

Can we Defuse the Digital Timebomb? Linguistics, Speech Technology and the Irish Language Community

Ailbhe Ní Chasaide, Neasa Ní Chiaráin, Harald Berthelsen, Christoph Wendler,
Andrew Murphy, Emily Barnes, Christer Gobl

Trinity College, Dublin, Ireland
anichsid@tcd.ie, neasa.nichiarain@tcd.ie
www.abair.ie

Abstract

Does speech/language technology represent a 'digital timebomb' - or an unprecedented opportunity - for minority and indigenous languages? For successful outcomes, technology development must address linguistic challenges, answer to the needs of the local language communities, enlisting them as a central partner in development. The Irish language ABAIR initiative is building (i) linguistic resources, (ii) core technologies, and (iii) applications for public, educational and access/disability use. The Government's *Digital Plan for Irish Speech and Language Technology* provides a model of the support needed by minority languages in the digital age, if the language is to feature in everyday community activities.

Keywords: Irish, speech technology, digital strategy, language community, education, disability

Résumé

An dainséar nó deis dúinn an réabhlóid dhigiteach? Do mhionteangacha agus do theangacha i mbaol, caithfidh forbairt na teicneolaíochta oiriúint do struchtúr na teanga agus do riachtanais an phobail. Sa tionscadal ABAIR a dhíríonn ar theicneolaíocht na hurlabhra, tá trí réimse taighde ar siúl: (i) bun-acmhainní teangeolaíochta, (ii) teicneolaíochtaí, agus (iii) áiseanna don phobal, do lucht an oideachais agus dóibh siúd faoi mhíchumas. Tá *Plean Digiteach do Theicneolaíocht Uirlabhra agus Teanga na Gaeilge* an Rialtais mar mhúnla don tacaíocht atá de dhíth ar mhionteangacha sa ré dhigiteach. Beidh toradh na hoibre ag brath ar chomhpháirtíocht idir theangeolaithe, theicneolaithe agus phobal na Gaeilge.

1. Timebomb or Unprecedented Opportunity ?

Speech and language technologies have become deeply embedded in our daily lives and are increasingly central to how we work, interact socially, access information and education. For a language that is not digitally available the language community is forced to switch to the major language for access to these resources. Increasingly the technology is seen as a 'digital timebomb' (Evans, 2018), which is forcing the shift from the indigenous to the major language (Mac Thomáis, 2018).

Precisely because of its increasingly central presence in our lives, digital technology also presents an unprecedented opportunity for the documentation, preservation and revitalisation of the endangered language. There is a growing awareness that the provision of speech and language technologies is one vital strand in ensuring language survival. The Irish Government is launching a *Digital Plan for Irish Speech and Language Technology*, which will provide a 10-year roadmap for research and development (see Section 4).

Irish is an endangered language (Moseley, 2013), spoken as a community language in small 'Gaeltacht' areas, mostly in the West of Ireland. Nonetheless, as an official language of the State, and since 2007 of the EU, it enjoys considerable State support and is a core curricular subject in Irish primary and second level schools.

In this paper, we describe the experience of the Irish language initiative ABAIR, which is developing speech and language technology for Irish (ABAIR, 2019). This entails a longterm strategy of (i) basic linguistic research, (ii) core technology building, and (iii) developing

applications which draw both on the linguistic resources of (i) and the technologies of (ii). The development of applications answers to the needs of the language community. The focus has been on providing not only public resources but also on tools for education and disability/access.

While technology development in major languages is typically driven by commercial considerations, for the minority and indigenous language it is important to focus on the community's own priorities. This can mean a different trajectory: for example, recognition and synthesis in call centres to handle large volumes of calls are hardly necessary, but their use in the teaching of the language can have far reaching implications in the minority language context.

2. Linguistic and Sociolinguistic Challenges

Each language presents its own complex linguistic and sociolinguistic set of challenges. Irish has a number of structural features that make it different from English and many of the European languages. It is a verb-initial language (VSO); it is highly inflected with alternations affecting both the beginnings and ends of words (see Figure 1a for an illustration of inflected forms for the word *bád* /bʲ a: d̪ˠv/ 'boat'). A striking feature of the language is its sound system, where there is a contrast of palatalized and velarized consonants (see Figure 1b) (Ní Chasaide, 1999). Simply put, there are twice as many consonants as in, for example, English. This quality distinction among consonants serves not only to differentiate words, but is intricately linked to the system of grammatical inflection. For example, switching

between a final velarized and palatalized consonant differentiates between the nominative and genitive case, as in *bád* /b^v a: ɫ^v/ vs *báid* / b^v a: ɫ^h/ (see first two items in Figure 1a). As the Roman alphabet does not provide different letters for this consonantal distinction, the orthographic conventions for signaling consonantal quality entail the use of adjacent vowel letters. The vowel letter closest to the consonant (before or after) indicates it as belonging to the palatalized or velarized set. This, of course, results in complex strings of vowels in orthographic forms. As the illustration in Figure 1c shows, many strings of vowels correspond to a single vowel phoneme (in this case /i:/) as many of these vowel letters are simply markers of consonantal quality. Structural aspects of the language such as these are important considerations in developing resources and technologies and they are crucial to many of the applications, particularly those targeting language teaching/learning. As in many minority languages, there is no spoken standard but rather strikingly different dialects. This has implications for technology development. In the case of Irish, it was necessary to consider multidialect provision from the outset, as this would be essential for its acceptability, uptake and usage by the community. The importance of this factor is not necessarily immediately obvious to an outside agency.

a)		b)						c)		
The word <i>bád</i> (boat)								Sound: /i:/		
		LABIAL		DENTAL	ALVEOLAR	ALVEOLO-PALATAL	PALATAL	VELAR	Orthography	Example
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^v /	p ^v	b ^v	t ^v	d ^v	n ^v	ɲ	ɟ	í	sí
báid	/b ^v a: ɫ ^h /	p ^h	b ^h	t ^h	d ^h	n ^h	ɲ	ɟ	í	sí

outside Ireland. This brings home the potential role of digital technology, revealing a global community we were hitherto unaware of, and creating broader networks which are a powerful source of strength and future support.

Current work on synthesis is extending coverage from the main dialects to those lesser spoken dialects to ensure they are not left behind by these technologies. A synthetic voice for the local dialect preserves a virtual speaker: a valuable resource for the local community trying to hold on to its language, now and for future generations.

Provision of child as well as male and female ‘speakers’ is a further priority. Furthermore, given the bilingual context in which Irish synthesis is used, and given the frequent code switching in Irish (a feature of all minority languages), work is underway to provide bilingual Irish+English voices (e.g. for bilingual websites) that can also provide codeswitching.

Dialogue systems and interactive games are future priorities (see below). Therefore, ongoing research is modelling the prosodic/voice quality modulations needed to capture the affective nuancing such applications require (Murphy et al. 2019 ; Yanushevskya et al. 2017).

3.2.2 Current Early Stage Developments

Development of speech recognition is underway and a preliminary system has been created. Success in this area will depend crucially on recordings from a very large number of speakers. To be useful in future applications, recognition for Irish must accommodate dialect diversity, native speakers/learners, children/adults. A targeted crowdsourcing initiative *Mile Glór* (‘ a thousand voices’) has been launched, which allows Irish speakers to record as many prompts as possible from materials that are appropriate to the speakers’ dialect. It identifies categories of speaker (child/adult, specific dialect etc).

Early research towards spoken dialogue systems is also underway and will progress as speech recognition becomes available. A pilot intelligent tutoring system, *Taidhgin* has been developed. Here the learner input is through text, which is then spoken aloud by a TTS voice, and spoken answers are delivered by a talking monkey. Evaluations in schools yielded very positive reactions and dialogue systems are a priority for future research.

3.3 Applications

As mentioned, applications target the broader public as well as specific education and disability/access needs.

3.3.1 Applications for the Language Community

The public webpage at www.abair.ie, has of itself become a public application. Parents use it to help their children with homework (being able to access the pronunciation of written words helps, overcome the obstacle of the complex mapping of orthography → sounds). Foreign users use it for this and wide diversity of reasons: e.g. to pronounce Irish names, to get authentic pronunciation of Irish utterances in a play, etc. An app version of the webpage is being launched, making public access simpler. An add-on feature to web browsers allows all online text to be read aloud. This gives ready access to the spoken realization of dictionary content, online newspapers, email, etc. It also allows the Government to fulfill their

statutory obligation to make information and webpages accessible to all, through the medium of Irish.



Figure 3: Homepage of the ABAIR.ie initiative

3.3.2 Educational Applications

The future of the language depends overwhelmingly on how successfully we can transmit it to the younger generation. Irish language teaching and learning can potentially be transformed by digital speech and language technology and much of ABAIR’s activity to date has explored how best to deploy all resources as they come onstream. Combining the core technology of (ii) and the linguistic resources of (i) presents unprecedented opportunities for linguistically informed content and technologically advanced platforms that are pedagogically effective, attractive and motivating for the language learner. Two of the educational applications piloted to date are reviewed briefly here, targeting early and more advanced/older learners respectively. These applications bring native speaker models of the language into the classroom, making the spoken language central for all aspects of language learning. They also put the power of multimodal games and digital technologies at the disposal of learners and teachers.

The game *‘Lón don Leon’* (‘lunch for the lion’) aims to develop phonological awareness (the ability to perceive and produce the palatalized/velarized distinction) and early literacy skills (an explicit grasp of the phonic basis for the spelling rules of the language, where vowel letters sometimes mark consonantal quality and sometimes represent actual vowel targets). Phonological awareness is trained through the use of minimal pairs, where the palatalized/velarized consonant contrast is essential to the differentiation of two words, here ‘lunch’ and ‘lion’. This is done by embedding them in newly composed songs, illustrations and stories – reinforcing learners’ awareness of the contrast and their ability to reproduce it correctly. Through games and language learning activities the learners’ acquisition is monitored. When the sound contrasts are adequately acquired, the orthographic letters are introduced. Differential colour coding of palatalized (orange) and velarized (blue) versions of consonant and vowel letters guide the young learner to an understanding of how they combine in the writing system. This game can be used in the home or in a classroom setting.

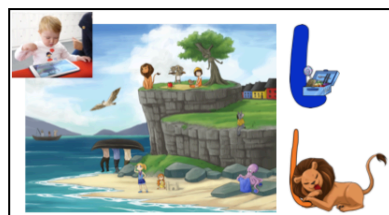


Figure 4: *Lón don Leon* phonological awareness game.

A second platform, *An Scéalai*, ‘the Storyteller’, targets

the older learner and has as an objective the parallel training of all four language skills: writing, listening, reading and speaking (Ní Chiaráin & Ní Chasaide, 2018; 2020). The cover page of *An Scéalai* is illustrated in Figure 5. On the face of it this platform presents as a writing tool, which includes spelling and grammar checkers. However, the TTS voice output is a central feature and the learner is led to have all written content spoken aloud. Prooflistening is used as the initial strategy for learners to correct their own written work. Here, constant reinforcement of the links of sound → orthographic sequence is essential and evaluations to date show it to be a highly effective strategy in developing learners' awareness of many types of errors. As an intelligent Computer-Assisted Language Learning (iCALL) application, there are many types of inbuilt prompts to review and correct written work. Many draw on the specific linguistic resources developed in the first research strand above. The prompting is carried out by an interactive dialogue partner who both speaks to the learner (using their preferred TTS voice/dialect) and provides spoken feedback along with written guidance. The dialogue partner can detect areas of recurring weakness, e.g. irregular verb conjugation, and offer the learner opportunities to carry out exercises targeting these. This platform is entirely suited to autonomous learning but is also set up to allow its integration in the classroom where the teacher can monitor their own students' progress and provide further oversight and feedback. This platform is currently under evaluation in second and third level educational settings in Ireland and abroad.

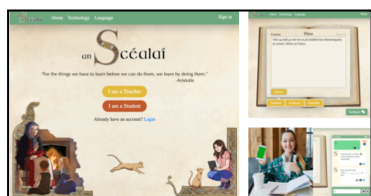


Figure 5: iCALL platform *An Scéalai* ('the Storyteller').

3.3.3 Disability/Access Applications

These technologies are particularly important for the inclusion of those with disabilities. Urgent requests from parents of schoolgoing children with visual disabilities led to the development of screenreading software, using the open source NVDA framework (McGuirk, 2015). Users control the spoken output in terms of speed and dialect, and simultaneous Braille output is provided (see Figure 6 below). Multimodal textbooks for the visually impaired have also been developed (Ní Chasaide et al. 2019).



Figure 6: visually impaired user listening & using the Liblouis Braille system simultaneously.

Current educationally-oriented multimodal platforms and games are also intended to provide for the assessment of phonological awareness and reading skills and eventually

as the basis for literacy remediation for those with dyslexia (Barnes et al. 2018). A 'design for all' approach for ABAIR applications aspires to include the widest possible groups of users, especially those with disabilities.

4. The Digital Plan for Irish

The Department of Culture, Heritage and the Gaeltacht is launching a *Digital Plan for Irish Speech and Language Technology*, which will seek to ensure that the power of digital technology is available to the Irish language community. It encompasses not only the development of specific core technologies (recognition, synthesis, machine translation, etc.) but also the range of applications and linguistic research that will enable the Plan to have a longterm impact. It also aims to set in place the infrastructure to ensure the future capacity to keep up with the rapid evolution of these technologies.

5. Central Role of Language Community

The extent to which digital technology serves to revitalize the language depends on the extent to which the language community make it their own. The ABAIR experience demonstrates how, over the years, collaboration with the language community has moved from informant → requesters of applications → partners in design, testing and dissemination of specific user applications. The *Digital Plan* acknowledges the importance of the native speaker Gaeltacht community in this enterprise, as the repository of living language 'experts' and seeks its engagement in every aspect. It further aims to ensure that native speakers are among the leading researchers in future digital development. It also seeks to engage in collaborative ventures with the broader Irish language community and user groups at home and abroad.

6. Conclusion : a Model for other Minority/Indigenous Languages

Although the Irish language is endangered and the Gaeltacht communities fragile, Irish is in a relatively privileged position of being a national language with EU support. Its survival to date owes much to State support. The work of ABAIR has only been possible due to State funding and enthusiastic support.

It is a basic principle of the ABAIR initiative that all outputs are freely available to the community. It is also a basic aspiration, reflected in the *Digital Plan*, that Irish developments might support the efforts of other language communities, many of whom enjoy little or no State support or recognition. The kinds of challenges faced by minority and indigenous languages are often rather similar and the solutions for one language may be exactly right for many others. It is thus imperative that we learn from each other and that we develop mechanisms for cooperation, sharing our experience, expertise and, where possible, resources.

7. Acknowledgements

This work is supported by the Department of Culture, Heritage and the Gaeltacht with National Lottery funds. This forms part of the Government's 20-year Strategy for Irish 2010-2030.

8. References

- ABAIR: the Irish language speech synthesizer - webpage with synthetic voices (2019). www.abair.ie
- Barnes, E., Ní Chasaide, A. and Ní Chiaráin, N. (2018) The design and pre-testing of literacy and cognitive tasks in Irish and English, *Literacy Association of Ireland 42nd International Conference, Dublin, Ireland*, pp1 - 11
- Dalton, M. and Ní Chasaide, A. (2005). Tonal alignment in Irish Dialects. *Language and Speech* 48, 441-464.
- Dorn, A. and Ní Chasaide, A. (2015). Sentence mode differentiation in four Donegal Irish varieties. *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK: the University of Glasgow. Paper number 0482, 5 pp. ISBN 978-0-85261-941-4. Retrieved from <http://www.icphs2015.info/pdfs/Papers/ICPHS0482.pdf>
- Evans, G. (2018). Report on Language Equality in the Digital Age. Retrieved online 22/11/2019 from: http://www.europarl.europa.eu/doceo/document/A-8-2018-0228_EN.html
- Mac Thomáis, U. (2018). “Tá guth-aithint an Bhéarla ag déanamh Béarlóirí de Ghaeilgeoirí, mé féin ina measc!” Article from [tuairisc.ie](https://tuairisc.ie/ta-guth-aithint-an-bhearla-ag-deanamh-bearloiri-de-ghaeilgeoiri-me-fein-ina-measc/). Retrieved from: <https://tuairisc.ie/ta-guth-aithint-an-bhearla-ag-deanamh-bearloiri-de-ghaeilgeoiri-me-fein-ina-measc/>
- McGuirk, R. (2015). *Exploration of the Use of Irish Language Synthesis with a Screen Reader in the Teaching of Irish to Pupils with Vision Impairment*. Unpublished M.Phil. dissertation. Trinity College Dublin, Ireland.
- Moseley, C. (Ed.). (2010). *Atlas of the world’s languages in danger* (3rd ed.). Paris: UNESCO Publishing. Retrieved from: <http://www.unesco.org/culture/en/endangeredlanguages/atlas>
- Murphy, A., Yanushevskaya, I., Ní Chasaide, A. and Gobl, C. (2019). The role of voice quality in the perception of prominence in synthetic speech, *INTERSPEECH 2019, Graz, Austria*, 2019, pp2543 - 2547
- Ní Chasaide, A. (1999). Irish. *Handbook of the International Phonetic Association: a guide to the use of the international phonetic alphabet*. Cambridge: Cambridge University Press, 111-116.
- Ní Chasaide, A., Ní Chiaráin, N., Berthelsen, H., Wendler, C., Murphy, A., Barnes, E. and Gobl, C. (2019). Leveraging Phonetic and Speech Research for Irish Language Revitalisation and Maintenance, *ICPhS the 19th International Congress of Phonetic Sciences, Melbourne, Australia*, pp994 - 998
- Ní Chiaráin, N. and Ní Chasaide, A. (2018). An Scéalaí: synthetic voices for autonomous learning. In: Taalas, P., Jalkanen, J., Bradley, L., Thoušny, S. (eds), *Future-proof CALL: language learning as exploration and encounters – short papers from EUROCALL 2018*, 230-235.
- Ní Chiaráin, N. and Ní Chasaide, A. (2020). The Potential of Text-to-Speech Synthesis in Computer-Assisted Language Learning: A Minority Language Perspective In Andujar, A. (ed) *Recent Tools for Computer-and Mobile-Assisted Foreign Language Learning*, Hershey, PA, IGI Global, pp.149-169.
- O’Reilly, M. and Ní Chasaide, A. (2016). Modelling the timing and scaling of nuclear pitch accents of Connought and Ulster Irish with the Fujisaki model of intonation. *Proceedings of the 8th International Conference on Speech Prosody*, Boston, Massachusetts, pp. 355–359. DOI: 10.21437/SpeechProsody.2016-73
- Yanushevskaya, I., Ní Chasaide, A. and Gobl, C. (2017) Cross-speaker variation in voice source correlates of focus and deaccentuation, *INTERSPEECH 2017, Stockholm, Sweden*, pp1034 - 1038