TUNEPAL - DISSEMINATING A MUSIC INFORMATION RETRIEVAL SYSTEM TO THE TRADITIONAL IRISH MUSIC COMMUNITY

Bryan Duggan, Brendan O' Shea

School of Computing
Dublin Institute of Technology
Kevin St
Dublin 8
Ireland
{bryan.duggan, brendan.oshea}@dit.ie

ABSTRACT

In this paper we present two new query-by-playing (QBP) music information retrieval (MIR) systems aimed at musicians playing traditional Irish dance music. Firstly, a browser hosted system - tunepal.org is presented. Secondly, we present Tunepal for iPhone/iPod touch devices - a QBP system that can be used *in situ* in traditional music sessions. Both of these systems use a backend corpus of 13,290 tunes drawn from community sources and "standard" references. These systems have evolved from academic research to become popular tools used by musicians around the world. 16,064 queries have been logged since the systems were launched on 31 July, 2009 and 11 February, 2010 respectively to 18 May 2010. As we log data on every query made, including geocoding queries made on the iPhone, we propose that these tools may be used to follow trends in the playing of traditional music. We also present an analysis of the data we have collected on the usage of these systems.

1. INTRODUCTION

There exist approximately seven thousand unique traditional Irish dance tunes [1]. Musicians playing traditional music have a personal repertoire of up to a thousand tunes. Many of these tunes are known by multiple names, while many more are known simply as "gan anim" (without name). In the past, commercial recordings of traditional music were accompanied by extensive sleeve notes providing biographic information on the tunes in the recording. In the modern age two trends have emerged. Firstly, the use of digital audio formats and digital downloading of music has meant that personal music collections do not contain this biographic data and many musicians are unfamiliar with the history and background to the tunes they are playing. This fact is compounded by the fact that although traditional tunes often have colourful and memorable titles (Table 1), there is nothing to link the title of a tune with its melody [2].

The second trend is the development of extensive; crowd sourced biographic references and discographies for tunes on websites such as thesession.org [3]. Linking the melodies of traditional Irish dance tunes to biographic data about the tune, including names, is the goal of an ongoing project at the DIT School of Computing.

Name
The Bucks Of Oranmore
Come West Along The Road
Repeal Of The Union
The Chicken That Made The Soup
More Power To Your Elbow
If It's Sick You Are Tea You Wants
The Night We Made The Match
Last Night's Fun
My Former Wife
The First Night In America

Table 1: Tune names taken from [4]

In our previous work [5-7], we described a proof of concept music information retrieval (MIR) system adapted to the characteristics of traditional Irish dance music that addressed this very problem. In this paper, we present follow up work in developing this research into robust and reliable tools that are now being used by thousands of musicians around the world. Specifically we present tunepal.org – a browser hosted query-by-playing (QBP) system and Tunepal for the iPhone/iPod touch – a QBP system that can be used *in situ* in traditional music sessions. As these systems log details of every query being made (including geotagging queries made on the iPhone), they represent a unique opportunity to analyse the *zeitgeist* of traditional music. In other words, to identify trends, popular tunes and tune types being played around the world.

Section 2 of this paper presents a brief overview of our previous work in this area. Section 3 presents the architecture of tunepal.org. Section 4 presents Tunepal for iPhone. Section 5 presents a summary of usage data collected from the two systems. Section 6 presents a summary and conclusions.

2. RELATED WORK

Our previous work describes Tunepal for Windows Mobile devices such as smartphones and PDA's [8,9]. This is a symbolic MIR system that allows musicians to search for tunes by name, retrieve the ABC notation [10] for the tune and playback the tune. Figure 1 presents screenshots of Tunepal running on a Windows Mobile smartphone. Our aim with this system was to facilitate musicians to

start tunes that they could recall the name of, but not the melody.

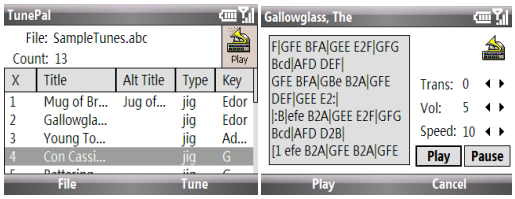

Figure 1: Screenshot of Tunepal for Windows Mobile

MATT2 (Machine Annotation of Traditional Tunes) is a standalone QBP MIR system for traditional Irish dance tunes initially developed for the tin-whistle and wooden flute [7] and subsequently enhanced to support queries on a range of traditional instruments including the uilleann pipes, concertina and fiddle [5,11]. MATT2 is based on two subsystems – a transcription subsystem and a matching subsystem. The transcription subsystem uses an onset detection function based on comb filters (ODCF) developed especially for the transcription of traditional music [12]. A harmonicity, pitch detection algorithm based on Klapuri's [13] multi-pitch estimator is used to extract frequencies from the FFT (Fast Fourier Transform) of a note frame. MATT2 incorporates Ornamentation Filtering (OF) to remove expressiveness from the transcription. The corpus used in MATT2 is Norbeck's reel and jig collection [14], which is pre-processed to expand parts, separate variations, remove ornamentation and normalise for register. This collection contains 1582 reels and jigs, with variations. Matching is achieved using the substring edit distance algorithm [15], with a cost function modified to take account of breath marks in the transcription. An evaluation of this system is presented in [11].

An enhancement to MATT2 is the TANSEY (Turn ANnotation from SEts using SimilarityY profiles) algorithm, named after the traditional flute player Seamus Tansey [6]. TANSEY is a segmentation algorithm to annotate tunes played in set (sequences of multiple tunes repeated multiple times and played segue). TANSEY makes use of *melodic similarity profiles* and can retrieve the start and end of each repetition of a tune, count the repetitions and retrieve the name and associated biographic data associated with each tune in a recording of a set of tunes.

3. TUNEPAL.ORG

Our first task in disseminating the work described in section 2 was to expand the corpus used in the experiments described in [5,11] to include a comprehensive collection of traditional Irish music from definitive sources available in ABC notation. The tunepal.org database contains 13,290 tunes drawn from community sources, such as the website thesession.org [3] and "standard" references including O'Neills Dance Music of Ireland [16] and Brendan Breathneach's Ceol Rince Na hÉireann series in five volumes [17]. Our corpus also includes collections of Welsh, Scottish and Breton music in addition to several different transcriptions of the same tune from the canon of Irish traditional music. Table 2 presents an analysis of sources of the tunes in the tunepal.org corpus.

Source	Count
thesession.org	9,310
Henrik Norbeck	1,474
O'Neills Dance Music of Ireland	994
Ceol Rince na hÉireann 1	73
Ceol Rince na hÉireann 2	192
Ceol Rince na hÉireann 3	37
Ceol Rince na hÉireann 4	220
Jonny O'Leary	196
Nigel Gatherer	794
Total:	13,290

Table 2: Sources of Tunepal tunes

In order to make the system easily accessible to traditional musicians without the necessity of installing software, a browser hosted version of MATT2 – tunepal.org was developed. For this version, the transcription algorithms were deployed in a Java applet, while the tune corpus and matching subsystems were hosted on a server. Figure 2 presents a screenshot of tunepal.org.

To find a tune, a musician records a query played on an instrument such as the concert flute, tin-whistle, uilleann pipes, accordion or concertina. An energy based

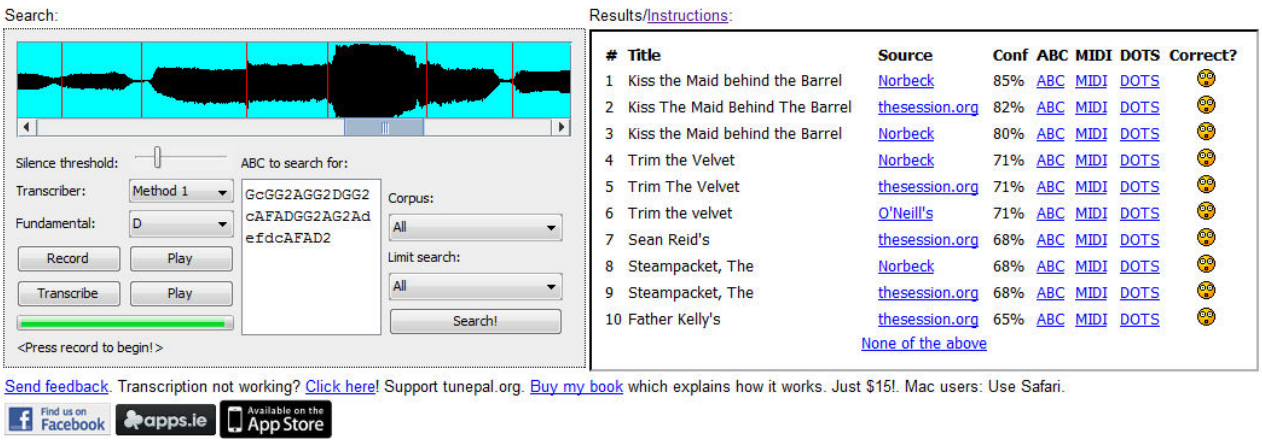

Figure 2: A screenshot of tunepal.org

silence detection algorithm removes silence at the start of recorded queries, which would affect the evaluation of the quaver length (a core element of our system). The user can then click the transcribe button and the system will extract the melody spelled in ABC notation from the recording [10]. tunepal.org differs from similar web based QBP systems such as Musipedia [18] in that traditional instrument queries are explicitly supported. Although Musipedia contains traditional Irish dance tunes as part of its corpus, it does not generate positive results when queries are played on the tin-whistle or wooden flute (as tested by the author).

Users are also offered the ability to change the transcription fundamental. This changes the frequencies used by the pitch spelling algorithm, so that tunepal.org can work with differently pitched instruments, such as Eb flutes and uilleann pipes pitched in B and C.

The query is then submitted to the matching engine, a J2EE web application; hosted on tunepal.org. The matching engine uses the substring edit distance algorithm against the corpus of *search keys* - strings of musical notes extracted from the tunes and normalised as described in [5,11]. These are stored in a MySQL database. For each submitted query, tunepal.org presents the ten closest matches in order of descending distance. MATT2 gives the correct tune as the closest match for 93% of queries in experiments using real-world field recordings of traditional musicians from sessions, classes, concerts and commercial recordings including solo and ensemble playing on traditional instruments recorded in a variety of real-world settings such as noisy public sessions [11]. In tunepal.org therefore we log the closest matching tune for a query in the database. tunepal.org incorporates a feedback system, so users can however proof listen to the results and give feedback as to which (if any) of the returned tunes was the correct one. We also store a confidence score for the match calculated as per (1), where q is the query length and ed is the minimum substring edit distance between the query and the closest match [6].

$$c = 1 - \left(\frac{ed}{q}\right) \quad (1)$$

Each tune in the database can be played, displayed in ABC notation or stave notation. Stave notation display uses ABCJS, an open source, browser hosted rendering engine for ABC notation [19].

tunepal.org was launched on 31 July, 2009. It runs on Windows, Mac and Linux systems. We promoted tunepal.org on popular traditional music discussion forums such as thesession.org and the chiff and fipple forum. tunepal.org has been quite successful and the site is now well known amongst traditional musicians having been profiled in a national newspaper [20]. At the time of writing (18 May 2010), 7,885 queries have been logged. A more detailed analysis of the usage of tunepal.org is presented in section 5.

4. TUNEPAL FOR IPHONE

Traditional Irish music is most commonly played by groups of musicians in a community setting known as a *session* [21]. Sessions usually take place in shared public spaces. It was felt important therefore that for this work to become ubiquitous, it had to be made available on a mobile handheld device. We therefore ported the functionality of tunepal.org to the iPhone platform. Figure 3 presents screenshots of Tunepal running on an iPhone.

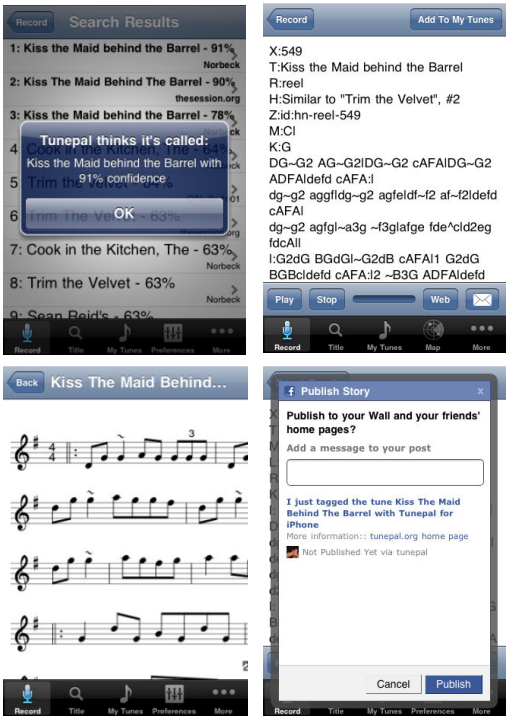

Figure 3: Screenshots of Tunepal running on an iPhone

Certain compromises were necessary in the iPhone version of Tunepal in order to make transcription speed acceptable. Firstly queries are limited to twelve seconds of audio (similar to Shazam [22]). Secondly, the sample rate is reduced to 22.05KHz and finally, onset detection is achieved using a combination of an STFT (Short-time Fourier Transform) with a Hanning window and a pitch speller instead of using ODCF. An STFT is carried out on the signal using a frame size of 2,048 samples, with a 50% overlap. This gives a frequency resolution of 10.76Hz, discriminant enough to detect pitches of traditional instruments without interpolation. Our harmonicity based, pitch detection algorithm [5] that analyses peak intervals in the frequency spectrum was ported to C++ for use in Tunepal for iPhone. Identified frequencies are then assigned pitch classes using the pitch spelling algorithm. A note onset is annotated when the pitch class changes in the time domain. The quaver length is determined using the fuzzy histogram clustering algorithm described in [5,7,11]. Ornamentation notes are removed from the transcription and long notes (crochets, dotted crochets) are

split into multiple quaver notes. The transcription string (a sequence of pitch classes) is then submitted to tunepal.org for matching.

Tunepal for iPhone uses the same back end database and infrastructure as tunepal.org and so has access to a corpus of 13,290 tunes. The iPhone version of Tunepal, returns the top ten closest matching tunes for a query with confidence scores. Similar to tunepal.org, we log each query, with the closest matching tune and confidence score. When a tune is matched both tunepal.org and Tunepal for iPhone offer the option to link back to the original source of the ABC notation on the internet. In the case of tunes indexed from the website thesession.org, this often includes extensive discussions on the origin of the tune, the source of transcription and recordings on which the tune appears (Figure 4).

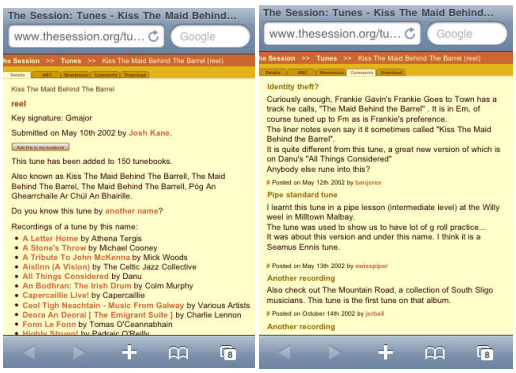

Figure 4: Biographic reference for the tune "Kiss the Maid Behind the Barrel" from the website thesession.org displayed on an iPhone

Retrieved tunes are stored in a “My Tunes” tab on the user’s device, in order of most recently tagged to facilitate future retrieval for learning purposes. Playback is achieved using ABC2MIDI [23] and the FMOD audio engine [24]. The iPhone version of Tunepal has one major advantage over tunepal.org and that is the ability for accurate geocoding (Figure 5).

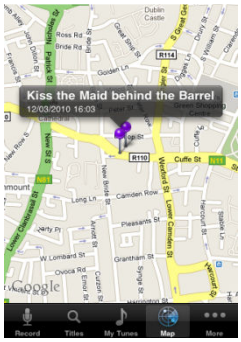

Figure 5: Geotagged tunes displayed within Tunepal on the iPhone

Therefore with the users permission, we geotag each query on the iPhone and store the longitude and latitude with each query in the tunepal.org database. This makes it possible for a user to track their queries on a map. Tunepal for iPhone was released on 11 February, 2010 and at the time of writing (18 May 2010), 5,866 QBP queries have been made, while 2,313 title searches were made (title searches were added as a feature on 13 February 2010). As the iPhone does not support programs written in Java, it was necessary to port the transcription subsystem of MATT2 and tunepal.org to a combination of C++ and Objective C. Tunepal for iPhone was listed in the top twenty cultural apps available on the iPhone by the Sunday Times (an Irish national newspaper) [25].

5. RESULTS

To date (18 May 2010) tunepal.org and Tunepal for iPhone have logged 16,064 queries since being released (Table 3).

Client	QBP
tunepal.org QBP	7,885
iPhone QBP	5,866
iPhone Title	2,313
Total:	16,064

Table 3: Queries logged from tunepal.org and Tunepal for iPhone

Table 4 gives the top ten tune types queried by users of tunepal.org and Tunepal for iPhone. The tunepal.org count was generated by counting the user verified tunes for each query. The iPhone count was generated by selecting the closest matching tune for each query.

	tunepal.org (Verified)		iPhone (QBP)	
#	Type	Count	Type	Count
1	Reel	521	Reel	1,594
2	Jig	240	Jig	913
3	Hornpipe	68	Hornpipe	211
4	Polka	57	Polka	116
5	Slip Jig	28	Waltz	111
6	Slide	23	Slip Jig	89
7	Waltz	20	Slide	56
8	Double Jig	13	Barndance	46
9	Barndance	9	Double Jig	38
10	Strathspey	7	Strathspey	18

Table 4: Top ten tune types queried by users of tunepal.org and Tunepal for iPhone

In order to minimise the effect of false positives on the iPhone counts, tunes with a confidence of < 65% are excluded. The cut-off of 65% was derived by stochastic sampling and proof listening. While this undoubtedly removes many true positives, it does eliminate most of the

false positives. The scores in Table 4 correspond broadly with the profile of tunes in most traditional musicians' repertoire, where reels and jigs assume prominence [26]. While it would be interesting to analyse the frequency that particular tunes appear in search results, more data is needed to make this analysis significant as the profile of tune appearances is in fact mostly flat, with the majority of tunes appearing only once or twice and even the top tunes appearing less than twenty times.

Table 5 gives a breakdown of QBP queries submitted by day of the week, though as these are in the local time of the server (the server is hosted in Ireland), there will be "bleed" from day to day due to the different time zones of users. Nevertheless, it is significant that weekends are more popular than weekdays for uses of tunepal.org, playing music being a leisure activity for many musicians. Tunepal for iPhone however demonstrates consistent usage across the week, which could be attributed to its portability.

	tunepal.org	iPhone	Total
Mon	999	793	1,792
Tue	1,039	862	1,901
Wed	985	728	1,713
Thurs	860	957	1,817
Fri	743	887	1,630
Sat	1,773	744	2,517
Sun	1,486	895	2,381
Total:	7,885	5,866	1,3751

Table 5: Analysis of queries by day of the week

Figure 6 better illustrates the trend towards high volumes of usage over the weekend, with significant usage on Monday and Tuesday, dropping off on Wednesday and Thursday to peak at the weekends.

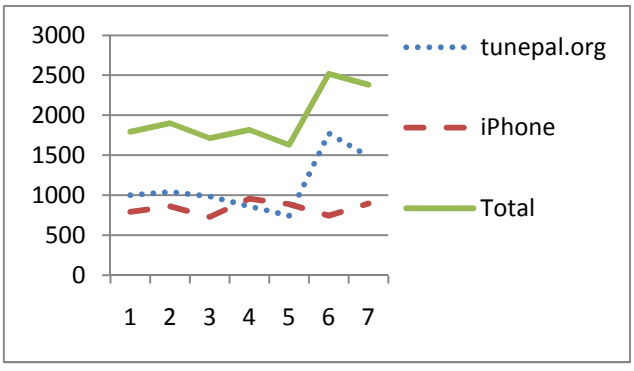

Figure 6: Plot of daily usage

We geotag queries generated by Tunepal for iPhone. An extract of this plot is given in Figure 7.

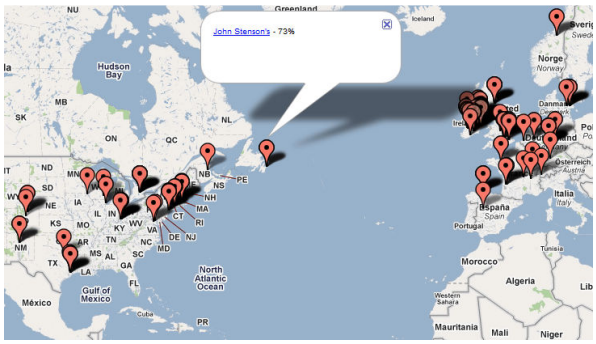

Figure 7: An extract from the worldwide geotagged query map

This is an optional feature that users must agree to; however 74% of queries made on an iPhone are geotagged. The realtime worldwide map of geotagged QBP queries can be viewed on a google map at the website http://tunepal.org.

Table 6 was generated by reverse geocoding the longitude and latitude from tagged queries to generate a profile of usage by country.

Country	Count
Ireland	1,276
United States	1,092
United Kingdom	393
Germany	179
Canada	122
Sweden	91
Spain	89
France	73
Netherlands	44
Australia	20

Table 6: Top ten countries for Tunepal for iPhone QBP queries

Although the amount of data collected is insufficient to draw any firm conclusions, it is nonetheless interesting to observe that the United States and the United Kingdom are significant sources for queries, these being major centers for the Irish Diaspora. This is a correlation we hope to explore in more detail in future work.

6. CONCLUSIONS AND FUTURE WORK

In this paper we presented two new QBP MIR systems for traditional music that developed from academic research. These tools have become popular, being used by musicians around the world to connect playing with tune names and biographic data. To achieve this, we use a corpus of 13,290 compositions collected by both the traditional music community and noted collectors such as O'Neill and Breathneach. Further we presented an analy-

sis of the data we have collected on the usage of these systems since being launched.

It is our aim to further disseminate these query-by-playing systems to the traditional music community by making them available on a greater variety of platforms such as the iPad, Android, Symbian, Maemo and Windows Phone 7 platforms. Usage of Tunepal is growing as are our usage logs. Once sufficient data is collected we hope to be able to mine these to gather new insights into musical trends and correlations that we hope to present in future work.

7. ACKNOWLEDGEMENTS

We are grateful for the support the School of Computing at the Dublin Institute of Technology who fund this work.

8. REFERENCES

[1] S. Driscoll, "A Trio of Internet Stars: ABC's," *Fiddler Magazine*, vol. 11, Summer. 2004.

[2] C. Carson, *Last Night's Fun: A Book about Irish Traditional Music*, North Point Press, 1997.

[3] "The Session."

[4] F. O'Neill, *The Music of Ireland*, 1903.

[5] B. Duggan, "Machine Annotation of Traditional Irish Dance Music," Dublin Institute of Technology, 2009.

[6] B. Duggan, M. Gainza, B. O'Shea, and P. Cunningham, "Machine Annotation of Sets of Traditional Irish Dance Tunes," *Ninth International Conference on Music Information Retrieval (ISMIR), Drexel University, Philadelphia, USA*, Sep. 2008.

[7] B. Duggan, B. O'Shea, and P. Cunningham, "A System for Automatically Annotating Traditional Irish Music Field Recordings," *Sixth International Workshop on Content-Based Multimedia Indexing, Queen Mary University of London, UK*, Jun. 2008.

[8] B. Duggan, "Enabling Access to Irish Traditional Music Archives on a PDA," *Eight Annual Irish Educational Technology Users Conference, DIT Bolton St, Ireland*, May. 2007.

[9] B. Duggan, "Learning Traditional Irish Music using a PDA," *IADIS Mobile Learning Conference, Trinity College Dublin*, Jul. 2006.

[10] C. Walshaw, "The ABC home page" Available: http://www.walshaw.plus.com/abc/.

[11] B. Duggan, M. Gainza, B. O'Shea, and P. Cunningham, "Compensating for Expressiveness in Queries to a Content Based Music Information Retrieval System," *2009 International Computer Music Conference*, Aug. 2009.

[12] M. Gainza, E. Coyle, and B. Lawler, "Onset Detection Using Comb Filters," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY: 2005.

[13] A. Klapuri, "Multiple fundamental frequency estimation based on harmonicity and spectral smoothness," *Speech and Audio Processing, IEEE Transactions on*, vol. 11, 2003, pp. 804-816.

[14] H. Norbeck, "ABC Tunes" Available: http://www.norbeck.nu/abc/index.html,

[15] G. Navarro and M. Raffinot, *Flexible Pattern Matching in Strings: Practical On-Line Search Algorithms for Texts and Biological Sequences*, Cambridge University Press, 2002.

[16] F. O'Neill, *The Dance Music of Ireland – 1001 Gems*, Chicago, USA: 1907.

[17] B. Breathnach, "Ceol Rince na hÉireann Cuid V [Dance Music of Ireland] Vol V," 1999.

[18] L. Prechelt and R. Typke, "An interface for melody input," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 8, 2001, pp. 133-149.

[19] G. Dyke and P. Rosen, "abcjs - Project Hosting on Google Code" Available: http://code.google.com/p/abcjs/.

[20] S. Long, "Film body makes case against bord snips," *The Irish Times*, Aug. 2009.

[21] H. O'Shea, "Getting to the Heart of the Music: Idealizing Musical Community and Irish Traditional Music Sessions," *Journal of the Society for Musicology in Ireland*, vol. 2, 2006, p. 1.

[22] Shazam, "Shazam - The amazing music discovery engine. Join our Community" Available: http://www.shazam.com/music/portal.

[23] S. Shlien, "The ABC Music project - abcMIDI" Available: http://abc.sourceforge.net/abcMIDI/.

[24] "fmod - interactive audio middleware" Available: http://www.fmod.org/.

[25] "The Arts World on iPhone," *The Sunday Times*, Mar. 2010.

[26] F. Vallely, *The Companion to Irish Traditional Music*, New York University Press, 1999.