Merja Koskela

Variation of Thematic Structure within a Text

1. Introduction

Previous research (e.g. Nwogu & Bloor 1991; Wiegand 1988; Dubois 1987) has shown that the thematic structure of texts reflects differences of genre. The explanation for this is that different texts belonging to different genres have different functions, which leads to certain characteristic types of thematic choices.

Dubois (1987: 90f.) among others states that the thematic progression of texts necessarily reflects the complexity of ideas presented in the text. Also Nwogu and Bloor (1991: 370) argue along the same line when they suppose that variations of thematic progression are constrained by e.g. purpose, audience and context. Even though there hardly exists a direct cause-effect relation between these factors and thematic structure, it seems reasonable to conclude that these factors are genre-specific and can affect the writer's thematic choices in a certain direction.

Following the same line of thought it can be stated that the thematic structure of text depens on a variety of interdependent factors, from the text's function to its contents. Therefore, it seems reasonable to assume that thematic stucture varies not only in texts representing different genres, but also within texts. For example, the introduction section of a scholarly article obviously has a different function from that of the method section.

The different functions of macrostructural sections of texts have been subject to a lot of research (see e.g. Lindeberg 1994). The primary interest, however, has been the rhetorical structure of text parts (see also Swales 1980). But other kinds of work has been carried out as well. Näslund (1992) for example has studied

the distribution of identity relations in different parts of Swedish LSP texts, and Mauranen (1993) has examined textual relations in academic writing. Of studies that focus on thematic structure the work of Weissberg (1984) should be mentioned. He has carried out a study of paragraph development models in scientific English.

The aim of my study is to find out if there are variations in the thematic structure in different macrostrucural sections of texts from two genres, science and popular science, and what kind of variation, if any, comes into question. The study will be supplemented by an analysis of cohesive devices that exist between themes, or themes and rhemes in order to give a more comprehensive picture of the eventual variations between different parts of the same text.

Before describing the material used in the pilot study a comment on some central concepts is in order. The term *genre* will be used here to describe intertextual differences, that is the difference between science and popular science. Therefore, it must be distinguished from the term *text type*, designating description, narration, exposition, argumentation, and instruction, i.e. intratextual differences (Werlich 1976). My definition of *theme* in this study is based on the Hallidayan (1985) concept of multiple theme. Basically, this means that I distinguish two kinds of thematic elements which are *basis* and *ideational theme* (see Koskela 1994: 635f.). The category of basis consists of elements other than the grammatical subject occupying the first position in the sentence. In practice, the category comprises Halliday's textual and interpersonal themes, mostly adverbials. The concept of *ideational theme* includes elements that concern the contents of the text. In other words, they specify and expand the the textual theme, understood as the constructional basis of the text. The concept of *rheme* is defined as the part of the sentence that follows the theme.

The idea of multiple theme is well suited to Swedish because the theme in Swedish is usually the first item at the beginning of

a sentence, and because the ideational theme is not an obligatory constituent in Swedish sentences. The identification of the sentential themes will, however, be carried out on a textual level with respect to the information previously provided in the text.[1]

(1) ***Vanligtvis*** formas **W-spetsen** rent mekaniskt genom slipning.
'The W-tip is usually formed mechanically by means of grinding.'

The above example stems from a text that introduces different types of microscopes. In the example sentence there is both a basis (*usually*) and a ideational theme (*W-tip*). Sentences where there is only a basis or only an ideational theme are also common in the material.

2. Material

The material of the study consists of 60 paragraphs from the introduction, method and discussion sections of Swedish scientific reports and popular scientific articles in the field of physics (chosen from a corpus of 20 texts). The scientific part of the corpus consists of texts written by experts to experts on topics like new types of microscopes, density measurements, measurements of atmospheric mercury etc. The texts have been published in different university series in Sweden, for example, the report series of *the Royal Institute of Technology* in Stockholm and *Lund Reports on Atomic Physics*. The popular scientific corpus consists of texts written by experts together with journalists to a nonspecialist audience. The texts have been published in the leading popular scientific magazine in Sweden, *Forskning och Framsteg*. The topics of the articles can be, for example, ways of describing atoms, super strings, different dimensions etc.

---

[1] In the examples the ideational theme is marked with bold face and basis with italics.

Because the macrostructures in the different genres differ, I have chosen for comparison those elements which are present in most texts. In both science and popular science there is always an introducing sequence (introduction in my categorization) as well as a concluding sequence (discussion in my categorization). After the introduction there usually follows a material/method section in the scientific texts and a section I have chosen to call a method/theory section in the popular scientific texts. The difference in naming is caused by the fact that the middle section in popular science generally presents central theoretical issues rather than the material used in an experiment.

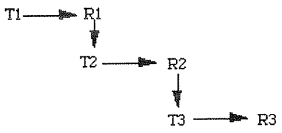
The analysis will primarily be carried out in sentences. With *sentence* I understand the main clause together with any dependent clauses. Accordingly, the thematic structure of subordinated clauses will not be considered separately in this study. However, the basic unit of analysis is the *paragraph*, understood as a typographically marked unit (see Enkvist 1991: 9) which is natural because the paragraph is a primary rhetorical and cognitive unit (see e.g. van Dijk 1981). However, only paragraphs with more than five sentences were admitted in the material in order to account for thematic patterns. The somewhat arbitrary restriction of the material to include only paragraphs with more than five sentences is necessary for the study of the types of thematic patterning because they normally need at least that much text for their realisation. The mean length of a paragraph in the whole corpus of 20 texts is 5.2 sentences in scientific texts and 3.9 sentences in the popular scientific texts, whereas the paragraphs studied show the following values: 7.7 sentences in scientific texts and 5.9 in the popular scientific texts. This difference can affect the results to some extent and must therefore be taken into consideration when explaining the results.

3. Method

The method of the study is an application of the method used by Weissberg (1984: 492) in the study of scientific research reports from three different subject fields. My analysis consists of two parts. In the first part the types of thematic patterning in each paragraph have been studied and mapped. With a thematic pattern I understand the types presented by Danes (1974). (For criticism see e.g. Heinemann & Viehweger 1991: 35; Gülich & Raible 1977: 86, Makovec-Cerne 1992: 436; Koskela 1995.)

According to Danes (1974: 118f.) there are three main types of *thematic patterning* (TP), namely linear (TPI), continuous (TPII), and a pattern with derived themes, also called the hypertheme pattern (TPIII). In addition, Danes mentions the exposition of a split rheme (TPIV) as a framework for combining the other types, and the thematic pattern with an omitted link (TPV). These patterns are found in different combinations in real texts.

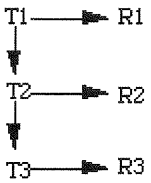
The linear pattern (TPI) can be illustrated as follows:



(2) 1. *Det* finns för närvarande inget sätt att identifiera vilket slags atom som mikroskopet avbildar.
2. **Detta** är ett allvarligt problem eftersom endast atomstrukturer hos elektriskt ledande ytor kan undersökas.
3. och **ytor** har mycket lätt för att förorenas av omgivningen.

'1. There is at the moment no way of identifying which type of atom the microscope reproduces.
2. This is a serious problem because atomic structures can only be studied on surfaces that conduct electricity,
3. and surfaces can easily become contaminated by the environment.'.'

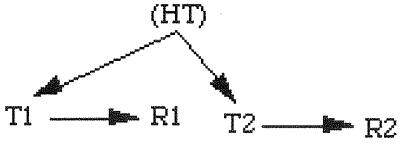
The constant pattern (TPII) can be found in the example below:



(3) 1. **Supersträngteorin** tycks vara fri från alla de obestämdbarheter och motsägelser som har vidlådit s k kvantteorier alltsedan de uppfanns på 1930-talet.
2. **Teorin** förenar på ett realistiskt sätt naturens krafter till en enda.
3. **Den** innebär också att man får en teori om tyngdkraften (gravitationen) i mikrokosmos, dvs atomernas värld.

'1. The theory of super strings seems to be free of all the indeterminacies and conflicts that have been a part of the so called quantum theories ever since they were invented in the 1930's.
2. The theory combines the forces of nature to a single one in a realistic way.
3. It also means that we get a theory of gravity in microcosmos, that is in the world of atoms.'

The hypertheme pattern (TPIII), where the thematic elements derive from a hierarchically superordinate theme, can be described as follows:



(4) Instrumenttyper för mätning av brytningsindex
1. **Refraktometer** bygger på mätning av kritiska vinkeln
2. **Interferometer** utnyttjar inferensfenomen.

'Types of instruments for measuring the refracting index
1. Refractometer is based on the measuring of the critical angle.
2. Interferometer exploits inferential phenomena.'

In the study I have categorized a paragraph as belonging to a particular pattern if at least three consecutive sentences (two links

between sentences) conform to that pattern. If the same paragraph includes different patterns, it has been classified as "mixed". A paragraph has been labeled "no pattern" if no sequences with three sentences conforming to a pattern have been found.

In the second part of the analysis I have studied four types of cohesive devices used in the thematic part of each sentence (after the first sentence) to refer to recoverable information earlier in the text. The cohesive devices analyzed are *identity*, *synonymy*, *pronouns* (explicit links), and *inference* (implicit links). From this part of the analysis I have excluded summative expressions (e.g. *these results indicate...*) analysed by Weissberg (1984) because they do not seem to be comparable with the cohesive devices. In my study these experessions will be categorised as inference.

4. Results

Part 1

In the first part of the analysis I have categorised the paragraphs in the material according to the dominant type of thematic patterning. This means that if there is a three-sentence sequence conforming to a pattern, the whole paragraph is classified as belonging to it.

Weissberg's (1984) results show that one of the three thematic patterns was more likely to occur than was no pattern at all. The only exception was the paragraphs taken from the part that Weissberg calls Methods/Materials section. In this section no pattern was more common than in the other sections. According to Weissberg's findings the linear pattern was the most common, especially in the introduction and discussion, whereas the mixed pattern had the second and the hypertheme pattern the third highest occurrences. The constant theme pattern was the least usual type of patterning. With respect to the sections Weissberg concludes that the hypertheme pattern is common in the Methods/Materials

section, as a way of creating a procedural description. Such a description of how the experiment was carried out is essential in scientific texts because there is a requirement to present the method so clearly that it will be possible for any other scientist to repeat the experiment. According to Weissberg's results the hypertheme progression is useful for such a purpose. However, also no pattern was very common in these paragraphs.

Table 1. Distribution by pattern of 30 paragraphs in three sections of scientific texts.

Scientific Physics				
Pattern	**Introduction**	**Material/method**	**Discussion**	**Total**
Some pattern	5	5	6	16
Linear	*1*	*0*	*3*	*4*
Constant	*1*	*0*	*0*	*1*
Hypertheme	*2*	*1*	*3*	*6*
Mixed	*1*	*4*	*0*	*5*
No Pattern	5	5	4	14
Total	**10**	**10**	**10**	**30**

In the whole material from the scientific texts, some pattern seems to be somewhat more usual than no pattern at all. However, this is true only for the discussion section where some pattern occured more often than no pattern. In the introduction and material/method sections no pattern was as usual as some pattern.

In this respect my results deviate from those of Weissberg in that the number of patterns in his study is higher. One reason for this could be that my interpretation of what actually is a pattern is different from Weissberg's. The interpretation always depends on the analyst's judgement on which relations constitute a relevant connection between a theme/a rheme and the following theme. The following example of a linear pattern illustrates what kinds of connections have been judged as relevant in my study.

(5) 1. **CERN** kommer att fördubbla den tillgängliga energin vid LEP-acceleratorn redan år 1994.
2. **Nästa steg** blir att i samma tunnel sätta in mycket kraftigare magneter.

3. *Det* är **de** som styr partiklarna i deras banor runt acceleratorringen.

'1. CERN is going to double the accessible energy in the LEP-accelerator as soon as 1994.
2. The following step will be to assemble much stronger magnets in the same tunnel.
3. It is those that control the particles on their orbits around the ring of the accelerator.'

Another obvious source for differences is, of course, the language studied: Weissberg's material was in English and my material is in Swedish. A third possible factor is that my texts represent physics, whereas Weissberg's material comes from agriculture, biology and engineering. Unfortunately, he does not report the proportions of the different fields of science in his material. Neither does he take into consideration the possible differences between the different fields of science.

In the paragrahps with some pattern in my material the hypertheme pattern was found to be the most common, but also the mixed type of paragraph (i.e. a paragraph where there is both the hypertheme pattern and either the linear or the constant pattern) was usual. The linear pattern accounted for the third highest occurence, whereas the constant theme pattern occured least frequently.

Even this result is partly different from Weissberg's findings, since the linear pattern was the most common in his material. Still, the constant pattern was the least usual in both studies.

The sections studied differ from each other with respect to the dominant pattern. In the introduction section all patterns occur, but the hypertheme pattern is the most common, whereas in the discussion section the linear pattern and the hypertheme pattern are the only types that occur, both three times.
Weissberg received the most interesting results from what he calls the Methods/Materials section. In my study the corresponding section, called the material/method section, differs from the rest of the sections in that the mixed pattern is clearly the most usual

one. This means that there are in total more sequences with three sentences conforming to a pattern in this section than in the other sections. One probable explanation for this is that this section has the longest paragraphs in my material (mean 8.9 sentences/paragraph).

However, Weissberg's (1984: 493) results showed that no pattern was the most common in this section together with the hypertheme pattern. Unfortunately it is not possible to control how long his paragraphs were, but the fact that the hypertheme pattern occured only once in my material/method sections lets us assume that the paragraph length is not the only reason for differences in the results.

Table 2 shows the distribution of the most frequent thematic patterns in the popular scientific texts in my material.

Table 2. Distribution by pattern of 30 paragraphs in three sections of popular scientific texts.

Popular Physics				
Pattern	**Introduction**	**Method/Theory**	**Discussion**	**Total**
Some Pattern	5	7	3	15
Linear	2	4	1	7
Constant	3	1	2	6
Hypertheme	0	1	0	1
Mixed	0	1	0	1
No Pattern	5	3	7	15
Total	**10**	**10**	**10**	**30**

In the popular scientific material there are 15 paragraphs where there is some thematic pattern to be found and 15 paragraphs where there is not any. Here the shorter paragrahps may again play a role. The sections, however, behave somewhat differently. Whereas the introduction conforms to the picture of the whole material, the method/theory section is clearly structured with some pattern, and the discussion section is structured more often without any pattern.

It is interesting to note that in the introduction section of the popular scientific texts the constant pattern is somewhat more common than the linear pattern. Weissberg's result show that the constant theme is the least frequent type in the introduction sections of scientific reports, and even in my scientific material there is only one instance of constant theme among the paragraphs.

The popular scientific method/theory section is most often constructed with a linear pattern, but all the other types of thematic patterning occur as well. The discussion section is mostly constructed without any thematic pattern, as in the example below:

(6) 1. **Instrumenten** är nu så känsliga att de kan känna av en vibration som består av en enda fonon, den s k kvantgränsen.
2. **Därmed** uppstår ett oväntat problem.
3. **Man** mäter vanligen två saker, amplituden ... och fasen ....
4. **Kvantmekaniken** säger att det inte går att samtidigt mäta både amplituden och fasen hur noga som helst.

'1. The instruments are at present so delicate that they are able to sense a vibration that consists of one single phon, the so called quantum border.
2. Here appears an unexpected problem.
3. Two things are usually measured, the amplitude and the phase.
4. Quantum mechanics maintains that it is not possible to measure both at the same time very accurately.'

In the few paragraphs where there is a pattern in the discussion section this is either the constant theme pattern or the linear pattern. In this respect the discussion section is similar to the introduction section where these two basic types of thematic pattern dominate.

To summarize, the clearest difference between the scientific and popular scientific material seems to be that in the cases where there is some pattern it is most likely that it is the hypertheme pattern or the mixed pattern in scientific texts, but linear or constant pattern in popular scientific texts. The reason for this

probably lies in the function of the texts. The hypertheme pattern is a tool for analysis and analysis is an important part of scientific writing. The linear pattern is often described as dynamic (Wiegand 1988: 32) and typical of explaining (Nwogu & Bloor 1991: 375), and the constant pattern is often used in describing. Both of these functions can be considered to be essential in popular science (Nwogu & Bloor 1991: 377). Presenting phenomena for a lay audience and making it interesting requires structures that are easy to process, and the basic types of thematic patterns seem to be well suited for this purpose.

On the whole, it becomes clear that the macrostructural sections differ from each other in their thematic organisation both in science and in popular science. In scientific texts the typical patterns are somewhat different in each section whereas the number of paragraphs with no pattern stays almost the same. In popular scientific texts the method/theory section is clearly different from the introduction and discussion sections, and the number of paragraphs with no pattern varies more than in the scientific texts.

Part 2

Weissberg completed his analysis of the thematic patterns in paragraphs with a study of intersentential linkage. My analysis of cohesive devices between themes or themes and rhemes includes four types of cohesive devices which can be divided into explicit links (identity, synonymy, and pronouns) and implicit links (different types of inference). Obviously, this categorization does not cover all possible intersentential relations. Hierarchical devices and contrast, for example, are left out of the analysis. However, to account for differences within texts from science and popular science these categories should be sufficient to reveal some characteristic tendencies.

Weissberg's (1984: 493f) results show that the most common type of cohesive device was repetition (53%), but also inference (bridging) was found to be usual, especially in paragraphs from the methods/materials section. In introduction paragraphs Weissberg found many subject pronouns and little inference. From this Weissberg concludes that introduction paragraphs should be easier for laymen to understand than the other parts of the scientific texts.

Table 3. Frequency of occurrence of four cohesive devices among 202 intersentential links, by section of the scientific reports.

ScientificPhysics							
Device	**Introduction**		**Method/Materia**		**Discussion**		**Total**
		%		%		%	%
Identity	19	32.2	36	45.6	16	25.0	71	35.1
Synonymy	5	8.5	6	7.6	5	7.8	16	7.9
Pronoun	7	11.9	5	6.3	3	4.7	15	7.4
Inference	24	40.7	23	29.1	37	57.9	84	41.6
Other	4	6.8	9	11.4	3	4.7	16	7.9
Total	**59**	**100.1**	**79**	**100**	**64**	**100**	**202**	**99.9**

Table 3 shows the results of the analysis of the cohesive devices in the 30 paragraphs from scientific reports. The frequencies differ somewhat with respect to macrostructural section, but regardless of the section the two most frequent types of cohesive devices were inference and identity.

The most usual type of cohesive device in the introduction and discussion sections is inference which accounts for over 50 % of all intersentential relations in the discussion section. A typical case of inference is one where the reader's previous knowledge, either general knowledge or specific knowledge of the research process, is used as the source. The following extract stems from a discussion section:

(7) 1. **Den primärt beskrivna metoden, uppsökande av absorptionsmaxmium** visade sig vara mycket starkt beroende på mätförstärkarens funktion, dvs. hur exakt man kunde sampla signalen från fotodioderna.

2. **Den preliminära avsikten** var att bygga två likadana mätkort och sedan dividera mätvärdena i datorn.

'1. The primarily descirbed method, the search for the absorption maximum turned out to be strongly dependent on the function of the amplifier, that is how exact the signal could be sampled from the photodiodes.
2. The preliminary aim was to build two similar circuit boards and then divide the measured values on them in the computer.'

In the introduction section inference is somewhat less frequent. My results also seem to support Weissberg's finding that subject pronouns are used more in the introduction section than in the other macrostructural sections. However, this result is in contrast to Näslund's (1992: 126) finding based on technical texts. According to Näslund, the phenomena discussed is related to individual human beings or the background situation in discussion sections, not in the introduction.

Regardless of the fact that the differences between the sections are too small for any definite conclusions, it is interesting to note that the result above seems to confirm the tendency that identity relations are used less in the beginning of a text than further on. This is probably because the beginning always focuses on introducing new referents (Näslund 1992: 125; Mauranen 1993: 91). Another tendency is that the amount of inference tends to increase from introduction to discussion, the method and material section being an exception. One explanation for this is that the shared knowledge between the writer and the reader increases during the reading. This means that the writer can expect more from the reader (Mauranen 1993: 139) in terms of what connections will be experienced as coherent.

The method/material section in my material is clearly more explicitly structured than the rest of the material. In these paragraphs 45.6% of all intersentential links are based on identity relations.

A typical case of identity is presented in the following example from a method/material section:

(8) 1. **Den kritiska ljusstrålen** går genom provet parallellt med gränsskiktet mellan prov och undre prisma.
2. **Denna ljusstråle** avböjs olika beroende på skillnaden i brytningsindex mellan prov och prisma.

'1. The critical light beam penetrates the specimen in parallell with the line between the specimen and the lower prisma.
2. This light beam deflects in different ways depending on the refracting index between the specimen and the prisma.'

The finding that there is a lot of identity in the method/material sections differs from Weissberg's results, but it seems motivated. Because the purpose of a scientific text is to present experiments so explicitly that they can be replicated, a greater linguistic explicitness is called for as well (cf. Mauranen 1993: 95). This adds to the frequency of identity in scientific texts.

Table 4. Frequency of occurrence of four cohesive devices among 146 intersentential links, by section of the popular scientific articles.

PopularPhysics							
Device	**Introduction**		**Method/theory**		**Discussion**		**Total**	
		%		%		%		%
Identity	26	54.2	21	45.7	18	34.6	65	44.5
Synonymy	5	10.4	3	6.5	8	15.4	16	11.0
Pronoun	6	12.5	8	17.4	8	15.4	22	15.1
Inference	8	16.7	8	17.4	11	17.3	27	18.5
Other	3	6.3	6	13.0	7	13.4	16	11.0
Total	**48**	**100.1**	**46**	**100**	**52**	**100**	**146**	**100.1**

In the popular scientific texts identity is clearly the most usual type of intersentential link in all sections studied (44.5% of all the links). The results also show that the explicit types of cohesive devices are much more frequent in popular science texts than in the scientific texts whereas inference occurs much less frequently in popular scientific text even though it has the second highest frequency of all cohesive devices.

The links in popular science are usually very explicit. The following example is an extract from an introduction to a popular scientific article. There is a pronoun linking the second sentence

to the first one, and a identity relation between the second and the third sentence:

(9) 1. **Väteatomen** är den enklaste av atomer.
2. **Den** har en kärna som utgörs av en enda positivt laddad partikel – protonen.
3.**Kärnan** omges av en negativt laddad partikel – elektronen.

'1. The hydrogen atom is the simplest type of an atom.
2. It has a nucleus that consists of one single positively charged particle – the proton.
3. The nucleus is surrounded by a negatively charged particle –the electron.'

The paragraphs from the introduction section have the highest frequency for identity (54.5%) and the lowest frequency for inference (16.7%). Subject pronouns creating linkage between themes or themes and rhemes occur less frequently in the introduction section than in the other sections studied.

In the popular scientific articles there is usually no explicit method description, because methodological details are not interesting for a lay reader. However, the essential features of the theories that the methods are based on must be explained. In these explanations identity is the most frequent type of cohesive device and pronouns are as frequent as inference.

Even in the popular scientific articles the discussion sections are similar to the introduction sections. The paragraphs in discussion sections are characterized by a high occurrence of identity (34.6%) and a somewhat lower occurrence of inference (17.3%). These paragrahps have the highest frequency of synonymy of all the sections (15.4%).

Synonymy can be used towards the end of a popular scientific text because the shared knowledge between the writer and the reader has increased during the reading. In the following example can the Riemannian geometry be interpreted as synonymous to advanced mathematics.

(10) 1. *Dessutom* knyter **strängteori** an till avancerad matematik på ett helt nytt sätt.
2. *Bland annat* kommer **s.k. Riemanngeometri** till användning...

'1. The string theory is also connected to advanced mathematics in a whole new way.
2. Among others can the so called Riemannian geometry be put into use...'

In the popular scientific introduction sections there are also many pronouns, but not as many as in the method/theory section.

5. Discussion

The present study set out to examine whether there are variations of thematic structure within texts. Thematic patterns in texts from two genres, scientific and popular scientific texts from physics were studied, and the study was complemented with an analysis of four cohesive devices.

The method of the study is based on previous research done by Weissberg, but the results of the study differ in many ways from his results. My study shows that there are variations in thematic patterning between the macrostructural sections of texts in both genres. In scientific texts the typical patterns differ in each section: hypertheme in introduction, mixed pattern in material/method and linear pattern and hypertheme pattern in discussion sections. In popular scientific texts the method/theory section is clearly linearly structured, whereas in the introduction and discussion sections the basic types of thematic patterning, the linear pattern and the constant theme pattern, prevail.

The results of the analysis of cohesive devices are predictable on the basis of the analysis of the thematic patterns: where there are basic types of thematic patterns, there are explicit cohesive relations. The analysis shows that in the scientific texts inference is the most common type of textual relation, whereas identity is in general most common in popular science.

These results confirm two expectations. First, that the different functions scientific and popular scientific texts have are reflected in the thematic structure of the text as intertextual differences. And second, that the macrostructural sections differ as to their functions to such a degree that the differences are reflected on the thematic structure. These intratextual differences can at least partly be explained by the use of different text types.

As for the method of study a word of criticism is in order. Especially two aspects, the unit of analysis and the use of clear patterns need to be reevaluated. Even though a paragraph is a primary rhetorical unit it does not have autonomy with respect to the rest of the text. Therefore the use of larger chunks of text for the analysis should be considered in order to account for relations that exist between thematic (and rhematic) elements that are located in different paragraphs.

Another question related to the unit of analysis is that only paragraphs with more than five sentences have been accounted for. Usually this is not a problem in scientific texts where the paragraphs are relatively long, but so long paragraphs cannot be considered typical of popular scientific texts. Even this calls for larger units of analysis.

As for the clear patterns it can be stated that the readers are no doubt able to experience thematic connections over longer stretches of texts than the previous sentence. Therefore an account where all thematic elements are categorized according to the type of thematic pattern would give more accurate information. This would even make the supplementary analysis of cohesive devices less important because the difference between implicit and explicit links would be revealed from the analysis of thematic structure.

Primary Material

The Scientific Texts:

Bergwall, Staffan (1989) Sveptunnelmikroskopet – ett nytt sätt att se in i atomernas värld. Teknisk rapport. Tekniska Högskolan i Luleå.
Edner, Hans & Olsson, Per & Ragnarson, Pär (1991) Undersökning av kvicksilveravgång från fibersediment med Lidar-teknik. Lund Reports on Atomic Physics 126. Lund.
Edner, Hans & Olsson, Per & Wallinder, Eva (1990) Mätning av kvicksilver med optisk teknik vid Råcksta krematorium. Lund Reports on Atomic Physics 118. Lund.
Edner, Hans (1989) Optiska metoder för mätning av kvicksilverinnehållet i rökgaser från krematorier. Lund Reports on Atomic Physics 110. Lund.
Eriksson, Jarl-Thure (1991) Effektivare energinät med supraledning. Elteknik. Utveckling & Miljö. (FUD). Vattenfall.
Jansson, Anders (1991) Bildkvalitetsmätning på bildskärmar. Ljusflödets temporala variationer från CRT-bildskärmar. Linjäritetsbestämning av fotokänsliga matriskameror. The Royal Institute of Technology. Department of Physics. TRITA-FYS 2109.
Lindgren, Bosse (1991) Från atomer till kluster. Kosmos pp. 19–42.
Ljunggren, Kristofer (1992) Metod för våglängdskalibrering av färgämneslaser. Lund Reports on Atomic Physics 132. Lund.
Persson, A. & Svanberg, S (1989) Högeffektlaserfacilitet vid Lasercentrum Lund. Lund Reports on Atomic Physics 109. Lund.
Sörheim, Bo (1989) Bestämning av densitet hos polyeten m.h.a. dielektricitetskonstant och brytningsindex. Lund Reports on Atomic Physics 97. Lund.

The Popular Scientific Texts:
(FoF stands for the magazine Forskning och Framsteg.)

Bergström, Lars & Johansson, Erik: Hur fick materien sin massa? FoF 3/93. 1993.
Bergwall, Staffan: Den svåra konsten att avbilda atomer. FoF 7/90. 1990.
Blomberg, Clas: Varför fungerar livets maskineri som det gör? FoF 6/90. 1990.
Brink, Lars: Supersträngar. FoF 5/87. 1987.
Johansson, Erik: Nu vet vi att universum är uppbyggt av tre familjer. FoF 2/91. 1991
Karlsson, Erik: På spionuppdrag i en metallklump. FoF 8/86. 1986.
Lindgren, Bosse: "Uppenbarligen spelar vår Herre fotboll". FoF 3/90. 1990.
Lindström, Ulf: Visst kan det finnas tio dimensioner. FoF 5/87. 1987.
Rosquist, Kjell: Universums gravitationsstrålning: Förr eller senare upptäcks den! FoF 2/88. 1988.
Skagerstam, Bo-Sture: Kosmiska strängar förklarar mycket – men finns dom? FoF 5/88. 1988.

References

Danes, Frantisek (1974) Functional sentence perspective and the organization of the text. In Frantisek Danes (ed.) *Papers on Functional Sentence Perspective* pp. 106-128. Hague: Mouton.

Dubois, Betty Lou (1987) A reformulation of thematic progression typology. *Text* 7(2): 89-116.

Enkvist, Nils Erik (1991) Discourse strategies and discourse types. In Eija Ventola (ed.), *Functional and Systemic Linguistics. Approaches and Uses* pp. 3-22. Berlin etc.: Mouton de Guyter.

Gülich, Elisabeth & Raible, Wolfgang (1977) *Linguistische Textmodelle*. München: Wilhelm Fink Verlag.

Heinemann, Wolfgang & Viehweger, Dieter (1991) *Textlinguistik. Eine Einführung*. Tübingen: Niemeyer.

Koskela, Merja (1994) Identifying the theme-complex in Swedish scientific and popular scientific texts. In Magnar Brekke, Øivin Andersen, Trine Dahl & Johan Myking (eds.), *Applications and Implications of Current LSP Research* pp. 630-638. Vol II. Bergen: Fagbokforlaget. .

Koskela, M. (1995; forthcoming). Thematic Patterning in Scientific and Popular Scientific Texts. In Hartwig Kalverkämper & Klaus-Dieter Baumann (eds.), *Fachliche Textsorten. Komponenten - Relationen - Strategien*. Gunter Narr Verlag: Tübingen.

Lindeberg, A-C. (1994). Rhetorical conventions in the discussion/conclusion sections of research articles in finance, management and marketing. In Magnar Brekke, Øivin Andersen, Trine Dahl & Johan Myking (eds.), *Applications and Implications of Current LSP Research* pp. 647-655. Vol II. Bergen: Bagbokforlaget.

Makovec-Cerne, Jasna (1992) Die thematische Organisation von Texten. Eine kontrastive Untersuchung. *Folia Linguistica* 26: 435-452.

Mauranen, Anna (1993) *Cultural Differences in Academic Rhetoric*. Frankfurt am Main etc.: Peter Lang.

Nwogu, Kevin & Bloor, Thomas (1991) Thematic progression in professional and popular medical texts. In Eija Ventola (ed.), *Functional and Systemic Linguistics. Approaches and Uses* pp. 369-384. Berlin etc.: Mouton de Guyter.

Näslund, Harry (1992) Identity relations and superthemes in Swedish LSP texts. In Ann-Charlotte Lindeberg, Nils Erik Enkvist & Kay Wikberg (eds.), *Nordic Research on Text and Discourse. Nordetext Symposium 1990* pp. 119-128. Åbo: Åbo Akademis Förlag.

Swales, John (1980) *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.

Van Dijk, Teun A. (1981) Episodes as units of discourse analysis. In Deborah Tannen (ed.), *Analyzing Discourse: Text and Talk* pp. 177-195. Washington DC: Georgetown University Press.

Weissberg, Robert (1984) Given and New: Paragraph Development Models from Scientific English. *TESOL Quarterly*. Vol 18, No 3: 485-499.

Werlich, Egon (1976) *A Text Grammar of English*. Heidelberg: Quelle & Meyer.

Wiegand, Ines (1988) Die thematische Progression als Kriterium zur Textsortendifferenzierung. In Rosemarie Gläser (ed), *Fachtext als Instrument und Resultat kommunikativer Tätigkeit* pp. 29-35. Leipzig: Karl-Marx-Universität.

Merja Koskela
Dept. of Scandinavian Languages
P.O. Box 297
65101 VAASA
FINLAND
E-mail: mko@uwasa.fi