



Teaching Lexicography or Training Lexicographers?

Michael Rundell

Michael Rundell has been a professional lexicographer since 1980, working on a wide range of English dictionaries. As Managing Editor at Longman Dictionaries (1984-94), he was responsible for running major projects including the *Longman Language Activator* (1993) and *Longman Dictionary of Contemporary English* (2/e 1987 and 3/e 1995). He had a leading role in the design of several lexicographic corpora (including the Longman Lancaster Corpus, Longman Learner Corpus, and British National Corpus) and was a member of the Advisory Committee for the forthcoming American National Corpus. As well as running training courses for Longman lexicographers, in both the UK and US, he has taught for many years on the Interlex and MA programmes at the University of Exeter. He is now a freelance consultant and divides his time between dictionary editing and lexicographic training. He is Editor-in-Chief of the forthcoming *Macmillan English Dictionary*. michael.rundell@dial.pipex.com

In an interesting paper at last year's Euralex Congress, Tamás Magay addressed the theme of teaching lexicography, and raised the issue of "where and how lexicography is taught in European universities" (2000). This is a highly relevant topic for almost any member of Euralex, but – as so often happens when this subject is discussed – no distinction was made between teaching people *about* lexicography, and training people to *be* lexicographers. The debate needs to move into this territory because there appears to be, in some quarters at least, an implicit assumption that the key to better dictionaries is simply more university-based training. To quote Magay: "Once we accept that dictionary-making is now a profession, it has to be taught at the highest level" (p.443). This is merely the latest airing of a view that has been floated at various times over the last 20 years (see for example the collection of papers edited by Robert Ilson in the mid-80s). But is it true?

The question is timely because there is currently a lot of activity in this area. The University of Exeter's Dictionary Research Centre (DRC) – which pioneered courses in lexicography in the UK, and has a well-established MA programme and a popular annual short course – is about to move to a new home at (very appropriately) the University of Birmingham. Meanwhile, the Lexicography MasterClass (LMC, consisting of Sue Atkins, Adam Kilgarriff and myself) is about to launch a one-week workshop at the University of Brighton's Information Technology Research Institute (ITRI), to be followed in 2002 by a new MSc programme in lexicography and lexical computing. Against this background, it seems incumbent on course-providers to be clear about the sort of knowledge and skills they are undertaking to equip prospective students with.

Lexicography is a legitimate academic subject in its own right, not least because of the wide range of "feeder" disciplines – lexical semantics, second-language acquisition theory, and computational linguistics, to name just a few – that supply a theoretical perspective against which lexicographic

issues can be intelligently addressed. Consequently, people working in many fields can derive great benefit from learning more about the theoretical, pedagogical, and computational underpinnings of the lexicographic process. This does not, however, turn them into lexicographers. As any practising lexicographer knows, the only real way to learn how to write dictionaries is to write dictionaries. Most good dictionary publishers, whether operating in the commercial or academic sphere, provide on-the-job training for their staff – typically some combination of initial "basic training" followed by ongoing feedback and mentoring over a long period. Against this background, what is the role for courses in lexicography provided by academic institutions?

Perhaps another way of framing this question would be to look at the skills a contemporary lexicographer needs and consider how far these are (or can be) supplied within dictionary publishing houses through the traditional "apprentice" model. Probably the most central task in general lexicography consists of analyzing very large quantities of primary data (from a corpus) and imposing some sort of order on it – or, more accurately, *discerning* the underlying order within it and then describing this in a way that is both useful and relevant to a particular group of dictionary users. The better one understands every stage in this process, the more successful one is likely to be. If you know, for example, where your corpus data comes from, why the corpus is designed in the way it is and how is it linguistically annotated, you are in a better position to conduct sophisticated searches, to perceive patterns and regularities in the mass of data, and to distinguish between what is relevant and what is marginal. Or again, the process of discovering meanings is likely to be more effective when gut instinct is complemented by some understanding of lexical relations or frame semantics. Traditionally, many lexicographers have acquired high levels of competence as data analysts, seemingly by osmosis and without necessarily being able to articulate the criteria that underlie their decision-making. But now,

with so much data to process, lexicographers need all the help they can get, and linguistic theory has an important role to play in informing the judgments that we make.

Dictionary publishers remain the primary suppliers of training in the practical skills of lexicography. But it is not realistic to expect them to deliver the full range of training needs – partly because they will not always have expertise in all of the relevant theoretical disciplines, and partly for straightforward business reasons. Though the best publishers set aside time and resources to train their staff, the normal pressures of deadlines take their toll. Meanwhile, changing patterns of work have reduced publishers' capacity to nurture new editors: large in-house dictionary teams (which provide a supportive learning environment for new lexicographers) are becoming more of a rarity. The arrival of email, intranets and inexpensive high-powered computing has had the effect of dispersing editorial teams, as experienced people increasingly opt to work from home, sometimes for several different publishers. Though this has many benefits, for publishers and employees alike, one of its less positive side-effects is that few publishers can now provide all the necessary training.

In tandem with these changes, we are also seeing (and not before time) the development of closer links between lexicography on the one hand and the natural language processing community on the other. The benefits of collaboration between these two sectors are at last beginning to be recognized, but this in turn requires lexicographers to learn yet more skills. This, then, is the environment in which the LMC has begun to plan new courses. It became clear to us that the best model for an effective training programme would be one that combined a grounding in relevant theoretical subjects with a strong element of hands-on learning of practical skills. Over the past four years or so we have, either collectively or individually, run a number of customized short courses for institutions in various parts of the world. This July, however, we are launching our first *general* programme in Brighton – lexicom@itri – a one-week training workshop.

The term "workshop" is deliberate, because a key feature of the programme is that lecture sessions alternate with periods in the computer lab, where participants will do practical tasks that relate directly to the subject of the previous lecture. There is an analogy here with professional disciplines such as law or medicine, where a good theoretical grounding is a starting point for the development of practical skills. And just as a detailed knowledge of legal precedent cannot in itself turn someone into an effective courtroom advocate, so a familiarity with the minutiae

of metalexicographic theory does not *per se* make someone a good dictionary writer. The practical and theoretical go hand in hand. The three areas we aim to cover in detail are: designing, building and working with a text corpus; creating and using a dictionary database; and, the process of writing actual dictionary entries. Along the way, we will take in issues such as corpus annotation (POS-tagging and other ways of enriching raw text data), smart approaches to data extraction (including some programming skills), and the principles of writing definitions.

The workshop is not aimed at complete beginners: rather, we expect participants to have some grounding in one of the three main disciplines involved (lexicography, linguistics and computer science), so that they can use the course to learn more about the other subjects that have a bearing on their work, and of course to contribute their own perspective in those areas where they are already experienced.

Lexicom promises to be a great learning opportunity for everyone involved – not least the tutors themselves, of course, since the best thing about lexicography is that there is always more to learn. When we planned the course, we decided that we would need at least 15 participants to make it worth doing. In the event, we have had to cap the attendance at 50 and create a waiting list for lexicom 2002. This level of interest suggests there is plenty of demand for courses of this type. We are currently developing, with several other colleagues at Brighton, a new MSc programme in Lexicography and Lexical Computing, which has been approved to start up in October 2002.

To return, finally, to the question we touched on earlier: what skills can participants in the lexicom workshop expect to go away with? It would be unwise for us to claim that we can turn people into lexicographers in a week. A more reasonable objective, however, is to give aspiring dictionary-editors enough grounding to know whether they have the potential to go further, and to enable established lexicographers to "raise their game" through a deeper and broader understanding of the diverse range of factors that contribute to making great dictionaries.

Our other main objective, of course, is that we all enjoy ourselves in the process.

References

- Ilson, R.** 1986. *Lexicography: an Emerging International Profession*, Manchester: Manchester University Press.
- Magay T.** 2000. 'Teaching Lexicography.' In Ulrich Heid et al. (eds), *EURALEX 2000 Proceedings*. Stuttgart: Universität Stuttgart, IMS.

Lexicom takes place on 16-20 July 2001. The other two members of the Lexicography MasterClass (LMC) are Sue Atkins and Adam Kilgariff. Sue Atkins is an immensely experienced lexicographer, a past president of Euralex, and the driving force behind the establishment of the BNC. Adam Kilgariff, a Senior Research Fellow at Brighton's Information Technology Research Institute (ITRI), is an authority on the interface between lexicography and language technology, and has published extensively in both areas. <http://www.itri.brighton.ac.uk/lexicon>