Chapter 10
The Truth about Metaphor

Gottlob Frege introduced into philosophy two doctrines whose subsequent influence, on analytic philosophers at least, has been momentous. One is the doctrine that to understand a sentence is to know how to set about establishing the truth-value of an assertion couched in those words. The other is the doctrine that a word has meaning only in the context of a sentence.

These doctrines make it hard to understand metaphor. They create difficulties, especially, for any theory of metaphor which assumes the intelligibility of talk about two *kinds* of *meaning*, literal and metaphorical. Frege's doctrines allow only one kind of meaning: one connected essentially with the issue of truth and falsity. But it is only when we take them literally that metaphorical sentences acquire a truth value: usually false.

Two obvious suggestions offer ways of getting around this. The first is the suggestion that metaphors are condensed or disguised *similes*: true or (sometimes) false statements about similarity. The second is the suggestion that a term used in a metaphorical context just has a different *literal* meaning from the one it has in a sentence intended to be taken literally.

Neither is very plausible. The first reduces a metaphorical comparison, 'Man is a wolf', to a literal statement of similarity, 'Men are like wolves'. But what similarity, exactly, is the statement supposed to be asserting? Perhaps that men prey upon one another, as wolves are supposed to do. But then 'prey' can only be meant metaphorically, at which point the puzzle revives. As Goodman says of the attempt to find a simile in the metaphorical assertion that a picture is sad, 'What the simile says in effect is that person and picture are alike in being sad, the one literally and the other metaphorically. Instead of

metaphor reducing to simile, simile reduces to metaphor; or rather, the difference between simile and metaphor is negligible.'[1]

An equally fatal objection to the second suggestion is that it is only if we take the terms which compose a metaphor in their plain everyday senses that they compose a metaphor. As R. M. J. Dammann puts it, 'If God is a rock in a different sense from that in which Gibraltar is one, then both are really, literally, rocks, just as what one buys in a sweetshop is rock, though a dissatisfied customer might call it such metaphorically.'[2] It cannot, in other words, be *literally* true that the process of understanding a metaphor involves changing the sense of its constituent expressions, otherwise metaphor would evaporate in the reader's grasp into literality.

What this suggests is that what makes a sentence metaphorical is not any change in the meaning of its constituent expressions, but the way in which those expressions, with their plain everyday meanings, are combined in the sentence. What we need, perhaps, if we are to begin to understand metaphor, is not a theory of metaphorical *meaning*, but an account of what it is to *assert* metaphorically: an account, that is, which would operate at the level of the sentence rather than at that of the word. But this returns us to Frege's two doctrines. If the meaning of a word is to be *wholly* identified with the contribution it makes to the determination of a truth-value for each sentence in which it occurs, then it is difficult to see how the constituents of a sentence could determine anything other than a plain, literal meaning for the sentence as a whole.

What I shall try to show in this essay is that this difficulty can be overcome: that we can give a satisfactory account of the semantic mechanisms of metaphor without giving up either of Frege's doctrines. As Frege thought, understanding a sentence *is* knowing how to set about establishing the truth or falsity of assertions couched in those words. But to say what, exactly, someone knows in knowing that demands, I shall argue, a more complicated story than Fregeans have supposed.

I

First, though, I want to consider a more radical strategy than mine for squaring Frege's doctrines with our experience of metaphor, proposed by Donald Davidson.[3] Davidson holds with Frege that the notion of meaning is wholly explicable in terms of the notion of truth, but draws from this premiss the conclusion that metaphor has nothing to do with meaning. 'The concept of metaphor as primarily a vehicle for conveying ideas, even if unusual ones, seems to me as

wrong as the parent idea that a metaphor has a special meaning' (p. 32).

Locke would have agreed with this, because Locke thought the functions of metaphor were decorative, forensic, or emotive – never cognitive. Davidson, however, is not here of one mind with Locke. Metaphors do yield cognitions of a sort, 'prompt insights': so much can be conceded to 'interactionist' theories of metaphor of the sort advanced by Max Black. It is just that the kind of cognition involved is in no way related to meaning. Metaphor is related to 'what it makes us see' rather as a physical event is related to 'what it makes us see'. 'Joke or dream or metaphor can, like a picture or a bump on the head, make us appreciate some fact – but not by standing for; or expressing, the fact' (p. 46). It follows, plausibly, that what a metaphor makes us see is not determinate in extent and may not all be formulable in words. 'How many facts or propositions are conveyed by a photograph? None, an infinity, or one great unstateable fact? Bad question. A picture is not worth a thousand words, or any other number. Words are the wrong currency to exchange for a picture' (p. 47).

Less plausibly, it follows that metaphor is not, strictly speaking, a form of communication between writer and reader: metaphors do not '*convey* ideas' [my italics].

> The central error about metaphor is most easily attacked when it takes the form of a theory of metaphorical meaning, but behind that theory, and statable independently, is the thesis that associated with a metaphor is a definite cognitive content that its author wishes to convey and that the interpreter must grasp if he is to get the message. This theory is false as a full account of metaphor, whether or not we call the purported content a meaning. (p. 46)

There is something not quite right about this. What is fishy is not what the passage asserts, which is unexceptionable, but what it presupposes. Nobody – or not many – would seriously want to contend that a metaphor conveys a 'cognitive content', if by that is meant a paraphrasable 'message' of some sort. But aren't there kinds of verbal communication which, while intimately dependent on meaning, do not involve the transmission of a 'cognitive content' in that sense?

Take jokes. Jokes *have a point*. The notion of a joke's having a point invokes a corresponding notion of relevance. We see the point of a joke only if we see what is relevant to its point and why. Not all of what gets said in telling a joke is relevant to its point. Some of it is mere scene-setting, and one way of being tedious is to resist all

attempts to explain the point of a joke by fixing on some bit of incidental scene-setting and demanding, impossibly, that it be shown how *that* is relevant to the point.

The mere fact, now, that some things are relevant and some not to understanding a joke suffices to give sense to a correlative concept of communication. If I tell a joke badly I fail to *communicate* to you the point of the joke. Notice, however, that this is a sense of 'communicate' which is not susceptible of Derridean deconstruction. Derrideans, unlike Davidsonians, doubt whether communication by means of language ever involves the transmission of 'a cognitive content which the author wishes to convey' and that the 'interpreter must grasp if he is to get the message'. Their claim is that understanding neither does nor can pass beyond the text itself to a reconstruction of the intentions or cognitive states of its author, and that any theory which claims otherwise must 'deconstruct' itself because in order to get itself formulated it must itself use language in ways which conflict with the thesis it is attempting to formulate. Jokes conform to the Derridean model in this sense, that the point of a joke is not something that the teller of the joke can intend, or fail to intend, to communicate. The point is inherent in the text of the joke, and the interpreter, in grasping the point, never has to refer beyond the text to the intentions or cognitive states of the teller. The whole transaction remains on the plane of 'textuality'. For all that, however, it makes perfect sense, for the reasons offered above, to speak of communicating or failing to communicate, the point of a joke. And, equally clearly, grasping the point of a joke depends upon first grasping the meanings of the words employed in telling it.

What about the other components of Davidson's string of putatively analogous cases: dreams, photographs, and bumps on the head? Here, where Davidson needs analogies, there are important disanalogies.

We can indeed be led by a bump on the head to 'appreciate some fact', but the fact in question will vary with the context and the head, and can be anything you please; whereas what we 'see' in the case of a joke is its point, which cannot be just anything you please, but is the same for everyone who understands the joke.

Dreams, according to Freudians and Jungians, have a point; but the point is in each case matter of dispute. Some of us may find one or another interpretation of a dream convincing, but none of us ever finds it uncontroversially evident, in the way that the point of a joke is, more often than not, uncontroversially evident.

Again, to say that a photograph has a point is just to say that it has a purpose. To know what that is one has to refer beyond the

photograph itself, to a text which it illustrates, a bit of personal experience which it records or a scientific enquiry which it subserves. The point of a joke, on the other hand, is an effect of language; as Derrida would say, an effect of *différance*: to grasp it we have no need to look beyond the text of the joke as told to the intentions or purposes of the teller.

We have, then, a substantial set of disanalogies between jokes on the one hand and dreams, photographs, and bumps on the head on the other. It terms of these disanalogies metaphors pretty clearly belong with jokes. Metaphors have point in the way that jokes do. Not everything counts as relevant to the point of a metaphor. Reading the opening lines of Sonnet LXXIII,

> That time of year thou mayst in me behold
> When yellow leaves, or none, or few, do hang
> Upon those boughs which shake against the cold,
> Bare ruin'd choirs where late the sweet birds sang . . .

might 'prompt the insight' that fan vaulting and deciduous trees resemble one another in the way their structure distributes stress. But this has nothing at all to do with the point of the metaphor, which is to forge connections, of what kind I shall discuss in a moment, between maturity and old age, summer and winter, poetry and the drought of poetry.

Again, to see that this *is* the point of the metaphor, we do not need to know that that is what Shakespeare intended the point to be, any more than to see the point of a joke we need to entertain hypotheses about the teller's intentions in telling it. The point of the metaphor, like that of the joke, is in some sense which we have yet to unravel 'there in the text'. And there is a sense in which the point is the same for every reader, that sense being that whatever Derridean 'supplement' of meaning a reader may draw from the potentialities of the text, there remains a distinction between developing the potentialities which the text offers to interpretation and simply reading into the text the content of some arbitrary private brainwave, such as the 'insight' about stress in fan vaulting and deciduous trees.

What then about the relationship between metaphor and meaning? Davidson's suggestion goes something like this: a metaphorical statement expresses a factual assertion, usually false. But the assertion in question, though false, has the interesting property of suggesting all sorts of thoughts and insights, even though none of these are actually *meant* (expressed) by the metaphorical sentence.

If what we have said is correct, this makes the connection between metaphor and meaning altogether too tenuous. The point of a

metaphor is not just any thought which it may suggest to an individual reader: it is 'there to be seen', as the point of a joke is 'there to be seen' by any reader or hearer.

In either case, therefore, someone who sees the point must be applying to the joke or the metaphor some more or less standard procedure of interpretation (what Derrida would call an 'organon of iterability')[4] which supplies the necessary constraints on relevance, and which must at some point engage with the meanings of the words in which the joke or metaphor is framed, because there is simply nothing else on which interpretation could go to work.

I conclude that Davidson is wrong about the relationship of metaphor to meaning, and that it is also a mistake to think that nothing is communicated by a metaphorical sentence. What is communicated is its point.

On the other hand I agree with Davidson that what is communicated is not a 'cognitive content', if by that is meant something which could be given complete expression in a proposition. And I agree with him on a more fundamental matter: his remark that 'as much of metaphor as can be explained in terms of meaning may, and indeed must, be explained by appeal to the literal meanings of words' (p. 41).

This, though, brings us back once more to Frege, who thought that the literal meaning of words consisted wholly in their power to direct the determination of the truth-values of sentences. If Frege was right, and words have meaning *solely* in relation to the determination of truth and falsity, then, it seems, *meaning* could not be the source, or even one source, of the constraints on relevance which, seemingly, have to be supposed to govern the interpretation of metaphor. For in the interpretation of metaphor truth and falsity are not at issue. That is why Davidson's account of metaphor, which simply takes it for granted that no such constraints operate, is so penetrating and so seductive.

Nevertheless, as I said earlier, I think that a way can be found to concede Frege's doctrine without conceding the more counter-intuitive parts of Davidson's. I will now try to say what that way is, and in the process to sketch the sentence-level, rather than word-level, account of the semantic mechanisms of metaphor promised earlier.

II

Frege held that to understand a sentence is to be in a position to determine its truth-value, or to be precise, the truth-value of the

corresponding assertion. Many more recent writers who take themselves to be saying the same thing put it this way: to know the meaning of a sentence is to know its truth-conditions.

I want to suggest that this is not just a piece of harmless shorthand, but one which involves a serious misrepresentation of the content of Frege's insight. The trouble is that the term 'truth-condition' can mean either of two quite different kinds of thing, but is almost invariably taken to mean only one of the two.

The ambiguity is not difficult to see, if we consider what is actually involved in knowing how to determine the truth-value of an assertion. Take, for example, 'There is a ruin', asserted of some object or other, call it O.

To assess whether 'There is a ruin' is truly or falsely asserted of O, I shall need, certainly, to be able to *recognize* whether O possesses the *characteristic features* of a ruin. And it seems clear that I could be taught to do this simply by being shown some ruins and taught to discriminate them from other superficially similar sorts of thing. I might, for instance, learn to discriminate the grassy mound which remains of a deserted croft from cairns and other kinds of stone heap by looking for a roughly rectangular shape and the foundations of the chimney.

But now, suppose I am quite ignorant of the English language (and perhaps a little stupid). What is to stop me from thinking that 'ruin' is just a word (perhaps a Gaelic one) for a particular kind of stone heap? Notice that the realization that the kind of heap in question originates in a house would not necessarily prevent me from thinking this. For I might think that 'ruin' is the name for a particular configuration of stones left by that kind of house, as 'dolmen' is the name for the most durable portion (the main burial chamber, consisting of capstone and uprights) of a passage grave, which also possesses certain easily recognizable characteristic features.

Thinking that, I am in the following position. I can assess the truth or falsity of some kinds of assertion made by means of the sentence 'This is a ruin'. To put it another way, I know some of the truth-conditions of 'This is a ruin' where 'truth-condition' is taken to mean something like 'criterion of recognition'. But I do not yet grasp the *meaning* that attaches to 'ruin' in English. If a fellow foreigner (whose English may be better than mine) says of Fountains Abbey 'There is a ruin', then, because the ruins of Fountains Abbey do not exhibit the characteristic features of the kind of object I have been taught to count as a ruin, I shall have no choice but to take him either to be speaking falsely or to be talking nonsense.

What do I need to be taught, if I am to acquire the capacity to

extend the application of 'ruin' from the remains of eighteenth-century crofts to the other kinds of object to which the term applies in everyday English?

It might look as if what I need is simply to be taught the characteristic features of those other kinds of ruin. But this suggestion will not do, for two connected reasons. The first is that there is no clear limit to the forms that ruins can take, and thus no clear limit to the sets of characteristic features which may come to be taken as warranting the assertion 'There is a ruin'.

The second is that if my grasp of the meaning of 'ruin' rests simply on a finite list of sets of characteristic features, then I shall not be able to extend the application of the term beyond the limits of that list, *except* by dint of linguistic stipulation. But clearly not every extension of the application of a term by competent speakers rests on linguistic stipulation. As Putnam[5] and others have insisted, there can be discoveries about the applications of terms. Thus, for instance, the archaeologists who first discovered the relationship of dolmens to intact passage graves *discovered thereby* that dolmens are a kind of ruin. The discovery of the relationship of dolmens to passage graves brought about, in other words, an extension of the list of sets of characteristic features warranting the assertion 'There is a ruin'. The characteristic features of dolmens now had to be added to the list. But no new linguistic stipulation was necessary to bring this about. It was a simple consequence of the discovery that dolmens are the remains of passage graves, though admittedly not a causal, but a linguistic consequence.

What I need to enable me to make comparable extensions of the application of terms (ones which will automatically match those made by other competent speakers, that is) is, then, not just more *criteria of recognition*: more sets of characteristic features of things, as it were, worthy-to-be-called-ruins. What I need is rather some principle regulating the admission of sets of characteristic features to the status of assertion-warranting considerations for the assertion 'There is a ruin'. Put another way, what I need is a principle which will tell me what is and is not relevant to the issue of whether or not something is a ruin. I need a *criterion of relevance*.[6]

Once we see that this is what is needed, it is not difficult to see what form it must take. I shall have hold of a suitable criterion of relevance as soon as I grasp that a ruin is the remains of a structure dilapidated beyond the point at which it can still be used to serve the purposes of the original structure. Call this *CR*.

At first *CR* might seem to be no more than a statement of some more properties of ruins, differing only in scope or generality from

the statement that ruined crofts are roughly rectangular grass-covered heaps of stones with the foundations of a chimney at one end. And of course *CR* does state some properties of ruins. Where it differs from the statement about the characteristics of ruined crofts is in the way it functions with respect to the procedures of determining truth-values for assertions made by means of the sentence 'There is a ruin'. The statement about ruined crofts gives some criteria by which ruined crofts can be recognized (states a criterion of recognition). *CR* by contrast does not say how any kind of ruin is to be *recognized*. It says what a ruin *is*. It does that by offering a criterion of relevance by appeal to which we can establish, for instance, the relevance to the issue of ruinhood of the features mentioned in the statement about ruined crofts. They are relevant because they are causally connected with properties characteristic of intact crofts, and because such features do not, in fact, come into being by any means other than the dilapidation across time of such a structure.

Once I know that ruins are structures so dilapidated as to be no longer usable for their original purposes, I can see why Fountains Abbey and a dolmen are just as properly called 'ruins' as the remains of crofts on which I first cut this particular linguistic tooth. I have a principle which enables me to extend the application of the term, without benefit of new linguistic stipulation, in ways which match, unsurprisingly, the extensions made by other competent speakers armed with the same principle of relevance.

No doubt, as Frege said, understanding the meaning of a sentence is knowing how to determine its truth-value. But the process by which we pass from a sentence to a truth-value turns out to be more complex than is customarily supposed, requiring access to criteria of two quite different kinds: criteria of recognition and criteria of relevance.

Should we, then, conclude that knowledge of meaning requires knowledge of both kinds of criteria? The answer is no, for the following reasons. Any account of what is learned in learning a language must meet the following conditions of adequacy: (1) What is represented as learned must be in principle finite. (2) What is represented as learned must be stipulative in character. I shall offer no argument for these conditions here, except to observe that they are commonplaces of theoretical linguistics.

It is to be noted, now, that criteria of recognition meet neither condition. On the one hand the list of features which can serve as diagnostic of a ruin is not in principle finite (for example, something as recondite as a crop mark seen from a helicopter can on occasion serve to warrant the assertion 'There is a ruin'). And on the other

hand the issue of what is and what is not a characteristic feature of ruins is not, in principle, the kind of issue that can be settled by linguistic stipulation. It happens to be the case that the characteristic features of dolmens are characteristic features of a type of ruin. But that this is so is not a consequence of the existence of a linguistic stipulation to that effect, but a consequence of its happening to be the case, as a contingent matter of fact, that dolmens are the remains of passage graves.

On the other hand, criteria of relevance meet both conditions. Firstly, such a criterion is in principle finite. To know that ruins are the remains of structures dilapidated beyond the possibility of serving their original purposes is to know all there is to know, linguistically speaking, about what a ruin is. Armed with that knowledge, a speaker is in a position to set about finding out for himself what the characteristic features of ruins happen, empirically speaking, to be. The class of criteria of recognition which he will arrive at in that way need not be, in principle, a finite class. But that fact has no tendency to make the criterion by which he determines membership of that class any the less a finite criterion.

Secondly, the content of a criterion of relevance is clearly going to depend entirely on linguistic stipulation. If English speakers were to begin using the term 'ruin' to pick out, say, a structure to any degree dilapidated, then the relevance criterion governing the use of the term would have changed, and changed by mere stipulation.

I conclude that, since meanings are presumably part of what is learned in learning a language, knowledge of meanings amounts simply to knowledge of criteria of relevance.

This implies that we can distinguish a speaker's linguistic knowledge from the general body of empirical knowledge in his possession and so is at odds with Quine's thesis that the two are in principle inextricable. The inextricability thesis, however, depends via Quine's notion of an 'analytical hypothesis' on the assumption that we can only set about disentangling linguistic from empirical knowledge if we have already managed somehow to make sense of the notion of analyticity, considered as a species of necessary truth.[7]

It is thus sufficient to evade Quine's strictures about inextricability to remark that the way of disentangling linguistic from general knowledge proposed here does not in any way require us to make sense of the notion of analyticity. A criterion of relevance, such as the statement that a ruin is a structure dilapidated beyond the possibility of serving its original purpose, is not an analytic truth, because it is not a necessary truth. It simply records the entirely contingent and empirical fact that speakers of English use a certain criterion for

determining what is and what is not relevant to the application of a certain term.

Back now to Frege. Frege held that a term has both a reference (*Bedeutung*) and a sense (*Sinn*), and that the sense is what picks out the reference, or to put it another way, determines the extension of the class of things to which the term applies.

Philosophers since Russell have taken it for granted that a Fregean sense is a description. But, lacking access to the distinction between the sort of description which serves as a criterion of recognition and the sort which serves as a criterion of relevance, they have assumed that the description in question is one which states the criteria by which we are to *recognize* something as falling within the extension of the term. Thus, the Fregean sense of 'ruins' would be some set of instructions for recognizing ruins.

Now plainly, if that is what a meaning (a sense) is, then to understand the meaning of words is simply to possess a great deal of homespun knowledge about what things – real, commonplace things – are like. On such a view, meaning is Gradgrind's province, not the poet's: a sort of plain man's Encyclopaedia of the Obvious. It is easy to see what, on such a view, the connection is between meaning, so conceived, and the literalities of Lockeian plain speech, but hard indeed to see what bearing meaning could have on metaphor.

The above arguments suggest that a Fregean sense is better regarded as a criterion of relevance. I am not of course suggesting that there is the slightest textual warrant for such an interpretation in Frege. What I am saying is that if 'sense' is regarded, in accordance with Frege's use of the word, as what picks out the extension of a name or the truth-value of an assertion, then a sense cannot be a criterion of recognition. It has to be a criterion of relevance.

This makes the determination of truth and falsity in principle a two-stage process. First there is the application of the criterion of relevance to select a criterion of recognition appropriate to a particular context, then there is the application of the criterion of recognition to assess the truth or falsity of a given assertion with respect to that context. Only the first stage involves the application of linguistic knowledge – knowledge 'of the language': of *meaning*.

Now, suppose there is no question of determining the truth or falsity of an assertion because, taken literally, it is either evidently false or evidently absurd. In that case there is no point in proceeding to the second stage of truth-determination. If meanings were just recognitional capacities, then to withdraw from stage two truth-determination would be to set aside any interest in meaning *per se*, since there would be no other purpose which a grasp of meaning

could be made to serve. But meanings are not just recognitional capacities, so withdrawal from stage two truth-determination still leaves us with the principles of relevance (the only principles deserving the title of *meanings*, in fact) which direct stage one.

In most cases a criterion of relevance will possess internal complexity: it will exhibit a schematic structure of some sort. There thus exists the possibility of interpreting a piece of text in a way which does not raise the issue of truth or falsity, by systematically mapping the subject-matter of the text on to the structural framework provided by one or more of the criteria of relevance associated with the words of the text.

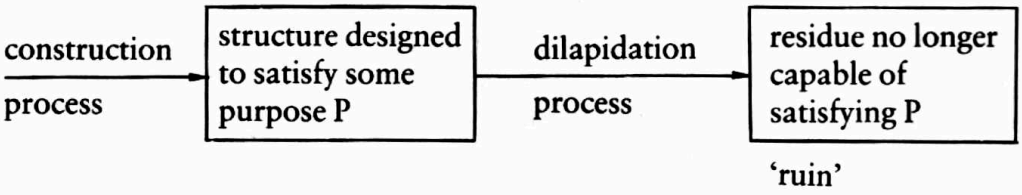
Not all, but much, of what we think of as metaphor involves the operation of this kind of semantic mechanism. I shall try to show in the next section how it works in a specific case. But if this is, or is often, what is involved in interpreting metaphor, then we shall have fulfilled two promises made earlier. We shall have shown metaphor to be a phenomenon which operates not at word-level, but at the level of the interpretation of whole utterances, including sentences, and one which does not involve any change in the meanings of individual words. And since, if the account is correct, one and the same set of criteria direct *both* the determination of literal truth and falsity *and* the interpretation of metaphor, we shall also have managed to preserve Davidson's plausible demand that, to the extent that metaphor is explicable in terms of meanings, it must be explicable in terms of the plain, literal meanings of words.

III

Let us return to lines 1–4 of Sonnet LXXIII. 'Ruin'd' is the key word of line 4, and the nub of the metaphor contained in the line. It refers back to the boughs of line 3, which it characterizes as 'bare ruin'd choirs'. What are we to make of the invitation to think of bare trees as ruins? Nothing, evidently, if we take the invitation literally. It is simply false that bare trees are or could be ruins. There is an older sense of 'ruin'd', accessible to Shakespeare and still active in Milton, to whose latinizing tendencies it was congenial, in which 'ruin'd' means 'fallen'. But that gets us no further: the boughs are bare, but hardly fallen, ones.

As long as we look in the lines for a true assertion, then, interpretation is blocked. But suppose we put the issue of truth and falsity on one side. We are still left with the meanings of the words, in the sense of the criteria of relevance governing their application.

We can represent the criterion of relevance which governs the application of 'ruin' roughly as follows:



Since the text balks any attempt to apply this schema to the determination of a truth-value, we move to another level of interpretation: we ask whether there is any way in which the subject-matter of the lines can be ordered analogically by reference to the schema.

To begin with, if the bare boughs of lines 3–4 are to be seen, in the terms offered by the above schema, as structurally analogous to ruins, there must be something which the reader can take as analogous to the process of dilapidation which produces the 'ruin'd' state (bareness) of the boughs. Such a process is ready to hand in the fall of the leaves, which is also invoked in line 4 by the older sense of the term 'ruin' mentioned earlier, in which it is an intransitive verb meaning 'fall'.

But introducing the fall of the leaves in this way locates it, not as a mere natural process of change, but as something which subverts and renders houseless some purposive order, which like the order of a household depends for its existence on the continued integrity of some physical fabric.

The word 'choirs', which carries the second metaphor of line 4, makes clear the nature of the purposive order in question. It is the order of sung offices which constitute the life of a cathedral or abbey, and are the purpose for which the building was erected. So we are to see the order of life in the summer words (the order which has been unhoused and rendered homeless by the falling of the leaves: the 'ruining' of the boughs) as analogous to the life of a religious building: the order of sung offices. Finally, on to this elaborate but quite determinate structure of analogies the poem maps the stages of a man's life, from prime (summer, the woods in full leaf, the order and ceremony of nature, conceived as analogous to the Christian ceremonial order, in full swing) to failing strength and approaching age, or perhaps loss of powers[8] (autumn, the leaves fallen, the birds fled, the choristers gone from a deserted and decayed church).

So much for the semantic machinery governing the interpretation of the lines as metaphor. However, the point of the lines, I take it, is

not merely to give the reader the opportunity to solve this little hermeneutical puzzle. Their point is to restructure the reader's feelings by opening, between commonplace words, channels of analogy through which feelings attached to one set of words and what they mean may flow and embrace other words, and what *they* mean. The effect of the whole nest of interlocking analogies is to open a passage for a certain way of feeling about ageing or loss of powers.

One powerful tradition of feeling sees the prime of life as an essentially flawed and sin-ridden mode of existence of which we are well rid in death, provided we have lived a Christian life. Sonnet LXXIII works against this. By presenting age and failing powers in terms of the analogy between bare boughs and deserted choirs it confers in retrospect upon the prime of human life and upon the natural order of the summer woods something of the spiritual authority associated with the Christian ceremonial order. It invites us to feel about the prime of life somewhat as we feel about the order of a cathedral: to feel towards it as towards something having that kind of richness, strength, and seriousness. The poem, ostensibly about death, is also about life.

Notice that there is in all this no question of the terms 'bare', 'boughs', or 'ruin'd' taking on new, 'metaphorical' senses distinct from the common everyday ones. On the contrary, it is precisely the reader's understanding of those commonplace senses which sustain his understanding of the metaphor.

At the same time something, a certain response to ageing and failure of powers, is communicated by the poem. So in a sense we can speak of the process of coming to understand the poem as having a cognitive aspect. The outcome of it is that we become acquainted with a certain structure of feeling; and acquaintance is surely a cognitive relationship. On the other hand, the process of becoming acquainted with that structure of feeling is not a matter of assimilating a propositional content. It just *is* the process of analogical restructuring, with the accompanying redirection of emotional responses, which we have just described. The nature of that process can, in a phrase of Davidson's, be 'brought out by using further words' – we have just done so – but the style of criticism which that enterprise yields is a purely explicatory, inherently non-reductive one. The 'further words' in question function not as a paraphrase, but in a way analogous to an explanation of the point of a joke. No words other than the actual words of the poem could paraphrase the poem, could express 'what it is the poem makes us see', because the process of analogical restructuring involved in 'seeing' that depends essentially upon access to the criteria of relevance which attach to

just those words, and on relating them to one another in an order determined by the precise order in which those words stand on the page.

IV

Hugh Bredin has recently proposed the following way of drawing the distinction between metaphor and metonymy:

> A metaphor is understood by virtue of its own semantic structure; if it asserts a similarity between things . . . this is something that may strike us for the first time, but which is none the less intelligible for that. But if I use the metonymic 'crown' or 'sceptre' for the institution of monarchy, the auditor must *already know* of the close connection between those objects and the institution. If he does not, he will fail to understand. Metaphor *creates* a knowledge of the relation between its objects; metonymy *presupposes* that knowledge.[9]

It is important to notice that the distinction drawn in this way is not a distinction between two senses of the notion *understanding a text*, or between two textual devices. Metonymy so defined is not a *textual* device at all, since our understanding of a metonym does not originate in the text, but in a piece of adventitious extra-textual knowledge.

At paragraph 4.026 of the *Tractatus Logico-Philosophicus* Wittgenstein remarks, 'The meanings of simple signs (words) must be explained to us if we are to understand them. With propositions, however, we make ourselves understood.' A profound point is at stake here, not just about the recursive character of the relationship between understanding a sentence and understanding its constituents, but about what it is to find oneself addressing a written text and *reading* what is written there. The point is that we are not *reading* if the purport of what is written has to be explained to us if we are to understand it. Nor, in such a case, does the writing in question constitute a *text*; for *ex hypothesi* there is no way in which an understanding of its meaning can be derived from the meanings of its constituent signs. That being so, we are free to ascribe to it any meaning we choose, as in the case of an invented word (a new 'simple sign'); but a string of marks which we are free to interpret in any way we choose is not a *text*, but only a string of marks.

If *reading*, engagement with a *text*, is to occur at any level of interest to literary criticism, then it seems that there must be some rhetorical devices for which understanding depends merely upon the reader's knowledge of the language in which the text is written, and

not upon knowledge of adventitious associations of which he may or may not happen to be aware. It would be nice if we could show that metaphor is such a device; and that is, roughly speaking, the task I have set myself in this essay.

Max Black's interesting and influential theory of metaphor fails, in my opinion, to engage with this aspect of the problem.[10] I agree with Black, and for that matter Goodman, that metaphor involves the perception of structural analogy. But Black identifies the analogically interacting structures as 'the systems of commonplaces' associated with metaphorically active terms. An 'associated commonplace' is just that: a bit of general knowledge which is associated with the word, either because of a relationship of natural causality or because of an habitual pattern of usage (like the habitual use of the phrase 'the Crown' to refer to the institution of monarchy), but which is in no way essential to the meaning of the term. The patterns of *incompatibility* set up between such 'systems of commonplaces' can, as Black suggests, lead us to focus on some subset of one or other set of connotations, but what we focus on in such a case cannot, in the nature of things, ever be something we didn't *know already*. The difference between metaphor and metonymy remains unaccounted for; and it comes as no surprise to find not merely Black's critic Paul Ricoeur, but Black himself, locating this as a difficulty for the theory. As Ricoeur puts it, 'The major difficulty (which, by the way, Black himself recognizes . . .) is that to return to a system of associated commonplaces is to address oneself to connotations that are already established. In one stroke, the explication is limited to trivial metaphors.'[11]

My object in this chapter has been to remove this difficulty. If the structures which metaphor analogically reinterprets are not structures *of associations*, then what is revealed by reinterpreting them will not necessarily be a 'commonplace': something already known to us. On my account grasping the meaning of a term is not a matter of knowing its familiar associations, but of grasping a rule for determining what is and what is not relevant to its application. So when we analogically reinterpret the internal structure of such a rule, as we do in reading Sonnet LXXIII, what we get is something new: a set of entirely *un*familiar associations between one set of things and another. Coleridge is vindicated: there *is* a difference between the mechanical operations of 'fancy', which receives 'all its materials ready made from the law of association', and the workings of the imagination, which 'dissolves . . . in order to re-create' (*Biographia Literaria*, ch. 13).

Those who defend Coleridge, it is often supposed, must turn their

backs on Derrida. On the contrary, it seems to me, one could hardly find a better instance of Derridean *dissémination* than the operation of the word 'ruin'd' in Sonnet LXXIII, or for that matter, in Owen Barfield's study of the history of the word in *Poetic Diction*.[12] Who would have thought the word 'ruin'd' had it in it? What could show better the power of a purely textual device to lead the mind away from all the customary associations of a word, into a new structure of contrast and collusion, of *différance*? English-speaking 'deconstructionist' literary critics often seem inclined to take Derrida not as proclaiming the priority of the text over what Husserl called the 'living present of consciousness', but as proclaiming the unlimited power of interpretation over the text. Thus Paul de Man takes the contrast between metaphor and metonymy (rightly) to be a contrast between necessity and chance, and comes down on the side of metonymy and chance, arguing that what we take to be metaphor is (often? always?) merely disguised metonymy.[13]

The difficulty with this is the one we have just been exploring. To the extent that a putative interpretation attaches to the text merely through the operations of chance, it is not, after all, an *interpretation*. No denial of this is to be found in Derrida, not merely because his thesis in no way requires such a denial, but because such a denial would be strictly incompatible with his position. Derrida's most fundamental claim is that meaning resides not in consciousness but in the text. His central argument for that claim is that any text is *itérable*: readable in the absence, or after the death, of its author. To maintain that argument he needs to be able to deploy the notions of *text* and of *reading* in ways which would be fatally undercut, in the way suggested earlier, by the claim that interpretation is not governed by what Derrida calls an *organon of iterability*,[14] but is a subjective, or private, or arbitrary process.