

Our talk is split into two parts, I will be talking about the computer from the point of view of publishing and more specifically lexicography - and as a user of computer systems, Ken Moore will be dealing with the hardware and software that we are using at Longmans, and the overall thinking that has gone into systems design.

Firstly, let me explain that Longmans has only recently been thinking of getting back into trade dictionaries - although I should mention that we were the publishers of Johnson's first English dictionary and also Roget's Thesaurus so that, at least as a company we are by no means newcomers to the game.

I would like to talk to you about two completely different kinds of dictionary - before moving on to some more speculative areas which we would hope would provoke discussion.

The first dictionary is entitled The Longman Dictionary of Contemporary English [code name LDOCE] whose characteristics I will summarise briefly as follows:

1. It has about 40,000 main entries, that is it is about the size of the Concise Oxford but it is intended primarily for the foreign learner of English as a medium of international communication.
2. Because it is intended for the Foreign learner it has two very distinctive features:
 - a) the text of all the definitions and examples is written in a limited vocabulary of approximately 2,000 words - that means that all concepts however complex have to be explained within this vocabulary.
 - b) almost every meaning of every word is preceded by a code which indicates the range of syntactic possibilities permitted. A glance at the coding table* will demonstrate how this system operates.

*see page 6

- c) there is a fairly massive amount of linguistic information attached to each meaning and to each entry word - most of which does not appear in the published work but is held on file for retrieval purposes, as for scholarly research. This information includes:
- i) Orthographic
 - ii) subject area
 - iii) national/regional restrictions
 - iv) frequency and temporal restrictions
 - v) usage level, attitudinal connotation or stylistic context
 - vi) semantic constraints

(show LDOCE proof with coding)

This is the information which is of most interest to us as computational linguists because it opens the doorway to the use of LDOCE as a reference dictionary to support such projects as

- a) automated citation reading (explain)
- b) automatic grammatical tagging and syntactic analysis

These developments are the first step along that long road towards the development of an effective machine-translation system, provided that a mirror image were produced in the target language, e.g. French, of the linguistic structure of the English dictionary. This is however a very long road.

(mention CALM)

I would now like to turn to a completely different dictionary programme, the one with which Ken Moore and I are chiefly concerned. This is a venture to produce a British-based dictionary for world markets from a best-selling American dictionary available to us on computer tape, together with a subject-field coding (or codings) attached to each definition. This is a work approximately twice the size of LDOCE and represents our first encounter (close but probably of the second kind) with in-House computing using the PDP 11.

The editorial features which mark this compilation out as a new development are summarisable as follows:

- i) the work, in its first phases, is entirely constructed by subject field - that is all the definitions in maths or medicine are edited as a separate entity before entering the total alphabetic arrangement.
- ii) the dictionary is based on a citation - reading programme for each of those subjects (explain)
- iii) all work is undertaken in direct interaction with an on-line system via input and retrieval VDUs, enabling consultation of the file to take place at any stage, and proof-listings to be generated without the expense of going through a composition system, either of portions of the whole text, or of any preceded selection of the data. We have thus virtually replaced manual filing systems for day-to-day working. A further feature here is that notes can be sent from one editor to another which are keyed to a particular field or entry.
- iv) We have evolved a system called vocabulary control which is intended to ensure that the language used to explain meanings is simpler than the word being explained.

(explain vocab. control and the concept of "the computer as proof reader".)

I would now like to talk about two further areas before stopping for questions. These are:

- a) file manipulation for editorial purposes
 - i) cross-referencing
 - ii) checking of limited vocabularies
- b) linguistic or publication spin-offs
 - i) thesauruses/semantic coding
 - ii) collocation programmes
 - iii) specialist dictionaries - e.g. special subject, derivatives, scrabble, rhyming.
 - iv) grammars, course materials
 - v) linguistic studies, e.g. orthography.

Codes	Table of		A	B	C	D	E	F	G	H	I	L
	or	used before a noun										
nominals with no letter are [A], adjectives with no letter [B]	[A1] the main difficulty	[B1] I'm happy	[C1] 4 boxes	[D1] Give me a book	[E1] the president elect	[F1] She was asleep	[G1] the committee	[H1] right through	[I1] He passed	[L1] She became tired		
0 need not be followed by anything												
followed by one or more nouns or pronouns												
followed by the infinitive without to												
followed by the infinitive with to												
followed by the -ing form												
followed by a finite clause												
followed by a wh- word												
followed by an adjective												
followed by a past participle												
needs a descriptive word or phrase												

Codes	Table of		N	P	R	S	T	U	V	W	X
	or	used in direct address									
vocative n	[N1] Good-bye, doctor!	[P1] The police are here!	[R1] God!	[S1] I shink!	[T1] She kicked the boy!	[U1] sugar is sweet	[V1] I want to see the man later	[W1] I saw the man later	[X1] I saw the man later		
plural n											
n that is a name or name-like											
singular n											
transitive v with one object											
uncountable n											
v with one object - verb form											
adjective and adverb											
noun											
preposition											
verb											
v with one object - something else											

Note: the numbers in the codes in this box have a different meaning

[W21] takes personal pronouns as objects

[W22] takes personal pronouns as objects

[W23] takes personal pronouns as objects

[W24] takes personal pronouns as objects

[W25] takes personal pronouns as objects

[W26] takes personal pronouns as objects

[W27] takes personal pronouns as objects

[W28] takes personal pronouns as objects

[W29] takes personal pronouns as objects

[W30] takes personal pronouns as objects

[W31] takes personal pronouns as objects

[W32] takes personal pronouns as objects

[W33] takes personal pronouns as objects

[W34] takes personal pronouns as objects

[W35] takes personal pronouns as objects

[W36] takes personal pronouns as objects

[W37] takes personal pronouns as objects

[W38] takes personal pronouns as objects

[W39] takes personal pronouns as objects

[W40] takes personal pronouns as objects

[W41] takes personal pronouns as objects

[W42] takes personal pronouns as objects

[W43] takes personal pronouns as objects

[W44] takes personal pronouns as objects

[W45] takes personal pronouns as objects

[W46] takes personal pronouns as objects

[W47] takes personal pronouns as objects

[W48] takes personal pronouns as objects

[W49] takes personal pronouns as objects

[W50] takes personal pronouns as objects

[W51] takes personal pronouns as objects

[W52] takes personal pronouns as objects

[W53] takes personal pronouns as objects

[W54] takes personal pronouns as objects

[W55] takes personal pronouns as objects

[W56] takes personal pronouns as objects

[W57] takes personal pronouns as objects

[W58] takes personal pronouns as objects

[W59] takes personal pronouns as objects

[W60] takes personal pronouns as objects

[W61] takes personal pronouns as objects

[W62] takes personal pronouns as objects

[W63] takes personal pronouns as objects

[W64] takes personal pronouns as objects

[W65] takes personal pronouns as objects

[W66] takes personal pronouns as objects

[W67] takes personal pronouns as objects

[W68] takes personal pronouns as objects

[W69] takes personal pronouns as objects

[W70] takes personal pronouns as objects

[W71] takes personal pronouns as objects

[W72] takes personal pronouns as objects

[W73] takes personal pronouns as objects

[W74] takes personal pronouns as objects

[W75] takes personal pronouns as objects

[W76] takes personal pronouns as objects

[W77] takes personal pronouns as objects

[W78] takes personal pronouns as objects

[W79] takes personal pronouns as objects

[W80] takes personal pronouns as objects

[W81] takes personal pronouns as objects

[W82] takes personal pronouns as objects

[W83] takes personal pronouns as objects

[W84] takes personal pronouns as objects

[W85] takes personal pronouns as objects

[W86] takes personal pronouns as objects

[W87] takes personal pronouns as objects

[W88] takes personal pronouns as objects

[W89] takes personal pronouns as objects

[W90] takes personal pronouns as objects

[W91] takes personal pronouns as objects

[W92] takes personal pronouns as objects

[W93] takes personal pronouns as objects

[W94] takes personal pronouns as objects

[W95] takes personal pronouns as objects

[W96] takes personal pronouns as objects

[W97] takes personal pronouns as objects

[W98] takes personal pronouns as objects

[W99] takes personal pronouns as objects

[W100] takes personal pronouns as objects