The indexing of Welsh place-names

Donald Moore

The indexing of Welsh place-names poses two special problems: certain initial letters which are subject to change (‘mutation’) according to fixed rules, and the use of hyphens. Both problems are discussed, and many examples cited. Lists and gazetteers that have been influential in attempts to achieve consistency in the spelling of Welsh place-names are cited. The merits and demerits of letter-by-letter and word-by-word sequences in this field are considered.

Introduction

Why should the indexing of Welsh place-names require special explanation? How could they differ from English place-names? After all, one might say that the indexer has only to record what an author wrote, and arrange the words in alphabetical order. The short answer is that Welsh place-names pose two sorts of problem not encountered in English.

First, the initial letter of a Welsh noun, common or proper, is not constant in all types of usage, but is liable to change according to certain linguistic rules. This phenomenon obliges the indexer to be aware of possible identities among apparently different words, and sometimes to make a personal decision on the alphabetical position of the main entry.

Secondly, the orthography of Welsh place-names often involves the use of hyphens between the elements which constitute the name. This convention has two main functions, which will be explained later; but for the indexer it means that the hyphenated components must be treated as separate words. The indexer will then be surprised to discover what is apparently the same place-name spelled as a single word, or even completely separated into two or more words. If a word-by-word sequence is being employed, the implications are obvious: two separate sequences could arise in what is logically one group of place-names. Since these difficulties will also affect the user, the indexer must take care to provide suitable cross-references.

Mutation

Let us begin with the problem of the initial letter. The Welsh language has a peculiarity known as ‘mutation’, which it shares with its kindred languages, Cornish and Breton. The first (and, occasionally, the last) letter of a word may change so as to indicate its relationship with a preceding, or following, word. The nature of the change depends to some extent on the sounds concerned, and also, in the case of nouns and adjectives, on the gender of the word (Welsh nouns are either masculine or feminine). This phenomenon is deeply rooted in speech and writing, and is an essential part of classical Welsh, giving a distinctive character to the strict metres of Welsh poetry.

Some examples are necessary to demonstrate the effects of mutation. The letter c (always hard in Welsh) may appear in three other forms: g, ngh and ch. Thus the noun caer (= fort) may appear in a Welsh text also as gaer, nghaer or chaer, depending on the associated words. A Welsh-language indexer would disregard the changes brought on by mutation, and index all forms under the ‘radical’ or dictionary form, viz. caer. This word, as might be expected, is a frequent element in place-names, where it is sometimes preceded by the definite article y, and sometimes not. In the former case, because it is a feminine noun, it becomes y gaer. This change is known as a ‘soft’ mutation. To denote a ‘large fort’, the adjective mawr has to be added, but it, too, must be mutated after a feminine noun, so that we have y gaerfawr. There is such a place near Merthyr Cynog in Brecknock, but it has lost its article in common usage and appears as Gaer Fawr. By contrast, there is a place of the same description near Llanfaredd in Radnorshire, known as Caer Fawr, where the initial mutation has not taken place! These are the present accepted forms, but the indexer cannot assume that they will have been so in older references, scattered in the transactions of antiquarian societies over the last hundred-odd years. A large-scale Ordnance Survey map may thus be necessary to determine whether similar-looking place-names denote the same or different localities.

To complete the story of caer, it must be said that where it is followed by a dependent noun, as opposed to an adjective, it does not require the definite article and is therefore not mutated. Thus Caer Gybi (the site of a Roman fort at Holyhead) means Cybi’s fort, where the name of Cybi (a Celtic saint) has been mutated after the feminine caer. The two other possible mutations of c do not occur after the definite article, so that we may say, in short, that caer names may be found under c or g, and that there may be doubtful identities necessitating a see also.

The word moel is an adjective meaning ‘bald’, but it is used also as a noun, often feminine, meaning ‘bare
mountain'. When the noun is regarded as feminine, the initial _m_ becomes _f_ (= English _v_); thus we have _Y Foel_ for a hill near Clynnog, Gwynedd. But thereafter consistency is hard to find. Almost all the _foel_ names have lost their article; they usually activate a 'soft' mutation in the following word, but curiously do not necessarily revert to _moel_ before a proper or common noun. There are many _moel_ names, some of which activate a soft mutation e.g., _Moel Ddu_ (_du_ = black), while others do not, e.g., _Moel Tywysog_ (_tywysog_ = prince). The question of how to find the authoritative version will be answered later, but even then it has to be said that practice regarding _moel_ and _foel_ is so variable that some cross-reference, in general or in particular, is inevitable.

Explaining the whole system of Welsh mutations would be a lengthy task, and fortunately it is not necessary for the present purpose. Suffice it to say that there are three categories: 'soft', 'nasal' and 'aspirate', depending on the sound involved. Various syntactical situations can bring changes about, and the learner can find tables in which they are set out. The indexer dealing with Welsh place-names in a predominantly English text will need to know about the 'soft' mutation, but only as it affects certain feminine nouns of topographical significance.

**Topographical names affected by mutation**

First, there are various words for rock, stone and heaps of stone (hardly surprising in a mountainous country): _carreg_ and _craig_, which may appear in their mutated forms of _garreg_ and _graig_, and _carn_ and _carnedd_, which may appear as _garn_ and _garnedd_ respectively. Again, there is some variability over the inclusion of the definite article _y_ with the mutated forms. _Carn_ and _carnedd_ usually refer to prehistoric burial places, but by extension may apply to the mountain on which a cairn stands. Again, there may be radical and mutated forms in use for the same place-name, while a particular place-name may be duplicated elsewhere. The only safe criterion of the identity of variants is a map reference.

Thus _Garnedd Wen_ (= white cairn) will be the same as _Carnedd Wen_ if it can be shown that both refer to the site at SJ 0035, near Llandrillo, Merioneth. These 'rock' names are often found in the plural: _cerrig_, _creigiau_, _carnau_ and _carneddli_, or _carneddau_, are plurals of _carreg_, _craig_, _carn_ and _carnedd_ respectively. These plural forms tend not to be used with _y_, but even if they were, there would be no mutation. Note that a word of similar appearance and meaning, _crug_ (plural _crugiau_), is masculine, and therefore will not mutate in the singular after _y_.

Other frequently-occurring elements in place-names are: _croes_ (= cross), _melin_ (= mill), _pont_ (= bridge), and _tre_ (= town or settlement). Because they are feminine, they become _groes_, _felin_, _bont_, and _dre_ respectively after the definite article _y_. The _f_ in _felin_ is pronounced _v_, as in English, and is found written _v_ in older Welsh orthography.

Place-names starting with _pont_ show an overwhelming preference for the unmutated form, even when linguistic logic would seem to demand otherwise. Thus we find _Pontnewydd_ (Gwent), meaning 'Newbridge'; we might have expected _Y Bontnewydd_ (n does not mutate). Near Caernarfon is found _Bontnewydd_, and near Llanaber in Merioneth is found _Y Bont-du_, meaning 'Blackbridge', where both observe the rule. On the other hand, where dependent nouns follow, the _p_ remains unmutated, e.g., _Pontsenni_, 'Bridge on the Senni', and _Pont-y-gwaith_, 'Worksbridge', as already demonstrated under _caer_.

One further mutation of a somewhat different kind remains to be considered, viz. where an initial _g_ disappears completely when its noun is preceded by the definite article. This does not occur often. A simple example is _Y Lasynys_; the unmutated form would be _Glasynys_, 'Blue, or Green, Island'. Less obvious is _Yr Orsedd_, representing _gorsedd_ (= tumulus), preceded by the article _y_, which here, as always before a vowel, takes an intrusive _r_ for euphony, becoming _yr_. The Anglicized form of this place-name is Rossett, a village in Clwyd.

So far we have met five versions of the soft mutation. There are altogether nine, but the remaining four are less likely to concern the indexer. An initial _f_ can be the mutated sound of _b_ as well as _m_ (the latter change has already been discussed under _moel_). For example, the word for a birchtree is _bedwen_; when preceded by the definite article it becomes _y fedwen_. In its collective or plural form, _bedw_, it occurs in place-names over much of Wales, but curiously it is treated as feminine, singular in most instances, giving rise to _Fedw Fawr_, 'Much Birch', in several counties. Again, the article has been lost. However, there are few _bedw_ forms, and these are distinct from the _fedw_ names. In a similar case, the word for a peak is _ban_, found in Brecknock as _bannau_ in the plural (= beacons) and as _Fan_ (or _Van_) in its mutated form, but without its article. But the indexer will not be troubled by a _ban_ alternative.

The letter _d_ at the beginning of feminine nouns can change to _dd_ (this is a voiced _th_ sound). A word meaning meadow, _dol_, may be found on its own as _Ddół_, but, more usually, unmutated in the plural, Dolau, or in compound, _Ddöl-y-bont_. Again, this is of academic rather than practical interest. The remaining mutations of the soft variety, _ll_ into _l_, and _rh_ into _r_, need not detain the indexer. Neither _ll_ nor _rh_ mutate after the definite article _y_.

**The Welsh Alphabet**

(28 letters)

A, B, C, Ch, D, Dd, E, F, Ff, G, Ng, H, I, L, Li, M, N, O, P, Ph, R, Rh, S, T, Th, U, W, Y.

The letter _J_ is used in borrowed words.
The vowels in Welsh are a, e, i, o, u, w and y (i and w are not found in English). On the other hand, there are is called for. There are certain sounds in Welsh which Welsh orthography

characters, or 'digraphs', are considered letters in no distinctive typographical characters in Welsh to c and / respectively (so indexers of older publications themselves; these are: ch, dd, ff, ng, ll, ph, rh, and th. The vowels in Welsh are a, e, i, o, u, w and y (i and w may also serve as consonants). The letters k and v were once frequently used, but have now been superseded by c and f respectively (so indexers of older publications need to be vigilant). The letters z, q and x are not part of the Welsh alphabet, and the letter j is used only in borrowed words. The remaining letters of the English alphabet are found also in the Welsh. Note that the values of the Welsh f and ff are exactly as in the English ‘of’ and ‘off’ respectively. Welsh is a phonetic language; its pronunciation and spelling are straightforward once the basic rules have been learned, and it does not have such a host of anomalies as English.

Before proceeding to the second part of this article, which deals with the ordering of Welsh place-names within a given letter, it is necessary to make some general remarks about their orthography. It must be admitted that the situation is still confusing, even to the Welsh themselves. Much has been done to achieve a consensus among Welsh-users, but there remain difficulties in Anglicized contexts. Ever since Welsh place-names have been reproduced in written form they have been subject to distortion and misinterpretation. To a great extent this is because they were often first recorded for official purposes in Latin, Norman-French or English, and by scribes who were not conversant with Welsh. This process continued in the early years of the Ordnance Survey, which was a formative period in the dissemination of standard forms of place-names. There is now liaison between the Ordnance Survey and Welsh official bodies, but it is not difficult to find numerous inconsistencies in the recording of Welsh place-names on Ordnance maps. Those who wish to trace the history of this matter may do so in a recent article in Archaeologia Cambrensis (1982) by J. B. Harley and Gwyn Walters.

Gazetteers

Not until 1958 was there an authoritative list of the principal Welsh place-names to guide current practice. This was prepared by the Language and Literature Committee of the Board of Celtic Studies of the University of Wales, and was published as A Gazetteer of Welsh Place-names, edited by Elwyn Davies. The list was based on the place-names to appear in the sixth edition of the Ordnance Survey one-inch maps relating to Wales, and it was intended to record the best Welsh forms of the place-names, without implying that they were the only ones or that they should generally supersede English alternatives. But the promoters did hope that where places had only Welsh names, any 'debased or incorrect versions' would be superseded by the Board's recommended forms, even in English writing.

In 1968 another and larger compilation of Welsh place-names was made, also based on the one-inch Ordnance Survey map, but recording the forms actually used, whether in accord with Welsh academic opinion or not. This important work of documentation was carried out by Gwyn Ellis of the Botany Department of the National Museum of Wales at Cardiff, but unfortunately was published in typewritten form and received only limited circulation. It was produced county by county (the pre-1974 dispensation), and, like the Gazetteer, gave four-figure map references to the place-names.

Eleven years after the publication of the Gazetteer, the Board of Celtic Studies sponsored another, longer list of place-names, this time of a historical nature, entitled Welsh Administrative and Territorial Units, Medieval and Modern, by Melville Richards (Cardiff, 1969). This was based on the orthographic principles enunciated in the Gazetteer, but contained also many obsolete and variant forms to guide the historical researcher.

These compilations were made during a period which saw a growing desire among writers for consistency in the spelling of Welsh place-names, and pressure from various quarters for the official use of Welsh in road signs. The latter meant the display of the distinctive Welsh form of a name where it differed from the English form (which had hitherto appeared alone on the signs). These Welsh forms were not academic inventions, but words that had been in daily currency among Welsh-users for centuries.

Bilingual and Welsh pairs

Since these alternatives may well give rise to cross-references in indexing, it is worth providing some details. There are three kinds of bilingual pairs. First, there are those of separate derivation, stemming from divergent historical tradition, e.g.: Cardigan/Aberteifi; Fishguard/Abergwaun; Brecon/Aberhonddu; Swansea/Abertawe; Newport (Gwent)/Casnewydd; Newport (Dyfed)/Trefdraeth; Montgomery/Trefaldwyn; St Asaph/Llanfyllin; Chirk/Y Waun; Holyhead/Caergybi (the Caer Gybi quoted earlier referred to an antiquity, a
Note that some Welsh place-names require the definite article (cf. Le Havre in French).

Secondly, there are pairs which express exactly the same meaning in both languages, e.g. Newtown (Powys)/Y Drenewydd; New Quay (Dyfed)/Ceinewydd; Golden Grove (Dyfed)/Gelli-aur.

Thirdly, there are pairs of related names. Some are parallel descendants in Welsh and English from a common, but different ancestral form, e.g. Penfro/Pembroke and Caerdydd/Cardiff. Other pairs consist of a pure Welsh form and an anglicised adaptation, e.g. Llanilltud Fawr/Llanwit Major (in the Vale of Glamorgan); Glandwr (near Swansea)/Landore; Llanbedr/Lampeter (there are two in Dyfed, distinguished by suffixes). Both Tenby (Dyfed) and Denbigh (Clwyd) are derived from the Welsh Dinbych, ‘Little Fort’; in Welsh the southern is differentiated from the northern by a suffix, giving Dinbych-y-pysgod (pysgod = fish, for Roman fort). Note that some Welsh place-names require the accent clearly on the last. This fact is now indicated in writing by inserting a hyphen between the penultimate and ultimate syllable; e.g., Bryn-mawr, this time a village; and Cefn-brith, a farm, as opposed to Cefn Brith, a mountain.

Problems of indexing

In any indexing task the first decision has to be whether to use a letter-by-letter or a word-by-word sequence. In a simple index comprising only geographical names, a letter-by-letter sequence is the most straightforward, since the separation of elements in a name by hyphens or spaces is irrelevant for this purpose, and a single sequence can be used for all names. This is the method of The Reader's Digest complete atlas of the British Isles (London, 1965); there, Cwmogwr, Cwm Owen, Cwm-parc and Cwmpenygraig all follow in the same single sequence, according to the English alphabet. The Gazetteer used the same method, but in Welsh, not English, alphabetical order. A point to remember when consulting a Welsh index is that the ‘digraphs’ (double letters), mentioned above, will count separately in the alphabetical order. The letters ch, dd, ff, ng, ll, rh and th follow c, d, f, g, l, r, and t respectively. Thus Aberdyfi would come before Aberddawan, and Bedwas before Beddgelert. As an initial letter, th will only be found in words assimilated from other languages.

However, if an index is to be a general one, referring to objects, topics and persons as well as place-names, the choice is more likely to be for a word-by-word sequence. This was done in the two volumes of the Index to Archaeologia Cambrensis published to date,* and in An Index to the transactions of the Carmarthenshire Antiquarian Society 1905–1977. Certain difficulties had to be faced. First there was the wide variety of spellings used over the years for many of the place-names. Some standard had to be adopted for the main entry and all variants cross-referred to it. If the Gazetteer form were adopted for the main entry, the consequence of the separation of elements, whether by a space or a hyphen, would necessitate one sequence for the separated names and another for the ‘ uninterrupted’ compounds.

How does this work out in practice? Let us take the commonest of all Welsh toponymies—the llan compounds. The word llan can denote a sacred enclosure, then, by extension, a church within it, and finally an associated settlement. It normally appears combined with the name of a saint, usually, though not always, a Celtic saint with local associations; hence Llanddewi, a church dedicated to Dewi (= David). Since llan is

feminine, the following noun or adjective is mutated, i.e., Dewi becomes Ddewi. In the Gazetteer there are 557 llan guide-words, some referring to more than one location; they cover fifteen pages. The second volume of the Index to Archaeologia Cambrensis also contains fifteen pages of llan names, this time in double columns. Convenience and common-sense suggest that it would be desirable to have all llan names in a single sequence. This was effected in the Gazetteer by means of a letter-by-letter system of indexing, rather than a word-by-word system. Thus it made no difference whether place-names had hyphens or not.

The same llan name may occur in different parts of Wales, and, curiously, with the stress on a different syllable. This is quite easy to indicate, as in the examples Llan-bedr and Llanbedr, which are two different places, one in Gwent and the other in Merioneth, the former with the accent on the last syllable, the latter on the first.

In a word-by-word system all hyphenated forms should be dealt with first, as if they were separate words, viz. Llan-bedr, Llan-crwy, Llan-dawg, Llan-faes, Llan-fair, Llan-fyrn, Llan-gain, Llan-gan, etc., etc. Then a new sequence would begin: Llanaber, Llanaelhaearn, Llanafan, Llanallgo, Llanandras, Llananno, Llanarmon, Llanarth, etc., etc. Should this be necessary, a suitable warning to the user would be helpful.

Curiously enough, in the Index to Archaeologia Cambrensis 1901-1960 (published well after the recommendations of the Gazetteer had been disseminated), it was possible to get almost all llan names into one sequence, on the grounds that they had not been hyphenated in their original chapter and verse. The convenience to the user was judged to outweigh the academic correctness of current orthography. What will happen in the next volume has yet to be decided!

There are numerous other toponymic elements which can give rise to the above situation, but the separation of sequences is not so serious for the user, since the entries are far fewer than those for llan, and thus can be perceived by rapid inspection. The most common elements which can give rise to hyphenated names are: aber (= mouth or confluence of rivers), alt (= steep hill) blaen (= summit or upper reach), bryn (= hill), caer (= fort), cefn (= ridge), coed (= wood), cwm (= valley), dôl (= meadow), gelli (= grove), glan (= bank), glyn (= valley), hen (= old), llwyn (= grove), maes (= field), nant (= brook), pant (= hollow), pen (= top), pont (= village), pont (= bridge), pwll (= pool), rhos (= moor), rhyd (= ford), tal (= end), tre (= town or settlement), tŷ (= house), waun (= moor) and ystrad (= vale). A few examples will suffice (accents have this time been inserted as reminders of the pronunciation): Aber-soch, Aberysychan, Bryntefi, Bryn-y-grôes, Cwmfelin, Cwm-pen-grâig, Maes y Gâer, Maesgânedd, Porthcôr and Porth Sgâdan.

Finally an exception: a circumflex has the effect of putting an accent on the final syllable, even without a preceding hyphen, e.g. Aberdar, Cwmbrân. The introduction to the Gazetteer declares ‘It is almost vain to seek rigid uniformity in a list such as this’, so the user is warned to expect anomalies.

Summary

After an exposition of such detail, it is worth summarizing the main points:

1) The initial letter of certain Welsh place-names may change or ‘mutate’ according to set rules when a definite article is placed (or assumed to exist) in front, and some place-names have come to possess two forms, obliging the indexer to make a choice for the main entry.

2) Some Welsh place-names have parallel versions of quite different appearance, which may need cross-references.

3) A letter-by-letter method of indexing will provide a single, straightforward sequence, despite the various ways of writing Welsh place-names, whether with hyphens or gaps, or as continuous words. This is true for the English and for the Welsh alphabet.

4) The orthography and make-up of Welsh place-names is nowadays likely to reflect the nature of the place concerned, as well as the way in which the name is pronounced. These characteristics may run counter to the indexer’s logic.

5) A word-by-word method, appropriate for mixed indexes, will often create two sequences of similar names, depending on whether their elements are separately written or not. This may mislead a user unfamiliar with Welsh orthographic conventions, and some rubrics may need to be inserted as warnings.

The whole field is increasingly attracting both scholarly and popular attention, and various related topics suggest themselves for further investigation—the etymology of the place-names, the difference between apparently synonymous elements, the role of the Ordnance Survey in fixing the form of place-names, and the presentation of bilingual road signs.

The question of the orthography of the Welsh language generally has not been discussed, though it naturally has an intimate connection with that of place-names. Again, there have been numerous attempts to improve existing practice and set agreed standards. The vital prerequisite for effective action was an organization to present a consensus of scholarly opinion for general acceptance, and this came into being in the Board of Celtic Studies of the University of Wales in 1919. It published its recommendations on general orthography in 1928, and these now hold the field. References to this and other topics raised above will be found in the bibliography which follows.

Bibliography


A list of Welsh place names (Welsh counties), in typescript, compiled by Gwyn Ellis (Cardiff: National Museum of Wales, 1968).


Index to Archaeologia Cambrensis 1901-1960, compiled by T. Rowland Powel, with lists and notes by Donald Moore (Cardiff: Cambrian Archaeological Association, 1976).


Yr Atlas Cymraeg (The Welsh-language Atlas), George Philip & Son Ltd. in collaboration with the Welsh Joint Education Committee, forthcoming.


The author wishes to express his thanks to the following individuals who kindly offered helpful comments on the draft of this article: Elwyn Davies, MA, MSc, PhD, LID; Professor R. Geraint Gruffydd, BA, PhD; Patricia Moore, BA, FSA; W. Gwyn Thomas, MA, FSA.

Commander Donald Moore, RD, BA, Med. FSA, RNR (Retd.), is a Past President of the Cambrian Archaeological Association; he has been Convenor of the Association’s Index Sub-committee for more than thirty years, and has seen through the press two volumes of the Index to Archaeologia Cambrensis. He is now supervising the compilation of a third volume, intended to cover the years 1961 to 1980. He has served on the Ordnance Survey Users Committee and the Welsh Atlas Panel of the Welsh Joint Education Committee. He retired in 1984 from the post of Keeper of Pictures and Maps at the National Library of Wales, Aberystwyth.

---

Apotheosis of the indexer

‘The undoubted literary champion of the cause of indexing’, we dubbed Bernard Levin in our last issue (The Indexer 14 (4), 279), and again he wields his mighty pen splendidly on our behalf in The Times of 23 Dec. 1985, under the heading, ‘Enter the lists for this noble minority’. In a full article he addresses himself to questions of the difficulty of indexing (‘it is an appalling and prolonged labour (no wonder Hercules published his autobiography without one)’); the qualities demanded (‘a very high level of technical skill’); the SI Register; rates of payment (‘the truly shocking level of payment that this very remarkable and responsible work commands’); publishers’ attitudes (‘it is high time general publishers took the view that the work of the indexer is as essential as that of the jacket-designer or for that matter the printer and the supplier of paper, and carried the cost’); quality and absence of indexes (‘I have read many books of outstanding quality, interest and value which have been seriously and irretrievably damaged by an inadequate index’). In conclusion, ‘This is a plea for an admirable profession, equipped with real skills, to be accorded both the respect and the reward that it deserves . . .’

Thank you, Mr. Levin, for gladdening our hearts with such a Christmas present.

Conferences to come

The next Cranfield International Conference, on Mechanized Information Transfer, will be held from Tuesday, 22 July to Friday, 28 July 1986, and will include a special session sponsored by the British Library.

The general theme of the conference will be research and developments affecting all aspects of the mechanized transfer of information from the originator to the end user. The areas expected to be covered include database creation; storage and dissemination; retrieval; influences on information flow. Details are available from Cranfield Conference Secretariat, The Institution of Electrical Engineers, Savoy Place, London, WC2R OBL.

The second International Expert Systems conference and exhibition will be held 30 Sept.—2 Oct. 1986 at Bloomsbury Crest Hotel, London, focussing on the practical applications of expert systems. Details from Expert Systems Conference Secretary, Learned Information Ltd, Besselsleigh Road, Abingdon, OX13 6LG, UK.