

Perspectives of Advanced Ear Training Using Audio Plug-Ins

PAVOL BREZINA, *AES Associate Member*
(pbrezina@ukf.sk)

Constantine the Philosopher University in Nitra, Nitra, Slovakia

Technical ear training is currently gaining more and more attention during the training of professional sound engineers. In recent years, a number of web applications and standalone programs have been created to provide a basic training interface for ear training. However, with the greater accessibility of audio plug-ins and the rising demand to use them in production, the question arises as to how these tools could be integrated into the ear training process. The aim of this article is to point out the specifics and possibilities of using audio plug-ins as training tools through the unique prototype of a proprietary, standalone host.

0 INTRODUCTION

Several scientific studies have highlighted the benefits of systematic ear training in technical ear training over three decades [1–6]. Nevertheless, there is still an insufficient number of complex training software solutions. Only gaining an acceptable level of experience through long-term practice as a sound engineer is not effective enough. The ability of sound engineers to assess the quality of a sound recording can only be most effectively developed through regular training [7], which is well grounded in research done by Quesnel and Woszczyk [8]. In addition, there must be made an improvement in the ear training system itself, and the methods used in the past should be adapted to the needs of the present. One kind of improvement should be the implementation of audio plug-ins as advanced training tools.

Considering the history of ear training, the comprehensive educational system created by Dave Moulton in the 1990s can be deemed groundbreaking. His publication *Golden Ears* [9] was one of the most successful methods of improving listening skills, also because it was available to the general public. The didactically oriented book publication was accompanied by eight CDs, which contained exercises focused on the frequency and dynamic spectrum, as well as reverberation and the delay of sound. The disadvantage of this type of training is that all changes in the examples are already fixed in the audio recording and the user does not have the opportunity to adapt the content of the training. Also problematic was the presence of im-

mutable musical demonstrations, which can become boring over time. However, given the possibilities of the technology of that time, it was a unique publication.

Software solutions in the form of standalone applications or websites are more complex and user friendly. The emergence of several portals offering adaptable exercises to improve listening skills has aroused an increased interest in ear training, not only from sound engineers but also from students of musically oriented disciplines, for whom nontechnical ear training is sufficient [10]. Among the first developers were Quesnel and Woszczyk with the *Timbral EarTrainer II* system, but this was not available to the public, and it was only tested within their own research [8]. A similar ear training tool that has institutional background was presented by Bassett and Martens [4]. Their software is focused on frequency ear training and is used in the SAE Institute.

On the other hand, special ear training software that is focused on dynamic ear training was presented by Martin, Massenburg, and King [11]. One of the first widely known and accessible ear training softwares was developed by the company *V-Plugs* [12], which later transformed its product *Mr. Soundman* into the extremely successful and globally popular *Soundgym* web interface [13]. In the field of standalone applications, there are products from *Harmann* [14] and *TrainYourEars* [15]. These software products are primarily focused on the field of ear training in the frequency spectrum. In practice, however, the sound engineer must take into account much broader sound attributes and elements. Based on the categorization of Corey [16], ear

training can be divided into five basic categories according to the nature of the sound attribute:

- Frequency spectrum analysis—Equalization
- Attributes of spatial perception—Delay and reverb
- Attributes of dynamic range—Compression/limiting and expansion
- Elements of sound recording quality—Distortion and noise
- Audio editing—Editing an audio recording

It is necessary to create specific exercises for the above categories, which also requires the creation of a specific graphic design of the training application. From the point of view of the versatility of a single software application that would perform exercises for all categories of audio attributes, this is an extremely demanding process, mainly because of the amount of programming work needed to create all necessary training modules for specific sound attributes and achieving a clear user interface. Probably also due to this facts, there is currently no such standalone application. Currently, the only comprehensive educational software is the Technical Ear Trainer web application (available at <https://www.webtet.net/apcl/#/>), created by Jason Corey.

The rest of this report is organized as follows. Sec. 1 describes the standardization of training means and possibilities of their extension. Sec. 2 introduces innovative methodology of implementation of ear training. Finally, Sec. 3 presents conclusions and visions of ear training.

1 STANDARDIZATION OF TRAINING MEANS AND POSSIBILITIES OF THEIR EXTENSION

Instruments for training listening skills are devices for adjusting the frequency spectrum, dynamics, spatiality, and sonic character. In the digital world, these are software modules, such as equalizer, compressor, expander, reverb, distortion, etc. The developers of software and web interfaces designed for ear training are based on certain standards. The advantage of such modules is that the change of sound is always without the intentional coloration of the audio signal, and training can take place very efficiently. Beginning sound engineers and students can more easily remember the nature of a given sound change.

In practice, however, sound engineers often encounter devices that have some kind of intentional coloration of the audio signal. It can be, for example, a specific equalization curve shape or a certain degree of harmonic distortion (a tube equalizer or compressor, for example). This is currently not only the case in recording studios, where hardware sound processors are dominant, but also in studios working only on a software platform. The era of digital software solutions has advanced to a level where more and more audio editing software vendors are bringing so-called hardware emulations. These emulations bring with them characteristic elements, such as specific equalizer frequency response curves of equalizers, compression curves of compressors, or specific distortion.

The existence of hardware emulations in the form of audio plug-ins opens up new possibilities in the postproduction of sound recordings and completely changes the normal view of a professional studio. Therefore, it is extremely important that those interested in using these new sophisticated audio plug-ins have the opportunity to conduct ear training through them. This method of ear training must be understood as an extension, and in no case should it replace basic ear training using training tools without the intentional coloration of the audio signal incorporated. For simplification, the following categorization is possible:

- Universal tools (without the intentional coloration of the audio signal incorporated)
- Dedicated tools (with the intentional coloration of the audio signal incorporated—the most common emulations of hardware devices)

1.1 Characteristics of Sonic Differences of Selected Universal and Dedicated Tools

Of all the sound-editing tools, the difference in the change in sonic character [17] on different equalizers and compressors is the easiest to demonstrate. This is true mainly because these tools are some of the most used in sound engineering, and their functionality is generally very well known.

The training of listening abilities through equalizers is aimed at acquiring the ability to quickly identify the frequency spectrum or a specific frequency on which a change has occurred. Universal tools are based on the preferred values given by the International Organization for Standardization (ISO) 266 standard [18] when defining the practiced frequencies. This is a breakdown that is defined as one-octave intervals (16 Hz, 31 Hz, 63 Hz, 125 Hz, 250 Hz, 500 Hz, 1,000 Hz, 2,000 Hz, 4,000 Hz, and 8,000 Hz). However, dedicated tools hardly have such a selection at all, and the choice of frequency grouping is often significantly different and focused on specific frequencies that manufacturers determine based on musical sound or spoken word editing requirements.

In addition to the differences in frequency distribution, there can also be fundamental differences between universal and dedicated equalizer tools in the form of an equalization curve. These differences can be caused for several reasons, and one of them is a specific curve shape. In this example, a dedicated tool simulating the behavior of the Solid State Logic equalizer curve was used. Fig. 1 shows the behavior of the bell-shaped curve for two different audio plug-ins. The settings are identical in both cases, but the graph shows deviations mainly in the shape of the curves and the location of the central frequency. This is slightly higher in the frequency spectrum in the case of a dedicated tool (audio plug-in).

Similar differences can be found and compared for different types of audio plug-ins and several types of equalization curves. For some dedicated tools emulating hardware equalizers, the presence of adjustable parameters may vary. An example is the ratio of bandwidth to center fre-

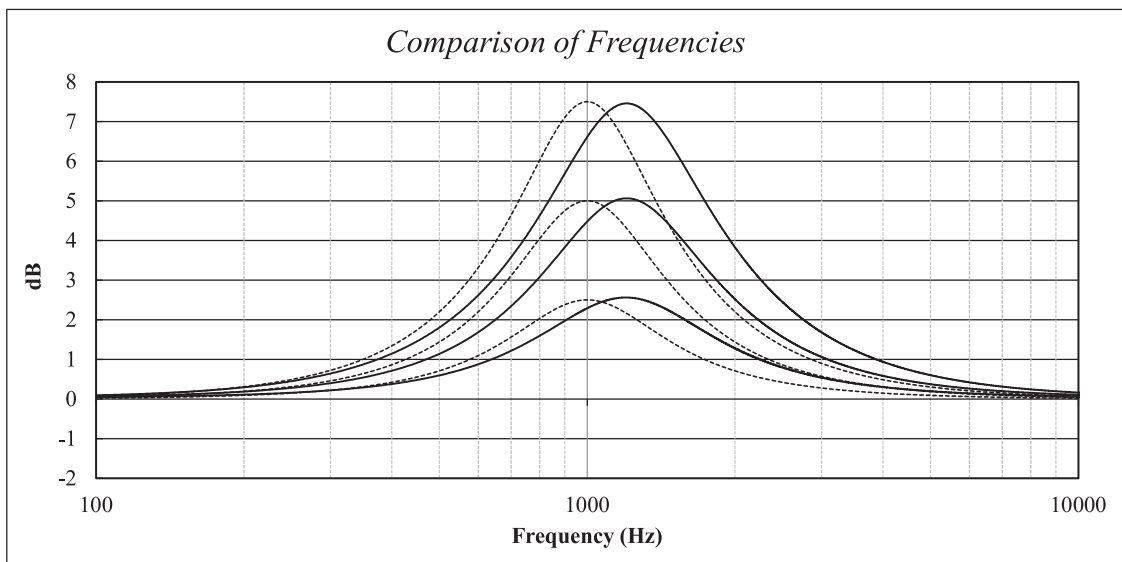


Fig. 1. Comparison of frequency curves of universal (dotted line) and dedicated (solid line) audio plug-in with identical settings—peak at 1,000 Hz; ratio of bandwidth to center frequency (Q) = 1.5; gain values are 2.5 dB, 5 dB, and 7.5 dB.

quency (Q), the behavior of which varies depending on whether the parameter is constant (constant- Q) or proportional (proportional- Q). Proportional- Q equalizers are almost exclusively of the hardware type, and in digital terms, they are mostly software emulations.

Fig. 2 shows a comparison of the curves of the universal and dedicated equalizer tools. The proportional behavior of the Q parameter with a dedicated equalizer is not only visible visually but also differs audibly. According to Pirkle, the design of the perfect constant- Q filter presupposes a lower error rate in the bass frequency band when adjusting the audio material [19]. On the other hand, a proportional- Q filter appears to be more natural and smoother in terms of processing musical sound material when the volume level changes [20]. Martin and Massenburg also point out that the issue of the Q parameter must be paid attention to during ear training in terms of resonance identification [21], which is especially true for hardware equalizer emulations.

These facts suggest that ear training through dedicated equalizers can help lead to a better and faster understanding of their behavior. In practice, the sound engineer can use the instrument more effectively, and the aural knowledge of the sound characteristics helps him or her to better choose these instruments for a particular situation.

A similar comparison of sonic character can be demonstrated for compressors. In addition to the standard function of adjusting the dynamic spectrum of an audio recording, they are often used to color the audio recording due to the characteristic saturation properties. This is also one of the reasons why the interest in their emulation in the form of audio plug-ins has been growing in recent years. Compared to universal software compressors, the sonic nature of dedicated audio plug-ins is richer and sound specific. The distortion–saturation comparison can be demonstrated by plotting curves through a total harmonic distortion (THD). Fig. 3 shows a comparison of the THD universal and ded-

icated audio plug-in across the entire frequency spectrum. A significant change in saturation is manifested from the limit of 500 Hz, and with a dedicated compressor, it is possible to observe a fundamental deviation from the course representing the universal compressor.

The variety of harmonic distortion and saturation affects the frequency characteristics of the sound being edited, and this distortion may behave differently at different settings. Therefore, for a sound engineer who wants to use these types of compressors in practice, ear training with their use is the most effective way to gain aural experience and sonic knowledge. As with equalizers, such training makes it easier to choose the right type of compressor.

Comparisons of sound characteristics can also be found in other types of dedicated audio plug-ins, such as reverb time devices (reverb, delay, chorus, etc.) or instruments designed specifically for saturation (for example, emulation of analog consoles). In general, it can be stated that thanks to ear training through dedicated audio plug-ins, the sound engineer builds skills to work with the user environment of the instrument and learns to use its parameters more quickly in conjunction with the sound concept.

2 PERSPECTIVE AND METHODOLOGY OF IMPLEMENTATION OF EAR TRAINING THROUGH AUDIO PLUG-INS

Despite the differences between the universal and dedicated audio plug-ins described above, it is a good idea not to limit ear training to the use of dedicated plug-ins. Sound engineers often have sound-neutral and versatile instruments in the portfolio of instruments they use. Getting to know these tools through ear training also has advantages, especially in the effectiveness of learning how to use the tool and its settings in practice.

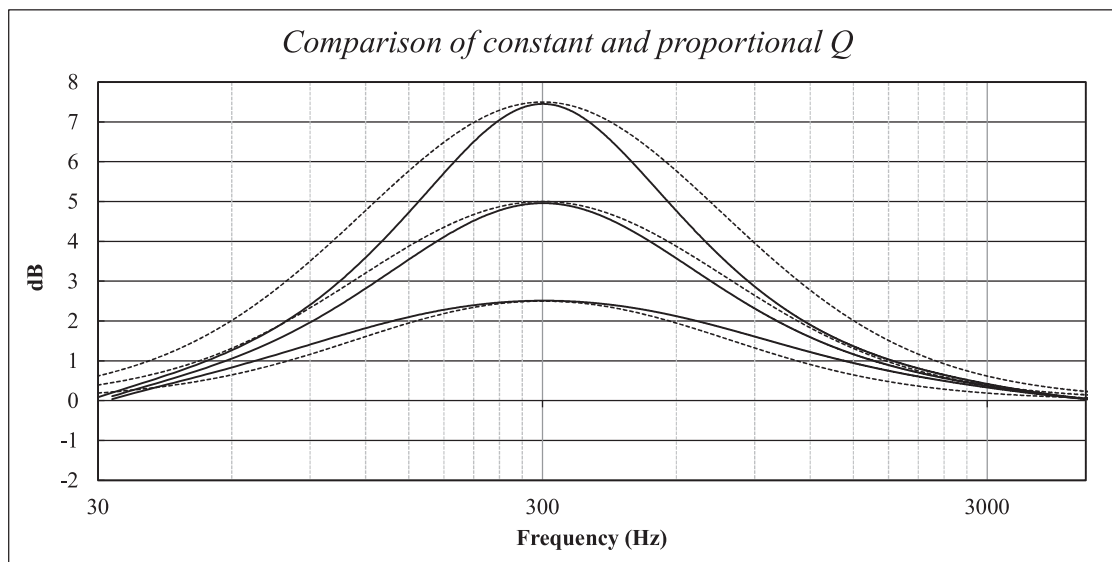


Fig. 2. Comparison of frequency curves of universal (dotted line) and dedicated (solid line) audio plug-in with identical settings—peak at 300 Hz; gain values are 2.5 dB, 5 dB, and 7.5 dB. The universal equalizer filter has the constant- $Q = 0.5$ and the dedicated equalizer has the proportional- Q without the option to select a value.

The vision of training through audio plug-ins requires the creation and programming of a standalone host manager module, through which it is possible to practice any audio plug-in within a single user environment. From a technical point of view, due to the diversity of audio plug-ins, this is an extremely demanding programming process, and it should be noted that so far, this type of software is not available to the general public even as commercial software or freeware.

The philosophy of how ear training works using audio plug-ins is easiest to demonstrate on a prototype that was created for this purpose. The first standalone host manager module, called Albrechtic Ear Training, was created

within the project Advanced application of information and communications technology in music education, solved at the Department of Music of Constantine the Philosopher University in Nitra, Slovakia (the author of this article is also the researcher responsible for the project). The same research team also created specific software for teaching music theory [22]. The architecture of the standalone host manager (Fig. 4) is based on two basic training methods defined by Corey [23] and which are also the most commonly used in practice:

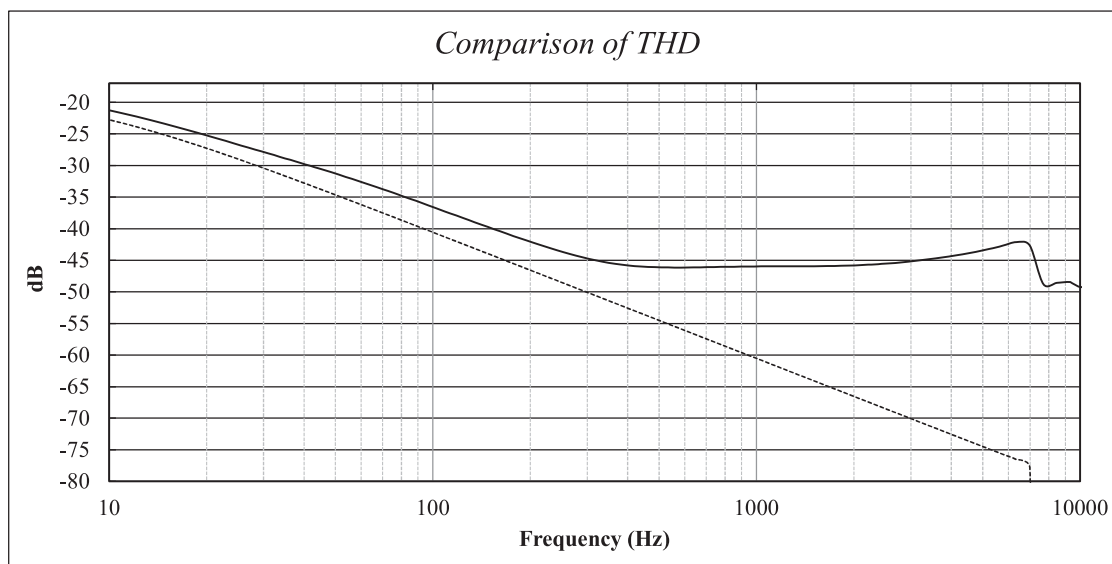


Fig. 3. Comparison of curves of total harmonic distortion (THD) universal (dotted line) and dedicated (solid line) audio plug-in with identical settings—ratio at 4:1; attack time = 10 ms; release time = 60 ms; gain reduction = -10 dB.

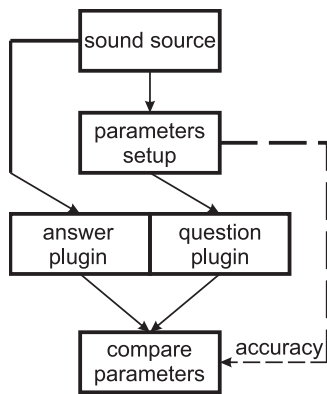


Fig. 4. Architecture of the question-and-answer training method with an implemented level of deviation in the answer (accuracy).

- Matching (sound processor settings by comparing a question and an answer)
- Matching Memory (sound processor settings without the possibility of comparing an answer with a question)

The question-and-answer training system requires the implementation of the simultaneous launch of two instances of the audio plug-in—through one, the question is played, and through the other, the answer is set. The adjustment of the difficulty of the exercise is set by the *Accuracy* parameter, which in percentage expression determines the width of the correct answer (scale of accuracy of the central frequency, scale of the reverberation value, scale of the compression ratio, etc.). In the Albrechtic Ear Training software, the *Accuracy* parameter is adjustable by percentage. The lower the percentage, the more accurately the value of the parameter being practiced must be entered in the response. This method of setting the difficulty is the most effective from the point of view of versatility, and it is applicable to all types of ear training.

The user environment of the Albrechtic Ear Training software is very simple, as the graphical interface of the given audio plug-in is dominant. For the versatility of the training, the program implements a choice between pink noise and a sound file (Fig. 5). In the case of training with audio plug-ins, pink noise can only be used for frequency spectrum exercises, and all other types of training (dynamic and time spectrum) must use music sound recording.

The training settings themselves are implemented universally and are based on the essence of the audio plug-in architecture. They are based on the principle of the possibility of automation of all parameters. Fig. 6 demonstrates a dialog box, in which the behavior of the required parameters of the selected audio plug-in is set. All available parameters of loaded plug-ins are always displayed in the parameter list. Subsequently, a parameter or a set of several parameters is selected to be part of the ear training. Due to the fact that absolute values in the range 0 to 1 are displayed when entering the values of the plug-in parameters, the numerical method of entering the values would be extremely confusing and demanding. The val-

ues are therefore entered directly by setting them in the user interface of the audio plug-in.

During practice, the values are selected at random. However, absolute random selection may not be effective for all types of audio plug-ins—these are mainly those that do not have fixed but continuous values in the parameters (frequency value, ratio value, reverberation length, etc.). It was therefore necessary to implement the category *Use only selected values*, in which it is possible to enter exact values, and only those will be randomly selected during training. If one of the parameters must be set to a certain value even while entering the answer (for example, a fixed level of the gain or threshold parameter), the *Use SW (software) for answer* check box is selected. The stated principle of operation and setting of exercises is universal and effective; in addition, the settings are easy to save and reopen. For each audio plug-in, it is possible to create a methodical set of exercises of varying difficulty exactly meeting the requirements of the tutor.

In addition to the already-mentioned advantages of using a special standalone host manager module, there are also several limitations that need to be taken into account and addressed in the further development of similar applications. An example is the use of audio plug-ins to adjust dynamics (mainly compressors), which do not have an autogain function. These types of plug-ins cannot be used without implementing a separate module to compensate for the volume level. Differences in the question-and-answer volume make it impossible to carry out full-fledged ear training. Another problem is the issue of audio plug-in formats that a standalone host manager will support. The most widespread format today is VST 2.0 (Virtual Studio Technology audio plugin interface by Steinberg), but Steinberg ended its support in 2018 [24]. In the future, only the VST 3.0 format should be preserved. If the host manager should also be used for the Macintosh by Apple Inc. (MAC) platform, it is necessary to consider the implementation of the Audio Unit (AU) format. The Albrechtic Ear Training software prototype currently only works on the Windows platform and only supports the VST 3.0 format for the above reasons. The freely available JUCE programming language (<https://juce.com/>), which has multiplatform support, was used to program the host manager. It is therefore not necessary to program separate versions for iOS and Windows. In the future, the research team will work with both of these platforms. Video demonstration of the program's functionality is available at <https://youtu.be/eO0s4Mc2sqE>. If you are interested in trying out the program, please contact the author of the article.

3 CONCLUSION

The variety of digital audio editing tools in the format of audio plug-ins currently places increased demands on the aural capabilities of sound engineers. In recent years, development companies have increasingly focused on the creation of dedicated audio plug-ins, especially emulations of real existing hardware devices. For less-experienced sound engineers and those who have never worked with hardware,

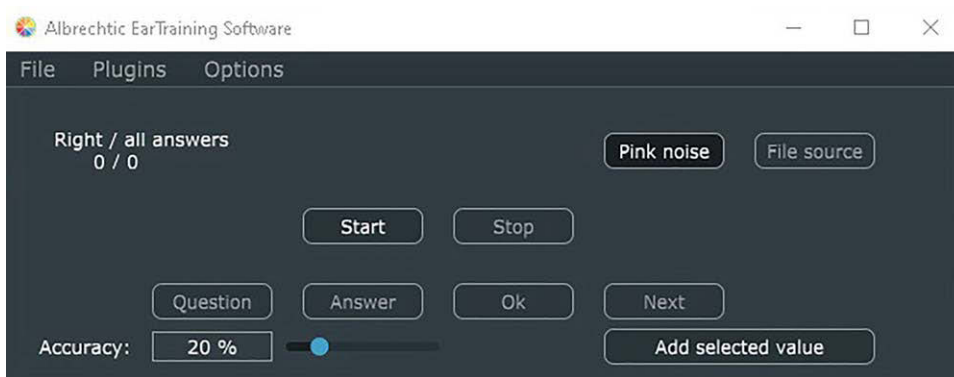


Fig. 5. Simple user interface of Albrechtic Ear Training with sound source selection and training interface.

using emulated audio plug-ins is often challenging. Some of the daily challenges are nontraditional shapes of frequency curves, specific behavior of dedicated tools for adjusting dynamics, and different parameters and properties of tools for adjusting the spatiotemporal characteristics of sound recording.

The only effective way to gain auditory experience and the skills of working with dedicated instruments is to include them in ear training. A better understanding of the

functionality of these sound editing tools helps to make faster and better decisions when using them, and it has an overall impact on the resulting quality in practice. Thanks to the open architecture of hosting managers, any existing listening training can be implemented for the sake of loading any audio plug-ins and creating listening training options. The methodology and creation of listening training courses will be the subject of further research by the research team



Fig. 6. Dialog window of the Albrechtic Ear Training program, in which the parameters of ear training are set. The lower part shows the graphical environment of the audio plug-in, which is used to enter parameter values (the equalizer plug-in was cropped intentionally to not show brand names).

on the basis of the responses and discussions of experts and teachers.

It is perhaps surprising that there is still no software solution in the commercial sphere for practicing listening skills through audio plug-ins. Hence, the aim of this article is to point out the possibilities of this type of training. For this purpose, a prototype of the standalone host manager Albrechtic Ear Training Software was also created, which should serve as a research tool and at the same time be an inspiration for potential creators of similar applications. In the next step, the software will be used for gaining training data. Results will be used for preparing case study about the effectivity of audio plug-ins-based ear training.

4 REFERENCES

- [1] E. B. Brixen, "Spectral Ear Training," presented at the 94th *Convention of the Audio Engineering Society* (1993 Mar.), paper 3474.
- [2] T. Neher and T. Brookes, "Training of Listeners for the Evaluation of Spatial Sound Reproduction," presented at the 112th *Convention of the Audio Engineering Society* (2002 May), paper 5584.
- [3] T. Brookes, R. Kassier, and F. Rumsey, "Training Versus Practice in Spatial Audio Attribute Evaluation Tasks," presented at the 122nd *Convention of the Audio Engineering Society* (2007 May), paper 7117.
- [4] M. Bassett and W. L. Martens, "SAE Parametric Equalizer Training: Development of a Technical Ear Training Program Using Max," presented at the 141st *Convention of the Audio Engineering Society* (2016 Sep.), eBrief 298.
- [5] T. R. Letowski, "Development of Technical Listening Skills: Timbre Solfeggio," *J. Audio Eng. Soc.*, vol. 33, no. 4, pp. 240–244 (1985 Apr.).
- [6] T. R. Letowski, "Anchor Effect in an Optimum Timbre Adjustment," *J. Audio Eng. Soc.*, vol. 40, no. 9, pp. 706–710 (1992 Sep.).
- [7] T. Schaller and I. G. Burleigh, "Using Web Audio API in Web-Based Tools for Ear Training of Sound Engineers," in *Proceedings of the AES 26th UK Conference: Audio Education*, pp. 105–110 (Glasgow, UK) (2015 Aug.).
- [8] R. Quesnel and W. R. Woszczyk, "A Computer-Aided System for Timbral Ear Training," presented at the 96th *Convention of the Audio Engineering Society* (1994 Feb.), paper 3856.
- [9] D. Moulton, *The Golden Ears* (KIQ Productions, California, 1995).
- [10] A. Nishimura, "Ear Training for Non-Technical Students Using a Web-Based Training System," in *Proceedings of the AES 50th International Conference: Audio Education*, pp. 1–7 (Murfreesboro, TN) (2013 Jul.).
- [11] D. Martin, G. Massenburg, and R. King, "Dynamic Range Controller Ear Training: Description of a Methodology, Software Application, and Required Stimuli," presented at the 143rd *Convention of the Audio Engineering Society* (2017 Oct.), paper 9888.
- [12] KVRAudio, "V-plugs – Mr. Soundman," <https://www.kvraudio.com/product/mr-soundman-by-v-plugs> (accessed February 14, 2020).
- [13] Soundgym, "SoundGym Ear Training," <https://www.soundgym.co> (accessed February 14, 2020).
- [14] Harmann, "Harmann's How to Listen!," <http://harmanhowtolisten.blogspot.com/2011/01/welcome-to-how-to-listen.html> (accessed February 14, 2020).
- [15] TrainYourEars, "TrainYourEars Ear Training," <https://www.trainyourears.com/> (accessed February 14, 2020).
- [16] J. Corey, "Beyond Splicing: Technical Ear Training Methods Derived From Digital Audio Editing Techniques," presented at the 123rd *Convention of the Audio Engineering Society* (2007 Oct.), paper 7212.
- [17] T. Letowski, "Sound Quality Assessment: Concepts and Criteria," presented at the 87th *Convention of the Audio Engineering Society* (1989 Oct.), paper 2825.
- [18] International Organization for Standardization, "ISO 266:1997, Acoustics—Preferred Frequencies" (Geneva, Switzerland, 1997).
- [19] W. Pirkle, *Designing Audio Effect Plug-Ins in C++: With Digital Audio Signal Processing Theory*, 3rd ed. (Routledge, New York, NY, 2019), <https://doi.org/10.4324/9780429490248>.
- [20] R. Miller, "Equalization Methods With True Response Using Discrete Filters," presented at the 116th *Convention of the Audio Engineering Society* (2004 May), paper 6088.
- [21] D. Martin and G. Massenburg, "Advanced Technical Ear Training: Development of an Innovative Set of Exercises for Audio Engineers," presented at the 139th *Convention of the Audio Engineering Society* (2015 Oct.), eBrief 217.
- [22] P. Brezina, A. Čierna, and M. Vozár, "Albrechtic—Software For Teaching Music Theory," in *Proceedings of DIVAI 2020: 13th International Scientific Conference on Distance Learning in Applied Informatics*, pp. 245–250 (Sturovo, Slovakia) (2020 Sept.).
- [23] J. Corey, *Audio Production and Critical Listening: Technical Ear Training*, 2nd ed. (Routledge, New York, NY, 2017). <https://doi.org/10.4324/9781315727813>.
- [24] Steinberg Media Technologies, "News: VST 2 Coming to an End," <https://www.steinberg.net/en/newsandevents/news/newsdetail/article/vst-2-coming-to-an-end-4727.html> (accessed February 14, 2020).

THE AUTHOR



Pavol Brezina

Pavol Brezina is Assistant Professor of Music and Sound Design in the Constantine the Philosopher University in Nitra, Slovakia. He is the program administrator for the undergraduate and graduate program in music and sound design. Pavol is a musician, composer, and sound engineer who specializes in popular, jazz, and classical

music composing and recording. He is also working as a sound engineer in his own recording studio and collaborating on tutorials for Vienna Symphonic Library software developers. His main research fields are music pedagogy and acoustics of historic spaces as a cultural heritage.