is: what does the term refer to? Sometimes the answer seems obvious. For example, ‘Winston Churchill’ refers to the famous statesman. Even this apparently obvious answer is rejected by those who take words to refer to ideas in the mind or to mental representations. But the received view in contemporary philosophy is that words refer to the world. This leaves room for different opinions about the reference of predicates. Some take ‘witty’ to refer to the property of Wittgenstein, some to the set of all witty things, and some to each witty thing separately.

The central question about reference is: by virtue of what does a representation have its reference? A theory of reference answers this question by explaining the relation of the representation to its referent. It has proved surprisingly hard to provide these explanations. There was a surge of interest in theories of reference in the twentieth century.

**VARIETIES OF THEORIES OF REFERENCE**

According to ‘description’ theories of reference, the reference of a representation is determined by certain descriptions – other representations – inferentially associated with it by competent speakers: these descriptions identify the referent. The simplest form of description theory, derived from the work of Frege (1893) and Russell (1912), specifies a set of descriptions each of which is necessary and all of which are sufficient for reference determination: for example, the references of ‘adult’, ‘unmarried’ and ‘male’ might be jointly sufficient and individually necessary for the reference of ‘bachelor’. According to another form of description
theory, the reference is whatever is picked out by a (weighted) majority of certain descriptions associated with the representation. On this ‘cluster’ theory, no one description is necessary for reference fixing (Searle, 1958).

Around 1970, several criticisms were made of description theories of proper names, such as ‘Winston Churchill’ (Kripke, 1980; Donnellan, 1972), and ‘natural kind’ words, such as ‘gold’ and ‘tiger’ (Kripke, 1980; Putnam, 1975). One important criticism is that the theories yield unwanted necessities. The description we inerrationally associate with ‘tiger’ is along the lines of ‘large carnivorous quadrupedal feline, tawny yellow in color with blackish transverse stripes and white belly’. So the objects that ‘tiger’ refers to must be four-legged and striped. Yet tigers need not be thus: a tiger might lose a leg; in a different environment tigers might not be striped. Another important criticism is that people who seem perfectly able to use words to refer are too ignorant to provide descriptions adequate to identify the referents. Thus, some who use ‘elm’ and ‘beech’ cannot supply descriptions that distinguish elms from beeches; many who use ‘gold’ cannot distinguish gold from fool’s gold. Worse, speakers are often so wrong about the referent that the descriptions they would provide apply not to the referent but to other entities or to nothing at all. Sometimes the whole speech community is ignorant or wrong about the referent. Thus, it was once common to associate ‘fish’ with ‘whale’. Description theories of these words seem to require too much knowledge, placing too great an epistemic burden on speakers.

Putnam added a further argument, built around the following fantasy. Imagine that somewhere in the galaxy there is a planet called Twin Earth. Twin Earth, as its name suggests, is very like Earth. In particular, each Earthian has a doppelgänger on Twin Earth who is a cell-for-cell duplicate of the Earthian. Twin Earth differs from Earth in one respect, however: the liquid that the Twin Earthians who appear to speak English call ‘water’ – liquid that is superficially indistinguishable from what we call ‘water’ – is not H₂O but a very different compound XYZ. So Oscar on Earth and Twin Oscar on Twin Earth refer to different liquids by ‘water’. Yet Oscar and Twin Oscar are doppelgängers, associating exactly the same descriptions with ‘water’ (which is more plausible if we place Oscar and Twin Oscar in 1750, before the chemical composition of water was known). So those associations are not sufficient to determine reference, and the description theory must be wrong. Indeed, nothing going on in the head is sufficient to determine reference. As Putnam put it, ‘meanings just ain’t in the head’: meanings require a relation to something external to the thinker.

This is not to say that description theories fail for all representations. They still seem plausible for ‘bachelor’, and perhaps even for ‘pencil’ and ‘pediatrician’ (cf. Putnam, 1975). But even where they work, description theories have a problem: they are essentially incomplete. Suppose that a theory claims that the reference of ‘bachelor’ is determined by the references of ‘adult’, ‘unmarried’ and ‘male’. We then need to explain the references of those words to complete the explanation of the reference of ‘bachelor’. Description theories might be offered again. But then the explanation will still be incomplete. At some point we must offer a theory of reference that does not make the reference of one word dependent on those of others. We need an ‘ultimate’ explanation of reference that relates some words directly to the world. Description theories ‘pass the referential buck’. The buck must stop somewhere if there is to be any reference at all.

‘Verificationist’ theories of reference implicitly acknowledge this point. They take a broader view than description theories of the required identification: speakers refer to whatever objects they would identify as the referents, whether by description or by recognition. Speakers recognize a referent by pointing it out in a crowd saying, for example, ‘that person’. But these theories still seem to place too great an epistemic burden on speakers: we can only dimly call to mind the appearances of many objects we refer to, and are often mistaken.

There has been disagreement over whether reference can be explained in non-semantic or non-mentalistic terms. Many think not: reference is, in a sense, irreducible. From a naturalistic perspective, this view is unacceptable: reference must be explained in scientifically acceptable terms, ultimately in physical terms. Attempted explanations have appealed to one or more of three causal relations between representations and reality. First, there is the historical cause of a particular token, a causal chain going back to the dubbing of the token’s referent. Theorists interested in this have emphasized the ‘reference borrowing’ links in the chain: in acquiring a word or concept from others we borrow their capacity to refer, even if we are ignorant of the referent (Kripke, 1980; Donnellan, 1972; Putnam, 1975; Burge, 1979). Second, there is the reliable cause of tokens of that type: a token refers to objects of a certain sort because tokens of that type are reliably correlated with the presence of those objects. The token ‘carries the information’ that a certain situation holds, in much the same
way that tree rings carry information about the age of a tree (Dretske, 1981; Fodor, 1990). Third, there is the teleological cause, or function, of tokens of that type, where the function is explained along Darwinian lines: the function is what tokens of that type do that explains why they exist, what the type has been ‘selected for’ (Millikan, 1984; Papineau, 1987).

What all these recent developments suggest is that the reference of any word or concept will be explained by some such ‘ultimate’ theory, or by a description theory, or by a theory that combines elements of both. And different sorts of representations may have different sorts of theories.

Finally, it should be noted that some have denied, in one way or another, the reality of reference (and hence the need for, or even possibility of, an explanation of it). Perhaps the most influential of these denials is the ‘deflationary’ theory of reference, which accompanies the deflationary theory of truth (Horwich, 1998).

THEORIES OF REFERENCE AND COGNITIVE SCIENCE

Cognitive science holds that humans (and some other things) have mental representations – concepts – the contents (meanings) of which determine behavior (Margolis and Laurence, 1999). The common view is that a concept (and the word that expresses it) typically refers, and its reference is central to its content and nature. So, on this view, theories of reference should bear importantly on theories of concepts. However, it is often not clear how they bear on a particular theory.

Many cognitive scientists take the inferential associations of a concept to constitute its content and nature. This could involve a rejection of the centrality of reference, but it usually does not. Most seem to hold the appealing Fregean view that a concept’s inferential associations determine its reference, thus endorsing a description theory of reference. The bearing of theories of reference is then obvious. However, some may prefer a ‘two-factor’ theory of content according to which reference constitutes one factor and inferential associations constitute an independent factor which does not determine reference. Most of what is said about concepts would then concern that independent factor and so would have nothing to do with theories of reference.

Finally, although psychological research on concepts is usually presented as if it concerned the nature of concepts, it may really concern something very different: the ‘structure of knowledge’. On this view, the research is not throwing light on what concepts are – on what constitutes their contents – but on what they do – on their cognitive role, in particular, their role in cognitive skills. Theories of reference have little bearing on this.

We can distinguish three particular implications of theories of reference for cognitive science.

First, it is common to have a holistic view of concepts: their contents are given by all or most of their inferential associations. When combined with the appealing Fregean view, this yields a holistic description theory of reference. A consequence of such a theory is that the mistaken views we tend to have about the putative referents of our concepts would lead to widespread reference failure: most of our concepts would not fit anything in the world. So the Fregean view counts against holism; and the holistic view encourages a two-factor theory of concepts.

Second, verificationism is common in psychology: a concept’s content is thought to be constituted by those features that lead us to apply it to the world. This application is naturally taken as determining reference. So the concept ‘bird’ refers to the objects that those who have the concept identify as birds. The recent developments in the theory of reference, particularly the view that reference can be borrowed, cast considerable doubt on verificationism: a person may have the concept ‘elm’, not by virtue of any ability to identify elms, but by virtue of being appropriately linked with a linguistic community.

Third, the simple Frege–Russell description theory calls to mind the ‘classical’ theory of concepts in psychology, and the more modern cluster theory calls to mind ‘family resemblance’, ‘prototype’ and ‘exemplar’ theories (Smith and Medin, 1981). In so far as the psychological theories are rightly construed as analogous to description theories, the recent developments cast doubt on them, particularly where they concern natural kind concepts like ‘bird’.

Still, this may well be the wrong construal of some of these psychological theories. Thus, prototype theory usually seems to claim, not that a concept refers to an object that has most of the prototypical features (the cluster theory), but rather that it refers to an object to the degree that it has those features. Thus, the concept ‘bird’ is said to refer to penguins to a lesser degree than to robins. But this could be so only if penguins were birds to a lesser degree than robins, not ‘fully’ birds. For clearly the concept ‘bird’ fully refers to birds, and so if penguins are fully birds then the concept ‘bird’ must ‘fully’ refer to them. So prototype theory
clashes disastrously with a biology which tells us that penguins are indeed ‘fully’ birds, as much so as robins. In the light of this, perhaps prototype theory should not be taken as a theory of the nature and reference of concepts but rather of their roles in recognition.

In conclusion, it is proving hard to discover fully satisfactory theories of reference. This has led some to pessimism about reference. This pessimism has a price: either we must abandon the view that reference is central to meaning (content), or, more radically still, we must abandon the idea that we have meaningful representations at all.

References


Further Reading


