Investigating the effectiveness of audio capture and integration with other resources to support student revision and review of classroom activities

Iain Stewart, Willie McKee

School of Engineering and Computing, Glasgow Caledonian University

Abstract

In this case study, the focus is on investigating the effectiveness of audio capture and integration with other resources to support student revision and review of classroom activities. At the time of writing the various techniques have been applied across two modules at Honours level. The audio is incorporated into video captures of PowerPoint presentations and podcasts and voice recognition allows the transcription and editing of the live lecture speech. The advantages and disadvantages of these different resources are evaluated from the student’s perspective as an aid to study and from the lecturer’s perspective in terms of the level of difficulty in implementing them, the specific hardware requirements and the impact they have on the lecturer’s performance.

1. Background

As lecturers teaching multimedia technology, we work with a wide range of media. We have recognised that it would be useful to integrate lecture capture, and in particular audio, from lectures into teaching and learning activity. Traditional podcasts have a number of limitations, particularly in areas where the lecture or presentation makes use of a range of visual resources. We were also interested in how effectively we could improve the accessibility of the spoken lecture or classroom session for deaf students. In this case study, we aim to investigate the effectiveness of audio capture and its integration with other resources to support student revision and review of classroom activities. It is worth noting that the students who were exposed to these activities were generally highly IT literate and used to accessing study resources in a range of media.

The project occupies the same “space” as the ASEL project\(^1\) in that audio is to be used to enhance the student learning experience. One of the points we had in mind all the way through the activity was the affordability of the techniques both financially and in terms of staff time.

2. Methodology

A number of different forms of audio capture systems were evaluated across a wide range of presentations (both in test situations and in live lectures), including directly wired, commercial radio microphones and Bluetooth microphones, as well as the built-in audio capability in camcorders. Audio was captured directly into PowerPoint, as a set of files associated with the presentation, and as a completely separate file for incorporation into several resources such as podcasts and videos. The captured audio has also been used for voice recognition to allow the transcription and editing of the live lecture speech into various text and captioning formats.

\(^1\) The ASEL project is being led by Will Stewart at the Universities of Bradford and project partners at the University of Hertfordshire. The ASEL project has developed, implemented and evaluated the use of audio within next generation technologies to support, enhance, and personalise the learner experience. http://aselproject.wordpress.com/
The audio was captured in a variety of ways:
- as the audio track of the video camera used to record the lecture
- as a separate audio track recorded via wired connection on the laptop used for the presentation
- as audio recorded via wireless headsets (both FM and Bluetooth)
- as audio embedded in the PowerPoint presentation
- as audio transcribed directly into text for those who could not hear.

The audio and other resources were then manipulated to develop materials for the students to use in their revision and study. This was normally made available via the Blackboard VLE.

3. Issues

One of the key elements was the usability of the equipment and whether it affected the presentation. Many staff simply do not wish to encumber themselves with extra equipment unless it is clearly advantageous to the students. Thus, experimenting to find a methodology which makes the recording as painless as possible is important and providing proof that the student experience is improved is necessary. Depending on which capture system was being applied, the lecturer would have different levels of setup requirements:
- for the capture of audio via the PowerPoint presentation there is a period of time required to set up the wired or wireless microphone and to make sure the audio files are recorded in the correct format. After this is done the lecturer can work as normal, with the one caveat that backtracking in the presentation overwrites any audio previously associated with the slide. This means that the lecture presentation has to be planned a little more carefully, and if it is expected that a slide will be referred to at several points throughout the presentation, it is best to include it multiple times
- with text transcription, there is the additional requirement to ensure that speech recognition software is running prior to the presentation starting
- if video capture is also required then there is an additional effort in setting up the video camera with a good viewing position to capture both the screen and the lecturer.

In general, for all of these setups it was found that it very quickly became habitual and that the extra setup time was minimal. The main issues that had to be overcome from the lecturer’s point of view relate to usability. For example:
- microphones which did not pick up enough to ensure a clear recording
- microphones which did not clip on to clothing securely
- head-mounted microphones which were uncomfortable
- wires which prevented easy movement (some presenters are very active)
- danger of damage (e.g. when a wire might pull a laptop off a stand)
- radio microphones which exhausted battery power quickly and switched off without warning. This caused several truncated recordings
- in the case of video recording, some lecturers may not feel comfortable about being filmed.

The views of the lecturing staff varied as outlined below:
- all staff who used the prototype felt that it was cumbersome. This has now been improved by the purchase and use of higher quality, lighter and longer-lasting wireless headsets and a significant simplification of the methodology and applications, based on much trial and error
- concerns over the effect of recording on the lecture delivery were addressed by setting up several experimental presentations
- in the case of the transcription of audio for deaf students, five different presenters utilised the application with a prepared and practised presentation, and were asked to complete questionnaires as summarised in Table 1.
Overall, the response to the pilot presentations was very positive. All presenters felt that they could handle the extra activity without significantly stunting their performance and that the effect on the students was worth the extra effort.

For the ‘audio capture to subtitles’ application, the training time required for effective translation was considered to be too great in some cases, although newer versions of the recognition software (such as Dragon V.10) are showing significantly shorter training times, even for those with strong accents.

For the voice recognition system, the accuracy of the captured speech is critical to its effectiveness. In an ideal environment with an active noise cancelling microphone and a system trained to a user’s voice, accuracy levels of 95% are possible. In the real world of lecture rooms, with poorer quality microphones, higher levels of background noise and varying room acoustics, there is a real issue in trying to maintain the accuracy of the speech transcription. As speech recognition is based on recognition of phrases, the loss of a word can result in the system losing the sense of the sentence.

Glass et al. [1] found that, when using a standard, unidirectional microphone attached to a video camera, errors of the order of 30% to 40% were to be expected. These expectations are realistic as the quality and omnidirectional nature of the microphone make it unsuited for such work. Wald [2] found the benefit of using a good quality dedicated noise cancelling headset microphone caused error rates to drop to 10 to 15% in the classroom environment and in a quiet office environment the error rate dropped to approximately 5%. In addition, the level of training of the voice recognition software for the speaker’s voice profile, as well as ensuring that the system is set up to recognise unfamiliar technical terms, has a significant impact on the performance of the system. It has been our experience that the recognition quality in such circumstances can be extremely high.

4. Benefits

The integration of digital audio with the PowerPoint presentation was found to be particularly effective. Informal discussions with the students identified that the ability to navigate directly to the relevant slide(s) within the presentation and to review the discussion was beneficial to study and review. It was noted that if the tutorial solutions were linked to the slide locations then the effectiveness of the presentation as a study tool was enhanced. It was also noted however that while the text cues were very useful and linked to the audio, in many cases the students found that it was not possible to relate the audio to the exact element of the presentation. In such circumstances the students found the video to be a more useful resource as it also captured gesture as well as audio.

It was also found that the knowledge that they were to be recorded added a certain amount of rigour to the presenters’ “performance”. It was clearly noticeable that they were better prepared, spoke in a more measured and clear manner and tended to have fewer unintended diversions from the subject at hand. However, none of the presenters felt that using this level of concentration detracted from their presentation. In fact, several commented that it encouraged them to perform at a higher standard and to note the outcome of the recordings to aid reflection upon both lecture content and personal skills.

Table 1. Lecturer questionnaire

| 1. I could easily start the application |
| 2. I could easily start the VR package to get subtitles |
| 3. I could easily select ADLIB mode |
| 4. I could easily start the presentation beside TalkShow |
| 5. I could easily change to AUTOCUE mode |
| 6. I could easily select a prepared AUTOCUE subtitles file |
| 7. I could easily choose the “step” mode (manual or RUN) |
| 8. I could easily choose the display mode and timings |
| 9. I could easily edit the ADLIB and AUTOCUE text |
| 10. I could easily close Talkshow and save the files |
5. Evidence of success

Figures 1 and 2 show the output of the lecturer and student questionnaires. It can be seen that almost no negative points were raised and thus, after the experimentation was complete, the lecturers found the recording mechanisms easy to use. The students’ responses are especially positive. Key issues identified on the effect on the presentation were:

- it did provide the intended extra “channel of communication”
- few other members of the audience felt inconvenienced
- many other members of the audience felt it was helpful
- staff members found it was generally easy to use (after practice).

![Figure 1. Lecturer responses (questions listed in table 1)](image)

<table>
<thead>
<tr>
<th>Question</th>
<th>Average Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The recording had no impact on the presentation.</td>
<td></td>
</tr>
<tr>
<td>2. I have reviewed the stored lecture.</td>
<td></td>
</tr>
<tr>
<td>3. The lecture recording played better when downloaded.</td>
<td></td>
</tr>
<tr>
<td>4. The quality of the stored lecture is sufficient.</td>
<td></td>
</tr>
<tr>
<td>5. I felt it useful that the lecture had been recorded.</td>
<td></td>
</tr>
<tr>
<td>6. I have found the recording useful for study.</td>
<td></td>
</tr>
<tr>
<td>7. The recording helped understand the presenter’s language.</td>
<td></td>
</tr>
<tr>
<td>8. The recording helped me understand the content.</td>
<td></td>
</tr>
<tr>
<td>9. The subtitles helped me understand the content.</td>
<td></td>
</tr>
<tr>
<td>10. I would like more lectures to be recorded.</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Student questionnaire
As the use of the audio linked to the presentations is intended to support revision and review, a further series of interviews with the students will take place just prior to their exams. This will be carried out during a revision class, the aim being to gain their perspective after having applied the resources to support their revision and study.

6. How can other academics reproduce this?

A number of different approaches have been covered in this case study. At the simplest level, incorporation of audio capture into PowerPoint presentations requires little modification of the lecturer’s normal activity (obviously assuming that a slide-based presentation is the core of the lecture activity). By simply taking advantage of the built-in capture tools in PowerPoint, a simple captured lecture can very quickly be made “on the fly”. There are a few issues to watch out for, for example, audio quality and the need to ensure that the file is saved correctly as well as the limitation on backtracking (as this overwrites the previously recorded audio).

There are other technical issues which need to be addressed, such as the need for easy set-up of the microphone systems to make the recording process much smoother. Ultimately, however, this simple system requires little more effort than a “normal” slide based presentation, yet adds significant extra value to the student learning experience.

At the other extreme, with the full use of speech capture and transcription, there are a number of more significant overheads. Firstly there is the need to train the software to the lecturer’s voice model. This requirement has reduced significantly in recent years but still remains. Secondly there is the need to speak more clearly and to avoid unstructured or colloquial speech as this tends to throw the recognition tool off track and result in poor transcription quality. There is a recognised effect in that people using the tool feel that they are dictating and do not speak as naturally as they would do normally. This effect wears off over time as experience with the system grows. Having said that, it should be incumbent for a lecturer to ensure that they speak clearly and use comprehensible English. The need for a high quality active noise cancelling microphone or headset is much more important as the quality of the overall transcription is directly dependent on the quality of the captured audio.

7. Reflections

One of the benefits of associating audio with the presentation slides is that it provides a simple mechanism to navigate through the audio. As the slides present visual cues to the content of the audio, navigating to a particular slide associated with a particular topic means that the student can jump straight into the audio section that they require without needing to scan through a large audio file to find the element of interest. This also allows the audio to be linked into tutorial activities more effectively as it is possible to link a particular tutorial question to a set of slides and hence the audio associated with that section.

Based on feedback from staff and students, the following key points were identified as being essential to the effective use of this technique: a “one size fits all” approach to the capture and presentation of lecture materials does not allow for the wide variety of learning styles and strategies that students use.
in their learning [3]. The challenge remains to develop a structured yet adaptable framework for the lecture content that will allow easy navigation by the student.

8. References

