ONLINE OR HARD COPY?

Marion M. Strachan
Oxford University Press, Oxford, United Kingdom

ABSTRACT: There is still an overwhelming demand for hard copy general dictionaries. However, the changed needs of various smaller user groups, not least scientists and technologists, are being increasingly recognised by lexicographers. Collaborative effort by the two parties *is* needed if up-to-date (and regularly updated), machine-readable scientific dictionaries are to be produced.

Let me start by defining my terms. For ease, I will use the phraseology of my title, but with modifications. 'Online' *may, for* certain people in certain circumstances, represent too great a narrowing of the field - so let us keep the broader term, 'machine-readable' , in mind. As for 'hard copy' - it is a useful term. It embraces what results from line printers as well as what results from hot metal, so I shall retain it. Nevertheless, I have a wholly subjective dislike of the term being used to describe an attractively printed book, even though the vast majority of these are now produced with the aid of computer technology.

Conventional dictionaries are, at the moment, hard copy. Only a fairly limited number of scientific and technical databases are currently online. This is not because lexicography has resolutely refused to abandon what Hartmann(7) calls its 'shoe-box file(s)'. On the contrary.

> 'The computer has been applied in numerous ways in the humanities, but nowhere has it been more useful than in lexicography. All phases of dictionary making can benefit. These range from the collection of the original research material to the production of new editions.' Alford(l)

The technology exists to produce any dictionary in machine-readable form and many of today's general dictionaries could actually (or potentially) be sold or leased in that form. So why are they still conventional books? (Indeed, why is 'dictionary' used to describe the printed work, and that only, and 'term bank' used to describe the online store of

specialist terminology? The dichotomy seems to me to be sad but significant, and that is something I will return to later).

Why are dictionaries hard copy? One major reason is obviously that, for five hundred years, print has been our pre-eminent way of communicating information and the habit is deeply ingrained in our culture. On a slightly different level, I, as a lexicographer working in the field of teaching English as a foreign language, cannot afford to ignore considerations of practicability, cost, and demand. Dictionaries for the same level of user are sold in many different countries. Some of those countries are technologically sophisticated and economically stable. Others, and they include major markets, are still seeking to acquire the basic technology most of us take for granted, and have difficulty in providing universal education for their populations, never mind buying books however cheaply we try to produce those books for them. The demand - from our point of view - has been, and still is, overwhelmingly for books. In addition - and, again, from the lexicography-cum-publishing point of view - by far the greater number of dictionary users at even fairly sophisticated levels are quite clear that they want their dictionaries to improve their treatment of specialist vocabularies of various kinds, but they do not feel a need for a really specialised dictionary. As our chairman and others have pointed out,

> 'In the English speaking world, applied linguistics is even today almost exclusively concerned with the problems of teaching strategies and methodology for the very simple reason that the teaching of English as a foreign language is such a major preoccupation in comparison with which translation or any other form of multilingual communication assume only a very minor role; most English speakers' information capacity is satisfied by publications in English.' Sager, Dungworth, and McDonald(12)

I have spent some time on the current working situation of a sizeable number of lexicographers to try to show you our present preoccupations and areas of activity. And, to sound a thoroughly realistic note: lexicography has to be paid for by somebody, and the area I have just described is not only worthwhile, it is profitable enough to encourage publishers to enter it.

However, none of this can yet have encouraged you very much. Your needs are very different because you, unlike many of our major user groups, have been, and are, on the immediate receiving end of two not-unrelated developments that have taken place in the last 30 years or so. Firstly, the computer which, 40 years ago was having trouble trying to do sums faster than the centuries-old abacus, has become a device capable of handling vast quantities of information with a speed and a relentless logic that the human mind cannot match. Secondly, scientists and technologists in particular have been on the receiving end of a multilingual information explosion. As a result, questions of practicability, cost, and demand for online information have altered radically for one group of users.

The industrialised nations have the infrastructure to support online term banks - namely, reliable telecommunications and power supply. At the same time, suitable hardware and a great deal of software are available. In real terms the cost of these two last continues to go down, so one can see why the chairman of this conference should feel frustrated.

> 'Online information retrieval has been operating with increasingly widespread success and efficiency for many years already, but when will publishers start compiling dictionaries for this market?' Snell(13)

Underlying the request for online instead of hard copy is, I suspect, the dissatisfaction which has led to the dictionary-term bank dichotomy to which I referred before. Online alone is not enough. A change of medium without a change of approach would leave you no happier. You have become increasingly dissatisfied with both dictionaries that appear not to have adapted in any significant way to meet your changed needs and with lexicographers who seem equally oblivious.

However, you can affect dictionaries and their makers. Uvarov and Isaacs(14) described how they (and others)

> "... attacked the Oxford dictionaries . . . for their lack of scientific words. ... Oxford were pretty cross about it .... In any case, Oxford employed a scientist, J.B. Sykes, to revise the <u>Concise</u> and ... the new edition (1976) (is) an excellent source of scientific information.'

For your purposes, this is not enough, but it shows that, if your needs are forcefully expressed in the right direction, lexicographers are prepared to respond. Leading on from that, lexicographers are increasingly aware that

> ' ... while until comparatively recently lexicographers had scarcely looked beyond one type of user - persons of cultivated literary tastes, sharing the same educational and linguistic background as themselves - more recently they have been led to acknowledge that the choice of linguistic information in a dictionary, and the means of access provided to it, will vary with the class of user for whom it is intended.' Dubois(4)

The reassessment of who uses a dictionary, how, and why is of central importance if there is to be any useful future collaboration between lexicographers and the scientific community. Unfortunately, there is a danger that the effects of the lexicographical debate will take too long to be expressed in practical terms. To offset that thought, it is worth pointing out a couple of areas in which lexicographers and terminology specialists are working along similar lines, although using different subject matter. In the EFL field, a great deal of attention has been paid in recent years to collocations, and to the ways in which collocability can be usefully indicated in dictionaries. The "extreme" form of collocability in general English is found in our thousands of idioms which constantly confuse foreign

learners. The approach which has been adopted in one dictionary on idioms, the <u>Oxford Dictionary of Current Idiomatic English(3),</u> with which I have spent a great deal of time recently, is in certain respects very similar to the philosophy of Eurodicautom as described by Goetschalckx(S).

> 'The phraseological approach is very important in the scientific and technical field where we have to deal with so-called special languages. Specialisation by our translators is not always possible. For this reason we have to offer full information on the use of the terms in this particular field and vary often also a lot of technical information. If these phrases of sentences are well chosen, they can cater for both the linguistic and technical information needed'.

Another example is the increasing readiness of lexicographers not to see words as discrete entities. An advanced EFL dictionary (such as the <u>Oxford Advanced Learner's Dictionary of Current English(9))</u> will not as a matter of course include much more cross referring between words linked as, say, synonyms or antonyms, or broader or narrower terms, than will a dictionary for native speakers who are assumed already to have acquired most of that sort of linguistic awareness. Indeed, a number of people have taken this approach to the point of abandoning the traditional alphabetical ordering in favour of an arrangement which better reflects the subject area covered. Hann(6) in a reassessment of dictionary/glossary compilation (he himself used both terms to refer to his work on a limited vocabulary drawn from computer science) evolves a hierarchical structure which is not unlike that of the British Standards Institution's Root Thesaurus(2). He concludes that

> '... this method of dictionary production has certain other advantages over the conventional lexicographical (i.e. alphabetical) approach in that specialist terminology is more systematically (and thus invariably more adequately) covered than is the case with alphabetical dictionaries, since important terms essential to the conceptual structure of the dictionary are not easily overlooked.'

(Incidentally, I personally do not feel that alphabetical order is sacrosanct. It is useful in that it is the least culture- and subject-specific approach readily available. However, if a given subject area benefits from a different approach, then use that approach - an alphabetical index can, in most cases, be provided to help the uninitiated).

In non-scientific fields lexicographers are working on areas which are also of relevance to your needs. Increasingly, lexicographers are aware that the needs of users are changing and should be met in new ways. But we face one vary considerable obstacle. With honourable exceptions, such as Professor Sager, relatively few applied linguists, and even fewer lexicographers, apply themselves to serious study of scientific terminology. I am not going to attempt a sociological analysis of the reasons for this. However, one reason can be fairly safely assumed. Most lexicographers are trained in the humanities and they simply so not feel equipped to cope with disciplines that seem to the non-scientist to be increasingly possessed of thought processes all their own. There is a serious breakdown in communication between scientists and non-scientists, and the latter are, often enough, so sure they cannot cope with esoteric and fast-developing disciplines that they see little point in trying. Perhaps this is why the sciences and technologies appear to be developing their own lexicography.

As I see it, you as a group are already well aware of, and trying to find solutions to, a number of problems.

- What the criteria for judging whether or not something (probably a document) constitutes an admissible source of terminology?
- Who is going to apply these criteria and process the terminology that results? Will it be done by one firm or translation bureau, by a national body, or by an international body?

- Will terms be presented singly, or will the contextual approach be adopted?

- Will subject fields be treated separately, or incorporated into a larger corpus?
- At what intervals will the material be updated, and who will decide what has become obsolete?

Lexicography by any other name still looks very much like lexicography. If I did not know I had just been describing the work of terminology specialists, I would assume that I had just outlines the problems of any lexicographer.

If lexicographers are to help scientific and technical users, then it must be a collaborative effort. Lexicographers will need subject specialists, and both may need the help of national and international bodies (Opitz(10)).

> 'Dictionaries are ultimately made for and judged by the needs of real people in real situations. The traditional typologies will have to be revised in the light of research into specific communicative conflicts. There will probably always be a case for the all-inclusive all-purpose dictionary, but the current trend towards specialisation should not be ignored. ... Another stimulus to professional lexicography has come from ESP and the analysis of terminological conventions of technical disciplines from navigation to low-temperature physics. Here too it will be necessary to move from...the single enthusiast to more collaborative ventures involving user tests and interdisciplinary contact. The promise of increased automation is likely to speed up these developments and may perhaps lead to completely new types of instant-access reference tools.' Hartmann(7)

Which brings me back to the immediate surface meaning of my title. Once the producer is reasonably confident that the material is what the user (in this case, you) really wants, how that material is presented largely depends on what the user is prepared to pay for - which is not itself something the lexicographer can decide. If dictionaries are to be relevant to you as a specialist group, the message is that they must be machine-readable. There is no other way that we have yet found of ensuring that they are sufficiently up-to-date for your needs. There are problems for publishers of reassessing pricing structures and marketing policies (see J. Page(11) for a discussion of this with reference to online information services), but these are secondary considerations. Online rather than, say, batch access to such dictionaries will be expensive. However, the speed and convenience may, for certain organisations, more than balance the cost. That is for you and your employers to decide.

So, if it is to be online (and that seems increasingly likely), will it be 'online or...'? I would be inclined to opt for 'online and...'. Those of you who want the super-duper solution may find it easier to achieve if you can call on the support (and finance) of others who may not be able to afford that level of sophistication. A multiple-access machine-readable dictionary might be the answer. Online search, batch search, and, perhaps loose-leaf hard copy as well. Loose-leaf hard copy is readily updateable and, while it is a medium less widely used in the UK than in either America or on the Continent, the necessary knowledge and experience are available. On a more mundane level, hard copy might prove its worth on those (fortunately infrequent) occasions when our gee-whizz technology falls prey to power cuts, especially if they result from planned industrial action or from the sort of winter we had last year, or, indeed, when there is a more localised disaster, such as a disk-head crash.

The need for much improved user-specific formats for dictionaries is felt by the customer and increasingly acknowledged by the lexicographer. The two debates we are interested in today are alive and well but perhaps still too inclined to live in their respective camps. Maybe we now need to do more direct talking with each other. If that step can be taken positively, it

> '...seems entirely possible today that increased international...contacts, communication technology, and training programmes will give us better products tomorrow.' Hartmann(8)

REFERENCES

(1) ALFORD, M.H.T. The computer and lexicography. ALCC Bulletin, 1(3), Michaelmas 1973, 8-9.

(2) BRITISH STANDARDS INSTITUTION. Root Thesaurus. London: British Standards Institution, 1981.

(3) COWIE, A.P. and MACKIN, R. Oxford dictionary of current idiomatic English Vol I: Verbs with prepositions and particles. London: Oxford University Press, 1975.

 COWIE, A.P., MACKIN, R. and MCCAIG, I.R. Vol 2: Phrase, clause and sentence idioms. Oxford: Oxford University Press (forthcoming mid 1983).

(4) DUBOIS, J. Models of the dictionary: evolution in dictionary design. Applied Linguistics, 2(3), Autumn 1981, 239-49.

(5) GOETSCHALCKX, J. Eurodicautom. In: B.M. Snell, ed, Translating and the computer. Seminar organised by the Technical Translation and Informatics Groups of Aslib. Amsterdam: North-Holland, 1979, 71-75.

(6) HANN, M.L. Computers and the production of systematic terminological glossaries. ALLC Bulletin, 5(1), 1977, 26-37.

(7) HARTMANN, R.R.K. Dictionaries, learners, users: some issues in lexicography. Applied Linguistics, 2(3), Autumn 1981, 297-303.

(8) HARTMANN, R.R.K. Advancing definitions. Times Literary Supplement.

(9) HORNBY, A.S.   Oxford advanced learner's dictionary of current English. 3rd edn. London: Oxford University Press, 1974, rev. 1980.

(10) OPITZ, K.   Technical dictionaries: testing the requirements of the professional user: In: R.R.K. Hartmann, ed, Dictionaries and their users. Papers from the 1978 BAAL seminar on lexicography. Exeter Linguistic Studies, Vol. 4. Exeter: University of Exeter, 1979. 89-95.

(11) PAGE, J. Main elements in costs and prices for online services: an overview.   In: EURIM 4: a European conference on the contribution of users to planning and policy making for information systems and networks, April 1978. London: Aslib, 25-27.

(12) SAGER, J.C., DUNGWORTH, D. and MCDONALD, P.F.   English special languages: principles and practice in science and technology. Wiesbaden: Brandstetter, 1980.

(13) SNELL, B.   Electronic translation?   Newsletter of the Translators' Guild of the Institute of Linguists, 5(1), Winter 1980, 15-16.

(14) UVAROV, E.B. and ISAACS, A.   Talk on the Penguin Dictionary of Science. Aslib Technical Translation Bulletin, 24(3), December 1978. 90-95.