Harold I. Brown

SELF-REFERENCE IN LOGIC

AND MULLIGAN STEW

The novel has always provided a vehicle for commenting on various aspects of human existence. We are familiar with the political novel, the historical novel, or the metaphysical novel, and in this sense Sorrentino's *Mulligan Stew*, with its running commentary on novels, novelists, critics and publishers, may be viewed as a critical novel. A critical novel, however, has a striking feature which it does not share with the other sorts of novels mentioned above in that a critical novel is itself a novel, and a member of the class of objects on which it is commenting. This is a structure that logicians call "self-reference" and the phenomenon of self-reference has provided a central theme in twentieth century logic and philosophy.

Early in the century logicians became interested in self-reference when they realized that it leads to some rather striking paradoxes, and the search for ways of eliminating these paradoxes became a major concern in the development of modern logic. The initial approach to the problem was to argue that, since self-reference yields paradox, there must be something vicious about any self-referential linguistic structure; thus logicians attempted to reconstruct language so as to eliminate all self-reference. Slowly, however, it became apparent that this

121

approach is too sweeping, for it is only in a few rather isolated cases that self-referential statements yield paradox, while in other cases self-reference provides the key to some profound discoveries. Indeed, within formal logic itself, one of the deepest and most significant results, the Gödel theorems, depend crucially on the use of self-reference. More recently, in *Gödel, Escher, Bach*, Douglas Hofstadter has presented us with a sustained meditation on the role of self-reference in music and art, as well as in logic, and has suggested that the capacity for self-reference may be one of the characteristic features of human intelligence. It is in this milieu that Sorrentino's self-referential novel appears.

In order to develop these ideas I will begin by sketching out some of the most striking paradoxes of self-reference in logic, exploring one major attempt to legislate away all self-reference, and outlining the ways in which the proofs of the Gödel theorems depend on the systematic use of self-reference. Secondly, I will consider some of Hofstadter's meditations on the pervasiveness of self-reference. Finally, using *Mulligan Stew* as a primary example, I will discuss the role of self-reference in literature, a field that Hofstadter does not explore.

I

When logicians describe a situation as "a paradox" they do not mean just that it is surprising, odd or unusual; rather, the force of this claim is that the situation in question leads to an explicit contradiction. To see how this develops as a result of self-reference I will discuss one version of the so-called "liar paradox," perhaps the oldest of the paradoxes of self-reference. Consider the following sentence, which I will label "*A*" for ease of reference:

A. Sentence *A* is false.

Sentence *A* is self-referential in that it makes a claim about itself, and the logical paradox arises when we ask if sentence *A* is true or false. Suppose on the one hand, that *A* is true. In this case the claim that *A* makes is correct, but the claim that sentence *A* makes is that that sentence *A* is false; thus if *A* is

122

true then A is false. On the other hand, if A is false, then the claim that it makes, i.e., that A is false, is incorrect, and sentence A is therefore true. Thus if A is false, it is true, and if A is true, it is false—this is the logical paradox.[1]

There are many other examples of logical paradox in literature, but I will mention only two which will serve to bring out specific variations on the main theme. The following pair of sentences B and C will serve to show that the paradoxical pattern can occur in cases of properly interrelated sentences, even though none of the sentences taken by itself involves any paradox or even any self-reference:

B. Sentence C is true.
C. Sentence B is false.

Here, again, the paradox results from considering if these sentences are true or false. Suppose, for example, that B is true; this requires that C is true; and this in turn requires that B is false. If, however, B is false, then the statement that B makes must be false, i.e., C must be false. But if C is false, then the statement that C makes, that B is false, must itself be false, and thus B is true. Once again, if B is true then it is false, and if B is false, then it is true. The same kind of result can be obtained by considering whether C is true or false.

One more example will show that paradoxes can arise without considerations of truth and falsity. The "Grelling paradox" arises from the observation that some words describe themselves while some words do not. The word "English," for example, is an English word, while the word "German" is not a German word. Similarly, "short" is a short word while "long" is not a long word. If we call words which describe themselves *autological* and words which do not describe themselves *heterological* then, since

[1] Throughout this discussion I am assuming what logicians call the "law of excluded middle," i.e., that every declarative sentence is true or false; there are two reasons for proceeding in this way. The first is that this principle was accepted in the discussions of self-reference that I am concerned with. The second is that although it may seem possible to avoid the paradoxical result by rejecting this law and thus refusing to move from the claim that a sentence is not true to the claim that it is false, or from the claim that a sentence is not false to the claim that it is true, there is a "strengthened liar paradox" that can be constructed for this case as well. Cf. Susan Haack, *Philosophy of Logics* (Cambridge, Cambridge Univ. Press, 1978), p. 140.

123

these two terms are themselves words, we may turn them back on themselves. Let us ask then, if the word "heterological" is itself autological or heterlogical. If it is autological then, by definition, it correctly describes itself, but if it correctly describes itself, then it is heterological. Yet if "heterological" is heterological, then it correctly describes itself, and if it correctly describes itself, it is autological. Thus we have a term which is autological if it is heterological and heterlogical if it is autological, the standard paradoxical pattern.

The traditional way of attempting to resolve these paradoxes in modern logic has been through the development of theories of language which eliminate all self-reference. The classic approach is Russell's "theory of types," but I will consider a later and rather less cumbersome device due to Tarski. The basic idea is that we must distinguish various levels at which our language operates. The lowest level would include the language that we use to talk about the world around us, and the key point to recognize about language at this level is that it contains symbols for the various objects we wish to talk about, but does not contain those objects themselves.[2] The English language contains, for example, the word "table," which is a symbol for an object in the physical world. The word "table" is not itself a table, and the tables which occur in the physical world are not denizens of the English language; leafing through a dictionary I will find such words as "table," "typewriter," and "person," but even an unabridged dictionary contains no tables, typewriters or persons. Suppose, now, that instead of talking about the world around me, I wish to talk about my language. Tarski's suggestion is that in order to do this I need a new level of language (what is usually called a "metalanguage"). The metalanguage for a particular language will not contain any of the words of that language, but it will contain names for those words. The standard convention for forming the name of a printed word (a convention that I have already been following throughout this paragraph) is to put the word in quotation marks. Thus the symbol '"table"'

---

[2] No evaluative sense should be attached to the use of "lower" and "higher" in this context. Rather, the idea is that higher levels are built on lower levels much as the higher stories of a building are built on the lower stories.

has the same relation to the symbol "table" as the symbol "table" has to a particular physical table, i.e., in each pair the former item stands as a name for the latter. It may seem, at first glance, that the previous sentence has multiplied quotation marks beyond necessity, but reflection will show that this is not the case. In any case in which one is talking about a symbol, one needs a name for that symbol, and the name is conventionally formed by putting the symbol in quotation marks, with the usual alteration between double and single quotes. If I wish to say "'Table' is the symbol for a particular kind of physical object," I am making a statement about a symbol, and that symbol does not occur in my statement, but a name for that symbol does, just as when I say "My table is gray" no table occurs in my statement, although a symbol for it does occur. The statement "My table is gray," is a statement about a nonlinguistic object (my table) and the statement occurs in our language. The statement, "'Table' is a symbol for a particular kind of physical object" is a statement about a symbol, and this statement occurs in the metalanguage. Finally, the earlier example, "The symbol '"table"' has the same relation to the symbol "table" as the symbol "table" has to a particular physical table," makes a statement about the symbol which stands for a symbol, and thus itself occurs in a meta-metalanguage. Note also that this repetition of my earlier example introduces yet a further level of quotation marks. This occurs because I am now talking *about* that earlier sentence, and I must therefore ascend one step up the metalinguistic ladder and introduce a symbol which will stand for that sentence; I have done so by placing the earlier sentence in quotation marks.

Technical devices aside, the point of all of this is the thesis that no word can describe itself and no sentence can be about itself, but rather that to talk about a word or a sentence we must ascend to a higher order language. Clearly, this will eliminate the paradoxes. The first example that we considered,

A. Sentence A is false,

is simply incoherent on this view. In this example the letter "A" is being used as a name for a particular sentence; it therefore occurs in a metalanguage which is used to talk about that sentence, not in the language in which the sentence occurs, and thus

125

the symbol "*A*" cannot occur in the sentence in question. In general, if I wish to say that a particular sentence is true or false I must do so in a metalanguage which contains a symbol for that sentence, and on any linguistic level it is illegitimate, i.e., it is a violation of the rules of logical grammar, to say that a sentence at that level is true or false. Put simply, no sentence can refer to itself.

The same approach will (in a slightly more complex fashion) eliminate the problematic pair:

> *B.* Sentence *C* is true,
> *C.* Sentence *B* is false.

In this case we can note that "*B*" and "*C*" are names of sentences and must occur in metalanguages used to talk about those sentences. Thus sentence *C* cannot occur on the same level as *B*, but only in *B*'s metalanguage, and since *B* is on a lower logical level than *C*, *B* cannot make a statement about *C*; this requires another statement in *C*'s metalanguage. Similarly, if we begin by considering *B* to be a statement in *C*'s metalanguage, we find that *C* cannot be a statement about *B*. The upshot of this is that the pair is incoherent; it cannot occur in a logically correct language, and in such a language the paradox cannot be formulated.

Finally, the terms "autological" and "heterological" do not occur on the same linguistic level as the terms that they describe, but on a higher level, and if we wish to ask questions about the nature of these terms, we must do so on yet a higher level. In no case can a term either describe itself or fail to describe itself; all such descriptions of terms require a distinct linguistic level.

The levels of language thesis is best understood as a proposal. It is not a fact about language which one discovers in the way one discovers, say, the order in which subject and verb typically occur, or whether a language has separable prefixes. The levels of language thesis is rather the suggestion that we should think about language and the ways language operates from a particular point of view. As with any such proposal, it is to be evaluated on the basis of how well it does the jobs for which it has been introduced, and by whether it generates more problems, or more serious problems, than it solves. For while the levels of language thesis will eliminate the paradoxes, the question that must be

126

asked is whether it does so in such a sweeping manner that, as one recent writer has suggested, "It falls foul of the 'don't cut off your nose to spite your face' principle...[3] Now the fact that the proposal does violate this principle can be seen by noting that, in the course of banning all self-reference, the proposal eliminates many quite innocent cases. In addition, there seem to be situations in which self-reference is both demanded and yields important and profound results. I will consider each of these in turn.

To begin with, there are a wide variety of self-referential sentences which make perfectly good sense, and yield no paradox: "This sentence is in English," "This sentence is in black type," "This sentence contains five words," and even "This sentence contains forty-seven words." The last example is false, but straight forwardly nonparadoxical. Even "This sentence is true," which is odd in that it gives us no indication of how we can determine if it is true or false, yields no paradox: it is true if it is true and false if it is false. Similarly, pairs of sentences such as

> D. Sentence E is false,
> E. Sentence D is in English,

yield no paradox. Nor do we generate a paradox by asking whether "autological" is autological or heterological, since this term is autological if it is autological, and heterological if it is heterological. One begins to suspect, then, that the proposal to eliminate the paradoxes of self-reference through the elimination of *all* self-reference is akin to the proposal to eliminate crime through the summary execution of all human beings.

More importantly, there are disciplines which demand self-reference. The case that leaps to the mind in the present context is the study of language and its logic, for any adequate theory in this field must apply to *all* language, including itself. Consider, for example, the claim that any piece of language can only be discussed in a higher order metalanguage—at which level in the linguistic hierarchy does this claim sit? If we take it literally, it must be only a claim about all language levels up to, but not including, the level at which this claim occurs, but then, it is not

[3] Haack, *op. cit.*, p. 139.

a general claim about all language. If, however, it purports to be a general claim, and if we take it literally, then it cannot be properly formulated. Similarly, an adequate theory of knowledge, one which aims at an understanding of the nature and limits of all human knowledge, must include itself in its own scope. One of the persistent problems with theories of the limitations of human knowledge is that they are often so restrictive as to entail that they themselves could never have been discovered.[4] And, to take one more example, attempts to understand the workings of the human mind and brain are intrinsically self-referential in that they constitute attempts by that mind and brain to understand themselves. With this last example we have clearly moved beyond our original concern with self-referential linguistic structures, but the observation that self-reference is a pervasive feature of knowledge suggests that linguistic self-reference is only one special case of a general feature of human thought, and this should serve to make us thoroughly sceptical about the desirability of seeking to eliminate self-referential language.

Third, the systematic use of self-reference has consistently yielded results of exceptional profundity and beauty. A particularly apt example comes from modern logic itself, for one of the deepest and most revolutionary outcomes of modern logic, the Gödel theorems, are thoroughly dependent on self-reference. There are two theorems that Gödel published in 1931 which have totally transformed our understanding of the nature of deductive systems. One of them states that no formal system sufficiently powerful to contain ordinary arithmetic can be proven to be absolutely consistent, i.e., any proof of consistency of a particular system must be carried out in a metasystem which is logically more sophisticated than the system in question, and whose absolute consistency has, therefore, not been established. (Note that this does not say that arithmetic is not consistent, but only that we cannot prove its consistency). The second theorem says that any system powerful enough to include arithmetic is incomplete in the following sense: there are theorems which are true given that system, but which cannot be proven in that

128

system. Now at two key stages the proofs themselves depend on the use of self-reference. I will not attempt to reproduce, or even outline the proofs here,[5] but I do want to describe the two points at which self-reference occurs in those proofs.

Gödel's consistency theorem is the exact opposite of the result he was aiming at, i.e., Gödel's original goal was to prove the absolute consistency of arithmetic. Gödel believed that he had found a way around the problem of the need for a logically stronger metalanguage in which to investigate arithmetic in that he developed a way of saying everything that he wanted to say about arithmetic *in* arithmetic itself. This involved associating a unique number with each statement about arithmetic and thus making arithmetic *its own metalanguage*; this is the first self-referential step. The second such step occurs in the way in which the completeness result is proven. In essence, Gödel constructs a sentence which says that the sentence associated with a particular number cannot be proven in the system, and on checking, it turns out that the number in question is just the number of that sentence. In other words, Gödel constructs a self-referential sentence which declares its own unprovability.

It should be clear at this point that although self-reference does, in some cases, yield troubling paradoxes, it is a fundamental and pervasive feature of human thought. Indeed, this holds not only in logic, mathematics, philosophy, and science, but in art as well; we will begin to examine the role of self-reference in art in the next section.

II

Douglas Hofstadter's *Gödel, Escher, Bach* provides an extended meditation on the varieties of self-reference. The central concept of the book is what Hofstadter calls a "strange loop": "The 'strange loop' phenomenon occurs whenever, by moving upwards (or downwards) through the levels of some hierarchical

---

[5] For excellent detailed discussions of Gödel's extremely complex proof see Ernest Nagel and James R. Newman, *Gödel's Proof* (New York: New York Univ. Press, 1958), and Douglas R. Hofstadter, *Gödel, Escher, Bach* (New York: Basic Books, 1979). For those who wish to tackle the original paper see Kurt Gödel, *On Formally Undecidable Propositions* (New York: Basic Books, 1962).

129

system, we unexpectedly find ourselves right back where we started."[6] Hofstadter also describes a structure in which strange loops occur as an "angled hierarchy." Strange loops occur whenever we have self-reference, and the occurrence of the strange loop is a major source of the air of depth and mystery that self-reference brings along with it. Let us see just how strange loops appear in some of the examples that we have already discussed.

Self-referential sentences provide a strikingly unusual use of language, for normally we use our language to talk about the world around us and our remarks direct the hearer or reader to various aspects of that world. The self-referential sentence, on the other hand, takes the reader no place. Even in as innocuous a case as "This sentence is in English" the sentence seems to direct the reader somewhere just as "*Mulligan Stew* is in English" does. But while the latter sentence leads us to Sorrentino's novel, the former leads us right back where we started, i.e., to the sentence itself; this is a simple strange loop. A slightly more complex example is provided by pairs of inter-referential sentences. Recall that in discussing how the language-metalanguage distinction would serve to eliminate the paradoxical pair that we examined earlier, it was argued that if we begin with the first sentence then we must take it to be in the metalanguage of the second sentence, but that beginning in the reverse order yields a reverse result as to which sentence is in the lower and which in the higher level language. The pairs seem to have a hierarchical relation, but in whichever way we travel through the hierarchy we end up with both sentences on the same level. Rather than taking this as a reason for rejecting the pair as nonsense, we can now see it as exemplifying a strange loop. Similarly, the structure of the proof of Gödel's incompleteness theorem is that of a strange loop: Gödel sets out to do metamathematics, i.e., to develop a theorem about arithmetic. In order to develop this theorem he finds it necessary to move down a level by reflecting the language in which he talks about arithmetic into arithmetic itself, and after working to a result in arithmetic, finds himself back on the level of metamathematics with a theorem about arithmetic. It is, perhaps, no coincidence that this procedure yielded

6 Hofstadter, *op. cit.,* p. 10.

130

a genuinely surprising result, quite at variance with traditional views of the nature of mathematics.

In the course of his book Hofstadter provides many other examples of such strange loops, e.g., a mind introspecting on itself and a government agency investigating itself, but the most sustained examples (other than those from Gödel) are taken from the other two figures mentioned in his title. In the case of Bach's music, the structure of the canon and the fugue involve the continual return of a theme in other voices, other keys, or playing other roles, such as that of harmony. Perhaps the most striking example of a musical strange loop that Hofstadter discusses is the "Canon per Tonos" from *The Musical Offering*; Hofstadter describes this piece as "an endlessly rising canon," and he uses it to introduce the idea of a strange loop:

> The uppermost voice sings a variant of the Royal Theme, while underneath it, two voices provide a canonic harmonization based on a second theme. The lower of this pair sings its theme in C minor (which is the key of the canon as a whole), and the upper of the pair sings the same theme displaced upwards in pitch by an interval of a fifth. What makes this canon different from any other, however, is that when it concludes—or, rather, *seems* to conclude—it is no longer in the key of C minor, but now is in D minor. Somehow Bach has contrived to modulate (change keys) right under the listener's nose. And it is so constructed that this "ending" ties smoothly onto the beginning again; thus one can repeat the process and return in the key of E, only to join again to the beginning. These successive modulations lead the ear to increasingly remote provinces of tonality, so that after several of them, one would expect to be hopelessly far away from the starting key. And yet magically, after exactly six such modulations, the original key of C minor has been restored![7]

Escher's drawings provide equally powerful examples of strange loops. In Waterfall[8] we have a picture of water falling from an upper level of a structure to a lower level. The water that has fallen runs off in a series of moderately winding channels. Each

[7] *Ibid.*
[8] *Ibid.*, p. 11. References to Escher drawings are to reproductions in *Gödel, Escher, Bach*, although they are available in many other places, e.g., *The Graphic Work of M. C. Escher* (New York: Ballantine Books, 1971).

131

section of the channel, taken by itself, seems perfectly capable of being part of a system used to carry off the water that has fallen, but when we follow the entire channel we find, surprisingly, that we have returned to the upper level of the structure and that the runoff channel is supplying the water to the waterfall. Similarly, in *Ascending and Descending*[9] we see two rows of cowled figures moving around an open balcony. One row is moving clockwise and is, at each stage, ascending a flight of stairs; the other group is moving counterclockwise and is descending the same flight of stairs. But whichever group we follow, we find that after a complete circuit of the balcony the figures are back at exactly the same height from which they began. Perhaps the archetypal example of a strange loop is Escher's *Drawing Hands*.[10] This drawing shows two hands, each holding a pencil and each drawing the other. If ever there was a clear example of a two level hierarchical structure one would think that it would be the relation between the artist and his drawing. The drawing is created by the artist, and in this sense is dependent on him, although the drawing exists apart from the artist once it has been completed. But in *Drawing Hands* we begin with an artist, move to the drawing, and then find ourselves back at the level of the artist. Each hand is both creator and creation, each is equally dependent on the other. and whichever way we follow the loop, we end up on the same level from which we began—much as occurred in the case of the paradoxical pair of inter-referential sentences.

The relation beteen the work of Gödel, on the one hand, and that of Escher and Bach, on the other, is equally intricate in Hofstadter's book. From one point of view the Gödel theorems and their significance are the central strand of the book, and Escher and Bach are used as means of illustrating ideas from formal logic; yet from another point of view it is the work of Escher and Bach that is central and the extended discussions of Gödel provide a commentary on the two artists. To see why this is the case, I must say a bit more about the impact of the Gödel theorems and the significance that Hofstadter wishes to draw out of them.

[9] Hofstadter, *op. cit.*, p. 12.
[10] *Ibid.*, p. 690.

Hofstadter's book is ultimately a meditation on the nature of consciousness and on the possibility of artificial intelligence, i.e., on the possibility of programming computers that are fully intelligent in the way in which humans are. Now several philosophers have maintained that one of the implications of Gödel's incompleteness theorem is that no such artificial intelligence is possible. The basic line of argument is that there is a crucial sense in which a computer program is equivalent to a deductive system, and since Gödel's incompleteness theorem shows that every deductive system (powerful enough to contain arithmetic) is incomplete, there will be intellectual tasks that any given computer program will be unable to carry out. The argument goes on to maintain that these will be tasks that human beings are capable of carrying out, and thus that the cognitive capacities of a human being must go beyond those of a computer. One example of such a situation might be the discovery of the Gödel incompleteness theorem itself. Hofstadter is not convinced by this line of argument, and he develops an extensive response to it, although any attempt to describe his counter-argument will take us too far afield and multiply the length of this paper beyond any reasonable bounds (I have not even stated the full argument against artificial intelligence, but only given an outline of it). What does bear mention here is that one strand of Hofstadter's response derives from a discussion of the relation between conscious, intelligent, creative thought and the underlying brain structure on which it depends, and it should not be surprising at this point to find that Hofstadter maintains that this relation is that of a tangled hierarchy. For while consciousness and its attendant phenomena are high level effects which are ultimately, through a series of levels, dependent on the physical structure and activity of our brains and nervous system, once consciousness has emerged, it is capable, through such capacities as reflection and choice, of affecting what is going on in its own physical substrate. Moreover, self-referential structures, rather than being merely a troublesome and eliminable source of logical paradox, are, for Hofstadter, essential to the very nature of mind:

My belief is that the explanations of the "emergent" phenomena in our brains—for instance, ideas, hopes, images, analogies, and finally consciousness and free will—are based on

133

> a kind of Strange Loop, an interaction between levels in which the top level reaches back down towards the bottom level and influences it, while at the same time being itself determined by the bottom level.[11]

We can now understand the sense in which a discussion of the implications of Gödel's incompleteness theorem provides the basis for a commentary on the work of Escher and Bach. For if self-reference is part of the very structure of human consciousness and creativity, then we can begin to grasp the peculiar fascination of self-referential art, since such art will reflect the structure of the consciousness which created it. Hofstadter develops this idea in the case of drawing and music, but there is a surprising lack of examples taken from literature, and this observation, after only one more brief detour, will take us to Sorrentino.

III

Literature provides a fertile source of examples of strange loops. Cervantes, for example, has Don Quixote encounter a copy of the supposed second part of the novel written by Avellaneda. The Don offers a number of criticisms of the apocryphal version of the novel and then alters his intention of taking part in the tournament at Saragossa. For the imitation Don has been in Saragossa, and Quixote reasons that if he never goes to Saragossa, this will demonstrate the falsity of Avellaneda's *Quixote*.[12] Later the Don meets a character from the apocryphal version and has this character sign a notarized deposition stating that the true Don Quixote, whom he has now met for the first time, is not the same as the one mentioned in Avellaneda's version.[13] Similarly, in *The Floating Opera,* Barth cuts across levels when he has Todd Andrews speak directly to the reader, suggesting that the reader peruse the story of the resumption of Andrews' affair with Jane Mack while Andrews goes to the men's room.[14] Borges is a

[11] *Ibid.,* p. 709.
[12] Miguel de Cervantes Saavedra, *Don Quixote*, trans. Walter Starkie (New York: Signet Books, 1964), part II, ch. 59.
[13] *Ibid,* part II, ch. 72.
[14] John Barth, *The Floating Opera* (New York: Avon Books, 1956), ch. 17.

134

master of level crossing, as is Pirandello, but in *Mulligan Stew* the technique is used at unprecedented length and the result is a novel which exemplifies a tangled hierarchy. To see why we must examine the structure of Sorrentino's tale.

The central character of the book is Antony Lamont, an *avant garde* novelist working on his fifth novel, initially titled *Guinea Red* although the title is later changed to *Crocodile Tears*, and most of what we are presented with in *Mulligan Stew* is filtered through Lamont. Sorrentino gives us chapters of Lamont's novel as they are written, including two versions of the first chapter, excerpts from Lamont's notebook, from his earlier writings, and from Lamont's side of his correspondence with various people. In addition, we get, through Lamont, a number of items written by other people. Many occur in Lamont's scrapbook, which seems to be a compilation of newspaper and magazine clippings, mail-order advertisements, and such. We also get some excerpts from the writings of Dermot Trellis, Lamont's archrival and the husband of Lamont's sister Sheila, and a small volume of poems, "The Sweat of Love," by Lorna Flambeaux. All of these items come through Lamont, although he never speaks as narrator of the book. Lamont's novel does have an explicit narrator, Martin Halpin, and all events in that novel are filtered through Halpin. Lamont's novel, like Sorrentino's, also contains a certain amount of material written by people other than Halpin, but we only get to see this material because Halpin offers it to us. This includes the correspondence between Ned Beaumont, Halpin's business partner, and Beaumont's lover, Daisy, and a number of mail-order advertisements for "He-man's Hot Photos" which would be quite appropriate for Lamont's scrapbook. Note also that Halpin and Beaumont are partners in a publishing business, surely not a random choice on Sorrentino's part.

In addition, Halpin functions as an explicit narrator in *Mulligan Stew*, for Sorrentino treats Halpin, Beaumont, and all of the characters in Lamont's novel as quasi-independent individuals who are employed by Lamont in writing his novel, somewhat in the way in which actors are hired to take a role in a play. The characters seem to have a life of their own quite apart from their jobs in Lamont's novel, and they are free to leave their jobs if they choose to do so. As long as these characters are actively

135

being employed in the novel, they are at the author's mercy and must do what he directs them to, and say whatever he puts in their mouths, even if they find it repulsive or embarassing. To refuse to do as directed is tantamount to running out on their jobs, and they are hesitant to do this because of the dire consequences it will have for their careers as novel characters. But these characters do have the power to leave a novel, and to this extent at least, the author is dependent on their good will. If an author's characters leave he will be unable to continue his work, just as Hubert in Queneau's *The Flight of Icarus* is unable to continue writing when he comes into his study and finds that his hero Icarus has fled.[15] The world in which the characters of Lamont's novel live when they are not working is wholly made up of characters from novels—Halpin and Beaumont meet several characters from other novels in the course of their adventures— and it is also furnished with scenes and objects from novels. From this character world we are presented with Halpin's journal, a journal that is written by the independent Halpin, that is not a part of Lamont's novel, and that provides, among other things, a vehicle for Halpin's own commentary on the novel in which he is being employed, on the relations between characters, novels and novelists, and on novels and novelists generally. In addition, Halpin's journal too contains a good deal of material written by other people which Halpin has chosen to include.

Characters and their world are, however, only quasi-independent. In the first place, while characters may have a degree of independence once they have been brought into existence by a novelist, their existence and experiences seem to be determined by the various novelists who employed them. Martin Halpin reminds us on several occasions that Lamont borrowed him from Mr. Joyce, and Halpin does not seem to have had any existence or any experience other than that provided by Joyce and Lamont:

> "An old gardener," so I have been for these thirty-odd years, an old gardener who has never gardened, never even seen, so far as I can remember, a garden—and happy *not* to have

[15] Raymond Queneau, *The Flight of Icarus*, trans. Barbara Wright (New York: New Directions, 1973).

136

seen one, by God! What a delight to reside in that quietly monumental world all this time, a small part of it, content behind the letters that form my name. In a way, I *was* the letters, no more.[16]

When Lamont uses Halpin in an orgy scene, the independent Halpin is thoroughly overwhelmed and delighted, this being his first sexual experience. Mr. Joyce has not provided him with any such experience, and although we are given no reason to believe that the Halpin of Lamont's novel has been similarly deprived, the character Halpin cannot have engaged in any sexual adventures until some writer has described him as doing so. Characters and their world seem, then, to be constituted only by what their employers have explicitly described. The main setting of Lamont's novel is a cabin in the woods in which Beaumont is dead and Halpin is waiting for the police. But when they are operating outside of their personae in Lamont's novel, Halpin and Beaumont both comment that, of course, the police are not on their way, since, in the novel, Lamont has never actually had Halpin make the call, and Halpin could not have made any such call since Lamont has not remembered to put (i.e., explicitly mention) a telephone in the cabin. The cabin initially is made up only of those parts which Lamont has described. This consists of two specific rooms and, since Lamont has hinted that there is more, other parts which, in the character world, are literally vague. Halpin even discovers, when he does leave the cabin, that its remainder is vague from the outside as well as from the inside. When, later in *Mulligan Stew*, Halpin and Beaumont find other parts of the cabin which seem not to have been there before, Beaumont suggests that the whole cabin exists in some other novel from which Lamont is stealing.[17]

This limitation of the characters' own world carries over to the situations that Halpin and Beaumont encounter when, exercising the freedom which they do possess, they decide to leave the cabin temporarily while Lamont is not working. The trees they

[16] Gilbert Sorrentino, *Mulligan Stew* (New York: Grove Press, 1979 London: Marion Boyars Publishers Ltd., 1980), p. 25.
[17] *Ibid.*, p. 241.

find outside of the cabin are typical, nondescript, generic trees, and the town they go to is not a complete town:

> Whatever seemed necessary to the builders or inhabitants had been built, and the rest forgotten or ignored. One street was lined with houses, all complete, and another, shady and well-paved, had on it nothing but a barbershop. On one small lane there was nothing but a front porch, a lawn before it, and a shade tree over all. The rest of the house was nowhere in evidence, yet the porch was complete with wicker furniture, lamps, a glider, even strips of fly paper hung from the rafters. A yellow insect light burned dully in the brilliant sunshine.[18]

The town exists "in a typescript locked away in a Poughkeepsie attic,"[19] and it is suggested that the typescript and trunk are themselves a part of the real world that exists apart from any novels and in which, ultimately, all novels exist. Thus as soon as Halpin steps out of the night scene of Lamont's novel he finds himself in bright sunshine—the same bright sunshine in which the porch with its insect light is located—and this real world is, at least initially, presented as one with which characters cannot directly interact. The sun does not change position during the time that Halpin spends out of the cabin, and in spite of the brilliance of the sun, Halpin casts no shadow. This disturbs him at first, but he reflects that, "After all, I am not real."[20] Still, it might be that the sunlight and the stationary sun exist only in some other novel through which Halpin passes on his way from the cabin to the trunk, for the sun certainly does exist in a novel (Sorrentino's), and Beaumont, who is generally portrayed as much more experienced and knowledgeable than Halpin, insists that characters are just as real as their authors.[21]

The characters that Halpin meets on his trips outside the cabin are also products of the novels in which they have worked. Clive Sollis, for example, a character from one of Lamont's earlier books, indeed, a character who existed only in the protagonist's

---

18 *Ibid.*, p. 153.
19 *Ibid.*, p. 154.
20 *Ibid.*, p. 152.
21 *Ibid.*, p. 242.

138

mind,[22] is a hopeless alcoholic because Lamont made him one. Sollis can hate what has been done to him and despise Lamont, but he has no more power to stop drinking without some novelist directing him to do so than Halpin has to have a sexual experience or see a garden without a novelist writing it down. And, unlike actors hired for a play, characters do not have names other than those that their authors gave them.

In spite of these limitations, characters have a substantial degree of freedom and independence. Halpin and Beaumont have the option of leaving the setting in which Lamont has placed them and travelling into town where they can meet characters from other novels. There is the risk that Lamont will come back to work and find them gone, but they are able to use Lamont's own limitations to protect themselves from this possibility. For while Beaumont is dead in Lamont's novel, Lamont keeps forgetting exactly where he put the body, and he has never given a detailed description of either of his two main characters. Presumably, then, Lamont does not know what these characters look like, or just where the body lies; if only one of them leaves at a time, and if the one who remains plays the part of the living Halpin, there is a good chance that, even if Lamont does come back to work, he will not notice that one of his characters is missing.[23] Further, while the characters are dependent for their situations and experiences on Lamont, they can act back on the author, as one would expect in a tangled hierarchy. They do have the option of quitting, an option which they eventually exercise, and presumably once they have gone, Lamont will be unable to continue writing without them. Moreover, the characters are not totally without ability to control their own destinies in the novel. Rather, Halpin maintains that there exists one situation, when an author decides to rewrite an entire section, in which the characters can influence the direction of a novel:

> When that occurs, there is a moment, *just* a moment, when one may assert oneself. Occasionally, the writer will allow this assertion to stand, and one's character is thereby subtly

[22] *Ibid.*, p. 4.
[23] *Ibid.*, p. 153.

139

changed. In fact, sometimes the whole world in which one is employed is changed.[24]

In addition, there are situations in which Sorrentino leaves it unclear in which direction the influence is flowing. I have already noted that Beaumont and Halpin, in their free time, find rooms and items in the cabin which have not been described by Lamont; but later in *Mulligan Stew* Lamont does write a chapter in which Halpin discovers exactly these rooms, and experiences the same surprise and the same feeling—that the rooms were not there before—that the character Halpin experienced when he discovered those rooms.[25] In the same chapter, which Lamont writes after the character Ned Beaumont walks out on the novel, Halpin finds that Beaumont's body has disappeared. And Lamont does eventually lose complete control of the novel and winds up denying that he has written a chapter which he finds in his study.[26]

There is another aspect of the characters' freedom that is of particular interest to us: they are capable of holding quite independent opinions about the quality of their authors' work, and of freely expressing those opinions in their own world. Beaumont and Halpin consider Lamont's story, characterizations, and use of language to be utterly abominable, and Halpin's journal contains a running commentary on Lamont's writing. The vague houses, incompleted towns, inconsistent chapters, and generic trees that novel characters encounter, and which are detailed in Halpin's journal, provide further commentary not only on Lamont, but on the lapses of authors generally. Virtually all of the characters that we meet through Halpin's journal have some choice remarks on the foibles of their authors, a process that culminates when Halpin and Beaumont meet three characters currently employed in Dermot Trellis' new novel. These characters offer an extended catalog of authorial clichés and oversights:

> What about the fellow who winced no matter what he said?
> Or the one who sighed all the time.

---

[24] *Ibid.*, p. 27.
[25] *Ibid.*, pp. 409-414.
[26] *Ibid.*, p. 400.

140

> I once worked a job where a young woman had to mastur-
> bate me—she neglected to open my trousers! A frightful mess.
> She laughed and laughed until *she* began having to go to bed
> every night with her shoes on.
>
> I knew a poor wench who was always forced to remove her
> stocking without first having to remove her shoes. She got
> an ulcer, I believe.
>
> The revolver with fifty bullets.
>
> Of course! But what of the pen that in the middle of the
> letter becomes a pencil?
>
> The steamship that is becalmed?
>
> Personally, I've *shaken* so many Martinis that I can't even
> look at one anymore. Absolutely barbarous.[27]

And so on for several hours.

We can now see the full development of the tangled hierarchy
structure of *Mulligan Stew*. To begin with, the novel is a satire
on *avant garde* novels, and as such, it must be an example of the
very sort of work that it is making fun of. Satirical novels can
be aimed at a wide variety of topics, but an author who under-
takes to write a novel that satirizes a novel, must produce a
self-referential structure. The task, we might note, is particularly
arduous since it requires that the authors exaggerate the foibles of
the type of novel being satirized and thus, if he is to succeed,
write a good novel by writing an extreme version of a bad novel.
Moreover, Sorrentino presents us with what could be a linear
hierarchical structure: Lamont is Sorrentino's character, Halpin
and Beaumont are Lamont's characters, and Clive Sollis is a
creation of one of Lamont's other characters. In such a structure
we normally find that the author tells us about his characters
but the characters do not turn back and comment on the author—
the latter, however, is exactly what is happening throughout
Mulligan Stew. Halpin, in particular, functions as narrator on
two distinct levels, that of Lamont's novel, and that of Sorren-
tino's. And Sorrentino develops this tangled hierarchy yet an-
other step when he begins the book with a set of rejection letters
from editors of various publishing houses. From the point of view
of a system of distinct levels and metalevels it makes no sense

[27] *Ibid.*, p. 272. The masturbation incident occurs in Sorrentino's own novel
Imaginative Qualities of Actual Things.

141

for a set of comments on a book to form a part of that very book, but we have seen that there is no reason to accept the demand that we keep our levels distinct, and it is exactly because of the looping structure of *Mulligan Stew* that those letters stand as an integral part of the novel. Indeed, we should recall here that letters also play a central role in the presentation of Lamont's life in *Mulligan Stew* as well as in Lamont's novel, and that Beaumont and Halpin are publishers. On the whole, then, *Mulligan Stew* presents us with a multi-level tangled hierarchy in which every level interacts with every other, and in which, in the course of attempting to move carefully from one level to the next, we can suddenly find ourselves on any level at all.

There is one more aspect of *Mulligan Stew* that should be addressed in the present context. I submit that this is an outstanding novel, and it is so against overwhelming odds. The characters in the book are, for the most part, bad writers, and a large part of the text of *Mulligan Stew* consists of their writings. As a result, taken piece by piece, *Mulligan Stew* contains a depressing proportion of overwhelmingly bad prose—but still the novel works and the bad prose contributes to its working. Part of the explanation for that must lie in the structure of the novel. Now if Hofstadter is right in suggesting that the structure of consciousness is itself that of a tangled hierarchy, and if a part of the fascination of works of art which exhibit this structure is that they reflect the nature of our consciousness, then one can begin to grasp how Sorrentino was able to forge a good novel out of bad writing, and appropriately admire the skill that he lavished in having his characters write so poorly.[28]

<div align="right">

Harold I. Brown
(*Northern Illinois University*)

</div>

142