

PART I

THE ENTERPRISE OF MORAL
PHILOSOPHY

I

PHILOSOPHY OF LANGUAGE IN ETHICS

1.1. ETHICS, or moral philosophy, is the point at which philosophers come closest to practical issues in morals and politics. It thus provides a major part of the practical justification for doing philosophy (H 1971c: 98). If, therefore, philosophy of language can be shown to have a crucial contribution to make to ethics, this greatly increases the practical relevance of the discipline. But it is very important to be clear about what the contribution is.

The following programme looks promising at first sight. Philosophy of language is concerned above all with the study of the concept of *meaning* in the various senses of that word. But the meanings of moral words and sentences, in at least some senses, determine the logic of inferences in which they appear. So a study of the meanings of moral words or sentences, or of what people mean when they utter them, ought to enable us to investigate the logical properties of what they say, and thus decide whether what they say is self-consistent, what it implies, and in general which arguments (in the sense of reasonings) are good ones and which are not. So philosophy of language, applied to moral language, ought to be able to provide a logical structure for our moral thinking. And since our moral thinking often founders for lack of such a structure, that would be no small gain.

There are a great many pitfalls to be avoided in carrying out this programme; but I shall argue that it is in principle a feasible one. So let us first consider some possible objections to it. I shall be in danger of being misunderstood if I do not make clear at the start that philosophy of language is not the same as linguistic philosophy. The former is a branch of philosophy, co-ordinate with philosophy of science, philosophy of law, philosophy of history, etc. To say that a philosopher is doing philosophy of language does not presuppose that he is doing it by any particular method, or in accordance with the tenets of any particular school. Philosophers of language can be realists or the opposite, intuitionists or the opposite, and so on. If anybody were to say, like Plato on some interpretations, that words have meaning because they *stand for* eternally existing non-sensible entities up in Heaven, he would still be doing philosophy of language, but would obviously not be a linguistic philosopher. But see H 1982*a*, esp. ch. 4, for a more 'linguistic' interpretation of Plato.

A linguistic philosopher is someone who believes in a particular way of doing philosophy (*any* kind of philosophy, not just philosophy of language), namely that which consists in studying the meanings of words that present philosophical problems, and so unravelling the problems. He will advocate, like Carnap (1932), an 'Überwindung der Metaphysik durch logische Analyse der Sprache'. To make my own position clear, I am a linguistic philosopher of a sort, but not of such an extreme sort as Carnap. I believe that metaphysics does not have to be overcome, nor even superseded; as inherited from Aristotle, it is a respectable and central branch of philosophy, and only certain bogus impersonations of it are suspect. Ever since Aristotle and before, it has used linguistic methods. A great many problems which are called 'ontological' are in fact to be resolved by careful attention to the words which give rise to them; and this is true above all in ethics. But I regard this, not as a way of overcoming metaphysics, but as a way of doing it competently—of mastering it, if we may so mistranslate 'überwinden'; and I believe that this way of doing it has yielded results when practised by all the great metaphysicians up to the present day. So I am not against metaphysics—only against some wholly spurious 'philosophical' and 'theological' activities which

have in recent times usurped the name; they would be better called 'mephistics', because they are attempts, like that of Mephistopheles in *Faust*, to get philosophers to sell their souls for fantasies.

I wish to consider two possible objections to the programme I projected at the beginning. The first says, 'Facts about particular languages, including facts about how people use words in particular cultures, are contingent facts. They therefore cannot be used to establish necessary truths such as we are looking for in ethics. We do not want to be told how particular people or cultures use the moral words; we want to be shown what *is* right or wrong, and to be shown by secure reasoning that this is necessarily the case.'

The second objection is related to the first: it says, 'Moral reasoning has to be concerned with moral facts, which are facts not about words but about the world—facts about the existence of moral values in the world. The study of words could never yield such facts.' Answers to both these objections can be given. For the first, consider the position of ordinary logic. It would be a mistake to suppose that logic discovers only contingent truths about language; but it is also a mistake to think that logic is independent of the study of language. It is a necessary truth that, in one common meaning of 'all' and the other words used, if all the books on the top shelf are by Wittgenstein, and this is a book on the top shelf, then this is by Wittgenstein. But in order to establish that this is a necessary truth, we have to be assured that the words are being used and understood in the senses that make it so. Logic is, at least in part, the study of the words which people use in their discourse, to ascertain which of the things they say are, as they use the words, necessary truths.

This does not make the truths of logic contingent. It is of course a contingent fact that people do use certain sounds with certain senses. But to ask in what senses they use them is to ask according to what rules or conventions, logical and semantical, they use them. And it is not a contingent fact, but a tautology, that anyone who is using the words in those senses will be committing logical errors if he does not observe those rules. To take the same example: it is a contingent fact that someone is using 'all' in the sense that he is. But it is not a contingent fact that, if he is using it in *that* sense (namely the sense in

which the above hypothetical is necessarily true), the hypothetical is necessarily true. What makes the sense *that* sense is that it is the sense which makes that hypothetical necessarily true.

1.2. Words, including words like 'all', have their meaning determined by the conventions according to which we use them. And the conventions are in part logical ones, which determine what implies what, what we can consistently say, etc. One is not being a conventionalist in any bad sense if one states the obvious truth that studying what the conventions are for the use of words like 'all' (i.e. what logical rules they are governed by, as people use them) is the basis of the discovery of these logical rules.

To this it may be objected that people do not *have* to use words in accordance with those rules. Humpty Dumpty was quite right (Carroll 1872: 196). 'All' *could* have meant the same as 'some' does now—which is to say that the rules which determine its meaning and the implications of propositions containing it might have been different, and like those which now determine the meaning and implications of 'some'. And Englishmen, Frenchmen, Germans, and Chinese use different sounds to express the same things. And the inventors of artificial languages like Carnap have a considerable liberty to invent new uses of words and symbols, and to invent, *pari passu* with this, new rules and conventions for their use. Here too, however, it has to be said that *if* a word is being used in any language (natural or artificial) to express the same meaning as a word in some other language, it is bound by the same logical rules. If it were bound by different rules, it would not express the same meaning. A word in Chinese is not the equivalent of 'all' unless, when used in the corresponding Chinese hypothetical about Wittgenstein, it makes it necessarily true.

So, if logic as a whole involves the study of words in this way, the same will be true of that branch of logic called theoretical ethics. I call theoretical ethics a branch of logic because its principal aim is the discovery of ways of determining what arguments about moral questions are good ones, or how to tell sound from unsound reasoning in this area. It is, in particular, a branch of modal logic. 'Ought', which we may take as the simplest example of a word used typically in moral discourse (a moral word, for short), expresses a deontic

modality, and this is shown by the fact that deontic logics can be systematized which are in all or nearly all respects analogous to the other kinds of modal logic (Prior 1955: III. i. 6). The same is even more clearly true of the word 'must': its use to express moral statements like 'I must not tell her a lie' is analogous in most ways to its use to express alethic modal statements.

If, as is beginning to happen, viable systems of deontic logic can be discovered which are adequate models of ordinary moral language, they will do as much for the understanding of moral arguments as ordinary logic does for the understanding of other arguments. So, although it is of course a contingent fact that English uses 'ought' to express the meaning that it does, it is not contingent that *any* language that has an equivalent sentence—i.e. a way of expressing the same thought—will be bound by the same rules of reasoning. And what the rules are, as the word is normally used, is discovered by asking how it is normally used.

As before, we do not *have* to use it in that way. But when we are arguing about moral problems we are arguing about whether to accept or reject certain moral judgements. Clearly, whether an argument is a good argument for accepting or rejecting a certain judgement will depend on what the judgement is. But *what* it is depends on what the words used in expressing it are being understood to mean. If they were being understood to mean something different, it would be a different judgement. But once we are committed to discussing whether to accept or reject *that* judgement (i.e. the judgement which those words express when they are taken in *that* way) we are committed to following the rules of reasoning which that way of taking them determines. To take the words in that way is to accept that the judgement (with or without additional premisses) logically implies such and such other judgements, is inconsistent with such and such other judgements, and so on. So the sense of the words, as before, determines which arguments about the questions we are asking are sound ones. Therefore, in order to determine whether they are sound, we have to examine the senses of the words, i.e. the rules for their use in arguments.

We can of course, as before, use words as we wish. But if we decide

to use words differently from how we were using them when we posed our original problem, we shall no longer be posing the same problem. We are free to pose different problems; and that is what we shall be doing if the words mean something different. To revert to our original example: if what we had been asking had been, not whether all the books were by Wittgenstein, but whether some of them were, it would not have been a reason for answering 'No' that one of the books was not by Wittgenstein. So if, when we said 'all', we had been using the word in the same sense as 'some' usually has, the reasoning we should have had to use in answering our question would have been different. In the same way, if 'ought' means to us what it does when we are asking our moral questions, we shall have in our moral reasoning to follow the rules (of implication, consistency, etc.) determined by *that* meaning of the word (by the fact that it is *that* question we are asking, and not a different question which would be asked by someone who uttered the same sounds but was using 'ought' in another sense). It is therefore in order, if we wish to determine what rules we have to follow, to ask in what sense the word was being used in our question. Indeed, to ask in what sense it is being used *is* to ask what the rules are.

All this is peculiarly true of words like 'ought', one of the most general terms used in asking moral questions. Such words, like other modal words, express *formal* concepts, in the sense that the rules for their use are exhausted by the implications and other logical properties that they give to the propositions containing them. This is not true of all words: for example, the formal logical properties of the words 'blue' and 'red' are the same, but 'red' does not mean the same as 'blue'. So their formal logical properties cannot exhaust their meaning. But if 'ought' is a purely formal word, then we should be able to discover all there is to be known about its meaning and the rules for its use by studying its logical properties. If true, this is, as we shall see, of fundamental importance for ethics. It means that, although in a sense it has semantical properties as well (its 'descriptive meaning'), these are not part of its meaning in the narrow sense (H 1986c), and do not affect at all profoundly the rules for reasoning about what we ought to do.

The answer to the second objection mentioned at the beginning is thus that, because the concepts studied by ethics are formal, there do not have to be moral facts in the world in order for us to develop a theory of moral reasoning, any more than there have to be logical facts to substantiate logical reasoning. The necessities which constrain our reasoning are formal necessities—which does not mean, any more than it does in logic and mathematics, that they cannot *in conjunction with* substantial non-moral information about the world, help us in deciding moral questions of substance. How this is to be done, we shall see later.

1.3. It is now time to ask how we can discover what these formal properties are. The first step requires us to anatomize language as a whole in order to see where in the anatomy such words as ‘ought’ belong. The most perspicuous way of doing this is by speech act theory. The term ‘speech act’ was brought into currency by J. L. Austin (1962: 41, 149), though he does not himself use the term very much, preferring more specific expressions. He can justly be regarded as the founder of speech act theory; but the idea that not all speech acts are of the same kind or obey the same rules has been used before and after him by Wittgenstein, Ryle, Searle, Habermas, and many others. In order to divide off speech acts of different kinds from one another, we need to articulate the sentences that are used to perform them. The main purpose of this is, if possible, to isolate the features of sentences which perform the various functions necessary for a complete speech act. Then we can see which features of a sentence are peculiar to a particular kind of speech act, and so mark the utterance of it as a performance of that kind of speech act; and which features are common to a number of different kinds of speech act. The best known marker of this sort is the sign of mood (e.g. indicative or imperative) which (to speak generally at first) marks off statements from imperations (if we may use that expression for speech acts typically expressed in the imperative).

We also need to be clear that the division of speech acts into kinds takes the form of a tree with genera, species, sub-species, etc. It cannot be assumed, for example, that there are no further subdivisions within the classes of statements and imperations, nor that

imperations may not belong, perhaps with moral judgements, within a larger class of prescriptions. Nor can it be assumed that a kind of speech act has to belong to one or other of these classes and cannot belong to more than one. The species and genera may not be mutually exclusive: perhaps moral judgements share some of the properties both of statements and of prescriptions. All this has to be investigated by the study of speech and language (I use these words to mark the distinction made famous by Saussure 1916).

A further necessary clarification can conveniently be made at this point. Austin used the term 'illocutionary force' to connote the property which distinguishes one speech act from another. Thus the statement that you are going to shut the door has a different illocutionary force from the command that you shut the door. But different writers since Austin have interpreted this distinction in different ways. Consider the two commands, that you open the door, and that you shut the window. Do these have the same illocutionary force, in that they are both commands, or different illocutionary forces, because they are different commands? It will make no difference to any argument, provided that we are clear about our use of the terms; but in what follows I shall myself adopt the second of these uses. I shall speak of these two commands as having two different illocutionary forces, though they belong to the same *type* of illocutionary force, namely the imperative. Similarly I can make two different statements, which have different illocutionary forces because their content is different, but have the same type of illocutionary force, namely what Austin called the constative (1962: 6 n.). This will be brought out if the sentences are articulated in such a way (as they are in most languages) as to distinguish the feature which marks the mood from the rest; the two commands 'Open the door' and 'Shut the window' share this feature, by which we recognize them as imperatives; but otherwise they differ.

The articulation of sentences, or the speech acts that they express, has to distinguish at least four functions (H 1989a). The first is the mood, already mentioned. I shall call the sign of mood the *tropic*. That mood is, or can be, part of meaning is evident from the fact that the Latin expressions '*i*' and '*ibis*' ('Go' and 'You are going to go') have dif-

ferent logical implications (H 1996*b*): the latter implies that you are going to leave this place; the former does not, because a command is not a prediction of its own fulfilment. Next, we have to distinguish the content of the speech act (for example what in particular is being stated to be the case, or commanded to be made the case). Thus the commands 'Open the door' and 'Shut the window' have the same tropic but different *phrastics* (using that term to denote the feature of the sentence, not necessarily a separate part of it, that indicates what is being e.g. stated or commanded). In a completely and perspicuously articulated language these functions would be assigned to different parts of the sentence.

The remaining two functions, which do not need to be discussed here, are those which would be expressed in a fully articulated language by the *clistic*, or sign of completeness, of the sentence, and the *neustic* or sign of subscription to a speech act by a speaker or writer. These signs are controversial, and many writers have denied the necessity of the latter in particular; but I shall not need to defend them for the purposes of the present argument (see H 1989*a*). Nevertheless, it is very important to distinguish between these different functions, as many writers (including myself in early days) have not (H 1971*c*: 21 ff.). In particular the tropic or mood-sign has to be distinguished from the neustic or sign of subscription, because one can mention or embed an indicative or imperative sentence, including its mood-sign, or use it mimetically (6.4, H 1989*a*)—e.g. on the stage—without making a statement or giving a command.

It will be asked at this point whether mood, as I am using the word, is a logical or a grammatical term. The answer is that it is both, but that we have to understand the difference between what are now often called surface and deep grammar, and used to be called grammatical form and logical form. If there is a difference between these two ways of making the distinction, it will not affect what I am now going to say. In history, grammar and logic grew up together, and metaphysics with them; and it has proved difficult to draw clear distinctions between these three. Even such diverse thinkers as Hegel and Carnap found it hard to distinguish between logic and metaphysics (Hegel assimilating the former to the latter, and Carnap, in

effect, the reverse—though he reserved the name ‘metaphysics’ for what I have called ‘mephistics’). And similarly deep grammar and logic are so intimately bound up with each other that it would be foolish to try to prise them apart. The difference between logic and surface grammar is what has made people think that there is a difference between grammar and logic as a whole.

There are indeed grammatical distinctions that have no logical significance, like that between strong and weak forms of the past tense (3.3). But mood is not like this; the distinction between the mood-sign and the rest of a sentence is as important logically as that between subject and predicate. These two have been both grammatical and logical terms, and rightly, because the grammar is a way of expressing the logic. In order to speak grammatically we have to be able to make, at any rate implicitly, the logical distinction; and when structural linguists construct their ‘trees’ (which in my school days was called ‘parsing’), they are using the logical distinction in order to mark off noun-phrases from verb-phrases.

There are complications here into which I shall not be able to go—for example, the false thesis held by many, including Aristotle (*An. Pr.* 43^a30), that there are terms which can occupy either subject or predicate places in propositions at will. The truth is that in ‘Red is a primary colour’ and ‘The book is red’, the word ‘red’ means different things, as is shown by the fact that we could rewrite the first sentence ‘The colour red is a primary colour’, but could not rewrite the second ‘The book is the colour red’. Similarly, in ‘Callias is a man’ we can substitute ‘human’ for ‘man’; but in ‘Man is an animal’ we cannot. As we have seen, if we alter the mood of a sentence, then by making the grammatical change we alter both its meaning and its logical properties; and this is enough to show that mood is both a logical and a grammatical category, without in this context distinguishing the two functions.

1.4. It is time to turn back to the question of what place moral judgements occupy in the anatomy of language, presuming that we have an adequate one. If it is adequate, it will at least distinguish between two genera of speech acts that I shall call the descriptive and

the prescriptive (I.6). All kinds of ordinary statements will belong to the former, and all speech acts which are typically expressed in the imperative to the latter. We must not presuppose that nothing except imperations belongs to the latter genus. We must not even presuppose that in order to give a command it is necessary to use the imperative. But let us now ask in a preliminary way whether moral judgements (for example those expressed with 'ought') are prescriptive or descriptive speech acts. The answer is that they are both, but that the distinction needs to be carefully preserved, because otherwise we shall not be able to understand the *different* features of 'ought'-sentences which link them to the two genera.

'Ought'-judgements are prescriptive, and in this respect like imperations, because in their typical uses agreement with them, if genuine, requires action in conformity with them, in situations where the action required is an action of the person agreeing. I deliberately say 'in their typical uses', because, as is well known, there are other uses, which have generated a vast literature. Such are uses by the weak-willed person, 'acritic' or 'backslider' who does not do what he agrees he ought to, because he very much wants not to (H 1963a: ch. 5, 1992e: ii. 1304), and by the 'satanist' who does what he agrees he ought not to, just because it is what he ought not to (H 1992d: 98). This is not the place to add to this literature; the point here is just that typical and central uses of 'ought' require compliance if they are to count as sincere. By contrast, constative speech acts require only accordant belief.

However, moral judgements are not just like ordinary imperations. They share with constative speech acts a very important feature, namely that when I say 'I ought to do that', I have to say it because of *something about* the act that I say I ought to do. This is a feature of all uses of 'ought', and not just of moral uses. It is true that imperations too are normally issued for reasons. But they do not have to be. If a drill serjeant is trying to see whether a new recruit will obey him, he may say to him 'Right turn', and may have no reason at all for saying this rather than 'Left turn'. But with 'ought' it is different. To take a non-moral example: suppose that instead they are doing a tactical

exercise and the instructor says 'You ought to attack on the right'. There has to be a reason in the facts of the situation why they ought to attack on the right rather than on the left (*FR* 3.3).

It is hard for Germans to appreciate this point, because the German word 'soll' can be used to translate both the English 'is to' (which can be equivalent to an imperative), and the English 'ought to' (which is a moral or other normative expression). Systems of deontic logic have sometimes been set up which fail to make this distinction, using a single symbol (for example 'O') for both 'ought' and the imperative. Since the logical behaviour of these is different (for example a 'square of opposition' which works for 'ought' does not work for imperatives—H 1967*d*), such systems start on the wrong foot. Confusion on this point can sometimes lead to treating the fact that one is commanded to do something (one *is to* or *soll* act in a certain way) as showing that one *ought* to act in that way. This can have grave political consequences (H 1955*b*).

Because moral judgements have to be made for reasons, the reasons being the facts of the situation, it is irrational to issue one having no regard for the facts (contrast the serjeant's command in the above example, which in no way convicts the serjeant of irrationality). It is indeed true that the choices expressed by imperative speech acts are normally required to be made for reasons if the chooser is not to be called irrational (H 1979*a*), and that even in this unusual case the serjeant *has* a reason for saying what he says (namely the intention to test the obedience of the recruit). But in this case he could have said 'Left turn' instead of 'Right turn' with equal rationality. It is the privilege of serjeants not to have reasons for this kind of choice.

Moral and other normative judgements by contrast cannot be arbitrary in this way. They have to be made because of the facts. This does not mean that the moral judgement *follows logically from* the facts (H 1963*b*: sec. 8). The facts do not *force* us logically to make one moral judgement rather than another; but, if we make one about one situation, we cannot, while admitting that the facts are the same in another situation, in the same breath make a conflicting one about the second situation. In the non-moral tactical example just used, the

officer could not say that there might be another tactical situation just like this one in which they ought to attack on the left rather than on the right. If the facts are just the same, they would supply a reason for making the same normative judgement. This is the basis for the feature of normative judgements called *universalizability* (H 1963a: ch. 2), and moral judgements share this feature (I.7).

1.5. Before assigning to moral judgements their place on this anatomy, there is an important distinction to be made, which in spite of a very clear statement of it by Austin (1962: chs. 9, 10), is still neglected by many, especially in connection with imperatives. It is encouraged by a too easy use of the term 'pragmatics' (6.5), and of the Wittgensteinian linking of meaning to use, by those who are not very clear about what exactly they mean by 'use' (see Austin 1962: 104). Austin distinguished between illocutionary and perlocutionary acts (6.4), the first being what we are doing *in* saying something (*in locutione*), and the second what we are doing or seeking to do *by* saying something (*per locutionem*). The 'pragmatics' and the 'use' of utterances are easily taken to mean the latter, especially in the case of imperatives; and so people slip into thinking that their meaning can be fully explained by giving their pragmatics or use, understanding by this their intended perlocutionary effect.

Besides the temptation just mentioned, there are others. Many logicians still hold the view, in spite of Austin and Wittgenstein, that there is only one kind of language-game or speech act that is respectable enough to be worthy of their attention, namely the constative. They sometimes cite Aristotle in their support (*De Int.* 16^b33 ff.). Others are so attached to truth-table and similar methods for setting up a logic that they cannot see how one could be set up that dealt with anything but true-or-false propositions. Others wish to define 'valid inference' as 'inference of such a form that no inference of that form can have true premisses and a false conclusion'.

Such writers exhibit the same sort of prejudice as has been in evidence in connection with the truth-condition theory of meaning. But there are many other ways of setting up logics, in particular that which starts from the notion of inconsistency. If we knew how to tell which speech acts were inconsistent with which, we could construct

a logic for those kinds of speech acts. And imperations can certainly be inconsistent with one another (for example 'Shut the door' and 'Do not shut the door'). The inconsistency lies here within what I have called the phrastic, which the imperative shares with its corresponding indicative; so the source of inconsistency is the same for both, and therefore so is the nature of the logical fault. In this case, though not always (*LM* 2.3, Searle and Vanderveken 1985: 152), the sign of negation is part of the phrastic. But there is nothing here to make us banish imperative speech acts from logic. Indeed, the rules of logic itself, for example formation rules and rules of inference, are imperations, and *they* have to be consistent.

But the greatest temptation to this way of thinking about imperations (that they have only pragmatics and no logic) is a confusion between illocutionary and perlocutionary acts. Here it is necessary to depart from Austin's view. He distinguished between *three*, not just two, kinds of act, the third being the locutionary (Austin 1962: 108). But if he thought that only the locutionary act had meaning—and I have argued elsewhere that this is a misinterpretation (H 1971c: 115 ff.)—he was clearly wrong; for, as we have seen, mood is part of meaning ('Go' and 'You are going to go' do not mean the same). Therefore, in order to understand what somebody meant, we have to know what mood his speech act was in. And this is to know something about its illocutionary force. It is therefore incoherent to say, both that locutionary acts are the sole repositories of meaning, and that one can specify the locutionary act without mentioning its illocutionary force. Meaning is, in part, illocutionary act potential (Alston 1964: 37 ff.). This does not necessarily imply that other elements in the illocutionary force cannot extend beyond the locutionary act as specified. It has been alleged, for example, that we could know what a person meant when he said 'The ice is thin', and thus know what locutionary act he performed, without knowing whether he intended it with the illocutionary force of a warning or a mere statement of fact. I would dispute this, but it would need too long an excursus into such notions as warning to settle the matter. I deal with 'warn' briefly in 3.3. It can at any rate be granted that, as Austin (1962: 32, 69) pointed out, there are often ways of making the illocu-

tionary force of our utterances explicit and thus disambiguating the sentence. We can do this by saying either 'I warn you that the ice is thin' or 'I affirm that the ice is thin'.

Be that as it may, the locutionary and illocutionary acts lie together on the other side of an important divide from the perlocutionary. For perlocutionary acts there can be no logic in a strict sense. The reason is that, as we have seen, logic is determined by the rules or conventions for the use of words, and perlocutionary acts (what we are doing or trying to do by saying things) need not be controlled by any rules or conventions of a logical sort (cf. Austin 1962: 118). It is true that what we can do by saying something depends on what the something is—i.e. on what we are doing in saying it—but it depends on much else; we have to size up the situation and think what would be the likely effects of certain utterances. Telling someone that the ice is thin may be a way of getting him not to go on the ice; but if he is a daredevil who does not fear cold water it may be a way of getting him to go on it. If he is a normal person who trusts us, it may be a way of getting him to believe that it is thin; if he is untrusting or countersuggestible, it may be a way of getting him to believe that it is not thin. And similarly with imperatives. Say 'Go on the ice' to a trusting child, and he may go; but say it to an untrusting or rebellious one, and it may make him do the opposite. Thus the same illocutionary act with the same meaning may have different perlocutionary effects, and this in itself shows that the perlocutionary effect or intended effect is not part of the meaning.

What may be called the 'verbal shove' theory of the meaning of imperatives has therefore to be rejected (*LM* I.7, H 1971c: 91 ff., 6.3). If 'pragmatics' is taken confusedly to cover both illocutionary and perlocutionary acts, we can say that to study the meaning of imperatives is to study their pragmatics; but only the illocutionary part of their pragmatics at the most. If we stray beyond this, we are no longer studying their meaning at all. Once we realize this, we shall not include as imperations speech acts which are clearly statements, such as 'There is dust on the table' said by a demanding lady to her housemaid. It has been alleged that this is really an imperation, because it is intended to *get* the housemaid to dust the table. It may indeed be so

intended; but that does not make it an imperation. It is a statement, which, in conjunction with an assumed standing order of the house (which *is* an imperation) that when tables are dusty she is to dust them, entitles the housemaid to infer the imperation that she is to dust the table. So, if the housemaid is both logical and obedient, saying this will get her to dust the table. But she has understood the meaning of the utterance perfectly well even if she is not obedient, and even if she has not heard of the standing order, and even if she is too stupid to think that there might be one. If she is stupid enough, she may not dust the table even if the tone of her mistress is menacing. She will not know what to do, because she has not been told that.

1.6. The relevance of all this to ethics is this. Moral judgements are, in a sense to be explained later, prescriptive, and therefore akin in some respects to imperations. The school of moral philosophers called *emotivists* (further discussed in Chapter 6) realized this. But, infected with the confusion about pragmatics that I have just been exposing, they were led into the error of thinking that the meaning of moral judgements had to be explained in terms of their perlocutionary effect (Urmson 1968: 29 ff.). This is evident from the title of the part of Stevenson's *Ethics and Language*, 'Pragmatic Aspects of Meaning' (1945: 37), which sets the tone for the whole book. But the same thought is to be found in Ayer (1936: ch. 6), and seems to be implicit in Carnap (1935: 23). It led people to look for the source of the meaning of imperations, and therefore of part of that of moral judgements, in their power of *getting* people to do things. But the perlocutionary act of getting them to do something is a quite different thing from the illocutionary act of telling them to do it (H 1951a). As we have seen, the latter may be a means of achieving the former; but this does not make them the same act in the sense relevant here. In particular, the illocutionary act of *telling to* is subject to logical control, just like the illocutionary act of *telling that*. In telling to, one must not contradict oneself, any more than in telling that; otherwise one is not telling people to do anything that they can do. But in getting to, including getting to believe that, one may contradict oneself if that is the most effective way of doing it.

The emotivists thus confused the essentially irrational or arational

perlocutionary act with the logic-governed illocutionary act. (6.3 f.) So they not only thought without good reason that there could be no logic of imperations, but, because of this confusion, tainted moral judgements with the same irrationality. I have even heard it argued that, because moral judgements are material for rational thought and imperatives are not, moral judgements cannot be imperatives. But the boot is on the other foot. Because imperations have to obey logical rules, the fact that moral judgements share some of their properties is no obstacle at all to the rationality of moral thinking. Therefore rejections of non-descriptivist ethical theories by aspiring rationalists on the ground that moral judgements could not be rational unless they were statements in the narrow sense—or constative, to use Austin's term (1962: 6 n.)—miss the point entirely. It can be allowed that in certain senses moral judgements can be called true or false (H 1976*b*); but even if they could not, their rationality would not be impugned. We shall see later that the prescriptivity of moral judgements, so far from being a bar to their rationality, is a vital ingredient in it (1.8).

But before showing this, it is time to ask in what sense moral judgements are prescriptive, and how their prescriptivity combines with their other features. And this cannot be clarified until we have explained what prescriptivity is. We have already used the word to describe the genus of speech acts to which imperations belong; they are the paradigm of it. The simplest way of characterizing this genus is to say that a speech act is prescriptive if someone who assents to it is not being sincere if he does not act accordingly (i.e. at the time and in the way specified), when he is the person whom it charges with fulfilling it, and is physically and psychologically able to do so (*LM* 2.2). But there are some ambiguities here which need to be unravelled. Expressions like 'the subject' and 'the addressee' (of an imperation) can mean three different things. They may denote the person to whom an imperation is spoken or written. Or they may denote the person or thing to which the grammatical subject of the sentence used refers. Or they may mean the person charged with complying with the imperation. These may all be different persons or things. If the *grande dame* in our previous example says to her butler 'The table

is to be dusted', the grammatical subject refers to the table; the person spoken to is the butler; and the person who is charged with complying is the housemaid (butlers do not dust tables).

In the present context it is the person charged that interests us. Let us call her, not the addressee or the subject, but the chargee. A prescriptive speech act is one such that, if I am the chargee, and I assent to the speech act, I cannot be assenting sincerely if I do not act accordingly. For example, if the above command is addressed to the housemaid, who knows that she is the person charged with dusting tables when they are to be dusted, and she assents by saying 'Very good, madam', she is not assenting sincerely if, though she could dust the table, she at once slinks off to bed without dusting it.

1.7. Are moral judgements prescriptive in this sense? Certainly not all are. The housemaid can assent to the judgement (even taken in a moral sense) that she ought to dust the table, and still slink off to bed. The question is rather, 'Is there an important class of moral judgements which *is* prescriptive, and if so what is the relation between those that are and those that are not?' It can be argued (but not here) that Plato (see H 1982a: 56, 66), Aristotle (*Eth. Nic.* 1143^a8, 1147^a25 ff.), Hume (1739: III. 1. 1), Kant (*Gr BA*36 f. = 412 f.) and Mill (1843: last chapter) all thought that moral judgements were typically prescriptive, though probably none of these thought that all were, nor that this exhausted their meaning, any more than I do (H 1998a). I have argued elsewhere that there is a prescriptive use of moral judgements, and that this is central in two senses. The first is that, if this use is explained, the others can be explained in terms of it and fall into place (*LM* ch. 11). The second is that, as I shall be saying later, their prescriptivity is a vital ingredient in moral reasoning (1.7: *MT* 6.1).

It was his recognition, inherited from Socrates and Plato, that moral and other normative judgements are prescriptive, that made *acrasia* or weakness of will a problem for Aristotle. If they are prescriptive, how could the housemaid assent to one and then slink off to bed? If Aristotle had been a pure descriptivist, as some of his pretended modern followers seem to themselves to be, there would have been no problem for him in the housemaid's backsliding. He devotes

half a book to resolving the Socratic problem (*Eth. Nic.* 1145^b21 ff.), because he, like Socrates, has to explain how one can accept a prescription and then not act on it. His explanation, though not completely adequate, is more subtle than that of Socrates. It consists in pointing out that the prescription in question is universal (the housemaid knows she ought to dust the table because she knows the universal rule of the household, and all households that are well ordered, that dusty tables ought to be dusted, and knows the particular fact that this table is dusty). Though his example is different, Aristotle could say that she can backslide from the universal rule because she is tired and wants to go to bed, and therefore ignores the particular fact, even though it is evident enough. This summary does not do justice to the subtlety of his solution to the problem, and I have myself suggested a more complex solution (*FR* ch. 5, H 1992e: ii. 1304). But the important thing is that there is a problem, which there would not be if we were descriptivists.

If, in spite of this alleged difficulty, we recognize that central cases of moral judgements are prescriptive, we still have to recognize also that they are not *purely* prescriptive. That indeed is the major part of the more complex solution to the problem of *acrasia*. As Aristotle (*Eth. Nic.* 1147^a31) and Kant (*Gr* sec. 2, para. 31) both saw, moral judgements are not merely prescriptive but *universally* prescriptive. And the universality of the moral prescription easily introduces a non-prescriptive element into its meaning. To explain this: if the housemaid accepts the universal rule that dusty tables ought to be dusted, this rule will assume for her (obedient girl as she normally is) the status of fact. That is, if ever she is tempted (as now) to neglect her duty, she will not be able to avoid thinking of the possibility that her mistress or the butler will notice the omission and punish her; and, if they do, *that* is a real enough fact. And so is the fact that she is frightened by the thought. Some people's attitude to morality is like that of the housemaid to the butler. Even when the housemaid has left that (or all) employment and has a house of her own—even when there is no longer a *grande dame* and the butler is out of work—she will not be able to escape the feeling of guilt caused in her by the sight of a dusty table for which she is responsible.

It is easy for the irreligious to proceed from this analogy to the thought that God does not exist, and that therefore everything is permitted. They should reflect on two things. The first is that, God or no God, the attitudes that make us revere the laws of morality are a social necessity; we could not live in communities without them. Kant may have carried this reverence to excess, and his moral law was no doubt too simple and rigid. But society would collapse unless children were brought up to feel bad when they do bad things; and we should not let psychologists convince us otherwise without empirical evidence to the contrary. The second is that a reflective critical morality can *justify* these laws or rules or principles and our attitudes to them. So even if there were no *grande dame* we would have to invent her. Critical moral thinking can also amend the principles if they are seen to be unsuited to our situation (*MT* 3.3).

The inescapable factuality or descriptivity of moral principles has a logical as well as a psychological basis (*MT* ch. 2). Moral judgements are like factual statements in many respects (on the face of it, they resemble each other more than either of them resembles imperations). It is easy, therefore, to think that they are like them in all respects. It is made easier still by the existence of a large class of moral judgements, referred to above, which are not prescriptive. The similarity is so great that I have thought it right to follow Stevenson (1945: 62 ff.) in using the term 'descriptive meaning' for the element in the meaning of moral judgements that makes them like constative speech acts. This is not the same as the *phrastic* referred to above (1.3); that is something else, which would indeed be part of moral judgements even if they were plain imperations, which they are not.

The element I am calling descriptive meaning can best be indicated by a non-moral example borrowed from Urmson (1968: 133). If you are meeting a girl at the station and do not know her by sight, I may enable you to recognize her by saying, among other things, that she has a good figure. To say this is to describe her, and my purpose has nothing to do with prescribing the acquisition of such figures. We all know what in our society counts as a good figure, so you will know what to look for. If your informant were a member of a society that thought fat girls more attractive, you would look for a different sort of

figure. Thus the descriptive meaning of 'good figure' is different in the two societies.

Because the standards or criteria for commendation vary from society to society and from century to century, whether we are speaking of moral or of other kinds of commendation, the descriptive meaning of words like 'good', 'right', 'wrong', and 'ought' can be relied on only within a certain circle; but within that circle it is reliable enough. Other evaluative and normative words have their descriptive meaning so firmly tied to them that it is hard to use them in communication between different societies; so that, if we were confined to the latter class of words (for example 'blasphemous' and 'cruel'), we might not be able to talk about values to those who did not substantially share our own values. We should have to fight one another. It is the existence of shared general value-words like 'ought' that makes peaceful discussion between cultures possible (H 1986c, 1993g, 6.9).

Moral judgements acquire a descriptive meaning, even without butlers to enforce them, because of an important logical feature that they share with other value judgements, called *universalizability* (FR 2.2). One way of approaching this is to say that all such judgements are made for reasons: that is, because of *something about* the subject of the judgement. The girl's figure could not be good if it were not good because of something about her measurements. A man cannot be a good man, if not because of the sort of man he is. An act cannot be wrong, if not because of something about it. They cannot be good or wrong just because they are good or wrong; there must be properties other than their goodness or wrongness which make them so. This feature of value judgements is sometimes called 'supervenience'. Causal judgements have it too: if an event causes another, there could not be a qualitatively identical situation in which the corresponding events were not conjoined and causally linked. This is the basis of the so-called 'covering law' theory of causal explanation (Hempel 1965: 345 ff.). And the notion has other applications too. But moral philosophers should not be misled by philosophers of mind and others who have borrowed the word and used it in another meaning which they have not made clear (H 1984b).

That moral properties supervene on non-moral properties means

simply that acts, etc., have the moral properties because they have the non-moral properties ('It is wrong because it was an act of inflicting pain for fun'), although the moral property is not the same property as the non-moral property, nor even entailed by it. Someone who said that it was an act of inflicting pain for fun but not wrong would not be contradicting himself, though most of us would call him immoral. Logic does not forbid the adoption of different moral standards by different people; it simply prohibits a single person from adopting inconsistent standards at the same time, and says that they *will* be inconsistent if he says conflicting things about situations which he agrees to be identical in their universal properties.

1.8. It has been disputed whether the universalizability of moral judgements is a logical feature of them, or embodies a substantial moral principle. A ground for holding the former view is that we react to breaches of the principle in the same sort of way as to breaches of logical principles. If someone says that there are two situations identical in all their universal non-moral properties, but says he thinks that the protagonist in one ought to tell a lie, but the protagonist in the other ought not, we are likely to be as nonplussed as if he had said that he thought that a rotating disc was both stationary and not stationary (cf. Plato, *Rep.* 436d). In either case there could be an explanation. In the second he might mean that the axis of rotation was stationary, so that the disc continued to occupy the same region of space, but that within this region it moved around its axis.

In the first case there could be many explanations, but none of them would impugn the universalizability thesis. The protagonists in the two cases might themselves have different characteristics. But when the thesis speaks of identical situations, it must be understood as ruling out this difference too. Another possibility is that in one case the person to be lied to is the mother of the protagonist, and in the other not. One can only have one (genetic) mother, and it might be thought that this makes a difference, because to tell lies to one's mother is worse than if someone else tells them to a person (perhaps even the same person) who is not his mother, however similar the situations. But relations can be universal properties (5.8), and the relation *being the mother of* is one such. The situations are different in

respect of a universal relational property, because in one the liar and the person to be lied to are related as mother to child, and in the other not.

Examples like this force us to make clear what the thesis means by 'universal property'. A simple, but for our present purposes sufficient, definition is the following. A property is universal if, in order to specify it, it is not necessary to mention any individual (for an apparent exception, in which the expression referring to the individual is preceded by 'like' or its equivalent, see 5.8 and *FR* 2.2).

It is sometimes claimed that the thesis of universalizability is inconsistent with the principle of the identity of indiscernibles. For it claims that, if there were two situations identical in all their universal properties, the same moral judgements would have to be made about both; but the principle of the identity of indiscernibles holds that there cannot be two situations, numerically different, but identical in all their universal properties. However, it has been convincingly argued that in this extreme form the principle of the identity of indiscernibles is not true (e.g. Strawson 1959: 119). It is true in less extreme forms, e.g. if it claims only that things identical in all their universal properties *and* in their relations to individuals must be numerically identical; but this obviously causes no trouble to the universalizability thesis.

There is a further problem about whether being actual as opposed to merely possible or hypothetical is a universal property (*MT* 6.4). If it were, a form of special pleading would become possible in moral reasoning, by which an aggressor could claim that he would never be actually in the position of his victim, and that this difference was morally relevant. It is perhaps best to follow those (e.g. Lewis 1973: 85) who claim that the actual world cannot be distinguished from possible worlds without a reference to individuals, namely those who are actual; but not to follow them into thinking that possible worlds have some real existence in limbo. In any case it seems that making moral distinctions on the ground of actuality would be rejected on logical grounds as we use words like 'ought'. If someone said 'I ought in the actual case, but someone else ought not in an identical hypothetical case', we should not understand what moral principle he was

invoking, because a moral principle which applied to the actual case but not to hypothetical cases exactly like it would not be counted by us as a moral principle, whatever our substantial moral views were, nor as any other sort of normative principle. This problem has analogies with the old one of whether existence is a property.

Those who think that the universalizability thesis is a substantial moral principle and not a logical doctrine will by this time be getting restive. They will think that we have fixed the logic so as to enable us to reach substantial conclusions in moral arguments. We must ask them to be patient until we have explained how the arguments work. Until then we can only point out that we would object to the above conjunction of moral judgements about the actual and hypothetical cases even if we knew nothing whatever about the substantial moral opinions of the person who made them; so it cannot be anything substantial that we are objecting to. The objection must therefore be logical. Suppose, even, that he also says that on other grounds he believes in complete impartiality between people, himself and others. It is not inconsistent to believe in impartiality between people, and still try to call the difference between actual and hypothetical morally relevant; for if it were relevant it could be used impartially between people. So we cannot be introducing a substantial moral principle requiring impartiality between people by insisting that actuality is not a morally relevant feature. On the problem of moral relevance in general see H 1978*b*: 73, *MT* 3.9.

We have found reasons for thinking that the universalizability thesis is a logical and not a substantial moral doctrine. The main ground on which people have thought otherwise is that the thesis does seem to have implications of a substantial sort for moral arguments, and there is some suspicion of a conjuring trick—of producing a substantial moral rabbit out of a logical hat. Moral philosophers have so often attempted similar tricks that one is right to be suspicious. For example they have sought to attribute a certain meaning on logical or conceptual grounds to phrases such as 'human needs', and have then gone on to extract substantial moral principles from these definitions (4.6). How we can allay this suspicion will not be clear until we have set out more fully the argument from formal logical or philosophy-of-

language considerations to an account of substantial morality. Here we must simply note that formal considerations are only one element in moral arguments. Others are the facts about situations, which are substantial, and in particular facts about people's *wills*, to use Kant's word; and these facts too are substantial (8.5 f.).

Let us try out this essentially Kantian method more clearly, and relate it to its basis in philosophy of language. If moral judgements are prescriptive, as has been argued, then in making one, I am asking that it be acted on, and, if sincere, must will this. But if they are also universalizable, I am, in making one, implicitly making identical judgements for all situations identical in their universal properties, no matter what role particular individuals, including myself, occupy in them. The question of what moral prescriptions I am prepared to issue thus resolves itself into that of what I am prepared to will for all situations of a given kind, no matter what role I occupy. Thus to issue a moral prescription I must accept the consequences (even the hypothetical consequences) of its being obeyed whatever role I occupy.

How constrictive this is will depend on what I will should be done to myself, were I in those various roles. The roles include the fact that the wills of the people in them are what they are. If I were in those situations, my will would in each case be the same as the present will of the person who is now in it, since the willing is part of the situation. So the question resolves itself into that of what I now will (NB not what I *would* will) should be done to me in those situations, in which I willed what they now will.

But here another factor enters, also obtainable from the logic of our language. By an argument which does not invoke universalizability, we can see that I must have as much regard to what I would will in those situations, as I do to what I now will. For if I do not, I am either not representing the situations fully to myself, or else not thinking of the person in them as *myself* (7.3, *MT* 5.4). To think of him as myself is to identify myself with his will. This is part of what we *mean* by 'myself'. Reflection on the meaning of 'myself' should convince us of this. The case is analogous to what I think about *future* states of myself which I expect to be actual. If I know what I shall then will, and am really thinking of the person in the situation as *myself*, and

do not irrationally discount the future, my will must be as strongly engaged as that of the future person who will be me. If anybody doubts this, he should arrange for himself to be whipped, and reflect on his state of mind just before it happens (cf. Aristotle, *Eth. Nic.* IIII5^a24). Failure to engage my will in this way is always due either to a failure of representation of the situation of the person that I shall be, or to a failure to think of him (or her) as *myself*.

Since for moral argument hypothetical situations are as relevant as actual, I have to will that the same should be done to me in them too. They will include all the situations in which I would occupy the roles of those affected by proposed actions of mine. I am therefore faced with the problem of finding a universal prescription for situations like that which I am presently in, which I can accept equally for all the identical situations that I could be in, in different roles. This in effect gives equal weight in my moral thinking to the wills of all those affected by my actions. The Kingdom of Ends is not really a kingdom, but a democracy with equality before the law. But if all wills have equal weight in proportion to their strength (for obviously how strong they are must make a difference) then the problem of doing justice between all these wills is to be resolved by choosing the moral prescription which maximally realizes the fulfilment of them, treating all impartially and giving them weight according to their strength (H 1996c).

1.9. This development of Kant's ideas thus turns into a kind of rational-will utilitarianism (see Chapter 8). He is, admittedly, selective with regard to the kinds of will that he is prepared to enfranchise: they have to be rational; but many utilitarians accept this. This shows the superficiality of the commonly accepted dogma that Kant and the utilitarians need to be at odds. If the two doctrines are sympathetically formulated, they are in agreement. The disagreement remaining is one *within* utilitarianism, as to whether any kinds of will are to be excluded from consideration. And such a formulation involves the use of insights from the philosophy of language. There is no space here to develop these insights further, nor to deal with other objections and difficulties. This must be left until later, and to my writings on the philosophy of education (e.g. H 1992d), in order to show how

in practice we manage to find a level of moral thinking more suited to us humans than the somewhat demanding level in which Holy Wills can engage.

It was Kant's predicament in between these levels (dare we say his insufficient grasp of an important difference between the levels?) that led him to try to justify what are only simple, general, prima facie intuitive principles (suitable to our human condition) *directly* by appeal to the Categorical Imperative; and this notoriously got him into trouble (8.4). The right way to try to justify them would have been to show that a Holy Will (perhaps God, whom Kant would have liked to believe in) would, by a use of the Categorical Imperative as it would be used by such a supremely rational will, select these simpler principles for the guidance of less rational wills than his own. But 'we have no intuition of the divine perfection, but can only deduce it from our own conceptions' (*Gr* BA92 = 443). We have no direct access to what a good God would will, so we have recourse to our own imperfect reason as the best means available to us.

In conclusion, we have to ask, in deference to an earlier objector, whether this development of Kantian ideas relies on resources lying beyond the philosophy of language, and in particular on antecedent substantial moral ideas and intuitions. Kant called his most-read work on this subject *Grundlegung zur Metaphysik der Sitten*. What has been sketched in this chapter is a kind of *Grundlegung zur Logik der Sitten*; and, as we have seen (1.3), logic and metaphysics are hard to tell apart. It certainly does not seem as if we have relied on extra-logical premisses. Anyone who doubts this should look for them. The argument has been generated using the following elements: first, the prescriptivity of moral judgements; secondly, their universalizability; and thirdly, the thesis that fully to represent another's situation to oneself one must come to have a will similar to his (or hers) for a situation in which one occupies his role. The last of these elements was obtained by considering what full representation of a hypothetical situation means, and what it means to think of a person in it as myself. All these are conceptual or logical moves, not involving appeals to substantial moral intuitions. Although, therefore, they can all be disputed, the disputes will be within the philosophy of language,

since the theses themselves belong to it. So at least we can claim to have shown the relevance of philosophy of language to ethics. But see 5.8 for further discussion of universalizability.