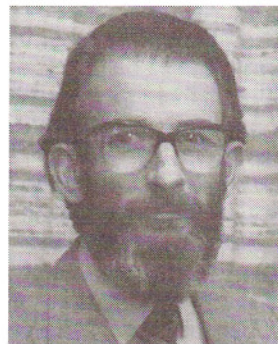

The encyclop(a)edic word game

DAVID CRYSTAL

*Tales of the unexpected
in editing a general encyclopedia*



FOR THE past three years I have been editing a general encyclopedia, and it probably shows. But it is over now: the work will be out in October of this year, a single-volume guide to life, the universe, and beyond, to be called the *Cambridge Encyclopedia*. It was born out of a close collaboration between the Edinburgh firm of W & R Chambers and Cambridge University Press, and will be published by the latter. The enterprise attracted me because it was the kind of intellectual offer one couldn't refuse: an opportunity to look systematically – albeit briefly – at, well, everything. At the very least, I thought, I would learn a lot; and it would be a change from doing linguistics.

In the event, I was right on the first count, and wrong on the second. I did learn a lot. But I found myself doing linguistics all the time.

I should be more precise: I found myself continually trying to solve problems which were more to do with language than with anything else. Now I had not expected this outcome. Like most people schooled in the British traditions of reference publishing, I had been taught to respect the dualism separating the two main genres of dictionary and encyclopedia. The former, we were given to understand, dealt with linguistic issues; the latter with 'reality', 'facts', 'knowledge'. I cannot remember a time when I did not worry about the artificiality of this distinction: there are for example, too many cases of

the *White House/Whitehall* type. (One would not expect to find names of streets or residences in a dictionary; but when people start saying things like *The White House is in confusion*, the meaning to be elucidated is indeed a lexicographical one.) And certainly, in recent years, British dictionaries have been showing signs of catching up with their American and European counterparts, and incorporating more encyclopedic information – the Reader's Digest *Great Illustrated Dictionary* being to date the clearest case in point. But in all of this, the traffic flow of information seemed to be one way – from encyclopedia to dictionary.

I was not expecting to find myself faced

DAVID CRYSTAL read English at University College London, and has since held posts in linguistics at the University College of North Wales, Bangor, and at the University of Reading, where he taught for twenty years. He works currently as a writer, lecturer, and broadcaster on language and linguistics, maintaining his academic links through an honorary professorship in linguistics at Bangor. He is the editor of 'Linguistics Abstracts' and 'Child Language Teaching and Therapy'. Among his recent publications are 'Listen to Your Child', 'Who Cares About English Usage?', 'Linguistic Encounters with Language Handicap' and 'Cambridge Encyclopedia of Language'.

with linguistic problems when dealing with general encyclopedic subject-matter – other than the obvious one, of needing to present the data in intelligible English. I confidently expected to find myself in a world of straightforward facts – of people, places, fauna, flora, and inventions, of chemical formulae and dates of birth, of historical events and mountain heights – and so I was, to a degree. But far more of my time, over the past three years, has been devoted to issues which proved to be linguistic, or semi-linguistic, in character. And although the rigours of the publishing schedule have made me unable to write my customary pieces for *ET* over the past 12 months, the issues with which *ET* routinely deals have never been far from my mind.

Not all aspects of encyclopedia work were affected. Indeed, of the two main editing tasks – decisions about coverage and decisions about treatment – the former seemed largely unaffected by linguistic considerations. How many entries should be included, given the size limitations necessarily imposed on a single-volume project? What kind of illustrations should be used, and how much space should be devoted to them? How were the entries to be distributed across the various branches of knowledge, given the subject-matter emphases that would enable this encyclopedia to make a distinctive contribution, as a reference book? Such matters involved an encounter with reality, the non-linguistic world, mediated by my team of contributors. But when it came to the question of treatment, it was a very different story.

The contrast was sometimes quite marked. I recall a visit to the Natural History Museum in London for an opening discussion with the team of contributors there. As I walked along the various corridors to the meeting room, the relevance of the real world could not have been more apparent, in the form of the models and pictures around me; but in the meeting room, all this was left behind, and we spent most of our time debating such matters as style, the discourse organisation of the entries, layout, and typography.

Cans of worms

A remarkable range of linguistic questions comes to light when working on an encyclo-

pedia, and each of them is a can of worms. For this paper, I will look briefly at five of them.

Spelling Any publishing venture needs to establish a house style for spelling variation, to ensure consistency in such matters as *judg(e)ment*, *focus(s)ed*, and the choice between *-ise* and *-ize*. In such cases, the decisions are largely arbitrary. But there are several factors which raise more complex issues, some of which require a commercial (marketing) resolution, some of which involve linguistic (often, specifically, socio-linguistic) considerations. Under the first heading, there is the major question of how to handle the difference between British and American spelling. How should this be reflected in the headwords: do we have *colour photography?* *color photography?* *colour/color photography?* *color/colour photography?* *colo(u)r photography?*. And should any of these alternatives be given within the body of the text? For any publishing enterprise, the answer to this question will be largely a commercial one, in the light of marketing considerations. Dual-dialect editions are the only means of solving this problem to everyone's satisfaction. But at the very least, problems of information retrieval must be anticipated, by having appropriate cross-references between differently spelled headwords (as with *(o)esophagus*, *f(o)etus*, *colo(u)r*, *p(a)ediatric*, and the like).

But not all spelling issues can be commercially resolved, nor is the regional factor the only one to be borne in mind. The divergent usage in cases like *p(a)ediatric* involves more than a matter of international dialect divergence. The universal provision of US texts using the *-e-* spelling means that British readers have become increasingly familiar with it, and usage has slowly increased during the present century, with younger pediatricians more likely to use it than their older paediatric colleagues. Historical and social factors are always present. The historical factor shows up clearly in the very name of the work: should it be *encyclopedia* or *encyclopaedia*? Both usages are now well-established in the UK, and as a present day encyclop(a)edist, I am happy to accept both. On the other hand, I do not feel happy about dropping the *-e-* when referring to, say, Diderot, or in talking about Ephraim Chambers' original

work. Chambers wrote an encyclopaedia, not an encyclopedia.

The social factors have far more worrying implications. There are some half a dozen ways of spelling the name Mohammed, for example, and they would be given different preference ratings by Muslims (Moslems). If the prophet is being referred to, *Muhammad* seems to rank higher than *Mohammed*, which is definitely higher than *Mahomet*, according to my informants. If it is not the prophet, other orders emerge, and these vary between localities (Afghanistan preferred spellings are different from those in North Africa, for example). Or again, *Mogul* may be a traditional form, but it carries colonial connotations which *Mughal* lacks. Or again, should it be *Quebec* or *Québec*? Or again, to take an example nearer home, should it be the *Stewart* dynasty of kings and princes (as preferred in Scotland) or the *Stuart* dynasty? Such cases raise the emotional tone of any spelling debate. They are of a different order from the 'standard' problems of the *-ise/-ize* type, and take up a great deal more time to resolve – and never to everyone's satisfaction.

Pronunciation This was an unexpected can of worms. The principle adopted was to give a guide, using a special transcription, whenever the pronunciation of the headword was not fairly obvious. There was no intention to give all the variant pronunciations, such as would be given in a dictionary. All the encyclopedia reader needs is an acceptable pronunciation, so that when talking about an entry, the version used will be recognisable. At the outset, I had not considered this to be a problem. And I was wrong. In particular, I had underestimated the difficulty of giving a pronunciation guide to place names and to people's names. In many cases, there was no information available at all. *Who's Who* manuals do not give pronunciations, and contributors often admitted to never having heard a particular person's name spoken aloud.

There were a surprisingly large number of words for which I realised I had only a visual recognition. If I had to say them, I would have had to introduce an apologetic remark (such as 'or however you say it') to my listener, who would, likewise, probably be uncertain. Having raised many of these cases with colleagues and friends, in recent mon-

ths, I am confident in saying that the extent of this problem is much underestimated. You can prove the point to yourself by looking at the words in Panel 1, taken from the beginning of the letter M. All you have to do is decide which pronunciation you would recommend, if you were editor. Note the cases where you are (a) quite definite about it, (b) totally unclear, and (c) where you are unsure, and would like to check in a dictionary first (but if you do, look in more than one dictionary, for disagreements abound).

After a few minutes of this kind of work, you lose control of your intuition, and become unsure of the pronunciations of words you have been happily using for years (are you *sure* it is [makia'veli] and not [matʃia'veli])? And there are some really thorny problems.

○ Your level of foreign language awareness will condition your views. If you have spent years learning French or Latin, then you may well wish to demonstrate this knowledge by insisting on the 'correct' pronunciation. On the other hand, you will then risk conflict with a popular pronunciation, which has anglicised the form. This can be seen in the case of *Mach*, where several sources gave *Mach* the person as [ma:x], but the concepts as [mak] or [ma:k]. One physicist, proud of his German, stressed the fricative ending in *Mach's principle*. The same point applies to the pronunciation of classical names, such as *Dionysus* and *Hades*: those who have a classical training will insist on a historical pronunciation (or, at least, one of the possibilities, depending on which system they learned).

How do you say?

1

M'Naghten rules	macaw
M'Zab Valley	Macbeth
Maastricht	Maccabees
Maazel, Lorin	MacCaig, Norman
McAdam, John	macchia
macadamia nut	MacDiarmid, Hugh
Macao	Macedonia
macaque	Macgillycuddy's Reeks
macaroni	Mach, Ernest
MacArthur, Douglas	Mach number
Macassar	Mach's principle
Macaulay, Rose	Machel, Samora
	Machiavelli

○ Standard pronunciations may vary, also, depending on the topic area. The same name may appear, for example, in classical mythology (as the name of a god) and in astronomy (as the name of a planet), but the recommended pronunciations are not always the same. Investigate Hyperion, Iapetus, Icarus, and Leda, for example.

○ Historical circumstances affect pronunciations: if a European scholar flees from Europe in the 1930s and lives the rest of his life in the USA, should the pronunciation be European or American? Should Tillich have a final fricative? Should Goepfert-Mayer have a rounded [oe] vowel? Ask the scholars themselves, you might suggest. But this does not always help, for some accept both, and some give up their preference, in the face of universal re-pronunciation.

○ When does a foreign name become

sufficiently anglicised that only the English form should be given. No problem with *Paris* ['paris], but is it *Munich* ['mju:nik] or ['mju:nix], *Badajoz* ['badadzoz] or [bada'hoθ], *Chania* ['xɑ:nia], ['kɑ:nia'], or [xan'ya]?

Add to these the problems of representing stress in words coming from languages where there is no comparable stress system, such as French, Japanese, Hindi, or Chinese (e.g. should one recommend Barrault ['barou] or [ba'rou]?) and the extent of the problems should be apparent. I have carried out no count, but I estimate that over a third of the headwords in the encyclopedia raise pronunciation problems of one kind or another. I never imagined there would be so many.

Capitalization Which words should have initial capitals? It is sometimes thought to be

President

2

... charity, founded in 1919, and having as its president the Princess Royal...

... charity, founded in 1919, and having as its President the Princess Royal...

... Indian philosopher, statesman, and president (1962...

... Indian philosopher, statesman, and President (1962...

... the country's first president (1975)...

... the country's first President (1975)...

... US Republican statesman and 40th president...

... US Republican statesman and 40th President...

... the domestic policies of US president Roosevelt...

... the domestic policies of US President Roosevelt...

... a department responsible to the president for the conduct of foreign affairs...

... a department responsible to the President for the conduct of foreign affairs...

... and his successor as president (1989)...

... and his successor as President (1989)...

... led to his being elected president of the colony...

... led to his being elected President of the colony...

... the constitution of 1987 provided for a president...

... the constitution of 1987 provided for a President...

... chairman of the Hawker Siddeley Group from 1935, and president from 1963...

... chairman of the Hawker Siddeley Group from 1935 and President from 1963...

... as president of the provisional government...

... as President of the provisional government...

... the first president of the Royal Academy (1768)...

... the first President of the Royal Academy (1768)...

... and became the only president to be re-elected three times...

... and became the only President to be re-elected three times...

... he became president of the National Union of Mineworkers...

... he became President of the National Union of Mineworkers...

obvious: proper names. Thus, there is no doubt over *Liverpool*, *Thatcher*, and *Church of England*. But linguistic studies have shown that there is no clear way to identify what is a proper name: is *Bible* or *bible*, *Moon* or *moon*? Take the latter case. Our moon is special, unique, so we may give it a capital. But will we do so when we write *full moon* or *blue moon*? And will *Toth* be a *moon-god* or a *Moon-god* (or, for that matter, a *Moon-God*)? The problem is one of gradience, from the clear-cut case where we are talking about a unique person, place or thing, to cases where we are talking about the class of entities. Thus, we have *President Kennedy*, at one extreme, and *The country is governed by a president*, at the other. But there are many intermediate cases. In Panel 2 is a list of the uses of the word *president* in one part of the *Cambridge Encyclopedia*. Which would you capitalize, and why? There is a space to tick one or the other. Pay particular attention to those cases where you make a tick without thinking twice and those where your pen hovers uncertainly over the paper. When you have made your decision, reflect on the second part of the task: ' . . . and why?' Is it possible to be totally consistent?

There are subtle constraints at work. Context seems important. Thus, *Indian President* is more acceptable than *Indian philosopher and President* (I assume that *Indian Philosopher and President* is unacceptable). The implied importance conveyed by a capital letter makes *President of the United States* more likely (in a general reference work, note) than *President of the National Union of Mineworkers* (the parenthesis is important, as in the NUM there may be a strong opinion that the order of priorities should be reversed). A provisional government presidency, being only provisional, might not merit capitals at all. What is plain is that no simple principle will work for all cases. All official titles should be capitalised? *He became Emperor of Rome*? *He became Emperor of all lands west of . . .*, *He was crowned Emperor*? *He acted as Emperor*? Or take academic titles. Dennis Gabor, for example, was a 'professor of physics', but one could not write this as *Professor of Physics*, for this was not his title: he was in actual fact *Professor of Applied Electron Physics*. To refer to his role briefly, one would have to avoid capitals altogether (unless one accepted *Professor of physics*).

Capitalization

Decisions about capitalization should not be taken lightly. The words involved are high frequency words in a general encyclopedia, and the overall visual appearance of the work is going to be radically affected by any decision, as can be seen by comparing two versions of the same entry, where the titles are first capitalized then printed in lower case.

Howe, Sir (Richard Edward) Geoffrey (1926–)

British Conservative statesman, educated at Winchester and Cambridge. He was called to the Bar in 1952 and became an MP in 1964. Knighted in 1970, he became Solicitor-General (1970–2), Minister for Trade and Consumer Affairs (1972–4), Chancellor of the Exchequer (1979–83), and Foreign Secretary (1983–9). In 1989 he was made Deputy Prime Minister, Lord President of the Council, and Leader of the House of Commons.

Howe, Sir (Richard Edward) Geoffrey (1926–)

British Conservative statesman, educated at Winchester and Cambridge. He was called to the bar in 1952 and became an MP in 1964. Knighted in 1970, he became solicitor-general (1970–2), minister for trade and consumer affairs (1972–4), chancellor of the Exchequer (1979–83), and foreign secretary (1983–9). In 1989 he was made deputy prime minister, lord president of the Council, and leader of the House of Commons.

Order of entries Even the elementary question of how to order entries raises linguistic issues. Often, in reference works, an arbitrary decision is made about which of the two main indexing conventions should be followed: letter-by-letter or word by word? Dictionaries use the former, as a general rule. What should an encyclopedia do? Contemporary practice varies. *The Macmillan Encyclopedia*, for example, follows the former principle; the *Encyclopedia Britannica* goes for the latter. An example of the alternative orderings is given in Panel 4.

There are merits and demerits in both approaches. The word-by-word approach has certain semantic advantages: entries related in meaning tend to cluster together (whereas the letter-by-letter principle splits these up in an arbitrary manner). This is both convenient (in the same way that a thesaurus puts together what belongs together) and potentially illuminating (*Mao Zedong* is adjacent to

Letter or word order

4

The difference between the two ordering approaches can be illustrated by the following sample of entries:

<i>letter-by-letter</i>	<i>word-by-word</i>
sea pen	sea pen
sea perch	sea perch
Searle, Ronald	sea robin
sea robin	sea slug
sea slug	sea spider
sea spider	sea urchin
SEATO	Searle, Ronald
Seattle	SEATO
sea urchin	Seattle

Maoism, instead of being separated from it by *Maori* and *Maori Wars*). On the other hand, the word-by-word approach is dependent on a principle which is not entirely systematic: the word-break conventions of the language. Is it *sea lion*, *sea-lion*, or *sealion*? Dictionaries provide different advice on such points. If *sealion* is taken as a single word, it will be listed well away from *sea otter* – and appear after it, in terms of strict alphabetical order, thus providing a possible problem of look-up (for readers who have not read the preface).

For my own project, I kept a note of the number of cases of the *sealion* type, and the number of cases where there was a natural sense clustering without any look-up problems. The latter far exceeded the former. I therefore opted for word-by-word ordering. At least this principle allows one to move a little way in the direction of what an encyclopedia is all about – the structuring of meaning.

The question of style

There is a noticeable tendency in reference book writing, especially that aimed at a general audience, to adopt what might be called a 'bland' style. I first noticed this several years ago, when I was asked to write something for a Readers' Digest publication. I produced a piece of work, and soon after was sent an edited version to check. This version was remarkably different. Whatever was idiosyncratic about my style had been eliminated, and in its place was a piece of simplified, slightly dramatised, clear and

compelling prose, which I had read a thousand times in waiting rooms. The RD editorial management team had done an excellent job, in turning my material into their product. The whole book, when it appeared, was written in the same style. Although there were dozens of contributors, it was not possible to discern a stylistic idiosyncrasy. Style, in this case, was not the man, but the manager.

I did not object. (It was not my place to object: I, like all the other contributors, had simply provided information, for which I had been paid. It was no longer 'my' material.) No evident disservice had been done to the subject-matter. A lot of the submitted material had been left out; and several qualifying words and phrases of an academic kind had been deleted, so that the language stated baldly what should have been said more circumspectly. There were now rather more half-truths than I would have liked to see; but there were no real errors of fact. I could see the point of such a book, and was glad that some systematic information on language matters had been included, in a work which would undoubtedly sell very well to a readership that might otherwise read little or nothing on linguistic topics.

But from a stylistic point of view, I found the whole enterprise extremely disturbing. Having spent a great deal of time analysing the properties of style (e.g. in *Investigating English Style*, 1969, written with Derek Davy), I had learned to recognise the enormous range of stylistic diversity which exists in English, and to understand a little of why it was there. Here, it was all gone. Although the book covered a wide range of subject-matter, it was no longer possible to differentiate between scientific, religious, legal, historical, geographical, business, or other kinds of style. Everything was the same. And in looking about, a similar blandness could be seen creeping into other contexts too – such as in the daily press, radio documentaries, and new translations of the Bible.

After I became editor of the *Cambridge Encyclopedia*, I looked for the first time at the style of the various general encyclopedias already available, and was struck immediately by the same bland, characterless impression. For the most part, the language was clear and functional – but at the same time it was faceless and boring. Biographies were treated

in the same way as towns and cities. There was little distinction between the different genres of topic entry. The language expressed a content, but did not convey anything of the stylistic character or identity of the genre to which the entry belonged.

It is easy to see why this kind of English prose style has developed. It derives from the application of two main criteria. First there is a concern for consistency: this involves both entry discourse conventions (e.g. ensuring that dates of birth and death are always given in the same place and in the same style) and language conventions (e.g. ensuring that capitalized or hyphenated words are given in the same way throughout). Secondly there is a demand for a basic level of clarity: this results in short entries, distinctive conventions of layout and typography, compact sentences, and the avoidance of complex syntactic constructions (such as the passive). Such considerations are crucial to the success of any reference work. An explicit set of treatment conventions acts as an aide-memoire to the editor (a check that all relevant information has been included for each entry) and facilitates information retrieval by the reader (who becomes increasingly familiar with the way the information in an entry is being presented). And to disregard the principle of consistency can lead to ambiguity: a typographic distinction may suggest a semantic distinction to the reader which the author did not intend. Also, apparent carelessness over points of detail inevitably carries with it the implication that perhaps the author has been sloppy over points of substance too. The need for clarity is presumably self-evident.

On the other hand, when it comes to a general encyclopedia, neither of these criteria is straightforward in application. One cannot be totally consistent. Different types of entry may require different types of convention, and special factors may intervene (as we have already seen in the discussion of capitalisation and pronunciation). (Indeed, the very relevance of the notion of consistency is debatable, in a book which is not being read from beginning to end, and where the majority of uses involve the looking up of a single entry.) Moreover, different levels of clarity may need to be invoked, when one is dealing with all branches of subject matter. In short single-paragraph entries, it is simply not possible to expound a concept in, say,

mathematics or physics with the same facility as one in cinema or recreation.

All of this raises the question of whether it is possible or desirable to incorporate into a general encyclopedia a principle of stylistic appropriateness for the various entries. Is it useful to build into the structure of the work an element of genre identity, so that entries from a particular domain could be seen to belong together, at a stylistic level? Their style, in other words, would be relatable to the stylistic norms encountered in other works within the same domain. I believe strongly in the desirability of this approach. I feel it adds interest, realism, and usability. To take just one example: if you were a pupil doing a chemistry project, and you wished to look something up in the encyclopedia, would you not be helped if the language of the entry bore a systematic resemblance to the style of the textbooks and other materials you had previously been exposed to? Or, putting this another way, if you had to contend with a stylistic shift, would not this be an extra barrier to intelligibility?

It is an attractive argument, but its consequences are by no means obvious in advance. How many entries would in fact be stylistically differentiated, if one were to introduce such a principle? Would it really be a noticeable feature of the work? How much distinctiveness is it possible to introduce into a short entry, in any case? One of my aims for the *Cambridge Encyclopedia* was to see how far this approach could be maintained without losing sight of the general need to have entries which generally conform to a single 'house-style' and which are accessible to a wide readership. How far it has been successful it will be for others to say. All that can be done at present is to illustrate some of the effects involved, to give the flavour of the enterprise.

A major dimension of the stylistic variation involved relates to the divergent discourse structure of the three main types of entry contained in an encyclopedia: biographical entries, gazetteer entries, and topical entries. (There are certain cases which fall uncertainly between these divisions: some archaeology entries, for example, are midway between places and topics – is Stonehenge a place or a topic? Movement originators sit uncertainly between biographies and topics – is Arianism best dealt with as part of Arius, or should it be Arius within Arianism? However, the vast

Discourse

The discourse structure of gazetteer and biographical entries is illustrated in the following two entries taken from the *Cambridge Encyclopedia* (the encyclopedia's layout and typographic conventions are not reproduced).

Niamey [neemay] 13°32N 2°05E, pop (1983) 399 100. Riverport capital of Niger; 800 km/500 ml NNW of Lagos (Nigeria); airport; railway terminus; university (1971); textiles, metals, food processing, ceramics, plastics, chemicals, pharmaceuticals; markets selling cloth, leather, iron and copper craftwork; national museum, zoo, botanical gardens. >> Niger [1]

Maclean, Alistair (1922–) British author, born in Glasgow, Scotland. He served in the Royal Navy (1941–6), and then studied at Glasgow. In 1954, while a teacher, won a short story competition, then wrote the novel *HMS Ulysses*, which became an immediate best-seller. He followed it with *The Guns of Navarone* (1957), and turned to the full-time writing of what he calls 'adventure stories', with fast-moving intricate plots and exotic settings, such as *Ice Station Zebra* (1963), *Where Eagles Dare* (1967), and *San Andreas* (1984). Most have made highly successful films. >> novel.

majority of the entries would be allocatable to one or other of these three categories without difficulty.) The discourse structure of the biographical entries (ignoring a few minor variations) is as follows (an example is found in Panel 5):

Surname, first name(s), title(s), byname(s), pronunciation (where needed), dates (birth/death), feast day (where relevant), nationality and characterization (e.g. 'British' author), where born, educational history, main positions, major contribution(s), where died (if relevant).

Entries for towns and cities, by contrast, have the following structure (see also Panel 5):

Placename in English, pronunciation (where needed), placename in local language (where relevant), latitude and longitude, population estimate, designation and location (e.g. 'River-port capital of Niger'), distance indicators (e.g. '800 km/500 ml NNW . . .'), historical background (where relevant), public transport facilities, educational facilities,

commercial status, range of economic products, buildings and other features of interest.

By contrast, topic entries, on the whole, lack any such fixed internal discourse structure. The nearest case is the natural history group (see below), where a finite set of formal and functional characteristics needs to be systematically expounded. Other cases include chemistry, where the opening part of most entries has a predictable structure (e.g. formulae, alternative nomenclature, data on boiling point or density), and many of the entries in astronomy (e.g. characteristics of constellations, planetary details) and history (data on battles and treaties). But in the majority of topic entries there is no predictability of discourse organisation. The possibilities of stylistic differentiation, accordingly, derive from specific features of linguistic structure. Here are a few brief examples of these features, followed by an extended illustration of one particularly important type.

○ An evaluative dimension is introduced into the various categories of entry in fine arts primarily in the citation of best exemplars of an artist or genre ('. . . while the rather frigid side of Neoclassicism is well exemplified by the sculpture of Canova . . .'), and an associated emphasis on the names of works and the listing of antecedents ('. . . a pupil of Giotto . . .').

○ A tightly constrained syntax is found in the mathematical entries, with a greater reliance on symbols and line drawings than in any other topic category, and a level of awareness presupposed which is also greater than in other categories (though the other 'hard' sciences – physics and chemistry – come close).

○ There is a terse style in the astronautics entries, reminiscent of the NASA dialogues between spacecraft and mission control, now a part of contemporary linguistic consciousness ('. . . A multiple outer-planet flyby mission . . .').

○ The legal entries display a carefully weighed language, with a balanced syntactic structure and judiciously selected lexicon ('. . . To succeed in negligence, the plaintiff must prove that the defendant owed him or her a duty of care . . .').

○ The anatomy and physiology entries dis-

Elliptical style

colugo [kuhloogoh] A nocturnal mammal, native to SE Asia; face lemur-like; large gliding membrane along each side of body, extending to tips of fingers, toes, and long tail; lives in trees; eats plant material; closely related to insectivores; only member of the order *Dermoptera*; also known as **flying lemur**. (Family: *Cynocephalidae*, 2 species.) >> Insectivore; lemur.

The above entry, as used in the *Cambridge Encyclopedia*, can be 'translated' into a non-elliptical style. The translation is 70 words long, instead of 51. I would argue that the extra length adds nothing. On the other hand, if this saving were lost for the whole of the natural history section of the encyclopedia (a major section, with space for over 3000 entries, using the elliptical style), it would mean the loss of several hundred entries.

colugo A nocturnal mammal, native to SE Asia, with a lemur-like face, and a large gliding membrane along each side of its body, extending to the tips of its fingers, toes and long tail. It lives in trees, and eats plant material. Colugos are closely related to insectivores. Also known as **flying lemurs**, they are the only member of the order *Dermoptera*. There are two species, belonging to the family *Cynocephalidae*.

or absorbed by atoms of certain type; spectral line corresponds to electron transition between two energy levels; . . .

It was not at all clear, at this point, whether this last clause meant 'the transition of an electron' or 'the transition of electrons'. Also, which 'spectral line' was being referred to? The need for articles and other grammatical determiners was evident, and it proved necessary to expand the style (after consultation with the physics contributor, to establish what meaning was intended), with the following result:

atomic spectra A set of specific frequencies of electromagnetic radiation, typically light, emitted or absorbed by atoms of a certain type. A given spectral line corresponds to the transition of an electron between two specific energy levels . . .

Similarly, there may be elements in gazetteer or natural history entries which do not work well elliptically. This is sometimes the case at the end of an entry, where there might be a

play descriptive detailed listings, sequences of events, and classifications, often involving special conventions of nomenclature and abbreviation ('. . . Pain is felt over the forehead, face, or jaw supplied by one or more of the branches of the trigeminal (Vth cranial) nerve . . .')

○ Several types of entry motivate the use of what might be called *elliptical* or *telegrammatic* style.

Elliptical style is found in contexts where the information to be presented is determinate and concrete, and where the omission of articles, auxiliary verbs, and other such functors would not cause ambiguity. In the real world, it is commonly seen in accounts of localities in gazetteers, touring guides, and holiday brochures, and thus would seem to be well suited to the treatment of localities in a general encyclopedia (cf. Niamey in Panel 5). When space is at a premium, a great deal more information can be presented than if an expanded style were used, and there is no risk of ambiguity. A second major context promoting the use of an elliptical style is natural history. Here too there is a focus on a specific range of discrete facts, such as data about species or genus, and the defining features of the fauna or flora. The elliptical style provides a compactness of expression which allows readers to focus immediately on the salient features; they do not have to dig through a mass of expository prose to find out what they want to know; and if the information is *not* there, this also can be seen straight away. Panel 6 shows a pair of natural history entries, one in the elliptical style used in the *Cambridge Encyclopedia*, which I have 'translated' into a non-elliptical style, keeping the content identical.

On the other hand, this kind of style is highly inappropriate for topics where abstract reasoning is involved (e.g. philosophy) or qualitative judgements (e.g. literature) or where a great deal hangs on the choice of function word (e.g. physics). As a matter of interest, the physics contributor to the *Cambridge Encyclopedia*, having read the editorial guidelines about style, tried out an elliptical style for the first draft of his entries. It did not work, as will be plain from the following extract:

atomic spectra Set of specific frequencies of electromagnetic radiation, typically light, emitted

Elliptical and expanded styles

A natural history entry showing the elliptical style followed by the expanded-style 'coda'.

carob An evergreen tree or shrub, growing to 10 m/30 ft, native to the Mediterranean region; leaves with 2-5 pairs of leathery oval leaflets; flowers lacking petals, borne in short catkin-like inflorescences; pods to 20 cm/8 in long, pendent, violet-brown when ripe; also called **locust tree**. It is cultivated for fodder, the pods containing a nutritious pulp rich in protein. The seeds are said to be the original carat weights used by goldsmiths. (*Ceratonia siliqua*. Family: *Leguminosae*.)
>> evergreen plants; false acacia; inflorescence [1]; protein; shrub; tree [1]

general observation made about the subject, as a kind of semantic 'coda' (see Panel 7).

Who, what, why, when . . . ?

There are several other features of the *Cambridge Encyclopedia* which raise interesting linguistic questions, such as the system of cross-referencing employed, and the relationship between verbal and non-verbal (pictorial) text: but a discussion of these must be left for another occasion. I have chosen to focus in this paper on what I feel to be the primary linguistic characteristic of the work, which is the stylistic variety of the entries. A consequence of this variety, of course, is that the criteria of consistency and clarity referred to above need to be re-evaluated. Looked at superficially, as one might when casually looking down a column, there seems to be no consistent style. But such an absolute notice of consistency, imposed on entries regardless of their purpose, is I believe an artificial and distorting concept. Rather, a higher order notion of consistency needs to be invoked – a

genuine attempt to match form to function, to choose language which is appropriate to a particular type of topic, and to find a level of clarity of expression which will meet (or at least not exceed) the expectations of the reader.

But what are these expectations? What sort of person is likely to look up the entries in a general encyclopedia? Who exactly needs to know, or be reminded of, the information the entries contain? I do not know, having carried out no survey. Are the only people who look up the more technical chemistry entries those who already have some kind of vested interest in chemistry – in other words, pupils (or their parents), students, professionals on the margins of the subject – people, in short, where one can assume a degree of prior knowledge? Who looks up what, why, when? The need for reader surveys is paramount. Until we have some systematic information of this kind, encyclopedia editors and publishers are working largely in the dark, and their policies and principles are largely informed guesswork.

All of this applied to my own case: the policy of maintaining in these entries the stylistic norms and conventions of expression from the various domains of the language, and introducing a level of exposition that makes varying demands on the reader, is based entirely on a set of assumptions about the intellectual and linguistic experience of those likely to consult the entries in the first place. A judgement has had to be made, for all subject areas. If this judgement turns out to be wrong, the editor will be the first to know, because people will write in and tell him, in no uncertain terms, that he (along with his contributors) has either (a) overestimated or (b) underestimated this experience. Such reactions will be invaluable, as they will provide a new kind of evidence bearing on the stylistic principle involved, which should then, as a consequence, become more powerfully and precisely interpreted. **ETI**