# 'THEME DYNAMICS' AND STYLE: AN EXPERIMENT

NILS ERIK ENKVIST

*Åbo Akademi*

## 1

In a bird's-eye view, my paper will try to relate two streams of modern linguistics to the study of literary texts. One of these streams is stylistics, the other is text linguistics.

Those who wish to study 'style' in operational terms must first explain precisely what they mean by 'style'. For the purposes of this paper I shall define the linguistic style of a text as the aggregate of those linguistically definable features whose density[1] in that text is significantly different from their density in another, comparable and therefore contextually related, body of text. In these terms, style becomes a differential, and stylistic investigation must be based on a comparison. We must find another body of text that we regard as worth comparing with our original text. This body of text may be called 'norm' as long as we remember that we are now using 'norm' as a technical term and divest it of all wider implications.

If we match the same text against different norms, the results may be different. Each comparison may yield a new set of 'style markers', that is, of features whose densities of occurrence are significantly different in the text and in the norm. The investigator must decide where he wishes to seek the stylistic norm against which he wants to view his text. Usually the decision is made on cultural and literary, and thus extralinguistic, grounds. Presumably such matching of text and norm has usually been an implicit or explicit step in all stylistic criticism. The choice of different norms helps to explain why different critics may legitimately arrive at different opinions about the stylistic characteristics of the same text[2].

[1] 'Density' here means the number of occurrences of a linguistic feature divided by a suitable measure of text length.

[2] These principles have been discussed at greater length in N. E. Enkvist, "On defining style" in John Spencer, N. E. Enkvist and Michael Gregory, 1964. *Linguistics and style*. London: Oxford University Press, and in N. E. Enkvist, 1973 *Linguistic stylistics*. The Hague: Mouton.

The second major stream that enters into my paper is text linguistics, that is, the branch of language study that investigates textual units larger than the sentence. A summary of the current state of text linguistics is unfortunately out of order in a short paper[3]. It is, however, obvious enough that text linguistics is one of the areas in which the study of language is about to peg new claims. A number of textual phenomena that linguists who worked with one sentence at a time used to dismiss as instances of 'free variation' are coming under intense grammatical, syntactic and semantic scrutiny, and a number of new regularities and trends are appearing under the linguist's newly refocussed microscope. The increasing impetus for text linguistics comes from several quarters. Grammarians are becoming more and more aware of the fact that the structure of an individual sentence cannot be fully explained without relating it to the textual environment. At the same time, the theoretical status of the sentence is under scrutiny, and some old problems have gained in clarity thanks to generative grammar and the recent attempts to find common ground for logic and grammar. Rules for processes such as pronominalization and the disambiguation of polysemic sentence constituents are being drawn up, and they too necessitate frequent glances beyond the confines of an individual sentence. Text linguists are also very much interested in finding out precisely how we are able to distinguish a coherent text from a string of unconnected, though otherwise well-formed, sentences. If making such a distinction is part of our linguistic competence — as it seems to be — it must be explained as part of that competence and therefore as part of grammar. In addition to well-formedness at the sentence level we must thus learn to define textual well-formedness or coherence. And once we can define conditions for textual well-formedness we may proceed to try to explain instances of aberration and textual deviance. Text linguistics also has practical implications. It may ultimately help us to teach people how to express themselves: it may become part of the new rhetoric. It also opens views for the study of literature. Absurd drama, for instance, achieves some of its effects precisely by departing from the normal patterns by which sentences are joined into a well-formed text. Here we have a textual parallel to the poetic use of deviant syntax.

The ways in which sentences join each other to form texts are thus becoming increasingly amenable to stringent description. Now any linguistic feature of a text is in principle capable of functioning as a style marker, if only its density in the text is significantly different from its density in a suitably selected norm. The question then arises, is it possible to find text-linguistic style markers? Is it possible to define the difference between a text and a norm in terms of such linguistic features as describe the ways in which sentences are strung together, and not only in terms of features within individual sentences?

[3] An up-to-date survey with a useful bibliography is in Teun A. van Dijk. 1972. *Some aspects of text grammars*. The Hague: Mouton.

## 2

To explore the answer, I have, on a modest scale and with very tentative methods, experimented with an approach I have given the working label 'theme dynamics'. First a few basic concepts and terms, some of which I have come to use in ways no less idiosyncratic than those in which they appear in the works of others.

In a well-formed text, every sentence must be linked to at least one other sentence of the same text. In every sentence of a well-formed text there must be at least one constituent which links that sentence to its textual environment. The means of such linking are many: repetition, pronominal or adverbial reference, synonymy, antonymy, hyponymy, the use of words from the same semantic field, and ellipsis. Let us use the term 'topic' for referentially or semantically linked constituents of different clauses or sentences. In the two-sentence text

Napoleon went to Tilsit. When he arrived there he was tired.

*Napoleon, he* and *he* form one topic, *Tilsit* and *there* another, *went* and *arrived* a third.

If we succeed in performing a topical analysis of our text, showing what constituents of each sentence are topically linked with what constituents of other sentences in that text, our next task will be to see what position each topic occupies within each sentence. If we adopt the traditional terms 'theme' and 'rheme' for the informational constituents of the individual clause and sentence, we may say that our topical analysis — the tracking of referential and semantic connections between constituents of different sentences — must be supplemented with a thematic analysis, that is, with an analysis of the thematic (including 'rhematic') function that each topic has within its own sentence.

After the topical and thematic analyses, the third step of the procedure should be a classification of the different patterns by which a topic can move from sentence to sentence within a text. Such a taxonomy may use syntactic terms — for instance, we might say that the object of one sentence is linked with the subject of the next — or it may build on thematic concepts such as theme and rheme[4].

The fourth step — in stylistic analysis the crucial one — will then be a comparison of the densities of different patterns of topical linkage in terms of thematic movement, always comparing the original text with a suitably selected

[4] Such a classification has been developed by František Daneš in "Zur linguistischen Analyse der Textstruktur", *Folia Linguistica* IV (1970). 72 - 78, and in "One instance of Prague School methodology: functional analysis of utterance and text" in Paul L. Garvin, ed., 1970. *Method and theory in linguistics*. The Hague: Mouton.

norm. If significant differences emerge, the text is characterized by theme-dynamic style markers. If no such differences can be found, the text and the norm are stylistically similar in this respect (which of course does not exclude the possibility that they may be radically different in an infinite number of other ways).

### 3

In such a brief and abstract outline, theme-dynamic analysis may sound as a simple and straightforward business. In actual fact the analyst must face a vast number of formidable problems, which I have only been able to solve by bold and arbitrary cutting of several Gordian knots. To begin with, topical analysis is often a very subtle task. Present-day semantics do not give us tools for work with sufficient precision. Some simple instances of reference, synonymy, antonymy and hyponymy may be obvious enough. But many others involve delicate semantic connections that different analysts readily view in a different light. At some such points, the unanimity of a team of analysts will increase after weeks and months of collaboration and discussion. At other points, different analysts with various temperaments will at best agree to disagree. Even the basic decision whether a topic in a certain sentence is 'given' or 'new' — in other words, whether the topic links the sentence backwards anaphorically, or forwards cataphorically — is often difficult to make.

In the second step, the one involving thematic analysis of the role of a given topic within its sentence, the difficulty results largely from the different ways in which theme and rheme (or 'topic' and 'comment', as many linguists prefer to call them with a terminology different from mine) can be defined. The conclusion of a scrutiny of available writings on these concepts, including particularly the studies of Functional Sentence Perspective of the Czech linguists[5], is that theme dynamics could well profit from a far more stringent and detailed thematic apparatus than that available so far.

In the third step, that involving a taxonomy of patterns of topical movement in thematic terms, the degree of accuracy and consistency will of course wholly depend on the degree of success of the earlier steps: topical analysis and thematic analysis. And it goes without saying that the fourth step, the stylistic comparison of text with norm, brings with it all the problems of stylistic analysis, including the quest for, and circumscription of, a sensibly chosen norm.

Finally, in the theme-dynamic analysis of many types of text, the very wealth of data will contribute to the investigator's troubles. Clauses and sen-

[5] As developed by František Daneš, Jan Firbas and others. See Zdenek Tyl, ed. 1970. *A tentative bibliography of studies in functional sentence perspective* Prague: Československa Akademie věd.

tences may be interwoven with most intricate patterns of warp and woof. Tracing different types of topical links through a text by joining them with lines, coloured differently for different types of linkage, will result in pretty but confusing pictures. Numerical coding for computerized treatment may be an avenue worth exploring. Even simple displays may, however, sometimes bring out suggestive differences between texts. Coloured tracings of topics will make some texts look like road-maps of the Netherlands and other texts like road-maps of Lapland, showing us something of their relative density of cohesion devices.

Obviously the practical analyst must seek his answer in simplification. As in all linguistics and all stylistics one ought to learn to pinpoint those features that are significant to one's purpose, and to dismiss the rest.

### 4

My own solution was to test some very rough, even oversimplified, methods in the hope that they might suggest improvements and better strategies.

In outline, the procedure was as follows. From a number of written prose texts, 100 macrosyntagm (MS)[6]—to—macrosyntagm transitions were analysed. The analysis began with the first word of the second MS of the text. Dialogue was excluded. The word was checked against the previous MS. If a topical link was found, it was registered; if not, the second word of the second MS was similarly checked against the preceding MS. If there was a topical link, it was noted down; if not, the third word was similarly checked, and so on. As soon as a topical link was found, the analysis moved on to the next MS-to-MS transition. Each MS-to-MS transition was thus classified just once, irrespective of how many topical links there may have been across the MS-to-MS juncture, and always by the first, leftmost constituent that had a topical link with the preceding MS. If a MS did not connect with the preceding sentence, the search continued with the sentence before that. Hopelessly ambiguous or difficult instances were simply omitted.

It the thematic analysis, every constituent from the beginning of a MS up to and including the subject of its main clause was classified as part of the theme. Every constituent to the right of the subject of the main clause was

[6] The term 'macrosyntagm' was borrowed from Bengt Loman (see Bengt Loman and Nils Jörgensen 1971. *Manual för analys och beskrivning av makrosyntagmer.* Lund: Studentlitteratur.) to avoid certain problems in the purely orthographic definition of sentences. Co-ordinate, complete main clauses capable of standing as well-formed sentences were counted as separate macrosyntagms irrespective of the author's own punctuation; a semi-colon thus usually ranked the same as a full-stop. Sentence fragments punctuated as sentences were also regarded as macrosyntagms.

regarded as part of the rheme[7]. Each transition was then classified under one of four headings[8]:

1. theme iteration, involving a topical link between the theme of the MS and a theme of a preceding MS;
2. theme progression, involving a topical link between the theme of the MS and a rheme of a preceding MS;
3. theme iteration, involving a topical link between the rheme of the MS and a rheme of a preceding MS; and
4. rheme regression, involving a topical link between the rheme of the MS and a theme of a preceding MS.

I must once again emphasize that the aim of this very rough procedure was to simplify, not to study exhaustively the cohesion devices of each text. (This was in fact done for some texts in separate 'microanalyses' to supplement the rough 'macroanalyses'.) Any illusions as to the universal applicability of the same procedure to the study of all texts were also quickly dispelled. The method must not be viewed as a single patent procedure recommended for the study of textual cohesion in general. It is rather a starting-point to be modified and developed according to the demands of each text and each textual comparison.

The reliability of the analyses was checked by re-analyses of some texts by the same analyst, and by separate analyses of the same texts by different analysts. The degree of variance varied considerably with the text. Some texts seemed to use theme-dynamic cohesion devices that were relatively unambiguous; repeated analyses yielded roughly the same results. The greater complexity of other texts was reflected in a greater discrepancy between the results of repeated analyses. Experience did, however, suggest that it would be possible to increase the agreement by construing more elaborate manuals with detailed rules for the procedure. Also, one analyst — a linguist by training — tended to focus his attention on formal links between topics, whereas another — a student of literature — was more readily prepared to accept subtle semantic links. Such discrepancies in approach could probably be minimized through intensive training, if such training seems worth while for some particular task.

It should also be remembered that the procedure — always starting from the left of the MS following the transition, and noting just one topical link for each transition — very much favoured the categories called theme iteration and theme progression. The statistics thus become loaded in favour of these two devices at the expense of rheme iteration and rheme regression.

[7] I am here consciously and deliberately using the terms 'theme' and 'rheme' in a very arbitrary, mechanical sense, without wishing to imply that such a definition has any virtues beyond the present experiment.

[8] In fact most analyses were done with a finer net involving 16 categories; they were later merged into 4 major ones.

One possible way of overcoming this bias — apart from a full 'microanalysis' that registers all the topical links — would be to supplement the analysis with another 'in reverse', that is, from right to left instead of from left to right.

5

Some examples of the results can be briefly summarized in tabular form:

Table 1

Distribution of theme-dynamic patterns in 100 macrosyntagm-to-macrosyntagm transitions

| Text | Theme iteration | Theme progression | Rheme iteration | Rheme regression |
|---|---|---|---|---|
| Boswell, *London Journal* | 64 | 27 | 3 | 6 |
| Hemingway, *Men Without Women* | 62 | 23 | 10 | 5 |
| Hemingway, *Islands in the Stream* | 62 | 22 | 9 | 7 |
| Hemingway, *The Snows of Kilimanjaro* | 57 | 20 | 16 | 7 |
| Hemingway, *The Old Man and the Sea* | 56 | 21 | 16 | 7 |
| "National Party Organization" | 58 | 37 | 4 | 1 |
| Zelda Fitzgerald, *Save Me the Waltz* | 56 | 14 | 23 | 7 |
| Gibbon, *Decline and Fall* | 48 | 36 | 10 | 6 |
| Stein, *Autobiography of Alice B. Toklas* | 48 | 15 | 35 | 2 |
| Orwell, *Shooting an Elephant* | 47 | 30 | 14 | 9 |
| "Origins of the British Welfare State" | 46 | 37 | 14 | 3 |
| Johnson, *Life of Milton* | 41 | 44 | 13 | 2 |
| "The Ruling Elite Model" | 37 | 38 | 19 | 6 |
| "Social Class" | 39 | 46 | 11 | 4 |
| "Social Cleavages and Political Parties" | 31 | 56 | 10 | 3 |

Here, Boswell's *London Journal*, an autobiography where the narrator is the main actor, has a high proportion of theme-to-theme links. The same preponderance of theme iteration is characteristic of certain passages in Hemingway. The samples from Zelda Fitzgerald and Gertrude Stein had a higher proportion of topical links between rhemes of successive sentences. The samples from Gibbon, Dr Johnson, and the articles on social science whose

titles have been given in quotation marks contained a higher proportion of links between the rheme of one sentence and theme of the next.

Two of the major stylistic poles that emerge from these materials might be characterized as (1) a thematically iterative, 'static' style, where the same topic is often repeated in successive sentences through thematic, and in some texts also rhematic, linking[9], and (2) a thematically 'dynamic' style, in which there is a relatively high incidence of thematic shift so that the themes of successive sentences tend to be different. Two brief passages may illustrate what I have in mind:

> *I* never had a fire in my bedroom, but one in my parlour in the morning and one in my dining-room in the evening. *I* had my own tea and sugar, and got in bread and butter and milk as I wanted it. In short *I* regulate everything in the most prudent way. At the end of the year *I* shall subjoin a succinct account of my expenses.

Here, Boswell linked his sentences with his ubiquitous *I*, usually in thematic position. As the 'dynamic' style seems to be common in sustained logical argument, it can be readily exemplified for instance from scientific textbooks:

> The ratio of the velocity of light in a vacuum to the phase velocity of light of a particular wavelength in any substance is called the *index of refraction* of the substance for light of that particular wavelength. *Index of refraction* will here be designated by n, stating, if necessary, the particular *wavelength* to which it refers. If no *wavelength* is stated, the index is usually assumed to be that corresponding to the yellow light from a *sodium flame*. The *sodium flame* was chosen because...

As my italics show, this text has frequent links between the rheme of one sentence and the theme of the next.

It is, of course, conceivable that two texts which yield the same proportions of the four cohesion devices — theme iteration, theme progression, rheme iteration, rheme regression — in fact cluster them very differently. One text may mix all devices evenly throughout the text, whereas another may consist of 'static' sections with high incidences of iteration, and 'dynamic' sections with high incidences of progression or regression. Therefore I have also experimented with methods of 'cluster analysis' to measure the degree of theme-dynamic homogeneity within a text. Various statistical approaches seem possible in the study of such clustering.

---

[9] An instance of theme iteration would be: "*John* came in. *He* went to the cupboard."; one of rheme iteration: "John loved *Margaret*. I also admired *her* enormously." Passivization would here result in theme progression: "John loves *Margaret*. *She* is also greatly admired by me.", or in theme iteration: "*Margaret* is loved by John. *She* is also greatly loved by me." The role of passives and of cleft constructions as thematic devices is, of course, well known by now.

## 6

In theme-dynamic analysis, figures are only as exact as the method, tempered by the investigator's shortcomings, allows. This is why I have so far refrained from statistical assessments of the significance of the differences between data from different texts, such as those cited in the above Table.

All the same I am tempted to regard the results as suggestive. They give a rough measure for the difference between a style such as, say, Boswell's or Hemingway's, and that of, say, scholarly argumentation. They also suggest one possible difference between Hemingway and Gertrude Stein in the way in which these two writers developed their argument, irrespective of the many other features they may share. Of course the interpretation of the data must build on sound common sense. For instance, there is a ceiling somewhere for theme progression. A text of any length built on theme progression alone would probably strike one as deviant because of its centrifugal tendency to escape further and further from the original subject and theme. A certain proportion of static iteration and dynamic progression, though within very wide tolerance limits, may therefore be a condition of textual well-formedness. At the moment, however, it would be hazardous to venture a guess as to how much centrifugality a text can tolerate before it strikes the receiver as deviant.

In sum, I think that a study of the ways in which the sentences of a text are interlinked can reveal certain stylistically relevant characteristics of that text. The confluence of stylistics and text linguistics may give the student of style a chance of defining textual style markers, which help us to describe more precisely certain types of stimuli actually present in the text, instead of merely describing their effects in subjective or impressionistic language.