Quantitative Studies of Literature.  
A Critique and an Outlook

W. van Peer

Dept. of Literary Theory, University of Utrecht, Muntstraat 4, 3512 EV Utrecht, The Netherlands

Abstract: The present paper is a critique of quantitative studies of literature. It is argued that such studies are involved in an act of reification, in which, moreover, fundamental ingredients of the texts, e.g. their (highly important) range of figurative meanings, are eliminated from the analysis. Instead a concentration on lower levels of linguistic organization, such as grammar and lexis, may be observed, in spite of the fact that these are often the least relevant aspects of the text. In doing so, quantitative studies of literature significantly reduce not only the cultural value of texts, but also the generalizability of its own findings. What is needed, therefore, is an awareness and readiness to relate to matters of textuality as an organizing principle underlying the cultural functioning of literary works of art.

Key Words: style/stylistics, quantitative studies of literature, ambiguity/disambiguation/figurative language, textuality, literary theory, relevance, parody, reader response (theory/research)

All too often perhaps it is forgotten that critique — any critique — must fulfil certain standards of adequacy. One is dealing with a genuine form of critique, I believe, only if the following principles are observed.

(1) The object or field being critically examined must be understood in depth and represented in all fairness. (This may demand some degree of empathy on the part of the critic).

(2) The rational analysis of its shortcomings and flaws must be unwavering, pitiless and remorseless. (This may require the critic to take a distanced stand, in which negativity is not shied away from).

(3) Some perspective on how these negative aspects of the criticism are to be remedied, or rather overcome, should form an integral part of the critique. (This will ask for the critic’s constructive attitude and spirit).

One could thus say that criticism only serves the advancement of knowledge if it incorporates elements bearing on each of these three principles. Obviously enough, (1) is not in itself a form of criticism at all. But neither is (2), if it is based on an incomplete understanding or on a misrepresentation of the object it purports to criticize, thus failing to fulfil condition (1). Although instances of principle (2) standing on its own may regularly be encountered, also in academic forms of criticism, it is by itself but an immature, often also a superficial and rather silly form of negativity. Concerning quantitative studies of literature, (Fish 1973) is a case in point. It attempts destruction rather than construction, thus also flouting principle (3). But worse even, it presents such an ill-informed caricature of the field that the subsequent analysis must necessarily remain shallow and misguided. Yet the field does have its shortcomings; I have said so myself in (Van Peer & Renkema 1984: 8). In what follows I attempt to
spell out that criticism in more detail. I propose to do so in three consecutive stages, each corresponding to one of the three principles mentioned before, but devoting proportionally more attention to the third one.

To start with the first principle, quantitative studies of literature must be understood against the background of the medium literature makes use of. That medium is language. But language can be talked and thought about. That is, language may be described. The discipline of linguistics and related research areas are but serious attempts to make such descriptions as objective and as insightful as possible. Furthermore, where things can be described in some generally objective sense, they can also be counted and counts can be compared. I take it that quantitative studies of literature do nothing less than that: trying to describe as accurately as possible the various linguistic elements and configurations one encounters in literary works of art. Language being a multi-levelled system of communication, such descriptions may take account of the various layers of linguistic organization: sounds or words, grammatical structure or semantic properties, discoursal recurrences or textual composition.

Because of the values conventionally associated with literary works of art such an enterprise may sometimes seem to go against the grain of our everyday experiences. But such a gap between everyday activities and concepts and methods of rational enquiry are by no means unique to literary scholarship. It is always there where cultural products are being examined in some spirit of scientific enterprise. There is, however, no a priori reason why literature, being made up of language would not be amenable to analysis from a quantitative perspective. The fact that in literature one encounters language in its most condensed form is yet another reason for applying the most sophisticated tools of description and analysis available. There can be little doubt that statistics and quantitative methods provide such a required level of sophistication.

However, these quantitative methods, when looked at in detail, also reveal their fundamental deficit. In doing so, we shall obey the second principle of criticism outlined above. In treating language as an object of description in the way it does, the quantitative approach simultaneously misses one of language’s most fundamental characteristics, i.e. its transient nature. The logical flaw involved here lies in an act of reification: what in fact is an ongoing and context-bound process, in which meanings are situationally constituted is reified into a distinct and isolated ontological category. Certainly such a mistaken view is not unique to quantitative studies of literature. Even in the related discipline of stylistics the emphasis has usually been on the text-as-product, so much even that (Enkvist 1988) has recently called for a more process-orientated approach in order to redress the balance. For quantitative studies of literature the problem of reification is deepened even further, however, precisely because of its quantitative methods. When stylistic features of a text have been transformed into numerical form, they acquire a status that actually prevents them from being perceived as language-for-communication as such. That is to say, in the very act of transforming textual qualities into counts, their essentially process-like character is irretrievably lost. Such a loss may be futile if the respective qualities are peripheral to the object under investigation (as is the case, for instance, in physics). Wherever such qualities belong to the very nature of the phenomenon being studied, however, the loss must be deemed detrimental.

Methodologically speaking it is important to point out that this problem cannot be solved by ever more sophisticated mathematical tools. The difficulty is rather of a categorical nature, in that it is the very act of transforming a text’s characteristics into numbers which renders an essential layer of its information inaccessible to further analysis. Thus no level of (mathematical) sophistication is able to overcome the problem that the processes of meaning constitution have been eliminated before the analysis is undertaken. By turning linguistic qualities of a text into numerical form, then, quantitative studies are involved in something much more serious than a problem of reification: they undermine the very foundations upon which the study of literature must necessarily be based.

This shortcoming of quantitative studies is quite apparent in one area particularly, i.e. the treatment of figurative language. Perhaps it is more correct
to speak of non-treatment in this respect, as it is one of the areas which has proved hardest to treat in any mechanically quantitative way. As far as I know there still is no computer program to automatically disambiguate figurative language as such (apart from dead metaphors or other kinds of ‘frozen’ figurative meanings). The initial stages of quantitative research in this area are still carried out by hand, thus demonstrating the fundamental shortcomings outlined in the previous paragraph: its neglect, in this case even its impotence, to come to terms with language as a partly open-ended process of communication. If the scoring of figurative language is done by hand, however, serious methodological problems arise. If, for instance, researcher A counts all metaphors in novels P and Q and then applies statistical analyses to the data thus obtained, it then becomes rather crucial to ask whether researcher B would have come up with the same numbers in the first place. The fact that in many cases an honest answer to this question will have to be ‘no’ corrupts the validity of the enterprise. (If the answer could be ‘yes’ then presumably we would be closer to a solution of how to have figurative meanings assigned automatically by machine.) Quantitative studies thus pretend to offer a high degree of objectivity, while in reality its procedures for analysis are insufficiently reliable.

There is a way out of this problem, one may imagine, by simply avoiding the study of figurative language altogether. What such a move would mean, however, may be appreciated from the list of literary devices that would then have to be eliminated from being studied. Such a list would certainly include the following ones: metaphor, metonymy, synecdoche, irony, symbol, hyperbole, litotes, pun, oxymoron, paradox, ambiguity, and many others. One could even wonder whether it would leave the stylistician much to do (apart from studying function words), as nearly all content words may under certain conditions or in specific circumstances acquire a new symbolic meaning which is not necessarily contained in the world’s literal sense. As is well known, a whale is not always simply a whale in a novel, but may develop into something which was quite unpredictable from its known biological properties. Eliminating figurative language from quantitative studies of literature may even involve whole genres, in which meaning is systematically located on another than the literal plane, as in satire, for instance. In short, then, it seems to be the case that a strategy of avoidance will not work, because its systematic implementation would rid the field of its own object of study. Literature relies for its functioning so heavily on devices involving figurative uses of language that any effort to avoid these would result in the abolishment of its own raison d’être.

Behind all this there lurks an even greater danger, i.e. that of bypassing figurative meanings without one’s being aware of it. In counting words as words only, one may easily overlook the fact that such words will have very different meanings and connotations in different literary works, or in the hands of different authors. Take, for instance, the English adjective ‘great’. Any quantitative description of the occurrence of this word in literary works may lead to ludicrous results if it treats the adjective at its face value only. As even a quick and superficial comparison of this word in the title of two famous English novels reveals, ‘great’ is not always ‘great’ at all. *The Great Gatsby* refers to a kind of greatness quite different from Fielding’s *Jonathan Wild, the Great*, showing the world to be a little more complicated than the quantitative student would hope.

Here we face the fundamental shortcoming of quantitative studies at a deeper, and hence perhaps even more devastating level of analysis, revealing the utterly naive view of language that underlies its practice. Without facing this problem, and without any serious effort to raise the solutions above the commonplace, quantitative studies run the risk of reducing literature, in the very act of its being studied, to a non-entity, i.e. literature deprived of its literary character. By removing all that is non-commonsensical in the language, by radically neglecting the multivalency of the linguistic sign-in-use, by eliminating the contextual factors that fundamentally shape its meaning prior to the analysis undertaken, the quantitative approach destroys the very ‘literariness’ of the texts under study. In the event, its own object of study vanishes completely. What comes out is not an analysis and a deeper understanding of the literary text, but the shaping and projection into some
strange non-literary version of that text, which is non-existent in the real world at large.

So far I have advocated a view of quantitative studies *ex negativo*, indicating where I think their basic shortcomings lie. This view can be summed up by saying that underlying these studies is a view of language which is so simplified and naive (and therefore so utterly mistaken) that is is untenable even by the most common standards of everyday observation, let alone by standards governing critical-analytic inquiry. The view has been put forward relentlessly enough perhaps, to allow us to move on to the third stage of our investigation. This involves a return to the principle of construction, attempting to further understanding in order to overcome the basic problems one is facing.

When quantitative studies are seen from this perspective, it emerges that in spite of the deficiencies outlined above, the approach nevertheless has a kernel of truth in it that has to be acknowledged. Since this may seem to come as an unexpected twist in the general reasoning, it may demand some more argumentation. That quantitative studies of literature treat literary works as ‘objects’ is true enough, thereby severely neglecting contextual factors. Yet at the same time such works can hardly be denied some kind of objective status. The fact that they are constantly being transmitted to new readers demands at least some degree of context-independence of their meaning. Clearly enough, if this were not the case, then communication on, say, *Madame Bovary*, would not be possible. The fact that readers from very different backgrounds may nevertheless understand each other without too many problems when this novel forms the topic of a conversation, means that there must be some level at which the information in the novel can objectively be established and exchanged.

The origin of this fact lies at least in part in the novel’s textuality. I am invoking a notion of textuality in which texts are seen as ‘extended speech acts’, the primary function of which is to be passed on through time and space. Texts belong to — in a sense they also constitute — cultural traditions, and consequently operate as cultural products which are constantly being (re)distributed throughout society and from one generation to a following one. No language can function in such a process unless it acquires some degree of autonomy from the context in which it is (re)produced. In the process of shaping a text, therefore, its language becomes pre-structured according to specific (generic) conventions. This causes the language to acquire a higher degree of object-quality. This materializing and objectifying of language is also true of ‘oral texts’, such as riddles, rhymes or tales. But their object-quality is further enhanced once writing is introduced; compare, to this end (Goody 1977, 1986). For the concepts of text and textuality referred to, see (Ehlich 1984, van Peer 1989).

There is a paradox here, however: the different levels of language organization (sound, grammar, meaning, illocutionary force) yield, on the one hand, a cline of increasing difficulty to objectively quantified analysis. That is, illocutionary force and meaning aspects of a text are much harder to quantify than issues of sound and grammar. The reason for this state of affairs is clear enough: meaning and force are not easily established out of context, while the description of sound and grammar may be done in a way which is more detached from the situation in which language is actually used; cp. the earlier reference to the treatment of figurative language. On the other hand, it is exactly these more abstract aspects of linguistic organization, such as meaning and force which are at the core of textuality and indeed of literature. The sounds and syntactic patterns observable in literary works of art may play a more prominent role than those discerned in everyday language, but they are hardly relevant without some kind of functionality of the work of art, which will be determined first and foremost by the semantic aspects and illocutionary levels. Thus two opposing forces are at work. One is a cline of ‘ease’ with which linguistic data may be quantified objectively. The other one is a cline of ‘relevance’ of linguistic qualities to textuality. Figure 1 shows these two forces to work in opposite directions.

The problem then becomes how to solve this paradox. Perhaps it is not exaggerated to say that the standard solution in quantitative studies has been, and still is, to largely neglect the textual dimension of the literary work, or to take that dimension for granted without paying too much
attention to it. However, while textuality forms the basis for the social functioning of the literary work, quantitative studies largely ignore its analytic significance. In the absence of a general cultural perspective, however, more problems are created than solved.

Perhaps it is good to look in somewhat more detail at this lack of fit between quantification and textual functioning. The paper by Norbert Bolz given at the conference, for instance, tries to answer the question how one can account for the different stylistic impressions made by Robert Greene's early euphuistic work and his later more unadorned descriptions of the Elizabethan underworld. Bolz thus attempts to establish a linguistic criterion correlating to an intuitive perception of stylistic differences. To this end, a presumably representative sample from both categories is scored for a relatively large number of linguistic variables of different degrees of complexity. A number of statistical analyses are then performed on these results, showing that even a small number of linguistic criteria suffices to discriminate between both styles. If indeed, as is claimed, a finite set of linguistic variables is able to predict membership to one of both styles with 95% accuracy, then this certainly must be acknowledged as a highly interesting and promising result. Yet one must not be blind to its limitations. For one thing, it is not at all clear what the relationship is between the predicting variables isolated in the analysis and the overall notions such as 'euphuistic style'. It is not difficult to disprove the existence of any direct relationship. Suppose, for instance, that Greene (or any other author, for that matter) would like to treat this euphuistic style ironically by having a character in a narrative speak euphuistically, his utterances being ironically commented on in brief asides by an omniscient narrator. Since the character's utterances would contain a high instance of the linguistic variables isolated in Bolz's analysis, they would (wrongly, in this case) be assigned to the class of euphuistic texts. Yet they are not really euphuistic prose at all, but a parody of it. That is, they resemble the euphuistic style, but in fact aim at quite a different effect. And a reader would be able to recognize those different aims and meanings. Obviously, what is needed here is a distinction between 'narrator text' and 'character text', a distinction operating at the level of textual organization.

I hope it will be clear that the example is not simply a matter for academic dispute: the history of literature is full of works which parody particular styles, conventions and traditions. Also it is not my intention to simply disclaim the value of the enterprise Dr. Bolz is engaged in. Anyone can see that the research he presents is of a high order of ingenuity, resulting in indeed remarkable achievements of descriptive delicacy. What I intend to point out, rather, is the limited use that can be made of the results, however interesting they may be in themselves. As the argumentation shows, this limitation is a consequence of the poor fit between the linguistic variables selected and the textual functioning of the two styles examined. Formulated in terms of Fig. 1, it is clear that Dr. Bolz's paper does present an attempt to relate detailed linguistic description to more general intuitions of style, thereby touching on the issue of textuality. The reason why the establishment of such a relationship does not succeed, lies, I should like to propose, in the one-sided attention which is nevertheless given to words and their attributes only. One could thus say that this type of quantitative research generally is trapped in the levels of grammar and lexis, unable to deal with more general (and more important) issues of meaning and force. While devoting all energy to the lower levels of language organization, at the expense of the textual dimension, quantitative studies sacrifice not only the cultural significance of literary texts, but also the generalizability of its own findings.

What is needed, therefore is both an awareness and a readiness to explicate and operationalize stylistic intuitions, not merely in order to compare
them to the linguistic structures encountered in the text, but also, and especially, to relate them to general matters of textuality. This may be done in generally two different ways. The first one consists in using quantitative methods, but applying them to the reactions real readers bring to the text. I have myself gone some way in using this method in Van Peer (1986). This does require the quantitative stylistician to adopt another perspective, however, bringing it in relation to response theory. Note that seminal work in this direction (albeit of a non-quantitative kind) had already been carried out in the English speaking countries as early as Richards (1929), but for some reason or other this line of thought has hardly been followed up. The other way to cope with the problem does not ask for such a profound change of perspective, but it will still demand a creative adaptation of methods and theoretical insights. It means nothing more nor less than that serious attention be paid to aspects of textuality as an organizing principle underlying the cultural functioning of literary works of art. New concepts, relating to the dimension of textuality, are needed to achieve this end, but initially a number of them may be derived from the theoretical study of literature. The difference between ‘narrator text’ and ‘character text’ invoked before is such a case in point. In a recent piece of research, for instance, — Van Peer & Andringa (forthcoming) — it was found that readers’ perceptions of sentence length did not relate in any direct way to objective measurements of a text’s real sentence length, but were in some complex way dependent on the alternation between dialogue (i.e. ‘character text’) and the narrator’s comments and descriptions. These are not the only aspects of textual organization that one may wish to investigate, however. Issues such as text deixis, turn taking mechanisms in dialogue, speech and thought presentation, parallelism, alternations in point of view, openings and closings, plot and the use of ‘editorial’ techniques by a narrator, all form fruitful topics to be investigated. The challenge will be how to develop reliable methods of quantification, but the rapid developments in computers will no doubt be helpful in this respect.

Two last remarks are in place here, one theoretical and one of a methodological kind. First, it is of considerable importance to relate such studies to general issues of literary theory. It is, in my opinion, an unfortunate and sad state of affairs that so very little of what goes on in quantitative studies bears any direct relevance to the theory of literature in general. Alternatively, a good deal of theoretical issues would benefit from a more rigorous approach, of the kind developed by quantitative studies of literature. Issues such as innovation and convention, tradition and creativity, ideological undercurrents and mimetic aspects of literature, literature’s moral and visionary power or its utopian qualities are far too important not to probe them in a rigorous way. And it will bring quantitative studies within the realm of central issues facing literary theory, especially the question how literary texts function in the lives of individuals and of social groups. Again, there is no a priori reason why these cannot be studied in a quantitative way. Such an attempt will no doubt have to face the fundamental problems outlined before. This concerns the second point I should like to raise, and which is of a methodological kind. It would be wise to adhere to a cyclic procedure for carrying out quantitative research, a rough outline of which might be formulated as follows. First, questions and problems related to the general theory of literature are isolated for closer study. Then a quantitative analysis of a representative text corpus bearing on that issue is undertaken. Finally the results obtained are interpreted with reference to the initial postulates and fed back into the general theoretical framework. Note that this methodology may be defended from yet another angle in this respect, i.e. the problem of over-production. The present situation, in which sophisticated concordances and analyses that one could only dream of twenty years ago, are now available even to students at the touch of a button, requires a keen sense of the relevance of the data generated.

In sum, then, one must conclude that quantitative studies of literature are not at the heart, but rather at the periphery of the theoretical and methodological occupation with literary phenomena. This is a position which the discipline does not, I believe, deserve. Some of the most sophisticated instruments of description belong to its standing analytical repertoire. And some of the
finest scholars work in this field. I therefore would hope that the present critique may lead to further developments, out of the periphery, and into the core of theoretical matters in the study of literature.

References