Musicology for the masses: Situating new audio technologies for musicology and music education

Dan Stowell, Mathieu Barthet, Simon Dixon and Mark Sandler
Centre for Digital Music
Queen Mary University of London
{dan.stowell, mathieu.barthet, simon.dixon}@eecs.qmul.ac.uk

1. INTRODUCTION

Music Information Retrieval (MIR) is a rich and burgeoning field, studying the information that can automatically be extracted and inferred from music audio, symbolic representations and metadata. It brings computational approaches to bear on musical questions, and can help us analyse large databases of music, or single pieces in fine detail. Some commercial successes are based on MIR technology (such as the Shazam music recognition service), and our own Sonic Visualiser software is increasingly being used in a variety of disciplines to analyse audio content.[1]

In the *Musicology for the Masses* project, we wish to increase the benefit of MIR techniques to society, and to help develop MIR in ways that connect with new use cases in real-world contexts.[2] This requires that we work with user groups directly, adapting our approach and conceptual toolset to that of the user groups. Connecting with user communities in this way is not just a way to disseminate research outputs, but to bring fresh ideas and perspectives into the research process.

In this paper we outline the results of the first phase of the project, in which we conducted observation studies with the identified target user groups. These observations are described in further detail elsewhere [1, 3, 4]; the outcomes feed into technical developments which are underway to bring MIR technology into productive contact with these contexts.

We chose to use an ethnographic approach for our empirical studies, to elicit a rich *thick description* of the way technologies and music-related concepts relate to each other in each specific context. Such an approach involves extended periods of contact with the studied context. In contrast to focus group or survey techniques, it can elucidate some of the underlying, unspoken structure of people's everyday interactions, which is important when planning a technical intervention into these interactions.

---

[1] http://sonicvisualiser.org
[2] http://www.elec.qmul.ac.uk/digitalmusic/m4m/

Here we report some findings from our observation study in schools, followed by those from our observation study of musicologists. In each of the two studies, observations were conducted by one observer with notebook and pen over a period of 3–4 months, followed by thematic analysis [2].

2. SECONDARY SCHOOL MUSIC LESSONS

We conducted observations in two comprehensive secondary schools in London, in various types of music lesson. Our findings can be grouped into two categories: issues around music-related concepts, and issues of technology & modality.

2.1 Musical concepts

One can get an initial impression of which musical concepts are discussed in music lessons by looking at the relevant syllabus document; but this would not give a clear idea of (for example) which concepts are broadly uncontroversial, and which need to be negotiated from different perspectives, nor how the concepts relate to each other and to the class environment. For example, musical notes (in the Western scale) were a common concept but never a source of difficulty: not only did they have clear and agreed names, they were easily anchored to concrete reality such as a key on a keyboard.

Conversely, negotiation of rhythm generally occurred without a stable set of labels or ways to refer to different rhythms: rhythms generally were included in discussion only by acting them out – whether on a drum, by clapping, or vocally. Acting out rhythms is an important part of music education, but discussion can be impeded if there is no shared set of common terms used as shorthand. In our observations we found a general tendency for rhythm talk to be limited by this lack of names, sometimes causing confusion or difficulties in remembering which is which.

The concept of musical "genre" has been the subject of debate in the MIR research community, and so the theme of "genre talk" that emerged from the analysis is pertinent. Genre-type terms were observed in many conversations, used to navigate known and unknown music – both in curriculum-oriented conversation and more informal conversation about music that people like or dislike. Note that the labels do not form a compact or mutually-exclusive set of categories (unlike the "record shop" approach to genre); instead they function more like landmarks, having particular traits which can be discussed and compared against other genres. This observation feeds into current discussions in MIR about the use of genre labels vs. tags, or even the abandonment of such labels in favour of similarity metrics. It tells us that genre terms are useful to users, yet their meaning can vary according to context, rather than being absolute.

2.2 Technological modes of use

From our observations we found a strong pattern in the modes that teachers and students use to negotiate music-related concepts: they use a wide variety of modes, both digital and otherwise, in quick succession and often in parallel. The classroom is a rich environment in which a wide variety of resource types can be called upon instantly (posters, screens, the video sharing platform YouTube, Wikipedia, mime, demonstration...), without necessarily planning in advance. Students also spend the majority of their time in exploratory and/or creative interactions, rather than on strictly bounded tasks.

These observations have key implications for any planned technical intervention: a technology is unlikely to be much used unless it can support exploratory interactions, and can form part of ad-hoc combinations of resources – one should not design a technological "bubble", aiming to build the full educational experience into a single system, rather a component to contribute to the rich learning environment.

One of the most-used technologies in the classes observed was YouTube. YouTube's breadth of coverage appears to be what supports its thorough integration into classroom practice: students and teachers often searched in YouTube without having checked in advance they would find something relevant, and almost always found a video which satisfied them. We observed YouTube being used as a component of various different activities, from demonstrating an example of a Brazilian carnival to playing a backing track for students to play along to, and more besides.

3. MUSICOLOGISTS IN THE BRITISH LIBRARY

In our second study we observed musicologists working with classical music recordings at the British Library in London, over a three month period. We focused on eliciting some of the strategies used by musicologists to explore musical documents, and the interactions with music-related technologies during this process.

3.1 Listening practices

The listening process lay at the center of the study of the recordings. We found that the musicologists commonly alternated two distinct but complementary listening practices.

In the first listening practice, the analysis of the recordings was performed exclusively through aural observations, without distraction from visual or other information ("closed listening"). The second listening practice was multimodal and characterised by the use of various music-related documents (e.g. the biography of a composer, information on the recording) and music representations (e.g. scores, feature visualisations) while listening. This listening practice can be described as an active process, since it does not just consist of receiving musical information, but is on the contrary based on a set of multimodal interactions between the listeners and musical documents. The advantages of using multiple modalities were an increased access to meaning, uncovering the context of a recording and the intentions of composers, conductors, or performers, and better understanding of the perception of the music. We suggest that software designed for assisting musicologists in their analyses of recordings should be in line with their listening practices by supporting both closed and multimodal listening.

3.2 Visualisation and content analysis

Even though, as educated and expert listeners, musicologists were able to perceive extremely fine details, visualisation and computational analysis conveyed empirical evidence which helped them to confirm and prove aural observations (*"The tools on one hand, I don't need them, I could describe that, on the other hand I can't prove it. This tool [Sonic Visualiser] is allowing me to express that in some way it [the finding] is objective."*). Visualisations and quantitative data retrieved through signal measurements were helpful in discussing, interpreting, or proving hypotheses about qualitative data collected through aural observations. Furthermore, these analyses enabled systematic comparison of the musical expression of various performers in different musical pieces (e.g. measurement of the rate and extent of the vibrato on long sustained notes based on spectrogram analyses). Slowed playback (time-stretching) used synchronously with spectrograms led to explanations of expressive effects which could not be reached through aural observations alone (*"You can only hear the pitch aspect of the vibrato as an educated listener with no software or technology."*).

Many of our observations also show that the visualisation and listening processes affect each other (cross-modal effects). For example, the spectrogram helps to hear vibrato much better by showing variations in partials of tones, the slowed playback of a tone helps to uncover that the vibrato is not constant, while the spectrogram aids in understanding that the perception of accurate pitch comes from the fact that the note starts with minimal vibrato. One counterpart of feature visualisations is to deemphasise what cannot be seen straightforwardly (the quieter instruments can be outshone by the singer, for instance). This in part explains why although visualisation is useful, software for musicologists should also facilitate the "closed listening" mode.

4. CONCLUSIONS

In this paper we have outlined some of the findings that emerged from ethnographic study conducted in school music lessons and in the British Library. In both contexts, such an approach yielded detailed information about the everyday interactions that go to make up the users' interaction with each other and with musical concepts. This knowledge steers the technical interventions we are making in the subsequent phases of the project. Further, we have been able to feed this knowledge back to the MIR community so that the impact of the studies is not merely contained within the project, but should inform future developments in the field.

5. ACKNOWLEDGEMENTS

6. REFERENCES

[1] M. Barthet and S. Dixon. Ethnographic observations of musicologists at the British Library: implications for Music Information Retrieval. In *Proc ISMIR*, 2011.

[2] R. M. Emerson, R. I. Fretz, and L. L. Shaw. *Writing ethnographic fieldnotes*. Univ Chicago Press, 1995.

[3] D. Stowell and S. Dixon. MIR in school? Lessons from ethnographic observation of secondary school music classes. In *Proceedings of ISMIR*, 2011.

[4] D. Stowell and S. Dixon. Integrating informal music technologies in secondary school music lessons. Submitted.