

Welsh Lexicography and Terminology: past, present and future (II)

Delyth Prys

My colleague, Geraint, has so far provided you with a broad overview of the history of lexicographical work in Wales to the present day as well as our dreams for the future.

I would like to provide you now with some more detail of the content of our work. The three major contemporary Welsh dictionaries cited by Geraint may be categorized as follows:

Geiriadur Prifysgol Cymru (The University of Wales Dictionary) is the definitive historical dictionary of the Welsh language. It is based at the National Library of Wales in Aberystwyth and should be completed this year after 80 years work. At present it is published on paper but electronic versions may in future be produced as CD Roms and/or in an on-line format. Further information on it can be obtained on the web at this address: http://www.wales.ac.uk/geiriadur/gpc_gwes.htm

The *Welsh Academy English-Welsh Dictionary* is an English-Welsh dictionary of contemporary language. Again this is currently available in paper form, but work has started at the University of Wales, Bangor to put it into a computer database so that various applications may be derived from it.

Y Termiadur Ysgol: Standardized Terminology for the Schools of Wales is a fully bilingual terminology dictionary containing both English-Welsh and Welsh-English sections. It was originally compiled in database form, and has so far been published by three methods: in paper form, as a CD Rom, and as an interactive on-line dictionary on the MELIN web site at <http://www.bangor.ac.uk/cyc/melin/index1.htm>

This terminology dictionary is different to the two first dictionaries mentioned above because it was based on *prescriptive* rather than *descriptive* principles.

- traditional lexicography is *descriptive* (it describes the language as it is, recording words and phrases as they occur in natural language)
- terminology work is *prescriptive* (it prescribes which words and phrases should be used as standard forms in a specific technical context, known as language for special purposes).

Descriptive Dictionaries

In natural language there may be a number of words with the same meaning. All these synonyms may be included in a traditional dictionary entry. Of course, lexicographers aren't always as objective and inclusive as they pretend to be, and you find

that even in dictionaries which purport to give a comprehensive list of words in a language, the editor has made certain judgements concerning 'correctness' or 'standard usage' and that this has coloured his or her judgement on the words to be included, and conversely which words to leave out. *The University of Wales Dictionary* is a historical dictionary of the Welsh language, and is very inclusive in its word list for the earliest Welsh manuscripts dating from the 8th century to 1800. However from 1800 onwards, due to the enormous proliferation of written evidence for Welsh after this period, it is far more selective, and criteria such as collecting words from the works of "the best prose writers" (whoever they may be) were used to choose words for inclusion.

Bilingual dictionaries are more prone than historical dictionaries to moving further away from being purely descriptive works. This is because they have to match words and concepts in two different languages and attempt to fill any gaps when there is no readily available equivalent in the target language. This may be done by either giving the nearest equivalent in that language or by coining a new word, and this was sometimes done in the *Welsh Academy English-Welsh Dictionary*. Neologisms in this dictionary are shown with an asterisk, but they are very few. What is more common is the inclusion of neologisms coined by others in the recent past, and in such instances it can be difficult to know whether they are indigenous words or recently coined words, and if they are recent introductions, to what extent they have become an accepted part of the language.

Prescriptive Terminology Dictionaries

Standardized terminology dictionaries such as *Y Termiadur Ysgol* on the other hand, admit openly that they are manipulating the language. If the editor is offered four competing terms for use in a specific area, he or she needs to recommend which one of the four should be used as the standard technical term, or if there is no term adequate for that use, they need to coin a new term for use in the language. This is done according to principles based on objective criteria. This did not necessarily mean that the other three forms are suddenly "wrong" or "incorrect" but that in the interests of consistency and clarity in technical writing, we should all use the same term as the standard form. The criteria we have developed for standardizing Welsh terminology is based on the International Standards Organisation terminology standards, especially ISO 704 Principles and Methods of Terminology and ISO 860 Terminology Work – Harmonization of concepts and terms. These may be summarized as follows: The term should

- be linguistically correct (in orthography and grammar)
- reflect, in so far as it is possible, the features of the concept given in the definition
- be concise
- be able to generate other forms (nouns, plural forms, verbs, adjectives)
- have one to one correspondence to a concept

To these may be added the most important criterion of all, that of common sense. As we say in Welsh *Trech gwlad nac arglwydd* – "The land is mightier than the lord", and in our terminology dictionaries we do not try to fly in the face of common usage if people have adopted a particular term, even if that is not the best term on offer.

Word based vs. Concept based Approach

Another major difference between traditional dictionaries and standardized terminologies is that the traditional dictionaries are word based. This means that you may have a word, such as *mole* in English, which has several different meanings (8 meanings are listed in the OED including: mole (the animal), mole (the chemical unit) and mole (the blemish on the skin) and all these are numbered together under the word *mole*. This has been done to some extent in large paper-based dictionaries in order to save space and provide a clear layout. Terminology work on the other hand is concept based. This means that the concept, or meaning of the term is central to the working process and presentation, and each concept has a separate entry in our database where the definition, synonyms, and corresponding terms in other languages are grouped together.

Early terminology dictionaries in Welsh tended to follow traditional dictionary format, but this does not help the user to understand that a word different meanings. It is especially important for translators to pick the right meaning when they translate a text, as some of the most serious errors in their work occur when they do not understand the concept to be translated. In our Terminology Centre we call this the "Concept of the Mole" after a famous Welsh miss-translation where mole was translated in a document on Chemistry as the 'little furry creature', rather than the 'chemical unit'.

Sometimes the definition in a bilingual dictionary would be given in the target rather than in the source language, as in this entry from an early Geography Terminology dictionary:

E. grain: 1. *grawn* (cnwd); 2. *græn* (mewn craig/pren); 3. *gronyn* (tywod etc.).

Apart from combining three different concepts under one header word, this pattern of bilingual entry does not help someone with insufficient knowledge of the target language to choose the correct translation. This was rendered in our standardized terminology dictionary for schools as

grain (food crop): *grawn*
grain (in rock, wood): *græn*
grain (=particle): *gronyn*

This simple method of denoting the different concepts on paper, using bracketed disambiguators as a means of distinguishing between concepts hides the fact that these are in reality three separate database entries. The full database entry may also be much more than the published part, including full definitions, notes and codes linking the concept to various coding systems and hierarchies. Our terminology databases also include all plural forms of nouns in full, and grammar information such as parts of speech. These may, or may not, be included in published dictionaries according to need.

Other new Terminology Dictionaries

This method of working was first developed by us for *Y Terriadur Ysgol*. This dictionary combined work on the main curriculum subjects in Welsh school education, including science, mathematics, history, geography, art, technology, computing, music and sports. Since then many other new standardized terminologies have followed. These are usually published as bilingual, English-Welsh and Welsh-English dic-

tionaries, either on paper, as CDs, on-line dictionaries, or a combination of these. They include terminology dictionaries for Nursing, Midwifery, Finance, Archaeology, Transport Law, the Environment, Social Work and Social Care, and more are being prepared. Any new terms that appear in them are passed over to our colleagues working on other language applications in Canolfan Bedwyr, so that they are then incorporated into new spell-checkers, thesauri, and so on.

We have always tried to follow international best practice in our work, including the work of the European Association of Terminology Associations and TechNet. Term Cat, the Catalanian Terminology Centre at Barcelona was a great help to us in our early days. We are now trying to share some of this expertise working with other Celtic languages, and are in the process of establishing Term Celt, a terminology association for Celtic languages. We also have an e-mail Welsh terminology discussion list which is archived on the web, and can be seen at <http://www.jiscmail.ac.uk/lists/WELSH-TERMAU-CYMRAEG.html> At present, this has 124 members, mostly professional translators and is a very lively forum in daily use by the profession in Wales.

The Electronic Revolution

The electronic revolution has made a very large difference to our methods of working. Electronic databases and computational aids are by now standard tools for producing dictionaries in all the major language. The *University of Wales Dictionary* and *The Welsh Academy English-Welsh Dictionary* were both created in the pre-computing era, but any new dictionaries – descriptive or prescriptive ones – produced by Canolfan Bedwyr from now on will use the latest computer technology. Already this has led to major internal reorganisation at the University as the various disparate projects dealing with lexicography, standardization of terminology, spell-checkers and other language utilities, computational linguistics and the Welsh language are brought together under the same roof. This, we believe, is vitally important in a minority language situation. We do not have money to waste, and need to recycle resources from one project to the next.

I will conclude therefore with a diagram showing how all the different dictionaries, terminology databases, spell-checkers, and other language utilities produced at the University of Wales, Bangor feed into each other. At the risk of over-simplification, this frugal use of resources enables us to produce cost-effective language tools for the Welsh language.

Without the electronic revolution, this would not have been possible.

