

A Utility Theory of 'Truth'

This paper is about some of the philosophical issues associated with taking the language of business entirely seriously. The approach it takes is to explore some relevant material from the philosophy of language - principally to do with truth, meaning and translation. The conclusion reached is that while it may be irrecoverably difficult to make sense of what 'taking seriously' would amount to, there are still some interesting consequences of the approach which might inform debates on business ethics and on the philosophy of language itself.

Introduction

We sometimes say of people things like this: 'He just says whatever he thinks he can get away with' or: 'She just says whatever will suit her'. When we do this, we are implying a special kind of dishonesty - a use of language to get what the user wants rather than to tell the truth. Sometimes we are kinder, and recognise the necessity of using language this way, when protecting oneself from an enemy or when negotiating a contract. In more extreme cases - e.g. the utterances of some politicians and advertisers - the public recognition that they must be dishonest under certain circumstances actually contributes to validating their special use of the language under those circumstances. What I want to do in this paper is consider the possibility of a language ("Utlish") where utility is the *principal* determinant of what the players say - in other words, to attempt to explore and to take completely seriously something like a utility theory of truth. In a language where utility is the principal criterion of correct usage, I say what it benefits me to say, and I expect all others to do the same. Speakers of Utlish, unlike the speakers of English in my examples, don't just talk this way for operational reasons. They recognise it as the *correct* way to speak - the equivalent, in some respects, of what an English speaker would call 'truthful'.

Issues

There are several reasons for trying to do this. One is that we commonly regard this kind of behaviour as 'dishonest'. This judgement, however, is a move in *our* language game, where the standards of truth are different. It *would* be dishonest to 'tell lies' in *our* language to get what we want. To say the same of Utlish is to miss the point - it's *meant* to work that way.

The second reason is that in order to engage in a serious ethical exchange with business people, we need to hear their language properly - this means avoiding making facile assumptions about what it is they are doing when they speak, and it means making a serious attempt to understand the underlying rules of their language. This, arguably, is part of what is involved in making space for 'otherness'.

A third reason for doing it is that, according to certain flavours of classical economic theory, we may, in a certain sense, already speak Utlish. Faced with a choice between different courses of action - even where one of these is telling the truth and another is telling a lie - we would be expected to choose that course which most enhanced our utility. If we do this completely consistently (as the theory must assume) then having a standard of truth which conflicted with utility would, by definition, be to have a standard of truth to which we could not generally adhere. Some writers have argued, in different ways, that truth telling and utility largely converge, especially when the general utility of things like having a language which works, or having a reputation for honesty, for instance, is taken into account. Arguments of this type - that truth telling has utility - are, however, different from arguments that there could be a language in which utility is the substantive and definitive test of truth. For such a language, to speak in a way which enhances the speaker's utility is to speak properly, to be a paradigmatic user of the language. It is not simply to, say, tell the truth because that's a particularly, or a generally, useful thing to do.

The Method

Let us, therefore, set up the following thought experiment. We take two languages – English and Utlish – that are similar in most respects, except that Utlish speakers, as a rule (literally), try to say things which maximise their utility whereas English speakers try to tell the truth in the conventional sense which we all recognise. I will call true in Utlish ‘true_(U)’ and true in English ‘true_(E)’.

I have to say at the outset that I don’t think this experiment will work. I can tell you this immediately, because a comfortable feature of academic writing is the absence of any stylistic or narrative convention which obliges me to keep the denouement a secret while I develop the plot. (After all academics read out of duty, and not for pleasure, so you will all read to the end even though you know how things turn out by paragraph five.) I think that the reasons that the experiment won’t work are interesting, however, and that justifies most of what appears from paragraph six onwards.

Theories of Truth

Before exploring the Utlish experiment itself, it’s worth making a few preliminary comments about theories of truth generally. Philosophers have struggled with the concept of truth since philosophy began, and, for about the same length of time, have found themselves facing a dilemma. The dilemma is that theories of truth which work often seem meaningless, while meaningful theories of truth routinely turn out to be false. As examples of the former, we have various minimalist theories derived from one version of Tarski’s ‘disquotational’ account. Examples of the latter are, for instance, the correspondence and coherence theories of truth, and the approaches suggested by the pragmatists.

Aristotle said that to tell the truth was ‘to say of what is that it is, or of what is not that it is not’ (Blackburn and Simmons, 1999, p1). There are, no doubt, scholarly reasons for believing that he was being absolutely serious here, and not simply making a satirical comment on the intractability of the issue. In fact, this elliptical, almost content-less quality, is something nearly all of the credible theories of truth share. For instance, most modern theorists take Tarski’s (1944) ‘disquotational’ account as a starting point. It is most concisely (although slightly misleadingly) expressed as: “X” is true if and only if X. In other words, the sentence “Snow is white” is true if and only if snow is white. It looks better written down than read – quotation marks verbalise clumsily. Both of these accounts seem to do little more than state the obvious. They both have the character of being impossible to deny without immediately raising questions about the meanings of the words they contain, but at the same time having a pretty high ‘so-what?-ness’ score. Perhaps, however, they do hint at the problems associated with stating a theory of truth in the language to which the theory applies.

In fact, Tarski’s account, at least, was never meant to work exactly in the way indicated – the sentence in the quotes should be regarded as a sentence in an ‘object language’ (the language under examination) and the whole statement should be regarded as a statement in a ‘meta-language’ – the language within which the examination is being carried out. This makes much more sense at a practical level – statements of this disquotational form are routinely encountered when discussing a translation, and are closely related to explanations of meaning (knowing the meaning of a statement is, at least partly, to know when the statement would be true).

Aristotle’s account of truth (arguably Tarski’s as well, depending on interpretation) is a type of ‘correspondence’ theory (Blackburn and Simmons, 1999, p1). Correspondence theories of truth are theories which make the truth of a statement depend upon whether it corresponds to some state of affairs in the world. In other words, ‘The cat is on the mat’ is true if and only if the cat is on the mat. (This *sounds* like Tarski, which is why Tarski may sometimes be described as a correspondence theorist). These theories all have a problem with what counts as ‘correspondence’. This clearly isn’t a simple relationship (like a mathematical correspondence) since the world doesn’t seem to have a linguistic structure to which our articulations can be matched up. A bigger, and completely unavoidable, problem is the one which arises when we try to think about how we would go about testing the truth of *our theory of correspondence*. Since we can only do this by referring to that theory for adequate criteria of truth, we find ourselves immediately begging the question.

As it turns out, all general theories of truth can seem to have this essentially paradoxical nature. In one way or another, they flirt with the same self-referential weakness as correspondence theories do. Consider any theory - theory 'T' - which states the conditions under which a sentence is true. Now consider the question 'Is "T" true?'. The difficulty is impossible to avoid: Any general account of truth from which criteria of truth can be derived – arguably, therefore, any account which has any content – is viciously circular.

A perverse, but interesting, way of addressing this paradox is suggested by the following Wittgensteinian (Wittgenstein 19XX) type of consideration: How can it make sense for someone to (seriously) ask the question 'What are the criteria for truth in this language?' since such a question comprises an implicit admission that they can have no reliable way of testing the truth of the answer. Another way of putting this is to observe that to be able to speak a language is to be able to speak *truly* in that language, so the question 'what is truth?' asked by a speaker of the language must always be at least partly rhetorical - they must know the answer at least well enough to be able to use the language to ask the question. Note that there is no difficulty associated with being able to do something (such as telling the truth) without being able to articulate that knowledge (in a theory of truth, for example). Knowing how to ride a bicycle seems to be an example of knowledge of exactly this type.

These kinds of consideration lend support to accounts of truth which depend upon coherence between different statements, rather than upon correspondence with reality. These accounts can avoid the paradox by what looks at first like a piece of slight of hand. For a coherence theorist, the truth of a statement depends upon its coherence with other statements which we already know to be true. This sounds question begging, because we need to have *some* true statements before we can establish the truth of any others. But the claim that we may *not* have some true statements turns out to be paradoxical as well, incoherent with itself, in fact, since it cannot claim that it may not, itself, be true. Once again, we are left with a necessary, but mysterious, concept of truth. But we have made a step forward – we at least now have an inductive definition: Since there must be some statement (T1) which is true, we can argue that any statement which is inconsistent with that statement must be false, and any statement which is true if and only if T1 is true must also be true. A candidate for T1 might be something like 'There are some true statements'. This approach obviously doesn't give us a criterion which can be applied to all the statements we would like to apply it to, but that shouldn't be surprising – there are, after all, many statements whose truth or falsehood we cannot absolutely decide.

This inductive approach is possibly the only one that can hope to produce a non-circular theory of truth with any content. Whether the candidate for T1 that I have suggested is the best, or only, one, one thing is clear: T1 must be some kind of statement which doesn't give any direct content to the concept of truth, but which cannot be false in the language in which it is stated. Clearly, also, T1 may be some conjunction or disjunction of statements which all have this quality (i.e. there may be many necessarily true statements with which all other true statements must be coherent). It seems likely, also, that whatever T1 is, it is going to have some of the characteristics of the example I suggested. That is, it will need to be a statement about what is necessary in order for it to be possible to make statements at all. Since we can't make a statement that statements are impossible (!), then whatever is necessary for the possibility of statements must be true in any language game.

The concept of truth, then, is closely connected with the notions of using language at all, and of using language correctly.

Before considering the consequences of this approach for the Utlish experiment, it is worth backtracking slightly and exploring another avenue of existing theory. The paradoxes associated with theories of truth, and the perverse 'obviousness' of such theories of truth as *can* be articulated, have sometimes been taken to suggest that general statements about truth in the language to which they refer may always be a kind of mistake (as some minimalist or 'deflationary' theorists suggest – see Horwich, 1998). What, though, about general statements about truth in another language? Can we express, in English, a general theory of 'truth in Utlish'? A difficulty with this is that it assumes we can separate two activities which may not be so easily separated. The first is translation, and the second is determining what counts as

'true'. It should be clear that these *need* to be separated if we are to be able to make statements about the concept of 'true' in a foreign language, because we need to be quite clear about what the speakers of this language *count* as true before we can work out what their theory of truth is. The difficulty is that we cannot translate the language in the first place without making assumptions about what the native speakers count as true. Broadly speaking, we test our translation against our *own* concept of truth – we would tend to judge that we had mistranslated if the results were false under circumstances where we expected them to be true.

Fortunately (for linguistic anthropology and for the theme of this paper) we only need to have *generally* consistent concepts of truth to make translation possible – we can tolerate occasional difficulties and differences, so long as, in general, what the translator and the native speaker count as true match up fairly well. In this way, we might learn how to speak a language by focussing on areas of agreement (e.g. where the ordinary physical world is concerned), and then go on to explore areas of disagreement (e.g. with respect to religious belief) once we are fairly sure we know how to speak the language well enough not to be confused by a bit of modal dissonance.

Clearly, the translation argument (initially put forward by W.V. Quine and developed by Donald Davidson) would seem to mean that we can't make sense of Utlish if the concept of truth implicit in the language is fundamentally different from ours. This might be the case if the *only* concept of 'truth' the Utlish speakers recognised was utility (especially if, prior to attempts at translation, we didn't know this). This would render the language incoherent to us, since we would have no usable test for a good translation. However – as indicated – it may be possible to have an intelligible Utlish in which there were only *some* areas where our concepts of truth diverged. In other words, the Utlish concept of truth and our concept of truth would have to result in a lot of statements shared between Utlish and English having the same truth value, and only a few counting as 'false' in one language and 'true' in the other.

This might work just fine, because it is, quite often, in our best interests to say what is (in *our* sense) true. In other words, utility and veracity don't automatically diverge. An Utlish user would, presumably, say the same thing as we would say if asked to describe to a doctor the symptoms of a curable disease. This consideration has led some theorists to propose 'pragmatic' theories of truth. These, of course, suffer from the same self-referential difficulties as all theories of truth which have definite content (see above). William James proposed a *general* 'pragmatic' theory of truth which slightly conflated utility and pragmatic acceptability. He suggested that 'true' was an adjective which we applied to our 'current working hypotheses', and that utility set the standard of which hypotheses work. This type of utility theory is subtly different from the utility theory of truth exemplified in Utlish, however, since 'true' in Utlish refers to statements which improve the utility of their individual users, not of language users in general. For Utlish users to be conforming to a Jamesian set of truth rules, it would need to be shown that the increasing the utility of individual language users was always at least consistent with increasing the utility of all language users.

Even where truth and utility converge, however, the Utlish speaker wouldn't make the statements which were both true_(E) and true_(U) because they were 'true' in *our* sense of true, but because saying these things would lead to an increase in the speaker's utility. In other words, they would say things which appeared to be true_(E) to the doctor in the example because that would lead the doctor to do the thing which was most beneficial to the Utlish speaker. This is the 'right' way to speak in Utlish.

If we wanted to translate Utlish into English (i.e. if we wanted to be sure that we were doing this properly), we might focus at first on those circumstances under which it was in the Utlish speaker's interest to say what an English speaker would regard as true. A difficulty with this possibility is that it depends upon our being able to fairly reliably distinguish those circumstances where the Utlish speaker's utility would be increased by their telling the 'truth' (in our terms). This would lead us into a translation problem very similar to Quine's problem - as outlined above. Faced with an expression which seemed to have been translated wrongly, we would always be in the position of trying to decide whether to revise our translation rule or to revise our ideas about the Utlish speaker's utility function.

There is a more serious difficulty, however, and it is one of the difficulties which led me to say, near the beginning, that I thought the Utlish experiment wouldn't work. The difficulty is that the Utlish speaker would need to have some concept in Utlish which was *similar* to the concept of 'truth' in English, in the same way as, in English we have the concept of 'utility' which is closely related to the Utlish concept of truth. The reason for this is that the Utlish speaker would need to know how to tell the truth (in our terms) when doing this would increase his or her utility. In other words, there would need to be, in Utlish, the concepts of, e.g. 'conformance to reality' and of 'true', corresponding (respectively) to the concepts 'true' and 'useful' in English. Looking back at the comments about translation above, it should be clear that a translator not previously primed by the supposition upon which the experiment described in this paper rests would simply translate the Utlish term for 'conformance to reality' into the English word 'true', and the Utlish word for 'true' into the English word for 'useful'. This translator might wonder about the ethical standards of Utlish speakers, rather than about their standard of truth. Given the further assumption that I made at the beginning, that Utlish was like English in all respects *except* for the Utlish users' conception of 'truth', it should be clear that to a listener who had not been warned in advance, Utlish and English would sound identical.

This consonance would extend even to statements appearing to be about conceptions of truth. Notice that it would almost certainly not be in a Utlish speakers interests to say 'My conception of talking properly is to say what I think will lead to an increase in my personal utility', and so this would be an invalid move in Utlish. It would be a *valid* move in Utlish to state a concept of truth which sounded to an English speaker like $\text{true}_{(E)}$, because it would probably be most in the Utlish speakers interest to have people believe this.

Consequences

One of the places where this train of thought is heading is this: On the basis of the kinds of discussion of translation outlined above, it should be clear that we always have some choices when we are faced with trying to resolve the ambiguities which might arise when translating a new language. In the example of Utlish, we can choose between hypotheses about (1) the Utlish concept of truth, (2) the meanings of some Utlish words and phrases and (3) the ethical values of Utlish users. There is no absolute rule about how to make this choice, and there is no 'fact of the matter' about which of the hypotheses is the correct one. Where does this leave us with the original problem of taking the language of business seriously? It is possible to interpret some writing on 'making room for the other' as arguing always in favour of hypothesis of type (1) or (2) rather than (3). In other words, we should not judge or doubt the ethical content of a 'foreign' culture where it is possible to construct a theory of meaning or truth for that culture which makes its ethical position intelligible to us. (It should be clear that this is an ethical position in *our* culture, with self-referential ramifications which cannot be developed here.)

This may also be a place to reflect on what the consideration of such an alien language game can tell us about our *own* truth values. One thing seems quite clear: this is that, within reason, a language such as Utlish could be made to 'work' in some sense. Arguably, it already does, in cultures (such as types of business culture) where language is used mainly 'operationally' – to achieve certain ends, rather than, say, to convey understanding. The fact that this kind of language 'works' is a challenge to our more elevated conceptions of truth, under which lie a prejudice that only a language game with a rigorous conception of truth can survive in the long term – can be ultimately playable. This is especially disturbing when we consider that the only general theories of truth which seem to have some hope of succeeding are some of the 'coherence' theories, which makes the truth of a statement depend upon how that statement fits into the language (coheres with other statements). This type of theory comes very close to saying that a statement is true if it 'works' in our language game. Since operationally effective statements 'work' in the type of business language game caricatured above, how can we argue that our conception of truth is better than the conception of truth exemplified in such an operational language? The answer I suggested above, that we must refer ultimately to the playability of the game, to statements about the possibility of language, ('This game is playable' is a statement to which all other statements in the game must cohere), doesn't offer much comfort.

Another, slightly different, way of asking this question is as follows: Is it possible, after all, to have a language 'game' where truth is not a relevant category? We have many other kinds of ordinary games where this is the case – where appropriate and inappropriate moves are clearly distinguished, but where truth doesn't play a part in making this distinction. The answer to this lies, again, with translation. We might be able to recognise such a game, and to work out its rules, but it is only when we can translate the moves in the game into statements in *our* language that we will regard it as a language, and this translation will depend on our having a conception of what counts as 'true' in the game. In other words, the idea of a *language* game without a conception of truth is incoherent.

The various hermeneutical and near hermeneutical descriptions of how rationality evolves (e.g. Habermas' conception of an 'ideal speech community') sometimes seem to suggest that the 'best' type of language games will develop under circumstances where there is the least likelihood of divergence between the utility of a statement to its utterer and the truth (as in 'truth(E)') of the statement. This is certainly a partial bowdlerisation of the hermeneutical theme, but I think that it represents a real problem and that it exists even for a more sympathetic representation of the hermeneutical argument. University researchers are, arguably, deliberately placed in a position where 'truth' and 'utility' converge for them. There is a 'truth' theory about why a society might invest in this type of activity, and a 'utility' theory. The truth theory is the somewhat Jamesian view that there is a pragmatic utility attached to good hypotheses, and that it is worth refining our working hypotheses to ensure that they have a continued utility in the future. The utility theory is that a cultural standard of 'higher truth' benefits the operational users of language within a society (those in powerful positions are most likely to use language this way). It benefits them because it creates a wider expectation that they are 'telling the truth' and therefore makes it easier to use language for control purposes. These users will, therefore, invest in the institutions which help to establish this cultural value.

I would like to finish with a personal reflection which cannot be part of an academic paper, but which it might be acceptable to air at this conference. When I first became an accountant, having taken a degree in philosophy, I found that it was very difficult to make some of the very general philosophical categories that I had learned to manipulate 'stick' to the new world that I had entered. This wasn't because they were 'irrelevant' or 'academic', but because I constantly found myself struggling with problems of interpretation. I might have been able, for instance, to bring ethical or epistemological considerations to bear on the activities and arguments of the people I found myself working among, if I could only come up with a reasonably unambiguous account of what it was that they were doing. The fact that I often couldn't led me to think very hard about what this fact might mean.

The fairytale about Utlish and English is designed to explore some of the conclusions that I reached. As writers on the various paradoxes of translation have pointed out, not only do irresolvable ambiguities arise when translating between two languages, they also arise when we try to decide what really counts as another language or as the same language. After all, each English speaker makes the assumption that other users of English are actually speaking English, and not some other language which sounds superficially similar but 'means' something completely different. Donald Davidson had to suggest a 'principle of charity' – the presumption that other people are doing more or less what they seem to be doing, and are, more or less, like us in their desires and actions – in order to reduce the scope for these kinds of uncertainty. Tomorrow, might we might wake up to find that, ever since we learned how to speak, we have been talking at cross purposes? I don't think this can be said, and I have suggested some reasons why in this paper. However, these reasons (which can hardly be called comforting) came to me only after spending many years working in the oil industry in Aberdeen, and many mornings when I awoke with almost exactly that gloomy suspicion.

References:

Blackburn S, and Simmons, K, (eds) *Truth*, Oxford Readings in Philosophy (Oxford University Press, 1999)

Davidson, D 'Truth and Meaning', *Synthese*, 17 (1967), 304-323

Horwich, P. *Truth* 2nd edition (Oxford University Press, 1998)

James, W 'Pragmatism's Conception of Truth', in *Pragmatism: A New Name for some Old Ways of Thinking* (Longans, 1907), 197-236

Tarski, A 'The Semantic Concept of Truth', *Philosophy and Phenomenological Research*, 4 (1944)

Tarski, A, *Logic, Semantics and Metamathematics*, Woodjer, J. H. (trans.) (Clarendon, Oxford, 1983)

Wittgenstein, L W, *Philosophical Investigations* (19XX)