Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display, Sydney, Australia, July 6-9, 2004

ANALYTIC LISTENING: A CASE STUDY OF RADIO PRODUCTION

Jim Barbour
Swinburne University of Technology

Lecturer in Radio, Audio and Multimedia
P.O. Box 218, Hawthorn, Melbourne, Victoria, 3122
jbarbour@swin.edu.au

ABSTRACT

Radio is the ultimate auditory display: all information is conveyed aurally to a listener. Sound from a radio invokes visual images in the imagination of a listener, the theatre of the mind. Listening to a radio voice creates images of a person talking from a studio live to their audience, providing information and entertainment. However, much radio is pre-recorded, produced in a radio studio at a different time and replayed on cue, often controlled by sophisticated computer systems. The art of radio production relies on developing excellent listening skills. The craft of radio production now relies extensively on digital audio technology. How has the process of radio production changed with the introduction of new technology, particularly the visualization of audio on a computer screen? Have these changes affected the listening skills of radio producers and has this affected the quality of radio production? This paper explores these issues based on individual interviews with experienced radio producers and the author's experience in training new producers.

1. INTRODUCTION

1.1 Listening

There are two principle modes of listening suggested by Williams [1]: holistic listening, where we take in the totality of the soundscape[1] without consciously listening to any one aural element, and analytic listening, where we focus our attention on one aural element in a complex aural environment. For example, when listening to an orchestral performance, holistic listening would refer to the process of allowing the complex aural production to wash over us as we enjoy the performance of the orchestra as a whole. However, switching our perception to analytic listening would allow us to focus on the sound of one trumpet, for example, and listen to the performance of the player, to analyze their pitch accuracy, their tone quality and subtle variations in their phrasing.

[1] A soundscape is an aural landscape, a three dimensional aural environment.

1.2 Foreground or Background

There are few people for whom sound becomes a passion that would lead them to develop excellent analytic listening skills. Research [2] suggests that almost everyone in Australia listens to the radio for some time every week, but they are listening holistically, with the radio being just another sound source in their environment. Very few radio listeners would be listening intently in a quiet room, sitting equi-distant from each of the stereo loudspeakers. Instead, radio is a part of their busy lives and their radio listening would vary between a foreground and a background sound. When an interesting song is playing or a voice attracts their attention, the radio would become a foreground sound, separate from other environmental sounds. When the song or the voice is no longer interesting, the radio becomes a background sound, blending back into the soundscape. However, while their listening focus is changing between foreground and background, it is principally holistic listening, because they are not analyzing the tonal quality of the voice or the balance of the voice compared to the music or sound effects that are part of the radio production. Instead they are listening to the message in the voice or the song.

1.3 Ear Training

To develop analytic listening skills, it is necessary to learn to separate the individual aural components within a complex soundscape. Ear training for radio producers is a process of learning to switch at will between holistic and analytic listening, to be able to identify and analyze individual aural parameters, but then to fit that sound back into the whole aural production and listen holistically as would a normal radio listener. If there was one primary skill great radio producers possess which sets them apart from their peers, it is their ability to switch effortlessly between holistic and analytic listening and to integrate the two modes of listening into their productions. For example, they are able to listen analytically to identify the timbres and textures of a voice and to articulate their requirements to the voice talent, to elicit the best vocal performance. The producer is also able to listen analytically and adjust the parameters of the audio technology to achieve the best

recorded sound. They are able to combine the voice with backing music and sound effects to enhance the message of the production without losing the clarity of the message or losing the impact of the music and effects.

1.4 Production Skills

Producing radio requires the development of creative and technical skills, combining the use of complex audio technology with artistic taste and judgement. It is necessary to practice excellent interpersonal skills when working with voice talent, creative writers, program directors and technical support staff. It is essential to maintain effective time management and attention to detail. It is also critical to learn the artistic and technical differences between good and bad sound and to understand how to operate the audio technology to obtain the highest professional standards of audio production. With the development of very fast processors and high quality audio interfaces, almost all audio is now produced with computers, including virtually all radio production. This has revolutionised the process and the pace of production, with many significant benefits for radio producers. However, there are several aspects of this computerisation that may lead to poorer quality productions and a deskilling of producers: the visualization of audio on a computer screen and the reduction of the human interface with the audio technology to a mouse.

2. THE DANCE OF RADIO PRODUCTION

2.1 Manual Interactivity

Audio technology has changed in ways that have completely altered the techniques for all audio production. For more than fifty years before the computerisation of audio at the end of the twentieth century, audio was recorded on magnetic tape, a process that involved complex electronic and magnetic technology and human interaction. Reels of tape were carried by hand to an expensive electro-mechanical machine, where a thin ribbon of plastic tape was laced around rotating bearings and operated by pushing buttons and flicking switches, accompanied by flashing lights and the smell of cleaning fluids. Recorded sound was edited by cutting the tape with a razor blade and sticking it back together. It was a tactile and mobile process, involving walking around, operating with both hands and lots of finger movements. While tape recorders are still used by many music producers, primarily for the aural quality of the sound reproduced from magnetic tape, very few radio producers now use tape recorders. It is certainly the case that most radio producers who have experience in both analogue tape production and digital computer production rarely miss tape recorders!

2.2 Movement and Memory

The sound sources in a radio studio were auditioned, altered and combined through a large mixing console, where each sound was available on a separate channel with many knobs, buttons and faders to change the parameters of the sound, turning it on or off, making it louder or quieter, or changing its timbre with more treble or less bass. Combining many sounds together, for example, two different voices, two or three different pieces of music and several different sound effects, involved moving different controls for each sound at different times, in an intricate dance of fingers and arms. The final mix of all the elements might take some time to learn the dynamic and timbral changes necessary for the best result, often repeating the movements again and again until muscle memory and subconscious interaction delivered the exact pattern. This mixing was often a long and complex process, but for many radio producers it was one of the more enjoyable parts of their working life. How have techniques changed and what are the consequences?

3. VISUALIZING AUDIO

3.1 Analyzing Audio

Perhaps the most significant change has been the visualization of audio on a computer screen. Prior to computers, sound was not a visual thing! To develop analytic listening skills involved focusing your attention on the aural environment, identifying a particular sound and learning to identify its component parts. How did the sound change over time? Did it get louder then softer? Did the timbre change from bright to dull? Did the pitch start low, rise to a high point then move lower again? One of the best ways to develop this skill was to shut your eyes and concentrate your attention, so that sound was the dominant sense. Now we can see sound on a computer screen, in lots of different colours, shapes and sizes and we can manipulate its aural characteristics based on what we see on the screen, rather than what we hear.

3.2 Audio Software

The software used in a radio production computer provides each individual sound with its own visual representation, which shows the changing loudness patterns over the duration of the production, see Figure 1. Each sound can be manipulated by the producer to change its loudness or tone at any time during the production, and these changes can be remembered by the software and repeated every time the production is played, a process called automation. A complex production is created adding one sound at a time, adjusting the starting and finishing time of each part and its loudness and tone until all the sounds are assembled.

The final mixing involved making subtle adjustments in loudness and tone to each sound until the correct balance between the elements is achieved and the final result recorded and stored ready for replay on-air. How do we judge the quality of a radio production and how is that quality affected by the visualization of audio on a computer screen?

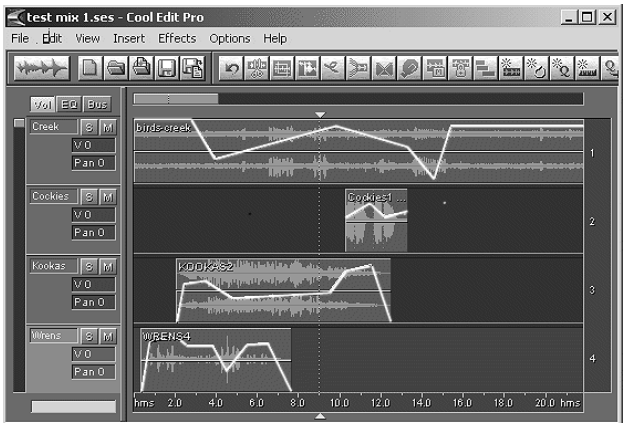

Figure 1: Four individual sounds, starting at different times, with the loudness changes shown drawn over the waveform of each sound.

4. ENHANCED PERCEPTION

4.1 Sensing Modes

The psychological theory of Gestalt [3] suggests that when we receive information from more than one sensing mode, overall perception is enhanced, that the whole will be greater that the sum of the individual parts. This would imply that we gain greater information about a source when we have visual and aural information simultaneously. This may be true for a musical performance where we can see as well as hear a person playing an instrument, where we may gain greater enjoyment from the performance because we can see the skill and effort of the performer. But it is not the case for radio production where the end result for a normal listener is an aural only sensing. Just as very few people can effectively lip read, there needs to be no visual sensing for us to understand spoken words.

4.2 Dominant Visual Focus

In interviews with radio producers, and from observations of trainee producers, the visualization of audio on a screen has become the dominant focus in a radio production studio. All the aural elements of a production are assembled ‘on the screen’, changes are made to elements based on the visual representations and the final balance is attempted while focused on the screen. The interface with the computer is primarily the mouse, which requires detailed visual attention and fine motor skills to manipulate the menus and settings which change audio parameters. It is difficult to use a mouse with eyes shut, but that would be exactly what is required to engage analytic listening. So what are the techniques experienced producers use to ensure high quality results?

5. HARNESSING ANALYTIC LISTENING

There are three principle areas of concern: timing, loudness and timbre. Each of these requires physical interaction with the software and analytic listening to achieve the correct aural characteristics in the final production. There are different ways that experienced producers approach these issues.

5.1 Timing

When will each sound element start in the production? Often the timing for the start of each sound will be dependent on other sounds occurring at the same time. For example, a music track may begin before the voice, with the voice timed to start on a particular beat in the music. When radio was produced on tape recorders, individual parts were transferred from CD to multi-track tape one at a time, with the second sound manually started while listening to the replay of the first sound. The process often needed to be repeated several times until the second sound did actually start at the correct moment. On a computer screen, the producer can zoom in visually to identify the beat where the sound should start and the second sound can be slipped left or right on the screen, representing earlier or later in time, until it matches the visual point. However, it is only by listening analytically to the replay can the producer be confident the timing is precisely correct. If the producer relies only on the visual alignment, it may be aurally incorrect. By looking at the screen while listening to the replay, it is difficult to separate the visual sense that it ‘looks correct’ from the aural sense that it ‘sounds correct’. The difference may be subtle, but is part of the magic in a great production.

5.2 Loudness or Dynamics

5.2.1 Masking

A more significant problem with computer visualization affects the loudness balance between the sounds. When two or more sounds occur simultaneously, one sound may overpower the other, known as masking. For example, a music track with bright instruments like saxophones or guitars will mask a voice if the instrument is playing at the same moment as the voice, with similar tonal characteristics and pitch ranges contributing to this masking. The voice will only be clear if the instrument is lower in volume or has its tone changed to reduce the clash with the voice. If the instrument in the track cannot individually be turned down in volume, then the whole

music track must be turned down to reduce masking. When mixing manually, the volume controls for the music track will be physically moved down to make the music quieter for the duration of the voice so that it is clear. The judgement of this volume change will be made by analytic listening, focusing on voice clarity.

5.2.2 Automation

When mixing with a computer, the changes in volume can be made using a mouse to mark points on the screen where the changes are necessary, and drawing a line representing volume between these points, see Figure 1. The volume for any sound can be made to increase or decrease and the computer will remember the changes and repeat them exactly every time the production is played. However, for many radio producers, the process of determining the amount of change will be estimated visually and checked while focussed intently on the visual representation, and it is very difficult to listen analytically while focussed visually.

5.2.3 Shut Your Eyes!

To engage analytic listening, many radio producers focus visually on nothing, like a spot on the wall midway between their loudspeakers, or better still, shut their eyes. The true balance between the voice and the music becomes much clearer, allowing greater accuracy in setting the volume changes. With the best audio software, it is possible to record the movements of a mouse or the movements of a manual volume control fader connected to the software. While the production is playing, the producer can close their eyes and focus their listening on the balance between the voice and the music, change the volume of the music as necessary and have the computer record those changes. When the best balance is achieved, it will undoubtedly be better than any balance drawn on the screen.

5.3 Equalization

5.3.1 Reproducing the Real

When a voice or any other instrument is recorded, the tone of the voice is changed by technology: the microphone, the recording medium, any electronic processing used and by the loudspeaker used for reproduction. Consequently, most listeners will readily differentiate between a recorded voice and a live voice. The art of recording seeks to use the best quality audio technology and the best recording techniques to attempt to make a recorded voice sound as natural or ‘real’ as possible. It is only by applying analytic listening and carefully learning the craft of recording that radio producers will achieve the best recorded results. The use of tonal change or equalization is an important part of this craft and is often achieved manually using tone controls on a physical mixing console. Particular bands in the frequency spectrum of the voice or instrument are selected and the volume in that band only can be increased or decreased to change the overall tone. By using analytic listening, the different bands in the frequency spectrum can be changed until the final result is the best possible for the radio production.

5.3.2 Visualizing Timbre

Equalization can also be applied in software. Using a mouse and watching on the computer screen, different frequency bands can be highlighted and changed as the producer hears the results. Once the desired tonal change is completed, it can be stored as a ‘preset’ in the software memory. The visual representation of a tone change is shown in figure 2. Changes in tone over time can also be automated, making the use of audio software a very powerful tool. However, there are two significant problems with this approach to equalization. Firstly, it is very tempting to set the tone change visually, drawing a ‘nice’ pattern on the screen, without absolute regard to how it affects the sound. The only way to achieve the highest quality in tonal change is to shut the eyes and employ analytic listening while making the variations to equalization. It is difficult to do this with a mouse as there are usually two or three parameters for each frequency band to be adjusted interactively, and each must be selected by the mouse to make a variation. An external manual controller linked to the software, with a separate physical control for each parameter that can be changed with eyes shut, will allow the greatest focus on listening.

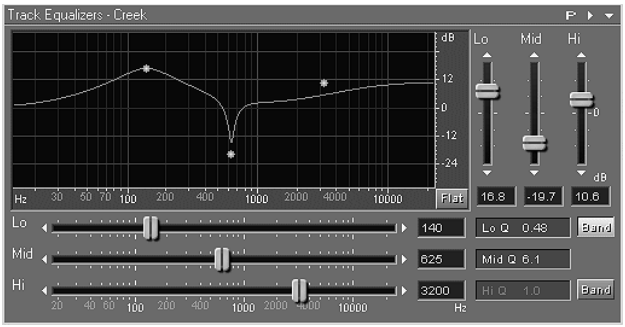

Figure 2: Visual representation of equalization, from low frequencies or bass sounds on the left to high frequencies or treble on the right

5.3.3 Preset Problems

The second potential problem is the use of memorized preset equalization patterns, where one pattern is used for a different recording. A vocal performer will never produce exactly the same tonal sound on a different day, or even on the same day at a different time, as their voice will change with tiredness and mood, and the audio technology will change with temperature and humidity. Every voice or instrument must be equalized using analytic listening for every separate recording to

achieve the best production. A preset suitable for one voice will never be the best for a different voice, and even if it is used as a starting point for fine-tuning the tonal characteristics, it will most likely influence the result in a potentially destructive way. Every recording must be approached from a neutral unaltered perspective and only the changes judged necessary by analytic listening should be applied.

6. MANUAL CONTROL INTERFACES

There are many companies now producing physical control devices that connect with the audio production software to allow manual interaction and control over audio parameters. [4] These manual control interfaces typically provide volume faders, on/off buttons, knobs for control of equalization, play/stop/record buttons and a rotating data wheel for slipping tracks in time or adjusting a parameter of the sound. Once the layout and operation of the manual control interface is learned, it is very easy to use the controller to perform almost all the functions necessary to assemble and mix a complex radio production without touching the mouse. Further, it is much easier to listen analytically when using the controller because we can use many fingers to change parameters with our eyes shut. While the interface is often expensive and only connects to expensive software, most successful radio producers have access to a professional control interface. It was suggested by several interviewees that a good interface provided a strong link between the eye, ear and touch, with the tactile mode being a link between visual perception and analytic listening.

7. PROFESSIONAL PRACTICE

7.1 Experience

However, despite access to a manual control interface, many interviewed producers use the mouse almost exclusively, leaving the controller virtually untouched! They believe that the pace of production and the quality of the parameter automation within the software allows them sufficient interaction to be able to engage analytic listening while still altering a parameter with a mouse. They also suggest that the pressure to produce more material leads to a greater reliance on the visual, and they considered the manual interface to be slower. The shift in listening between analytic and holistic, and the shift in focus between listening and looking are very rapid and based on many years experience. There is very detailed knowledge of the software and its operation and years of training in the use of audio technology. More important perhaps is the detailed ear training gained from analyzing the quality of individual sounds, the effect that changing one parameter will have on a sound and the way individual sounds will combine in a complex production.

7.2 Studio Layout

To achieve analytic listening, the techniques employed varied between producers, and the physical layout of their studio influenced this process. Virtually all radio production is currently produced in stereo, with the ideal listening position equi-distant between two loudspeakers. The studio is acoustically treated to provide an environment similar to an average home listening room, though very much quieter. For most producers, the screen for their computer is placed midway between the loudspeakers, so that when looking at the screen, the sound is balanced on either side. While this is an apparently logical placement, it might be considered exactly the wrong place when analytic listening is critical!

It is very difficult to focus on the sound coming from the two loudspeakers when the computer screen fills the visual space. In studios where this is the placement, producers adopt several different techniques for analytic listening: some close their eyes, some ‘glaze over’ or de-focus their vision while others look over the screen to a spot on the wall beyond. Several radio producers interviewed have chosen to place the computer screen to one side, away from the stereo listening position. In this position, they would typically look at the screen to select the parameter to adjust for an individual sound, and then they would turn away from the screen and engage analytic listening while performing the change, with the automation remembering any movements.

7.3 Loudspeaker References

It is common for radio producers to have more than one pair of loudspeakers to listen to, since the differences in radio listening environments are very great, for example, in the bathroom, in the car, at work, etc. Usually producers will have one pair of very high quality loudspeakers for normal listening and one pair of smaller, poorer quality loudspeakers for checking under more adverse conditions. Several producers have set up one pair of loudspeakers either side of the computer screen and a second pair where they turn away from the screen. They would replay a final mix on the second pair of loudspeakers and engage analytic listening to check its quality.

8. CONCLUSIONS

Developing strong analytic listening skills is a critical part of the training necessary to be a successful radio producer. Humans have a very strong preference for visual sensing over all other modes [5]. To learn analytic listening, producers must disengage visual sensing.

From the interviews with successful radio producers, they all consciously engage analytic listening at different times during production, and there were a variety of techniques to disengage visually from their productions, including shutting their eyes, looking away and de-focusing. All may be effective depending on personal preferences.

The interface between the producer and the software used for radio production is important, and most interviewed producers favour using the mouse almost exclusively, despite often having an expensive manual control interface available to them. While using the mouse requires focused viewing to select and manipulate software parameters, these producers have a well-developed ability to engage analytic listening while performing their mouse actions.

Experienced producers have spent many hours experimenting in the studio, manually changing controls and carefully listening to the sound and how it changes, playing around with the technology and learning what will work and what does not. Less experienced radio producers who rely on visual perception, use only a mouse and work quickly will struggle to produce good quality productions. New producers can only learn the interaction between audio technology and sound by experimenting while practicing analytic listening.

9. AKNOWLEDGEMENTS

The author would like to thank radio producers at the following radio stations and radio production studios in Melbourne for the time and support they have given to this research project: Nova100, GoldFM, ABC Triple J, 3AW, mcm entertainment, Rowland Productions and several independent producers.

10. REFERENCES

[1] Shiela M. Williams, 'Perceptual Principles in Sound Grouping', in *Auditory display : sonification, audification, and auditory interfaces*, editor: Gregory Kramer, Addison-Wesley, 1994, ISBN: 0201626039

[2] Radio research information available at Nielsen Media Research: www.nielsenmedia.com.au

[3] W. Viney, *A History of Psychology*. Allyn and Bacon, Massachusetts, 1993, ISBN 0205141749

[4] Manual interfaces are produced by many companies, including Digidesign, Steinberg, Mackie and JLCooper.

[5] S. Goren, L. Ward, J. Enns, *Sensation and Perception*. Harcourt Brace College Publishers, 1993. ISBN0155001035

About the Author:

Jim Barbour:
Jim has extensive experience working in the audio industry in Australia and overseas, with more than twenty years as a music producer and recording engineer and many years as a production manager in commercial radio. After Electrical Engineering studies at Adelaide University in South Australia, Jim worked overseas and in Australia in recording studios and conservatories of music, producing audio for CD, vinyl, radio, film and television releases and is still active in music performances and recordings. Credits include recording engineer for Men at Work's 'Business As Usual', Sound Supervisor for the Australian Opera, film music awards and advertising awards for radio production. Over the last decade, Jim has designed and delivered training programs at Swinburne University of Technology in Radio and Multimedia. Jim is a member of the Audio Engineering Society, Australian Forum for Acoustic Ecology, and the Australasian Computer Music Association. Jim holds a Master of Communications with a thesis titled *Digitizing Acoustic Space: the Impact of new Digital Audio Technology on Aural Perception*, and has a Graduate Diploma of Education in Information and Communications Technology from Melbourne University.