



British-Irish Council 4th IML Languages Ministerial

Friday 10th March 2017, Edinburgh

**Report of the Promoting Our Languages Through Technology
Conference held in Dublin on 9-10th November 2015**

Ministers are invited to

- **note the report of the Conference in Dublin in November 2015;**
- **recognise that technology can be used to promote preservation, engagement and the use of IML languages across BIC Member Administrations;**
- **reflect on future requirements for further developing and fostering use of language technology; and**
- **agree to the publication of this report on the website of the British-Irish Council.**

INTRODUCTION

1. On 9-10 November 2015, the Irish Government hosted the British-Irish Council's Indigenous, Minority and Lesser-Used (IML) Languages conference on Promoting Our Languages Through Technology in Dublin.
2. The conference provided an opportunity for those working in both the areas of language promotion and in the field of language technology (LT), to demonstrate how technology can be used to promote preservation, engagement and the use of each of the IML languages across BIC Member Administrations. Delegates and speakers highlighted the current use of technology for their respective languages and also outlined future requirements in further developing and fostering the use of language technology. These contributions provide a basis for recommendations for the role of policy in supporting and developing IML languages through technology. The body of this paper summarises the main discussion points raised at the Conference and a more detailed note of presentations is set out in Annex 1 (see below).

BACKGROUND AND SUMMARY OF THE CONFERENCE

3. Administrations nominated speakers and invitees from their jurisdictions. In total, 86 delegates attended over the one and a half day event, with representation from each Member Administration, from the USA and Canada. The conference was officially opened by Mr. Joe McHugh TD, Minister of State for Gaeltacht Affairs from the Irish Government.
4. Day one (morning sessions) focused on reports on current research and development in the area of language technology (LT) for the relatively more highly-resourced languages in the group. Specific case-studies were presented as examples of successful approaches to LT in Welsh, Irish and also Scots Gaelic, and experiences from which other groups could benefit. The afternoon session focused on the less-resourced languages such as Jèrriais, Manx, Cornish, Ulster Scots and northern dialects of Irish. Each of these talks highlighted the important role technology plays in making language learning resources accessible to learners, and in bringing geographically dispersed linguistic communities together.

Day one also involved a workshop, dividing the delegates into groups of 3 to discuss (i) how IML language communities could work together and (ii) what can be done to make IML languages more attractive to multinational companies.

5. Day two focused on how technology can be used to normalise the use of IML languages in our respective societies, where English is the predominant and easy “default” language. Presentations reported on studies of Welsh speakers’ engagement with technology that provides linguistic options, and on a 5-year strategy to promote and normalise the use of Irish by RTÉ, the Irish National Television and Radio broadcaster. In addition, suggestions were made from a media production perspective for reaching the wider linguistic community through virtual online environments. Two case studies were also presented on the use of technology through both Irish and Scots Gaelic amongst young people.

COMMON THEMES ARISING FROM PRESENTATIONS

6. In her closing remarks, the rapporteur highlighted some of the common themes that had arisen through presentations and follow-up discussions during the event. The details of these points are discussed below.
7. A recurring theme, which arose from experiences shared by LT groups with past successes in research, was the strong need for collaboration. As highlighted by Prof. Kevin Scannell, due to the fact that there are such a high number of components required for technological advances for any language, it is too difficult for each language community to address them alone. This is based on a lack of resources in terms of both funding and skilled researchers, both of which are challenges for low-resourced languages. Going forward, therefore, collaboration will play a significant role in overcoming this lack of resources. Such collaboration can be realised through cross-language projects and even cross-research group projects within one language group. The involvement of indigenous SMEs can also bring value to such collaborative efforts. The need for joint funding proposals, as outlined by Delyth Prys, is therefore a practical approach to furthering work in this area.
8. A related issue is the need for the practice of developing open-source technology. If linguistic communities and LT researchers are to benefit from the research already carried out in this area, there is a need for LT resources to be developed under some type of open-source licensing model. This approach will avoid situations where funding is spent on “reinventing the wheel”. Open-source licences on software/tech can ease the integration process, and there are a number of licences that can be considered, where the developers’ contributions are protected and recognized. This open-source culture reinforces an approach whereby research and development arising from public funding should be for the benefit of the public.

9. One of the challenges shared across research groups to date is the lack of consistent, directed funding for minority language technology. Until now, research groups have received limited funds for short-term projects but it is clear that this kind of piecemeal approach is not enough. This model contributes to partly-researched or partly-developed technologies that rarely fully advance from pilot phase, and also to the loss of valuable human resources who will seek the financial security of long-term contracts. The work involved in this area of research requires such a niche set of skills that this type of loss can be hugely detrimental to any project. Therefore, there is a clear requirement for ongoing financial support in conjunction with a definitive strategy. In relation to this, there should also be ongoing government support for further education opportunities to equip new researchers with the appropriate skills to continue this work.
10. The very existence of an event such as this conference highlighted the importance of cross-language sharing. Throughout both days, delegates from each language group learned something valuable from the technological approaches taken by others working to promote their own language. Again, the open-source nature of some existing technologies allow for easy porting to another language, thus enabling a lower-resourced language to benefit from the skills and research of another better-resourced one. The sharing of information across language also highlighted the need for a dedicated forum for language promoters and technologists to exchange ideas. Suggestions were made to set up a social media interest group, a mailing list, and possibly a dedicated workshop similar to the Celtic Language Technology Workshop.
11. The need for tech industry support and engagement was also a commonly recurring theme throughout the event. The general consensus is that governments should consider imposing obligations of corporate social responsibility onto these large corporations to help our indigenous languages. There are a numbers of ways in which industry partnerships can help. A guest speaker from Google highlighted how the translation of Gmail into Irish, Welsh and Cherokee is just one way large multinationals can endorse and promote the use of IML languages. In addition, the provision of translated interfaces, localised installations, translated web content and even hosting language promotion events or hackathons are all examples of ways in which these companies can support and give validity to IML languages. This pressure of course should be extended to indigenous industries who can use their own technologies to promote their native language.
12. Following from point 10, community engagement in creating language resources

was also highlighted as a significant requirement in the future. Crowdsourcing efforts (e.g. for data annotation, translation, user testing) are invaluable for the development of LT in the context of minority languages. According to experiences of many delegates, such projects, when managed and monitored adequately, are successful in both reducing development time and also in creating a positive reinforcement for speakers who will see the important role they have played in promoting and preserving their own language.

ANNEX 1

NOTE OF CONFERENCE CONTRIBUTIONS

AT THE CUTTING EDGE: PERSPECTIVES FROM THE WORLD OF HIGH TECHNOLOGY

Laura Toby, Google Ireland, gave us an insight into Google's philosophy regarding their development of language technology. Their focus is on global coverage and enabling users to write and add content in their language. She also discussed some minority language initiatives at Google: Gmail in volunteer languages (e.g. Irish, Welsh, Cherokee), Google Translate, Endangered Languages Project & Dynamics of Language. The success of these projects relates to the involvement of volunteers from those IML communities. All of these highlighted the importance of both industry and community engagement with language promotion through technology.

MINORITY LANGUAGES AND TECHNOLOGY: OPENNESS AND CROSS-LANGUAGE COLLABORATION

Professor Kevin Scannell, Department of Mathematics and Computer Science, St Louis University, USA, discussed the benefits of open source software, open databases, and cross-linguistic collaborations for the development of language technologies. These points were illustrated through two case studies: first, the language revitalization work currently underway at Biigtigong First Nation in Ontario, and second, new Scottish and Manx Gaelic resources aimed at Irish speakers interested in learning these languages.

Based on his vast experience in this area, he put forward a number of recommendations aimed at governments, academics, and industry partners to put the development of language technology for Irish and other minority languages on more solid footing: i.e. Government investment; recognition of community contributions; tech industry support and cross-language collaboration.

DEVELOPMENTS IN MACHINE TRANSLATION

Dr. John Judge, Dublin City University and Dr. Elaine Uí Dhonnchadha, Trinity College Dublin, gave a joint presentation that both explained the workings of machine translation MT systems, both statistical and rule-based. They highlighted the strengths and weaknesses of both, thus providing motivation for a combined approach. They also

explained why statistical MT systems are more effective when tailored to deal with texts of a specific nature (domain-specific) and such custom built systems are far superior to free online translation systems.

Based on on-going projects, we also heard of the benefit of developing tailored SMT systems to meet the translation demands of the Irish government. The need for a translator's involvement in the post-editing process was outlined, noting that MT is not intended to replace translators, but instead take the repetitiveness out of the task and make their work more efficient. This talk also focused on the open-source nature of most SMT systems and the general culture for openness and sharing in the MT research community. It also showed how important collaboration between the two universities is for the future of Irish machine translation.

TECHNOLOGY IN ACTION: CASE STUDIES

Professor Ailbhe Ní Chasaide, Trinity College Dublin, spoke about the ABAIR project, which has seen the development of Irish text-to-speech synthesis (TTS) along with a number of speech-enabled applications. Such technology provides access for learners at all levels to native speaker voices, and therefore reaches not only those who are school learners lacking linguistic support at home, but also learners worldwide who are too geographically removed from linguistic communities.

This work illustrates the potential importance of speech and language technology for endangered and minority languages. From her team's experiences, she noted the need for tailoring technology to minority languages, as the requirements of those speakers often differ from those of major world languages. Further, the user's involvement is also crucial in the development of speech and language technologies.

TECHNOLOGY IN ACTION: CASE STUDIES

Donald W Morrison, Stòrlann Nàiseanta na Gàidhlig, spoke about Stòrlann and web-based support for the Gàidhlig education community in Scotland.

Stòrlann Nàiseanta na Gàidhlig was formed in 1999 to support the Gàidhlig education system in Scotland and to co-ordinate the production and distribution of curriculum materials for Gaelic education. Over time the organisation has extended its function to include all areas of Gàidhlig medium and Gaelic learner education and has a key role in the development of education related terminology, support for parents, and the provision of professional development opportunities for teachers. The organisation harnesses the potential of web-based social media and online resources in text, audio and visual

formats to achieve maximum reach and support for the wider Gaelic education and learner community throughout Scotland.

He identified some of the challenges and key messages as follows:

Challenges:

- Limited bandwidth and broadband speeds
- Inconsistencies in firewalls and service providers (Local Authorities)
- Equipment – availability and standards of provision
- Access – plethora of permissions issues and restrictions in schools

Key messages:

- Collaborative practice essential
- Use popular/existing/known platforms where possible
- Invest and prepare now, in advance of future demands

TECHNOLOGY IN ACTION: CASE STUDIES

Gareth Morlais, Welsh Government, informed us of the Welsh Government Welsh-language Technology and Digital Media Action Plan and £750k Fund 2013-16, which set out to do five things:

1. Marketing and raising awareness
2. Motivating the main technology companies
3. Encouraging the development of new Welsh-language software applications and digital services
4. Stimulating the creation, sharing and consumption of Welsh-language digital content
5. Supporting good practice in the public, private and third sectors

He also gave us insight into 25 successful Welsh-language technology projects, which have led or will lead to diverse new services such as: a mobile phone live translation app, Welsh-language Minecraft walkthrough online training programme, a program that aggregates Welsh-language tweets and surfaces trending topics in Welsh, new openly-licensed speech and language technologies from Bangor University's Language Technology Unit and the addition of Welsh to the interfaces and notifications of some iOS apps.

Welsh was noted amongst the delegates as a blue-print for success in minority language technology, and Gareth attributed continued success to being open to collaboration and sharing infrastructural language technology with other language groups.

SMART TECHNOLOGY: THE WAY FORWARD

Delyth Prys, Bangor University, spoke about the Language Technology Unit's work in developing speech-recognition technologies, and the urgent need for this type of technology in a future where people will engage with technology more through voice controlled commands. She noted that, if IML languages are to survive this change in our daily lives, the appropriate technology needs to be available. Delyth also highlighted that research in speech recognition is an undertaking that requires substantial expertise and investment, and cannot rely solely on volunteer work.

The need for openness and sharing was also a key point of this talk, highlighted by a current project which makes tools and resources for Welsh freely available on-line, allowing them to be shared between languages. This also led to the announcement of current efforts to build a consortium of partners for European innovation funding to accomplish this type of sharing and collaboration across research groups and also extending to the inclusion to SMEs, thus producing a positive impact on our economies.

SMART TECHNOLOGY: THE WAY FORWARD

Michal Boleslav Měchura spoke about the need to concentrate on tools such as spellcheckers, dictionaries, speech synthesizers and localized software which facilitate users in creating original content in a minority language and which normalise the presence of the language in public.

He also highlighted his own concern with misuse by the public and sometimes by public bodies of free online translation tools, such as Google Translate, and how it can lead to the production of poor quality linguistic content. He also acknowledged that machine translation, within the correct post-editing environment, is valuable. In addition, he suggested that there should be greater focus on natural language generation systems.

CHALLENGES FACING LOW COHORTS OF SPEAKERS

Tony Scott Warren, L'Office du Jèrriais, gave us an overview of the status of Jèrriais, a highly endangered language, spoken in the Island of Jersey. Estimates from 2011 showed that around 2,700 people – 2.1% of the population – had the ability to speak it. 77% of them were aged over 60. He discussed the role of L'Office du Jèrriais in the

development of a teaching programme in school, in the production of publications, promotion of the language through the media and organisation of promotional activities.

Tony also reported on how L'Office du Jèrriais engages the public through an active blog, Facebook, Twitter, YouTube and Soundcloud, and on their on-going efforts to assess technological developments for use in the maintenance of Jèrriais. They see a clear need for collaboration with or leveraging better-resourced minority languages of the group.

CHALLENGES FACING LOW COHORTS OF SPEAKERS

Adrian Cain, Isle of Man, spoke about the challenges of developing resources for Manx. While the number of speakers is growing, they have found that they need to be innovative and creative with the limited resources they have. He noted that technology has a key role to play not just in providing accessible material for learners and speakers, but is equally important for providing the language with credibility. He then provided an overview of how Culture Vannin has used technology to support Manx, especially to facilitate the learning of Manx language, Manx music and songs online. He outlined the difficulties involved when competing against larger (possibly more popular) organisations when applying for funding, and noted the importance of identifying a point of difference or a niche demand in order to generate the required credibility when acquiring funds for IML technology.

CHALLENGES FACING LOW COHORTS OF SPEAKERS

Gordon McCoy, Northern Ireland, gave us an insight into the use of technology among learners of Irish in East Belfast. He explained that as learners of Irish at Turas, the Irish language project of East Belfast Mission, are often ideologically and geographically separate from other learners or fluent speakers Northern Ireland, the majority of them have little or no opportunity to speak Irish outside the classroom. Social media sites and apps in particular are proving beneficial in closing this gap. Tailored language technology that facilitates this digital interaction is therefore clearly needed.

He also highlighted the issue with the lack of sites, apps and interactive content available to these learners in their own Ulster dialect of Irish. This further reinforces the clear requirement for awareness of varying contexts and user needs that need to be considered when developing speech and language technology.

CHALLENGES FACING LOW COHORTS OF SPEAKERS

Mike Tresidder, Cornish Language Partnership, provided an overview of examples of how information technology has supported the Cornish language movement. While there is a growing interest in learning the Cornish language, there is no sizable geographical concentration of language activists, technology has offered opportunities to overcome this. For example, through online learning platforms such as www.learncornishnow.com and www.saysomethingincornish.com have proven successful, with the latter collecting user registration data to allow tracking of the success of the programme. Through a university collaboration project, a virtual online learning environment has also been set up aid pre-school teachers.

It was also noted, however, that the pedagogic viability of these online language learning sites still need to be assessed, and that further research is required to understand a low uptake of engagement by the established Cornish language community.

CHALLENGES FACING LOW COHORTS OF SPEAKERS

John Murray, Ulster-Scots Community Network, described the many challenges faced in the promotion of engagement with Ulster-Scots. These include limited access to key written materials (with the use of Ulster-Scots being a more oral tradition), regional variations/differences of opinion on dialects, perceptions of others (on its dialect/language status), political overtones, lack of government intervention on decisions of investment, a lack of standardisation which deters academic involvement and finally, the significant lack of suitable education resources.

Despite these challenges however, we heard about the large role that technology plays in making Ulster-Scots accessible to learners and speakers worldwide. Based on online radio broadcasts (funded by the Ulster Scots Agency), there are an estimated 10,000 speakers among the wider Ulster Scots diaspora in countries such as the U.S, Australia and Western Europe. From a TV perspective, the language also featured in a fully animated series of Seamus Heaney poems narrated by Billy Connolly. Apps and social media have also proved successful in facilitating networking and accessing other US communities and speakers. In addition, the digitisation of key texts such as poems and letters, and of tape recorded interviews, for example, has increased the amount of available written texts to learners of Ulster Scots. Finally, QC (quick response coding) has also been made available at a number of sites and heritage trails that allows for scanning of QC code so that visitors can hear the content in Ulster Scots. All of these highlight the benefit of engaging with technology for promoting a language with very

limited resources, and yet highlights the need for more sophisticated tools for continued future use.

HOW CAN YOU USE SOMETHING IF YOU DON'T KNOW IT EXISTS?

Rd. Jeremy Evas, Cardiff University, Wales, discussed behavioural change theory in relation to the update of computer interfaces in minority languages. Results of recent research on the interaction of users with Welsh interfaces (on websites, cashpoints, Operating Systems and Microsoft Office tools), showed that there is generally low awareness of the availability of Welsh language interfaces for these technologies. However, when they are presented correctly to users (enforcing a choice of language, and therefore highlighting the availability of a Welsh language option), engagement and use can increase substantially. Such approaches can serve to reinforce normalisation of Welsh language in everyday life.

ROLE OF MEDIA IN PROMOTING LANGUAGES THROUGH TECHNOLOGY

Rónán Mac Con Iomaire, RTÉ (Raidió Téilifís Éireann), spoke about a recent 5-year Irish Language Action Plan, launched by RTÉ, to increase the presence of the Irish language in mainstream media. The decision was taken to not draw such a distinct divide between the major language (English) and the minor language (Irish), but to put Irish at the centre of everything thus normalising the use of Irish for the average consumer of RTÉ media – that being through television, radio, online content or social media. Such an approach highlights the importance of national media bodies' engagement with minority languages through all possible outlets, in particular as part of an effort to encourage an increase in uptake of language use amongst those in the wider public. Rónán noted that while the 2011 census reports over 1.7 million people use Irish outside of the education system, in reality this figure is only around 77,000. The gap serves to highlight, that while there is a clear positive disposition towards the language, there is still a large number of willing speakers that need support and encouragement in using Irish more frequently in their daily lives. This approach to normalisation through technology hopes to encourage this subgroup to increase their engagement with the language.

ROLE OF MEDIA IN PROMOTING LANGUAGES THROUGH TECHNOLOGY

Ciarán Ó Cofaigh, ROSG, comes from a media production background. In this talk, he highlighted the growth of the international Irish language community, and the difficulties for learners who are geographically removed from native Irish speakers, or contexts where Irish being spoken in everyday use is the norm, such as the Gaeltacht areas.

Thus, there is a clear need for providing an online forum or platform on which disparate communities can connect through Irish. He suggests exploiting ongoing developments in the area of online virtual 3D worlds and tailoring such technology to the Irish language communities, thus providing learners with the opportunity to learn Irish through ‘one to one’ interactions with other learners worldwide.

LANGUAGES AND TECHNOLOGY FOR YOUNG PEOPLE

Janice Feighery, Techspace, delivered a presentation on the work of Techspace in engaging young people with technology, while using the Irish language. The aim of the movement is to help equip youths with skills for the technological age, and to date has impacted over 2,500 young people. A branch of this work has focused on translating technical content of learning resources into Irish – both documents (Adobe Youth Voice) and online videos (dubbing video content). This project is funded by the Department of Arts, Heritage and the Gaeltacht and has proven successful in inspiring young Irish people to influence the use of their native language in technology.

LANGUAGES AND TECHNOLOGY FOR YOUNG PEOPLE

Alan Esslemont, MG ALBA, provided an overview of the success of LearnGaelic.scot, a website for Scots Gaelic learners of all levels. This project highlights the benefits of collaboration as it is funded and run in a broad partnership involving MG ALBA, BBC, Bòrd na Gàidhlig, Sabhal Mòr Ostaig and the Board of Celtic Studies Scotland. The website provides linguistic resources such as a dictionary and thesaurus, along with content for language learning support. In addition it provides listings of all Scots Gaelic language courses within Scotland.

Alan noted the benefits of using social media to reach out to language learners and potential website users. Website statistics show that an increase of users reflects the positive uptake and success of this project. He also asserted that new technology is central in the creation of effective non-linear audio-visual tools for language learners.