

Pádraig Ó Mianáin

The New English-Irish Dictionary

1. The project

The New English-Irish Dictionary (NEID) is the first major English-Irish Dictionary since 1959, when de Bhaldraithe's *English-Irish Dictionary* was published. The NEID is being produced and funded by Foras na Gaeilge, the cross-border body with responsibility for the promotion of the Irish language throughout the island of Ireland, with a project budget of €6m. The project started in 2000 and is scheduled to be completed at the end of 2015. The project is divided into three major phases:

Phase 1: Planning and design (began 2000, completed 2007). As the field of large-scale bilingual lexicography in Ireland had been dormant for such a long period, the project relied heavily on outside expertise in the initial stages. The planning and design phase was carried out by Lexicography Masterclass and delivered key elements such as a headword list, a draft style guide and, most importantly, the English and Irish corpora which would underpin the entire project.

Phase 2: Compilation, Writing and Editing (began 2008, to be completed 2015). This phase concerns the production of the final text of the dictionary, and was further divided into three subphases:

2A: The English-language content (began 2008, completed 2011). Lexicography Masterclass also supplied the English-language database on which the source language content of the dictionary is based. This is a lexicographical database of c. 60k headwords and 150k sense units derived from a systematic analysis of the project's 1.7 billion word English corpus. Each entry contains a detailed description of headwords in terms of senses, phrases, collocations, grammar, usage, register etc, with over 640,000 examples taken from the corpus as evidence. The final product, *Database of Analysed Texts of English* (the DANTE database) is a 16 million word, target-language neutral, lexicographical database a demo version of which can be accessed online at www.webdante.com. This is a unique resource for lexicographers working on monolingual or bilingual English dictionaries, linguists, researchers and teachers.

2B: The Irish-language content (began 2009, completed 2012). In the translation phase, Irish-language equivalents were provided for the English-language material. Apart from terminology, the Irish-language material was initially translated without recourse to existing reference sources, in order to ensure currency and originality in the translations. Also, the most frequent 1000 words in English were each translated by three translators, one from each of the main dialects. The translation team was recruited and managed directly by Foras na Gaeilge and was comprised of three full-time and a large number of part-time translators. In this phase nearly 4 million words in Irish were added to the English database to produce a 20-million-word bilingual corpus.

2C: Editing the final entries (began 2011, ongoing). In this phase the editing team first select the final English-language content for each entry and subsequently the equivalent Irish-language content. The editing team of three full-time editors and a number of part-time editors was directly recruited and managed by Foras na Gaeilge.

Phase 3: Publication (began 2011, ongoing). In this phase the final entries are prepared for publication, first online and ultimately in book form. The online and mobile versions of the dictionary, containing a third of the eventual content, were launched in January 2013 at www.focloir.ie and will be followed by an app version. The full dictionary will be online by the end of 2014 with a printed version to follow in 2015. NEID will eventually contain some 100,000 sense units (c. 40,000 headwords) and will be available in both electronic and book formats.

2. Practical challenges

The project to produce a modern bilingual English-Irish dictionary faced a number of significant practical challenges. Some of these challenges may be common to other minority languages, while others stemmed from the specific situation of Irish-language lexicography and in particular the difficulties arising from the shortage of available personnel with the appropriate skills and experience necessary for such a large-scale project at editorial and managerial level.

- **The absence of adequate electronic corpora of Irish.** Modern lexicography is largely corpus-based and relies heavily on the ability of computers to analyse huge digital corpora. Unfortunately such corpora did not exist in the Irish language, and although the project produced a new 30 million word Irish-language corpus, it falls far short in terms of size, content and representativeness when compared to the norm in other projects internationally, as exemplified in the 1.7 billion word corpus of English produced specifically for the NEID during Phase 1 of the project.
- **Qualified, experienced staff.** Finding qualified and/or experienced translators and editors has been a major obstacle throughout the course of the project. From a purely lexicographical point of view, there is a critical shortage of people who have the lexicographical expertise and the necessary knowledge of the Irish language; as mentioned previously, the last major English-Irish dictionary was published in the 1950s and the last Irish-English dictionary in the 1970s. Even from a translation point of view, while there are many qualified and skilled translators working in a professional capacity, the concentration on currency in both source and target languages of the current project required competency in registers not normally central to academic or professional activities and which are beyond the scope of what professional translators deal with on a daily basis.
- **An innovative approach to dictionary compilation.** The approach followed in the NEID project is groundbreaking in that the final content of the dictionary is entirely derived from a unique lexicographical database. A more traditional, straightforward approach would have been to start with a substantial English-language dictionary framework from a recent bilingual dictionary by a major international company, then

have a team of translators ‘fill the blanks’, so to speak, to provide the target language content; this model would also have required a much smaller editorial team. Working from the lexicographical database, however, required all translators and editors to exercise a higher level of judgement throughout, highlighting even more the lack of experienced staff available to the project. In the long term, however, the approach adopted should have major advantages for any future projects.

- **Technical challenges.** The technical/electronic working environment of modern lexicography posed a significant challenge, both at the organisational level dealing with various databases hosted externally and at individual level where a lot of training was required to enable the staff, most of whom were working remotely, to use the systems as efficiently as possible and as quickly as possible.
- **Management of staff and processes.** The varied nature of the project and the overlapping of the various phases and sub-phases required careful organisation and management, particularly as a lot of the processes in the translation and editing stages were being developed from scratch. There were a number of recruitment drives due to turnover of part-time personnel and to meet project deadlines, and there was also a significant amount of coordination involved with staff from other departments within Foras na Gaeilge who were seconded on a temporary basis in the period leading up to the launch.
- **Training and upskilling of staff.** The shortage of experienced people at all levels in the project required a significant and continuous programme of training and monitoring as the processes for each phase of the project were being implemented. Unfortunately, much of this training effort was expended on part-time staff who may not be available to work on future projects.

3. Linguistic challenges

The project also faced a number of challenges arising from the particular situation of the Irish language.

- **Lexical gaps as a minority language.** Bearing in mind that the last major English-Irish dictionary was in 1959, the emphasis in NEID on contemporary English, as in all modern bilingual dictionaries, meant that a large proportion of the Irish-language translations could not be sourced from existing sources and thus had to be produced first-hand. Also, as with many minority languages, many words and phrases that become current in the major language would often remain untranslated among native Irish speakers, and such material is often difficult to translate in a natural register that would not offend a native speaker's ear.
- **Official Standard(s) or Dialects.** There are three major dialects of modern Irish and many subdialects within them as well as dialects emerging outside the traditional Gaeltacht (Irish-speaking) areas. None of these dialects is universally accepted as standard, and although there has been an Official Standard of grammar and orthography since 1957, which is an attempt at a compromise between the dialects, it has never been fully accepted among the native speaker population. Given the nature of the translations required to capture the essence of the living language, there is at times a

tension and a divergence between dialect and Official Standard. Also there have been two separate revisions of the Official Standard since the writing phase of NEID started. The Revised Official Standard was published as recently as August 2012 without any involvement of senior project staff

- **‘Book-Irish’ or ‘Real Irish’**. As with any language where the standard norm is not accepted by the majority of its speakers, the prescriptive versus descriptive question is particularly pertinent to NEID. There is a constant debate within the language community over the validity of current language as it is generally spoken by native speakers, especially where high-level or ‘more correct’ usage is being replaced by common and simpler forms, or where the influence of the English language is being reflected more and more in day-to-day usage of Irish.

4. The aim of NEID

Ultimately the NEID project aims to produce a practical resource that will cater for a range of users from the native or highly proficient Irish-speaker to learners both within and without the education system. The increasingly democratic nature of modern dictionaries, where currency of language dictates coverage, is reflected in NEID in both the English-language content, as evidenced in the corpus analysis, and the Irish-language content, where priority is given to natural or ordinary expression over the more formal register that would have been the norm in previous Irish dictionaries. In summary, the stated aims of the NEID project are to provide:

- an accessible guide to self-expression in correct and contemporary Irish for learners; and
- an authoritative guide to current and best practice in the Irish language for native or fluent speakers.

5. Further information

Further information regarding the NEID project can be found at the following websites:

The New English-Irish Dictionary: www.focloir.ie

The Database of Analysed Text of English: www.webdante.com

Foras na Gaeilge: www.gaeilge.ie