

THE DEVELOPMENT OF AN ONLINE COURSE IN DSP EARTRAINING

*Øyvind Brandtsegg,
Sigurd Saue,
John Pål Inderberg,*
Department of Music
Norwegian University of
Science and Technology (NTNU)
Trondheim, Norway
oyvind.brandtsegg@ntnu.no

Axel Tidemann,
Dept. of Computer and Information Science
NTNU
Trondheim, Norway
tidemann@idi.ntnu.no

Jan Tro,
Dept. of Electronics and
Telecommunications
NTNU
Trondheim, Norway
jan.tro@iet.ntnu.no

Victor Lazzarini,
Department of Music
National University of Ireland Maynooth
Maynooth, Ireland
victor.lazzarini@nuim.ie

Håkon Kvidal,
Norwegian Academy of Music
Oslo, Norway
hakon.kvidal@nmh.no

*Jøran Rudi,
Notto J. W. Thelle,*
Norwegian center for technology
in music and the arts (NOTAM)
Oslo, Norway
joran.rudi@notam02.no
sirnotto@notam02.no

ABSTRACT

The authors present a collaborative effort on establishing an online course in DSP eartraining. The paper reports from a preliminary workshop that covered a large range of topics such as eartraining in music education, terminology for sound characterization, e-learning, automated tutoring, DSP techniques, music examples and audio programming. An initial design of the web application is presented as a rich content database with flexible views to allow customized online presentations. Technical risks have already been mitigated through prototyping.

1. INTRODUCTION

The subject of Digital Signal Processing (DSP) is traditionally taught with emphasis on the mathematical foundation for signal processing. As a consequence of modern methods of music production musicians and music producers utilize DSP in their daily work. They usually need a more practical understanding of the subject. Typically these groups work intuitively with what they can relate to through hearing, and eartraining is an important component of any musician's education. Traditional forms of eartraining relate to musical melodies, intervals and chords. Within DSP, eartraining has no tradition, but we propose that it can be approached in a similar fashion as one would approach eartraining in music. Especially, the way eartraining is taught within jazz, where there is a focus on internalizing the knowledge

to a degree where it can be used spontaneously in improvised music performance.

On this background, NTNU has started a course in DSP eartraining as part of the master's degree in music technology. The students shall learn to aurally recognize common signal processing techniques. This is combined with a practical implementation of audio effects utilizing the same techniques. Hence the course seeks to combine "hearing" and "doing".

1.1. Goal

The current project aims to create an online course for DSP eartraining. The digital course material can be used for independent study or qualify for attendance at an on-site exam at the host institution (NTNU). The material will also form the basis for a course taught to master students at NTNU, and relevant parts of the material will be used in teaching at the partner institutions. The course language for the digital version is English.

1.2. Method

In the first phase of the project, a sketch of the course material was made with a simplified technical implementation. This material was used as basis for a workshop where the pedagogical methods and technical content was discussed and refined. Experts from a wide range of relevant fields were invited to the workshop, including eartraining, jazz education, acoustics, signal

processing, musicology, artificial intelligence, programming and music production.

In this paper, we report the main findings of the workshop. A revised version of the course material is being made based on the results of the workshop, and will finally be subjected to critical evaluation by all partners in the project. The course material will be tested in practical use through teaching at NTNU before a final version is settled on.

2. METHODS FOR EARTRAINING

Part of the background for the idea of this course is rooted in the desire to enable an intuitive and immediate relationship between the student and the techniques at hand. This immediacy is thought to enable a fluent, creative workflow in productions utilizing the techniques studied. A similar motivation is found in the use of deep learning strategies in the jazz education program at NTNU. This method of eartraining has been practiced for approximately 30 years at this institution and it is heavily based on physical interaction with the material to be learned, for instance through clapping and singing. This internalization of the material is considered crucial, and the “learning by doing” approach enforces deep learning. Since it is difficult to “sing” digital processing routines, other methods must be developed to internalize the knowledge inherent to the subject. As a step in the process of developing a methodology for eartraining in the subject of DSP techniques, the workshop included a session on eartraining in the context of jazz improvisation. As a courtesy to the reader unfamiliar with jazz education, the rest of this section may serve as a glimpse into it.

One of the historically basic tenets of jazz has been the development of each musician’s individual voice. If our goal is the development of individual voices there should be as many different voices as there are players. We try to show to the students how we can use small parts of the jazz history to develop musical individuality. We imitate history, not to create copies but to establish a fertile ground for individual development.

We experience our musical environment in very different ways depending on psychological, physiological, social and cultural prerequisites. The process of learning how to play is rarely starting out with a clear conception of how we want to play. It’s a process of self-discovery through concentrating on parts of our musical environment. It’s a work of trial and error, trying out different musical ideas and concepts to discover our own individual voices. Quite often it’s a matter of finding out “where we don’t want to be” through a process of elimination. To succeed in making those musical choices we try to embody our internal ear through imitation of what each individual (student, musician, producer, etc.) considers good music. This imitation could include musical time and rhythm, form, beat, space, colour of sound, melodic lines and modules, scale forms, vertical harmonic changes, tone modulation, attack, incidental noise and so on. The internalization of all those musical parameters is experienced by using the voice, feet tapping, hand clapping and dancing.

The music is taught by demonstration and by copying the master’s playing. The student is copying *what* the master is playing and *how* he is playing the music. This methodology involves the student directly in the sound and the feel of the music. This process develops the student’s self-confidence and stimulates the student to make individual choices. We are learning the rules of how to play good music more on an intuitive than an intellectual level.

We have different ways of experiencing the “correct rules”. Our individual search for ways to organize our musical source material can at the same time stimulate musical communication and interplay on complex levels. We are jointly contributing to our fellow musician’s musical awareness, and are as such aiming for a common *aural awareness*. Aural awareness, as discussed by Pratt [1] may constitute yet another method for reflection on and training of the abilities sought for in the DSP course. The aim for eartraining as we see it is a process starting with the different personal perceptions of an aural phenomenon, converging towards a common and objective understanding of the phenomenon and the theory involved, and then using this common ground as basis for development of a personal style of expression.

3. TERMINOLOGY

As part of learning *how each processing technique sounds*, it is pertinent to try to describe in words the different qualities of a sound. Verbal description of sound characteristics has always been a challenging task especially when it comes to quantitative definitions based on mathematical descriptions. The task increases in complexity when we have to accept that the same adjective may be used in different contexts: for describing a single instrument tone, the sounding concert hall, a complex synthesized sound and an artificial reverb. On the acoustical scene the term *warm sound* is often described as *having a lot of energy in the lower frequency region* without stating exactly where and how much. This may work as a global quasi-quantitative term as it could be used in all four contexts, and relates to resonance in a certain frequency area. Other adjectives like *edgy*, *biting*, *reedy* and *pure* may not have the same kind of global significance.

If we consider a distinction between sounds with a clear tone sensation (stationary frequency, perceived pitch, imagination of an acoustical instrument) and sounds without this sensation (noise, synthesized sounds, electronically processed sounds), we may consider Zwicker’s *roughness (rauhigkeit)* [2] with its contrary attribute *smoothness* as an important timbre space factor in the description of the tone sensation. This effect is mathematically described based on modulation by single frequencies with a transition area strongly related to the hearing mechanism’s *critical bandwidth*. A detailed and comprehensive discussion is given in [3]. Even without a mathematical foundation it seems tempting to use this *roughness-smoothness* sensation in the description of non-stationary synthesized sounds as well.

Some examples of sound descriptions and tests of subjective attributes are shown in the following.

3.1. Single instrument tones

Single instrument tones have been thoroughly examined mainly with the focus of defining timbre factors in a multi-dimensional timbre space. Kendall and Carterette [4] interpreted a two-dimensional timbral domain for wind instruments as having a principal dimension of *nasality* versus *richness* and a secondary dimension of *reediness* versus *brilliance*. In acoustical terms these adjectives relates to specific spectral qualities (*richness* of partials, *brilliance* with lot of upper harmonics) except for *reediness* that maybe more related to the excitation process with reeds. An example list of 61 adjectives can be found in the appendix of [4].

3.2. Room acoustics

After the ingenious introduction of reverberation time (RT) by Sabine [5], room acoustic parameters have been developed and

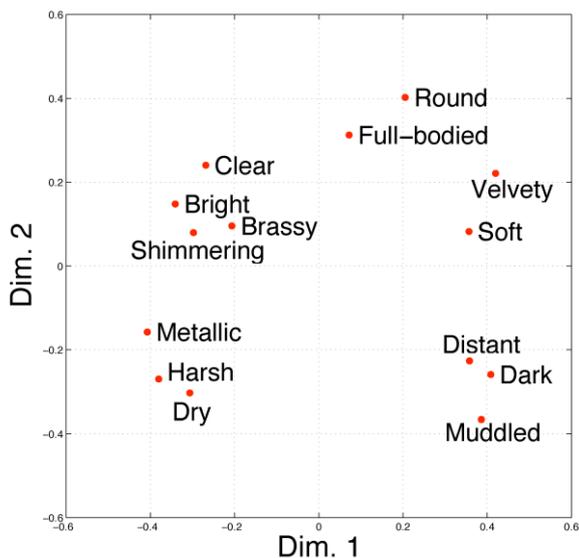


Figure 1: Dimension 1 and 2 of a multidimensional semantic space of piano timbre descriptors (from [8] with permission from the author).

been strongly based on physical parameters. Important room measuring parameters are defined by international standardization [6] and include the five acoustic quantities *sound strength* (subjective level of sound), *early decay time* (perceived reverberance), *clarity/definition/center time* (perceived clarity of sound), *early lateral energy fraction* (apparent source width) and *late lateral sound level* (listener envelopment). A detailed and extended terminology description for evaluation of concert hall measurements can be found in [7].

3.3. Performer’s experience.

As part of his PhD research, Bernays studied how piano performers can control timbre nuances [8]. This study includes mean evaluation of familiarity with 14 selected piano timbre verbal descriptors (in descending order *soft, bright, round, clear, harsh, dry, dark, full, velvety, metallic, shim, distant, brassy* and *muddled*). Five terms to best describe the whole semantic space are *bright, dark, dry, round* and *velvety*. The relations to acoustical characteristics are not defined even if his analyzed dimension 1 can be interpreted as an inverse frequency scale (see Figure 1).

3.4. Parameter orthogonality

Is it possible to analyze and treat musical features (pitch, timbre, dynamics, rhythm, etc.) as independent factors?

Referring to Houtsma [9] in music-related studies timbre has always been treated as a multidimensional continuum in which any point is potentially meaningful. It has been established by rating and multidimensional scaling techniques that the space can be adequately described in four subjective dimensions (dull-sharp, compact-scattered, colorful-colorless and full-empty) which are linked to physical dimensions such as spectral energy

distribution, amount of high-frequency energy in the attack, and amount of synchronicity high-harmonic transients.

Houtsma is concluding by stating “because of their subjective nature, the parameters pitch and timbre should never be presented as independent variables in perception studies. Doing so would amount to describing one unknown in terms of other unknowns”.

Stepanek is stating: “Musicians internal imagination of timbre supports orthogonal dimensions, but their saliency or relationship in real sounds is sound context dependent (for example depends on pitch or type of the instrument — violin, organ, etc.)” [10].

3.5. Amplified effects

A study by Dempwolf et.al [11] presents the results of a listening test employing eight attributes for the description of the perceivable timbral changes caused by effect units and amplifiers for electric guitars. Eight attributes (*aggressive, smooth, broken, fuzzy, crunchy, singing, warm* and *transparent*) were selected for listening tests. Appropriate terms to describe guitar distortion were *aggressive, smooth, warm, fuzzy, transparent,* and (partly) *broken*.

3.6. Concluding remarks on terminology

As a basic rule we have to make sure that adjectives will not be misinterpreted, i.e. we have to avoid terminology ambiguity.

If possible adjectives should be explained with common basic characteristics in the frequency and time domain. To clarify and pinpoint the use, each adjective should also be accompanied by at least two sound examples where it appropriately can be applied.

As available technology evolves the introduction of new adjectives should be acceptable for the music technology community as a whole.

4. E-LEARNING

Developing teaching material is about creating guidelines for learning activities that lay the foundation for new insight, skills and knowledge [12]. The learning material provides a base from which learners can venture forth on various learning journeys, and so the participants themselves make a good share of the final decisions about activities and even learning outcomes [13]. It’s the student’s own activity that fosters learning.

4.1. Science and humanities

Analyzing DSP by ear requires theoretical knowledge, auditory representations and skills to aurally diagnose the sound. It’s a combination of scientific, logical thinking and esthetical, sensory-based approach. The project is a prime example of music technology as an integration of science and humanities [14]. It is important to be aware of the two different approaches and balance them, as the logical thinking might determine the analytical process and displace the esthetical approach [15].



Figure 2: Prototype of automated tutor - determine which effect is applied.

4.2. Self-regulated learning

E-learning may provide the freedom to study when and wherever you like, but this is also one of the main challenges for participants. According to Whipp and Chiarelli [16] the students who self-regulate their learning and reflect on their own learning strategies are the ones who succeed. Such strategies may be daily logons, scheduling when to work with the course and submit assignments, coordination of off-line and online work and planning for technical problems. Acquiring knowledge through an online course requires self-discipline and learning material with a well-designed progression.

4.3. Motivation

Tasks that are challenging but manageable are crucial for motivation [17]. Finding such tasks is traditionally the teacher's responsibility but the lack of a teacher in calls for other ways of adapting the difficulty of the tasks to the skills of the participant. The interactive system monitors the participant's performance, and adjusts the progression accordingly. A fast learner will have a steeper learning curve and advance to more complex assignments while those who need more training get the opportunity to do more tasks on the same level. This is further discussed in the next section about the automated tutor.

4.4. Feedback

The learner gets immediate response to his answers which usually has a positive impact on learning. The objective of the tasks is however not to find the right button but to acquire strategies and methods for diagnosing a processed sound by ear. To do so you may need some guidance. In traditional eartraining this is taken care of in a communicative process involving a teacher. In our course this has to be safeguarded by the instructional texts.

Feedback fertilizes motivation and the fact that someone sees the learner by evaluating his or her performance usually boosts motivation. Personal feedback based on the learner's strengths and weaknesses is however hard to achieve in an automatic course. It may, however, be possible to develop a system that creates individual feedback based on the learner's performance

and that creates assignments and tasks with an adaptive difficulty.

Tests with beta versions of the learning materials show that the users experience some kind of a game-feeling. This usually fuels the motivation and the perceived playfulness contributes to a sense of achievement. This coincides with the findings of Lioa & Hsieh [18] which stated that e-learning may "increase students' satisfaction with learning and since satisfaction has a very important mediating role, students' performance would be expected to increase".

5. AUTOMATED TUTOR

As with virtually any subject matter, "learning by doing" is the best way for students to acquire new knowledge. Having a tutor to aid the student is a clear benefit, but for an online course this is more challenging due to resource and physical limitations. On the other hand, a course on DSP techniques is very well suited to the online domain, since DSP theory must eventually be implemented on a computer for the results to be heard. To utilize this advantage and overcome some of the limitations, we aim to create an intelligent automated tutor. The tutor's interaction with the student would be through generating tasks. The tutor can monitor the progress of the student and provide feedback on the strengths and weaknesses of the student. This allows the tutor to adapt to the student and generate individual tasks. For instance, if the student seems to have a good grasp of the course material, the tasks can become more difficult. Vice versa, if the student struggles to understand basic concepts, the tasks can become simpler. Student activity logs permit statistical analysis, so the course teacher (in the physical or virtual domain) could see which subjects should receive more educational focus.

As a proof of concept, a very simple adaptive tutor was implemented. The tutor had three tasks. Common for all the tasks was that the student should listen to a clean and a processed version of a source sound, and identify which processing had been applied to the sound. As a basis for the tasks, a set of audio files and audio effects were prepared. All effects included in the course are available for automatic generation of tasks. The source files are organized as a bank of clean instrument and vocal recordings. The selection of source sounds is done randomly for each listening task.

To encourage training, the listener would choose when to advance to the next type of task, allowing repeated work on similar tasks. The three task types were as follows:

- 1) *Determine which effect is applied.* The tutor applies a randomly selected effect to the chosen audio file. The listener is presented with 5 different choices of what effect it could be (see Figure 2). If the listener chooses the wrong effect, the list of effects is decremented, to make it easier to select the correct effect. If a correct choice is made, a new file and corresponding effect is selected, and a new task is generated.
- 2) *Determine the center frequency of a bandpass filter.* The tutor selects a center frequency of a bandpass filter (the Q is hard coded), and presents three choices to the listener¹. If the answer is wrong, the choices are decremented. If the answer is correct, the tutor randomly goes up or down in relation to the previous frequency, selects a new audio file and applies the bandpass filter with the new center frequency. The distance between the available options to the listener becomes narrower for each correct answer, making the task to determine the center frequency harder for each iteration.
- 3) *Determine the effects chain applied.* This task is a more adaptable version of task 1, where the tutor first asks the listener to identify which two effects are applied to a source sound, with 5 options. When a wrong answer is submitted, one option is removed. If the student answers correctly on the first attempt, the number of effects is incremented. If the listener needs several attempts to guess which effects are applied, the number of effects stays the same for the next task. However, if the listener cannot figure out which effects are applied, an effect is removed when generating another task. In other words, the task becomes simpler if the listener cannot determine which effects are applied, and harder if the listener chooses the right answer on the first attempt, i.e. shows a good grasp of which effects are applied.

To make this proof-of-concept implementation, a collection of DSP routines was assembled using Csound [19] as the synthesis and processing engine. Several of the authors are already familiar with this language and it was chosen as a common starting ground. As Csound is a text based language for specification of audio processing, it was possible to write a computer program in Python² that automatically generates Csound code based on a specification of effects and synthesis techniques needed in each listening task.

For more complex tasks, the challenge will be how effects can be combined in a meaningful way, to avoid concatenating effects with similar characteristics, e.g. a flanger and chorus effect. This also holds true for what range of parameters are eligible for each effect. We are currently looking into these issues.

6. DSP TECHNIQUES

As part of determining the focus and scope of the course under planning, it is relevant to look at the selection of specific DSP techniques and decide which techniques would *have* to be covered, and what could possibly be left out, either for clarity of presentation or due to resource limitations. A Master's level DSP

training course taught to music technology students at NTNU forms a basis for the online course under development. In this course the following subjects are currently included: A brief introduction to listening for specific frequency bands and an aural awareness of the differences between classic waveform shapes like sine, triangle, sawtooth and square waves. Basic sample manipulation techniques (downsampling, bit reduction) and panning, time varying delay based effects (chorus, flanger, phaser), modulation techniques (FM, AM, RM), distortion and waveshaping, basic filter theory, reverb, physical modeling (waveguides), granular synthesis and processing, convolution and effects in the frequency domain. The focus is on aural awareness of the effect of each processing method, and the students are also encouraged to use the DSP templates in designing their own effects for practical use in productions. For this purpose, a simple VST effect wrapper for DSP code is introduced.

During the workshop a number of effects were noted missing from the curriculum, from the point of view of the studio practitioner. These were identified as EQ (as a practical application of the filter theory already in the curriculum), dynamics processing, envelopes and envelope followers, pitch detection, and the vocoder (as a practical example combining envelope followers and filters).

Further, it was discussed if the course should cover effects processing only, or also include synthesis techniques. Digital signal processing comprises a wide variety of different techniques, most of which modulate signals that are input externally. We are, however, also dealing with signals generated internally. Algorithmically generated signals are generally defined as being synthesized. The distinction between audio synthesis and audio processing is not always trivial to make. Almost exclusively, various processing techniques are integrated in synthesizers in order to make the generated signals manageable and interesting. In other words, there is a gray area between synthesis and processing which is not easily defined. As a rule of thumb, one can say that synthesis involves algorithmically generated sounds, as opposed to sounds that are processed. Synthesizers do not need audio inputs—processors always do.

The basic consensus regarding the question of *what to include* was that a broad and balanced survey of processing techniques is needed for the course to provide a stable platform for further development for the student. As the online course model can provide selectable alternate and flexible routes through the curriculum, the project progresses with the intention of covering all of the above techniques. Specifically adapted courses may be constructed using the same corpus of material, and the current Master's level course at NTNU can be seen as one such route.

7. MUSIC EXAMPLES

In order to augment the musical relevancy of the course, the content will be supported by music examples from different genres, ranging from music where technology is used solely for production purposes, to genres where composition process and sonic results are fully technology dependent. Students will find material to which they have an affinity, and this will increase their learning. Genres should include for example pop/rock, hip hop, techno, metal, jazz, as well as electronica and electroacoustic music. We aim to include larger sections of pieces and songs, and discuss what has been achieved musically through the signal processing. The experienced relevancy of the course will be low if the selection is poor in either clearly audible signal processing or experienced musical relevance.

¹ Similar to Harman International's «How to listen»-system: <http://www.harmanhowtolisten.blogspot.com/>

² Python Programming Language – Official Website: <http://www.python.org/>

In sum, the music selection is crosscut by several considerations, and needs to focus on a wide area of musical inclinations. Finding clear examples of music where several types of effects have been chained together is a challenge, and it is likely that we would have to compose a number of these examples especially for this project.

Ideally, the learning situation should encourage students to go beyond mere recognition of DSP techniques, to also encourage them to test their theory by recreating musical elements from examples, and to add their own creativity in the process, in a type of action-learning tied to actual music content. This would entail changing parameter settings in DSP algorithms, and being able to hear their changes and how they are included in and affects the total musical context. In order to realize this, a model with several musical “tracks” is needed, of the same type as the method appearing at the website of Cambridge Music Technology³. However, time restraints makes it difficult to realize this ambition during the first development run of the course, and it will be necessary to first focus on having the students correctly identify signal processing methods in commonly known music. An example could be the type of variable filter center frequency that Jimi Hendrix used in some of his guitar playing. True, this example is from the analog domain, but the DSP implementation of this effect is common.

8. AUDIO PROGRAMMING

In the process of developing the infrastructure for this course, one important aspect was to provide interaction in the algorithm examples. This connects to the “learning by doing” strategy deemed as a basic principle for this course. It was decided that the best way to provide these was to embed a music programming system on which the different effects could be realized. For these, it was clear that an open source solution was required, as (a) integration with the online component would be facilitated, (b) correctness of code could be verified (if some doubt arises) and (c) no licensing costs would be incurred.

The choice of such a system would depend on a number of factors:

1. The software should be capable of implementing all of the DSP examples used in the course.
2. The system language should provide a clear translation of the algorithms and flowcharts.
3. The general aspects of programming such language should not involve a too steep learning curve.
4. There should be clear routes for integrating such system to implement the whole set of examples in an online environment.

Of all the various open-source solutions, three stood out as clear possibilities: Csound [19], Pure Data (PD) [21] and Faust [22]. These were the ones that initially more or less matched the criteria above, especially item 2, which was where other languages were deemed inappropriate.

Each one of these systems represented a different approach: Csound is a mature language of the MUSIC N paradigm, with a host Application Programming Interface (API) that could simplify integration; Pure Data is a graphical programming language that could perhaps facilitate flowchart translation; Faust is a rela-

tively newer language, purely functional and designed for DSP programming.

After a review of the features of the three languages, it was decided that Csound was the better match for the implementation of the course examples. Pure Data, while apparently simpler for straightforward flowcharts, presented difficulties for more involved examples. Deployment of PD-based examples would depend on the libpd, a separate project that adds an API layer to the software. The uncertainties involving its integration via libpd were considered another risk factor in deployment. These were concentrated into two points: how to use libpd in the implementation and whether libpd, as a separate project, would be kept in sync with the development of the PD system/language.

Faust turns out to be an interesting possibility from the implementation point of view, especially now with a HTML5/Javascript output being developed for its compiler. However, it involves a somewhat steeper learning curve for students. There are also question marks on its capabilities to provide all course examples; for instance, spectral processing is not available in the current released version of Faust.

Csound, on the other hand supports well all of the requested criteria. The test implementation of the automatic tutor also showed that automatic generation of Csound code could be done in a straightforward manner. With regards to point 4, integration/deployment, a Csound engine can be delivered as a native module to Java applications deployed via the Java Web Start (JAWS) mechanism and the Java sound API. A proof-of-concept prototype was implemented, deploying a small Csound Java application via an internet link to computers without any prior Csound installation.

The plans for deployment involve a small Java application for each example with user-interface controls such as sliders and buttons, providing both sound control interaction and code examples.

Based on these considerations, the first version of the online course will utilize Csound, but there is nothing preventing the number of supported languages to be expanded later on. A modular structure (as described below) allows for generating new “views” of the course material where other programming languages can be used to demonstrate each processing technique.

9. PUTTING THE PIECES TOGETHER

The online DSP eartraining course should provide multiple views on the topics it covers: text, figures, sound, source code and so on. All this material needs organization. We would also like to offer the option to customize the course to fit a certain target audience. A teacher using the course material might want to focus on just a subset of topics and perhaps avoid math or source code. A flexible course organization should support topic selection and optional elements. Figure 3 provides an overview of how the course is organized at different levels.

A key concept is the *lesson*. It represents a single topic such as for instance “Flanger” or “Bit reduction”. It typically contains text blocks, figures, signal plots, mathematical formulas, sound examples and code examples. The elements within a lesson are considered separate units and can be arranged in various layouts. Advanced text blocks or code examples can be tagged as optional and only displayed if switched on as global course preferences.

Lessons are grouped in *modules* representing a natural higher level of organization, such as “Time-varying delay effects” containing lessons on “Flanger”, “Chorus”, “Echo”, and so on. The complete course package holds a large and expandable set of

³ Mixing Secrets For The Small Studio – additional resources: <http://www.cambridge-mt.com/ms-mtk.htm>

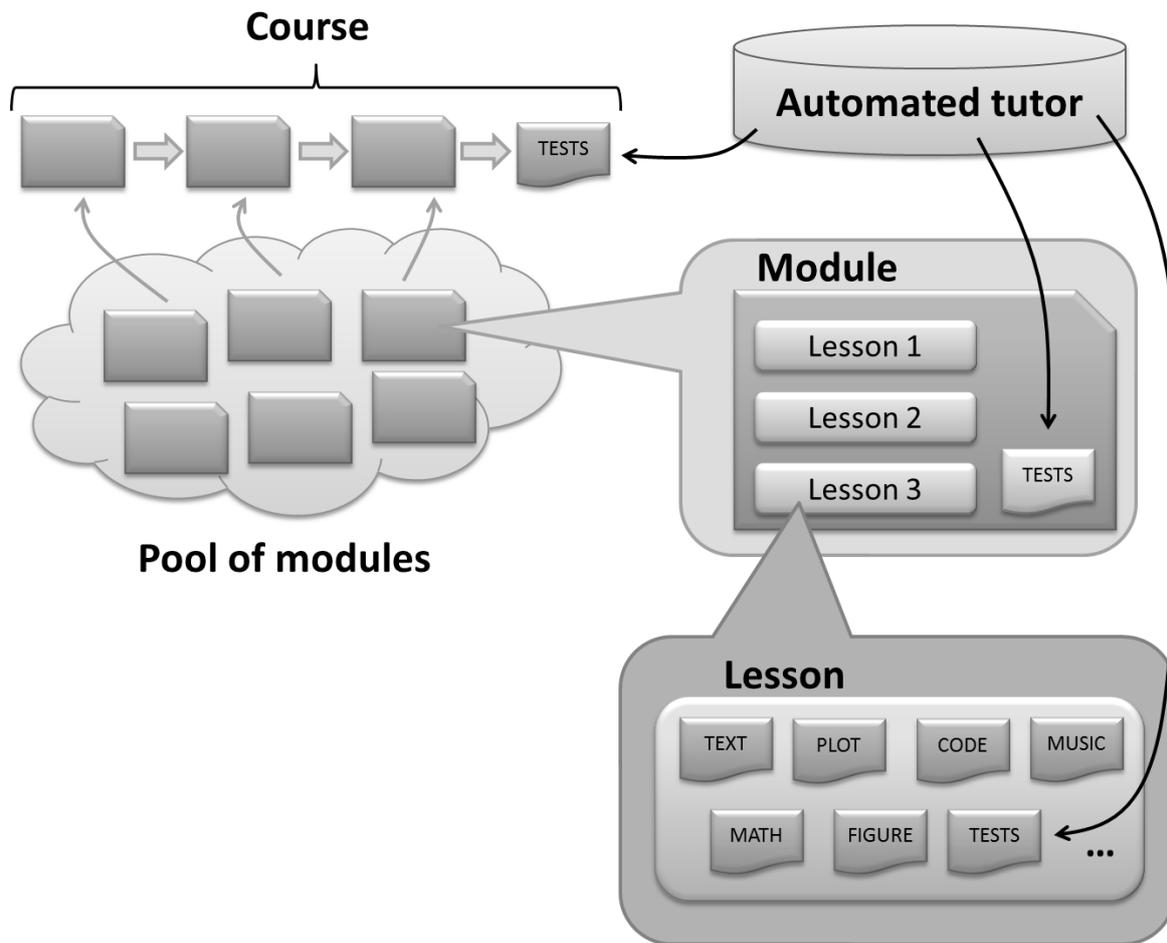


Figure 3: A schematic overview of the course design. A course is basically an adaptable selection of high-level modules. Each module is subdivided into lessons containing a variety of learning elements (text, plots, code, sound, etc.). The automated tutor can provide a large number of interactive tests at any level of the course.

modules. When designing a specific course a teacher may pick any number of modules. Hence the concept of *course* here represents a particular path through the set of modules.

Listening tests, music examples, literature and web references are linked into each lesson, but should also be separately available for summary and repetition. The automated tutor plays a particularly important role: It organizes listening tests and reference matching tasks in a separate pool. *Reference matching* in this context means tuning DSP effects on a music track to match a processed reference. It comes in two flavors, as a sandbox to play around with effects and effect parameters, and as a test to check the student's ability to find and configure the effects used on a music example. A Csound-based DSP widget will be used to enable this feature.

All tests should be properly tagged to support flexible querying: For instance all tests related to a specific lesson, a random selection of tests within a module, or tests from all modules that the student finds difficult. The latter requires that the system keeps track of user history including course progress and test performance. This is part of the automated tutor concept. In addition the system must have some sort of user management to allow

registration, user logging and forums. Interfacing with social media such as Facebook and SoundCloud is also on the feature list.

The requirements listed above add up to a quite complex web application. We considered various alternatives for implementing the online course. Content management systems (CMS) such as WordPress, Joomla and Drupal⁴ have large user communities and well-designed templates to get you started. However, the ease-of-use quickly deteriorates when trying to implement complex applications. Learning management systems (LMS) like It's learning and Fronter⁵ were also considered, but they enforce too much structure and are not suitable for our purposes.

We have chosen Django, a high-level Python framework for building web applications [20]. It is object-oriented and flexible. In fact, Django allows you to build your own custom content management system from scratch. The initial learning curve is steeper, but it has great support for modularity and adaptability that should pay off in the long run.

⁴ Content management systems (CMS): Wordpress (<http://wordpress.org/>), Joomla (<http://www.joomla.org/>), Drupal (<http://drupal.org/>).

⁵ Learning management systems (LMS): It's learning (<http://www.itslearning.eu/>), Fronter (<http://com.fronter.info/>).

The actual implementation involves constructing the course model outlined above, providing administrative interfaces for adding content, and designing various views and presentation templates. The views are in essence mechanisms for selecting model data based on user queries. Most of the work should be straightforward, but there are a few challenges that need to be resolved:

- Interfacing Csound: A prototype has already been tested, but awaits integration into the Django-based system.
- Streaming audio for music examples with options for time-line annotation (similar to SoundCloud⁶): Technology has not been determined yet.
- Integrating the automated tutor: The tutor incorporates both interactive tests and tracking of user performance. As shown earlier a prototype is already implemented, actually using Django as presentation platform.
- Supporting user customization, i.e. making course design available to teachers.
- Connecting to social media.

We already have a good grip on most of these challenges through prototyping, and feel confident that the course will be up and running by year end 2012. An early version should be ready for demonstration at the DAFx 2012 conference.

10. CONCLUSION

This paper reports from a workshop on DSP eartraining arranged as preparation of an online course on the subject. As should be evident the workshop covered a broad range of topics presented by specialists from many different fields. The discussions helped establish a strong foundation for continued work.

The immediate goal of creating an online course in DSP eartraining is definitely within reach, with only a few issues left for further elaboration. The production of course material is a continuous process, adding to an ever-growing database of modularized content pertinent to audio signal processing. The idea of providing multiple and customizable views to this very rich collection of material, enables us to restrain the presentation while the underlying content database can grow far beyond the bounds of the planned eartraining course.

This could be the first step towards an online, interactive knowledge repository covering audio signal processing from multiple viewpoints.

11. ACKNOWLEDGEMENTS

The authors wish to thank Norgesuniversitetet (Norway Opening Universities) for financial support.

12. REFERENCES

- [1] G. Pratt, *Aural Awareness: Principles and Practice*. 2nd ed. Oxford, UK: Oxford University Press, 1998.
- [2] E. Zwicker, *Psychoakustik*, Springer-Verlag, 1982.
- [3] E. Zwicker and H. Fastl, *Psychoacoustics*, 2. Aufl. Springer-Verlag, 1999.

- [4] R. Kendall og E. Carterette, "Verbal Attributes of Simultaneous Wind Instrument Timbres II," in *Music Perception*, vol. 10, no. 4, pp. 469-502, 1993.
- [5] W. C. Sabine, *Collected papers on Acoustics*, Harvard University Press, 1922
- [6] *ISO 3382-1: Acoustics - Measurements of room acoustics parameters: Part 1: Performance spaces*, International Organization for Standardization, 2009.
- [7] L. Beranek, "Subjective Rank Orderings and Acoustical Measurements for Fifty-Eight Concert Halls," in *Acta Acustica*, vol. 89, no.3, 2003.
- [8] M. Bernays, "Verbal expression of piano timbre. Multidimensional semantic space of adjectival descriptors," poster at International Symposium on Performance Science (ISPS), Toronto, 24 August 2011.
- [9] A.J.M. Houtsma, "Pitch and Timbre; Definition, Meaning and Use," in *New Music Research*, vol. 26, no. 2, 1997.
- [10] J. Stepanek, "Musical Sound Timbre; Verbal Description and Dimensions," in *Proc. of Digital Audio Effects (DAFx-06)*, Montreal, 18-20 September 2006.
- [11] K. Dempwolf et al., "Verbal Attributes Describing Distorted Electric Guitar Sound," in *Proc. of Digital Audio Effects (DAFx-10)*, Graz, 6-10 September 2010.
- [12] B.G. Wilson, "Designing E-learning environments for flexible activity and instruction" in *Educational Technology Research and Development*, vol. 52, no. 4, pp. 77-84, 2004.
- [13] J.R. Hill and M.J. Hannafin, "Teaching and learning in digital environments: The resurgence of resource-based learning," in *Educational Technology Research and Development*, vol. 49, no. 3, pp. 37-52, 2001.
- [14] H. Kvidal. (2009). *Undervisning i musikkteknologi – Kasusstudier av to undervisningsfag* [Online]. Available in Norwegian: http://www.musikkteknologi.no/undervisning_i_musikkteknologi
- [15] M. Knakkegaard, *IO - om musikkteknologi, musik og teknologi*, (in Danish), Odense: Odense Universitets Forlag, 1992.
- [16] J.L. Whipp and S. Chiarelli, "Self-regulation in a web-based course: A case study" in *Educational Technology Research and Development*, vol. 52, no. 4, pp. 5-21, 2004.
- [17] G. Imsen, *Elevers verden: innføring i pedagogisk psykologi*, 4th ed. (in Norwegian), Oslo: Universitetsforlaget, 2005.
- [18] P.-W. Liao and J.Y. Hsieh, "What Influences Internet-Based Learning?" in *Social Behavior and Personality*, vol. 39, no. 7, pp. 887-896, 2011.
- [19] R. Boulanger, Ed., *The Csound Book*. Cambridge, MA: MIT Press, 2000.
- [20] *The Django project*. Django Software Foundation [Online]. Available at <https://www.djangoproject.com/>
- [21] M. Puckette, *The theory and technique of computer music*, New York: World scientific Publ., 2007.
- [22] Y. Orlarey, D. Fober and S. Letz, "Faust: an Efficient Functional Approach to DSP programming" in *New Computational Paradigms for Computer Music*, G. Assayag and A. Gerzso, Eds., Éditions Delatour, 2009.

⁶ SoundCloud: <http://soundcloud.com/>