Auditory Performance of Individuals with Reduced Hearing Capability in Virtual Reality Environment

Hung Tran
CAE USA Inc.
Tampa, FL
hung.tran@caemilusa.com

The effort to apply Virtual Reality (VR) technology to advance the fields of medicine, education, engineering, and entertainment is currently underway. The military uses VR to carry out training such as aircraft maintenance or virtual war scenarios. Despite important progress made in display technologies, a complete immersion in VR space is not possible without an auditory representation of the simulated environment. Sound is important in an immersive virtual environment because it enhances the sense of presence (Freeman, Lessiter, 2001) and improves situational awareness by providing feedback for situations that are not in the listener's field of view (Kukka et al., 2016). Additionally, the degree of presence experienced by an individual affects the performance of the training tasks (Stevens, Kincaid, 2015).

In order to synthesize a functional auditory environment, it is important to obtain a better understanding of how ears receive and process sound. To assess human training performance, it is essential to understand how the perception of the simulated sound environment is impaired for individuals with reduced hearing capability, which can be caused either by natural factors (e.g., age-related hearing loss [presbycusis]) or by external agents (e.g., noise-induced hearing loss). This paper provides a literature review for identifying limitations in auditory perception for individuals with sensorineural hearing loss (SHL). We reviewed hearing impairments due to SHL, which will likely affect auditory performance in a virtual environment. Using an ecological approach to explain the relationship between hearing impairments and auditory demands, we analyzed how this relationship affects human training performance. Finally, we provide guidelines to effectively design and implement VR environments, while taking into account human auditory performance, including individuals with reduced hearing capability.

ABOUT THE AUTHOR

Hung Tran is a Tactical Systems, Aircraft Sensors and Interoperability Group Leader at CAE USA. He joined CAE USA more than 20 years ago and has worked on the modeling and simulation of several Electronic Warfare (EW) systems. Hung is the lead designer of the Computer-Generated Forces (CGF) currently used by all United States Air Force C-130J Weapons System Trainers. Hung holds a Bachelor of Science in Electrical Engineering from Polytechnique Montréal at University de Montréal in Quebec, Canada.

Auditory Performance of Individuals with Reduced Hearing Capability in Virtual Reality Environment

Hung Tran
CAE USA Inc.
Tampa, FL
hung.tran@caemilusa.com

INTRODUCTION

The term "virtual reality" leads us to immediately think of the *head-mounted display* device. However, a broader meaning of this term refers to all technologies used to create a virtual environment presented to our senses in a coherent manner. To be coherent, the virtual environment must provide sensation or illusion of sensation that is similar to those we experience in the real world. The use of virtual reality (VR) for training and learning is an extension of flight simulators originally developed for pilot training many years ago. Today, VR technology is applied in the fields of entertainment, education, medicine, engineering, and much more—ranging from complex surgical procedures to military mission-rehearsal operations. One of the first areas for practical application of VR was in military training. Military VR training applications include flight simulators, battlefield and war scenario simulation, medic training, and training exercises. VR quickly became the preferred approach for training because it offers many important advantages over legacy training approaches. Besides being cost effective, VR presents a safe virtual environment in which to train complicated tasks or dangerous procedures, allowing students to learn from mistakes that—in real life—could present a major risk of injury or death. Thus, when compared to traditional training techniques, VR training represents a far safer environment for learning and training.

A functional VR environment is not necessarily an exact copy of the real world but a virtual space of activity that provides an operator the necessary cues and response options to perform specified tasks the same way as in the real world. *Immersion and presence* represent the primary and interconnected characteristics of a VR environment. The term "immersion" refers to the technology employed to create the virtual environment. The higher fidelity a system can be in relation to the real-world equivalent sensory modalities, the more immersive it is. Interactive devices, such as data gloves and head-mounted displays, are used to interact with the environment, facilitating tactile, visual, and auditory sensing. The fidelity of application interfaces directly affects the immersion of the virtual environment (Fuchs, Moreau, & Guitton, 2006). In contrast to immersion, the concept of "presence" in VR has been defined as the degree to which the users feel that they are physically somewhere other than where they currently are. Thus, the notion of "being present" in a VR world is a psychological, perceptual, and cognitive *consequence of immersion*. One of the principal external determinants of presence is the extent of sensory information presented to the user within the simulated environment; that is, the greater the number of sensory inputs provided to different modalities, the greater the sense of presence (Sheridan, 1992). The perception aspects that directly affect the design of the VR include visual perception, auditory perception, haptic perception, and kinesthetic perception. While a large proportion of VR work is focused on visual display, auditory display has received very little attention despite the fact that it has a significant effect on presence and awareness (Freeman & Lessiter, 2001; Kukka et al., 2016). Imagine if the auditory background of our every day is removed—we will feel less "connected" to the world (i.e., less presence). Where our eyes are completely blind to the "rear" half of the world, our ears do not present such limitations. When we hear something behind us, not only are we *aware* that the object exists in the environment, but often we can also identify *what it is*. Therefore, in a VR environment, it is possible to control the audible properties of objects in the manner that our hearing mechanism will allow us to perceive the existence of these objects, as well as what they look like. Furthermore, research in VR demonstrates that the addition of audio display enhances visual perception and human performance (David et al., 1999).

Does presence have an effect on human performance while performing a training task in a VR environment? The objective of using a VR environment for training is to optimize the transfer of knowledge and skills to the trainee. Transfer is positive when the trainee performs the real-world task better after training. Some studies have shown a positive correlation between presence and transfer: higher presence is associated with higher transfer (Winn, Windschitl, Fruland, & Lee, 2002; Mikropoulos, 2006). A study of the relationship between presence and performance in VR training found evidence that higher presence in virtual simulation training results in higher human performance

(Stevens & Kincaid, 2015). The result of this research is important because VR designers could potentially maximize the trainee's sense of presence to increase training performance. The extent of sensory information presented to the trainee within the training environment has been identified as the determinant of the sense presence (Sheridan, 1992). While a large proportion of presence studies have focused on the visual modality, some studies have demonstrated that addition of audio display in the VR environment significantly increases the sense of presence (Hendrix & Barfield, 1996; David et al., 1999). Furthermore, in order to maximize human training performance in VR, designers must take into account the design constraints imposed by human sensory and motor physiology. Without a foundational knowledge in these areas, the VR system used for training will likely be incompatible with the users. Consequently, for VR training environments where the production of a *virtual auditory environment* is required, an understanding of the constraints and limitation of the human auditory system is essential.

In this paper, first, we will briefly review how ears receive and process the sound. Second, we will explain how sensorineural hearing loss can impair this process in terms of functional limitations in a VR environment. Third, we will present an ecological framework that explains how hearing capabilities affect human performance in a VR environment. Finally, we will provide a guideline to effectively design and implement an auditory environment, while taking into account human auditory performance, including individuals with reduced hearing capability.

HEARING IMPAIRMENT ASSOCIATED WITH SENSORINEURAL HEARING LOSS

Human ears convert sound pressure into neural signals, which in turn lead to a perceptual experience. The ear is divided into 3 main parts: the outer, the middle, and the inner part. The *pinna*, part of the outer ear, collects sound energy and guides it into the *ear canal*. After traveling down the ear canal, the sound waves cause the eardrum to vibrate. The *eardrum* is a cone-shaped membrane and is part of the middle ear. The function of the middle ear is to convert the vibrating air into vibrating liquid of the inner ear. The inner ear contains the *vestibular organs* and the *cochlea*, which are the sense organs for hearing. Sensorineural hearing loss (SHL) occurs when there is damage to the inner ear or from the nerve pathway from the inner ear to the brain. There are several causes of sensorineural hearing loss, but the most common are due to either a prolonged exposure to intense sound or the natural aging process (presbycusis). Therefore, we will all acquire some degree of hearing loss at some stage in our life. Sensorineural hearing loss affects our auditory perception. A considerable body of knowledge is available in the literature for describing the hearing impairment associated with SHL. The deterioration of the hearing perception due to a sensorineural hearing loss is manifested in various ways. It is beyond the scope of this paper to describe in detail the effect of hearing impairment on the auditory capability. The purpose of this paper is to summarize the aspects of hearing impairments that will likely affect the VR experience, considering that these aspects cause significant changes in the processing of auditory signals by the inner ear (Hétu & Tran, 1996).

Loss of Absolute Sensitivity

The human hearing system is sensitive to sounds with frequency ranging from approximately 20 Hz to 20 kHz. Absolute sensitivity, typically expressed in decibel (dB) sound pressure level (SPL), represents the signal detection threshold of the hearing system, which is the lowest level of sound pressure required to produce an auditory sensation in a silent environment. This threshold of hearing varies with the frequency of the sound. The ear is more sensitive in the middle of the frequency range than the lower and upper extreme (International Organization for Standardization [ISO] 226:2003). ISO 28961:2012 provides detailed descriptive statistics and percentiles of the hearing threshold distribution for normal persons in the age range of 18 to 25 years old. Figure 1a shows the absolute hearing thresholds for normal hearing listeners.

Absolute sensitivity can be determined by a hearing test. The result of a hearing test is often displayed in the form of an *audiogram*. Based on the hearing thresholds obtained from the audiogram, the hearing healthcare professional can quickly determine whether the listener is suffering a hearing loss and how serious the hearing loss is. The audiogram illustrates the hearing capability by showing the hearing threshold at various frequencies, typically ranging from 250 Hz to 8 kHz. Hearing thresholds, as shown on an audiogram, are thresholds normalized to the values of a normal population (ISO 389-7:2005). Normalized hearing thresholds are measured in dB Hearing Level (HL). The value of a normal hearing population has a hearing threshold value of 0 dB HL; consequently, this value increases as the hearing loss increases. Figure 1b shows typical hearing thresholds of a normal hearing listener and a listener with a moderate noise-induced hearing loss (Tran & Hétu, 1996).

The audible frequency range changes dramatically for SHL listeners. Healthy young children may have a full hearing frequency range up to 20 kHz. From the age of 20, the upper limit drops to 16 kHz and continues to diminish gradually. This is commonly known as *presbycusis*, which is a condition of hearing loss due to the effect of age. Furthermore, the reduction of the audible frequency range is also accompanied by a reduction of sensitivity at all frequency ranges. The yearly rate of change in hearing level due to the effects of noise exposure and age can be estimated from a mathematical model put forward in the ISO 1999:2013.

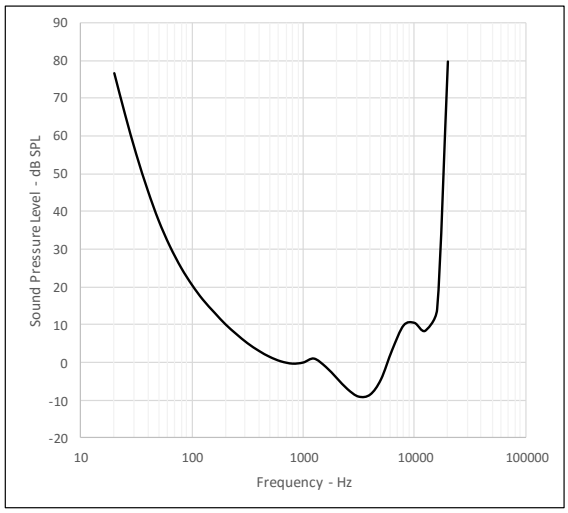

Figure 1a. Absolute Threshold - Normal Hearing Subjects (ISO 226)

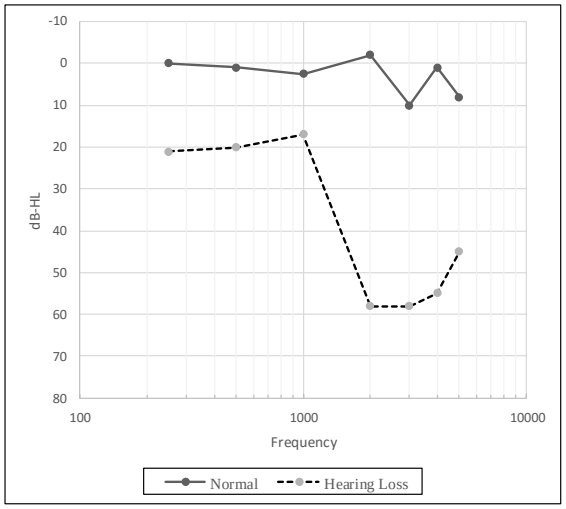

Figure 1b. Typical Hearing Level of a Normal Hearing Listener and a Hearing Loss Listener (Tran, Hétu, 1996)

Deterioration of the Frequency Selectivity

Hearing loss at the cochlear level affects not only the absolute threshold but also the *frequency selectivity*. Frequency selectivity is one of the most important characteristics of the auditory system because it represents the ability to separate or resolve multiple spectral peaks in a complex sound and it contributes to the detection of acoustic signals in a noisy background. This phenomenon is called *masking*. The most obvious example of masking is when a low-intensity sound is not audible in the presence of another loud sound, but masking can also occur when two sounds have similar frequencies. The basilar membrane within the cochlea of the inner ear acts as a frequency analyzer. Each part of the basilar membrane corresponds to a certain frequency range. High-frequency sound stimulates the parts of membrane close to the outer ear and low-frequency sound stimulates responses of the membrane at the end of the cochlea. Since different sections of the basilar membrane correspond to different frequencies, those sections actually act as a series of band-pass filters commonly known as *auditory filters.* The frequency selectivity of the auditory system is characterized by measuring the bandwidth of these auditory filters. To characterize the frequency selectivity of the auditory system, the Equivalent Rectangular Bandwidth (ERB) of the auditory filter at different frequencies is measured (Glasberg & Moore, 1990). Clinical

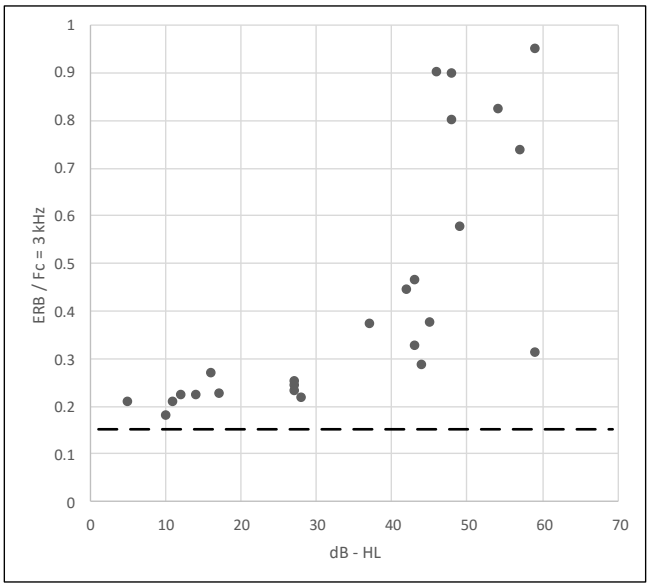

Figure 2. Equivalent Rectangular Bandwidth of Noise-Exposed Industrial Workers (Laroche et al., 1992)
The dashed line represents the expected normalized EBR for normal hearing listeners (Glasberg & Moore, 1990)

methods to measure ERB were developed and successfully used for normal hearing and noise-induced hearing loss subjects (Tran & Hétu, 1991). Abnormally wide ERBs indicated a deterioration of the hearing frequency selectivity due to SHL. Several studies have been conducted on the relationship between loss of sensitivity (determined by an audiogram) and loss of frequency selectivity (characterized by wider ERBs). The results of these studies indicated only a moderate correlation between the two measurements, even when well-trained listeners were employed. Figure 2 represents the relationship of the absolute threshold (dB - HL) and the measured ERB for a group of industrial workers (Laroche, Hétu, Tran, Josserand, & Glasberg, 1992). The hearing condition of these workers, determined by their audiogram, is categorized as normal (HL < 25 dB) to moderate hearing loss (HL > 25 dB). The result can be separated into a horizontal portion of ERB ~ 0.25 below the hearing threshold of approximately 30 dB HL and an upward sloping line above 30 dB HL. For normal hearing listeners, the measured ERBs tend to stabilize around a mean value of 0.23, but when the absolute threshold is greater than 30 dB HL, the measured ERB increases substantially. The spread of measured ERBs is remarkable, for listeners with a hearing threshold approximately 50 dB HL, the range of EBRs spreads from 0.2 to nearly 1.0. This observation led to the conclusion that, while both loss of sensitivity and loss of frequency selectivity are consequents of sensorineural hearing, they represent two completely different aspects of hearing perception. The first one represents the ability to detect acoustic signal in silence and the latter represents the ability to detect acoustic signal in a noisy background. The measure of one cannot be used to predict the other; therefore, any sound design strategy for the VR environment must consider these two hearing capabilities independently.

Compression of the Loudness Function

Loudness is defined as the "subjective intensity" of the sound. This sensation, which is related to the sound pressure, allows us to distinguish between soft and loud sound. The perceived loudness varies in function of the sound frequency. Therefore, it is possible that a sound with larger pressure amplitude is perceived softer than one with a lower pressure amplitude. To quantify loudness, scientists define *phon* as the unit of loudness that is exactly equal to the sound pressure of a 1 kHz tone. Consequently, 40 phons is the loudness of a tone of 1 kHz with sound pressure equal to 40 dB SPL. In other words, a loudness level of 40 phons is equivalent to "as loud as 40 dB SPL of 1 kHz tone." Using this definition, we can measure the sound pressure of tones at different frequencies that produce a loudness equivalent to 40 phons. Figure 3a shows the equal loudness contour for a loudness level of 0, 40, and 80 phons (ISO 236: 2003).

Although the perceived loudness is related to the sound pressure (i.e., loudness increases with sound pressure), there is not a simple one-to-one relationship—double the sound pressure will not double the perceived loudness. To determine how loudness varies with sound pressure, the level of 40 dB SPL of a 1 kHz tone had been proposed to be the reference level for loudness sensation. *Sone* was defined as the unit for loudness sensation, consequently the loudness level of a tone of 1 kHz with sound pressure of 40 dB SPL is equal to 1 sone. For normal hearing listeners, an increase of approximately 10 dB SPL will double the loudness sensation. For instance, a sound with a sound pressure of 40 dB SPL has to be increased to 50 dB SPL in order to double the loudness sensation, which then corresponds to 2 sones. Figure 3b depicts the loudness function of a 1 kHz tone for a normal hearing listener and a listener with SHL. As shown in this figure, the effect of hearing impairment results in a "compression" of the loudness function, especially in the frequency regions where there is loss of sensitivity (Hellman & Meiselman, 1990). This condition is known as *loudness recruitment;* it is manifested by an abnormally rapid growth of perceived loudness above the hearing threshold, which slows down to a normal growth as the intensity of the signal increases. In other words, when presented with the same variation of sound pressure, listeners with a hearing impairment will experience a rapid increase of loudness sensation as compared to normal listeners.

Another consequence of loudness recruitment is the reduction of the hearing dynamic range— the difference between the loudest sound and softest sound. As opposed to people with normal hearing, individuals with hearing loss experience a reduced dynamic range; the loudness of a sound goes from very soft to uncomfortably loud within a small range of increase sound intensity. Loudness recruitment is due probably to a combination of the loss of sensitivity and the decrease of frequency selectivity. As explained previously, hearing loss listeners are characterized by a loss of absolute sensitivity and an increase of the auditory filter bandwidth. Because of the loss of sensitivity, sound cannot be detected until the sound pressure level reaches the absolute threshold; beyond the absolute threshold, loudness level increases rapidly because of an abnormally wider bandwidth of the auditory filters (Zwicker, 1990).

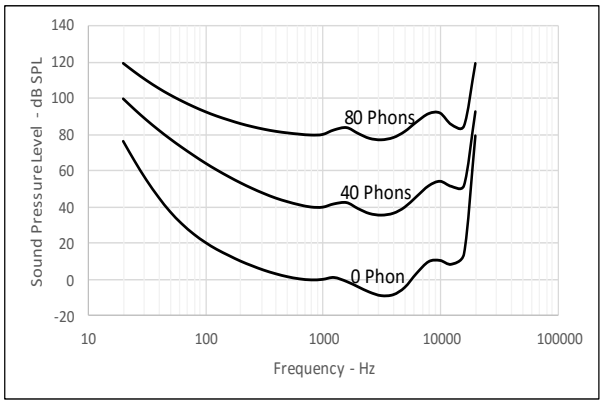

Figure 3a. Equal Loudness Curves (ISO 226-2003)

Figure 3b. Loudness Function of a Normal Hearing Listener and a Hearing Impaired Listener

Loss of Spatial Resolution

Auditory localization is the element of auditory perception that is the most critical to human effectiveness and safety. Sound can normally be heard much earlier than the source of the sound can be seen. In order to localize a sound, the auditory system relies on binaural and monaural acoustics cues. The "duplex theory" was first proposed by Lord Rayleigh around 1906. Our two ears are separated by a relatively large head, thus for a sound that originated from the horizontal plane, the onset time (time to reach the ear) for that sound is different for each ear. This is referred to as the interaural time difference (ITD). However, these localization cues are effective only for sound with frequencies lower than approximately 1.5 kHz. Another mechanism involved in the sound localization is the interaural level difference (ILD), which represents the difference of sound intensity that reaches each ear. The combination of ITD and ILD represents the binaural cues that the auditory system uses to localize sound originated from the horizontal plane. Psychophysical experiments have shown that the duplex theory is incomplete because there are gaps in frequency that neither the ITD nor ILD can be used as cues for localization. In addition, sounds originated from the front and the back of the sagittal plane produced the same values of ITD and ILD. Therefore, ITD and ILD alone cannot explain our capability to distinguish sound that comes from the front versus the one that comes from the back of the sagittal plane. It has also been shown that sound can be localized with only monaural cues. Monaural cues consist of direction-dependent sound described by head-related transfer functions (HRTFs)—the transformation of the sound from the free field to the eardrum (Zhong & Xie, 2014). These spectral cues are essential to resolve front–back confusions and to localize sounds in the vertical plane elevation (Blauert, 1996). The ability to localize sounds is substantially degraded in hearing impaired persons (Akeroyd, 2014). As explained previously, our hearing systems rely on the ITD and ILD to localize sounds from the horizontal plane. Therefore, asymmetrical hearing loss (i.e., unilateral hearing loss) decreases the localization performance in the horizontal plane (Comalli & Altshuler, 1976). Likewise, this decrease presents when peripheral asymmetry is artificially introduced by earplugs or headsets. Many studies have reported that symmetrical hearing loss has little effect on localization performance on the horizontal plane (Abel & Hay, 1996). In contrast, the effect of bilateral hearing loss affects sound localization performance in the vertical plane, especially for listeners with hearing loss in high-frequency regions (Noble, Byrne, & LePage, 1994).

The effect of age on the performance of localization is noticeable. In an extensive study, Abel, Giguère, Consoli, and Papsin (2000) investigated this effect. Seven groups of 16 listeners, aged from 10 to 81 years, participated in this study. The results indicated a decrease in performance as early as in the third decade. The decrease was largest for low-frequency sounds and smallest for broadband noise. Another extensive study (Gagné, Tran, Denis, & Leblanc, 1998) investigated the performance of localization with 21 normal hearing listeners and 21 SHL listeners. Four different types of sound were used in this study: pure tones with constant pressure level, pulsed sounds, amplitude modulated sounds, and frequency modulated sounds. The results obtained for the group of normal hearing listeners indicated:

1- For the horizontal plane, errors of localization are intra-quadrant (i.e., less than 90 degrees). The rate of localization success is approximately 70%. Sounds with frequency of modulation are the best condition for localization, with a greater than 90% rate of success.

2- With a sound test that contains only one single frequency component (e.g., pure tone), the rate of success is better when the test is performed without a background noise.
3- With a sound test that contains multiple frequency components (e.g., frequency modulated sound), the rate of success is excellent for both testing conditions, with or without background noise. Furthermore, a signal-to-noise ratio of +12 dB SPL is enough to ensure very good performance of localization in the horizontal plane.

Generally, for all test sounds and for both testing conditions, with and without background noise, the performance of localization decreased for the group of hearing loss listeners when compared to the groups of normal listeners. The observation of these studies supports the idea that the hearing system localizes sound based on ITD and ILD indices and has the ability to analyze the frequency content of an incoming sound. The degradation of this ability to localize sound limits the amount of spatial information that hearing impaired listeners can obtain from the real world. Obviously, this limitation will eventually actualize when hearing impaired listeners evolve in a virtual environment.

Head-Mounted Display

The ability to localize sound as discussed in the above section is applicable when sound is originated from the free field. Our hearing system perceives sound generated from the free field significantly differently than if it was generated through headsets, which is the case of the head-mounted display. Typically, when we listen to the music from headphones, the music appears to be originated from within the head. For VR environments that use headphones to deliver an acoustical display, it is a simple process to manipulate the indices IDT and ILD to simulate the direction of the sound. However, as explained previously, these indices are incomplete for sound localization and efficient only for sounds originated from the horizontal plane. The ability to localize sound can be preserved by "externalizing" the sound environment. The process of externalizing sound is known as *virtual auditory space*. To create a virtual auditory space, it was assumed that if the acoustics information at the listener's eardrum is the same regardless if the sound was originated from the free field or from the headphones, the listener experience should be also the same. The transformation of the acoustical information from the free field to the eardrum is mainly due to the outer ear, with the pinna acting as a sound spectral filtering system. Beside the effect of the outer ear, the head shape, shoulders, and torso also affect the transformation from the free field to the listener's eardrum. Audio reproduction through headphones simulating a virtual acoustical environment as if it was generated from the free field is possible by applying the HRTF to the virtual sound. Briefly, the localization capability can be achieved using headphones by creating a "virtual auditory space" (Fink & Ray, 2012). For applications where localization does not need high accuracy, an average HRTF can be used to externalize sounds (Jones, Stanney, & Foaud, 2005).

Based on the above literature review, aging is associated with a decrease in auditory capability, whether detecting a sound in a noisy background or localizing an aural cue. The *2015 Demographics: Profile of the Military Community* report, published by the Department of Defense (DoD), states that more than one-quarter (28.1%) of Active Duty personnel are 31 years of age or older. This demographic is more likely to demonstrate decreased auditory performance while performing a training task within a VR environment than their younger counterparts.

PERFORMANCE IMPAIRMENT ANALYSIS: AN ECOLOGICAL APPROACH

An important aspect influencing human performance in VR environments is the effect of user differences. Auditory capability is very different for normal listeners when compared to hearing loss listeners. Many studies have evaluated the performance of virtual audio display with normal hearing listeners, but little information is available on the effect of hearing loss on human performance in VR environment. We found a considerable amount of data gathered on the effects of SHL on auditory performance in the literature. A major portion of this work characterized alterations in basic auditory functions, such as loss of hearing sensitivity, deterioration of the frequency selectivity, temporal resolution, and compression of the loudness function in a number audible frequency regions. Yet, very little work discussed how these hearing impairments affect human performance, especially within a VR environment.

Partial loss of hearing functions is referred to as an *impairment*, according to the World Health Organization's (WHO) nomenclature of the effects of chronic disorders (WHO, 1980). Systematic characterization of impairment provides a description of the pathological condition and offers clues to help understand the underlying auditory mechanism. However, these impairments, as such, may be inconsequential if they do not result in reduced auditory capability that interferes with the ability to perform a specific task. *Disability* emerges only when a dysfunction is actualized. When

an individual with an impairment is confronted with a specific demand in a given environment and is limited or prevented from performing that specific task, that person, consequently, experiences a *handicap*. For example, during a training exercise, a trainee is required to detect a warning sound generated from the simulator flight deck. If the trainee has a reduced hearing capability and the level intensity of the warning sound is set too low in relation to the trainee's hearing function, the trainee will not be able to detect this sound warning. In this case, the detection of a warning sound represents an *auditory demand*, and the reduction of hearing capability represents an *impairment*. This specific impairment is only actualized because the training task "detecting the warning sound" is required. Therefore, the trainee experiences a disability situation because it prevents the trainee from successfully performing the training task. Based on this illustration, performance impairment is viewed with an *ecological* perspective that focuses on the interaction between individual capacity and the training environment. This perspective leads us to analyze performance impairment in terms of a mismatch between capacities and demands. In other words, it allows us to describe the effect of hearing loss within the VR environment in which hearing-impaired trainees evolve. Figure 4 illustrates how hearing disorders can result in a hearing disability that affects human performance.

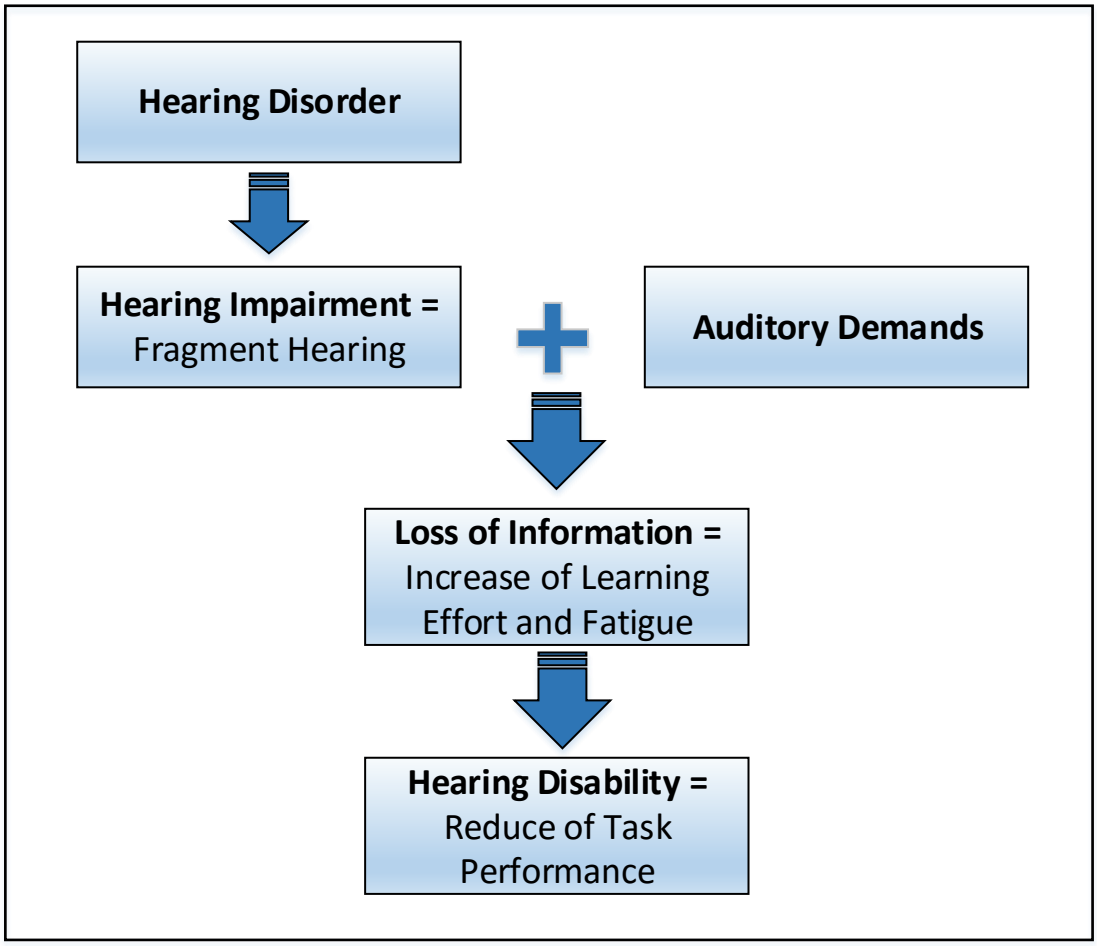

Figure 4. Diagram Illustrating How a Hearing Disorder Can Result in Reduction of Task Performance

In the next section of this paper, we will examine how this ecological framework is used to analyze the mismatch between the hearing impairment and the auditory demands, and what actions we can take to enhance human performance experience in a virtual environment where an auditory environment is present.

MATCHING DEMANDS WITH AUDITORY CAPACITIES: A GUIDELINE FOR SIMULATING SOUND IN A VIRTUAL REALITY ENVIRONMENT

As explained previously, hearing impairments may be inconsequential if they do not result in reduced capability to perform a specific training task. The ecology of human performance is a framework to consider desired performances, such as the relationship between human skills (e.g., abilities, experiences), the tasks to be performed, and the environment (real-life or virtual environment). In this context, the performance impairment is explicitly viewed within an ecological perspective that focuses on the interaction between individual capacity and environment demands. It is impossible to fully understand human performance without an understanding of the environment and the context within which the task was performed. This perspective leads us to analyze performance impairment in terms of compatibility or mismatch between demands and capabilities. In other words, it allows us to describe the effect of hearing loss within the environment in which hearing-impaired users evolve. This strategy acknowledges the person's handicap, so the training environment is designed in the way that it will not interfere with the expected performance. Figure 5 represents the ecological framework to characterize auditory demands and match them to the auditory capabilities. The various tasks that are performed, whether they involved detection, recognition, or localization, call on different specific capacities, depending on the exact nature of the sound signal's action on the auditory system. This ecological approach was proposed and used for adapting industrial work settings, in terms of incompatibility between workplace auditory demands and residual hearing capacities (Hétu, 1994; Tran & Hétu, 1996).

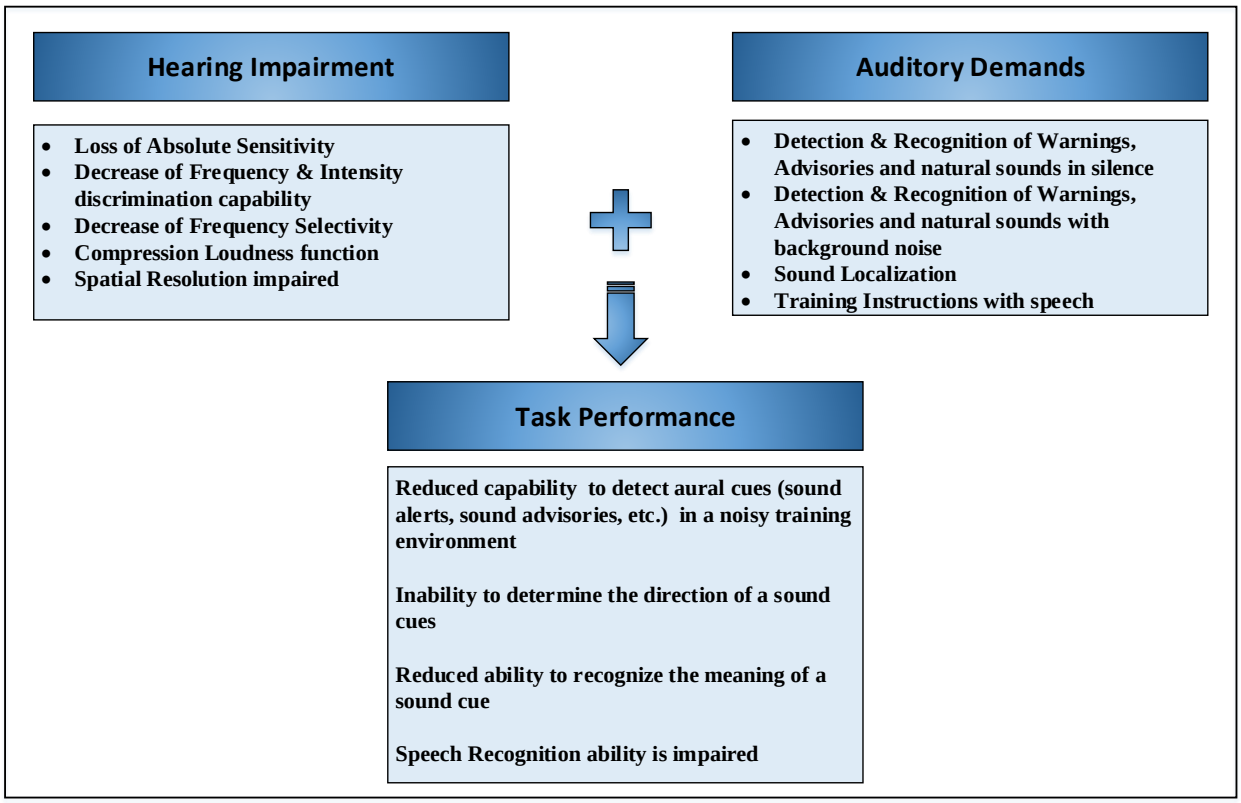

Figure 5. Ecological Framework for Analysis of Possible Mismatch Between Hearing Impairment and Auditory Demands

Human perception of the acoustic environment is very complex. Phenomena, such as cocktail party (Bronkhorst, 2000) or auditory scene (Bregman, 1990), involve mechanisms that go beyond the peripheral aspect of the hearing system and are not yet fully understood by researchers. For the purpose of this paper, a simplified view of a simulated acoustic environment is a composition of two main components: the background acoustic environment and the aural cues that convey "useful" information. For example, in a flight simulator, engine noise represents the acoustic background and cockpit-warning sounds are aural cues; in an urban simulation, the cacophony simulated the background acoustic environment, and the warning sound of an incoming ambulance represents the stimulated aural cue. For a virtual

environment that requires the production of auditory display, human training performance can be enhanced via optimized designs by controlling either the background acoustic environment, the aural cues, or both. Whether for normal listeners or for hearing loss listeners, maximizing the efficiency of the information conveyed in VR will require developing a set of guiding design principles that take into account both conditions: hearing limitations of users and the auditory demands. Table 1 provides such guidelines that can be used to achieve these ends.

Table 1. Guidelines for Designing Simulated Acoustical Environment within the VR Environment

G1. Match stimulus intensity with hearing capability. Moderate hearing loss that affects auditory perception and speech recognition can often mitigate by *increasing stimulus intensity*. Use ISO guidelines to estimate hearing thresholds: ISO 28961:2012 provides references to obtain the absolute hearing thresholds of normal listeners. ISO 9189:2013 provides a model to estimate the decrease of hearing threshold due to the effect of noise-induced hearing loss and the effect of age.

G1.1. A comfortable sound level is approximately 10-15 dB* SPL above the hearing thresholds.

G1.2. Avoid a wide range of variation in sound intensity to control the effect of loudness recruitment.

G2. Control the effect of background noise. The decrease of frequency selectivity affects the capability to detect acoustic information within a noisy background. Therefore, when possible, *keep background noise at a minimum*. Control the signal-to-noise (S/N) ratio as follows:

G2.1. The S/N ratio represents the level of an aural cue over the level of the masked threshold. An appropriate S/N ratio level improves the capability to detect and recognize simulated aural cues within a noisy background, such as warning sounds from a simulated flight deck or warning sounds generated from medical equipment in a simulated of operating room. Ensure that the S/N ratio exceeds at least +13 dB*, but no more than +25 dB* (ISO 7731:1986). The level of masked threshold for both normal and hearing loss listeners can be estimated using existing a computerized model such as DETECTSOUND (Tran, Hétu, & Laroche, 1992).

G2.2. Avoid the usage of pure tone (sound with a single frequency component) for sounds to be detected in a noisy environment.

G3. Optimize the capability of localization. The capability of localization decreases as hearing loss increases. Optimizing the localization can be achieved with the following guidelines:

G3.1. Avoid the usage of pure tone (single frequency component) for sounds to be localized.

G3.2. Generally, to localize sounds originated from the vertical plane, the hearing system uses cues from the sound frequency spectrum. Therefore, the usage of acoustic signals with multiple frequency components, such as frequency of modulation (FM), will optimize the localization.

G3.3. Use signals with a duration of more than 1 second. The duration of sound improves the localization due to the head movement.

G3.4. Limit the signal envelope onset and offset to 20 msec or less.

G3.5. Limit the signal frequency spectrum between 300 Hz and 3000 Hz. This is the most efficient frequency zone for the hearing system to localize acoustic signals.

G3.4. For a head-mounted display device, sound localization can be preserved by using the average HRTF for externalizing simulated acoustical environment.

G4. Optimize speech recognition. Speech recognition is a combination of both sensory and cognitive processes. Sensory hearing factors alone, as the ones discussed in this paper, cannot account entirely for the human ability to perceive and recognize speech. Nonetheless, dramatic effects on speech perception due to hearing loss have been reported in many studies. For a virtual environment that uses speech information, such as training instructions, the following guidelines can be used to optimize speech perception and recognition:

G4.1. Speech recognition can be significantly enhanced with instructions semantically well-structured.

G4.2. Present speech at a reasonable and consistent pace.

G4.3. Whenever possible, repetition enhances speech perception and recognition.

**Note*: The level of signal and sound pressure are quantified in 1/3 octave band.

CONCLUSION

Training within the VR environment involves two main factors: first, the hardware and software used to create the VR environment and, second, the training tasks to be performed. Therefore, to justify the use of VR technology for training, when compared to alternated approaches, the training task performance should improve over time when transferred to the real-world task. In this paper, we presented an analysis of how hearing impairment can result in a disability of human training performance in VR environments. The analysis is based on an ecological framework that focuses on the relationship of hearing impairments and auditory demands. A mismatch between auditory demands and hearing capability inevitably affects training task performance in a cascading fashion: *hearing impaired −> fragment hearing −> loss of information −> increase learning effort and fatigue −> pace of learning −> performance impaired.*

For a virtual reality environment that requires the simulation of an auditory display, human training performance can be enhanced via optimized designs. Maximizing the efficiency of the information conveyed in VR requires development of a set of guiding design principles that take sensory limitations into account. Based on the proposed ecological framework, this paper provides guidelines to match the auditory demands with the altered auditory capacities. These guidelines can be used to effectively design and implement a simulated auditory environment, while taking into account human auditory performance, including individuals with reduced hearing capability.

Virtual environment designers tend to focus on the replication of the real world (the technology) and neglect the aspect of human capability (the human factor). An ecological approach, as explained in this paper, is advocated, paying attention to interactions within the system: trainee abilities, training environment, and the interface between the two. Finally, researchers need to focus significantly on addressing a number of human factors if VR systems are to be effectively used and adapted to the user's capability. We believe an ecological model, as the one presented in this paper, is well suited to describe and explain the relationship between constraints imposed by human sensory and human performance when performing a training task. Moreover, this framework can further expand to include other human factors, such as vision, memory, and attention.

REFERENCE

Department of Defense (2015). 2015 Demographics: Profile of the Military Community. Office of the Deputy Assistant Secretary of Defense for Military Community and Family Policy

International Standards Organization ISO 7731. (1986). Danger Signals for Workplaces – Auditory Danger Signals.

International Standards Organization ISO 532. (1989). Acoustics - Method for calculating Loudness level.

International Standards Organization ISO 7196. (1995). Acoustics - Frequency-weighting characteristic for infrasound measurements.

International Standards Organization ISO 226. (2003). Acoustics – Normal equal loudness level contour.

International Standards Organization ISO 389-7. (2005). Reference threshold of hearing under free-field and diffuse-field listening conditions.

International Standards Organization ISO 28961. (2012). Acoustics - Statistical distribution of hearing thresholds of otologically normal persons in the age range from 18 years to 25 years under free-field listening conditions.

International Standards Organization ISO 9189. (2013). Acoustics - Estimation of noise-induced hearing loss.

Abel, S. M, Giguère, C, Consoli, A, Papsin, B. C. (2000). The Effect of Aging on Horizontal Plane Sound Localization. *Journal of the Acoustical Society of America*, 108, 743–752.

Abel, S. M, Hay, V. H. (1996). Sound Localization: The Interaction of HPDs, Aging, and Hearing Loss. *Scandinavian Audiology*, 25 (1), 3–12.

Akeroyd, M. A. (2014). An overview of the major phenomena of the localization of sound sources by normal-hearing, hearing-impaired, and aided listeners. *Trends in hearing*, vol. 18.

Bregman, A. S. (1990) Auditory scene analysis. MIT Press: Cambridge, MA.

Bronkhorst, A. W. (2000). The Cocktail Party Phenomenon: A Review on Speech Intelligibility in Multiple-Talker Conditions. *Acta Acustica united with Acustica*, 86: 117–128.

Comalli, P. E.; Altshuler, M. W. (1976). Effects of Stimulus Intensity, Frequency, and Unilateral Hearing Loss on Sound Localization. *Journal of Auditory Research*, 16, 275–279.

Davis, E. T., Scott, K., Pair, J., Larry F. Hodges, L. F., Oliverio, J. (1999). Can Audio Enhance Visual Perception and Performance in a Virtual Environment? *Proceeding of Human Factors and Ergonomics Society Annual Meeting*. Vol 43, Issue 22.

Fink, K. J., Ray, L. (2012). Tuning Principal Component weights to individual HRTFs. *Acoustics, Speech and Signal Processing, IEEE International Conference*, 389-392.

Freeman, J., Lessiter, J. (2001). Here, There and Everywhere: The effects of multichannel audio on presence. *Proceedings of the 2001 International Conference on Auditory Display, Espoo,* Finland, 231-234.

Fuchs, P., Moreau, G., Guitton, P. (2006). Virtual Reality: Concept and Technologies. *CRC Press/Balkema*.

Gagné, J. P., Tran, H., Denis, S., Leblanc, M. (1998). Conception de signaux sonores et sur la mesure inductive de la capacité de localization auditive des avertisseurs sonores de danger en milieu industriels. *Institute de Recherche sur la Santé et Securité au Travail*.

Glasberg, B. R.; Moore, B. C. J. (1990), Derivation of auditory filter shapes from notched-noise data, *Hear. Res.* (47), 103–138.

Hellman R. P., Meiselman C. H. (1990). Loudness relations for individuals and groups in normal and impaired hearing. *Journal Acoustical Society of America*, 88:2596-2606.

Hendrix, C., Barfield. W. (1996). The sense of presence within auditory virtual environments. *Presence Teleoperators and Virtual Environments.* 5(3): 290-301.

Hétu, R. (1994). Mismatches between auditory demands and capacities in the industrial work environment. *Audiology*. 33:1-14.

Hétu, R., Tran, H. (1996). Psychoacoustic Performance in Workers with NHIL. Scientific Basic of Noise-Induced Hearing Loss. *Thieme Medical Publishers, Inc*. Chapter 22. 264-285.

Jones, D. L., Stanney, K. M., Foaud, H. (2005). An Optimized Spatial Audio System for Virtual Training Simulations: Design and Evaluation. *Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display*, 223-227.

Kukka, H., Goncalves, J., Wang, K., Puolamaa, T., Louis, J., Mazouzi, M., Barco, L. R. (2016). Utilizing Audio Cues to Raise Awareness and Entice Interaction on Public Displays. *Proceedings of the 2016 ACM Conference On designing Interactive Systems*. Brisban, QLD, Australia. 807-811.

Laroche. C., Hétu, R., Tran, H., Josserand, B., Glasberg, B. (1992). Frequency Selectivity in workers with noise-inducted hearing loss. *Hearing Research*, 64, 61-72.

Mikropoulos, T. (2006). Presence: A Unique Characteristic in Educational Virtual Environments. *Virtual Reality*, 10**,** 197-206.

Noble, W.; Byrne, D.; LePage, B. (1994). Effects on Sound Localization of Configuration and Type of Hearing Impairment. *Journal of the Acoustical Society of America*, 95, 992–1005.

Sheridan, T. B. (1992). Musings on telepresence and virtual presence. Presence: *Teleoperators and Virtual Environments*, 1 (1), 120-126.

Stevens, J. A., Kincaid, J. P. (2015). The Relationship between Presence and Performance in Virtual Simulation Training. *Open Journal of Modelling and Simulation*, 2015, 3, 41-48.

Tran, H., Hétu, R. (1996). La planification de la signalization acosutique en milieu industriel: Critère de conception des avertisseurs sonores de danger. *Canadian Acoustics, 24(2), 3-17.*

Tran, H., Hétu, R., Laroche, C. (1991). The choice of a procedure for measuring Signal Detection Capabilities in Noise with a view to developing a clinical test. *Journal of Speech and Language Pathology, 15(2), 21-33*.

Tran, H., Hétu, R., Laroche, C. (1992). Computerized assessment and prediction of the audibility of sounds warning signals for normal and hearing impaired individuals. *Mattila, M., Karwonski, W. (eds)*. Computer Application in Ergonomics, Occupational Safety and Health. Elsevier Science Publishers. 105-112.

Winn, W., Windschitl, M., Fruland, R., Lee, Y. (2002). When Does Immersion in a Virtual Environment Help Students Construct Understanding? *Proceedings of the International Conference of the Learning Sciences*, Mahwah, 497-503.

World Health Organization. (1980). International Classification of Impairment, Disabilities, and Handicaps. Geneva, WHO.

Zhong, X. L., Xie, B. S. (2014). Soundscape Semiotics - Localization and Categorization. *InTECH Publishers*. Chapter 6.

Zwicker, E., Fastl, H. (1990). Psychoacoustics – Facts and Models. *Springer Serie in Information Sciences*.