



O Maolalaigh, R. (2016) DASG: Digital Archive of Scottish Gaelic /
Dachaigh airson Stòras na Gàidhlig. Scottish Gaelic Studies.

This is the author's final accepted version.

There may be differences between this version and the published version.
You are advised to consult the publisher's version if you wish to cite from
it.

<http://eprints.gla.ac.uk/116766/>

Deposited on: 30 November 2016

DASG: Digital Archive of Scottish Gaelic / Dachaigh airson Stòras na Gàidhlig

1. Introduction

The *Digital Archive of Scottish Gaelic / Dachaigh airson Stòras na Gàidhlig* (DASG) website resource was officially launched at the University of Glasgow on Wednesday, 29 November 2014. DASG, a British Academy recognised project since 2007, is an online repository of digitised Gaelic texts and other lexical resources for Scottish Gaelic. It can be accessed free of charge at www.dasg.ac.uk. It currently contains two main components, *Corpas na Gàidhlig* and *Faclan bhon t-Sluagh*. Together, these two resources provide fingertip access to the riches of Gaelic language, literature and culture, and make them more accessible to a world-wide audience.

Corpas na Gàidhlig aims to provide a comprehensive searchable electronic full-text corpus of Scottish Gaelic texts for students and researchers of Scottish Gaelic language, literature and culture. It aims to bring together printed and unpublished texts dating from the twelfth century to the present day. The online Corpus currently contains almost 20 million words from printed texts, and thus represents the largest online corpus available for the study of Scottish Gaelic.¹ Texts will continue to be added to the Corpus, including, in due course, transcriptions of spoken Gaelic.

The first 30 million words of *Corpas na Gàidhlig* will also provide the textual basis for the interuniversity project, *Faclair na Gàidhlig* ('Dictionary of the Scottish Gaelic Language'), which will be edited on the same principles as the *Oxford English Dictionary* and the *Dictionary of the Older Scottish Tongue*, both of which provide a historical lexical reference for their respective languages. When the dictionary has completed its first pass through the alphabet, DASG will be used to keep it updated. The dictionary, in turn, will future-proof DASG by providing a linked explanatory resource. Partners in the *Faclair na Gàidhlig* project are the Universities of Glasgow, Aberdeen, Edinburgh, Strathclyde and Sabhal Mòr Ostaig UHI. For further information about *Faclair na*

Gàidhlig, see Mackie and Pike (2010), Gillies and Pike (2012: 253–59), Pike and Ó Maolalaigh (2013: 299–313) and www.faclair.ac.uk.

Faclan bhon t-Sluagh (lit. 'words from the people') is a lexical resource based on the Fieldwork Archive of the *Historical Dictionary of Scottish Gaelic* (HDSG), which is described briefly in section 5 below. These vernacular materials consist of questionnaires, wordlists, and words and phrases transcribed from speech and sound recordings, collected throughout Gaelic Scotland and in Nova Scotia between the 1960s and 1980s. They uniquely describe traditional Gaelic life and society, and many of the headwords are accompanied by magnificent hand-drawn illustrations; see figures 3 and 4 below.

The DASG website also has a social media dimension, having its own blog and links to Facebook and Twitter.² The DASG blog has a 'Word of the Week' feature ('Facal na Seachdain') which highlights interesting words from *Faclan bhon t-Sluagh*, *Corpas na Gàidhlig* and other sources. It is currently organised by Shelagh Campbell and Shona Masson.

The DASG website also hosts a new project, *Seanchas* ('the Global Gaelic Jukebox'), a collaborative research project that aims to facilitate access to ethnographic Irish and Scottish Gaelic sound and film recordings, held internationally at a variety of archives and institutions.³ The project aims to provide a new online directory of Gaelic recordings, particularly those pertaining to the folklore and ethnology of Gaelic communities in Scotland, Ireland and elsewhere. It is directed by Dr Sìon Innes, Lecturer in Celtic and Gaelic at the University of Glasgow, and Dr Barbara Hillers, Lecturer in Irish Folklore at University College Dublin. The *Seanchas* project receives funding from Colmille, a partnership between Bòrd na Gàidhlig and Foras na Gaeilge, and the University of Glasgow.

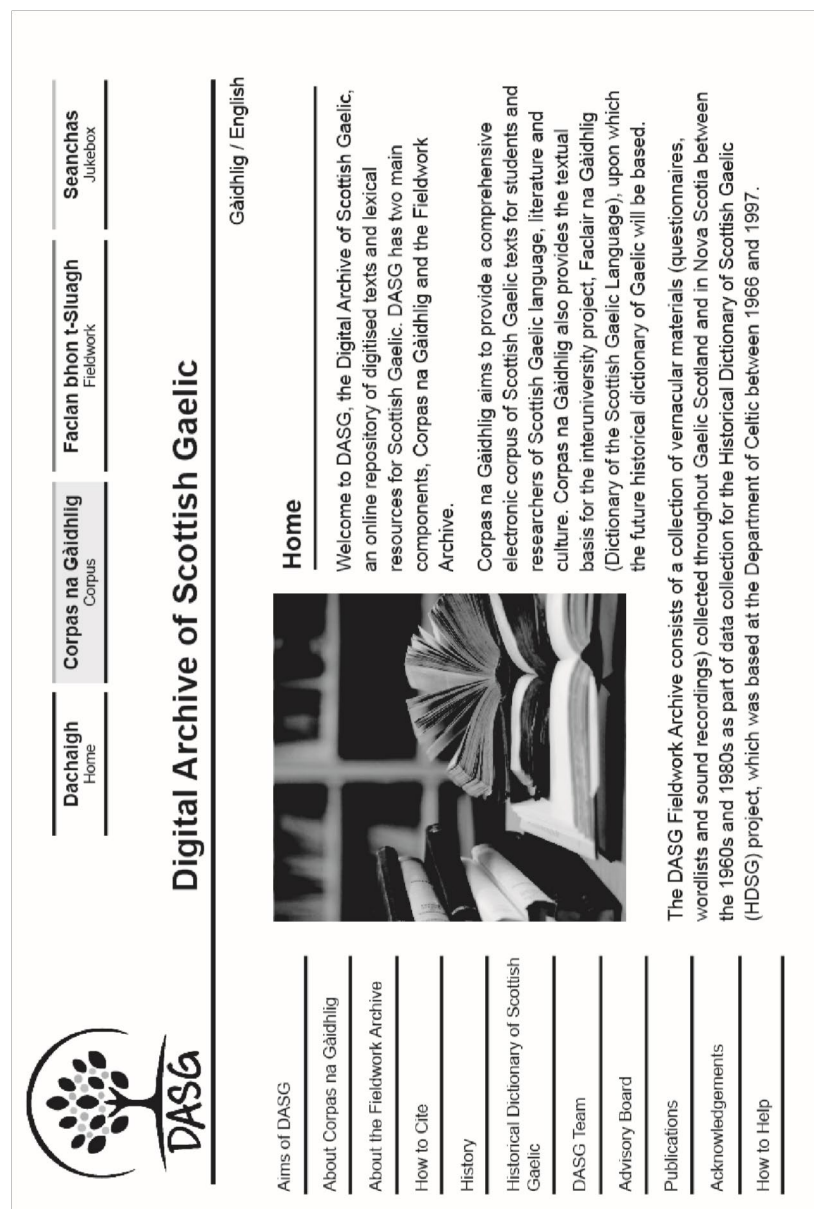


Figure 1: The DASG homepage

2. DASG: Background

The *Digital Archive of Scottish Gaelic* (DASG) project was established by Roibeard Ó Maolalaigh in 2006 and is a recognised British Academy project based within Celtic and Gaelic in the School of Humanities / Sgoil nan Daonnachdan at the University of Glasgow. It was founded initially in order to digitise valuable parts of the archive generated by the *Historical Dictionary of Scottish Gaelic* (HDSG) project, which was established within the then Department of Celtic at the University of Glasgow in 1966. The first stage of work concentrated on digitising the Fieldwork Archive associated with HDSG. This task was undertaken between 2008 and 2014 by Olga Szczesnowicz, who captured the data in Word and ASCII files. These in turn were transferred to XML by Wojtek Dziejma, Stephen Barrett and Dr Mark McConville. The sound recordings associated with the Fieldwork Archive (described further below in sections 4 and 6) were digitised by Uist Digitisation Centre in 2011. It is intended to transcribe these recordings and make them available via the DASG website in due course.

Corpas na Gàidhlig was established by Roibeard Ó Maolalaigh in 2008 as a constituent project of DASG. It was founded in order to create a comprehensive electronic corpus of Scottish Gaelic texts (a) for students and researchers of Scottish Gaelic language, literature and culture; (b) for the purposes of providing the textual foundation for *Faclair na Gàidhlig*, and (c) as a resource which would facilitate corpus planning and corpus development technology for Gaelic.

DASG has been funded directly and indirectly by: the University of Glasgow; The British Academy; the Robert Leith Thomson Endowment, University of Glasgow; Urras Brosnachadh na Gàidhlig / Gaelic Language Promotion Trust; Comunn na Gàidhlig; and the Scottish Funding Council; Bòrd na Gàidhlig; the Arts and Humanities Research Council (AHRC) and the Economic and Social Research Council (ESRC) through *Faclair na Gàidhlig*.

3. Corpas na Gàidhlig

The first phase of *Corpas na Gàidhlig* began by digitising over 200 texts from all periods of Gaelic literature and to include a wide variety of genres, including poetry, prose, song and folklore. These texts have been prioritised in order to provide the main textual basis for the interuniversity dictionary project, *Faclair*

na Gàidhlig. It is envisaged, as *Corpas na Gàidhlig* progresses, that a broad range of other texts will be added, and in time, that speech and video materials will also be represented by both text and sound and audio-video files. In the long term, the Corpus and other DASG resources will be used to update the *Faclair na Gàidhlig* dictionary. For further information about *Corpas na Gàidhlig*, see Ó Maolalaigh (2013a: 115–21) and Pike and Ó Maolalaigh (2013: 313–20).

The web interface for *Corpas na Gàidhlig* has been developed by Stephen Barrett with much valuable input from Dr Mark McConville and other members of the DASG team. *Corpas na Gàidhlig* is powered by CQPweb, a web-based graphical user interface for IMS Open Corpus Workbench (CWB), which in turn is a collection of open-source tools for managing and querying large text corpora ranging from 10 million to 2 billion words (OCWB). Corpus Workbench is a powerful research tool which facilitates simple and advanced searches and the manipulation of large sets of data.

The current *Corpas na Gàidhlig* website allows for four main initial search options: standard query, restricted query, word lookup and frequency list. With the standard and restricted queries, users have the further options of including or excluding lenited forms, accented vowels, upper case letters and the number of results displaying in results' pages. The results of queries are listed in consecutive lines within the immediate context in which the search word(s) occur(s). Further context can be viewed by clicking on the highlighted search word. Metadata for the textual source from which examples are cited is available by clicking on the source text number (the 'filename' in figure 2). This metadata was compiled by Dr Catriona Mackie during 2005–08 as part of the foundation project for *Faclair na Gàidhlig*, based at the University of Edinburgh and funded by the Leverhulme Trust (Gillies and Pike 2012: 258). Query results can be ordered and reordered in a variety of ways and can be downloaded with typical settings for use in Word, Excel, FileMaker Pro and other programs. Lists of collocations can also be generated and downloaded. Restricted queries can be performed according to any one, or combination, of date of language, geographical origins, register and / or short title of textual source. The word lookup option allows for a variety of sophisticated searches of word forms or partial word forms. For instance, it is possible to search word forms ending with, beginning with or containing any string of letters. The usual wild-cards of Simple Query language can also be used (see Hoffman et al. 2008). The final option, the frequency list, allows users to display the frequency of words or certain word forms. It is also possible to generate a word frequency list for all word forms occurring in the Corpus. Choosing this

option illustrates that the top twenty most frequently occurring words in the Corpus are: *a, an, air, agus, na, e, do, gu, tha, ann, am, iad, mi, is, le, bha, mar, cha, sin, nach*, the vast majority of which are closed set function words; the most commonly occurring lexical word is the noun *duine*. It is important to note that the Corpus is not yet tagged, and it is therefore impossible to distinguish between different lexical and grammatical categories, e.g. *an* (definite article), *an* (preposition), *an* (possessive pronoun), *an* (relative particle).

Corpus na Gàidhlig has the power to transform our understanding of linguistic patterns in Scottish Gaelic which were previously opaque to us.⁴ As well as facilitating the new historical dictionary of Scottish Gaelic, this resource enables us for the first time to contemplate and envisage the possibility of providing intermediate and advanced grammars of Scottish Gaelic language. The power of *Corpus na Gàidhlig* can be illustrated with two examples. It has long been known that there is variation between *fios* and *fhios* in the common Scottish Gaelic construction *tha / chan eil f(h)ios aig Calum* ('Calum knows / doesn't know') but it has not been possible to say when or how each variant is used. The Corpus shows us for the first time that the lenited variant *fhios* occurs far more frequently after *eil / bheil* than it does following any other verbal forms. This provides new valuable information about the modern Scottish Gaelic language. However, this pattern may also have highly significant implications for our understanding of the historical development of this construction in Irish, Manx as well as Scottish Gaelic (Ó Maolalaigh 2013b).

Similarly, an examination of the variation between *b-* and the absence of *b-* before vowels following the distributive adjective *a b-uile* ('every') shows that *b-* is used proportionally more frequently with the time noun *oidhche* than with any other noun beginning with a vowel, i.e. *a b-uile b-oidhche* ('every night'). This in turn suggests that the non-leniting aspect of *a b-uile* in Scottish Gaelic (which contrasts with leniting *chuile* and *dy chooilley* in Irish and Manx respectively) derives from an older genitive of time used adverbially, i.e. that Scottish Gaelic *a b-uile bliadhna* and *a b-uile b-oidhche* derive ultimately from *cacha b-uile bliadhna* and *cacha b-uile b-aidche*, which may in turn partially account for the Scottish Gaelic nominative form *bliadhna* (earlier nominative *blíadain*) and the Scottish Gaelic, Irish and Manx forms *oidhche*, *oíche* and *oie* respectively (earlier nominative *adaig*); see Ó Maolalaigh (2013c) for further details. These examples suffice to illustrate that *Corpas na Gàidhlig* has the power to elucidate patterns not just in the modern language but also in earlier forms of the language.

4. Faclan bhon t-Sluagh

Faclan bhon t-Sluagh (lit. ‘words from the people’) is a resource based on the Fieldwork Archive generated by *Historical Dictionary of Scottish Gaelic* (HDSG) project. It consists of materials gathered from four types of paper and orally recorded materials dating from the 1960s, 1970s and 1980s:

- thematic questionnaires
- wordlists
- cassette and reel-to-reel tapes of recorded speech
- excerpts of tape-recorded interviews
- papers slips and transcriptions based on fieldwork materials and recordings

There are 25 different types of thematic questionnaires, each relating to different lexical domains in traditional Gaelic society. Their titles and the numbers of filled in returns in each category (with a total of 194) are given in the following table:

Questionnaire type	Gaelic (G) / English (E)	Number
Peat Working / Mòine	G+E	23
Cattle / Crodh	E	17
Ecclesiastical Terms / An Eaglais	E	15
Sheep / Caoraich	E	14
Shellfish / Maorach	G	14
Agriculture / Àiteach	G+E	11
Wool Working / Obair na Clòimhe	E	10
Fishing Tackle / Acainn Iasgaich	G	9
House and Furnishings / Taigh Gàidhealach	G	9
Personal Appearance / Coltas an Duine	G	8
Lobster fishing / Iasgaich a’ Ghiomaich (E)	E	8
Personality / Nàdur an Duine	G	8
Recreation: Toys, Games, Contests / Curseachadan: E Dèideagan, Geamaichean is Farpaisean	E	8
Food and Drink / Biadh is Deoch	G	7
Weather / Side	E	7
Death and Burial / Bàs is Adhlacadh	G	6

Landscape Features / Cruth na Tìre	E	5
Senses / Faireachdain	G	4
Names and Uses of Medicinal Plants / Blàthan- Leighis	E	3
Boat-Building / Togail Bhàtaichean	G+E	3
Boats / Eathraichean	G	2
Herring fishing / Iasgach an Sgadain (E)	E	2
Piping / Pìobaireachd	G	1
Joinery / Saoirsinneachd (G)	G	0
Masonry / Clachaireachd (G) ^s	G	0
TOTAL		194

Table 1: Thematic questionnaires in HDSG / DASG Archive

In addition to these 25 questionnaire-types, there is a large number of wordlists and tape recordings, which both amplify and complement the questionnaire materials. The semantic domains in the wordlists include the following:

agriculture, beasties, birds, boats, carts, cattle, clothing, creatures, cures, deer, domestic articles, drawings and explanation, earmarks (on sheep), farm or croftwork, female personal names, fishing, nets, fishing-tackle, flowers, forestry work, grammar, literature, human body, disease, human nature, knitting, land usage and apportionment, landscape, line fishing, thatching, Norse mill, peat working, piping, place names, plants, plough, proverbs & expressions, riddles, sea, seashore, seaweed, sheep, sheepdogs, shellfish, survivals in Scots, wild flowers, wool working.

The words are often placed in context by the provision of phrases, sentences, idioms and proverbs, which are all vital in allowing us to gain a fuller understanding of individual lexical items. Some of the fieldworkers provided phonetic transcriptions using IPA symbols. Some of the documents also contain spectacular hand-drawn illustrations, which serve to further illustrate visually the semantics of individual words; see figures 3 and 4.

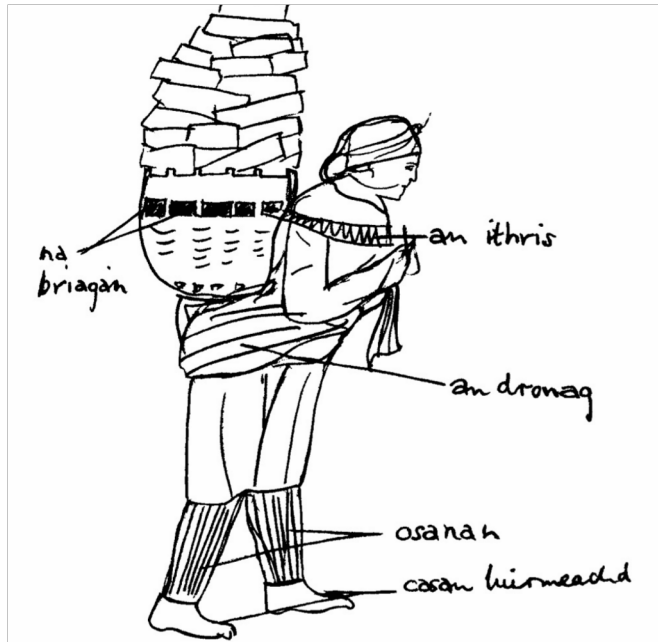


Figure 2: *Faclan bhon t-Sluagh*: Taking peats home (Point, Lewis)

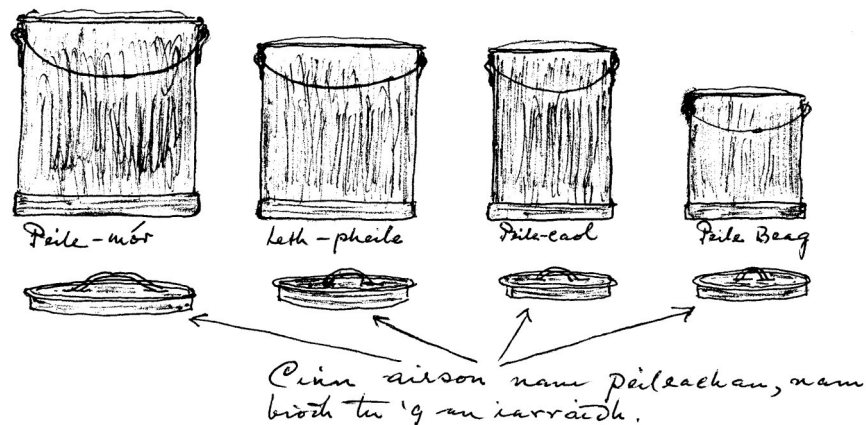


Figure 3: *Faclan bhon t-Sluagh*: peilichean (Uig, Lewis)

These materials can be accessed and viewed in two ways: (a) by searching for individual words or groups of words (either in Gaelic or English); and (b) by

browsing the documents themselves.

These rich materials provide a valuable window on traditional Gaelic society and life. Numerous words and senses are found that do not appear in standard Gaelic dictionaries. Some words, however, merit further investigation and need to be treated with caution. Two examples should suffice to illustrate the point. In materials collected from a Barra informant in 1988 by Ailig O'Henley we find two collocations containing the same generic noun element *foras*, which means 'basis, foundation': *foras feasa* with the meaning 'encyclopedia', and *foras focail* meaning 'etymology'. Neither of these noun phrases are found in modern Scottish Gaelic dictionaries, including Dwelly's *The Illustrated Gaelic-English Dictionary* ([1901–11] 1977) or, as far as I am aware, in the nineteenth-century dictionaries which preceded it and upon which Dwelly drew.

For those with an historical knowledge of Gaelic language and literature, these examples are quite arresting as they have notable resonances stretching back to the seventeenth and fourteenth centuries respectively. The phrase *foras feasa* is found in the title of Dr Seathrún Céitinn's (c. 1569 – c. 1644) early seventeenth-century chronological history of Ireland, *Foras Feasa ar Éirinn* (Keating [1634] 1902–14), which has been described as 'the most influential of all works of Gaelic historiography' (Welch 1996: 202) and which was completed c. 1634. *Foras Feasa ar Éirinn* (lit. 'a foundation of knowledge about Ireland') is certainly encyclopaedic in its range and coverage, and one can see how the sense 'encyclopedia' might easily develop, especially within a culture in which the work was well known and respected, and manuscript copies were in wide circulation.⁶ Knowing these earlier connotations makes *foras feasa* infinitely more preferable to *leabhar-eòlais* or *leabhar mòr-eòlais* as a modern Gaelic equivalent for 'encyclopedia'.

Foras focal (lit. 'foundation of words') occurs in the first line of the metrical glossary, 'Foras focal luaidhtear libh', which sets out to explain difficult Gaelic words and which is said to have been composed by the Irish poet, Seán Ó Dubhagáin (d. 1372), in the 14th century; see Russell (1988: 8) and references therein. One of the copies of this glossary survives in a manuscript made by Eoghan Mac Ghilleoin (Hugh or Ewen MacLean), schoolmaster at Kilchenzie, for Lachlan Campbell (Lochlan Caimbel) on the third of October 1698 in Campeltown, Kintyre. This manuscript, now in Trinity College Dublin (TCD ms. 1307 (H 2.12 (6))), was sent to Edward Lhuyd by Campbell as a gift in 1702 (see Campbell and Thomson 1963: 10; Sharp 2013: 246). *Foras focal* is in fact also explicitly referred to in a letter from Lachlan Campbell to Lhuyd, dated

16 April 1705 (Sharp 2013: 261). For more information on Lochlan Campbell (1675–1707), who was educated at the University of Glasgow and who is said to have been tutor in the household of Archibald Campbell (1658–1703), 1st Duke of Argyll, see Sharp (2013: 244–46 *et passim*).

Foras focal occurs with the meaning ‘etymology’ in Lhuyd’s Irish-English dictionary (Lhuyd [1707] 1971: s.v. *foras* ‘A law; also a foundation’), no doubt deriving from the earlier metrical glossary whose first line contains these words. It also occurs in William Shaw’s *Galic and English Dictionary* with the meaning ‘etymologicon’ (Shaw 1780: s.vv. *foras*, *focal*; etymology), doubtless from Lhuyd (1707), but does not otherwise occur in Scottish Gaelic dictionaries.⁷ Could it be that these two collocations are genuine survivals in the vernacular tradition? If so, as well as providing further important evidence for the permeable nature of the boundary between literary, scholarly and vernacular Gaelic traditions (cf. Gillies 2011; Ó Maolalaigh 2013c: 85), it would also reaffirm the extraordinary value of the HDSG fieldwork materials. However, the form *foca(i)l*, with *o* in the Barra form, is perhaps suspicious since *faca(i)l* with *a* might be expected for Barra and most modern Scottish Gaelic dialects (Borgström [1937]: 121, §169);⁸ the *o* here is strongly suggestive of literary and / or high register provenance.⁹ Similarly, the genitive form *feasa* is unusual for vernacular Gaelic. As it happens, both forms, *foras feasa* ‘a general or fundamental account, **an encyclopaedia**, a history’ and *foras focal* ‘an **etymological** dictionary’ (emphasis mine) occur in the Rev. Fr Patrick S. Dinneen’s *Foclóir Gaedhilge agus Béarla / An Irish-English Dictionary*, which was first published in 1927 (Dinneen [1927] 1953: 479, s.v. *foras*). Although *foras focail* (‘etymology’) is not an exact match here, it could be that the Barra forms derive ultimately from Dinneen’s dictionary or from interaction between Irish and Scottish Gaelic speakers with an interest in Gaelic scholarship, hence the need for caution in assessing some of the word forms in the Fieldwork Archive.¹⁰

Finally, it should be noted that there is a great deal of variation in the spelling of words in the Fieldwork Archive. These have not yet been standardised although it is intended that standardised forms for headwords will be supplied in due course. It is also hoped in the future to offer a means of searching for variant spellings automatically. For an edition of some Tíree questionnaires relating to the senses, cattle, sheep, personal appearance and the weather, which illustrate the value of the HDSG fieldwork materials, see Ó Maolalaigh (2008).

5. Historical Dictionary of Scottish Gaelic (HDSG)

The history and background to the *Historical Dictionary of Scottish Gaelic* (HDSG) project deserves a detailed treatment; only a brief outline is offered here. The HDSG project was established in the Department of Celtic, University of Glasgow in the year 1966 by Professor Derick S. Thomson with Kenneth MacDonald as Editor. Thomson described the task and progress of the dictionary project in 1969 as follows:¹¹

The aim is to excerpt exhaustively from all Gaelic printed books, from MSS., collections of private papers, newspaper and B.B.C. files, archives of sound recordings, and current speech in all the Gaelic areas, both in Scotland and in Nova Scotia. Printed books up to and including the year 1801, in which the final part of the Gaelic translation of the Old Testament appeared, will be completely excerpted. Thereafter excerpting will be selective to some extent, but there will be extensive checking to prevent later first occurrences of words slipping through the net. Questionnaires on a large range of topics are being drawn up, circulated to interested persons, and followed up in the field. This work has already begun, pilot questionnaires (e.g. on peat-cutting, seaweed, household utensils and furniture etc.) having been tried out in schools over the Gaelic area, and by individual collectors. Recordings of many kinds of conversation will be made, so that not only the vocabulary, but also the syntax of current Gaelic, will be illustrated. The collection will also include place-names and other proper names, and although many of these will not appear in the printed Dictionary, an index of them will be available to scholars. Part at least of the excerpting of early texts will be computerized, and it is hoped in this way to provide, in due course, a valuable tool for the future editors of these texts, and for researchers into such topics as word frequency and stylistics. The first computer outputs have already come in, and there will be a steady flow of these from now on.

The Editor of the Dictionary is Mr Kenneth MacDonald, who is a native Gaelic speaker and a graduate of Glasgow and Cambridge. To date about 130 voluntary readers have been recruited, and it is hoped to double this number eventually. All of the Celtic Department’s staff of seven are actively engaged, in one way or another, in the project. A

fulltime field-worker [i.e. Angus John Smith] was appointed in September 1967, and after a period of training he has now been working in the field for some time, surveying to date the southwestern, western and central districts (Arran, Kintyre, Knapdale, Mid-Argyll, Cowal, Appin, Lochaber and Perthshire). He is at present working in the Western Isles. A part-time collector [i.e. Duncan MacQuarrie] has been at work during the summer of 1968 in Mull and Morvern. Another is surveying collections of books and MSS. in Argyll, and a third is marking texts for subsequent excerpting.

An index of printed Gaelic books is being compiled, and an index of Gaelic MSS. and other archives. A considerable number of books (including a good many rare ones), and some MSS., have been gifted to the Department in connection with the work on the Dictionary, and we are in the process of building up a first-rate collection of Gaelic materials for research purposes.

Completed slips from the voluntary readers have been coming in steadily since November 1966. At the last count, approximately 100,000 slips has been received. Eventually the collection of slips will probably amount to roughly 1½ million. Private collections of rare vocabulary are coming in, and many more are expected. Though it is rash to make estimates now, it is anticipated that the final tally of head-words in the printed Dictionary will exceed 100,000. (Thomson 1969: 280–81; cf. Thomson 1967)

It was estimated that this ambitious and complex project – an early example of crowd-sourcing – would take a mere ‘twelve years’ to complete (Thomson 1969: 280).¹² However, given the available resources and competing demands on staff time in a busy teaching department, the project was never brought to completion. This is not altogether unsurprising given that the *Dictionary of the Older Scottish Tongue* took over 80 years to be compiled (Dareau 2012).¹³ With the retirement of Departmental staff associated with the dictionary, the project was effectively suspended in 1996–97 leaving in its wake a unique and unparalleled resource for Scottish Gaelic students and researchers. The broad aims of HDSG are now encompassed within the new interuniversity project, *Faclair na Gàidhlig*, whose main objective is to produce the much needed historical dictionary.

Under the direction of Professor Cathair Ó Dochartaigh, Anja Gunderloch was employed during 1997–98 to assess the remaining archive; in the absence of adequate documentation of work carried out by and for HDSG, she produced a number of valuable investigative reports from which I have partially drawn in the following section.

6. The HDSG / DASG Archive

The HDSG (now DASG) Archive consists of:

- the paper slip Drawer Archive¹⁴
- thematic questionnaires
- wordlists
- cassette and reel-to-reel tapes of recorded speech
- transcriptions of excerpts of tape-recorded interviews

The Drawer Archive contains an estimated 550,000 paper slips, of which there are four types: (a) slips that were excerpted from printed sources (the majority of slips); (b) slips that were excerpted from manuscript sources; (c) computer generated slips; and (d) slips derived from fieldwork and vernacular oral materials.

Some 312 key published works (including book, journal, magazine and newspaper titles) were mined, and of these a total of 216 were comprehensively excerpted onto dictionary slips. This excerpting from books was conducted by a group of more than 100 voluntary collaborators from Scotland, England, Ireland and Canada, the most prolific of which were John MacLean (Oban), the Rev. Robert Leith Thomson and Alison Fergusson; project staff also participated in excerpting.

Excerpting from manuscripts seems to have been carried out mostly, if not solely, by Ronald Black during 1973–77, who concentrated on Classical Gaelic manuscripts. It is interesting to note from Thomson’s 1967 and 1969 account of progress with the dictionary project that computerisation was already conceived of and apparently underway by 1969.¹⁵ Computerisation work was carried on by Donald Meek, who was Assistant Editor between 1973 and 1979, with the assistance and support of Dr William Sharp. Richard Cox was also involved in the computerisation of texts at various points during the 1980s and 1990s.

The main HDSG fieldworkers were Angus John Smith (1967–c. 1973),

Angus Martin (c. 1974–c.1978), Duncan MacQuarrie (1968), Duncan MacLaren (1974–75), Susan Norwood (1969), Alan Boyd (1981), Alick O’Henley (1987–88) and Ian Quick (1983). The regional coverage of this activity and extant materials may be briefly summarised as including: Argyllshire, Arran, Assynt, Barra, Benbecula, Eriskay, Gigha, Harris, Islay, Kintyre, Lewis, Lochaber, Morvern, Mull, Nova Scotia, Perthshire, Raasay, Ross-shire, Skye, Sutherland, Tiree, North Uist, South Uist.

The Archive has a total of 86 tape recordings (31 reels and 55 cassettes) directly related to the fieldwork collections of the dictionary. More than half of these were made by Angus Martin and relate mostly to Southwest Argyll, particularly Kintyre. A total of 22 recordings was made in Cape Breton, 20 of these by Angus John Smith, who spent the Summer months of 1971 in Cape Breton as a fieldworker, and a further two recordings by Sister Margaret Beaton. In addition to these tapes there are 29 reels recorded in Nova Scotia. These belong to the Major Calum I. N. Macleod Bequest which was donated to the Department of Celtic at the University of Glasgow following the Major’s death in 1977. All of these recordings have been digitised and it is hoped at a later stage to transcribe them and make them available via the DASG website.

HDSG was funded by the Carnegie Trust for the Universities of Scotland, The British Academy, the Modern Humanities Research Association, the Highlands and Islands Development Board and a number of other bodies including the Pilgrim Trust. For more information about HDSG, see Thomson (1967; 1969), Macdonald ([1987] 1994: 62), Ó Maolalaigh (2008: 474–75), Mackie and Pike (2010: 101), Gillies and Pike (2012: 236–59, esp. 251–53).¹⁶

7. DASG: A Research Resource

DASG resources have underpinned or informed a number of lexical and grammatical studies, e.g. domestic furnishings and utensils (Quick ([1988]); *smùid* (Meek 2005a); *sgaoil* (Meek 2005b); *suacan* / *suachdan* (Ó Maolalaigh 2005); *iorram* (Ó Maolalaigh 2006); Tiree lexis and lexicology (Ó Maolalaigh 2008); Gaelic words for ‘snowflakes’ (Ó Maolalaigh 2010); *mearan*, *mearanach*, *dàsachdach*, *dàsan(n)ach* (Ó Maolalaigh 2014); possessive constructions (Bell 2012); Gaelic numerals ‘three’ to ‘ten’ (Ó Maolalaigh 2013a); *fios* ~ *fhios* variation (Ó Maolalaigh 2013b); *a b-uile* (Ó Maolalaigh 2013c).¹⁷ We are keen for researchers to use and acknowledge use of the DASG research resources; we intend to list

publications utilising DASG on our website.

8. How to Help

DASG is interested in collecting and digitising as wide a variety of textual genres as possible. We are particularly interested in receiving digitised copies of Scottish Gaelic texts and / or permissions to digitise further texts. We are also interested in gathering further lexical data for Scottish Gaelic. If readers have texts or data that they would like to offer to us, please contact us at mail@dasg.ac.uk and tell us more about the texts (e.g. who wrote them, are they literary texts, letters, spoken texts, broadcasts, etc). If we can use texts and we have the capacity to add them to the Corpus, we will provide forms for completion, including a Copyright form for us to obtain the required permission, and Author and Text forms to collect more detailed information.

Acknowledgements

I am grateful to Olga Szczesnowicz, Stephen Barrett and Lorna Pike for their comments. I am also grateful to the editors of *Scottish Gaelic Studies* for including this brief background paper as a means of raising awareness among the scholarly community about the new DASG website and the resources it contains.

Endnotes

- 1 For a recent overview of existing Gaelic textual corpora and resources, see Ó Maolalaigh (2013a: 113–15).
- 2 The Facebook page may be accessed at <https://www.facebook.com/DasgGlaschu>, Twitter at @DASG_Glaschu and the blog at <http://dasg.ac.uk/blog/gd> (Gaelic) and <http://dasg.ac.uk/blog/en> (English).
- 3 For a recent description of Celtic folklore collections in Harvard libraries, see Innes and Hillers (2011).
- 4 On the use of a corpus for a very successful investigation of register variation in Scottish Gaelic, see Lamb (2008).
- 5 These last two questionnaire types are extant only as sample questionnaires; no filled-in examples exist.
- 6 For an edition of *Foras Feasa ar Éirinn* (FFÉ), see Keating ([1634] 1902–

- 14). A copy dating to 1647 of *FFÉ* is housed in the National Library of Scotland (NLS Adv. ms. 72.1.43 and continued in Adv. ms. 72.2.1). Another copy of *FFÉ* is found in NLS Adv. ms. 72.2.8, from which the Lochaber born Aberdeen librarian and classicist, Ewen MacLachlan (1773–1822), transcribed extracts, which survive in NLS ms. 72.3.5 (*Leabhar Caol*). See MacKinnon (1912: 122, 126, 128, 258).
- 7 'Forus fios air Erin. Notitia Hibernia, the name of an Irish book' also occurs in Shaw (1780: s.v.) but not with the meaning 'encyclopedia'.
- 8 However, the existence of doublets reflecting vernacular and literary forms respectively is not unknown, e.g. *iutharn(a) ~ ifrinn* (MacInnes 1977: 42); *lighiche ~ lèigh* (Gillies 2004: 254).
- 9 *Focal*, the historical form, is the norm in Irish and also occurs in some northern Gaelic dialects, e.g. East Sutherland (Dorian 1978: 90), although it is not clear whether the *o* in such cases represents a secondary development from *a < o*.
- 10 Jackson's decision not to collect lexical data, especially by uncontrolled means, as part of the Gaelic linguistic survey was due to his experience of one potential informant admitting that he had used an English-Gaelic dictionary to fill in a postal questionnaire (Ó Dochartaigh 1994–97, i: 66; cf. p. 35).
- 11 Thomson (1969) represents an updated and slightly expanded version of a note which had previously appeared in a Scottish number of *Forum for Modern Language Studies*, published by the University of St. Andrews (Thomson 1967). I am grateful to Honor Hania, College Librarian for Arts and Social Sciences at the University of Glasgow for helping me to trace Thomson (1967).
- 12 cf. *The Glasgow Herald* (29 September 1966), p. 15; *The Glasgow Herald* (26 April 1967), p. 10. In 1971, there was talk of the dictionary taking 'eight more years to complete' (*The Glasgow Herald* (5 July 1971), p. 3).
- 13 The long-term nature of dictionary compilation was later acknowledged by Professor Thomson, who is quoted in a newspaper article, published in 1988, as saying: 'There are some sanguine university administrators who seem to think that dictionaries can be completed in five years, or at a push, seven, but they take decades.' (*The Glasgow Herald* (18 February 1988), p. 14).
- 14 The paper slip Archive has been crucial to the progress of *Faclair na Gàidhlig* as it has provided the data for the editorial foundation and training materials for future Gaelic lexicographers. Thomson (1967: 286) refers to plans for computerising early texts; Thomson (1969: 281) refers to 'the first computer outputs [having] already come in'.
- 15 See also *The Glasgow Herald* (29 September 1966), pp. 8, 15; *The Glasgow Herald* (26 April 1967), p. 10; *The Glasgow Herald* (5 July 1971), p. 3; *The Glasgow Herald* (3 March 1972), p. 10; *The Glasgow Herald* (2 July 1972), p. 6; *The Glasgow Herald* (18 February 1988), p. 14; *The Glasgow Herald* (24 February 1988), p. 10. I am grateful to Dr Aonghas MacCoinnich who gathered these references as part of the University of Glasgow *Sgeul na Gàidhlig* project.

- 16 *Corpas na Gàidhlig* has other, initially unexpected, uses. It can be used, for instance, to find place-names with generic place-name elements, e.g. *Cill* + specific. DASG therefore complements resources such as *Saints in Scottish Place-names*, which was launched in Celtic and Gaelic at the University of Glasgow on 4 November 2014, and which is available at <<http://saintsplaces.gla.ac.uk/index.php>>.

Bibliography

- BELL, Susan (2013). 'An Seilbheach Gàidhlig: Rannsachadh le Corpas', in *Rannsachadh na Gàidhlig* 6, ed. by Nancy McGuire & Colm Ó Baoill. Obar Dheathain: An Clò Gàidhealach, 253–73.
- BORGSTRØM, Carl Hj. ([1937]). *The Dialect of Barra in the Outer Hebrides*. Oslo: [Aschehoug].
- CAMPBELL, J. L. and THOMSON, Derick (1963). *Edward Lhuyd in the Scottish Highlands 1699–1700*. Oxford: Clarendon Press.
- DAREAU, Margaret G. (2012). 'Dictionary of the Older Scottish Tongue', in *Scotland in Definition: A History of Scottish Dictionaries*, ed. by Iseabail Macleod and J. Derrick McClure. Edinburgh: John Donald, 116–43.
- DASG: Digital Archive of Scottish Gaelic / Dachaigh airson Stòras na Gàidhlig. <www.dasg.ac.uk> [accessed 6 November 2014].
- DINNEEN, Patrick S. ([1927] 1953). *Foclóir Gaedhilge agus Béarla / An Irish English Dictionary*. Dublin: Irish Texts Society.
- DORIAN, Nancy (1978). *East Sutherland Gaelic*. Dublin: Dublin Institute for Advanced Studies.
- DWELLY, Edward ([1901–11] 1977). *The Illustrated Gaelic-English dictionary*. Glasgow: Gairm.
- FACLAR NA GÀIDHLIG. <<http://www.faclair.ac.uk>> [accessed 13 November 2014].
- GILLIES, William (2004). Varia II: Early Irish *lia*, Scottish Gaelic *liutha*, *Ériu* 54, 253–56.
- GILLIES, William (2011). 'The Gaelic of Niall Mac Mhuirich', *Transactions of the Gaelic Society of Inverness* 65, 69–95.
- GILLIES, William and PIKE, Lorna (2012). 'The Last Hundred Years', in *Scotland in Definition: A History of Scottish Dictionaries*, ed. by Iseabail Macleod and J. Derrick McClure. Edinburgh: John Donald, 236–59.
- HOFFMAN, Sebastian; EVERT, Stefan; SMITH, Nicholas; LEE, David and BERGLUND Prytz, Ylva (2008). *Corpus Linguistics with BNCweb: A Practical Guide*. English Corpus Linguistics, Vol. 6. Frankfurt: Peter Lang.
- INNES, Sìon and HILLERS, Barbara (2011). 'A Mixed-media Folklore Trove: Celtic Folklore Collections in Harvard Libraries', in *Proceedings of the*

- Harvard Celtic Colloquium 31, ed. by Deborah Furchtgott, Matthew Holmberg, A. Joseph McMullen, Natasha Sumner. Cambridge, MA, USA: Harvard University Press, 173–208.
- KEATING, Geoffrey ([1634] 1902–14). *Foras Feasa ar Éirinn / The History of Ireland*, 4 vols, ed. by David Comyn (vol. i (1902)), Patrick S. Dinneen (vols. ii, iii (1908); vol. iv (1914)). London: Irish Texts Society.
- LAMB, William (2008). *Scottish Gaelic Speech and Writing: Register Variation in an Endangered Language*. Belfast: Cló Ollscoil na Banríona.
- LHUYD, Edward ([1707] 1971). *Archaeologia Britannica*. Dublin: Irish University Press.
- MACDONALD, Kenneth D. [1987] 1994. ‘Dictionaries, Scottish Gaelic’, in *The Companion to Gaelic Scotland*, ed. by Derick S. Thomson. Oxford: Blackwell, 61–63.
- MACINNES, John (1977). ‘Some Gaelic Words and Usages’, *Transactions of the Gaelic Society of Inverness* 49 (1974–76), 428–55.
- MACKIE, Catriona and PIKE, Lorna (2010). ‘Faclair na Gàidhlig: The Advent of an Historical Dictionary of Scottish Gaelic’, in *Rannsachadh na Gàidhlig: Fifth Scottish Gaelic Research Conference*, ed. by Kenneth E. Nilsen. Sydney, NS: Cape Breton University Press, 100–07.
- MACKINNON, Donald (1912). *A Descriptive Catalogue of Gaelic Manuscripts in the Advocates’ Library Edinburgh and Elsewhere in Scotland*. Edinburgh: W. Brown.
- MEEK, Donald E. (2005a). ‘Smoking, Drinking, Dancing and Singing on the High Seas: Steamships and the Uses of *smùid* in Scottish Gaelic’, *Scottish Language* 24, 46–70.
- MEEK, Donald (2005b). ‘The Spread of a Word: *scail* in Scots and *sgaoil* in Gaelic’, in *Perspectives on the Older Scottish Tongue: A Celebration of DOST*, ed. by Christian J. Kay and Margaret A. MacKay. Edinburgh: EUP, 84–111.
- OCWB. *The IMS Open Corpus Workbench*. <<http://cwb.sourceforge.net>> [accessed 9 November 2014].
- Ó DOCHARTAIGH, Cathair (ed.) (1994–97). *Survey of the Gaelic Dialects of Scotland*, 5 vols. Dublin: Dublin Institute for Advanced Studies.
- Ó MAOLALAIGH, Roibeard (2005). ‘A Gaulish-Gaelic Correspondence: *s(o)nxt-* and *suac(hd)an*’, *Ériu* 55, 103–17.
- Ó MAOLALAIGH, Roibeard (2006). ‘On the Possible Origins of Scottish Gaelic *iorram* “rowing song”’, in *Litreachas & Eachdraidh. Rannsachadh na Gàidhlig 2, Glaschu 2002 / Literature & History. Papers from the Second Conference of Scottish Gaelic Studies, Glasgow 2002*, ed. by Michel Byrne, Thomas O. Clancy and Sheila Kidd. Glasgow: Department of Celtic, 232–88.
- Ó MAOLALAIGH, Roibeard (2008). ‘“Bochanan Modhail Foghlaimte”: Tíree Gaelic. Lexicology and Glasgow’s Historical Dictionary of Scottish Gaelic’, *Scottish Gaelic Studies* 24, 473–523. Also available at: <http://eprints.gla.ac.uk/4830/1/4830.pdf>
- Ó MAOLALAIGH, Roibeard (2010). ‘Caochlaideachd Leicseachail agus “snowflakes” sa Ghàidhlig’, in *Cànan agus Cultar / Language and Culture*, ed. by Gillian Munro & Richard A. V. Cox. Edinburgh: Dunedin, 7–21.
- Ó MAOLALAIGH, Roibeard (2013a). ‘Corpas na Gàidhlig and the Numerals “three” to “ten”’, in *Language in Scotland: Corpus-based Studies*, ed. by Wendy Anderson. Scottish Cultural Review of Language and Literature series. Amsterdam: Rodopi, 113–42.
- Ó MAOLALAIGH, Roibeard (2013b). ‘Léargas Stairiúil ó Mhalartaíocht Shioncrónach i nGaeilge na hAlban: *tha fios agam ~ tha fhios agam*’, unpublished paper delivered at Teangeolaíocht na Gaeilge, University College Dublin (April 2013).
- Ó MAOLALAIGH, Roibeard (2013c). ‘Gaelic *gach uile* and the Genitive of Time’, *Éigse* 38, 41–93.
- Ó MAOLALAIGH, Roibeard (2014). ‘Am Buadhfhacal Meadhan-Aoiseach *Meranach* agus *mearan*, *mearanach*, *dàsachdach*, *dàsan(n)ach* na Gàidhlig’, *Scottish Studies* 37, 183–206.
- PIKE, Lorna and Ó MAOLALAIGH, Roibeard (2013). ‘Faclair na Gàidhlig and Corpas na Gàidhlig: New Approaches Make Sense’, in *After the Storm: Papers from the Forum for Research on the Languages of Scotland and Ulster Triennial Meeting, Aberdeen 2012*, ed. by Jenet Cruickshank and Robert McColl Millar. Aberdeen, 299–337. Also available at: http://www.abdn.ac.uk/pfrlsu/uploads/files/Pike%20O%20Maolalaigh_2.pdf
- QUICK, Ian ([1988]). ‘The Scots Element in the Gaelic Vocabulary of Domestic Furnishings and Utensils’, in *Gaelic and Scots in Harmony: Proceedings of the Second International Conference on the Languages of Scotland*, ed. by Derick S. Thomson. Glasgow: Department of Celtic, University of Glasgow, 36–42.
- RUSSELL, Paul (1988). ‘The Sounds of Silence: The Growth of Cormac’s Glossary’, *Cambridge Medieval Celtic Studies* 15 (Summer 1988), 1–30. *Saints in Scottish Place-names*. Glasgow. <<http://saintsplaces.gla.ac.uk/index.php>> [accessed on 9 November 2014].
- SHARP, Richard (2013). ‘Lachlan Campbell’s Letters to Edward Lhwyd, 1704–7’, *Scottish Gaelic Studies* 29, 244–81.
- SHAW, William (1780). *A Galic and English Dictionary*. London: W. and A. Strachan.
- THOMSON, Derick S. (1967). ‘The Historical Dictionary of Scottish Gaelic’, *Forum for Modern Language Studies* 3.3 [Scottish Number] (July 1967), 286–87.

THOMSON, Derick S. (1969). 'The Historical Dictionary of Scottish Gaelic',
Lochlann 4, 280–81.

WELCH, Robert (ed.) (1996). *The Oxford Companion to Irish Literature*. Oxford:
Clarendon Press.

University of Glasgow

Roibeard Ó Maolalaigh