KRISTA VARANTOLA: POPULARIZATION STRATEGIES AND TEXT FUNCTIONAL SHIFTS IN SCIENTIFIC/TECHNICAL WRITING

Background.

Quantitative linguistic information of the characteristics of SL (special language) style, particularly of sci/tec texts, has often been controversial. It could also be claimed that conflicting or misleading information of this type has been due to corpora which have been either too large or too small. Results based on small samples of material are not generalizable even if they provide a good intuitive basis for further research. It is a tempting mistake, possible also in this study, that has already been made in a number of qualitative analyses. Such analyses are based more or less exclusively on intuition that is proved correct by suitable examples. With small sample generalizations, then, we encounter such claims as:

- the most common tense in sci/tec discourse is the present/past/present progressive
- there is no unified scientific style
- there is a unified scientific style
- there is the style of physics and the style of chemistry but no common style for the two
- it depends on writers’ individual styles whether they resemble each other or not

Thus it becomes obvious that the topic, its treatment, the publication in question, subject-, culture- or country specific conventions, individual style and idiosyncrasies or mannerisms play an excessive role in short extracts used as data for generalizations.

The large corpus represents different problems. Such corpora certainly offer a good overview of sci/tec style and a fair idea of the conventions prevailing in a particular field whether narrow
or wide, i.e. the controlling and standardizing effects of the particular community (national or international).

If, however, the corpus is large, and heterogeneous, embracing material from specialized texts, textbooks, encyclopaedias, manuals, operating instructions, etc., important information of field-internal variation becomes diluted by the sheer volume of data. The linguistic information made available gives a good general point of reference and shows the main quantifiable differences between special and general language. Didactically or pragmatically, however, the results are not very useful. They cannot be directly applied in the training of future translators or students of science.

Instructions such as

- Use the passive in 25-30% of your finite clauses
- Avoid pronominal reference, other than it and they
- Make your subject NPs 3-4 words long on average and NPs in other functions 2-3 words longer

are unthinkable.

Obviously, a different approach is called for if the following types of questions are asked:

What happens to a specialized text when it is redirected to a new audience?

More specifically, what happens when a technical exposition is turned into an application for a patent, into a maintenance manual, operating instructions for non-experts, a popularized newspaper article, a memo for political authorities to base their decisions on, etc.?

The number of possible textual realizations is obviously high and mixed even if it is possible to group them under more abstract umbrella terms such as text type or text form.

The aims of this study.

This study is an attempt to consider the process described above from the following angle:

What happens to SL texts when their communicative target is re-assessed or when special subject matter is presented with different communicative functions in mind? What kind of hypotheses could be reached and what kind of linguistic criteria would be relevant?

Are there consistent linguistic features that can be used in contrasting special texts with popularized texts? What SL features seem to be most resistant to a change in communicative purpose? Do any valid linguistic advice patterns emerge that could be used to give SL translators, popularizers, bureaucrats, scientists, etc., advice on how to check texts linguistically?
In other words, is it possible to identify linguistically features that characterize a popularization as good, bad or indifferent, or a text dealing with special subject matter as an authentic or unauthentic text? To what extent are such features - always assuming they exist - "international"?

Reassessment of the communicative target.

Douglas Biber (Linguistics 1985:338) makes a distinction between micro- and macro-scapic textual analysis. A microscopic analysis provides a detailed description of the communicative functions of particular linguistic features, such as the function of passives in two scientific articles or the role of quotation marks and italics in representing speech in writing. Macroscopic textual analyses, on the other hand, seek to define the overall textual dimensions of variation within a set of texts.

I shall not go deeper into Biber’s elaborate factor analysis but only stress that his detailed analysis of textual dimensions provides an excellent source of information and novel ideas even for a small scale study.

My own approach might be described as combining the macroscopic and the microscopic. It starts at the textual level, from text strategic considerations and abstractions, moves on through actual textual realizations to syntactic and lexical choices.

Naturally, the more concrete levels do not exist in isolation. There is obviously a constant interplay between textual, syntactic and lexical choices which all together account for the actual realization of a particular surface text.

Let us imagine a situation where we have to change scientific journal articles in a very special field (one in astrophysics and another in theoretical physics) into a popular article comprehensible at least to the educated layman. (A good description of how this is done in practice and what kind of factors are involved may be found in Dubois 1986).

One of the first strategic choices could be to change an expository/descriptive + argumentative text into a purely descriptive piece of writing or perhaps into a descriptive + narrative text in order to add a more personal or didactic touch. If our task were different, say, to change a technical description of a new invention into a patent application, the text strategic choices would be much more strictly controlled by convention, legal demands and practice and there would be very little left to genuine free choice. Similarly, if we were to turn a description into operating instructions, our strategy choices would also be very limited, but in a different way.

To return to the first metamorphosis; in the popularization of very specialized information, the strategy choices would probably centre around the problem of reassessing the new audience. The
assumption is that it takes an expert to popularize highly specialized subject matter. If this is done by a layman, for instance by a journalist working alone, the resulting text is usually inadequate or simply incorrect. (See, however, Dubois 1986 for examples of how laymen can cope with such problems.) The reassessment process could be formulated as follows.

What is the basic text type required for the new version? What does the new audience know about the matter in advance? How much and what should they know after reading the popularization, i.e., what kind of information is irrelevant because the new reader has no chance of understanding the issues without expert background? How should the remaining information be simplified or made more concrete?

I analyzed, from the layman's angle, extracts from two sets of texts for this purpose, two special texts and their popularizations, one on astrophysics and one on theoretical physics, and tried to find out what had happened in the popularization process.

The texts were:


Notice that at least one author is the same in both versions. The text extracts were about 2,000 words long and come from the same point in the discourse, in terms of the treatment of the subject matter.

Textual level changes.

The new text strategies of the popularizations could be described as descriptive + narrative. Prominent textual realization level features were:

The building up of extralinguistic knowledge and simplification.

Redundancy was built in through repetitions, paraphrases and in-text summaries, often by means of rhetorical questions or other figures of speech that acted as some kind of mnemonics. Simplification of subject matter was achieved through visual illustration, everyday analogies or other concretization, shorter paragraphs that dealt with one main issue only. The CIO (crucial information only) principle was relaxed in the popularization. The lacking extralinguistic background was built up through careful
definition of new concepts. All concepts were overtly interrelated, i.e., their relation to previous concepts was carefully explained. No concepts cropped up out of the blue as is often the case in specialized texts where a high degree of extralinguistic inference is acceptable. Also, typographic means could be used to help the reader to go back to a concept and its definition if necessary.

Level of elaborateness.

The popularizations were less elaborate, less exact. Details were avoided and only main points concentrated on. Debatable points were occasionally brought up but not discussed. All these aspects were reflected as a much shorter text than the original and as a rather general level discussion. (Cf. Källgren 1979:7 for an opposite case. She defends the rights of specialists to use specialist language. She points out in the preface to her doctoral thesis that, had the thesis been written for a lay reader, it would have been a different and a much longer work.)

Syntactically and lexically this meant that exact quotations became rare; exact references were dropped as well as mathematical formulae; numerical hedging was the rule and verbal hedging became more common. (See also Dubois, forthcoming, for the use of hedges in scientific contexts.)

(Numerical hedging sometimes results in bizarre translations. In one case, a Swedish newspaper report dealing with bank robberies described the stolen amounts as about 100,000 SEK and 200,000 SEK. These were in a Finnish newspaper referred to as 71,500 FMK and 143,000 FMK, respectively.)

"Rules of thumb" were created lexically, i.e., if something was the outcome in the majority of cases, it became the norm; minority outcomes were ignored. E.g.

Supernovae of type II are generally believed to occur at the end of the evolution of more massive stars 1-10$^3$ in the range of $8$ to perhaps $100$ solar masses (Mo). When the nuclear fuel in the core of the star is exhausted, the centre of the star begins to undergo gravitational collapse, and falls inward until the core rebounds, or "bounces"... (Nuclear Physics 481-482)

The death of a large star is a sudden and violent event. The star evolves peacefully for millions of years, passing through various stages of development, but when it runs out of nuclear fuel, it collapses under its own weight in less than a second. (Scientific American, May 1985:40)

Whereas the sun could have a lifetime of $10$ billion years, a star 10 times as massive can complete its evolution 1,000 times faster. Loosely speaking, when ordinary matter is compressed, higher density is achieved by squeezing out the empty space between atoms. (Scientific American, May 1985:42)
Control of the rate of information and information density.

Filler material was frequent. It helps to control and regulate the rate of new information. It gives the reader a break by dealing with something less weighty, e.g., background information or historical facts, personal experience, anecdotes or case studies. The same method can be used in specialized contexts if the text form allows or demands it, e.g. in professional journal reports that are no less specialized than many scientific articles, an objective, impersonal style may be modified in a more personal, even subjective direction.

Lexical and syntactic changes.

How are these trends reflected at lexical and syntactic levels? The terminology used is affected in the following ways: the term depth is lower. Term use is more general and closer to everyday usage, e.g. supernovas vs. supernovae, a large star vs. more massive stars, in the range of \( 28 \) to perhaps 100 solar masses (\( M_\odot \)). Very special terms are avoided. Term use is also looser, in the style of "the whale is the biggest fish in the ocean, although it is not strictly speaking a fish." The rate of terms is lower and when the terms are introduced they are also defined and their relation to the other terms is explicitly stated, e.g.

...becoming a white dwarf: a burned-out star that emits only a faint glow of radiation.

Chandrasekhar showed there is a limit to how much pressure can be resisted by the electrons' mutual repulsion. ...Equilibrium is possible only if the mass is less than a critical value, now called the Chandrasekhar mass.

I have left the discussion of the syntactic level features of popularizations till this stage because the surface level probes resulted in an intriguing and confusing picture, on the basis of which no clear-cut trends are immediately discernible.

In his article on "Spoken and written dimensions in English" Biber (1986:391 ff.) has shown that one of the three textual dimensions, dimension two, consists of such co-occurring features as nominalizations, prepositions, passives, specific conjuncts, it-clefts etc. According to Biber, these syntactic features are typical of "a highly abstract, nominal content and a highly learned style" (p.395). Important features with negative weight (features that are rare or less frequent) in this dimension are place adverbs, time adverbs, relative pronoun deletion, third person pronouns etc.

Dimension one, on the other hand, consists of such co-occurring features as yes-no-questions, that-clauses, final prepositions, the pro-verb do, contractions, I/you, general hedges etc. The negative cluster of features in this dimension includes high word length and type/token ratio. If we then study how various text
types are located along these dimensions we notice that spoken language forms receive high scores along the first dimension, spontaneous and planned speeches, broadcasts lie in the middle, whereas official documents, academic prose and press reports, etc., receive very low scores for this dimension. Academic prose and official documents stand very high on the second dimension, press reports closer to the middle, and impromptu speech (telephone conversations) right at the bottom (Biber 1986:398-399).

If we bear these trends in mind when we look at the two popularizations and at a few syntactic variables in them, our diffuse picture becomes somewhat clearer.

The syntactic variables that I studied systematically were sentence length, the use of finite and non-finite clauses, finite relative clauses, direct questions, the use of passives and noun phrase length and modification. All these features are known to be style markers in sci/tec writing, in one way or the other, on the basis of previous literature. The obvious question, then, was whether changes in the use of these syntactic markers reflected a simplification of style in the popularizations as compared to the original specialized texts.

It must be emphasized that all comments refer to the overall use of the syntactic features and not to instances of one-to-one comparisons, whether a certain passive construction has become an active construction in the popularization, for instance. Another word of caution is also in place here. It has been emphasized in several studies that mechanical syntactic simplification is no "real simplification." It does not increase the readability of professional texts (cf. Ulijn & Strother, forthcoming) and may actually be dangerous for the rhetorical build-up of texts and distort the logic (Wood 1982:124-126). These observations do not, however, prevent the study of what might be called "natural syntactic simplification" that originates at the text strategic level.

Sentence length was not a criterion used by Biber because he compared spoken and written language characteristics in his article and in such a case sentence length does not qualify as a valid criterion. Sentence length is, however, a well-known style marker within written language text types, the great dividing line being between informative and imaginative prose, or nonfiction and fiction (cf. e.g. Marckworth & Bell 1967:375)

The hypothetical average sentence lengths for the text extracts studied should be around 20 to 30 plus words, i.e., if the texts were typical representatives of scientific and informative prose.

This was also true, but there was a great deal of variation. The average sentence lengths in the original texts were 30.9 and 35.5 words (the supernova set is always mentioned first and the computation set second) and 20.9 and 22.4 words in the respective popularizations. It could thus be claimed that simplification had taken place in terms of sentence length. We can also speculate
that this could at least partly be due to the editorial policies of Scientific American, which could also be behind the homogenization that could be observed in the popularizations. The popularizations were more similar in terms of sentence length than the originals.

Simplification tendencies could also be discerned in the number of finite clauses per sentence. The originals showed more complicated overall sentence patterns than the popularizations. The average number of finite clauses per sentence in the originals was 2.6 and 2.4; in the popularizations it was 1.9 and 2.2. Here the simplification patterns varied slightly. In the supernova set syntactic simplification had affected particularly the number of finite clauses, whereas in the computation set the number of non-finite clauses was clearly reduced in the popularization. The figures were 0.4 and 0.9 non-finite clauses per sentence in the original and 0.3 and 0.4 in the popularizations. In other words, the original computation text had almost one non-finite clause per sentence; this had been felt to be too much for the popularization.

Finite relative clauses are obviously felt to be more explicit than non-finite postmodifiers or premodifiers. Their share was also clearly higher in the popularizations.

Direct rhetorical questions are helpful, too. Their use is not a necessity, but they help to summarize the information given and they also act as some kind of in-text mnemonics and explications. This means was fully exploited in the computation popularization, e.g.

Is the classical thermodynamical analysis in conflict with quantum theory?...
Does this mean there are no physical limitations to computing? Far from it...
For example, do elementary logic operations require some minimum amount of time? What is the smallest possible gadgetry that could accomplish such operations?...
At the other extreme, how large can we make a computer memory? How many particles in the universe can we bring and keep together for that purpose?... (all on the same page, p. 45)

Passives, the well-known trademark of sci/tec writing, showed contradictory trends in these extracts. The average rates in the originals were 18% and 29% and in the popularizations 21% and 14%. Thus, in the supernova popularization the passive rate was actually higher than in the original. It was also relatively low in the original. In the computation set, on the other hand, reducing the passive rate was used as a simplification method.

A few comments might be appropriate before going to the last criterion applied, variation within the noun phrase types. I have discussed the syntactic differences between the two sets as if they were deliberate strategies and conscious processes. They may be to a certain extent, particularly if the publication for the
popularizations has consistent policies, as seems to be the case with Scientific American. Such policies can, however only play a minor part in the outcome. It would seem more likely that the syntactic choices predominantly reflect, at the surface level, the deeper level textual strategies chosen by the authors.

Another factor to be borne in mind at this stage is that the figures given are based on a relatively small corpus and as such have only limited validity. They are, however, in my opinion relevant in the sense that they provide a background for hypotheses that can be tested with larger corpora and also compared with larger-scale studies of the type Douglas Biber has conducted.

At the constituent level, longish, structurally complex noun phrases and nominalizations are another well-known style marker of sci/tec writing. Therefore an obvious hypothesis is that even they would be simpler, in a number of ways, in the popularizations. In this context, the NPs were only studied in terms of length and overall modification patterns. (For more thorough analyses, see e.g. Aarts 1971 and Varantola 1984.)

The hypothesis proved correct for NP lengths. The average lengths in the originals were 4.93 and 5.73; in the popularizations 4.03 and 4.54. Another hunch was that noun premodification of the type N + head noun would be more common in the originals because of its inherent compactness, nonexplicitness and the greater demands it makes on readers' extralinguistic knowledge. This hypothesis was not confirmed by the data. The N+N type was least common in the original supernova text, about 17% of NP types, and about equally common in the remaining three texts, close to 30%. The opposite was true of postmodification. It was most frequently used in the supernova original, in over 30% of the NPs, and again at about the same rate in the rest of the texts, in around 20% of cases. A clear difference was, however, discernible in the use of simple vs. modified NPs. Simple NPs were about 10% more frequent in the popularizations.

The use of passives and NP types, together, produce the confusing and diffuse syntactic picture mentioned earlier. Some results are close to expectations, some clearly contrary to them. The small size of the corpus can naturally be blamed for some of the discrepancies but I think that there is more to it than mere corpus size.

When a large scale corpus is analyzed we obtain reliable average values for the occurrence of various phenomena in the whole corpus and its various text types or genres but relatively little information about the tolerance levels and limits within a text type or genre. Another thing we know very little about, except perhaps intuitively, is reader reaction at either end of the allowed variation - scale.

There have been strong and sometimes unwarranted reactions against gobbledygook, against special jargon, too complicated legalese, officialese etc. There are also plain language movements
in various countries that seriously aim at greater readability of various types of official writing to the public. Also more balanced accounts have recently appeared about good and bad special language use (cf. English Today 5.7-9, 1986, particularly Aitchison; and also Varantola 1984 and 1986).

The following hypotheses arise from the above considerations. For certain style markers, known both for their frequency and lack of it in certain LSP text types, there may be tolerance levels with relatively clear upper and lower limits. For example, if in a popularization the rate of passives exceeds the upper limit and the noun phrase structures or overall sentence patterns become too complicated, readers will react, criticizing the text as a bad popularization. Exceeding the upper limits is, however, not the result of a conscious choice, but a reflection of wrong text strategies at a deeper level.

The syntactic realizations are naturally not the only reason for communicative failures. Wrong lexical/terminological choices that reflect incorrect presuppositions on the writer's part, a messy textual build-up or an excessively high information content can play a part, too. These hypotheses only suggest that, syntactically, misjudgment of textual strategies can be reflected as an exceeding of tolerance level limits.

Similarly, if the rate of passives becomes too low or the noun phrases or sentence patterns too simple in a special text, these choices also reflect a misjudgment of text strategy. The expert reader will react against the simplicity of the text and think it is not intended for an expert but for a lay reader. Violation of the lower limits would thus make the expert feel that his expertise is being underestimated.

If we can postulate reasonably reliable levels for various syntactic style markers, then we can also benefit from this knowledge at a concrete level, such as giving special field translators, LSP students or popularizers applicable quantitative information of special language characteristics. We can explain what kind of textual strategies lie behind the actual frequencies of syntactic categories and whether these strategies appear acceptable or not.

It is clear that no accurate upper and lower limits can be given or even suggested at this stage. The relevant syntactic features are certainly interdependent (cf. Biber 1985; 1986); their acceptable occurrence levels depend, in addition to text type, on the special field in question, the conventions of a particular (scientific) community or culture. Different languages certainly also reflect existing different traditions.

Internationalisms.

We can look at this last issue from a different angle, too, and ask optimistically if the popularization strategies could be similar in different languages and would therefore also result in
similar surface structures irrespective of language but within language-specific means. In other words, would it be possible to set a hypothesis of the above kind also for other languages than English? Would the principle of "fitness of manner to matter", (Wells, 1960: 215), manifest itself in special language texts written, say in Swedish), manifest itself in special language texts written, say, in Swedish or Finnish?

It has been suggested before that special language phenomena become more "international" at the higher levels of syntax and semantics, because the way of reasoning, logical consistency, the aim at objectivity, exactness, compactness etc. call for similar syntactic means (cf. e.g. Hoffmann 1976:372 ff., Ulijn 1979:153 and Karcsay 1977). At the same time, however, it has been emphasized that we should be very careful not to overgeneralize. Similarities may be found that help translators in their work (Karcsay 1977, Finlay 1973:23), particularly in areas that are internationally homogeneous such as science and technology. On the other hand, similarities can be very misleading within, for example the social sciences. It has also been claimed that academic traditions, in say German or English-speaking countries, are so different that they result in different academic writing styles.

I made a few, more or less random, checks with a couple of Swedish and Finnish texts, both specialist and popular texts from the fields of medicine and law to see whether any of their linguistic characteristics showed the same tendencies that were present in my English material. The texts were not interdependent in the same way as the English texts but the topics were the same in both the specialist and the popular texts.

At the lexical level the tendencies were much the same as those observed above. The terminological variables were modified according to the background knowledge of the readers. At the syntactic levels the tentative results showed equally confusing trends as in the English texts. The proportions of style markers did not change systematically (e.g. sentence length, NP characteristics) but seemed again to move within tolerance limits. In some cases sentence lengths were shorter in the popular texts, in others again, other means were used to make the topic more palatable for a general audience. In one legal text (in Finnish), for example, the sentences were actually shorter, on average, than in the popular text which also dealt with the making of contracts. The average sentence lengths conformed, however, closely to the averages available for Finnish informative/non-fiction writing (cf. Hakulinen et. al. 1980).

It is obvious that much more extensive and carefully conducted studies are needed for cross-language comparisons than this superficial and preliminary attempt. It would, however, seem beneficial even in cross-language approaches to start from the textual level and text strategies to see what effects they have at the more concrete linguistic levels. It is only after this stage that further contrastive conclusions are possible and can be used for pragmatic purposes, e.g. in translation training or LSP teaching.
Text functional aspects.

Text strategic decisions come first in the hierarchy of actual text planning but they in their turn are based on more abstract text functional considerations of the communicative purpose of a certain text.

The function of sci/tec writing in its purest form is perhaps best described as informative, objective, factual and non-committal. What happens, then, if in an actual text other functions, such as normative, directive, regulatory, persuasive or didactic functions are superimposed on the basic functions? Secondly, what type of prototypical style markers are most resistant to changes in text function? Does nominal style, for example, give way to a more verbal style in popularizations and are passives or terms avoided in less-technical contexts?

This starting point seems to deviate somewhat from the above text strategic approach, but the difference is not very great. Text strategies were considered in connection with concrete popularizations; here we start from a higher level of abstraction. Text functional decisions are in their turn concretized as strategies for actual texts such as popularizations, instructions or manuals etc.

In an earlier study (Hiltunen & Varantola 1987), I studied the differences between "pure" engineering texts and corresponding patent texts and pointed out that "the patent texts include definitions and detailed explanations that would be out of place in a normal professional engineering context". In other words, the legal function of the patent text took over also textually and the basically informative function of the technical content was overruled by the restrictions the legal function imposed on the text. It is probably for this reason that patent texts have been considered to represent a particular special area, that of legal engineering (cf. Langevin 1965).

Material.

In order to study text functional variation from a broader angle I chose a few texts that may be said to deal with the same overall topic - transport, more specifically, aviation, except for one text that deals with rail transport.

Two of the texts deal with specific passenger aircraft. The first is taken from a professional aviation journal, Flight, January 13, 1972, the other from a general periodical, Newsweek, May 27, 1985. The third text, about aircraft in general, comes from an encyclopaedia, Science and Technology Illustrated. The fourth text comes from question time in the House of Lords, Hansard, November 28, 1984. The fifth text is a statutory text, The Rules of the Air and Air Traffic Control Regulations 1981. The sixth text, Railway Accident, does not deal with aviation but with an official inquiry. It is a Department of Transport report on a derail-
ment that occurred in 1982. The report was published in 1984. I shall refer to the texts as the professional, general, encyclopaedia, question, statute and report texts.

Analysis of text functions.

The professional text is the most prototypical technical text of the set and shows a high rate of passives, nominalizations and complex noun phrases that are often very specialized terms. Yet, it is not merely a technical description but has other functions as well. It has been written as a personal account of a test flight. This means that the report is chronologically ordered, from the beginning of the flight till the end. It contains non-technical information about place, weather conditions and attitudes, e.g.

The inevitable Los Angeles smog blurred the outlines of the city and softened the shoulder of the mountains as the Beech Queen Air climbed out of Burbank and carried me over the hills to Palmdale, the Lockheed establishment on the edge of the Mojave desert. This was the business end of the trip, and the prospect of flying the shiny new TriStar made the long trip out worthwhile.

These general passages are, however, few in number and as soon as technical matters are introduced, the language becomes highly specialized and full of sci/tec style markers, e.g.

For the take-off, the weight was 353,000 lb, which included 15 flight-test crew, just over 100,000 lb of fuel and 4,900 lb of water ballast. This weight gave a c.g. of 25 per cent SMC, a V1/Vr speed of 140 kt with a 10° flap setting, and a V2 of 150 kt, in the prevailing temperature of 15° C with no wind. Just prior to take-off, the Go around/take-off switch was depressed momentarily, and Take-off appeared on the AFCS mode-annunciating indicators.

We can summarize by saying that the terminology is very demanding. No overt definitions are given. It is taken for granted that the readers come from a peer group. Nevertheless, the personal approach results in a number of personal pronouns, time and place specifications, attitudinal and colloquial vocabulary that would be atypical for a prototext, e.g.

By this time the traffic pattern at Palmdale was getting rather crowded with military traffic, but I decided to try an automatic approach and landing...

One felt very comfortable flying around the circuit because...

...and initially I made the mistake of following the flight director, not realising... For me, this highlighted two things.
I have flown many aircraft during their precertification stages but I have never met one so polished and accurate at such an early stage in the development.

The only criticisms I would make are that for my money the aileron loads are a little on the light side, the nosewheel steering too heavy,...

The average sentence length was normal for this type of informative writing. The second description of a passenger plane, the Newsweek report, is much less detailed. It keeps firmly to the general superficial level of knowledge and the commercial aspect is considered more central:

Each 747 model is designed to meet specific operational requirements: the 747SP has a shorter fuselage and carries fewer passengers but offers more range, almost 20 & better fuel economy and shorter take-offs.

The GE CF6-80C2 - an advanced-technology derivative of the CF6-80 - claims increased fuel economy with noise and weight reductions.

However, this text is a more typical prototype text than the previous one. There are many typical sci/tec style markers present. The average sentence length is normal, passives abound, the approach is impersonal, objective and the noun phrases heavy. It is the quality of the noun phrases that makes the text general. Their interpretation does not require expert knowledge because they remain at a relatively low term depth level, e.g. an advanced-technology derivative. The term use is also looser and less controlled in the sense that terms are used with more indeterminate general-language style meanings (Cf. de Beaugrande 1987 and forthcoming).

The encyclopaedia text makes an adult, possible a younger reader too, feel that the text is intended for youngsters or "simpler" readers. This is not the official aim. The preface only claims that the encyclopaedia has a novel approach. What, then, causes this feeling of naiveté of a specific target age group? The style is informative and the average sentence length within normal limits but the style is also rather too eloquent at times. The text is a feature article on aircraft but it starts very slowly and generally. The fillers are too long so that the adult reader loses patience, if his intention is only to obtain factual information about aircraft:

The airplane is a tool of transportation, and seldom in the long history of transportation has anything held such potential. The airplane promised speed. It promised convenience. It promised access to distant places. It promised to cancel out mountains, desert, oceans.
To say that the airplane changed the world is probably an understatement; it might be more truthful to say that it created a different world... It has gone far beyond the age-old dream of soaring with the birds; it faces us now with an inescapable reality all its own.

But there are discrepancies. Some portions of the text more closely resemble a factual encyclopaedia style and in reality demand more background knowledge than the fillers lead one to expect, e.g.

Top: From left to right, the position of rudder and aileron control surfaces during the four basic aircraft maneuvers: left bank, right bank, climb, and dive.

... the retractable landing gear and the variable-pitch propeller. Removing the heavy wheels and struts from the slipstream and fairing them into the wings and fuselage significantly increased an airplane's cruising speed and range. The variable-pitch propeller was to the airplane what a gear shift is to a car; it provides different gear ratios for different jobs.

The approach is quite mixed. The technical information is not backed up by enough in-text clarification if the text is intended for young readers who know as little about aviation as the overlong colourful fillers seem to suggest. It may be that in this case a more comprehensive linguistic analysis would reveal that certain critical linguistic limits have not been observed and that the text does not fulfill its communicative purpose properly. For example, there might be too many unjustifiable fiction-style characteristics in a non-fiction context which will not satisfy reader expectations.

The fifth text comes from a totally different context, an oral context where political decisions are made and where the function of the text is to further a cause and influence the decisions of the other participants. In this case the issue is the purchase of a new basic trainer aircraft for the Royal Airforce. The Lords do not pretend to know anything about the technical aspects of this particular plane. It is national interests and British industry that they are lobbying for. Therefore, although the topic is a particular aircraft, it is only referred to in the vaguest of terms as

... RAF should have an aircraft meeting its specifications

... that 70 per cent of the airframe will be built on British territory ... with the exception of the Canadian engine.

... first, the aircraft had to obtain cost effectiveness; secondly, that the aircraft had to have a good performance; and thirdly, that it had to have a long fatigue life.

... that the RAF should have the best aircraft for the purpose...
For the purposes of this Rule, "special VFR flight" means a flight made in Instrument Meteorological Conditions or at night in a control zone or in a control zone notified for the purposes of Rule 21 of these Rules in respect of which the appropriate air traffic control unit has given permission for the flight to be made in accordance with special instructions given by that unit instead of in accordance with the Instrument Flight Rules.

In order to comply with the Instrument Flight Rules, before an aircraft either takes off from a point within any controlled airspace or otherwise flies within any controlled airspace the commander of the aircraft shall cause...

We can see that at the lexical level, the term use is painstakingly careful, in the sense that no gaps are allowed. Abbreviations, acronyms and other short cuts are notably absent. All prevailing conditions are clearly stated and all terms overtly defined. Inferred relations that would be typical in a proper technical context are not permitted.

It is clear that the way a pilot knows the above rules is much more pragmatic; the conceptualization of the rules is certainly quite different from the official statute style and terminology. The last text, an official report of a derailment accident, is both a technical account of what happened and a chronological narration of the phases of the inquiry process. It therefore contains exact time and place references in a fashion similar to that of the first text. The only difference is that the time and place references in this official report are more frequent and accurate and are given in terms of minutes and miles. Cause and-effect relations are also stated in a very precise fashion, more like legal style than technical, e.g.

Linsdale Tunnel lies on the West Coast Main Line (WCML) approximately 1/2 mile north of Leighton Buzzard and is 262m long.

Although this train should have departed from Euston station at 22.55 on Wednesday 8th December its departure was delayed because of defective heating on two of the vehicles, which were eventually detached from the train. The train departed at 01.19 on Thursday 9th December.

The technical descriptions, however, are prototypical:

Track-circuit Block regulations apply to all the lines. The lines are electrified on the 25kv AC overhead system. Standard 4-aspect colour-light signals are provided, controlled from Bletchley Signal Box.

The locomotive No 81016 weighs 79 tonnes, has a Bo-Bo wheel arrangement and a maximum permissible speed of 100 mile/h.
This text has very few, if any, characteristics of sci/tec language. The few technical references are extremely general and the rest of the text is typical general political language which in this context does not even attempt to penetrate the issue at a deeper technical level.

I should point out in this context that I have taken no stand in this study as to whether a text is a special text or not. If we had followed the very strict definition of Sager et al. (1980:69), only the first text would have qualified so far because it represents peer group writing at specialist level. Here the issue is, however, not special language as such but the ways in which special subject matter is dealt with in various social contexts. We move on to the results of parliamentary work, to a statutory text that deals with air traffic regulations.

The text, even though it deals with technical subject matter, is clearly a special legal text. Again, as with patent texts, legal textual conventions dominate and the technical style is modified accordingly.

This appears for example in the stringent definition of the concepts used. The definitions are verbal and, if necessary, also supported by detailed drawings in order to leave no room for misinterpretations, e.g.

"Apron" means the part of an aerodrome provided for the stationing of aircraft for the embarkation and disembarkation of passengers, the loading and unloading of cargo and for parking.

"VFR Flight" means a flight conducted in accordance with the Visual Flight Rules in Section V of these Rules.

In this text, the overall sentence patterns and average sentence length give away the legal character of the text. Sentence patterns are complicated, with deep embeddings and leftbranching and nesting subordination. The average sentence length is clearly over the 20-30 plus words per sentence limit which seems to form the tolerance level for sci/tec writing, e.g.

An airship, while moored within the United Kingdom by night, shall display the following lights-

(a) when moored to a mooring mast, at or near the rear a white light of at least 5 candela showing in all directions;

(b) when moored other than to a mooring mast-

(i) a white light of at least 5 candela showing through angles of 110 from dead ahead to each side in the horizontal plane;

...
The most atypical feature from the point of view of a prototypical technical text is the high proportion of indirect speech in the report. Its presence is due to the reporting of the evidence given during the inquiry process, e.g.

Davies said that as he went back through the train to collect the equipment he needed to protect the train he checked that the passengers had not been injured and found that just one or two had been shaken up.

At the linguistic level this results in above-average frequency of pronouns, proper names and general language. Because the text is an official report it also fulfils the conventional conditions, such as the usual polite phrases at the beginning, a careful summary with evaluation of the findings and recommendations for future measures at the end.

The average sentence length, however, remains within the normal limits for informative writing, i.e., in this sense the text has not become a legal text, stylistically speaking.

Summary.

It has been suggested that special language style is often very formulaic and even easier to learn for a foreign subject specialist than the foreign language general styles. In other words, it is easier to learn the conventions of purer text types than the conventions of types with more mixed functions. Mixed types require a much higher standard of language skills, a much more profound knowledge of stylistic variation and requirements. The tentative findings in this analysis seem to support this view. The sci/tec style markers, impersonal and heavily nominalized style, are at their strongest and most formulaic in the prototypical categories. A shift in the textual function then superimposes other text strategies and requires the writer, for example, to make his term use more general and less explicit to suit a popularization. He will also have to add redundancy to the text and define the interrelationships between the concepts explicitly.

The writer will have to reassess the level of involvement in the new readership and adjust the information content and density accordingly. If the text is an interface text, e.g., a semi-legal style text, legal style conventions will have to be brought in. If the text has statutory power the legal conventions will take over and technical style markers will play a minor role.

The prototype feature that seems to be strongest is, as might be expected, the nominal style that reflects the way the subject matter of science and technology is conceptualized, as small separable entities that together form the whole process or piece of equipment that is the topic of the discourse. An impersonal approach is not such a strong feature and is easily overridden if the text type requires a personal and evaluative strategy. This does not necessarily make the text less special, as we may observe from the first text.
Average sentence length and overall sentence patterns are good indicators of the genre. If the limits are exceeded at either end of the scale, we may presume that the text has either become a fiction-style text or a legal-style text unless it is to be accounted a failure within the genre of informative, non-fiction writing.

All these considerations are but tentative hypotheses so far. There is no large scale quantitative analysis to support them. It would, however, appear possible to check these ideas by studying available large computer corpora from this text-hierarchical angle; to start with a text-functional analysis, relate functional observations to text-strategic decisions and through them to linguistic surface-level realizations. The top-down analysis could then be turned around. If interdependencies between the levels could be established, we should then know that certain surface level features are systematically related to certain higher level choices.

If information of this kind were available it would give us new intuitions about textual parameters, reliable and operational criteria and interlingual similarities and differences.

In LSP contexts, knowledge of this kind would be very helpful for all kinds of text production, for different SL text types, translations, modifications, etc. It could help avoid breaks in communication by showing how the text should be pitched, at the various stages, for a particular communicative purpose. It could also show that no change in pitch affects one level only. We cannot make a text more palatable by modifying vocabulary, syntax or textual parameters alone, but by paying attention to all hierarchically-ordered textual variables, from abstract text-functional decisions to surface-level choices.

References

DUBOIS, B.L. Forthcoming. "Something on the order of around 40 to
44 minutes. Imprecise numerical expressions in biomedical
slide talks".


HAKULINEN, A., KARLSSON, F. and VILKUNA, M. 1980. Suomen teksti-
lauseiden piirteitä: kvantitatiivinen tutkimus. Department
of General Linguistics. University of Helsinki. Publication
No. 6. Helsinki.

lish and the English of Technology. Two fields of LSP in
contact". Neophilologica Fennica. Helsinki. 23-35.

HOFFMANN, L. 1976. Kommunikationsmittel Fachsprache. Eine Ein-
führung. Berlin.

KARCSAY, S. 1977. "Theoretical and methodological differences of
sociological and technical-scientific Translating". Babel
23:3. 116-119.

KALLGREN, G. 1979. Innehåll i text. En genomgång av faktorer av
betydelse för texters innehåll, uppfyggna och sammanhang.
Lund.

LANGEVIN, R.A. 1965. "Communication between the legal and scien-
tific professions". In L.E. Allen and M.E. Caldwell (Eds.).
Communication Sciences and Law: Reflections from the Jurime-
trics Conference: 39-44.

MARCKWORTH, M.L. and Bell, L.M. 1967. "Sentence-length distribu-
tion in the corpus". In Kucera, H. and Francis, W.N. Compu-
tational analysis of present-day English. Providence. Rhode
Island.

SAGER, J.C., DUNGWORTH, D. and MCDONALD, P.M. 1980. English
special languages. Principles and practice in science and
technology. Wiesbaden.

ULIJN, J. 1979. "Le Registre scientifique et technique et ses
constantes et variantes supralinguistiques". Fachsprache
1:4, 126-153.

ULIJN, J. and STROther, J. Forthcoming. "Is syntactic simplifica-
tion a real simplification of EST texts?"

VARANTOLA, K. 1984. On noun phrase structures in engineering Eng-
168.

VARANTOLA, K. 1986. "Special Language and General Language".
Trans. Publication of the School of Translation Studies.
University of Turku. 223-241.

WELLS, R. 1960. "Nominal and verbal style". In Sebeok, T.A. (Ed.)

WOOD, A.S. 1982. "An examination of the rhetorical structures of
authentic chemistry texts". Applied Linguistics 3:2. 121-
143.