Gledhill Christopher. 1999b. The phraseology of rhetoric, collocations and discourse in cancer research abstracts. In C. Barron and N. Bruce (réds.), *Knowledge and Discourse* : *Proceedings of the International Multidisciplinary Conference,* 18-21 June 1996. Hong Kong : University of Hong Kong.

The phraseology of rhetoric, collocations and discourse in cancer research abstracts

Chris J. Gledhill
St Andrews University

Abstract
1 Introduction
2 The textual nature of science
3 Towards a phraseology of cancer research
4 Popularisation
5 Conclusion: the discourse of cancer research
References

Abstract

This paper describes the development of scientific ideas in cancer research articles. The paper argues that language is a key process in the innovation as well as dissemination of scientific knowledge. The aim is to ground linguistic analysis in the specific context of a discourse community, firstly to find a basis for teaching English for Specific Purposes (ESP, Swales 1990) but also to improve the textual analysis offered by current linguistic studies of large text corpora (Sinclair 1991, Stubbs 1982, 1996). The paper is based on the working context of the Pharmaceutical Sciences Department at Aston University. Fifteen researchers in the department were shadowed and their research was observed to give a contextual backdrop to an analysis of the discourse of cancer research. In a corpus of 150 research articles, phraseology is seen to vary systematically in rhetorical sections and the concept of phraseology is postulated as the preferred way of expressing a delimited set of semantic and communicative roles. Our findings also indicate that research articles use language to create science by reformulating data as research models and by altering the established patterns of phraseology. Some samples of our wider results are developed here. Contrasting new ideas with the accepted way of saying things (the phraseology of cancer research) is seen as a key mechanism in science writing and the dissemination of science. Science should therefore not be seen as an ideology transmitted via language, but as a discourse mediated by the mechanisms of textual reformulation and innovation.

1 Introduction

This paper sets out to explore the relationship between language and science. In the first instance, I argue for the centrality of language in science. I follow this by discussing briefly some salient results from a large scale study of language in cancer research articles. This establishes a groundwork for describing the role of language as a mechanism of change in this specialist discourse. Finally, I set out the wider picture of a discourse of cancer and with implications in terms of the symbiotic relationship between empirical science, science writing and science policy.

2 The textual nature of science

Latour and Woolgar (1972) characterised scientific activity as the manufacture of written text, and postulated that written material is as much valued by the scientific community as the actual physical compounds they are manipulating in the laboratory. Similarly, Nystrand (1982) has argued that it is text that provides a basis for exchange and individual promotion in the scientific group. The philosopher of science, McKinney (1991) equally emphasises the discursive nature of science and mentiones the famous case of *polywater*, the Soviet wonder substance which inspired a new but short-lived chemical paradigm. While chemists in America took the Russian findings on faith, they started to question their own beliefs about molecules and water. As new publications were produced to deal with the phenomenon, the new science was dashed by the Soviets' admission that their findings were based on nothing more than uncleaned petri dishes. The phrase 'self-perpetuating discourse' appears to be an appropriate term to describe this kind of human activity. The property of recursion, of self-replication and change from within is perhaps a fundamental property of all discourses. On the fundamental relationship between language and science, relativist and hermeneutic philosophers (Wittgenstein 1957, Heidegger 1966, Gadamer 1976) rejected the idea that science can be described metalinguistically or in terms of truth values. Instead, scientific truth cannot be anything but 'rooted' (Heidegger's term) in its culture. Gadamer has perhaps the clearest views on rhetoric and science:

> Even Descartes, that great and passionate advocate of method and certainty, is in all his writings an author who uses the means of rhetoric in a magnificent fashion. There can be no doubt about the fundamental function of rhetoric within social life. But one may go further, in view of the ubiquity of rhetoric, to defend the primordial claims of rhetoric over modern science, remembering that all science that would wish to be of practical usefulness at all is dependent on it. (Gadamer 1976: 68)

The central role of text in the dissemination of science should not therefore be dismissed because language is the medium of scientific innovation; I claim below that language provides the key mechanisms for change in science. One piece of evidence for this is the shifting nature of much scientific communication. Just as spoken discourse is ephemeral, change in society is reflected in our shifting goals and needs, processes of information access change form or take on new roles. As new text forms appear, old forms change or disappear. Atkinson (1992) has demonstrated this in his analysis of the history of the *Edinburgh Medical Journal*. Essentially, the *EMJ* evolved from a personal, replicable explanation to an impersonal, un-replicable report. The diversification of texts is mirrored by increased specialisation in fields of research resulting in a kind of discourse evolution where one field is seen to develop or expand while others split and diversify (Sager et al. 1980:xviii). One of the results of this is an increasing array of competing types of message form, including abstracting indexes, internet bulletins, automatic search indexes, and interactive self-updating databases (Jennings 1990).

In theory, the time spent by researchers on reading articles and keeping up with their fields is bound to increase. In practice, textual format and reading techniques adapt to minimise the effects of textual inflation. Miniaturisation has had its effects on the lexicogrammar (Halliday 1985). For example, my own study on the graphic interface between language and science (Gledhill 1995b) revealed codes which have a knock-on effect on the pronominal system of reference in chemistry. Similarly, Jaime-Sisó (1993) has demonstrated that newer fields of biology use clause structures in titles which have come to replace old-style nominal titles. In the same way that electronic mail has made transactions more informal and immediate, so the range of message forms reflects varying levels of formality and consolidation of scientific knowledge. Verifiable science (known as primary science) is

presented in high prestige refereed journals, such as *Trends in Pharmacology*, while 'popularised' articles appearing in *Nature* and *The New Scientist* or the medically oriented *Lancet* represent a considerable time-lag between discovery and established knowledge. Interestingly, statistical studies of citation indexes show that tertiary research (that is accelerated research that has not been peer-reviewed) tends to have higher citation scores than primary research because primary research, albeit scientifically more stable, is already out of date (Williams 1996). Within a list of research journals, therefore, a certain dynamic hierarchy can be seen to form which has a great deal to do with prestige and established scientific doctrine. This tension coupled with the pressure to publish findings before competitors must affect the quality of research. It also means that what we know as popular science, apart from the normal distortions, can hardly be said to be cutting-edge.

One process underlying textual diversification lies in the fact that the majority of 'publication' is not public at all. Many texts are circulated to an exclusive number of specialists within an institution or between institutions and funding organisations, and this forms what is called the 'Grey Literature' (Auger 1989). As information seeps out from institutions, the dividing line between genres such as exploratory bulletins, grant proposals and polished refereed research articles becomes blurred. Another process that has to be considered is the editing and reviewing process itself. As we know, one text represents many drafts of reformulations, even before the editor has had a chance to change things. Mainstream science periodicals send articles to more specialised journals or simply reject them. This is what Swales (1990) has called the 'traffic-cop' function of the peer-review process. In his procedural study of the rejection of five scientific articles in molecular microbiology, Myers (1990) identifies key processes involved in the calibration of a research article. Myers exploits the circularity of scientific observation. The argument runs like this: only a rigorous experiment can show the true nature of the phenomenon, but the rigour of the experiment can only be judged on *whether* it reveals the true nature of the phenomenon. His argument is that since this is patently circular, consensus about scientific knowledge must be negotiated socially. One way in which this is done is by changing the level of claims (from the big claim of *originality* to a smaller one *following on from previous work*). Here decisions about where the research fits into a new discipline are made in the conceptual framework and more specifically in the terminology chosen and the terminological changes proposed by the reviewer. Myers cites one author who was obliged to write *reproductive processes* instead of *reproductive behaviour* to fit in with the new field of physiology (1990: 52). At a more rhetorical level, the separate presentation of a hypothesis and data indicates to Myers that a writer had not evolved the 'sufficient syntax' to connect the two conceptually (1990:54). Finally Myers finds that all forms of discourse share fundamental principles and all have a role to play in the social construction and renegotiation of knowledge:

> Though scientific texts come out of an unusual social structure, and thus are different in some details from texts in other discourses, they are not doing something fundamentally different from other texts... Science uses our language and despite attempts to purify it, it is still loaded with social and political implications. (1990:258)

This last comment is indicative of where I must part company with the constructivists. Certain linguistic structures are appropriated and used in unique ways, which is quite a different process from identifying features of the general language as indicative of rhetoric from the wider speech community. Myers' work is perhaps the most comprehensive work on this subject, and yet even he admits that there was no possibility of systematically identifying areas of change in syntax or lexis in his text sample. While the arguments of constructivism have been well rehearsed, in this paper I argue that the paradigm needs to pay more attention to the linguistic mechanisms that play a role in the interaction between the scientific discourse

community (as defined by Swales 1990) and the wider speech community. To simplify somewhat, the constructivist view (including that of such luminaries as Kaplan & Grabe 1992) is that science is entirely a product of social interaction, a property of the 'relevance of context' (as Weingart puts it, 1993). According to Weingart, science is a product of institutional and political conflict where language is considered in terms of metaphors and topic shifts. Weingart's valuable message is that science is a product of conflicting discourses and of self-perpetuating myths engendered by the scientists or the public authorities. But there is little discussion of the mechanisms of change in discourse, nor of the mechanisms of interaction between discourses. The linguist is left with the feeling that while the discourse community has been well characterised, the discourse itself has been neglected.

Because those linguistic mechanisms that are identified are termed 'special' or 'discourse specific', it is largely assumed that science creates its own new forms of language which escape the descriptive framework of linguistics. In their well-known work on ESP, Sager, Dungworth & McDonand (1980) describe the main function of language in science as a metafunctional tool: to express abstract thought and argumentation that cannot be set out in other scientific forms (such as graphic or tabular form). In short, the constructivist claim amounts to claiming that there is a 'special language', a language that belongs outside the sphere of reference of the general language. While the creation of new languages can be easily equated with the creation of new ideas, I do not think the 'special language' route is the best way to describe what is happening. The view that there are discrete languages which are more or less translatable is difficult to hold when language is seen to exist in symbiosis with the experimental or hard science. Here I argue that instead of its use as a metalanguage for abstract representation, it is the inherent adaptiveness of language that lends itself to the creation of new ideas. Instead of claiming scientific 'uniqueness' we should instead emphasise the paraphrasing ability of both linguistic and non-linguistic forms of scientific communication. Yet even in linguistic fields which have been traditionally close to the sciences, such as terminology, this rational Cartesian view of science is disappearing. For example, Sylvia Pavel's terminographic work on the life of phrases in fractal geometry has demonstrated that expressions borrowed from one sphere bring along their own grammar, passing on a community's world view rather like one would pass on a virus (Pavel 1990). Terms and concepts (the *foci* of traditional terminology) are no longer seen as hermetic and well defined but open and negotiable, usually within the space of the same specialist text. Below, I take this idea one step further: the grammatical level is seen to be infused with lexical patterns that are part of an unrecognised system that in turn forms part of the phraseology or 'preferred way of saying things' in cancer research.

I argue, much as Stubbs has argued in critical linguistics (Stubbs 1994, 1996) that ideology is inherent in the grammar of science. Some aspects of this were discussed above in relation to phraseology. New thought is said to be a juxtaposition of the old, a new reformulation of the given knowledge base. Successful scientific language similarly appropriates metaphors and phraseology from other competing discourses, formulates its own novel phraseology and propagates its own structures outwards. The hypothesis is that science is enacted on the basis of successful linguistic communication, and that science as a cultural activity is not embodied in any text but is instead an enactment of text.

3 Towards a phraseology of cancer research

The *Pharmaceutical Sciences Corpus* (PSC) was developed as part of my PhD thesis (Gledhill 1995b). The PSC includes 150 research articles written or used by fifteen researchers at Aston University's Pharmaceutical Sciences Department. The texts were analysed using a concordancer (see an example below): a program that enables the linguist to

see patterns across thousands of words that would not ordinarily be evident just by 'manual' searching. Having shadowed the researchers, conducting protocol analysis on the ways they use and interact with these texts, it became possible to identify key structures and expressions and to envisage their meaning within the texts as a whole. The researchers gave invaluable insights about how textual clues led them to read different sections of research articles, and how graphic chemical structure diagrams were integrated into the text.

In this type of analysis, there are serious consequences for those linguists who use register distinctions such as 'written / spoken' 'formal / informal' and so on. In terms of text typology, the cancer research article cannot be slotted into one specific register, because it turns out that, as I mentioned above, such articles are used in a number of different ways, at various levels of expertise. It is really the *use* of the texts that should be categorised, and this may involve the text's function as index (reference) or report, argumentation or explanation. This is why Swales' (1990) term *genre* is preferable in this context. A genre is defined by the discourse community that actually uses it, it is neither defined by linguists nor by the internal linguistic characteristics of the text. As we have seen, these characteristics will change over time or even within the same text. In view of the multi-faced functions of these texts, the ethnographic approach of English for Specific Purposes has led to a rhetorical perspective for each and every linguistic expression. For example, instead of seeing the 'passive' as a characteristic of Abstracts and Methods sections (where it was particularly prevalent), the passive is placed in a system of alternate expressions where its full meaning emerges in contrast with other paraphrases. Sometimes the choice of the expression is evidence of textual argumentation, at other times it reveals grammatical differences that are particular to the cancer research article genre and to the ideological stances that the scientists typically make in their writing.

For example, one telling characteristic of the 150 abstracts studied in the corpus is the typical use of defining relative clauses introduced by 'who'. A concordance of a representative sample of the word *who* is listed below:

The Phraseology of *who* in Relative Clauses

subjects	*who* receive active management
patients	*who* had received active management
% of those	*who* had taken aspirin,
subjects	*who* took part in radiation studies
patients	*who* showed positive response to the administration of AZT
those	*who* progressed slowly
patients	*who* succumbed to the cancers
patients	*who* had tumours,

As can be seen, *who* refers back to animate referents, and is typically followed by the verb *receive*. *Patients, subjects, physiological groups* are all seen to *'receive'* drugs and to '*take part in*' the scientists' tests; the most typical complements of the verbs being: *experiments, trials and studies*. In cancer research, the only discourse participants to be '*given drugs*' are mice and rats. In terms of alternate expression, therefore, patients are never *given drugs*, they *receive* them. Neither are they *seen to die* or *observed to get better*, and *tumours* are seldom *observed* in the patients. Instead patients *succumb, progress,* or *have tumours.* Just as Stubbs (1994) has pointed out in his approach to critical linguistics, this systematic phraseology reveals an ideological orientation which is deeply rooted: *patients* (even when they are objectified as subjects, or *'the control group*') are conceived of as active and willing *participants*, concomitant with the distancing of the researchers from any clinical or empirical process. None of my specialist informants were aware of this convention, but the question is:

are we ever aware of all of the ideological orientations we give away in our language? Whether or not this ideological stance is transferred out again into the wider speech community is not clear, but the tendency to express in the medical spoken register the fact that patients observed to have a disease '*present with disease Y*' would appear to be a parallel development.

The passive is one of the most stereotypical characteristics of scientific writing, and yet when scrutinized it also displays phraseological characteristics that delimit its use with certain types of lexis. In the examples below, (from the PSC corpus of 150 Methods sections) the past passive is used exclusively with research- or empirically oriented verbs, while the present passive is reserved for clinical and biochemical processes:

The Phraseology of the Passive

		(empirical process)
food intake	*was*	increased
values	*were*	indicated
a sample	*was*	indicated
cell lines	*were*	determined
glucose levels	*were*	enumerated
		(clinical process)
samples	*are*	fractionated
brains	*are*	homogenized
slices	*are*	incubated
glucose	*is*	synthesized
cells	*are*	resuspended

Exceptionally, for some reason *mice* as subjects command the past passive:

		(clinical process)
mice	*were*	bred
		fed
		injected
		starved
		necrotized
		shot [meaning: injected]

These patterns are not accidental. There is now a body of linguistic theory that sees such patterns as central to the way discourse is *construed*, or to reformulate Halliday (1985), how we build and interpret the world though discourse. This neo-Firthian view of language sees the semantics of the word as textually distributed, and syntax as intimately linked with lexical knowledge. Fillmore, Kay and Connor (1988) write of lexical knowledge in terms of:

> ...phenomena larger than words, which are like words in that they have to be learned separately as individual facts about pieces of the language, but which also have grammatical structure [and] interact in important ways with the rest of the language. (1988:504)

What role does phraseology play in cancer research? I suggest that phraseology is a key process in recontextualising and reformulating *instantial knowledge*. Instantial knowledge is a systemic term for knowledge as it is embodied and represented in the language (Halliday 1985). In the specific context of cancer research articles, such knowledge involves knowing for instance which tense to use in expressing biochemical and research processes, which valencies to adopt when in relative clauses, which verbs to use when dealing with specific classes of subjects, complements and so on. Instantial knowledge, represented in the

formulations of phraseology, can be seen as a central factor in the process of writing and reading in this specialist field. To know one's field, of course, means to know this phraseology. In this regard, Francis (1993) has argued that such knowledge is a key mechanism by which we progress from cognitive plans to linguistic form:

> As communicators we do not proceed by selecting syntactic structures and independently choosing lexis to slot into them. Instead we have concepts to convey and communicative choices to make which require central lexical items, and these choices find themselves syntactic structures in which they can be said comfortably and grammatically (1993:122).

Given this view, that meanings acquire their own wordings, we can conceive of phraseology as the set of linguistic forms motivated by rhetorical aims and which further shape the discourse. It follows that the phraseological units we have identified are formulated by previous discourse and must be acquired or learnt by the community. None of my expert informants were able to recall any editorial policy that prohibits *patients* from *being given drugs*. Clearly, any changes in phraseology introduced by the author must have consequences for formulation throughout a running text, and I have demonstrated one aspect of this in the analysis of reformulation elsewhere (Gledhill 1995b). One example of reformulation within a text shall suffice here. We can consider the development of a claim in one of my informant's texts from *Tetrahedron Letters* (henceforth TL, author: Dr John Gardiner). Not all expressions can be followed throughout a text, but, if the reader is familiar with the scientific implications of what the author is doing, the central argument can be traced to a small number of key expressions. Hence the key role of the expert informants in my study.

TL's claim is based on the creation of a new compound which may be used to fight AIDS and other diseases related to cancer. But the novelty of his claim derives from his new use of a well-known chemical process known as 'total synthesis'. The new process is referred to in the Title and Abstract as *Total synthesis of the antiviral agent d4T (from crotonaldehyde)*. Phraseologically, we can see that a central expression is 'unpacked' during the course of the text. *Synthesis* is first expressed nominally (as a noun) and then expanded using prepositional phrases (usually post-qualifiers). The central claim has to be built over the course of the article. It is expressed in the Introduction together with its material source: *synthesis .. from crotonaldehyde.* Then sentences #7 and #8 (of Gardiner's text) introduce qualifying phrases (using *from*): #7 … *syntheses from nucleoside starting materials... syntheses from non-chiral pool materials*. #8 *... syntheses of the anti-AIDS drug AZT ... from the inexpensive achiral starting material, crotonaldehyde*. In the Results section, the drug *d4T* is removed from the original formulation by qualifiers indicating its function and derivation instead, and the head noun and modifier are further pre-modified, this time with evaluative epithets (underlined): #22 … *the efficient synthesis of a range of important antiviral modified nucleosides from cheap achiral starting materials.* This is in fact the opposite process to Halliday's grammatical metaphor (Halliday and Martin 1993). Grammatical metaphor is identified as a process of nominal building from verbal expressions (such as *glass breaks –>glass breakage*). One explanation may be that we are dealing with a reporting genre rather than a textbook explanation (as in Halliday's case).

This kind of additive nominal packing is consistent with the general patterns observed in the other texts (Gledhill 1995b); however TL also displays another pattern of grammatical metaphor which is less clear-cut but fundamental for my notion of reformulation. Throughout the text, the key process of synthesis is lexically reformulated to express a different stage of the chemical reaction, with a general elaboration towards the Methods section, and general reconstruction towards the Results and Discussion sections:

```
#7 A number of syntheses...
#8 ... novel and versatile synthetic routes...
```

#9 ... epoxy alcohol can also be elaborated in six steps...

The six steps are then presented as a cycle of very specific nominalised (Martin 1991) material processes:

#11 Ring opening of the epoxy alcohol...
#13 Cyclization .. proceeded in near quantitative yield...
#14 Combination of the glycosides obtained from this reaction...
during the ring-openin
#15 Acetylation of these alcohols...
#17 Treatment of the seleno compound
#18 Deacetylation was effected...g reaction...

The process is then re-generalised, again nominally:

#20 This route provides d4T in six steps...

And in #22, the process is reformulated first as a *methodology* and then as a *strategy*, concepts with clearly a much broader semantic scope than *synthesis:*

#22 The completion of this total synthesis... establishes this methodology as a general and versatile strategy towards the efficient synthesis of a range of important antiviral nucleosides...
#23 Further work on the extension of this methodology... is under way.

Text TL provides us with a very clear pattern of lexical reformulation. A scientific innovation is presented to us as a 'given' piece of instantial knowledge: it is at first nominally packed, congruently unpacked, and at the same time it is reformulated in terms of increasing claims from *synthesis (*via *steps* and *route)* to *methodology* and finally to *strategy*. While the first three items are types of topical reformulation, the last two are anaphoric nouns refocusing the methodology as a higher order of research activity, as claims. What we have here, then, lies at the boundary between grammatical metaphor (Halliday and Martin 1993) and the cohesive use of discourse items in the structure of the text. I have observed similar processes in other articles from the PSC corpus (reported in Gledhill 1995b). There are thus two mechanisms of reformulation:

Reconceptualisation

where an anaphoric noun encapsulates previous discourse in terms of the generally accepted knowledge structure (i.e. reformulating *synthesis* as *this methodology* or *this route*).

Recontextualisation

where a previous term is compared or juxtaposed with a new term such as *establishes this methodology as a new and versatile strategy.*

To conclude this section, I hope to have demonstrated here the general phraseology of two structures in cancer research writing: the relative clause and the passive. These are perhaps extreme examples, but I have argued elsewhere that there are many more structures that display the typical phraseology of science writing (Gledhill 1995b). I have also attempted to show here that phraseology operates at a textual level. I would claim that reformulation is a key process in the construction of scientific claims in these texts. It consists of the systematic

reconceptualisation and *recontextualisation* of expressions in the gradual progression of ideas within a text.

While we have seen one way scientific knowledge is reformulated within text, it would be helpful to examine how primary science is reformulated in other contexts, and especially in the popular reporting of science.

4 Popularisation

While I was conducting my survey, the Pharmaceutical Sciences Department had a number of 'breakthroughs'. In particular, Professor Michael Tisdale's research on *cachexia* came to fruition, and was accorded coverage in the local and national press, as well as interviews on local radio and television. Tisdale's research was funded by the charity Cancer Research Campaign, and it turned out that the publicity was not generated by the media but instead an organised publicity exercise staged by the head office in Manchester. Journalists were invited to hear that Tisdale (and his co-workers) had identified a key mechanism in the process of tumour growth, a process leading to a fuller understanding of cancer. However, somewhere along the line, this became a 'cure for cancer story', one that is repeated approximately once a month, in alternation with the UK's other main research charity: the Imperial Cancer Research Fund. The publicity highlights one institution at a time, as well as one group of researchers, and to be fair it is clearly a vital psychological tool in finding funding and encouraging future research. Unfortunately, however, the 'the cure for cancer' turns out to be a catch phrase, much like the expressions *fight, struggle, battle, war ... against cancer*. As such it expresses a powerful metaphor and reveals a profound psychological desire to undo a perceived injustice.

Without a large enough corpus, it is difficult to say whether the phraseological characteristics I identified in the primary research corpus are carried into the popularised versions. In the fourteen newspaper accounts of this 'breakthrough' however, we can see certain types of reformulation from the highly specialised original (namely article in the *Journal of the National Cancer Institute, Biochemistry Journal* and the more generalist *Trends in Pharmaceutical Sciences*). One such popular account was the *Daily Telegraph*'s '*Cancer discovery by farmer scientist*', and others are listed in Appendix 2. Using Gamson and Modigliani's (1989) frame approach to changing topics in the media, I categorised the typical metathemes that occur in these 'breakthrough' articles. This reveals an interesting set of oppositions that are always present at some level in the argumentation of such newspaper articles:

Fear / Hope: The breakthrough exploits our natural fear of disease and eventual hope of treatment. The emphasis is on a cure, even when this is only one potential application.

National / Local: The breakthrough emphasises the national or local role in scientific exploration. This level corresponds entirely to the locality of the newspaper.

Guilt / Responsibility. The breakthrough relieves our guilt that we cannot directly contribute to the solution, emphasising the role of charity (not industry) and the idea that the money is safe in experts' hands. Meanwhile, they are increasing their status: by acquiring money, titles and awards.

Curiosity / Ignorance: The breakthrough assures us with the general feeling that we can get to grips with science, that it's really quite straightforward and that we're getting the nutshell here.

Hero / Individual. The breakthrough assures us that the homely individual is still capable of saving the world, that the scientist is unchanged and human. The lab is de-emphasised.
These initial findings suggest that much of the science is lost to discussion of the personalities involved (most especially the role of charities and fund-raisers and the lifestyle of the senior researchers), the national or local role of the research group (depending on the national or local scope of the reporting newspaper) and finally issues not presented in the original research such as the meaning and fear of cancer. This framework can be formalised by a move-analysis of the texts, and I found the following obligatory rhetorical structure (with steps that are optional in terms of their presence and order) (following Swales' model 1990):

Move 1 Break the news

Step 1 Formulate breakthrough as the end of a struggle
Step 2 Formulate breakthrough as hope for sufferers
Step 3 Link expert source with Step 2

Move 2 Explain the news in scientific terms

Step 1 Explain the nature of the breakthrough as
a) novel approach to old problem, or
b) discovery of a new substance
Step 2 State the effect of discovery on sufferers
Step 3 Establish the extent of discovery for sufferers
Step 4 Establish the extent of change the discovery may effect

Move 3 State the news in human terms

Step 1 Identify expert source as: a) homely or b) discoverer
Step 2 Quote the expert source

a) hoping that discovery will effect change
b) linking discovery with future funding
c) aiming for direct and indirect applications
d) stating how the discovery explains previous data

Step 3 Emphasise the role of fund-raisers / sponsors as well as expert team.

Those parts of the articles which do report the actual research findings emphasise the novel approach to what is presented as an old problem and the discovery of a new substance (even though in this case there is no new substance, and the real discovery is the characterisation of a process, namely *cachexia*). The nearest we get, in fact, to a reformulation of the original findings is what I have called the 'news' statement (Move 2, above) in most of these articles. As listed in Appendix 2, these statements are almost uniformly very short in length, although many in fact do not concern themselves with the findings as they are presented in the original research. These one-sentence 'news' reports are instead concerned with the one of the five frames set out above, most usually a combination of 'national / hero' or 'local / individual', although one or two themes were quite unexpected (such as the *Daily Mail*'s emphasis on the diet of Eskimos). Just as the research articles are heterogeneous in nature, a superficial classification of the newspaper articles ranges from the title page 'national breakthrough' or (if

a local paper) 'local breakthrough'. The breakthrough also found its way into an editorial and a couple of specialist correspondent reports (c.f. Appendix 2).

Given this contextual and rhetorical backdrop, I turn to my original enquiry, namely the phraseology of popularisation. I did not expect much phraseology to have 'infected' the popular versions. But this is in fact not the case. The result reported in *Biochemistry Journal* is originally formulated as:

> The reason for depletion of host tissues is not known, but is thought to arise from differences in metabolism in the tumour-bearing state. (Biochemistry Journal)

This is a key point in the text and clearly depends on a long build-up of methodology. Indeed, when I asked Michael Tisdale to split the article into the 'classic' Introduction / Methods / Results / Discussion sections, which was not required for this publication, this one sentence was the only one which qualified for him as 'Results'. Tisdale explained that the term *depletion* is clearly identified with cancer-related *cachexia*. Similarly, the phrase *differences in metabolism* hides the complex discussion of possible difficulties in interpreting metabolic data in the brain. And yet, at this point in the text, the results are felt by the reader and the author to come together. In the newspaper articles, the breakthrough does not have the same long lines of argumentation, since the perspective of the article is concerned with the five frames of reference I set out above. Yet, when the time comes to explain the 'science' behind the breakthrough, the research is surprisingly intact. The first example from the *Independent* suffices to show that similar processes of reformulation are taking place:

> A substance found in fish oil is to be used in the treatment of cancer, following new evidence that it can shrink solid tumours and may halt the dramatic weight loss associated with the disease. (The Independent)

I have emphasised the expressions and phraseology which are reminiscent of the original texts. The degree of compaction is high (the use of *'of'*, post-modifying reduced relative clauses), as is grammatical metaphor (one passive construction and the use of nominals: *treatment, new evidence, weight loss*) as well as a projected clause following '*evidence that'*. Here is a different formulation from *The Times*:

> British scientists have isolated a substance responsible for weight loss in cancer patients, and have shown that its action can be controlled by fish oil.

Here the term *substance* is qualified by a very typical Title-like expression straight from the original article, much as it is in the *Independent*. The process of 'empirical report' is expressed in the same way as it is expressed in RA Discssion sections, the process which involves the substance is reformulated as a nominal metaphor (*action*) and the passive expression *'action can be controlled'* is also a typical passive structure from the Discussion section where the modal *'can'* finds its only use - in expressing possible applications. This suggests that stereotypical traits of scientific writing such as nominalisation, passivisation and embedded qualification cannot be simply seen as characterisations of the scientific language or research article register, but must be seen as important processes that govern our way of writing these thoughts. What has been carried through in terms of the cancer story is supplemented by other facets of the story that are implicit in the original formulation of the research. What is interesting is that we find no evidence of 'unpacking' of Halliday's grammatical metaphor - at least in the serious accounts of the breakthrough. We do have to recognise that there is little

actual reformulation of scientific ideas in these texts, in other words the long strands or argumentation (or lexical chains: Hoey 1991, Hasan 1989). It may be that argumentation is in fact a redundant feature of the primary science texts, and that while their research can be summarised in a few phrases (as is evidenced by the increasing 'miniaturisation' of Abstracts), their expository nature provides a rehearsal of the scientific argumentation which is necessary in order to negotiate the main research claims in the wider scientific community. It would be wrong therefore to see the newspaper versions of science breakthroughs as the 'secondary' effects of primary science: they are valid stories in their own right, with their own phraseology and which belong to a different discourse. This point about discourse is an important note to end on here, since we intend to say something about cancer research beyond the artificially restrictive textual confines of the research journal and the newspaper headline.

5 Conclusion: the discourse of cancer research

We have seen that discourse communities adapt linguistic processes from the general language, make them into their own and produce a 'preferred way of saying things'. Discourse communities regulate communication within their community by peer-review of research articles. However, the relationship between discourse community and speech community is perhaps more complex than Weingart's (1993) model would suggest. Weingart considers high profile cases of what is called 'big science', where scientists are in direct contact with political authorities usually in the case of war, or where, in the case of the USSR's Lysenko, the scientist *was* the political authority (Brickman R. & Jasandoff 1982). The interplay is clearly between more groups than just scientists on the one hand and politicians on the other. Since we all belong to different discourse communities, we all master and recognise many varying genres. I suggest that the role of the media, and intermediary genres such as the charity press together with scientists' own excursions into the wider speech community are alternative areas to observe the initial spread of scientific discourse. We don't have to look at newspaper articles. But scientific discourse works in even wider circles than that. Once a scientific metaphor has been adopted, let's say; that moral purity is the equivalent of health (Bonafous's (1991) link between *sain et saint* or *heil und heilig*), it can be replicated in religious discourse, in advertising, in teaching, and back again in medical discourse. As Halliday and Martin say:

> Every text, from the discourses of technocracy and bureaucracy to the television magazine and the blurb on the back of the cereal packet, is in some way affected by the modes of meaning that evolved as the scaffolding for scientific knowledge... In other words, the language of science has become the language of literacy. (Halliday and Martin 1993:11)

Halliday and Martin argued that the influence of scientific discourse is pervasive in society, especially in advanced and higher education. Their thesis has been to alert educational authorities to this influence so that students from non-literate backgrounds can deal with technical English. The work of their colleagues (Kress et al. 1995) at the Institute of Education (London, UK) suggests that a major area of concern is primary and secondary education, where the role of explanation in the science classroom is essential in the initial building of scientific competencies such as recontextualising and taxonomising; processes which we began to explore above.

In my analysis of cancer research writing above, the reader will note that I avoided analysis of the word cancer itself. Cancer is found to be a diffuse concept rather than a distinct entity or a terminological fact. Instead of seeing scientific fact as necessarily open to rhetorical interpretation, I argue that a researcher's perspective of cancer will emphasise and use entirely different aspects of a complex concept, as well as extend the concept to other uses. My own work, and that of sociologists of science has demonstrated that scientific

activity resides just as much in the production and use of texts as it does in laboratories. In what is termed the indexical function, research articles are treated as repositories for instructions, in the researchers' own words, like 'recipes'. The indexical and referential use of the research article demonstrates that it is an integral part of laboratory activity. So similarly, the signalling of successive research claims, the taxonomising effect of grammatical metaphor and the changes in wording seen in the language of these texts also initiate action by the reader, even if this is conceptual: they augment the readers' knowledge and teach new phraseology. At any point where the reader acts upon the wording of a research article, he or she can be said to be engaged in scientific activity. Thus I argue that science, embodied in complex concepts such as cancer, is language-like: science is not transmitted to minds via language because it is *already a form of language*.

To use a common metaphor from popular science (Cavalli-Sforza & Felman 1989), if we conceive of genetics or language as a set of instructions couched in a discrete combinatorial system, the relative success of one set of instructions (such as a *gene*, or a *phrase*) can be said to be down to their function outside the code rather than to any inherent properties of the code. The instructions or the codes in which instructions are set do not constitute scientific activity: it is the enactment of the code which constitutes science. Evoking natural selection here may appear far-fetched, but the analogy I think is striking: successful scientific ideas are only replicated by attracting readers rather than being intrinsically useful. The burden of usefulness is ultimately placed on the reader, and one point of interaction between the reader and text is the use of reformulation and phraseology. The point of interaction is discourse, and this can be embodied by the phrase 'science as an emergent property of discourse'.

References

Atkinson D. 1992. "The evolution of medical research and writing from 1735 to 1985: the case of the *Edinburgh Medical Journal* in *Applied Linguistics* Vol. 13/4: 337-374.

Auger C.P. 1989. *Information Sources in Grey Literature,* London: Bowker-Saur

Baker M., Francis G. and Tognini-Bonelli E. (eds.) 1993 *Text and Technology* Amsterdam: John Benjamins

Brickman R. and Jasandoff 1982. *Chemical regulation and cancer: a cross-national study of policy and politics*. Washington: US National Technical Information Service

Cavalli-Sforza l. and Felman M. 1989. *Cultural Transmission and Evolution.* Princeton New Jersey, Princeton University Press

Gadamer H.G. 1976. "On the scope and function of hermeneutical reflection." in D.E. Linge (ed. and tr.) *Philosophical Hermeneutics,* University of California Press.

Gamson W. and Modigliani A. 1989. "Media discourse and public opinion on nuclear power: a constructionist approach". *American Journal of Sociology* 95: 1-38

Gledhill, C. 1995a "Collocation and genre analysis. The discourse function of collocation in cancer research abstracts and articles." In *Zeitschrift für Anglistik und Amerikanistik.* Vol. 1, 1-26.

Gledhill, C. 1995b. *Scientific innovation and the phraseology of rhetoric. Posture, reformulation and collocation in cancer research articles.* Unpublished PhD. Thesis, Department of Languages and European Studies, Aston University.

Francis G. 1993 "A corpus-driven approach to grammar." in Baker et al. (eds.) 1993: 137-156

Halliday M.A.K. 1985 *Introduction to Functional Grammar* London: Edward Arnold

Halliday M.A.K. and Martin J. 1993 *Writing Science: Literacy and Discursive Power* London: Falmer Press

Jaime-Sisó M. 1993. "The new role of titles in research articles." Unpublished MS, 5th International Systemic Workshop on corpus-based studies, Universidad complutense de Madrid, 26-29 July 1993

Jennings E.M. 1990 "Paperless writing revisited." in *Computers and the Humanities* Vol.24: 43-48

Kaplan E. and Grabe W. 1992 "The fiction in science writing." in Schröder (ed.) :199-217

Kress G., Martins I. McKillicuddy K and Ogborn J. 1995 *Explanation in the science classroom.* Proceedings of the mid-project consultative meeting, Feb. 1995 Institute of Education, University of London

Latour B. and Woolgar S. 1972 *Inside the Laboratory* New York: Garland Press

Martin J.R. 1991 "Nominalization in science and humanities: Distilling knowledge and scaffolding text." in E. Ventola (ed.) 1991: 307-337

McKinney M. 1991 "Experimenting on and experimenting with: Polywater and experimental realism." *British Journal of the Philosophy of Science* Vol. 42: 295-307

Myers G. 1988 "The social construction of science and the teaching of English: An example of research." in *English Language Teaching Documents: Academic writing process and product*:143-150

Myers G. 1990 *Writing Biology: Texts in the Social Construction of Scientific Knowledge* Milwaukee: University of Wisconsin Press

Myers G. 1991 "Lexical cohesion and specialized knowledge in science and popular science texts." in *Discourse Processes* Vol. 14/1: 1-26

Myers G. 1992 "Textbooks and the sociology of scientific knowledge." in *English for Specific Purposes Journal* Vol. 11: 3-17

Nystrand M. 1982 *What Writers Know. The Language, Process and Structure of Written Discourse* New York: Academic Press

Pavel S. 1993a "Neology and phraseology as terminology-in-the-making." in H.B. Sonneveld & K.L. Loening (eds.) 1993: 21-34

Pavel S. 1993b "La phraséologie en langue de spécialité. Méthodologie de consignation dans les vocabulaires terminologiques." Unpublished MS, Secrétariat d'État du Canada: Direction de la terminologie et des services linguistiques. Canada

Sager J.C. 1990. *A Practical Course in Terminology Processing* Amsterdam: John Benjamins

Sager J. C. 1991. "A theory of text production, modification, reception." in H. Schršder (ed.): 34-57

Sager J., Dungworth D. & P.F. McDonald, 1980. *English Special Languages: Principles and Practice in Science and Technology* Wiesbaden, Oscar Nadstetter Verlag

Schröder H. (ed.) 1992. "Subject-oriented Texts: Languages for Special Purposes and Text Theory." Berlin: Mouton de Gruyter

Sonneveld H.B. and Loening J (eds.) 1993. "Terminology. Applications in interdisciplinary communication." John Benjamins: Amsterdam.

Stubbs M. 1982. "Written language and society: Some particular cases and general observations." in M. Nystrand (ed.) 1992: 31-55

Stubbs M. 1994. "Grammar, text and ideology: computer-assisted methods in the linguistics of representation". in *Applied Linguistics* Vol.15/2 :201-223

Stubbs M. 1996. *Text and Corpus Analysis: Computer-assisted studies of language and culture*. Oxford: Blackwell.

Swales J. 1990. *Genre Analysis: English in Academic and Research Settings* Cambridge: Cambridge University Press.

Ventola E. (ed.) 1991. *Functional and Systemic Linguistics: Approaches and Uses*, Den Haag: Mouton de Gruyter

Weingart P. 1993. "Science abused? Challenging a legend", in *Science in Context,* Vol. 6/2: 555-567.

Williams I. 1996 "IFs and buts. Impact factors of journals may affect decisions on resource allocation". in *Chemistry in Britain*, February 1996: 31-33.

Appendix 1: Topical breakdown of the Pharmaceutical Sciences Corpus.

Oncology (Cancer Research Total=83 articles)

Topic	No. of Articles	Explanation.
Chemotherapy:	26	Chemico-toxic effects on cancer.
Carcinogenesis:	18	Processes that activate cancer.
Histopathology:	12	Metabolic effects of tumours.
Immunohistochemistry:	11	Organic resistance to tumours.
Cytogenetics:	10	Genetic characteristics of cancer.
Cancer Epidemiology:	2	Population study of carcinogenesis.
Radioimmunology:	2	Radio-toxic effects on tumours.
Histology:	1	Organic properties of tumours.
Immunology:	1	Organic resistance to tumours.

Pharmaceutical science (Medicinal Chemistry Total*=63)*

Structural chemistry:	18	Processes of chemical interaction.
Organic Chemistry:	15	Functions of organic compounds.
Toxicology:	13	Effects of drugs on metabolism.
Pharmacology:	9	Effect of drugs on disease.
Enzymology:	8	Organic compounds in the metabolism.

General Medicine (Total=4)

Epidemiology:	1	Population study of disease.
Gynaecology:	1	Population study of fertility.
Patient Care:	1	Hospital management of disease.
Virology:	1	Population study of rubella virus.

Appendix 2: Expression of a scientific breaktrhough. From the Research article (RA) to the Popular / Newspaper article (NA).

RA - Work in Progress
Journal: Trends in pharmacological sciences
Author: Michael Tisdale
Title: Newly identified factors that alter host metabolism in cancer cachexia.
Extract (statement of findings):
Progressive weight loss is a characteristic feature of malignant diseases and some studies suggest that nearly 90% of patients are affected...

RA - Classic IMRD
Journal: Journal of the National Cancer Institute:
Author: Michael Tisdale.
Title: Lipolytic factors associated with murine and human cancer cachexia.
Extract (statement of findings):
Recently, considerable attention has been directed toward the isolation and identification of the factors responsible for the complex metabolic changes associated with cancer cachexia...

RA - Experimental.
Journal: Biochemistry Journal
Author: Michael Tisdale
Title: Metabolic substrate utlization by tumour and host tissues in cancer cachexia.
Extract (statement of findings):
Biochemical changes in host tissues frequently occur in cancer patients, and depletion of host adipose tissue and muscle protein in cancer patients is an important parameter determining overall survival...

NA - National Breakthrough
Paper: Daily Telegraph:
Headline: Cancer discovery by farmer scientist.
Extract (statement of news)
Scientists are a step closer to developing an early detection test and possible treatment of four of the most common and intractable cancers.

Paper: The Independent:
Headline: Chemical in fish oil to be used to treat cancer.
Extract (statement of news):
A substance found in fish oil is to be used in the treatment of cancer, following new evidence that it can shrink solid tumours and may halt the dramatic weight loss associated with the disease.

Paper: Today
Headline: Fish oil clue brings hope of cancer cure.
Extract (statement of news):
A crucial breakthrough in the use of fish oil to treat cancer brought fresh hope to millions of sufferers last night.

Paper: The Times:
Headline: Fish oil raises hope of cancer treatment.
Extract (statement of news):
British scientists have isolated a substance responsible for weight loss in cancer pateints, and have shown that its action can be controlled by fish oil.

NA - Local Breakthrough
Paper: Birmingham Post:
Headline: Midland team may have cancer cure within year.
Extract (statement of news)
Birmingham scientists believe they are on the verge of beating cancer.

Paper: Evening Mail (Birmingham):
Headline: Cancer find 'will aid funds'.
Extract (statement of news)
A breakthrough in the hunt for a cure for cancer would boost a £1 million appeal for a new experimental treatment centre in Birmingham, fundraisers said.

Paper: Evening Telegraph (Warwickshire):
Headline: Cancer team leader's hope.
Extract (statement of news)
A Warwickshire scientist is spearheading research that could lead to a major breakthrough in the diagnosis and treatment of cancer.

Paper: Warwick Courrier:
Headline: Breakthrough in fight with cancer.
Extract (statement of news)
A Claverdon Professor may hold the key to the breakthrough in a cancer cure.

NA - Correspondent Report
Paper: The Guardian:
Headline: Fish acid may help cancer victims.
Extract (statement of news)
A substance found only in oily fish may help to fight one of the main symptoms of cancer as well as leading to new forms of treatment for some of the most resistant tumours.

Paper: Daily Mail:
Headline: Have Eskimos found the key to cancer?
Extract (statement of news)
Doctors are constantly telling us to 'eat like the Italians' - lots of olive oil, pasta, fruit and vegetables. But they may have to add 'and a little like an Eskimo'.