Reconceptualizing the Theory and Generic Scope of Unreliable Narration

Ansgar F. Nünning

DOI: 10.4000/books.pufr.3950
Publisher: Presses universitaires François-Rabelais
Place of publication: Presses universitaires François-Rabelais
Year of publication: 1999
Published on OpenEdition Books: 1 June 2017
Serie: GRAAT
Electronic ISBN: 9782869064720

Electronic reference

This text was automatically generated on 10 May 2019. It is the result of an OCR (optical character recognition) scanning.
Ever since Wayne C. Booth first proposed the unreliable narrator as a concept, it has been considered to be among the basic and indispensable categories of textual analysis. Hardly anyone to date has modified or challenged Booth's well-known formulation, which has become the canonized definition of the term: "I have called a narrator reliable when he speaks for or acts in accordance with the norms of the work (which is to say the implied author's norms), unreliable when he does not" (1961, 158-59). According to Booth, the distinction between reliable and unreliable narrators is based on "the degree and kind of distance" (155) that separates a given narrator from the implied author of a work. A comparison of the definitions provided in standard narratological works, in scholarly articles, and in glossaries of literary terms shows that the great majority of narratologists have followed Booth, providing almost identical definitions of the unreliable narrator.

What most critics seem to have forgotten, however, is that Booth himself freely admitted that the terminology for "this kind of distance in narrators is almost hopelessly inadequate" (1961, 158). There is indeed a peculiar discrepancy between the importance generally attributed to the question of reliability in narrative and the unresolved issues surrounding the concept of the unreliable narrator: "There can be little doubt about the importance of the problem of reliability in narrative and in literature as a whole.... [But] the problem is (predictably) as complex and (unfortunately) as ill-defined as it is important" (Yacobi, 1981, 113). Booth's canonical definition does not really make for clarity but rather sets the fox to keep the geese, as it were, since it falls back on the ill-defined and elusive notion of the implied author, which hardly provides a reliable basis for determining a narrator's unreliability.

The thesis of this article is that the concept of the unreliable narrator needs to be radically rethought because, as currently defined, it is terminologically imprecise and theoretically inadequate. The postulation of essentialized and anthropomorphized
entities designated "unreliable narrator" and "implied author" ignores the complexity of the phenomena involved and stands in the way of a systematic exploration of the cognitive processes which result in the projection of unreliable narrators in the first place. It would arguably be more adequate to conceptualize unreliable narration in the context of frame theory as a projection by the reader who tries to resolve ambiguities and textual inconsistencies by attributing them to the narrator's "unreliability". In the context of frame theory, the invention of "unreliable narrators" can be understood as an interpretative strategy or cognitive process of the sort that has come to be known as "naturalization." 1 Paraphrasing one of Malcolm Bradbury's observations in Mensonge, his hilarious satire on deconstruction, I would like to suggest that it is high time to dismantle many of the premises of realist theories of unreliable narration, and to convert the foundations of some others.

The first part of the article is devoted to giving an assessment and critique of the standard notions of the unreliable narrator. The second part outlines a radical reconceptualization of unreliable narration. It is argued that a number of empirical frames of reference and literary models serve as the modes of naturalization by means of which readers (like critics and most theorists of unreliable narration) account for contradictions both within texts and between the world-model of texts and their empirical world-models. The third part gives a brief outline of the generic scope of unreliable narration, arguing that it is unjustifiable and counter-productive to limit the study of this phenomenon to narrative fiction. The final section will then provide a brief summary and suggest that much more work needs to be done in this particular field of narratology.

I. A critique of conventional theories of unreliable narration

A brief look at conventional accounts of the concept "unreliable narrator" may be in order so as to distinguish the approach argued for in this essay from the general approach in narratology. Let us begin by asking just what it is that we know about the mysterious unreliable narrator and by presenting a critique of traditional theories of unreliable narration against the background of five hypotheses.

The definition provided by Gerald Prince in his Dictionary of Narratology will suffice to indicate what is usually meant by the term "unreliable narrator": "A narrator whose norms and behavior are not in accordance with the implied author's norms; a narrator whose values (tastes, judgments, moral sense) diverge from those of the implied author's; a narrator the reliability of whose account is undermined by various features of that account" (Prince, 1987, 101). Despite the good job Prince does in summarizing the communis opinio on the subject, this definition of the concept comprises an unholy mixture of vagueness and tautology. Nonetheless, most theorists and critics who have written on the unreliable narrator take the implied author both for granted and for the only standard according to which unreliability can be determined.

One of the central problems in defining unreliable narration is the unresolved question of what standards allow the critic to recognize an unreliable narrator. The usual answer to the question "Unreliable, compared to what?" is woefully inadequate and untenable, because it specifies just one basis for recognizing the narrator's unreliability, namely the
ill-defined concept of the implied author. The trouble with all of the definitions that are based on the implied author is that they try to define unreliability by relating it to a concept that is itself ill-defined and paradoxical. Curiously enough, however, even the most sophisticated recent articles retain the notion of the implied author. In what are arguably the best critiques of orthodox theories of unreliable narration to date, the articles of Tamar Yacobi (1981, 1987) and Kathleen Wall (1994), the authors hold on to the implied author as though he, or rather it, was the only possible way of accounting for unreliable narration.

Critics who argue that a narrator's unreliability is to be gauged in comparison to the norms of the implied author just shift the burden of determination onto a critical catch-all term that is itself notoriously ill-defined. The tenacity with which narratologists have clung to the implied author in their attempts at defining unreliability suggests, as Mieke Bal (1981b, 209) observes, that the implied author is "a remainder category, a kind of passepartout that serves to clear away all the problematic remainders of a theory." Introducing the implied author has certainly not managed to clear away the problems of defining unreliable narration.

Some narratologists have pointed out that the concept of the implied author does not provide a reliable basis for determining a narrator's unreliability. Not only are "the values (or 'norms') of the implied author... notoriously difficult to arrive at," as Rimmon-Kenan (1983, 101) observes, but the implied author is itself a very elusive and opaque notion. One might go much further than Rimmon-Kenan and suggest that the implied author's norms are impossible to establish and that the concept of the implied author is dispensable.

From a theoretical point of view, the concept of the implied author is also problematic because it creates the illusion that it is a matter of a purely textual phenomenon. But it is obvious from many of the definitions that the implied author is a construct established by the reader on the basis of the whole structure of a text. When Chatman (1990, 77) writes that "we might better speak of the 'inferred' than of the 'implied' author," he implicitly concedes that one is dealing with something that has to be worked out by the reader. Being a structural phenomenon that is voiceless, the implied author must be seen "as a construct inferred and assembled by the reader from all the components of the text" (Rimmon-Kenan, 1983, 87). Toolan has made the sensible suggestion that one should look at the implied author not as a speaker but as a component of the reception process, as the reader's idea of the author: "The implied author is a real position in narrative processing, a receptor's construct, but it is not a real role in narrative transmission. It is a projection back from the decoding side, not a real projecting stage on the encoding side" (1988, 78).

The most controversial aspect of the concept of the implied author is that it carries far-reaching, though largely unacknowledged theoretical implications. First, the concept of the implied author reintroduces the notion of authorial intention, though through the backdoor. As Chatman (1990, 77) has pointed out, "the concept of implied authorship arose in the debate about the relevance of authorial intention to interpretation." Providing "a new link to the sphere of the actual author and authorial values" (York, 1987, 166; cf. Yacobi, 1987), the implied author turns out to be little more than a terminologically presentable way of making it possible to talk again about the author's intention: "The concept of 'the implied author', with its air of being an inference from the work and thus as it were, like plot, an objective feature of the work, enables Booth to talk about the author under the guise of still appearing to talk about the work" (Baker,
1972/73, 204f.; cf. also Juhl, 1980, 203). Second, representing the work's norms and values, the implied author is intended to serve both as a yardstick for a moralistic kind of criticism and as a check on the potentially boundless relativism of interpretation. Third, the use of the definite article and the singular misleadingly suggest that there is only one correct interpretation: "The very fact that Booth and Chatman speak of the implied author already implies, suggests the existence of one ideal interpretation of the narrative text" (Berendsen, 1984, 148). In short, the concept of the implied author appears to provide the critic again with a basis for judging both the acceptability of an author's "moral position," about which, according to Booth, a writer "has an obligation to be as clear [...] as he possibly can be" (1961, 389), and the correctness of an interpretation.

12 The lack of terminological clarity and the problematic theoretical implications associated with the notion of the implied author have led some narratologists to argue that the concept should be abandoned. Some theorists have recognized that it has not fulfilled the promise "to account for the ideology of the text" (Bal, 1981a, 42) and is not capable of doing what it was supposed to do: "It not only adds another narrating subject to the heap but it fails to resolve what it sets out to bridge: the author-narrator relationship" (Lanser, 1981, 49f.).

13 Whether or not narratology is really well served with such a problematic concept as the implied author, be it of the personalized and anthropomorphized or the depersonified variety, is an open question. Recently, some prominent narratologists have again emphatically come out in favour of the implied author, while others have argued just as strongly against the concept. But given the fact that phenomena like norms and values, structure, and meaning are central problems in literary criticism and will continue to occupy the attention of theorists and critics alike, they probably should not be allowed to disappear behind a concept like the implied author, which is ill-defined and potentially misleading. As I hope to show below, the implied author is neither a necessary nor a sufficient standard by which to determine a narrator's putative unreliability.

14 Despite what common sense would appear to tell us, definition is a problem with the unreliable narrator because most theories leave unclear what unreliability is and whether it involves moral or epistemological shortcomings. Most definitions in the wake of Booth have emphasized that unreliability consists of a moral distance between the norms of the implied or real author and those articulated by the narrator while other theorists have pointed out that what is at stake is not a question of moral norms but of the veracity of the account a narrator gives (cf. Toolan, 1988, 88).

15 In most work on the unreliable narrator, it is also unclear whether unreliability is primarily meant to designate a matter of misrepresenting the events of the story or whether it consists of the narrator's dubious judgments or interpretations. Rimmon-Kenan's (1983, 100) definition is a case in point. She simply leaves open whether unreliability is to be gauged in comparison to the accuracy of the narrator's account of the story or to his or her commentary and judgments: "An unreliable narrator... is one whose rendering of the story and/or commentary on it the reader has reasons to suspect." The "and/or"-construction sounds very open and flexible, but this is a bit too nonchalant. Most would agree that it does make a difference whether we have a deviant narrator who provides a sober and factually veracious account of the most egregious or horrible events, which, from his point of view, are hardly noteworthy, or a normal narrator who is just a bit slow on the uptake and whose flawed interpretations of what is going on reveal that he is a benighted fool. Lanser (1981, 170ff.) provides an answer to
the question of how we may classify narrators "with respect to 'reliability'" by positing three axes between the poles "dissimulation vs. honesty", "unreliability vs. reliability" and "narrative incompetence vs. narrative skill.

Conventional theories of unreliable narration are methodologically unsatisfactory as well because they either leave unclear how the narrator's unreliability is apprehended in the reading process or they provide only highly metaphorical and vague explanations of it. The metaphors that Chatman uses in order to explain how the reader detects the narrator's unreliability are a case in point. He resorts to what is arguably one of the two most popular metaphors in this context — that of "reading between the lines." Chatman (1978, 233) argues that readers "conclude, by 'reading out,' between the lines, that the events and existente could not have been 'like that,' and so we hold the narrator suspect." Leaving aside for the moment that the repeated use of inverted commas in definitions is not particularly reassuring, I just wish to suggest that such observations fail to shed much light on how a narrator's unreliability is apprehended in the reading process.

The second metaphor that critics and theorists continually employ in order to account for unreliable narration is that something is going on "behind the narrator's back" (cf. Riggan, 1981, 13; Yacobi, 1981, 125). Chatman (1978, 233), for instance, suggests that the implied author establishes "a secret communication with the implied reader. Riggan (1981, 13) not only uses almost exactly the same phrase, but he also states quite unequivocally that "the presence of the implied author's hand is always discernible behind the narrator's back" (77). He does not, however, bother to enlighten the uninitiated as to how the hand of the omnipresent implied author behind the narrator's back may in fact be discerned. Such metaphors, though vivid, provide only very opaque explanations of unreliable narration. From a methodological and theoretical point of view, they amount to nothing more than a declaration of bankruptcy. With regard to the question of how readers know an unreliable narrator when they see one, these metaphors are unenlightening.

To explain the mechanisms that stand behind the impression that a narrator is unreliable, it is not necessary to postulate an implied author but simply to have recourse to the concept of structural or dramatic irony (cf. Booth 1961, 255). The structure of unreliable narration can be explained in terms of dramatic irony and discrepant awareness because it involves a contrast between a narrator's view of the fictional world and the contrary state of affairs which the reader can grasp. The reader interprets what the narrator says in two quite different contexts. On the one hand, the reader is exposed to what the narrator wants and means to say. On the other hand, the statements of the narrator take on additional meaning for the reader, meaning that the narrator is not conscious of and does not intend to convey. Without being aware of it, unreliable narrators continually give the reader indirect information about their idiosyncrasies and states of mind. The peculiar effects of unreliable narration result from the conflict between the narrator s report of the facts" on the level of the story and his own interpretations. The narrative not only informs the reader of the narrator s version of events, it also provides him or her with indirect information about what presumably "really happened" and about the narrator's frame of mind.

If one gives up the notion of the implied author, then it is necessary to modify Booth's (1961, 74) explanations of the unreliable narrator in such a way as is already suggested by his definition of the implied author as "the core of norms and choices." Unreliable narrators are those whose perspective is in contradiction with the value and norm system
of the whole text or with that of the reader. The phenomenon of unreliable narration can be seen as the result of discrepant awareness and dramatic irony.

The general effect of what is called unreliable narration consists of redirecting the reader's attention from the level of the story to the speaker and of foregrounding peculiarities of the narrator's psychology. Wall (1994, 23) argues very convincingly that unreliable narration "refocuses the reader's attention on the narrator's mental processes." What is needed therefore is a more systematic exploration of the relation between unreliability and characterization. In the only available article on the subject, Dan Shen (1989, 309) has shown that "deviations in terms of reliability may have a significant role to play in revealing or reinforcing narratorial stance" and "in characterizing a particular consciousness." In unreliable narration it is often very difficult to determine whether what the narrator says provides facts about the fictional world or only clues to his distorted and evaluating consciousness. Consequently, the answer to the question "reliable, compared to what?" may vary dramatically depending on whether the standard according to which we gauge the potential unreliability of the narrator involves the events or the narrator's subjective view of them.

In sum, the link that theorists have forged between the unreliable narrator and the implied author deprives narratology of the possibility of accounting for the pragmatic effects subsumed under the term of unreliable narration. The critic accounts for whatever incongruence s/he may have detected by reading the text as an instance of dramatic irony and by projecting an unreliable narrator as an integrative hermeneutic device. Culler (1975, 157) has clarified what is involved here: "At the moment when we propose that a text means something other than what it appears to say we introduce, as hermeneutic devices which are supposed to lead us to the truth of the text, models which are based on our expectations about the text and the world." This, of course, raises the questions of what kind of models are involved in the cognitive processes that lead to the projection of an unreliable narrator.

II. Reconceptualizing conventional theories of unreliable narration

Heeding Harker's (1989) call for a radical reorientation, I will try to outline a model-oriented approach to how texts that display features of unreliable narration are read. I will contend that we can define unreliable narration neither as a structural nor as a semantic aspect of the textbase alone, but only by taking into account the conceptual frameworks that readers bring to the text. If we are to make sense of unreliable narration at all, it would be wise to begin by looking at the standards according to which critics think they recognize an unreliable narrator when they see one.

Determining whether a narrator is unreliable is not just an innocent descriptive statement but a subjectively tinged value-judgment or projection governed by the normative presuppositions and moral convictions of the critic, which as a rule remain unacknowledged. Critics concerned with unreliable narrators recuperate textual inconsistencies by relating them to accepted cultural models. Recent work on unreliable narration confirms Culler's hypothesis about the impact of realist and referential notions for the generation of literary effects. Culler (1975, 144) argues that "most literary effects, particularly in narrative prose, depend on the fact that readers will try to relate what the
text tells them to a level of ordinary human concerns, to the actions and reactions of characters constructed in accordance with models of integrity and coherence."

Riggan's monograph on the unreliable first-person narrator provides a case in point. Despite its insights into a broad range of texts, it suffers from all of the theoretical shortcomings outlined above. A look at Riggan's typology of unreliable narrators provides insight into the basic mechanisms that are involved in the projection of an unreliable narrator. Riggan distinguishes four types of such narrators, which he designates as "picaros," "madmen," "naïfs," and "clowns."

These typological distinctions can best be understood as a way of relating the text to accepted cultural models or to literary conventions. What critics like Riggan are doing is integrating previously held world-knowledge with textual data or even imposing preexisting conceptual models on the text. The models used to account for unreliable narration provide a context which resolves textual inconsistencies and makes the respective novels intelligible in terms of culturally prevalent frames.

It is these models which determine the perception of narrators designated as "unreliable," and not the other way round. The information on which the projection of an unreliable narrator is based derives at least as much from within the mind of the beholder as from textual data. In other words: whether a narrator is called unreliable or not does not depend on the distance between the norms and values of the narrator and those of the implied author, but between the distance that separates the narrator's view of the world from the reader's or critic's world-model and standards of normalcy, which are themselves, of course, open to challenge. It is thus necessary to make explicit that customary presuppositional framework on which theories of unreliable narration have hitherto been based.

An analysis of the presuppositional framework on which most theories of unreliable narration rest is overdue, since research into unreliable narration has been based on a number of highly questionable conceptual presuppositions, which as a rule remain implicit and unacknowledged. The general notion of unreliability presupposes some sort of standard for establishing whether or not the facts or interpretations provided by a narrator may be held suspect. The violations of norms which interest critics and theorists "are only made possible by norms which," as Culler (1975, 160) wittily observes, "they have been too impatient to investigate in detail." These presuppositions about unreliable narration need to be made explicit and clarified because they provide the key for reconceptualizing unreliability.

Among these underlying (and unwarranted) presuppositions on which the concept of unreliable narration relies, one might distinguish between epistemological and ontological premises, assumptions that are rooted in a liberal humanist view of literature, and psychological, moral, and linguistic norms — all of which are based on stylistic and other deviational models. An analysis of the presuppositional framework on which most theories of unreliable narration are based reveals that the orthodox concept of the unreliable narrator is a curious amalgam of a realist epistemology and a mimetic view of literature.

The epistemological and ontological premises consist of realist and by now doubtful notions of objectivity and truth. More specifically, the notion of unreliability presupposes that an objective view of the world, of others, and of oneself can be attained. In contrast to the ideal of objective self-observation, it needs to be emphasized that "a maximally
objective view of oneself can be attained only by others" (Fludernik, 1993, 53). The concept of unreliable narration also implies that human beings are principally taken to be capable of providing veracious accounts of events, proceeding from the assumption that "an authoritative version of events" (Wall, 1994, 37) can in principle be established or retrieved.

Theories of narrational unreliability also tend to rely on realist and mimetic notions of literature. The concept of the unreliable narrator is based on what Yacobi (1981, 119) has aptly called "a quasi-human model of a narrator" and, one might add, on an equally anthropomorphized model of the implied author. Amoros (1991, 42) has provided a convincing critique of this general tendency of allocating human features to the narrative agent.

In addition, theories of narrational unreliability are also heavily imbued with a wide range of unacknowledged notions that are based on stylistic deviation models or on more general notions of deviation from some norm or other. The notion of unreliability presupposes some default value which is taken to be unmarked "reliability." This is usually left undefined and merely taken for granted. Most critics agree, however, that reliability is indeed the default value (cf. Martinez-Bonati, 1981). Lanser (1981, 171), for instance, argues that "the conventional degrees zero [are] rather close to the poles of authority," and Riggan (1981, 19) observes that "our natural tendency is to grant our speaker the full credibility possible within the limitations of human memory and capability." To my knowledge, Wall is the first theorist of unreliable narration who sheds some light on the presuppositions on which this "reliable counterpart" of the unreliable narrator rests when she argues that the reliable narrator "is the 'rational, self-present subject of humanism,' who occupies a world in which language is a transparent medium that is capable of reflecting a 'real' world."

Vague and ill-defined though this norm of reliability may be, it supplies the standard according to which narrational unreliability is gauged. If one takes a close look at the presuppositional framework on which theories of unreliable narration are based, one can further elucidate the assumption that an unreliable narrator departs from certain norms. What is involved here are various sets of ill-defined and usually unacknowledged norms, which can, however, theoretically be distinguished.

One of these sets of norms includes all those notions that are usually referred to as "common sense." Another set encompasses those standards that a given culture holds to be constitutive of normal psychological behaviour. Thirdly, the habit of discussing the stylistic peculiarities of unreliable narrators shows that linguistic norms also play a role in determining how far a given narrator deviates from some implied default. Finally, many critics seem to think that there are agreed-upon moral and ethical standards that are often used as frames of reference when the question of the possible unreliability of a narrator is raised.

One of the main problems with all of these tacit presuppositions based on unacknowledged norms and notions of deviation is that the establishment of norms is much more difficult than critics want to make us believe. Fludernik (1993, 349), for instance, argues that the "explicatory power of stylistic deviation breaks down at the point where one can no longer establish a norm, or where deviations from the norm are no longer empirically perceptible."
In both critical practice and theoretical work on unreliable narration, however, these different sets of norms are usually not explicitly set out but merely introduced in passing, and they seldom if ever receive any theoretical examination. Let me give one typical example: in what is the only book-length study of the unreliable first-person narrator, Riggan (1981, 36), for instance, suggests that the narrator's unreliability may be revealed by the "unacceptability of his [moral] philosophy in terms of normal moral standards or of basic common sense and human decency." By saying this, he lets the cat out of the bag in a way that is very illuminating indeed.

Phrases like these unwittingly reveal the real standards according to which critics decide whether a narrator may be unreliable: It is not the norms and values of the implied author, whoever or wherever that phantom may be, that provide the critic with the yardstick for determining how abnormal, indecent, immoral or perverse a given narrator is, but "normal moral standards," "basic common sense" and "human decency" — in other words: unreliable, not in comparison to the implied author, but unreliable in comparison to what the critic takes to be "normal moral standards" and "common sense."

The trouble with seemingly self-explanatory yardsticks like "normal moral standards" and "basic common sense" is that no generally accepted standard of normality exists which can serve as the basis for impartial judgments. In a pluralist, postmodernist, and multicultural age like ours it has become more difficult than ever before to determine what may count as "normal moral standards" and "human decency." In other words: a narrator may be perfectly reliable compared to one critic's notions of normality but quite unreliable in comparison to those held other people. To put it quite bluntly: a pederast would not find anything wrong with *Lolita*; a male chauvinist fetishist who gets his kicks out of making love to dummies is unlikely to detect any distance between his norms and those of the mad monologuist in Ian McEwan's "Dead As They Come"; and someone used to watching his beloved mother disposing of unwelcome babies would not even find the stories collected in Ambrose Bierce's "The Parenticide Club" in any way objectionable.

There are a number of definable textual clues to unreliability, and what is needed is a more subtle and systematic account of these signals. Unreliable narrators tend to be marked by a number of textual inconsistencies. These may range from internal contradictions within their discourse over discrepancies between their utterances and actions (cf. Riggan, 1981, 36, who calls this "a gaping discrepancy between his conduct and the moral views he propounds"), to those inconsistencies that result from multiperspectival accounts of the same event (cf. Rimmon-Kenan, 1983, 101; Toolan, 1988, 88).

The range of clues to unreliability that Wall (1994, 19) simply refers to as "verbal tics" or "verbal habits of the narrator" (20) can and should be further differentiated by specifying the linguistic expressions of subjectivity. Due to the close link between subjectivity, on the one hand, and the effect called unreliability, on the other, the virtually exhaustive account of categories of expressivity and subjectivity that Fludernik (1993, 227-279) has provided are also extremely useful for drawing up a list of grammatical signals of unreliability, which can be further differentiated in terms of the linguistic expressions of subjectivity. The "establishment" of a reading in terms of "unreliable narration" frequently depends on the linguistic and stylistic evocation of a narrator's subjectivity or cognitive limitations (cf. Fludernik, 1993, 280).
Despite the above list of textual clues to unreliability, it needs to be emphasized that the problem of unreliable narration cannot be resolved on the basis of textual data alone. In addition to these intratextual signals, the reader also draws on extratextual frames of reference in his or her attempt to gauge the narrator's potential degree of unreliability. The term "unreliable narrator" does not designate a structural or semantic feature of texts, but a pragmatic phenomenon that cannot be fully grasped without taking into account the conceptual premises that readers and critics bring to texts. Consequently, it seems doubtful whether the term unreliable narrator can be defined, as Zimmermann (1995, 61) has recently maintained, solely on the basis of what she calls "intratextual dissonances."

What is needed instead is a pragmatic and cognitive framework that takes into consideration the world-model or conceptual information previously existing in the mind of reader or critic. It is necessary to take into consideration both the world-model and norms in the mind of the reader and the interplay between textual and extratextual information. Coming to grips with narrational unreliability is impossible if one conceives reading as being a mere "bottom-up or data-driven process" just as if one conceives it as being nothing but "a 'top-down' or conceptually driven process" (Harker, 1989, 471).

Developing a viable theory of unreliable narration that accounts for the complex meaning effects subsumed under the concept of unreliable narration presupposes an "interactive model of the reading process" (Harker, 1989, 471) and a reader-oriented pragmatic or cognitive framework (cf. Fludernik, 1993, 51). It is only within an interactive model of the reading process that an adequate theory of unreliable narration can be elaborated. Fludernik's (1993, 353) explanation of irony illuminates how this might be conceptualized: "textual contradictions and inconsistencies alongside semantic infelicities, or discrepancies between utterances and action (in the case of hypocrisy), merely signal the interpretational incomptability... which then requires a recuperatory move on the reader's part — aligning the discrepancy with an intended higher-level significance: irony."

An interactive model of the reading process alerts theorists of unreliable narration that the projection of an unreliable narrator depends upon both textual information and extratextual conceptual information located in the reader's mind (cf. Harker, 1989, 476). Detaching the text from the reader and ignoring the world-models in the reader's mind has resulted in the aporias outlined above. On the other hand, one should beware of throwing the baby out with the bathwater by rejecting textual data as a legitimate basis for explaining unreliable narration.

Pragmatics and frame theory present a possible way out of the methodological and theoretical problems that most theories of unreliable narration suffer from because cognitive theories can shed light on the way in which readers naturalize texts that are taken to display features of narrational unreliability. To offer a reading of a narrative text in terms of unreliable narration can be thought of as a way of naturalizing textual inconsistencies by giving them a function in some larger pattern supplied by accepted cultural models. Culler (1975, 138) clarifies what "naturalization" means in this context: "to naturalize a text is to bring it into relation with a type of discourse or model which is already, in some sense, natural or legible." The concept of unreliable narration, for instance, provides the reader with a general framework which allows him or her to "treat anything anomalous as the effect of the narrator's vision or cast of mind" (Culler, 1975, 200). To my knowledge, Wall (1994, 30) is the only theorist to date who has at least briefly
discussed the relation between naturalization and unreliable narration: "Part of the way in which we arrive at suspicions that the narrator is unreliable, then, is through the process of naturalizing the text, using what we know about human psychology and history to evaluate the probable accuracy of, or motives for, a narrator's assertions." She is certainly also right when she suggests that this kind of naturalization "is so much a part of our reading strategy with respect to both characters and narrators that, in all probability, we do not notice it."

Noticing and clarifying those unacknowledged frames of reference provides the clue to reconceptualizing the whole notion of unreliable narration. The question of whether a narrator is described as unreliable or not needs to be gauged in relation to various frames of reference. More particularly, one might distinguish between schemata derived from everyday experience and those that result from knowledge of literary conventions.

A first referential framework should be based on the reader's empirical experience and criteria of verisimilitude. These frames depend on the referentiality of the text, the assumption that the text refers to or is at least compatible with the so-called real world. Whether a narrator is taken to be reliable or not depends, among other things, on such referential frameworks as the reader's or critic's

- general world-knowledge,
- historical world-model or cultural codes,
- explicit theories of personality or implicit models of psychological coherence and human behaviour,
- knowledge of the social, moral or linguistic norms relevant for the period in which a text was written and published (cf. Yacobi, 1987),
- the reader's or critic's psychological disposition, and system of norms and values.

Whether a narrator is taken to be reliable or not depends, among other things, on such referential frameworks as the reader's general world-knowledge. Deviations from what is usually referred to as "common sense" or general world-knowledge may indicate that the narrator is unreliable. Secondly, narrators who violate the standards that a given culture holds to be constitutive of normal psychological behaviour are generally taken to be unreliable. What is involved here is psychological theories of personality or implicit models of normal human behaviour. In order to gauge the potential unreliability of the fictitious child-molester Humbert Humbert, the narrator of Nabokov's Lolita (1955), it does not suffice to look at textual data alone, because the process of character constitution during the reading process is inevitably influenced by the reader's implied personality theory, as Grabes (1996) has convincingly demonstrated. Thirdly, generally agreed-upon moral and ethical standards are often used as frames of reference when the question of the possible unreliability of a narrator is raised.

When a narrative text violates one or several of these normative presuppositions, the reader can always resort to one of these frames of reference in order to naturalize the text. As the reader relates discrepancies to these frames of reference, he or she brings the text into a context of coherence. Note that the choice of a particular frame of reference brings about a change in the mode of reading.

A second set of models brought into play in order to gauge a narrator's possible unreliability involves a number of specifically literary frames of reference. These include, for example:

- general literary conventions (cf. Amorós, 1991),
• conventions and models of literary genres,\textsuperscript{10}
• intertextual frames of reference — that is, references to specific pre-texts,
• stereotyped models of characters such as the picaro, the \textit{miles gloriosus}, the trickster,
• and last but not least the structure and norms established by the respective work itself.

The generic framework determines in part which criteria are used when a narrator's potential unreliability is gauged (cf. Yacobi, 1987, 20f.). A narrator that is considered to be unreliable in psychological or realistic terms may appear quite reliable if the text belongs to the genre of science fiction.

Both the concept of unreliable narration and the various types of unreliable narrators that have been proposed can be seen as modes of naturalization. These are based on widely accepted cultural frames which not only link a high number of disparate items, but also resolve whatever conflicts he or she may have noticed. The reader can try to account for textual inconsistencies by reading the text as the utterance of an obtuse, morally peculiar, or psychologically disturbed (i.e., unreliable) narrator. In this process, accepted cultural models of "deviant" but plausible human attitudes or behaviour are made use of, and the text begins to become naturalized.

The postulation of an unreliable narrator can be understood as a "mechanism of integration" (Yacobi, 1981, 119) in that it resolves whatever textual contradictions or discrepancies between the textual data and the reader's world-knowledge there may have been and leads to a synthesis at a higher level. Although relying on the implied and/or real author as the ultimate reference-point on which "reliability-judgments performed by the reader" (Yacobi, 1987, 22) depend, Yacobi (1987, 24f.) comes to a similar conclusion: "The hypothesis of a fictional reporter's unreliability is a mechanism for reconciling textual incongruities by appeal to a deliberate tension between the viewpoint of this informant (a character, narrator, dialogist, monologist) and that of the implied author who created him for his own purposes." In calling the source from which the utterances emanate an unreliable narrator, the critic not only makes peculiar features readily intelligible, but she or he also specifies how the text as a whole should be read. In the pragmatic context provided by frame theory unreliable narration can be explained as "an interpretive procedure" (Yacobi, 1981, 121): "as the result of interpretative work brought to bear on the juxtaposition between the wording of the text and the (by implication incompatible) cultural or textual norms of the text \textit{as constructed by the reader} or implied as values shared by the reader and the realistic textual world" (Fludernik, 1993, 440; emphasis added). Conceived in this way, the projection of an unreliable narrator is not only informed by textual data, as Chatman and other proponents of the implied author would like to make us believe, but also by the conceptual models or frames previously existing in the mind of the reader or critic.

III. Reconceptualizing the generic scope of unreliable narration

So far, the focus of the discussion of unreliable narration — not just in the present article, but in literary studies at large — has been almost exclusively on narrative fiction (cf. Jahn 1998). The generic scope of the phenomenon in question, however, extends far beyond first-person narrators in novels or short stories. The following brief outline of the
broader generic scope of unreliable narration will have to be provisional and programmatic because no general overview of the subject is currently available.

Just as the history of unreliable narration does not begin with modern fiction, the use of unreliable narrators is also not confined to narrative fiction; it rather extends to a wider range of genres. The subgenres known as the dramatic monologue (cf. Bennett, 1987) and the memory play (cf. Brunkhorst, 1980) are cases in point. These hybrid genres cut across established generic categories of poetry, drama, and narrative: with its limitation to a single speaker usually revealing key episodes of his or her life, the dramatic monologue combines poetic diction with dramatic presentation and story-telling elements; similarly, the memory play is a type of drama with distinct narrative features.

The dramatic monologues of nineteenth-century English literature provide ample evidence of the use of unreliable narration in poetry. There are many noteworthy examples of such unreliable narration in Victorian poetry, the most famous of which are probably Browning's "My Last Duchess" (1842) and Tennyson's Maud: A Monodrama (1855). Both structurally and thematically, these poems display almost all of the features of unreliable narration that have been discussed: they involve first-person speakers whose disturbed perceptions, egotistic personalities, and problematic value-systems lead the reader to question the accuracy of their accounts. In Maud the monologist's strong bias results from a high degree of emotional involvement, from his divided loyalties, and from his overt partiality. Similarly, the speaker of Browning's "The Bishop Orders his Tomb at Saint Praxed's Church" (1845) is an unreliable narrator if ever there was one. The bishop unwittingly reveals that he has fathered several bastards and that even on his deathbed he is thinking of nothing other than material wealth and sexual joy. The study of a host of other Victorian poems — e.g. Browning's Porphyrias Lover (1836) and John Davidson's "The Testament of an Empire-Builder" (1902) — and of many of Rudyard Kipling's Barrack-Room Ballads would also benefit from the application of the conceptual tools developed for the analysis of unreliable narration.

The same is true for the "memory play," which typically features an unreliable first-person narrator. Many post-war English plays prove those critics and theorists wrong who, like Elam (1980, 111), maintain that drama is "without narratorial mediation." However, the study of both unreliable narration and point of view or focalisation in drama has received hardly any attention to date. In the only available article on the subject, Brian Richardson (1988, 194) has convincingly shown that the deployment of narratorial mediation and the appearance of unreliability in plays call "for the kind of analysis of point of view usually reserved for modern fiction."

Such memory plays as Tom Stoppard's Travesties (1974) and Peter Shaffer's Amadeus (1979), which feature Henry Carr and Antonio Salieri respectively as narrators, demonstrate that post-war English playwrights make very subtle use of unreliable narration. In the stage directions of his play, Stoppard explicitly draws attention to Carr's unreliability, something which results from the old man's poor memory and his reactionary prejudices: "the scene (and most of the play) is under the erratic control of Old Carr's memory, which is not notably reliable, and also of his various prejudices and delusions" (Stoppard, 1974, 27). The main reasons for Salieri's unreliability are his limited knowledge, the high degree of his emotional involvement, and his problematic valuesystem. In Amadeus dramatic irony results primarily from the tension between what the audience sees and what Salieri describes, while Travesties contains a wide range of textual clues to Carr's unreliability.11
Like many contemporary English novels, these memory plays call into question conventional notions of unreliable narration because they question the existence of a fundamental difference between "reliable" and "unreliable" narration (cf. Wall, 1994, 23). The fact that many recent English novels and plays challenge realist notions of truth and objectivity seems to confirm Wall's (1994, 37) view that we perhaps "need to re-think entirely our notion that unreliable narrators give an inaccurate version of events and that our task is to figure out 'what really happened'." It could be argued that the unreliable first-person speakers in plays such as Beckett's *Play* or Stoppard's *Travesties*, and novels such as Julian Barnes' *Talking It Over* or Kazuo Ishiguro's *The Remains of the Day* are ultimately not at all unreliable: the stories they tell may not provide objective renderings of events, but they depict, in a very truthful way, the illusions and self-deceptions of the narrators themselves.12

Despite its brevity, this sketch of the generic scope of unreliable narration may serve to show that this feature is not confined to narrative fiction. Rather, such hybrid subgenres as the dramatic monologue and the memory play demonstrate that unreliable narration appears crossgenerically. However, the use of unreliable narration in genres other than narrative fiction has yet to receive the scholarly attention it deserves.

It needs to be emphasized that narrative theory could and should be applied to both narration in drama (cf. Richardson, 1988, 198) and to such hybrid genres as the dramatic monologue and the ballad. The application of narrative theory to genres other than fiction could open up new directions of research in an age of literature that has, after all, become noted for the blurring of genre distinctions. Since both the crossing of the boundaries between fiction, drama, and poetry and the phenomenon that has come to be known as "intermediality" have become hallmarks of contemporary English literature, literary studies would arguably stand to gain by applying the categories and methods developed for the study of one genre (e.g., narrative fiction) to the study of other genres and media. If criticism and theory want to keep up with such innovative literary developments as the blurring of generic boundary lines, critics should not forget the insights which the "cross-generic" application of genre-specific theories affords.

IV. Conclusion

Let me conclude with a few brief indications of some of the new territories to be explored that are opened up by such a cognitive framework for the analysis of unreliable narration. Firstly, it can bridge the gap that has separated narratology and cognitive theory for much too long — to the detriment of narratological inquiry, one might add. Secondly, such a cognitive reconceptualization can be usefully applied in the as yet unwritten narratological history of the development of unreliable narration. Thirdly, a cognitive theory of unreliable narration may be useful for understanding how readers make sense of a narrative as a whole. Lastly, only if we take into consideration both the cognitive strategies and the culturally accepted models and frames that readers and critics, usually unconsciously, deploy when they naturalize texts in terms of unreliable narration will we be in a position to assess possible links between the historically variable notions of subjectivity and the equally changing uses of what has come to be known as the unreliable narrator.
Many recent novels do indeed suggest that there is something to Wall's (1994, 22) hypothesis "that changes in how subjectivity is viewed will inevitably be reflected in the way reliable or unreliable narration is presented." Contemporary British fiction, for example, often challenges the processes of naturalization involved in the projection of unreliable narrators and calls into question conventional notions of unreliable narration. Wall (1994, 18) has demonstrated that Kazuo Ishiguro's *The Remains of the Day* not only "challenges our usual definition of an unreliable narrator," but also "deconstructs the notion of truth, and consequently questions both 'reliable' and 'unreliable' narration and the distinctions we make between them" (23). The same point could be made with respect to many other post-war novels that employ first-person narrators including William Golding's *Free Fall* (1959), Anthony Burgess' *Earthly Powers* (1980), Nigel Williams' *Star Turn* (1985), William Boyd's *The New Confessions* (1988), and Jeanette Winterson's *Sexing the Cherry* (1989). Graham Swift's short stories and novels, for example, both foreground and challenge the problematic notions of truth, objectivity, and reliability on which theories of unreliable narration are based (see Nünning, 1993b). The fact that many recent novels and short stories challenge the usual definitions of an unreliable narrator confirms Wall's (1994, 37) view that we really "need to re-think entirely our notion that unreliable narrators give an inaccurate version of events and that our task is to figure out 'what really happened.'" Many historians have at last begun to reject the noble dream of objectivity, and it seems to be high time that narratologists did the same. But much more work needs to be done if we want to come to terms with the complex set of narrative strategies and reading processes that, ever since the good old days of Wayne C. Booth, have been subsumed under the wide umbrella of the term "unreliable narration."

From the point of view of the proposed cognitive theory of unreliable narration, the answer to the question "Unreliable, compared to what?" can be summed up in one brief sentence: unreliable, not compared to the implied author's norms and values, but to the reader's or critic's preexisting conceptual knowledge of the world and to his or her (usually unacknowledged) frames of reference. This answer and the above hypotheses are not, however, meant to be the last word on the unreliable narrator, but rather the first word on a radical reconceptualization of the subject. If, however, we are to make sense of unreliable narration at all, we would be wise to give up the implied author and instead take into consideration the unacknowledged standards and frames of reference according to which critics think they recognize an unreliable narrator when they see one.

---

**BIBLIOGRAPHY**
Works Cited

Secondary sources


---

Recent Trends in Narratological Research


NOTES


2. Cf. Bal (1981b, 208ff.); Jakobsen (1977, 182ff.); Toolan (1988, 77ff.); as I have proposed elsewhere (Nünning, 1993a, 1997b), instead of setting up an "implied author" and an "implied reader," it would be more sensible to conceptualize a textual level that encompasses the entirety of the structural properties of a work.

3. Cf. Chatman (1990, 81, 90ff.) and Nelles (1993), who does not even question the concept. Nelles (1993, 22) argues that the implied author and the implied reader "each has its distinctive function: [...] the implied author means, the implied reader interprets." By contrast, Diengott (1993a, 1993b) and Nünning (1993a, 1997b) are highly suspicious of the concept and try to demonstrate the problems it raises for narrative poetics.

4. Lanser (1981, 170ff.) and Hof (1984, 55) have suggested that one should distinguish between an unreliable narrator and an untrustworthy one. Hof shows that the report a narrator gives of the events may be heavily flawed although the narrator himself may appear to be absolutely trustworthy. And Lanser (1981, 171) suggests that "a narrator may be quite trustworthy in reporting events but not competent in interpreting them." Cf. Lanser's (1981, 40) observation that "a narrator can be perfectly reliable with respect to the 'facts' of a given story, but unreliable regarding opinions and judgments about the story world." It seems doubtful, however, as Manfred Jahn has pointed out to me (personal communication), whether anything is gained by substituting two ill-defined terms for one problematic category.

5. This approach is indebted to Jahn's valuable suggestions for a cognitive narratology (see Jahn, 1997) and to Fludernik's work on "natural" narratology (see Fludernik, 1993, 1996).


7. Cf. Amorós (1991, 56ff.) and Culler (1975, 140), who distinguishes "five levels of vraisemblance, five ways in which a text may be brought into contact with and defined in relation to another text which helps to make it intelligible."

8. Cf. Culler's (1975, 140) notion of "a general cultural text: shared knowledge which would be recognized by participants as part of culture," and Chatman's (1978, 149) observation that what Booth called moral norms are really "general cultural codes."

9. Cf. Culler (1975, 190, 225, 237) and Wall's (1994, 21) observation that unreliable narrators "reflect some reasonable model of human fallibility" and that readers use their "knowledge of psychology" (29) in evaluating a narrator's motives and behaviour. For the relevance of the reader's implied personality theory for the process of character constitution during the reading process, see Grabes (1996).

11. Other examples of plays which violate naturalistic stage conventions by relying on unreliable narration would be Samuel Beckett’s *Krapp’s Last Tape* (1958) and Harold Pinter’s *Landscape* (1968), the latter being composed of alternating and independent acts of narration spoken by two characters.

12. Cf. Rabinovitz (1983, 67), who argues that "the unreliable narrator in *Murphy* is in an ultimate sense not at all unreliable; for it depicts, in a truthful way, the illusions and deceptions of the outer world."

ENDNOTES

* The present article is a completely revised and extended version of a paper read at a symposium on "Narratologie, Rhetorik, Textlinguistik: Ein Integrationsversuch," held at the University of Dortmund in February 1996. I should like to thank Jon L. Erickson (University of Cologne), Luc Herman (Antwerp), Andreas Höfele (Heidelberg), Manfred Jahn (Cologne), and the convenors and participants of the Dortmund Symposium, especially Monika Fludernik (Freiburg), who also brought Yacobi (1987) to my attention, Gerald Prince and Susan Lanser for their stimulating comments and valuable suggestions.

ABSTRACTS

The paper argues that the concept "unreliable narrator" needs to be radically rethought because, as currently defined, it is terminologically imprecise and theoretically inadequate. The first part of the article is devoted to giving an assessment and critique of the standard notions of the unreliable narrator, arguing that the postulation of essentialized and anthropomorphized entities designated "unreliable narrator" and "implied author" ignore the complexity of the phenomena involved and stands in the way of a systematic exploration of the cognitive processes which result in the projection of unreliable narrators in the first place. The second part outlines a radical reconceptualization of unreliable narration. It is proposed that it would be more sensible to conceptualize the relevant phenomena in the context of frame theory as a projection by the reader who tries to resolve ambiguities and textual inconsistencies by attributing them to the narrator's "unreliability." In the context of frame theory, the reader's projection of "unreliable narrators" can be understood as an interpretive strategy or a cognitive process of the sort that has come to be known as "naturalization" (cf. Culler 1975; Fludernik 1993, 1996). A number of empirical frames of reference and literary models can be seen as standard modes of naturalization by means of which readers account for contradictions both within texts and between the world-model of texts and their empirical world-models. The final section gives a brief outline of the generic scope of unreliable narration, arguing that it is unjustifiable and counterproductive to limit the study of this phenomenon to narrative fiction.
AUTHOR

ANGAR F. NÜNNING

Professor of English and American Literary and Cultural Studies at the Justus-Liebig-Universität Giessen. He has published books on narratology, Virginia Woolf, the historical novel and historiographic metafiction, and he has recently edited an encyclopedia of literary and cultural theory. His articles on narratology, the eighteenth and nineteenth-century British novel and literary theory (e.g., the New Historicism, gender studies, the history of mentalities, radical constructivism and postcolonial criticism) have appeared in numerous scholarly journals and edited books.