

Towards a Philosophy of Language Management¹

Hacia una filosofía del manejo de la lengua

David CRYSTAL
University of Wales
(Bangor, Wales, UK)

Abstract

Over the past twenty years, in a number of different domains, I have been preoccupied by the relationship between content and language. Finding myself the editor of a family of general encyclopedias in the late 1980s brought an encounter with “knowledge” which had to be integrated with my professional linguistic concerns. This has since developed to include issues in document classification, search, e-commerce, and Internet security. Other directions of integration emerged in higher education, notably the need for a synergy between linguistic and cultural studies and between language and literature. And accompanying all this has been a major change in public attitudes to language, following the reaction against institutionalized linguistic prescriptivism and the evolution of a fresh understanding of the relationship between standard and nonstandard language. The varied nature of these examples suggests the need to consider the question of integration at an appropriately general level, and it is this - on analogy with established domains such as the philosophy of science or the philosophy of religion - that the title of my paper is intended to address.

Key Words: content-based; linguist; sociolinguistic; psycholinguistic; non-standard English.

Resumen

Durante los últimos veinte años, en una variedad de campos distintos, me ha preocupado la relación que existe entre el contenido y la lengua. Al editar un grupo de enciclopedias generales hacia finales de los años ochenta llegué al encuentro del “conocimiento” que debía ser integrado a mis preocupaciones lingüísticas profesionales. Desde entonces, este conocimiento se ha desarrollado hasta incluir temas como la clasificación de documentos, la investigación, el comercio electrónico y la seguridad en internet. Otras directivas de integración han surgido de la enseñanza en instituciones de formación superior; particularmente, la necesidad de hallar una sinergia entre estudios lingüísticos y culturales y entre lengua y literatura. Todo esto ha venido acompañado de un cambio importante en las actitudes del público hacia la lengua, seguido de la reacción en contra del normativismo lingüístico institucionalizado y la evolución de una nueva comprensión de la relación entre una lengua estándar y una no estándar. La variada naturaleza de estos ejemplos sugiere la necesidad de considerar su integración en un nivel general apropiado, y es este tema, en analogía con campos establecidos como la filosofía de la ciencia y la filosofía de la religión, el que mi ponencia pretende tratar.

Palabras Claves: basado en contenidos; lingüista; sociolingüística; sicolingüística; Inglés no estándar.

¹ This paper is based on a keynote presentation at the conference Integrating Content and Language in Higher Education ([ICLHE 2006](#)), Maastricht University, Maastricht, the Netherlands, 28 June 2006.

INTRODUCTION

Integrating content and language. It is the story of my life—and especially during the last 20 years or so of it. I have learned how difficult it is to do. And also how worthwhile it is to do it. I cannot address the conference themes to do with the implementation of this aim in educational programs, for I have not worked in that context for some time. But I can certainly affirm the wisdom of integrationist thinking, as I have to deal with its imperatives every day. The aim of my paper, then, is to report on this experience, and to suggest a direction for future thinking which I hope will have some relevance for educational settings.

In the beginning, I was just a linguist. This is “just” in the sense of “uniquely” not “merely”.

The study of language is no second-rate subject. On the contrary, it is, to my mind, the queen of intellectual endeavors, because it underpins the efforts of all other subjects to express themselves. Whatever the subject-matter, if the language is not right, practitioners or investigators do not have a satisfactory encounter with the domain of knowledge they wish to explore. The buzz-word of the 1990s was “knowledge-management”. Without a satisfactory linguistics, there is only knowledge mismanagement.

Knowledge, information, conceptualization, thought—content. Whatever you call it, when you first encounter it, as a manager, you are inevitably struck by the sheer size of it. If you type a search enquiry into Google, you might get back 42 million hits. This is Web content. The totals are meant to impress; in fact, they profoundly depress. For which of us could ever do something sensible with 42 million hits? Admittedly, a great deal (some think as much as 60 per cent) is duplication—the same record repeated with little or no change on dozens of sites), and a significant number of other sites, chiefly of a pornographic nature, strain the concept of “knowledge” somewhat. But we are still left with a huge amount of content to manage—more, in fact, than the contents of all the libraries in the world.

I first engaged with this content long before Google—indeed, five years before even the World Wide Web came to be. In 1986 I was asked to edit the first edition of *The Cambridge Encyclopedia*—a general-purpose reference work consisting eventually of some 25,000 alphabetical entries on people, places, and topics from every domain of knowledge, and an extensive tabular Ready Reference section providing lists of political leaders, sports results, Nobel prizes, and so on. Most of the data was provided by a group of some 350 specialists, from scientists to sports journalists, supplemented by a small in-house team who looked after current affairs. Twenty years on, and that book has gone through four editions, and has morphed into *The Penguin Encyclopedia*, whose third edition comes out next month. In addition, a whole family of associated reference books has appeared—biographical encyclopedias and dictionaries, fact finders, concise editions, pocket books. The entire database which underlies these works is now held electronically and most of it is available online (at <http://www.findout.tv/>). It is kept up-to-date each day by the in-house editorial team. If you die in the morning, and are famous enough, your demise should be in our database by the evening.

That is what I mean by content. And when I first began to perform my role as editor I thought it would be a totally different world to the one I had previously inhabited as a linguist. I soon realized I was wrong. My linguistic background interacted with the content in all kinds of ways. And, in generalizing from my experience, it seems to me that there will always be this interaction, whatever the nature of the content and the occasion of its use.

ACCESSIBILITY AND ACCOUNTABILITY

To see this, we must first appreciate that content is created for a purpose, its originator having a particular audience (readership, customer ...) in mind. The first thing we have to do, therefore, is decide on a language level to ensure accessibility—something, of course, which is usually done in association with the institution (publisher, course leader, materials provider...) through whose auspices the content is being organized. This language level has to be operationalized, by which I mean specific linguistic conventions need to be established, and this can be quite a long list. In the case of written content, we have to make decisions about graphic design, typeface, and graphology (regional spelling, punctuation, capitalization); in the case of auditory content, the decisions are to do with regional accent, speaker gender, and prosody (intonation, loudness, speed of delivery, tone of voice); in the case of electronic content (graphic or auditory), we additionally need to take account of the strengths and limitations of the technology, such as screen size, color, and animation—particularly important if we anticipate that our content will be downloaded (to iPod, mobile phone ...). In all of these cases, we need to decide whether to use a restricted grammar and vocabulary—something that is critical with young learners, but is in principle an issue for content-users of all ages.

We must reach our audience, but we must also be true to the content. The second consideration, therefore, is to respect the linguistic identity of the domain of knowledge that is being communicated. It is now a commonplace of sociolinguistics and stylistics that varieties of language exist—that domains such as science, law, religion, journalism, advertising, and so on each operate with linguistic preferences that define their communicative identity. Accordingly, an article on a scientific topic in an encyclopedia should read in a scientific way, whatever the level of accessibility we have adopted. One on art history should reflect the stylistic norms of that subject. There is, for example, a principle of impersonal expression which underlies most scientific writing, which motivates the use of such constructions as the passive and a disinterested terminology; by contrast, there is a subjective element implicit in writing on the arts which requires a more personal syntax and a more emotive vocabulary. Entries on natural history (accounts of birds, flowers ...) involve the descriptive listing of defining features. Gazetteer entries about countries, cities, and the like warrant an elliptical style, summarizing location, population, size, status, points of interest, and so on. Each content domain has its stylistic identity which needs to be respected if the plausibility of an entry is to be maintained. An entry written in the wrong style becomes unrealistic and may attract censure or ridicule. And at the level of schoolwork or examinations, work which does not respect stylistic conventions may be penalized—for example, introducing too much “first person” into a scientific report or an essay on geography.

Content providers, then, need to bear in mind the linguistic level of the audience while reflecting the linguistic level of the content source. This has to be an axiom of any integrationist approach. Not surprisingly, the balance between accessibility and accountability is not always easy to achieve. It has to be worked at. A huge amount of my editorial time is devoted to arriving at the required integration. For example, my chemistry consultant submits an entry on solvents. I have no doubt that this is an excellent entry on solvents, but there is one problem: I cannot understand it. It is too technical. So I guess at what it means, rewrite it into intelligible English, and return it to the consultant. He then tells me that it is now excellent English but rotten chemistry. In simplifying the language I have distorted the content, unacceptably. He rewrites my rewrite, and makes it good chemistry again. But in so doing he has made the content less accessible. I rewrite his rewrite, and resubmit. He does the same. Sometimes this kind of exchange goes through half a dozen reiterations before

we are both satisfied. The entries usually get longer, in the process. Unpacking a technical term using familiar language invariably adds to length. That is why ELT dictionary entries which use a defining vocabulary (such as the *Longman Dictionary of Contemporary English*) are typically longer than those which do not.

Integrating accessibility and accountability involves a sociolinguistic or stylistic perspective. The more we know, as editors, about the language (or style) of chemistry, the more we will be able to achieve by way of a mutually satisfactory outcome. Sociolinguistic awareness enhances accountability. But we also need to adopt a psycholinguistic perspective, for this is especially relevant when the focus is on accessibility. Take the “order-of-mention” principle, much studied in relation to child language acquisition. This states that, in a sequence of clauses, the easiest sequence to process will be the one where the order of the events as reported in the clauses exactly corresponds to the order of the events as they happened in life. Thus the simplest sequence is of the type:

- A happened and then B happened and then C happened

Breaking or reversing the order of mention is possible through a variety of linguistic devices, such as temporal conjunctions and adverbials.

- Before A happened, B happened.
- A happened. Years earlier, B had happened.

Numerous experiments have shown that the sequences which break order-of-mention are more difficult to process—and in children’s language are acquired later than the basic “and ... and ... and” pattern.

We can apply this principle to the notion of textual accessibility. Faced with a text which, as part of its stylistic norms, routinely breaks order-of-mention, we can improve levels of comprehension by replacing the more difficult order by the easier order. For example, historians, perhaps surprisingly, routinely deviate from basic order-of-mention. We might expect them to slavishly follow chronological order, being historians; but they do not. Any stylistic analysis of a history book will quickly bring to light such sequences as the following (I take this example from a junior school text):

In 1666 there was a great fire in London. The year before, there had been a great plague. The fire put paid to the plague.

A surprising number of children make errors when asked the question: “Which came first?” They will say: “the fire”. “And which next?” “The plague”. Then, faced with the question, “So, how did the fire eliminate the plague”, there is confusion. It is not difficult to restore order-of-mention, so that this confusion is avoided:

In 1665 there was a great plague in London. The following year, there was a great fire. The fire put paid to the plague.

This is also a regularly used style in history books, so we are doing no disservice to the genre if we use it. It is much easier to understand. Accordingly, as a teaching solution, it has a great deal to commend it.

But it is not, of course, the whole solution. We cannot routinely replace all reverse orders-of-mention by their un-reversed counterparts. That would distort the sociolinguistic identity of historical exposition. It would also be doing students a disservice. If we always simplify the language of their texts, we reinforce the gap between the language of the classroom and the language of the real world. At some point they have to bite the linguistic bullet, and get to grips with reversed order-of-mention. It is up to the teacher and course-book writer to decide when this might be best. But it needs to be done, to avoid their students remaining in a world of cotton-wool language.

Order-of-mention is just one psycholinguistic principle. There are many others. The end-weight principle is one: restructuring sentences so that subjects are short and the bulk of the information comes towards the end of the sentence. The line-break principle is another: avoiding awkward hyphenations and syntactic divisions at the end of a line of print (especially important in early reading materials). Indeed, the analysis of most grammatical topics will suggest alternatives which affect the processing complexity of a sentence.

However, few of these topics have been given the requisite research, from this point of view, so the goal of a fully psycholinguistically informed language-teaching syllabus is still some way off.

CULTURAL PERSPECTIVES

Integrating language and content also requires a cultural perspective. This was a lesson I learned very early on in the encyclopedia world. Cultural differences affect the content of a surprising number of domains. By “culture” here I mean the totality of features arising out of where we were brought up and where we live which make our behavior and our view of the world different from those brought up and living in other places. Language is an important element in these differentiating factors, both in relation to “what we say” and “the way that we say it”.

Culture affects content in all three main domains of encyclopedic knowledge: places, people, and topics. In an entry on a country or city, for example, which features of the locality should be mentioned? An account written by an outsider differs—often in important ways—from that written by a local. The outsider tends to focus on the tourist spots; the local person more on cultural history. Things can go horribly wrong, especially if a part of the world is in dispute. Encyclopedias have been banned from India or Pakistan, depending on how their map of the sub-continent represents the disputed territory of Jammu-Kashmir. And I know of one well-known dictionary which had to be withdrawn from the shelves of Thailand because it gave undue prominence to the red-light area of Bangkok. When you are an encyclopedia editor, you soon get used to being in receipt of letters—often, letter campaigns—arguing for an alternative description of a part of the world. “Arguing” is a euphemism. Threats are not uncommon. Is it Arabian Gulf, Persian Gulf, or The Gulf? Is it Macedonia or (as a result of pressure from Greece) the Former Yugoslav Republic of Macedonia? Is it Sea of Japan or (preferred by the Koreans) East Sea? Most editors end up following UN guidelines. But that does not satisfy everyone, and the letters keep coming.

Biographical entries are even trickier. What is important in a person’s career that should be mentioned? Compare the *Cambridge Dictionary of American Biography* with, say, the *Chambers Biographical Dictionary*, and you will see far more space devoted to a person’s educational background (BA, MA, PhD ...) in the former book. Americans seem to be much more interested in your “school” than Brits are. Or again: Should an entry on a person refer to a person’s color or religion? If so, how is the fact to be expressed: “black” has different cultural values in (for example) West Africa, South Africa, the USA, and Britain, and may demand alternative phrasing (such as “African-American” in the US). Even the name of a person, or the person’s by-name (nickname), can hold cultural values. If someone is known in one country as the “slayer of Turks”, and you give prominence to that fact, do not expect your book to be reviewed well in Turkey.

You also have to keep an eye on the time. A description of one person as a “rebel” or a “guerrilla” may suffice while the person is out of power, but if times change, and the rebels win,

then you had best alter the designation to “freedom-fighter”, or some such, if you want your content to be acceptable in their country. And your description of the now-defeated politicians may have to change, again to suit the mood of the time. This applies as much to past rulers of Chile as to the present ruler of Zimbabwe, and to the way in which entries describe members of controversial political groups, such as the IRA.

Topics are no less sensitive. Whether a battle is called a victory or a defeat is often a matter of cultural opinion. Whether a battle is even worth mentioning is a matter of cultural opinion. Look at an American encyclopedia and you will see details of virtually every battle of the American Revolution; you will not see these in a British encyclopedia of corresponding size. Most entries in politics, religion, economics, the social sciences, and the arts raise similar issues. Names of peoples are especially culture-bound. We might see nothing wrong with describing the members of a certain Ethiopian group as *Galla*, until we learn that this usage is now pejorative and they prefer to go by the name of *Oromo*. Many peoples of the world have more than one name, depending on whether the name comes from the group themselves or from the people who conquered them.

Here we encounter a different aspect of the content-provider’s perception of audience: we have to take a view about the relevance of the entry to the reader’s interests and priorities. It is this consideration which would, for example, make it important to include in a general encyclopedia entries on national television presenters (usually unknown outside their own country), and why an entry on the harp in a Welsh encyclopedia would be twice as long (at least) as one in an English encyclopedia.

The link between content and language can be especially striking, when we take culture into account. Take this series of extracts from local English-language newspapers:

- **Japanese English (The Daily Yomiuri, 1993):**
 - “Wakonahana, facing one of the few rikishi smaller than himself, had little trouble with No. 6 maegashire Mainoumi, who could use none of his tricks against the technically-sound sekiwake.”
- **South African English (The Sunday Times, 1974):**
 - “It is interesting to recall that some verkrampste Nationalists, who pose now as super Afrikaners, were once bittereinder bloedsappe.” [= *It is interesting to recall that some bigoted Nationalists, who pose now as super Afrikaners, were once die-hard members of the United Party.*]
- **American English [baseball game report, 2003]:**
 - “Brown was hit in the helmet by a Jim Taylor pitch in the top of the eighth inning and was down at home plate for three minutes.”
 - “Jack Sprat’s Restaurant—a business studies case-study on a fat-free food chain.”
- **British English (various news reports, 2003):**
 - “If they can challenge Microsoft, it would be Jack the giant-killing with a vengeance.”
 - “Police in hunt for Sussex Ripper.”
 - “Without functioning lifts, they’ll be running up and down the building like Jack and Jill.”

These are not just problems relating to the spread of English around the world. Any language used as a *lingua franca* in different countries will display similar variations. So examples of this kind raise a general question: How are we to deal with them? Quite plainly, the linguistic dimension is not enough. I will never understand the world of sumo wrestling if I approach it with a dictionary and try to decode it item by item. But if I approach the subject culturally, by which I mean going to sumo contests, reading up on its history, and so on, then the linguistic

dimension will be painlessly assimilated. And the same point applies to South African politics, American baseball, and British criminal history.

The three “Jack” examples are of a rather different kind, as they assume knowledge of children’s literature, in the form of fairy-tales and nursery rhymes. Most people would make a guess at the context for the Microsoft example, but the Jack and Jill extract makes no sense unless you know about the terrible accident that befell this couple as, one day, they went up a hill to fetch a pail of water. The restaurant name, likewise, assumes you know that the individual in the nursery rhyme was someone who lived according to a strict anti-fat-eating regime. This is trickier information to handle. It is the kind of knowledge that is unselfconsciously assimilated as part of the business of growing up as a young child, so there is no natural context in which an adult would encounter these stories. Fortunately there are a few ELT dictionaries (such as the *Longman Dictionary of English Language and Culture*) which provide information about them.

Traditional language-teaching materials have always included cultural elements, of course. Simply by going on a visit to the seaside, or the zoo, or wherever, content would be given a cultural perspective. But such cultural encounters were serendipitous. a means to an end of illustrating the use of the language—a particular set of vocabulary items, or a particular kind of grammatical structure. My point is that the cultural perspective should be an end in itself—treated more systematically and centrally, and not left to be picked up randomly. How often do we encounter a systematic account of a country’s political parties, say, in a language-teaching context, or its legal system, government structure, newspapers (quality vs. tabloid, left vs. right), public holidays, or literature? I have been learning French, for example, for most of my life, and I am still very confused by French political parties. (But then, my French friends inform me, so are they!)

LITERARY PERSPECTIVES

Integrating language and content means taking on board a perspective from cultural studies. And now that literature has been mentioned. I suppose if any subject demonstrates the need for integration between language and content, it is that. It is a great shame that the two subjects do not interact more often. We now have forty years of stylistic research into the relationship between language and literature, innumerable case studies investigating the literary use of language, and several demonstrations of the value literary texts can have in language teaching. Yet in many institutions the gap between linguistic and literary studies seems as wide today as it ever was. However, not for much longer, I suspect.

The focus on sociolinguistic and stylistic variation, already alluded to, is all-encompassing, and includes literature as much as the professionally orientated content domains illustrated above. Indeed, literature is the domain par excellence for a variationist perspective, because it is mimetic of all human experience. Any subject-matter is able to turn up in a literary work, and thus any associated language. To understand the language of literature we must be prepared to encounter whatever linguistic expression the author has decided to include—and this might be science, natural history, journalism, sumo wrestling, nursery rhymes, or whatever. A great deal of the difficulty involved in interpreting James Joyce (in *Ulysses* and *Finnegan’s Wake*) is because his text is packed full of allusions to multiple content domains. The more we are aware of the stylistic conventions associated with these domains, the fuller our understanding of the literary work will be. And that is what a linguistic perspective can provide.

There is a second motivation for making such a perspective routine in literary study. Authors do not just include the conventional language of content domains; they play with it. Robert Graves once said that “a poet has to master the rules of English grammar before he attempts to bend or break them”. We can adapt this observation to the present paper, and generalize it, by saying that all literary authors have to master the linguistic rules associated with content domains in their language before they attempt to bend or break them. The point is: bend and break them they often do. A scientist character in a play or a novel might stop talking like a scientist. A news-reader might be made to speak in an inappropriate way. If bending and breaking has taken place, we need to recognize it, if we are to appreciate it. And that also is what a linguistically informed literary or theatre criticism can do.

The point has been recognized for decades, and yet literature and linguistics have stayed well apart. But something is happening today that is changing things, and forcing a rapprochement between these two fields of study. A greater social realism has emerged in literature and theatre, as authors attempt to represent a world where diversity is privileged, and where societies are becoming ethnically more complex. There are fresh opportunities for the expression of new as well as established community identities. Suhayl Saadi’s novel *Psychoraag* is a good example, with its portrayal of Asian people in Glasgow. They speak in a Glasgow-Urdu way—a variety of English that I had not encountered before I read that novel. And the whole book is written in this new variety.

Saadi is just one of many authors who are exploiting new varieties of non-standard English to express local identities. It is not a new phenomenon, as such, but until the later decades of the 20th century, manifestations of non-standard English tended to be occasional and character-motivated—as in the Scottish dialogue of the novels of Scott, the use of Yorkshire English in Emily Bronte, or the Southern US dialogue in the novels of Mark Twain. But the last half-century has changed all that. We now find whole novels written in non-standard English, such as those by Saadi, Irvine Welsh, and Roddy Doyle. Poetry likewise, as in the Caribbean writing of Benjamin Zephaniah and John Agard. And there are many examples of non-standard English being the dominant voice of radio and television plays, as well as films.

Literature is reintroducing us all to the dynamic role of the non-standard language in the linguistic life of a community. In English, since the Middle Ages, literature has come to be increasingly dominated by the standard variety, defined primarily with reference to spelling, punctuation, and grammar. Now, non-standard English is reasserting itself, chiefly through the representation of cultural diversity. We see it in such art-forms as Indian cinema, the West African novel, and writing in the “new Englishes” of the world, such as *Singlish* (Singaporean English). But it has a home-grown dimension, too, with the increasingly cosmopolitan character of English-speaking society all over the world being expressed in diverse dialects. Britain of course has no monopoly of ethnic diversity. And although no other language matches English for its global spread, the emergence of “new Frenches”, “new Germans”, and suchlike are already in evidence in the arts media, expressing new social realities and using new language in which to do it.

And something else has happened in the UK, and to varying extents in some other countries—a movement away from institutional prescriptivism. The prescriptive approach to language—insisting that one variety of a language is intrinsically superior to all others—began in the UK during the eighteenth century, and became a routine part of the school curriculum, dominating the linguistic intuitions of generations of schoolchildren, and making them feel inferior about their local accents and dialects and about informal styles of speaking and writing. This curriculum disappeared from British schools in the 1960s. A new language-awareness

curriculum was introduced in the 1990s, which lacked the prescriptive biases of earlier times. Today, all dialects of English—national and international—are seen to be equal, though some (and specifically the standard dialect) are acknowledged as being, from a social point of view, more equal than others. There is no longer a negative attitude towards non-standard varieties of English. The relationship between the standard and non-standard language is being re-evaluated.

INTERNET PERSPECTIVES

Once upon a time—before the 1990s—it would have been quite a difficult matter for “new voices”, especially those reflecting minority or low-prestige communities, to be given a public presence. A minority language or dialect had to attract the attention of editors in broadcasting and the press, and compete for space with majority languages in their standard varieties. The pressure on non-standard varieties to conform was immense. There was a logic behind this state of affairs: most people listen to, watch, or read the majority languages; and the standard variety of a language evolved in the first place in order to achieve maximum coverage as a *lingua franca*. Little chance of a minority language or a non-standard dialect achieving good coverage in such a climate.

The Internet has changed everything. It has added a third medium - the electronic medium - to the two we have worked with for centuries, spoken language and written language. It has added a cluster of new varieties to those already existing in speech and writing, notably email, chat, instant messaging, and blogging. It has added a range of new linguistic conventions to our repertoire, such as the hypertext link, pop-up advertising, cutting-and-pasting, and simultaneous conversations (in chat rooms). It is enabling language change to operate at a faster pace than ever before. And it is giving us the opportunity to encounter new varieties and languages. Today, if you have an electricity supply and a computer terminal, it is the easiest thing to give your minority language or non-standard dialect a global presence. So we are going to encounter content in an increasingly diverse range of languages and dialects. The Internet is becoming increasingly multilingual as time goes by—and we ain’t seen nothin’ yet, for such languages as Chinese are still only in the infancy of their Internet existence.

Internet content raises fresh and as yet unresolved questions of management. The sheer scale of the Internet forces us to rethink our notions of relevance—how do we impose some order on our 42 million hits? The uncontrolled nature of much of the Internet forces us to rethink our notions of sanction—can we allow people to say anything they want or must their utterances be moderated or censored in some way? The accessibility of the Internet forces us to re-examine mechanisms of monitoring—issues to do with security, fraud, terrorism, and child protection.

There is still a huge amount of naivety when we look at the way in which content providers manage their portfolios. To take an example from the field of advertising, not long ago, an Internet news website reported a story of a street stabbing in Chicago. The ads down the side of the page said “Buy your knives here”, “You can buy knives on eBay”, and the like. You can see what has happened. The primitive software has found the word knife in the news report and associated this with a cache of ads which also contain the word *knife*. The software isn’t sophisticated enough to perform a content analysis of the news story in order to work out that it is a story about a crime—and therefore, if there are any ads relevant at all, they should be to do with crime prevention. The general approach is referred to as contextual advertising.

This is in fact the area in which I have found myself working a great deal in recent years. It is not only contextual advertising, of course; the whole field of search is implicated. A similar

lack of linguistic sophistication appears in e-commerce: each time you type an enquiry into a commercial site and it fails to recognize your purchase, even though you know the site must contain the product. One site would refuse to accept such search terms as mobile phones and cell phones—only cellular phones would lead you to the relevant pages. This is hardly an efficient—and certainly not a commercially sensible—way of proceeding.

Or take an example from search engines. If you want to find out what a depression is in the domain of meteorology, and you type the word depression into Google or some other search engine, you will be lucky if you get relevant results in the first few dozen hits. Google will find millions of hits for depression, but your opening pages will all be to do with depression in the psychiatric sense. You will be flooded with advice about how to manage your condition. Scroll down a few pages and—if you have the patience - you will eventually find a site to do with climate. Again, this is not an efficient way of proceeding.

Linguistically sophisticated content analysis, with a particular focus on lexicology, is what is required to sort out such problems, and it is one of the fastest growing and most exciting areas of what I call applied Internet linguistics.

LANGUAGE MANAGEMENT

In this paper I have chosen examples of the interaction between language and content from a wide range of topics—indeed, bearing in mind the general encyclopedic perspective with which I opened, the widest possible range of topics. Each example raises its own issues, and especially raises questions about aims and methods—in short, about how these issues are to be managed.

I use the term language management to include all the principles and procedures we need to have in place in order to look after a language or languages, within and across communities. The data which informs a theory of language management come from a wide range of sources, such as language policy and planning, language teaching and learning, language pathology, psycholinguistics, multilingualism, sociolinguistics, stylistics, and lexicology. In addition to the domains already reviewed in this paper, we must also consider such topics as local interest in a language's accents and dialects, concerns about clarity of expression (as in Plain English campaigns, or debate about the accessibility of the language of science), purist anxieties about language change, and public concern over the maintenance of linguistic standards. At present, each of these topics is the focus of very different professions, journals, societies, and interests. Language and literature, as suggested above, are usually taught in totally different university departments. In schools, there is usually little connection between what goes on in the Modern Languages department and what goes on in the department which teaches children oracy and literacy in their mother-tongue.

For me, the relationship between content and language requires a comprehensive frame of reference in which the theories, methods, and findings of the various professional domains dealing with language are interrelated. It is something that we all do serendipitously. Ideas from foreign-language teaching and learning have informed mother-tongue teaching and learning, and vice versa. Ideas from mother-tongue teaching have informed work in language pathology. But there is no systematic approach, in which the principles and procedures of all those interested in the study of language, professional as well as amateur, are identified, compared, and evaluated within a single frame of reference. This is what I am thinking of when I talk about a philosophy of language management. I look with admiration at the way the philosophy of science has illuminated our understanding of scientific enquiry, or at the way the philosophy of religion has

given us insight into different religious traditions. We are all, in our different ways, language managers, and the language business, it seems to me, deserves no less attention. It is a noble aim, and I would like to think that the present conference is a step towards its achievement.

BIODATA

David CRYSTAL works from his home in Holyhead, North Wales as a writer, editor, lecturer, and broadcaster. Born in Lisburn, Northern Ireland in 1941, he spent his early years in Holyhead. His family moved to Liverpool in 1951, and he received his secondary schooling at St Mary's College. He read English at University College London (1959-62), specialized in English language studies, did some research there at the Survey of English Usage under Randolph Quirk (1962-3), then joined academic life as a lecturer in linguistics, first at Bangor, then at Reading. He published the first of his 100 or so books in 1964, and became known chiefly for his research work in English language studies, in such fields as intonation and stylistics, and in the application of linguistics to religious, educational, and clinical contexts, notably in the development of a range of linguistic profiling techniques for diagnostic and therapeutic purposes. He held a chair at the University of Reading for 10 years, and is now Honorary Professor of Linguistics at the University of Wales, Bangor. These days he divides his time between work on language and work on general reference publishing. For more information on David Crystal, please see <http://www.davidcrystal.com/>.