Updating and Expanding a Retro-digitised Dictionary: Some Insights from the Dictionary of the Irish Language

Sharon Arbuthnot
Queen’s University, Belfast
Electronic Dictionary of the Irish Language (eDIL)
www.dil.ie
Webstatistics for 2017

### Monthly history

<table>
<thead>
<tr>
<th>Month</th>
<th>Unique visitors</th>
<th>Number of visits</th>
<th>Pages</th>
<th>Hits</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 2017</td>
<td>18,846</td>
<td>26,525</td>
<td>185,370</td>
<td>286,618</td>
<td>3.36 GB</td>
</tr>
<tr>
<td>Feb 2017</td>
<td>9,020</td>
<td>12,279</td>
<td>91,194</td>
<td>136,268</td>
<td>1.54 GB</td>
</tr>
<tr>
<td>Mar 2017</td>
<td>19,533</td>
<td>27,659</td>
<td>229,722</td>
<td>338,846</td>
<td>3.92 GB</td>
</tr>
<tr>
<td>Apr 2017</td>
<td>17,135</td>
<td>24,367</td>
<td>283,669</td>
<td>382,999</td>
<td>3.56 GB</td>
</tr>
<tr>
<td>May 2017</td>
<td>16,413</td>
<td>23,051</td>
<td>169,916</td>
<td>262,846</td>
<td>3.14 GB</td>
</tr>
<tr>
<td>Jun 2017</td>
<td>14,205</td>
<td>20,627</td>
<td>154,400</td>
<td>232,485</td>
<td>2.49 GB</td>
</tr>
<tr>
<td>Jul 2017</td>
<td>14,832</td>
<td>21,691</td>
<td>171,704</td>
<td>248,582</td>
<td>3.23 GB</td>
</tr>
<tr>
<td>Aug 2017</td>
<td>16,526</td>
<td>23,488</td>
<td>138,783</td>
<td>213,850</td>
<td>2.53 GB</td>
</tr>
<tr>
<td>Sep 2017</td>
<td>18,494</td>
<td>25,359</td>
<td>165,812</td>
<td>245,946</td>
<td>2.82 GB</td>
</tr>
<tr>
<td>Oct 2017</td>
<td>17,995</td>
<td>26,231</td>
<td>210,439</td>
<td>306,289</td>
<td>3.65 GB</td>
</tr>
<tr>
<td>Nov 2017</td>
<td>18,552</td>
<td>26,792</td>
<td>212,623</td>
<td>314,364</td>
<td>3.72 GB</td>
</tr>
<tr>
<td>Dec 2017</td>
<td>18,931</td>
<td>27,603</td>
<td>159,825</td>
<td>250,096</td>
<td>2.87 GB</td>
</tr>
<tr>
<td>Total</td>
<td>200,482</td>
<td>285,672</td>
<td>2,173,457</td>
<td>3,219,189</td>
<td>36.83 GB</td>
</tr>
</tbody>
</table>
# Sources of Visitors by Country

<table>
<thead>
<tr>
<th>Countries</th>
<th>Pages</th>
<th>Hits</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>99,654</td>
<td>113,600</td>
<td>710.45 MB</td>
</tr>
<tr>
<td>Ireland</td>
<td>21,786</td>
<td>30,443</td>
<td>268.05 MB</td>
</tr>
<tr>
<td>Great Britain</td>
<td>16,106</td>
<td>22,106</td>
<td>195.02 MB</td>
</tr>
<tr>
<td>Germany</td>
<td>6,986</td>
<td>12,105</td>
<td>156.54 MB</td>
</tr>
<tr>
<td>Netherlands</td>
<td>2,363</td>
<td>3,508</td>
<td>37.19 MB</td>
</tr>
<tr>
<td>South Korea</td>
<td>1,579</td>
<td>1,726</td>
<td>167.00 MB</td>
</tr>
<tr>
<td>Japan</td>
<td>1,476</td>
<td>1,895</td>
<td>49.35 MB</td>
</tr>
<tr>
<td>France</td>
<td>1,172</td>
<td>2,049</td>
<td>42.39 MB</td>
</tr>
<tr>
<td>Canada</td>
<td>1,140</td>
<td>2,075</td>
<td>37.85 MB</td>
</tr>
<tr>
<td>China</td>
<td>938</td>
<td>1,423</td>
<td>47.43 MB</td>
</tr>
<tr>
<td>Others</td>
<td>8269</td>
<td>15299</td>
<td>225.33 MB</td>
</tr>
</tbody>
</table>

*Full list*
The 23 original fascicles of the Dictionary of Irish Language and the Compact Edition (foreground)
Limitations of the Paper Dictionary

- Expensive
- Unwieldy
- Difficult to know where to look to find forms
- Required expertise in the language
- Definitions, translations, grammatical information, etc. not searchable or collatable
- Restricted user-base
Tagging in TEI-conformant XML
| 3 gáeth | n m. and f.(?) the sea, a stream, an estuary | gáethd i. fairge, Eg. Gl. 374. O'C. | gáeth i. mór, H.3.18, 468. gáeth a shallow stream into which the tide flows and which is fordable at low water. O'Don. Suppl. 'gáeth (f.) a stream left at low water,' Conelys. ros-gab eile darr gáeth nglass, Écrit vii 224 § 59. tar gáeth glás ba trom a trom, O'C. 1180 - i (H.3.18, 468), cf. Met. Dinds. i 36.30 n. Anm. dí Libba on gáeth grinn nodasailte, Rawl. 78a38 = on gáeth grind LL 135b84 (Trans. RIA 2 Ser. i 226). cor' bailead Aed ruad 'con gáeth, Met. Dinds. iv 8.35. andsom gobél gáeth a 'gáeth' is most troublesome in a strait, Tec. Com. § 17. is desin atá in gáeth glán | loc% maic Mairead Muman, Met. Dinds. iv 68.151. ro ba tarcasal mar... ar crudh-gheathbhaínl cabhchr, MR 288. ag gáeth an chorríngin the inlet of Carrigín (a place-name) FM i 122.13. See also MR 288. O'Don. Suppl. |
| 4 gáeth | n stealth (?): mní gáethóide i. mní do ní gáethóideachto at milteasann gan fios, s.v. gáeth i. táidhe, O'C. |
| gáethach | adj (2 gáeth) o.à windy: errach trim gáethach (gáethach LL), ZCP xi 82 § 25. Tec. Com. § 36. gabél gáethch, CRR 10.2. ón tsín gáethcháin, IGT Decl. ex. 646. indí fuara garba goethacha, FA 30 (LB) (IT i 190.31). Títhe 611. FM i 364.7. |
| gáethaire | n a vent: fion agach na bháthaire, Job xxxii 19. |
| gáethamlacht | n(2 gáethi) i. (med.) flatulence: innrbaídigh in gh, cí b'ís an bhais ghaill, O’Gr. Cat. 228.22. |
| gáethi | n(1 gáeth) i. cleverness: ‘Etáin... artriúil (i. rothu-laing) mór ngaibhthi’... la gáethi (i. la glicus) F., ZCP v 524.23. |
| gáethlachd | adj(gáethch) i. l. gáethochd (gl. paluster). Sg. 54b5. |
Aims Achieved by Digitisation

- Highly inflected and unpredictable forms now locatable
- Possible to search English definitions and translations
- Opened up to wider audiences
- Directly linkable from other internet sites
- Specialist filters for e.g. grammar and other languages
Issues Remaining after Digitisation

Structural Issues
• Citations not in chronological order
• Dead hyperlinks
• Extraneous material
• Annoying inconsistencies in labelling

Fundamental Issues
• Typos
• Misrepresentations of manuscript readings
• Misunderstandings of material & incorrect assumptions
• Incomplete
PACDIL: An experiment in crowd-sourcing

About PACDIL
Contents of this screen
1 Purpose of the web-site.
2 Copyright - important notice to users and contributors
3 Searching
4 Entering special characters
5 Making Contributions
6 Obtaining a Password
7 Logging In
8 Contributions are to Come from Published Sources Only
9 Making a New Contribution
10 Entering Details of the Published Source
11 Adding a New Source
12 Correcting a Previous Contribution
13 Partners in the PACDIL project
14 Other Sites of Interest

1 Purpose of the web-site
PACDIL is an automatic website, designed for the collection of published additions and corrections to the Royal Irish Academy's Dictionary of the Irish Language (DIL). Many established scholars of medieval Irish currently have hundreds (if not thousands) of such additions and corrections entered into their own personal copies of DIL. PACDIL is designed to pool those resources and make them available to a wider audience. It is intended that the website will become a living repository for published additions and corrections noted by scholars all around the world, each with their own particular specialisation in the field of medieval Irish linguistics.
Expanding and Correcting: Phase 1 (2007-2013)

- Scholarly work published since 1932
- c. 5,000 interventions
- New entries created
- Citations added
- Citations reassigned
- Errors deleted
- Mistranslations corrected
- Homonyms separated
Expanding and correcting: Phase 2 (2014-2019)

- Editions of texts published since 1932
- Editors’ notes found to be inadequate and sometimes unreliable
- Texts had to be read afresh
- Manuscripts consulted in cases of doubt
New Entries and the eDIL Working Database

New entries:
• Follow structure of existing entries
• Citations in chronological order
• Notable additions in the domains of medicine, wildlife, linguistic terminology

Retained in the database:
• Possible scribal errors
• Unassimilated loans
Loanwords beginning with ‘alien’ consonants placed under letters of the Irish alphabet in the original dictionary

eDIL will follow attested forms and create e.g. a section for words beginning with q-

Existing entry:

**cuartíd** n (Lat. quartus + Idus): i quartīd mārtā, ZCP xii 368.8 = ... qart-id ... `le quatrième jour des ides de mars', RC xlii 128.6.

New eDIL entry:

**quart-** num (Lat. quartus) **fourth**. Precedes and forms compd. with qualified noun: (a) **gramm.**.: for quart-choibeden (sic leg.) ... for cet-choibeden = quartae conjugationis, Lib. Hymn. 11 gl. 74.75. (b) **in dating:** i quartīd mārtā, ZCP xii 368.8. for quart-id a comdal 'their convention is on the fourth of the ides', Gorm. Nov. 10. hi quart decimm Kalaind Mái 'on the 14th day before the Kalends of May', PH 1399.
Headwords II: Spelling for Date

Headword standardised according to conventions of Old Irish (c. 600-900) but examples are 17th/18th century

EDIL aims to give headwords in accordance with the period of earliest attestation

Existing entry:

bélaites
n [u, m.] (bél + aites) oral tradition: seanchus
γ beuloideas Éireann, Keat. i 48.16. is follus cheana as bealoideas cháich a coitchinne from general tradition, 23 E 26, 286.15.

New EDIL entry:

béaloides
n [u, m.] (bél + aites) oral tradition: seanchus
γ beuloideas Éireann, Keat. i 48.16. is follus cheana as bealoideas cháich a coitchinne from general tradition, 23 E 26, 286.15.
<table>
<thead>
<tr>
<th>Entry</th>
<th>Entry List</th>
</tr>
</thead>
<tbody>
<tr>
<td>caillech</td>
<td></td>
</tr>
</tbody>
</table>

n. f. (1 caille)

(a 1) Lit. veiled woman (see Éigsie xxviii 74-85); in sense married woman: ba bec ind immirgini nucad ann i. vii. mbaí ταρο oçus g-sea-rein a caillech ταρο ingen Fing. R. 507 (Easanda Tige Bocel). Brigid bráthir a caillech som Scano 224, cen aile díb, cen Ína dois caillechta so, v.l) Cumman ταρο ταρο ταρο ταρο g-tárrach CGSH § 722. 6. In phr. caillech aiththre, penitent wife, a married woman who becomes celibate: inar find ... cacha calidí aiththre Cán Ad. 14 § 24. (a) nun, woman in religion; in ch. taided Currech (Brigí), Thes. 3 349. 3 (Hy. v). a. n. fail di chaillí i mchíd i. ... / Maire Ó SanctBrigí, 349. 11 (b) a. ch. Vita Br. 26. Brigid ... / cenn calidí c. n-Enn. Fí. Fé. 1. sencion calidídi 'senior nun' (opop. to mac-c.), Mon. Tull. 151. 7 (§ 62). clérig ταρο ταρο ταρο ταρο caillecha, Laws ill. 14. 12. do cleiríchib ταρο ταρο ταρο ταρο, Érfu xii 60 § 8. cell ch. c. LU 1651 (RC a 452). tri hairgarla ecalise: c. trí clocco, eoc. Triads 66. cenn c. abbess, Trip. 2. 1000. c. teach nábacht Cíle dara do ghabháil ... ταρο ταρο ταρο ταρο ταρο do breith a mbroid, ALG l 130. 22. c. leightinn lectrix, Trip. xii 15 (of. fer léiginn). In later lang. c. dub: pían na manach ... ταρο ταρο ταρο ταρο ταρο nubh. Aisil. Tund. 130 § 4. ag sathruigh-adh beannaighthe bharnabach ταρο ταρο ταρο ταρο ταρο ταρο ταρο, RSClára 118 § 16. D. ... do écc ina caillí chuidib, FM ill. 252. 6. ord onóirch caillech \textit{ndup the Order of S. Cúran}, Fi. Ears 164. 19.
Factual and Cultural Obsolescence

*Taprobâne* pn (Lat. Tāprobanē, Gr. Ταπροβάνη) *Ceylon:* i tîr thall Taprobâne, *Sits* 1912, 439 § 4 (see also p. 442). óta inis Taprofâni anair, airm i turgaib grian i n-airthiur in domain, *TTr.* 633.

écond adj o-á **wanting in sense, without reasoning power; mentally defective, irresponsible; as subst. one legally incapable; a minor, an imbecile** (see *ZCP* xiv 379 § 43 . 380 ) : dia athair nírbó chialla | ferg fír mac nóí certbliadna...dáig rofírtir ba hécond, *Met. Dinds.* iv 132. énirt cach n-écond *Arch.* iii 228.85 . écond cach fuíthi *ib.* 90.

*rebrad* o,m. (reb+ -rad); also u,m. (regarded as a vn. formation), *sport, diversion, frolicking* ...
(b) ra érig Hectoir ... ar a baethair baísi γ ar a luagaill rebartha, *LL* 230a16 ( *TTr.* 951 ). in dernais lanamnas rebartha ri mnaí aile?, 273b21 (of an unnatural union).